Faculty of Humanities

Department of Archaeology

Evaluating network science in archaeology.

A Roman Archaeology perspective

by

Tom Brughmans

VOLUME I: THESIS

Thesis for the degree of Doctor of Philosophy

October 2014

ABSTRACT

In the last decade, network science techniques have become more common in archaeology. But is this just a short-lived trend caused by the popularity of these techniques in other disciplines, the availability of big data, and the rise of the World-Wide Web and online social networks? Or does network science hold the potential of making a unique contribution to our discipline? This PhD project aims to evaluate this potential.

It starts with the claim that the unique contribution of network science to archaeology lies in its ability to deal with network data. It evaluates this claim through a literature review and three archaeological case studies. Two of the case studies concern research themes in Roman archaeology, to focus the evaluation process on one archaeological subdiscipline rather than tackling the full diversity of research themes and data types in archaeology. This thesis takes a Roman archaeology perspective to evaluating network science, and further discusses the implications of its results for the archaeological discipline at large. This project not only aims to evaluate whether network science does indeed offer a unique contribution, but also whether archaeologists have need for such techniques. Do archaeologists commonly ask research questions that are best addressed through the analysis of network data? Does network science allow archaeologists to ask new questions? Are archaeologists commonly confronted with network data?

The literature review reveals that formal network techniques have been used in archaeology since at least the 1960s, but have only recently become more commonly applied. Moreover, it concludes that the methodological range of network techniques applied in archaeology is limited and that its place in the archaeological research process is ill-defined. It suggests these issues can be tackled by taking a broad multi-disciplinary scope, by letting the archaeological research context dominate the adoption and development of network techniques, and by working explicitly through a network science research process. These three lessons subsequently influenced the selection of three archaeological case studies that have the ability to address this project's research questions, to evaluate possible particularities in the use of network science for Roman archaeology, and that have the potential to allow the application of network techniques that have never before been applied in archaeology.

The first case study presents a citation network analysis of the adoption, use and adaptation of formal network techniques in archaeology. This case study can be seen as a network analysis of the literature review, using exploratory network techniques and visualisations. It illustrates how an exploratory network analysis allows one to gain insights that cannot be obtained from a close reading in a literature review, and how some of the conclusions of the literature review can be identified or reproduced through the use of exploratory network techniques. The second case study is an analysis of inter-settlement visibility in Iron Age and Roman Southern Spain. It introduces a statistical modelling approach (Exponential Random Graph Modelling (ERGM)) for simulating dependence assumptions surrounding visibility networks: theoretical assumptions formulated by archaeologists as hypotheses of why lines of sight between settlements matter. Finally, the third case study presents an agent-based network model of tableware distribution in the Roman Eastern Mediterranean. It focuses on the potential of network science methods for performing confirmatory analyses by evaluating two complex hypotheses surrounding the functioning of the Roman trade systems.

This thesis concludes that archaeologists commonly formulate (implicitly or explicitly) dependence assumptions, that they are confronted with network data, and that in such cases network science techniques can offer methodological advantages over other methodological tools used by archaeologists. This conclusion firmly establishes network science in the archaeologist's methodological toolbox. The main contributions of this PhD project are, therefore, i) to argue why network science offers archaeologists unique and necessary methodological advantages, ii) to provide three critical practical examples of how it can be applied, iii) to offer a number of suggestions which could guide the future archaeological use of network science, iv) to introduce innovative methods of citation analysis and statistical network modelling to our discipline, and v) to emphasise the importance and possibility of falsifying and abstracting Roman archaeology hypotheses as transparent and comparable conceptualisations.

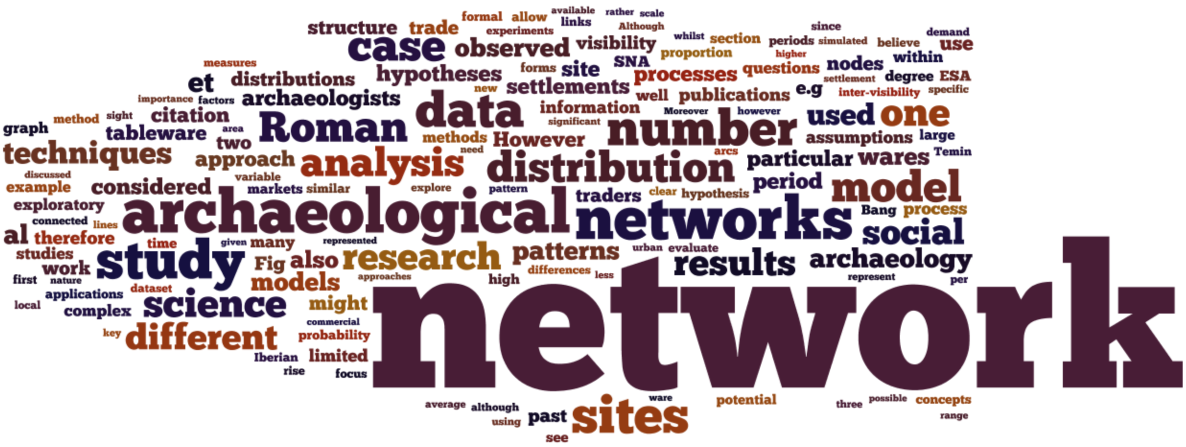


Fig. 1. Word cloud of the full text of this PhD thesis. It's about networks!

TABLE OF CONTENTS

CHAPTER 1. Introduction: why evaluate network science in archaeology? 16

1.1. Aims 16
1.2. What is network science? 17
1.3. Why does archaeology need network science? 24
1.4. Research questions 27
1.5. Why a Roman archaeology perspective? 28
1.6. The structure of this PhD 29

CHAPTER 2. Literature review: the roots and shoots of archaeological network science 35

2.1. Introduction 35
2.2. Origins in graph theory? 36
2.3. Social network analysis 43
2.3.1. An introduction to SNA 43
2.3.2. An early archaeological model 44
2.3.3. Diffusion of material and immaterial resources 46
2.3.4. Network centrality 48
2.3.5. Affiliation networks 51
2.3.6. Ego-networks 52
2.4. Discussion: beyond SNA 54
2.5. Complex networks and physics 56
2.5.1. Small-world networks 57
2.5.2. Scale-free networks and power-laws 58
2.5.3. Small-world and scale-free networks in archaeology 60
2.5.4. Discussion: complex networks beyond popular models 62
2.5.5. Dynamic network models in archaeology 64
2.6. Network science in Roman Archaeology 68
2.7. Conclusion: the need to evaluate network science in archaeology 71
2.8. Implications for developing case studies 74

CHAPTER 3.... Case study 1: a citation network analysis of the adoption, use, and adaptation of formal network techniques in archaeology 79

3.1. Introduction 79
3.2. Citation network analysis and archaeology 80
3.3. Data collection, citation network creation and sample size issues 82
3.4. Methods 85
3.5. Results 88

3.5.1. Local scale: network of archaeological applications ... 89
3.5.2. Meso scale: ego-network of archaeological applications ... 94
3.5.3. Global scale: combined dataset ... 102
3.6. Discussion ... 104

CHAPTER 4. ... Case Study 2: understanding inter-settlement visibility in Iron Age and Roman Southern Spain with exponential random graph models for visibility networks ... 113

4.1. Introduction ... 113
4.2. Visibility networks ... 114
4.2.1. The study area ... 114
4.2.2. Why study visibility networks in Iron Age II and Roman Southern Spain? ... 120
4.2.3. Why processes of emerging inter-visibility? ... 124
4.2.4. Hypotheses of visibility network creation processes ... 125
4.2.5. Previous archaeological applications of visibility networks ... 129
4.2.6. What does inter-visibility mean? ... 130
4.3. Data: topography, sites, chronology, and attributes ... 132
4.3.1. Digital Elevation Model ... 132
4.3.2. Settlement data ... 132
4.3.3. Chronology ... 134
4.3.4. Site attributes ... 135
4.4. Method: an approach for exploring changing visibility networks ... 136
4.4.1. Probable viewsheds ... 136
4.4.2. Observation point locations ... 138
4.4.3. Visibility network creation ... 139
4.4.4. Exploratory Network Analysis ... 140
4.4.5. Exponential Random Graph Models ... 140
4.4.6. Software ... 150
4.5. Results of the combined approach ... 150
4.5.1. Exploratory network analysis ... 152
4.5.2. Exponential Random Graph Models ... 159
4.6. Discussion: towards a better understanding of urban connectivity in Iberian and Roman Southern Spain 166

CHAPTER 5. ... Case study 3: Bang's Roman Bazaar and/or Temin's Market Economy? An agent-based network model of tableware distribution in the Roman East ... 175

5.1. Introduction ... 175
5.2. Tableware studies, distribution patterns, and processes ... 180
5.2.1. Introduction ... 180
5.2.2. Studying Roman pottery distributions ... 181
5.2.3. Distribution processes ... 182

5.2.4. Being wrong can be useful: questions, methods, and issues for a new approach 199
5.3. Tableware distribution patterns: data and the state of the art 202
5.3.1. Tableware fabric-groups in the Roman East 202
5.3.2. The ICRATES database 204
5.3.3. Issues with the data.. 205
5.3.4. Chronological overview tableware distribution in the Roman East 207
5.4. Exploratory data analysis 212
5.4.1. Introduction 212
5.4.2. Typologies and functional categories 212
5.4.3. Gaussian distribution 213
5.4.4. Number of sherds per ware.. 214
5.4.5. Number of forms per ware 215
5.4.6. Number of sites per ware... 216
5.4.7. Frequency distributions 217
5.4.8. Exploratory network analysis results.. 221
5.4.9. Conclusion.. 223
5.5. Agent-based model... 225
5.5.1. Model description... 226
5.5.2. Results 241
5.5.3. Conclusions 255
5.6. Discussion 256

CHAPTER 6. Discussion: addressing challenges in archaeological network science.............266

6.1. Lessons learned from the case studies... 266
6.1.1. Case study 1: methodological implications 266
6.1.2. Case study 2: methodological implications 267
6.1.3. Case study 3: methodological implications 269
6.1.4. Implications for Roman archaeology 270
6.2. 'We shall overcome': challenges and suggested solutions.. 272
6.2.1. Method 273
6.2.2. Data 277
6.2.3. Space 279
6.2.4. Process.. 280
6.2.5. Past social networks 283
6.3. A 'guide to good practice': specifications for future archaeological network science...................... 284

CHAPTER 7. Conclusion: the potential of network science in archaeology...........................290

7.1. The contribution of this PhD.. 290
7.2. Where to go from here? Archaeological network science beyond the "hype" 292

8. Bibliography 297

9. Appendix I: literature review 324

9.1. Publications used to compile figure 5 324

10. Appendix II: case study 1 332

10.1. Bibliography local scale 332
10.2. Bibliography tables 4-6 336

11. Appendix III: case study 2 340

11.1. Sites index: dates, urban status, location, road and river networks 340
11.1.1. All 190 sites included in this case study in alphabetical order 340
11.1.2. All 190 sites included in this case study sorted according to their project ID 345
11.1.3. All sites with an urban status 355
11.1.4. Sites on the Via Augusta 356
11.1.5. Sites on the river network 356
11.2. Cumulative viewsheds 358
11.3. Results of the node-based network measures 363
11.3.1. Iberian period network up to 20km 363
11.3.2. Republican period network up to 20km 365
11.3.3. Early Imperial period network up to 20km 367
11.3.4. Middle Imperial period network up to 20km 369
11.3.5. Late Imperial period network up to 20km 370
11.4. ERGM results 372
11.4.1. Goodness of fit of Bernoulli random graph models 372
11.4.2. ERGMs without attributes 384
11.4.3. ERGMs of Early Imperial attributed networks 399

12. Appendix IV: case study 3 417

12.1. Lists of sites and forms used 417
12.2. Exploratory network analysis case study 2 426
12.2.1. Introduction 426
12.2.2. Similarity of forms' distributions: the Brainerd-Robinson coefficient 427
12.2.3. Creating and exploring similarity networks 430
12.2.4. Network structure and sensitivity analysis 431
12.3. Agent-based model code 449

13. Glossary 470

BOX SUMMARIES

Box 1. Summary: what is network science? ... 24
Box 2. Summary Chapter 1 ... 32
Box 3. Summary chapter 2 ... 76
Box 4. Network science research process case study 1 ... 88
Box 5. Summary chapter 3 ... 110
Box 6. Network science research process case study 2 ... 151
Box 7. Summary of chapter 4 ... 172
Box 8. Research outline of case study 3 ... 180
Box 9. Network science research process ABM case study 3 ... 242
Box 10. Summary of chapter 5 ... 264

DEFINITIONS AND ABBREVIATIONS USED

This thesis uses concepts, techniques, and abbreviations uncommon in archaeology and adopted from other disciplines. It was therefore decided to include an elaborate glossary that explains these, see section 13 at the end of this thesis.

ELECTRONIC SUPPLEMENTS

The electronic supplements to this thesis include a PDF version of this PhD as well as the main data and network files used to obtain the results presented here. Three folders contain the electronic supplements of the three case studies. Each folder includes a 'README' file describing the files and how to use them. All spreadsheets are also provided in the .csv format.

Case study 1:

1) 'case-study1_network_arc-list.xlsx': this spreadsheet is an arc list which has two columns: the starting node of an arc and the ending node of an arc.

2) 'case-study1_network_attributes.xlsx': this spreadsheet is a list of attributes for each unique node mentioned in the arc-list. It has five columns: the unique ID of the node, the network scale in which the node is included in the PhD analysis (local, meso, global), the first author, the publication year, the journal title. The latter three columns hold unedited data derived from Web of Knowledge.

Case study 2:

1) 'Case-study2_network.xls': this spreadsheet includes the network used in case study2. The first section (starting at NETWORK) lists each arc, the node it departs from, the node it arrives at, its probability, its length, the geographical coordinates of the from node, and the geographical coordinates of the to node (coordinates in WGS1984_UTMzone30N). The second section (starting at ATTRIBUTES) lists each individual node (siteID, which can be used to map attributes of each unique ID to the FROM nodes and TO nodes), the site name, and the period in which it was occupied (0 indicates is was not occupied, 1 indicates it was occupied).

2) 'Case-study2_site-locations.kml': this file can be opened with Google Earth or GIS and includes the locations used for each site in the case study.

Case study 3:

1) the 'nw' folder contains the nw extension for Netlogo. The entire folder will need to be placed into the 'extensions' folder inside the 'netlogo' installation folder on your computer in order to use the ABM created for this case study.

2) 'Case-study3_ABM.nlogo' is a Netlogo ABM model. In order to run the ABM you will need to download and install Netlogo: https://ccl.northwestern.edu/netlogo/download.shtml The model can be run by first clicking the 'setup' button (takes a long time) and then clicking the 'go' button.

3) 'Case-study3_tableware_matrices.xlsx': this file holds all information used in the exploratory data and network analysis of the ICRATES database.

DECLARATION OF AUTHORSHIP

I, Tom Brughmans, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Evaluating network science in archaeology. A Roman Archaeology perspective

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Either none of this work has been published before submission, or parts of this work have been published as:

Chapter 1: section 1.2 is an adaptated part of ‘Brughmans, T., Collar, A., & Coward, F. (in press). Introduction: challenging network perspectives on the past. In T. Brughmans, A. Collar, & F. Coward (Eds.), *The Connected Past: challenges to network studies of the past*. Oxford: Oxford University Press’. The first draft of this publication from which the text used here was drawn was written by myself, and therefore does not include reworkings by the co-authors of this publication.

Chapter 2: this chapter is an adapted, updated, and extended version of ‘Brughmans, T. (2013). Thinking Through Networks: A Review of Formal Network Methods in Archaeology. *Journal of Archaeological Method and Theory*, *20*, 623–662. doi:10.1007/s10816-012-9133-8’. This publication and therefore this chapter benefited greatly from the detailed comments by Barbara Mills, Claire Lemercier, and three anonymous reviewers.

Chapter 3: this case study is an adapted version of ‘Brughmans, T. (2013). Networks of networks: a citation network analysis of the adoption, use and adaptation of formal network techniques in archaeology. *Literary and Linguistic Computing, The Journal of Digital Scholarship in the Humanities*, *28*(4), 538–562. doi:10.1093/llc/fqt048’. This publication and therefore this chapter benefited greatly from the detailed comments by Almila Akdag, Paul Reilly, and two anonymous reviewers.

Chapter 4: this case study is an adapted version of two publications combined: ‘Brughmans, T., Keay, S., & Earl, G. P. (2014, in press). Introducing exponential random graph models for visibility networks. *Journal of Archaeological Science*.’ and ‘Brughmans, T., Keay, S., & Earl, G. P. (2014, in press). Understanding inter-settlement visibility in Iron Age and Roman Southern Spain with exponential random graph models for visibility networks. *Journal of Archaeological Method and Theory*.' The final versions of these publications were used as a basis for this chapter. Although my co-authors directed the project on which this case study is based and collected the settlement dataset used, written contributions to the final versions of these publications by the co-authors are limited to: a reformulation of the origins of the settlement dataset (section 4.3.2), mentions of the aims and approaches of the 'Urban connectivity' project, grammatical and vocabulary changes, and corrected use of jargon. These publications and therefore this chapter benefited greatly from the detailed comments by Viviana Amati, Barbara Mills, Fiona Coward, Anna Collar, and four anonymous reviewers.

Chapter 5: this case study has not been submitted for publication yet.

Chapter 6: parts of this chapter were adopted and highly extended from: 'Brughmans, T. (2014). The roots and shoots of archaeological network analysis: A citation analysis and review of the archaeological use of formal network methods. *Archaeological Review from Cambridge*, *29*(1), 18–41.'

Chapter 7: section 7.2 is a highly modified version of part of a first draft written by myself of the following co-authored publication: 'Brughmans, T., Collar, A., Coward, F., & Mills, B. J. (2014, in press). Introduction: positioning networks in archaeology. *Journal of Archaeological Method and Theory*'.

Signed:

Date: 23-10-2014

ACKNOWLEDGEMENTS

First of all, I would like to thank Simon Keay and Graeme Earl for their continuous professional and personal support. Ever since I moved to Southampton to start my MSc they have been guilty of constantly pushing me out of my academic comfort zone. They provided me with opportunities that really challenged me at the time but continually boosted my skills and confidence. If I am ever accused of being over-confident, it is their doing. My special thanks also go out to Jeroen Poblome for his constant support and for providing me with a fascinating dataset to play with. I would like to thank the examiners of this thesis, Fraser Sturt and Andy Bevan, for a critical and constructive discussion during the viva voce which will help me improve my future work on what’s presented here.

This thesis is really the result of an emerging community of academics obsessed with networks. I had the honour and pleasure of meeting archaeologists, historians, classicists, computer scientists, physicists and network scientists from all over the world who all share my passion for networks. In particular, I would like to thank the members of the Networks Network discussion group, the CAA community, Marco Büchler and his colleagues in Leipzig, Johannes Preiser-Kapeller and his Austrian colleagues, Marten Düring and his Historical networks group in Germany, Tim Evans and Ray Rivers for their feedback and interest in my work, John Terrell and colleagues at the Chicago Field Museum, Barbara Mills and Matt Peeples for critical expert feedback and for working with me to spread our message to the American archaeology/anthropology communities, and Claire Lemercier and all the great French sociologists and historians that I met thanks to her at a French SNA summer school. The most important event during my PhD, however, is undoubtedly The Connected Past conferences I organised with Anna Collar and Fiona Coward (and later with Claire and Tim). Both know how much this project means to me and how grateful I am to be able to pursue this with them. Without this supporting community this thesis would be a wasted effort.

Cat, Clara, Eleonora, Elly, Emilie, Gareth, Grant, Hatti, James, James, Karen, Leif, Lucy, Maria, Nicole, Pete, Phil, Rachel, Rodrigo, Sam, Sara, Sarah, Scott: we are all in the same boat, thanks for four good years!

I would also like to thank the World Universities Network for funding a research stay at The University of Auckland, where I had the pleasure of working with David O’Sullivan and Ben

Davies (thanks for all the fascinating discussions about work and life), and of interviewing Geoff Irwin and Ethan Cochrane (I really enjoyed that!). The Academia Belgica en Roma is thanked for funding a research stay in Rome which allowed me to lay the foundations of my Roman case studies. Being able to devote myself full-time for four years to this work would not have been possible without the financial support of The University of Southampton and the Faculty of Humanities (and Eleonore Quince for all her hard work in making things happen), and the University of Leuven Sagalassos Archaeological Research Project.

Chapter 2: I would like to thank Leif Isaksen, Claire Lemercier, Barbara Mills, Johannes Preiser-Kapeller, Iza Romanowska, John Terrell and three anonymous reviewers for the many helpful comments on earlier versions of the text.

Chapter 3: I would like to thank Almila Akdag, Ethan Cochrane, Graeme Earl, Tim Evans, Simon Keay, Paul Reilly, Ray Rivers, Matteo Romanello, Iza Romanowska, Fraser Sturt, John Terrell, David Wheatley and two anonymous reviewers for commenting on and discussing the work presented here, and/or reading an early draft.

Chapter 4: the ‘Urban Connectivity in Iron Age and Roman Southern Spain’ project directed by Prof. Simon Keay and Dr. Graeme Earl was funded by the UK Arts and Humanities Research Council (AHRC) between 2002 and 2005 with subsequent support by the University of Southampton and institutions in Seville. I would like to thank Mari Carmen Moreno Escobar, Viviana Amati, two peer-reviewers; Cat Cooper for help with the maps, Leticia López Mondéjar for bibliographical suggestions, Iza Romanowska for advice and reading an early draft, Matt Peeples for suggestions concerning sensitivity analyses, Pablo Garrido González for expert advice on the study area, and Dave Wheatley for his ArcGIS Python script for probable viewsheds.

Finally, my fiancée Iza for everything.

CHAPTER 1

INTRODUCTION:

WHY EVALUATE NETWORK SCIENCE IN ARCHAEOLOGY?

1. Introduction: why evaluate network science in archaeology?

1.1. Aims

This PhD project aims to make a methodological contribution to the archaeological discipline at large and to Roman archaeology specifically by evaluating the potential of network science as a method in archaeology. This will be achieved by:

- evaluating previous archaeological applications of network science (chapters 2-3);
- identifying methodological challenges of the application of network science in archaeology and their implications for Roman archaeology (chapters 2-6);
- addressing some of these challenges through three archaeological case studies (chapters 3-5);
- evaluating network science techniques that have not previously been applied in archaeology and are (a) developed and/or selected in light of specific archaeological research contexts, research questions and data, and (b) draw on a critical understanding of existing methods in multiple disciplines (chapters 3-5).

This thesis therefore aims to provide an overview of the advantages and disadvantages of network science in archaeology (chapter 6), which might allow archaeologists to better evaluate the potential contribution a network science method offers within particular research contexts. Such research is a useful exercise for the archaeological discipline due to the utility network science has shown in other disciplines and the increasing popularity of network science in archaeology, coupled with the current uncertainty surrounding the role network science could play within the archaeological research process.

It is easier to write of the necessity of a PhD project than to prove this statement. This is not a straightforward process and throughout the following chapters I will evaluate different aspects of my argumentation through literature review and case studies. However, arguably the most crucial starting point is the definition of network science I have adopted in this thesis. In the rest of this introduction I will introduce this definition and argue how it implies the potential of an innovative method for archaeology. This will lead to a number of more specific research questions that require evaluation. The introduction is concluded with a description of the structure of this PhD and my approach to achieving this research agenda.

1.2. What is network science?

To argue that network science holds the potential to make unique methodological contributions to archaeology implies that one believes network science is somehow different from other methods already present in the archaeologist's toolkit. However, network perspectives themselves are by no means a unified body of theories and methods that perfectly complement each other. Network perspectives range from the highly quantitative to very qualitative, from working on a small scale to those functioning best on a large scale, from scientific to philosophical, and from an explicitly present-day perspective to attempting to recreate past perspectives. Each one of these is a valid way of thinking about past human behaviour and behavioural change (or rather every configuration or combination of these perspectives). This diversity of perspectives runs the risk of fostering the false impression that network theories are somehow fundamentally different from network methods, that network "thinking" and network "doing" can be easily separated. As with any research perspective, however, there can be no "doing" without "thinking" and vice-versa. Although this sounds like an obvious point to make, I have observed through the literature review presented in chapter 2 that network theories and methods have often been applied individually with little discussion of the implications of doing so. I will also argue below that what I understand as the fundamentals of network science makes it impossible to separate these two parts of the process.

Despite this diversity, network perspectives nevertheless share some common features, allowing one to talk about a science of networks. I will argue here that these features set network perspectives apart from others, and that they may be the source of the potential for a better understanding of the past. I will start with a descriptive definition of these features followed by a more formal definition.

Both social network analysts and physicists studying large empirically observed networks (the two scientific communities that have been most prominent in network science at large and that have been particularly influential to archaeologists (see literature review chapter 2)) agree that a few fundamental features and assumptions set network perspectives apart from others. In network perspectives, entities of research interest are never studied in isolation. Instead it is assumed that the relationships these entities are engaged in are fundamental for understanding their opportunities and behaviour. These entities could be anything the researcher considers interesting within the context of their research aims: a technological innovation, molecules, neurons in the brain, objects, individual humans, archaeological sites, islands, modern countries, or even entire planets. Note how the physical size of the entities does not really matter. Indeed, they do not even need to be physical. Anything can be usefully considered an entity of research interest if it allows the researcher to answer their research questions. This feature has led some to argue for the potential of network perspectives to cross different (spatial, social, conceptual) scales of analysis (e.g. Knappett 2011). The relationships between these entities could be equally diverse: a recorded action of transmission, spatial proximity, a physical connection such as a road, friendship, political alliance, being a member of an institution, the presence of an object on a site or the morphological similarity of objects. Network perspectives have the flexibility to incorporate multiple entities and relationships within a single research framework, as long as the researcher considers it useful to do so. The above definition implies that network perspectives not only aim to trace patterns of relationships between entities, they aim to explore the implications of doing so. To give an example inspired by classical antiquity: in order for a Roman senator called Cicero to climb the Roman political ladder and become consul, he is not solely dependent on his life-story, his family history, or his wealth; equally important are his political alliances and his popularity among the different communities of voters, since their actions and alliances will affect Cicero's political opportunities and will influence his actions (for a network analysis of Cicero's social network see Alexander and Danowski 1990; Brughmans 2012).

This definition also implies that entities and the way in which they relate will always need to be definable, their conceptual and/or physical boundaries will need to be tied down (if only just hypothetically or temporarily). This assumption does not mean concepts with fluid boundaries cannot exist, that concepts cannot have multiple meanings, or that concepts cannot change their meaning, performance or nature through time. It is merely an assumption that simplifies aspects of a complex world because it is useful to do so, and because by doing so the network perspectives allow one to better understand this complexity (such assumptions that simplify a complex real-world feature are present in almost every workable theory or method).

This descriptive definition already highlights some of the defining features of network perspectives. However, it remains to be shown why network methods cannot function without theories and vice versa. I find the more formal definition of network science formulated by the editors of the journal '*Network Science*' in their introduction to the first issue is a particularly suitable way to explain this (Brandes et al. 2013). Network science is seen as the "study of the collection, management, analysis, interpretation, and presentation of relational data" (Brandes et al. 2013, 2). This simple definition explains why so far I have refrained from using the term *network analysis*, or indeed *social network analysis*. Network perspectives allow for far more than merely the *analysis* of relational data, and as I have argued above they are not just concerned with social entities. It is, therefore, more suitable to refer to the collective of approaches that share these fundamental features as network science, or more generally as network perspectives. Brandes and colleagues further refine their definition of network science by stating it "is the study of network models" (Brandes et al. 2013, 4). They see network models as consisting of different elements and processes as illustrated in Figure 2, which can be considered an abstract representation of the key features of a network science research process. As such, it is a useful formal tool for further explaining the key combining features of network perspectives.

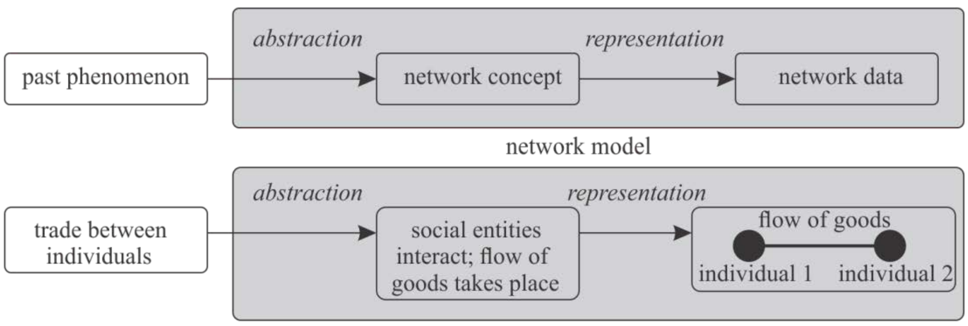


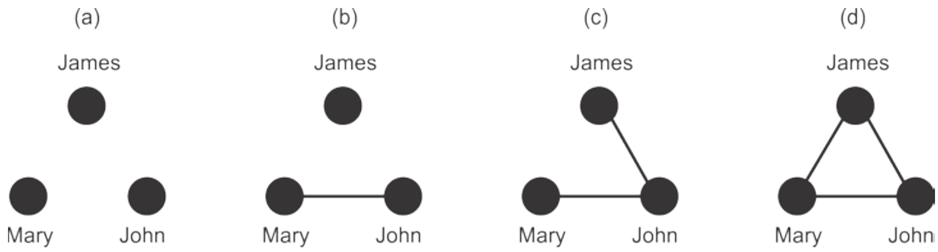
Fig. 2. At the top, an abstract representation of a network model (adapted from Brandes et al. 2013, Fig. 1). Every network perspective for the study of the past includes these elements and processes. At the bottom, an example of a network model to explore a particular phenomenon.

Archaeologists as well as historians aim to understand past phenomena, whether they are past networks of some sort that are hypothesised to have existed (e.g. a road network or a social network, e.g. see case study 3 in chapter 6), or aspects of human behaviour that translate less straightforwardly into network concepts (e.g. communication and trade, e.g. see case studies 2 and 3 in chapters 5-6). In figure 2, the past phenomenon we are interested in is clearly separated from the perspective we use to understand it. This highlights an epistemological issue every research perspective struggles with and that is particularly critical with network perspectives since it uses intuitively appealing concepts for the study of the past: the past networks we are interested in should not be confused with the network perspective, concepts and data we use to understand it (Knox et al. 2006; Riles 2001; Isaksen 2013). Instead of drawing a one-to-one relationship between the past phenomena and networks, the network perspective requires us to go through a process of abstraction in terms of network concepts and a process of representation as network data.

This model represents a common research process and is by no means unique to network perspectives. Yet I believe there is a need to emphasise this here, since the literature review presented in chapter 2 shows that sometimes these steps are not given much explicit thought, or in some cases there is evidence of a lack of awareness of their existence in archaeological network studies. Moreover, the model presented in figure 2 enforces a directionality to the research process that is commonly reversed in archaeology: archaeologists often start by describing the archaeological record and from there try to evaluate what phenomena can be addressed through this data. Both directions (from phenomenon to data, and from data to phenomenon) are valid approaches and I believe they never exist in isolation and, therefore, influence each other. However, I will argue throughout this thesis that following a research process as the one described in figure 2 reveals the untapped potential of network science for archaeology, more so than the reversed process of working from data to phenomenon that is more common in archaeology. Nevertheless, every scholar working with a network perspective does network modelling. The unawareness of the network modelling process is what leads to the false impression that network theories and methods are easily separated; as if a network analysis of archaeological data could be performed without reference to a theoretical network perspective, or as if a theoretical assumption about the structure or functioning of a past phenomenon could be formulated without any reference to data (or in the absence of empirically observed data, the formulation of specifications of what this expected data would look like). On the one hand, a more methodological network approach is aimed at testing a hypothesis, which cannot be achieved without thinking of the past phenomena as network concepts and formalising assumptions about how relationships affect the behaviour of entities and the evolution of the network. Alternativelly, it is aimed at representing archaeological or historical sources as network data, where the interpretation of the results in particular requires an inverted network modelling process to obtain useful insights about past phenomena. On the other hand, a more theoretical network approach might allow one to describe the hypothetical structure of a past phenomenon or the processes functioning as driving mechanisms of network evolution or node behaviour. Either way, to evaluate the probability of the hypothesis a representation of the network concepts employed will be necessary (to compare with empirical observations or with simulations of the hypothesis). Only in cases where falsification of hypotheses is not considered necessary we can argue that there can be network theory without representation as network data (although then a representation of a hypothesis as abstract network data can still be considered useful for communication purposes). Most recent archaeological network studies seem increasingly aware of this (e.g. case studies in Knappett 2013a). According to this definition one practical consideration becomes particularly important: the need to clearly define the archaeological research question or hypothesis, the entities and relationships that are considered, as well as stating the scholar's assumptions of how these relationships affect the behaviour of entities.

One crucial issue remains unexplained: what is network data? How exactly do we represent not only entities and relationships, but also the assumptions that govern their behaviour? Once again I will follow the reasoning of Brandes and colleagues (2013), who argue that network data are different from other types of data (they consider unrelated entities in data tables, dyadic data, and network data), or at least it is useful to consider it differently from other data types. This is because of the nature of the assumptions one makes about the dependencies existing in different data types. These so-called dependence assumptions in network data will be discussed in practice in more detail in the second case study in chapter 4. For now, it suffices to explain the difference between network data and other types of data through two examples: one social and the other spatial.

In a first example we consider the social relationships between a group of individuals, both men and women (Fig. 3). When we consider unrelated entities in data tables we can collect some information about these entities which is represented as attributes (Fig. 3a): we know one individual is female and two are male; the female is called Mary whilst the males are called John and James; Mary is an archaeologist with an income of X, John is an archaeologist with an income of X, and James is an accountant with an income of Y. If we just consider these individuals on their own we can already formulate assumptions of how some attributes depend on each other. For example, Mary's and John's salaries may be dependent on their job as archaeologists. We could also consider dyadic data, which assumes some attributes are only relevant to pairs of individuals and not merely to those individuals in isolation. Figure 3b represents Mary and John as a dyad, both are now connected because we know they are a couple. Here we make a new type of dependence assumption that the 'couple' attribute of both Mary and John is dependent on each other: their relationship cannot be understood by merely the attributes of either one of the partners. Finally, in figure 3c we have network data, where every line represents a friendship relationship (i.e. Mary and John are friends; John and James are friends). In network data we make the additional dependence assumption that one relationship will affect the existence of another relationship. It is up to the researcher to decide what the nature of this assumption is, depending on their research questions. For example, one might assume that a pair of unrelated individuals who have a common friend are more likely to become friends themselves at some point in the future. In this case our assumption about the friendship relationship between John and James might give rise to the friendship between Mary and James in the future (as in Fig. 3d). However, these three people do not live in isolation, they are part of a wider community whose friendships are likely to also be governed by the same assumption (Fig. 3e). This example illustrates how the dependence assumptions formulated by scholars might be used to represent network evolution in network data.



	Gender	Profession	Income
Mary	*F*	*Archaeologist*	*X*
John	*M*	*Archaeologist*	*X*
James	*M*	*Accountant*	*Y*

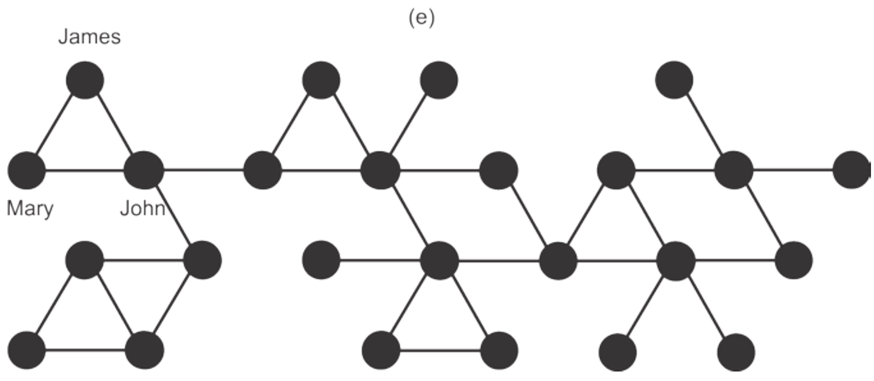


Fig. 3. Different data types representing individuals and social relationships: (a) individuals in isolation, characterised by their attributes which can be listed as a data table; (b) dyadic data, pairs of individuals are characterised by attributes that can only be understood with reference to both individuals (e.g. a romantic relationship); (c) network data, representing friendship relationships. When one formulates the dependence assumption that a pair of individuals who have a common friend might become friends in the future, then (d) can be seen as a future development of (c); (e) network data, showing Mary, John and James as part of a wider social context governed by the same dependency assumption as in (c).

In a second example we could consider a road network between towns A, B and C. In figure 4a we see that roads connect town A with town B and town B with town C. This means that all road-bound traffic between towns A and C will need to pass through town B. We could imagine a scenario in which the inhabitants of town B levy a toll on traffic through their town, or that the direct distance between A and C is shorter than passing through B, or that the inhabitants of C have family ties with the inhabitants of town A. If we consider any one of these scenarios it will be likely that this network will change into that of figure 4b, where a new road is built connecting towns A and C. The researcher explores their hypotheses of the processes that govern these scenarios by formulating assumptions, in this case that a direct road between A and C will emerge in the future with higher probability if the road via B becomes unappealing for some reason. Such an assumption does not merely change the structure of the network, it also affects the opportunities of the nodes involved. The traffic passing through town B will presumably decrease, which might affect the commercial opportunities of its inhabitants.

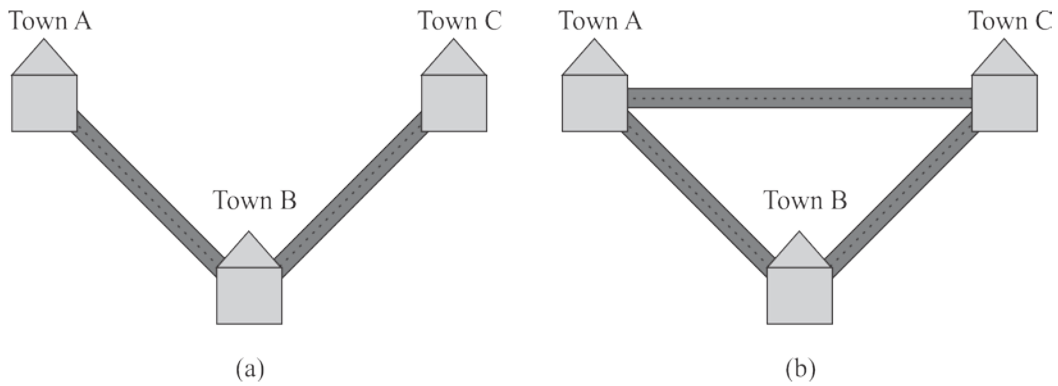


Fig. 4. A representation as network data of towns and the roads connecting them. The assumption that the road from town A to town C via town B becomes unappealing (for whatever reason the researcher considers meaningful) will increase the probability the network (a) will evolve into network (b). This change will affect the opportunities of each town in terms of controlling the flow of resources (goods, people, information), and might in turn trigger network change.

In network data our assumptions about what it means for a pair of nodes to be connected are explicitly expressed. They might affect the existence or absence of all possible relationships in the network and represent our hypotheses of how the structure of the network changes over time. The importance of dependence assumptions formulated by researchers to address their research questions and how these are represented in network data are fundamental for understanding the potential of network science. In the study of the past, network representation can never take place in laboratory conditions, in an interpretative vacuum if you will. Scholars always make assumptions about what relationships between entities mean and what kind of behaviour they allow for. This is exactly what, I believe, makes network perspectives different from other perspectives and why it holds potential for the study of the past: we formulate and represent our assumptions about what it means for nodes to be engaged in a relationship, and by doing so we can evaluate how these relationships lead to observed network patterning, how they affect nodes' opportunities and behaviours, to better understand the processes driving structural changes in the network.

Box 1. Summary: what is network science?

Network science is (see Brandes et al. 2013):

- based on an assumption that the relationships entities are engaged in are fundamental for understanding their opportunities and behaviour;
- the study of the collection, management, analysis, interpretation, and presentation of relational data;
- the study of network models.

Network models constitute:

- the research process whereby the phenomena under study is abstracted in terms of network concepts;
- network concepts are in turn represented by network data, or specifications are formulated of what network data of defined concepts would look like.

In network data dependence assumptions exist between:

- values of attributes of individual entities (as in tabular data);
- values of attributes of pairs of entities (as in dyadic data);
- tie variables (i.e. the presence of one tie might influence the presence of another).

1.3. Why does archaeology need network science?

There are multiple answers to this question, some more convincing than others. Firstly, the last decade has seen a surge in the use of formal network methods in archaeology, as will become evident from the literature review (chapter 2). This increased popularity might be an indication that archaeologists are aware that their existing methodological approaches are not sufficient to answer some of the questions they are interested in. However, the popularity of a method alone is not an argument for it being innovative and useful. Secondly, this surge goes hand in hand with the increasing availability of big data in archaeology, more powerful computers, and user-friendly network science software. Archaeologists have turned to network science as one of the approaches that becomes more useful when applied to large datasets. But just the ability to do something we could not do before does not mean it makes for a good argument for why it should be done. I find the use of fractals in archaeology a very good example of this problem, where it is clear that fractals offer a new approach that has proven its value in other disciplines, but we are also still waiting for a convincing archaeological use of fractals that has led to new insights into the human past (for a review see Brown et al. 2005). Thirdly, the increase in the use of network science in archaeology is not just guided by blind optimism but has also given rise to a number of critical reviews, listing a large number of challenges archaeologists face when adopting network techniques developed in other disciplines for different purposes (Isaksen 2013; Knappett *in press*; Sindbæk 2013). As I mentioned above, most crucial among these challenges is that the position of network science in the archaeological research process is not well understood. These challenges give me a reason to evaluate network science for archaeology, but they do not offer an argument as to its usefulness. Equally, there have been claims that formal network methods offer an approach that is metaphorically close to the way archaeologists commonly abstract past phenomena, and that they offer a degree of flexibility in formally analysing these phenomena. I believe this argument is suitably summed up by Carl Knappett (2011) in his recently published book '*An archaeology of interaction: network perspectives on material culture and society*', where he writes of the advantages of networks:

1. "they force us to consider *relations* between entities"
2. "they are inherently spatial, with the flexibility to be both social and physical"
3. "networks are a strong method for articulating scales"
4. "networks can incorporate both people and objects"
5. "more recent network analysis incorporates a temporal dimension" (Knappett 2011, 10).

These five arguments are part of the answer to the question of why archaeology needs network science. However, I claim that there is a more fundamental argument that incorporates all of the above and should form the starting point for identifying how network science offers a new and useful method for archaeology. This argument is adopted from the definition of network science given in the previous section, and it forms the central claim in this PhD project:

Archaeologists are confronted with network data; network science offers the suite of techniques necessary to deal with network data.

Network science is therefore not merely the *analysis* of networks or the study of *social* networks, as the popular term Social Network Analysis (SNA) implies. It is not useful just because it is a hot topic or because we have big data and computing power, nor does it merely concern the representation of archaeological data as network data. Rather, network science concerns the study of the management, representation, and analysis of network data, and network data represent our theoretical statements about why relationships matter. The starting point of my PhD is, therefore, the claim that network science only allows archaeologists to do something other methods do not succeed in doing when one is confronted with theoretical assumptions of how relationships affect each other that can be represented as network data. This claim will be explored and tested through a literature review and three case studies. The claim is purposefully formulated as an extreme hypothesis to offer clarity and focus to its evaluation in the thesis, although I recognise that the answer will not simply be “true” or “false” but rather “true is some cases, false in others”.

Before I continue it is important to stress that this PhD project aims to show the innovative aspects of network science in archaeology. However, I would like to emphasise three reasons why this focus should not necessarily drive *all* future archaeological network science. Firstly, network science should not be considered to replace an archaeological research process. Rather, it provides a set of techniques that might prove useful at different stages of the process. It should be clear that the abstraction of past phenomena into concepts is something all archaeologists do, and that it is archaeological theory and reasoning that should motivate the formulation of the assumptions why relationships matter. It is exactly this problem of where network science can contribute in the archaeological research process that I aim to address in this PhD project by a strict focus on its innovative aspects. Secondly, network science techniques might incorporate other techniques that are more frequently used by archaeologists, or they might be a minor part of commonly used methods. For example, a Harris matrix can be considered a network representation of the theoretical assumptions known as the laws of stratigraphy. Formal methods should be selected for their ability to perform necessary tasks no other method can do, and they can complement each other in such cases. Thirdly, there is no need to restrict the archaeological use of network science to merely its innovative factors. For example, the representation of archaeological data as network data and the use of exploratory network techniques can sometimes be a form of exploratory data analysis (EDA) as will be illustrated below in the third case study. It is true that statistical techniques can be used in some cases that might be better than network techniques at addressing a certain archaeological question with a given dataset. However, the process of representing archaeological data as networks, to visually explore them, and to be forced to think about relationships and their implications can sometimes lead to new insights and questions, even though it was not the innovative aspects of network science that led the scholar to this.

So far I have restricted my discussion of the potential of network science to a technical definition. Similar definitions could be formulated for other formal methods used in archaeology. For example, GIS is the study of the management, representation, and analysis of spatial data, where one assumes that spatial data are somehow different from other types of data and merits the development of a methodological toolkit dedicated to its study. However, identifying the contribution of network science to archaeology does not merely depend on arguments why it is different from, say, GIS because it deals with different data types. GIS is commonly used in archaeology because we frequently deal with spatial data, because we ask research questions that require the analysis of spatial data, and because archaeologists find visualisations of the spatial distribution of archaeological data useful for visual exploration and communication. Can similar arguments be made for network data?

1.4. Research questions

This PhD project will evaluate the abovementioned claim to novelty and the possible need for network science in archaeology by addressing this question, formulated as a series of more specific research questions.

1. Do archaeologists commonly ask research questions that are best addressed through the analysis of network data? What kinds of archaeological research questions are well suited to explore from a networks perspective? Does network science allow archaeologists to ask new questions?
2. Are archaeologists commonly confronted with network data? What data are typically well suited to submit to a network science approach?
3. Are network techniques and visualisations useful exploration and communication tools in archaeology?
4. What are the limitations of formal network methods in archaeology?
5. Do the typical research questions and datasets in Roman archaeology pose particular advantages or disadvantages?

Addressing these research questions will lead to a better understanding of the advantages and limitations of network science in archaeology, to a better positioning of network science within the archaeological research process, and will result in specifications that allow archaeologists to evaluate when network science techniques are suitable for addressing their research questions.

1.5. Why a Roman archaeology perspective?

It was decided to approach the aim of this PhD project through Roman archaeology case studies for a number of reasons. Most importantly, evaluating the potential of network science for the entire archaeological discipline would require the development of a much larger number of case studies addressing some of the possible particular challenges posed by different archaeological subdiscplines. Instead, it was decided to make the contribution of this PhD project more tightly focused around Roman archaeology case studies, whilst the more general implications of this work can be extrapolated to archaeology as a whole. However, the case studies were carefully selected to address a number of more general challenges of relevance to archaeology: the use of material datasets as proxies for human behaviour, creating similarity networks, dealing with long time-spans and limited chronological accuracy, working with bad samples. A more practical consideration was the availability of Roman archaeology datasets from the 'Urban connectivity' (see case study 2) and 'ICRATES' projects (see case study 3), my familiarity with these datasets through my education and work before the PhD, and the expert support these research contexts offered.

A Roman archaeology perspective also offers some advantages to this project. Firstly, in addition to the material data types available for most of archaeological periods, Roman archaeologists can also draw on literary sources. Although these are not directly the focus of this PhD project, they do allow Roman (and later periods) archaeologists to formulate complex research questions, and offer the information needed to suggest detailed hypotheses of past phenomena. Roman archaeologists frequently discuss and incorporate the role of individuals, communities, institutions, as well as materials in such hypotheses. This allows me to evaluate how network science techniques developed to study modern social networks can be modified or applied to incorporate the archaeological record. An example of this is offered by the complexity and diversity of hypotheses surrounding the study of the Roman economy, discussed in case study 3. However, the ability to ask complex research questions thanks to the diversity of data types has also in some cases led to strong differences in the research traditions of Roman and prehistoric archaeologists. This has resulted in some cases in research themes, more common in prehistoric archaeology, to be neglected in Roman archaeology, and creates the impression of radical changes between the Roman and prehistoric periods. Case study 2 offers an example of this, where visibility has been hypothesised as an explanatory variable for settlement location by Iron Age archaeologists, but has been almost completely ignored by Roman archaeologists. This case study will allow me to evaluate how network science techniques can help bridge the gap between different research traditions by drawing on both research themes and data from both prehistoric and Roman archaeology. Finally, the critical analysis of the diverse and large datasets and the formulation of complex hypotheses in Roman archaeology did not go hand in hand with methodological developments to deal with big data and to test hypotheses. Case study 3 illustrates the need for such methods most clearly. This offers the ability to evaluate the potential of network science techniques as both exploratory tools for dealing with large datasets, as well as confirmatory tools for testing or falsifying complex hypotheses in Roman archaeology.

1.6. The structure of this PhD

In line with my aims set out at the start of this introduction I believe that the research questions can be addressed through a literature review and archaeological case studies. This thesis includes three archaeological case studies. The case studies are selected for their ability to address some of these research questions, to evaluate possible particularities in the use of network science for Roman archaeology, and for their potential to allow the application of network techniques that have never before been applied in archaeology. The complexity of the network science techniques used will increase from one case study to another, i.e. the first case study introduces techniques and concepts applied in the second and third case studies, and so on. Each case study is complemented by boxed summaries, which clearly state the phenomena studied, the network concepts these are abstracted into, and how these are represented as network data. Therefore, by explicitly working through the network modelling process as described in the definition of network science introduced above in three independent case studies, I will be able to discuss the advantages, disadvantages and the diversity of network science approaches in chapter 6.

Chapter 2 provides a literature review of the archaeological use of formal network methods. It will trace back its roots to the first use of graph theory in archaeology in the late 1960s, as well as describe the more recent surge in the use of network methods stimulated by the popularity of SNA and complex network modelling in physics. This chapter will partly address the first two research questions by providing examples of archaeological research questions that are best addressed through network data. It will also address the fifth research question by reviewing the use of formal network methods in Roman archaeology. Moreover, this chapter will identify a large number of challenges archaeologists have been confronted with. It will, therefore, provide the framework within which the importance of addressing particular challenges within this PhD project becomes clear.

Chapter 3 presents the first case study: a citation network analysis of the adoption, use and adaptation of formal network techniques in archaeology. This case study can be seen as a network analysis of the literature review presented in chapter 2, using exploratory network techniques and visualisations. Indeed, this case study illustrates how an exploratory network analysis allows one to gain insights that cannot be obtained from a close reading in a literature review, and how some of the conclusions of the literature review can be identified or reproduced through the use of exploratory network techniques. I, therefore, consider it a particularly relevant example for addressing the third research question. It is also useful for introducing many of the formal network techniques in practice in this PhD. Unlike the other two case studies it does not involve any material culture or Roman archaeology research questions. Through the citation practices of archaeologists, which are in this case study considered as formal representations of academic influence, I will evaluate whether and how exploratory network analysis and visualisation can add new insights about the adoption of formal network methods to the close reading of a literature review.

Chapter 4 presents the second case study: understanding inter-settlement visibility in Iron Age and Roman Southern Spain with exponential random graph models for visibility networks. This case study will put the definition of network science and network data introduced above to the test most explicitly, by adopting a statistical modelling approach for simulating dependence assumptions surrounding visibility networks: theoretical assumptions formulated by archaeologists as hypotheses of why lines of sight between settlements matter. In doing so it will also for the first time introduce and evaluate a key method in network science for its potential in archaeology: Exponential Random Graph Modelling (ERGM). The approach taken here can be considered a transition from the exploratory approach taken in case study 2 to the confirmatory approach taken in case study 3, since it uses a statistical approach that very much relies on the availability of good empirical data. This case study will therefore address the first two research questions, as well as contribute to answering the fifth research question given its focus on Roman archaeology.

Chapter 5 presents the third and last case study: “Bang’s Roman bazaar and/or Temin’s market economy? An agent-based network model of tableware distribution in the Roman East”. In this last case study I focus on the potential of network science methods for performing confirmatory analyses by evaluating two complex hypotheses surrounding the functioning of the Roman trade systems. The two tested hypotheses concern the actions of individual social agents and their roles in giving rise to an archaeologically observed distribution pattern of tablewares. However, unlike in the previous two case studies it is less straightforward to identify the phenomena of interest in the dataset used and a confirmatory approach that draws on agent-based computational network modelling is applied instead. This case study will emphasise the importance of selecting the right network science technique for addressing archaeological research questions. Moreover, it will address the fifth research question most directly of all case studies by emphasising the importance of clear formulation of the conceptualisation and representation as data in Roman archaeology models, and of the largely unaddressed need in Roman archaeology for formal hypothesis testing.

Chapter 6 will draw on all previous chapters to systematically address all of the thesis’ research questions. A long list of methodological, data, spatial, and processual challenges identified in practice through the case studies or through the literature review will be discussed. This will be intertwined with a discussion of how the case studies presented here succeeded or not in overcoming some of these challenges, how they added new challenges to the list, and how this PhD suggests refocusing future efforts of overcoming the many remaining challenges archaeological network scientists face. The lessons learned through this PhD project will finally result in a list of suggestions for future archaeological network science: the start of a ‘guide to good practice’.

Box 2. Summary Chapter 1

Aim of the thesis:

- To make a methodological contribution to the archaeological discipline at large and to Roman archaeology specifically by evaluating the potential of network science as a method in archaeology.

The key argument:

- When archaeologists are confronted with network data a suite of techniques is needed that is designed for this purpose, these techniques are offered by network science.

Research questions:

1. Do archaeologists commonly ask research questions that are best addressed through the analysis of network data? What kind of archaeological research questions are well suited to explore from a networks perspective? Does network science allow archaeologists to ask new questions?
2. Are archaeologists commonly confronted with network data? What data are typically well suited to submit to a network science approach?
3. Are network techniques and visualisations useful exploration and communication tools in archaeology?
4. What are the limitations of formal network methods in archaeology?
5. Do the typical research questions and datasets in Roman archaeology pose particular advantages or disadvantages?

CHAPTER 2

LITERATURE REVIEW:

THE ROOTS AND SHOOTS OF ARCHAEOLOGICAL NETWORK SCIENCE

2. Literature review: the roots and shoots of archaeological network science

2.1. Introduction

Network science in archaeology has become more common in recent years. This statement is best illustrated by the recent increase in the number of archaeological publications using network techniques (Fig. 5). However, figure 5 also shows that network techniques have been used in archaeology at least since the 1960's. What were the methodological influences of these early adopters, and how did they in turn influence more recent archaeological network science? Is the use of network techniques in archaeology in the last decade fundamentally different from how it was used before? What were the motivations for archaeologists to adopt and adapt network techniques developed in other disciplines? The recent increase in the archaeological use of network science techniques might create the impression that most of these techniques have only been applied relatively recently in archaeology, or that the development of new techniques in the last decade made network science more useful for addressing archaeological research questions. In order to evaluate this impression one needs to go back to the roots of archaeological network science, to trace the multi-disciplinary developments which influenced archaeologists and stimulated the adoption and adaptation of network techniques. Through a close reading of published examples of the archaeological use of network techniques, I will illustrate that this impression is only partly true. I will reveal how very early archaeological applications introduced network techniques to our discipline that are still the most popular in more recent archaeological applications. This literature review will also highlight a number of challenges that many archaeological applications have in common, some of which I aim to overcome through the case studies in chapter 3 to 5. I start my review with the archaeological use of the branch of mathematics underlying much of network science: graph theory.

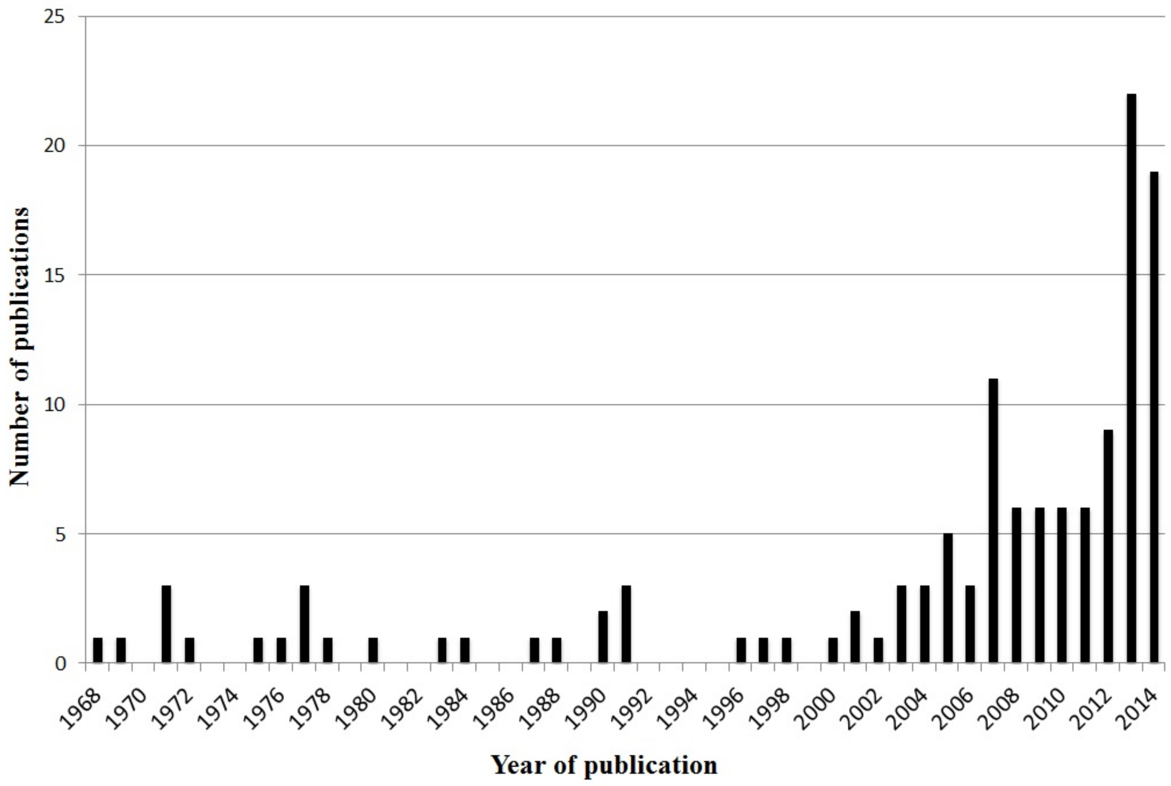


Fig. 5. Histogram showing the number of published examples of the archaeological use of network techniques. Network techniques were used sporadically since the late 1960's but only become more commonly used in the 2000's. The full list of publications included in this histogram is provided in Appendix I.

2.2. Origins in graph theory?

The branch of mathematics concerned with the study of graphs is called graph theory, which is generally considered to be one of the major mathematical foundations of SNA (Barnes and Harary 1983; Wasserman and Faust 1994, 15) and underlies much of the work done in network studies in physics. It is therefore not surprising that archaeological applications influenced by SNA and physics are also strongly rooted in graph theory. However, as I will illustrate in this section, it would be wrong to claim that formal network methods were introduced to the archaeological discipline through graph theory.

The reason why graph theory and networks are such a happy couple lies in the fact that graph theory not only offered network analysts a way to visualize networks as vertices and lines, a representation that came to dominate the way we think of networks nowadays, but it also brought with it a descriptive and mathematical system. Harary, Norman and Cartwright (1965, 3) described the potential of graph theory for SNA: (1) graph theory provides a vocabulary of concepts that can be used to describe properties of social structure, (2) it gives us a set of mathematical operations to quantify and measure these properties, and (3) given this vocabulary and mathematical operations it allows us to prove theorems about social structure represented as graphs (Wasserman and Faust 1994, 93). It is crucial to stress that in SNA graphs are used as models for social networks, which means that the nodes of graphs always represent social entities like individuals, communities or organizations and that the ties always represent relations with a social connotation like co-membership in organizations, kinship ties or proximity to social entities (Wasserman and Faust 1994). The social nature of graphs is an assumption in SNA that underlies the creation and interpretation of graph theoretic techniques developed by social network analysts. In network science in general the social nature of graphs is far less ubiquitous and less entangled with formal methods which, as I will illustrate below, is a key distinguishing feature between the two most influential network traditions.

Graph theoretic techniques have been used in archaeological research since at least the 1960's and have given rise to a number of interesting quantitative approaches to archaeological data. Very few of these influenced later archaeological network analyses directly, however, and in most the graph was merely used to visualize relationships rather than analyze them. For example, in Doran and Hodson's 1975 monograph titled '*Mathematics and Computers in Archaeology*' a graph was introduced as "a set together with a relationship which may or may not exist between each pair of its elements" (Doran and Hodson 1975, 13). The authors largely limit their excursion into graph theory by introducing a number of graph theoretic concepts rather than elaborating on the mathematics of specific graph theoretic techniques. It seems that Doran and Hodson were mainly interested in the graph as a visualization of archaeological relational data or concepts that stimulates visual exploration. A number of scholars have also used graph theory and matrices for seriation (Kendall 1969; Kendall 1971a; Shuchat 1984) and Santley (1991) used graph theory to explore aspects of Aztec regional economic organization. Clive Orton in his '*Mathematics in Archaeology*' (1980) did not elaborate on graphs explicitly, although he did suggest the graph as an alternative visualization for dissimilarity matrices (Orton 1980, 44-47). His dissimilarity graph introduced one of the key features of "graph drawing aesthetics" (Nooy et al. 2005, 14), namely that "each object can be thought of as a point in a space, closer to objects which are more similar … and further from objects which are less similar" (Orton 1980, 45). Unlike Doran and Hodson, however, Orton stresses one of the weaknesses of graph visualization by showing that relational space often cannot be represented using nodes and links without considerable simplification. The studies of Oxfordshire parish registers by David Kendall (1971b) and Robert Hiorns (1971) also make limited use of graph theory.

Both studies explored the degree of relatedness between parish communities in Oxfordshire using the same data sources. Robert Hiorns used marriage registers to investigate the effects on the relatedness of parishes' populations caused by movements between these parishes due to marriages (Hiorns 1971). The results of iterations of two mathematical models were visualized as graphs representing the hypothetical relatedness between parishes. These results were described and compared visually with graphs created from the marriage registers. David Kendall on the other hand explored the spatial relatedness between the same Oxfordshire parishes by using a multi-dimensional scaling algorithm known as MDSCAL (Kendall 1971b) to calculate the hypothetical location of villages whose location is no longer known (so-called *lost villages*) relative to the spatial location of known villages. John Terrell (1976; 1977a; 1977b) also used graphs to explore spatial relationships. Influenced by the geographers Chorley and Haggett (1967; Haggett 1965), Terrell developed Proximal Point Analysis (PPA) as a graph theoretical approach to think through interactions between island communities. This approach was later applied by Hunt (1988) for the study of exchange networks between Lapita island communities. In all of these examples graphs were largely used to visually compare results and to explicitly address interactions between people, data or places.

Mitchell Rothman's (1987) study of regional survey data from Middle Uruk south-western Iran is also largely restricted to a visual comparison of graphs. Compared to Hiorns and Kendall, however, Rothman attributes a more central role to graph analysis in his arguments by presenting graph theory as an ideal method for the analysis of settlement pattern data. Rothman sums up a number of advantages of graph theory which include statements like "elements of the structure of settlement can be described objectively and analyzed using simpler and more appropriate assumptions than those of many currently used models", "[graph theory] can deal with the magnitude and direction of the movement of goods, information, or people between individual sites in a settlement system" and "[graph theory converts] a variety of empirical detail of regional systems into mathematical matrices ideal for the flexible, verifiable analysis of system characteristics and for objective comparison with other patterns" (Rothman 1987, 74). Both the descriptive and the analytical power of graph theory are stressed but for neither of them are the author's arguments very convincing. Although Rothman is keen to point out the objective nature of graph theoretical techniques and the associated vocabulary, his discussion of what he calls "simpler and more appropriate assumptions" does involve a straightforward and seemingly restrictive social interpretation of graph theoretical concepts, which makes his use of graph theory very deterministic and simply prevents it from performing one of its main functions: comparing a variety of empirical data (the third advantage quoted above).

The graph theoretical work by the geographer Forrest Pitts (1965; 1979) on the Medieval river trade network of Russia seems to have been more influential to later archaeological network analysts (e.g. quoted by Isaksen 2007; Isaksen 2008; Peregrine 1991). Pitts was interested in exploring the connectivity of Moscow based solely on its position within the network of medieval trade routes to test the statements by Russian historians that the dominance of Moscow was at least in part due to its strategic position. His 1965 article is not very specific about the graph theoretic terms used. He does not define *connectivity* for example, but nevertheless suggests a measure for connectivity based on the graph diameter (the maximum number of steps between any pair of points in a connected network; Newman 2010, 140). In his 1979 article Pitts modified his method to calculate what are essentially the *betweenness centrality* values of towns along the river trade network. At this time he was a prominent member of the still very young SNA community and it is therefore more likely that the influence of Pitts' early graph theoretical work on archaeological network analysts was the result of the influence of SNA, rather than graph theory on the archaeological discipline.

One of the archaeologists influenced by the work of Pitts, Rothman and Irwin-Williams (discussed below) was Peter Peregrine who explored the evolution of the prehistoric centre Cahokia along the Mississippi, Missouri and Illinois rivers by applying "the graph theoretic concept of centrality" (Peregrine 1991, 68). Peregrine aimed to mathematically test the hypothesis proposed by other archaeologists that Cahokia evolved into a major center thanks to its position near the confluence of major rivers, which allowed it to exercise control over riverine exchange in the Mississippi Basin. For this purpose he visualized the rivers as a graph where nodes represented river heads and junctions, and ties represented the rivers themselves. Peregrine used three centrality measures, as developed and described by the social network analyst Linton Freeman (1979), to analyze his graph and Cahokia's position on it. Peregrine therefore makes use of both graph visualization and analysis techniques, contrary to most of the studies described above.

The earlier article by Pitts, along with the network models for geography described by Chorley and Haggett (1970), dominated the graph theoretical techniques applied to Geoffrey

Irwin's (1978) study of the development of a Papuan settlement and interaction system. Irwin was interested in exploring the role of Mailu Island as a manufacturing and trading centre, which in 1890 AD had an atypical and more prominent economic development compared to other sites in the study area. He assumed that an effective communication network was of importance and decided to use graph theory, alongside other techniques, to explore consecutive hypothetical versions of this network for Mailu's prehistoric period (before 1890 AD). The centrality of nodes on these hypothetical networks was explored using the connection-array connectivity (the total number of alternative paths from a node) and short-path array connectivity (the path with a minimum number of links) measures as introduced by Pitts (1965) and discussed by Haggett (1970, 636-637). Nodes were ranked according to the results of these measures, suggesting that Mailu was mildly more prominent than other sites but not as much as its clearly advantageous position in 1890. The connectivity results were compared with a measure of accessibility by weighting the same networks with the actual distances between nodes, which seemed to lead to similar hypothetical inferences about the centrality of Mailu. Rather than having any predictive value, Irwin argues that the strength of this graph theoretical approach lies in making explicit and exploring the structure of an archaeological hypothesis. This study by Irwin makes clear that decisions made during the creation of networks dominate the choice of graph theoretical techniques as well as the usefulness of the results one can expect.

The work by anthropologist Per Hage and mathematician Frank Harary is exceptional since it concerns a multi-disciplinary collaboration and the adaptation of graph theoretic techniques to address anthropological research questions. This is rare and did not occur again to my knowledge until the multi-disciplinary collaboration between archaeologist Carl Knappett and physicists Tim Evans and Ray Rivers (discussed in more detail below). Moreover, the work by Hage and Harary wished to challenge the use of graph theory as merely a visualization technique, as in most of archaeological work described in this section: "We wish to emphasize right at the outset that the ultimate value of graph theory for anthropology will depend not just on the use of its pictorial representations, but also on the application of its theorems" (Hage and Harary 1991, 2). The duo was successful at putting this statement into practice. A good example is their discussion of the anthropological use of the mathematical minimum spanning tree problem and their application in a case study of the Lakemban *matanitu* or chiefdom (Hage and Harary 1996). The Lakemban *matanitu* is described as "a hierarchically structured island network based on relations of kinship, alliance, and conquest" (Hage and Harary 1996, 70). The authors argue that the origin of this network should be seen as a process of growth that can be best represented as a graph theoretic algorithm. In particular, they believe that the hypothesis by Thompson (1940) of the origins of the Lakemban *matanitu* implies the minimum spanning tree algorithm of Boruvka (1926). Thompson (1940, 214-215) writes: "gradually the small, poor islands became dependent upon the larger, richer islands like Lakemba and Kambara. There arose small chiefdoms, within which the weaker islands stood in tributary relationship to the stronger". Boruvka's algorithm represents a step-by-step process for creating a unique minimum spanning tree graph between a set of points following simple rules: small trees of minimum value are built and then joined.

The books and articles by Hage and Harary (1983; 1991; 1996) present a wealth of such examples, and I believe their work can be seen as an earlier attempt to critically explore the potential of network science (although they called it graph theory) for the anthropological/archaeological discipline: Hage and Harary acknowledged that theoretical dependence assumptions are commonly formulated in anthropology/archaeology and that these can be formalized and tested using graph theory. However, their work had very little following in anthropology and archaeology, and even in Pacific archaeology which was Per Hage's field of expertise (Cochrane pers. comm.; Irwin pers. comm.). One could argue that the mathematical detail and the scientific process of hypothesis testing made their work less accessible or appealing to archaeologists. Indeed, their work is more commonly cited in SNA than in anthropology/archaeology, and I believe this is where the real reason for their limited exposure in our discipline lies. Some of their work was more concerned with refining existing graph theoretical algorithms or developing new algorithms, which are perfectly valuable research objectives in mathematics and SNA. Although archaeological case studies invariably served as a starting point and inspiration for this work, the added value of their work for the archaeological discipline was not always clear. For example, Hage and Harary (1996, 75-87) critique the Renfrew-Sterud method of close-proximity analysis (Renfrew and Sterud 1969), arguing that the method is too complex, that it should be performed by a computer, and that there is not just one but many minimum spanning trees that can be created from the data used by Renfrew and Sterud. However, the alternatives suggested by Hage and Harary did not lead to a reinterpretation of Renfrew and Sterud's results, and the contribution of this case study to the archaeological discipline seems limited to a cautionary methodological note when using minimum spanning trees (Hage and Harary 1996, 84). I believe Hage and Harary should be considered pioneers in archaeological network science (see also tribute to Hage by Jenkins 2008) and this PhD project in particular should learn from their experiences: multi-disciplinary collaboration enables critical and innovative applications, but the contribution to our understanding of past phenomena should be clear to an archaeological audience.

All of the archaeological studies discussed in this section used graph visualization or analysis techniques for different purposes and with varying success. Most of these graph theoretic applications, however, show similarities with SNA (excluding studies on seriation) by stressing the importance of attaching explicit social assumptions to graph theoretic concepts. Rothman, for example, introduced graph theory as a subset of network analysis (Rothman 1987, 74). It is not clear if he was referring to work by the growing SNA community, but the social interpretations he attributes to his graph theoretic vocabulary seem to indicate that he was at least thinking in terms of past social networks. Although Peregrine considers his work to be graph theoretical it is clearly influenced by developments in SNA through the works of Freeman (1979), Hage and Harary (1983) and Pitts (1965; 1979). However, none of these early archaeological applications seems to have had a significant impact on later archaeological network-based research.

This section on graph theory raised three issues:

1. the research potential of graph theory as an alternative approach for the visualization and analysis of social or geographical hypotheses in archaeology has been recognized at least since the 1960’s;
2. in spite of the obvious similarities in approaches and the relevance to archaeological network analysts, the research potential illustrated by early graph theoretical work in archaeology has not been very influential to more recent network applications in the discipline;
3. as a result, the introduction of graph theory and SNA into the archaeological discipline happened largely independently and, unlike social network analysts, archaeologists did not collaborate with graph theorists to develop mathematical techniques tailored for their needs (the work by Hage and Harary is a notable exception). The specific graph theoretical techniques underlying network-based work in archaeology were developed in SNA and physics, and adopted into the archaeological discipline without much reference to their graph theoretical roots.

2.3. Social network analysis

Many archaeological network scientists have been strongly influenced by SNA, most of whom performed their SNA-related research only within the last ten years (e.g. Graham 2006a; Hart and Engelbrecht 2011; Isaksen 2007; 2008; Jenkins 2001; Mills et al. 2013; Mizoguchi 2009; Munson and Macri 2009). SNA has a long history throughout which a very large variety of network-based methods and applications was developed. This diversity is not reflected in the archaeological literature and I will argue that it is worth exploring this since it might lead to original and valuable archaeological applications. In this section I will briefly introduce the development of SNA and discuss some popular or promising research themes and techniques.

2.3.1. An introduction to SNA

Social network analysis developed as a major research perspective in the social and behavioral sciences from its precursor, sociometry, which involves the measurement of interpersonal relations in small groups and was founded by Jacob Moreno after his invention of the sociogram in the early 1930's (Moreno 1934; 1946; 1960; Moreno and Jennings 1938). The sociogram is a means for depicting the interpersonal structure of groups as nodes and edges in two-dimensional space, like graphs. According to Linton Freeman (2004, 30), sociometry "was the first work that included all (…) of the defining features of social network analysis". Later social network analysts built on Moreno's work as well as on the pioneering efforts by a group of Harvard scholars in the late 1920's to the early 1940's (Freeman 2004, 43-64). Graph theory, statistical and probability theory, and algebraic models in particular found a place early on in mainstream social network methods (Wasserman and Faust 1994, 10-17). SNA methods and applications have been further formalized by a number of extremely influential books throughout the last two decades (Carrington et al. 2005; Scott and Carrington 2011; Wasserman and Faust 1994), with contributions being largely limited to a dominant group of key players in the SNA community. The evolution of formal network methods within the SNA community is documented in the journals *Social Networks* (Elsevier) and *Connections* (INSNA), both first published in 1978.

Wasserman and Faust (1994, 4) have formulated a list of principles shared by SNA applications that clearly specifies the extent of the social assumptions of SNA:

- "Actors and their actions are viewed as interdependent rather than independent, autonomous units

- Relational ties (linkages) between actors are channels for transfer or ‘flow’ of resources (either material or nonmaterial)
- Network models focusing on individuals view the network structural environment as providing opportunities for or constraints on individual action
- Network models conceptualize structure (social, economic, political, and so forth) as lasting patterns of relations among actors”

These principles illustrate the main difference between SNA and other network-based approaches, namely a restriction to social units as well as its implications. It is concerned with exploring social relationships as media for the flow of resources between active individuals, corporations or communities. The focus on social entities in a networks perspective has proven useful for addressing a wide range of research questions in the social and behavioral sciences. Wasserman and Faust (2004, 5-6) provide a list of topics network analysts are traditionally interested in, including the diffusion and adaptation of innovations (Rogers 1979; Rogers 1995; Valente 1995; Valente 2005), belief systems (Erickson 1988), markets (White 1981), exchange and power (Markovsky et al. 1988), and occupational mobility (Breiger 1981).

2.3.2. An early archaeological model

From the above it becomes clear that formal SNA methods have been around in some form or other since at least the 1930’s and more coherently since the 1970’s, yet it seems that archaeologists have only recently been interested in using SNA in their own research. This late adoption becomes even more striking when one considers how prominent anthropologists were in the SNA communities prior to 1970 and even after that (Freeman 2004; Johnson 1994; Mitchell 1974; Wolfe 1978; 2011). These anthropological network studies addressed many research themes that are of great interest to archaeologists (for reviews see Johnson 1994; Wolfe 2011) and may have stimulated archaeologists to adopt SNA, especially in the US. There is at least one notable exception to this trend, however, which I believe might help us understand the limited use of SNA techniques in archaeology before the 2000’s.

The potential of SNA for archaeology was clearly recognized no later than 1977 in Cynthia Irwin-Williams’ (1977) network model for the analysis of prehistoric trade. She argued that in archaeology the treatment of the exchange of material goods and services has tended to be simply descriptive and that a network model might provide a quantitative framework for this subject. Irwin-Williams limited her model to the exchange relations connecting settlements (described as network points), although it could easily be applied at different levels of archaeological analysis. The author still provided no less than six examples of measures of linkages within contemporaneous archaeological networks: "(1) within assemblages from a given settlement, the presence or absence of objects originating at another point; (2) the proportion of specific exchange goods from a particular origin to local goods of the same class; (3) the proportion of goods of the same class originating at various different points; (4) the directional dominance of the flow of goods, that is, the import-export ratio between settlement points; (5) the number of classes of objects exchanged between points; (6) the kinds of classes of objects exchanged between points" (Irwin-Williams 1977, 142-143). She goes on to suggest the seven network-based approaches that make up her model: (1) three "points of view" for exchange networks are given: global (the whole network), zonal (part of a network, defined in geographical, cultural or other terms), and anchored (a so-called ego-network focused on one point and its direct neighbors); (2) networks can be visualized as node-link diagrams and matrices; (3) the network density measure is introduced (this measure was adopted from Haggett and Chorley (1969) as well as Mitchell (1969)); (4) a "first order star" network is introduced as the ego and its direct neighbors, and a "first order zone" is introduced as the relations between all members of the "first order star" (a description adopted from Barnes (1972)); (5) she introduces uniplex and multiplex links as relationships through which one or more classes of goods may circulate (adopted from Kapferer (1969)); (6) it is possible to differentiate zones with maximum internal linkage bounded by zones of relative low density and few multiplex relations (an idea adopted from Kapferer (1969)); (7) an effective network is characterized by large channels, multiplex linkages, and relatively great density, whilst relations within an extended network will be more attenuated and probably more specialized (ideas adopted from Epstein (1969)). The author went on to argue for the potential of this network model as it might be applied to research on ancient Puebloan society in northwestern New Mexico, but she sadly did not elaborate on the results of this case-study.

Some of the types of relationships mentioned by Irwin-Williams have formed the basis of later archaeological network analysis (e.g. Brughmans 2010; Graham 2006b; Sindbæk 2007a; 2007b) and the network analytic approaches she suggested are now part of the core set of network techniques used by archaeological network analysts. The network "points of view" are at the very least implicit in most archaeological network analyses and graph and matrix formats are ubiquitous in these applications. I believe it is fair to argue that the ego-network approach, link multiplexity and the identification of *zones* with certain topological features, all introduced by Irwin-Williams, have only recently been given more attention (e.g. Munson and Macri, 2009). In light of all this it seems striking that SNA techniques have received so little attention from archaeologists and, even more remarkable, that the clear potential illustrated by Irwin-Williams' model has influenced so few archaeological network analysts directly (including Peregrine 1991; Rothman 1987; Branting 2007). The reason for this might lie in the limited availability of cheap and potent computing power and large digital datasets in the late 1970's. Although Irwin-Williams clearly illustrated the potential of a networks approach, she did not illustrate how this should be applied to complicated networks in a large dataset. Many of the network techniques she described provide rather obvious results when applied to smaller datasets and often only reveal their real analytical strengths when applied to large datasets. As a consequence the more widespread adoption of SNA by archaeologists was delayed until the necessary computing power was more generally available. However, this does not explain why the network concepts introduced by Irwin-Williams did not break through as more qualitative or small-scale approaches to think with. Knappett (2011, 17-18) suggests that one reason for the reluctance of the New Archaeology to adopt networks might be that connectivity was generally conceived as interactions at the borders of zones around sites rather than as concrete geographical connections between sites.

The work by Irwin-Williams illustrates the core argument of this literature review: the potential of formal network methods for archaeology was discovered decades ago, but only a limited segment of this potential has been explored so far.

The rest of this section will introduce SNA research themes and analytical techniques that have either been particularly influential to archaeological network scientists, or because they lend themselves particularly well to exploring archaeological data and addressing archaeological research questions.

2.3.3. Diffusion of material and immaterial resources

The diffusion of material resources and information might prove of particular interest to the archaeological discipline. According to the social networks perspective, social relations are channels of social contagion and persuasion, and as such instrumental to the diffusion process (Nooy et al. 2005, 161; Valente 2005, 98). Diffusion techniques focus on exploring the relation between structural positions of actors and the moment at which they adopt an innovation. Most interestingly, the structure of the diffusion of innovations shows similarities to the spread of an infectious disease: the number of initial adopters is very limited, then large numbers adopt, and finally the growth rate decreases. Typical examples of this are network studies that explore the structure of the world-economy (Nooy et al. 2005, 29-57; Snyder and Kick 1979). Calculating the density of social networks, the number and relationships of components, the centrality of key nodes, and the adoption rate allows network scientists to explore the structure of diffusion (Nooy et al. 2005, 161-183; Rogers 1995; Valente 1995). Some of these measures have been used in a study of Roman pottery distributed from place of production to place of deposition where we have stressed their potential for exploring large archaeological datasets and geographical hypotheses (Brughmans, 2010; and see also case study 3 in chapter 5). This study also made it clear that there can be no straightforward and standardized interpretation to the results provided by SNA measures and that these results therefore require a re-contextualization in a wider socio-political, archaeological and historical framework.

This issue is also very prominent in another archaeological example of the study of past diffusion processes: Shawn Graham’s (2006a) analysis of Roman itineraries. Graham created a network of towns connected by the routes between them as mentioned in the Antonine Itineraries (a collection of route descriptions within the Roman Empire). He went on to calculate the average shortest path length, the density (referred to as cohesion by Graham) and fragmentation curves[1] for a number of regions on this network. He concluded his SNA approach by stating that the structures identified have implications for how information was disseminated and he explores this aspect more explicitly through an agent-based model, which will be discussed further on in this article. Graham seems to be very much aware of the fact that the results of his SNA approach inform him of the structure of a particular data source, rather than the actual structure of past road networks. This implies, however, that the SNA techniques used are not necessarily linked to the study of past processes of the diffusion of information. In Graham’s study the results will at most reveal hypotheses on the spread of information as implied by this ancient source. The potential of network techniques for studying the spread of an innovation is further evaluated in the archaeological citation network analysis presented in case study 1 (chapter 3).

[1] The fragmentation curves represent the number of nodes that can be removed before the network falls apart in different components (unconnected parts of a network).

2.3.4. Network centrality

Centrality measures are arguably the most popular tools in the social network analyst's arsenal. These allow for the identification of nodes that have better access to information and enhanced opportunities to spread information because of their central position or role as a necessary go-between in a social network. In SNA, centrality is commonly applied in analyses of the structure of organizations (Michael and Massey 1997; Nooy et al. 2005, 123-137). The identification of key people in the communication network of an organization, for example, will help tailor an optimized business-specific flow of information (Burt 2011). Degree centrality, closeness centrality, betweenness centrality (Freeman 1979; Fig. 6) and eigenvector centrality (Bonacich 1972) are by far the most widely applied measures in archaeology (e.g. Bernardini 2007; Isaksen 2007; 2008; Jenkins 2001; Mills et al. 2013; Mizoguchi 2009; Peeples 2011a; Peregrine 1991; Phillips 2011). These have been recently extended with measures for group centrality and centrality in two-mode networks (Everett and Borgatti 2005), of which there are no published archaeological examples yet to my knowledge.

In their work on the socio-political interactions of the Classic Maya, Munson and Macri (2009) used a renormalized degree centralization calculation (Butts 2006) that incorporates the size and density of the network and can therefore be used to compare different networks. The eigenvector centrality measure also takes the overall network structure into account (Bonacich 1972; Hanneman and Riddle 2005, 68-70; Newman 2010, 169-172) and can therefore be considered a good example of how SNA techniques can contribute to the search for global structure in networks, something physicists are traditionally concerned with (see below). In a recent study of ceramic networks in the Late Hispanic U.S. Southwest, Mills and colleagues (2013) have argued that eigenvector centrality also provides a more accurate reflection of complex flow processes measured by similarities in ceramic assemblages since it assumes that each node affects all of its neighbors at the same time (Borgatti 2005, 62).

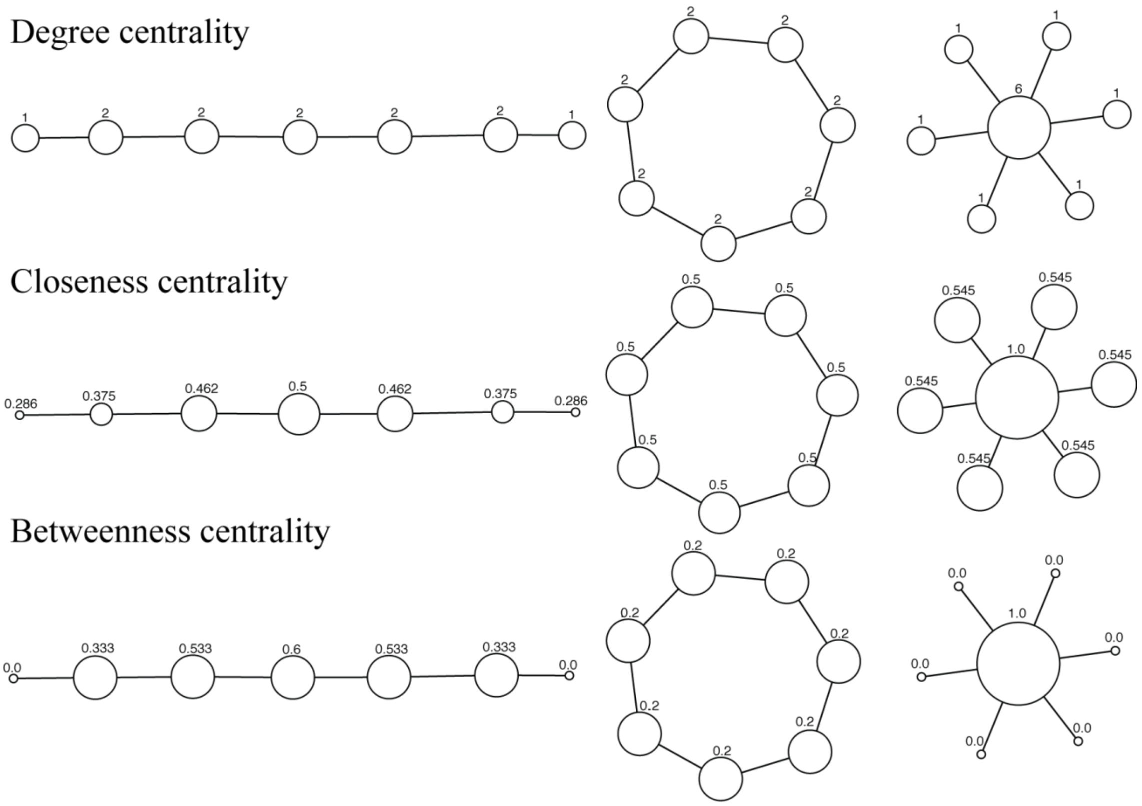


Fig. 6. An example of degree-, closeness- and betweenness centrality for three undirected network structures. Node size and labels indicate centrality values. Degree centrality is the number of relationships a node has, closeness centrality is "the number of other vertices divided by the sum of all distances between the vertex and all others" (Nooy et al. 2005, 127; Sabidussi 1966) and the betweenness centrality is the proportion of all shortest paths between pairs of other vertices that include this vertex (Nooy et al. 2005, 131).

Centrality measures have been used by archaeologists to explore properties of ancient transport networks (e.g. Isaksen 2007, 2008; Jenkins 2001) in a very similar way to Peregrine's (1991) earlier work discussed above. David Jenkins (2001) aimed to analyze the significance of locational advantage relative to the Inka road network, administrative centers, productive enclaves, and storage sites. A network of 54 administrative, storage or other sites connected by roads was created using published studies of the Inka road network as well as site reports and Spanish chronicles. Jenkins is very much aware that he is not studying the Inka road system itself directly. As I have stressed above for the work of Graham and Irwin, it is crucial to keep the network as a technique and as a past phenomenon clearly separated (Knox et al. 2006). Jenkins explored the structure of this hypothetical model using three centrality measures developed by Freeman (1979) and critically discussed their meaning in this specific archaeological context. To explore a hypothesis of wealth finance Jenkins (2001, 671-675) created a directed network representing the potential flow of what he defines as prestige goods from their origins at the periphery of the network towards the capital of Cuzco at the core. The latter network seems to be more a visualization of his hypothesis of wealth finance than anything else, given that the creation of this network was not discussed and he did not analyze it using centrality measures (which would require the use of modified graph theoretical algorithms).

Another example of the archaeological use of centrality techniques is Koji Mizoguchi's (2009) study of the emergence of a centralized hierarchy in Japan's initial Kofun period. Contrary to Jenkins, Mizoguchi compared the results of no less than six centrality measures: degree centrality, closeness centrality, betweenness centrality, eigenvector centrality, Bonacich power centrality and reach centrality. His descriptions of these techniques are taken from the SNA handbook by Hanneman and Riddle (2005) and he calculated them with the SNA software package UCINET (Borgatti et al. 2002). The author aimed to test the hypothesis that the relationships between social groups in the initial Kofun period were more significant to the emergence of interregional hierarchy than attributes of these groups, such as the dominance over the exploitation of raw materials. To this purpose he created two networks, one for the initial Kofun period and one for the earlier Yayoi V period. Sites were clumped together per region and represented as nodes. Mizoguchi drew edges between nodes according to the presence of non-locally produced prestige goods. The presence of non-local pottery and locally made pots according to non-local stylistic traditions in particular were considered evidence for interregional interactions. Although his use of centrality measures allows for an interesting and innovative evaluation of his archaeological hypothesis, his subsequent interpretation of the centrality results (see Mizoguchi 2009, 24) does reveal an issue related to the definition of network elements and their structural properties and the adoption of standard interpretations of network measures. The artifact distributions from which the network edges are created are assumed to be significant proxies for interregional interaction and socio-cultural dependencies. Using such hypothetical networks of things to explain the emergence of an interregional hierarchy involves a significant leap of faith, for which Mizoguchi relies entirely on the centrality measures. However, the authors from whom Mizoguchi adopted the descriptions of these centrality indices themselves stress that "the definitions of what it means to be at the center differ. It is more correct to describe network approaches this way -- measures of centrality -- than as measures of power" (Hanneman and Riddle 2005, 62. For a good example see Osa 2003). I would therefore argue that Mizoguchi successfully explored the structure of a hypothetical interpretation of a hypothetical network. The centrality measures allowed him to identify problems surrounding his hypothesis and data (Mizoguchi pers. comm.) rather than to test his very interesting hypothesis, let alone refute alternative hypotheses like dominance over resources.

2.3.5. Affiliation networks

A significant part of the social contexts in which individuals are embedded is shaped by their affiliations. When social network analysts examine affiliation networks, yet another popular topic, they assume that membership of an organization or participation in an event is a source of social ties (Fig. 7; Nooy et al. 2005, 101; Wasserman and Faust 1994, 30, 291-343). Directors of different corporations, for example, might share information and make professional decisions at gatherings of the clubs they are members of. Alternatively, academics might be influenced by the novel ideas of other researchers at the conferences they attend. Affiliation networks are traditionally visualized as two-mode networks. The use of two-mode networks is not restricted to affiliations, however, since modes could represent two sets of actors (Wasserman and Faust 1994, 39-40) or indeed any data types (e.g. Brughmans et al. 2012). A growing set of metrics to analyze two-mode affiliation networks is being developed (Everett and Borgatti 2005; Faust 2005). These techniques hint at the existence of the layered and heterogeneous nature of social relationships and could be considered a first step to exploring the complex web of interlocking contexts that make up social networks.

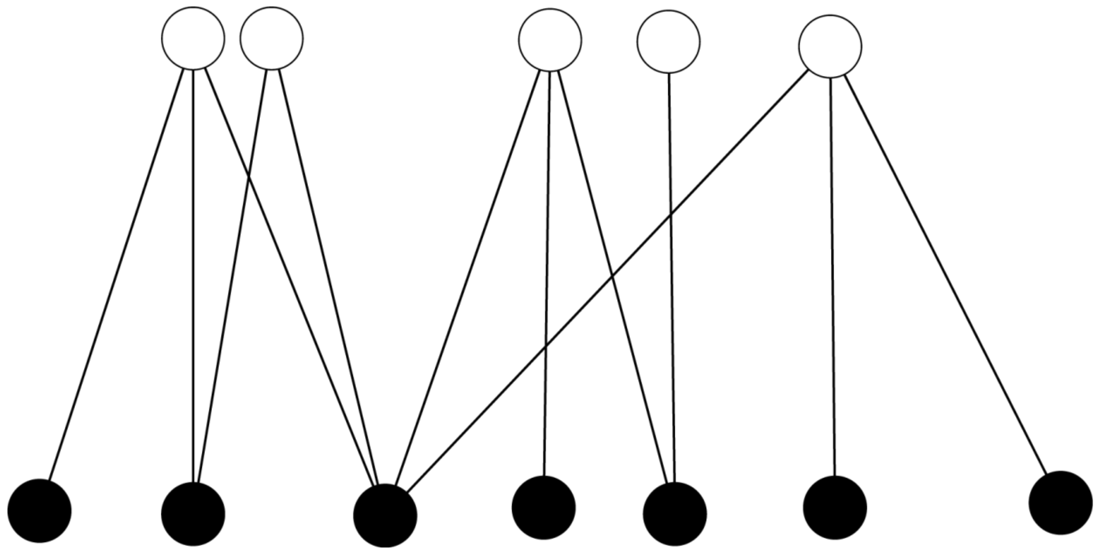


Fig. 7. Example of a two-mode network where white nodes can only connect to black nodes and vice versa.

Published archaeological examples of affiliation networks are few. However, the approach holds great potential for dealing with the complexity of past social interactions by mapping broad generic (e.g. known social, geographic or political entities) or small specific contexts (e.g. typologies, stratigraphic contexts) explicitly as affiliations. This potential is clearly illustrated by Carl Knappett's (2011) recent use of affiliation networks. On a more practical level it can also help archaeologists deal with networks of multiple data types. Phillips (2011) in his work on lithic raw material consumption in the Kuril Islands linked specific obsidian source groups with sites in affiliation networks. Phillips further explored the degree of association between the two classes in his affiliation networks by using correspondence analysis and the Jaccard similarity coefficient. Whilst these studies were limited to exploring relationships between sites and a single data type, Søren Sindbæk (2007b) in his work on Early Medieval communication and exchange networks illustrated how sites can be seen as affiliated to multiple data types in their assemblages. Sindbæk points out the exploratory nature of his approach, stressing that there is no direct relationship between shared artefacts and specific past processes (Sindbæk 2007b, 66).

2.3.6. Ego-networks

As a final example of social network analysis themes and tools I will introduce the concept of ego-networks as a technique to study the social environment surrounding individuals. An ego-centered approach focuses "on the position of one person in the network and his or her opportunities to broker or mediate between other people" (Nooy et al. 2005, 144; for early examples of ego-network applications see Boissevain 1973 and Bott 1957). To this aim, ego-networks are constructed consisting of one node (called the 'ego'), its neighbors, and the ties among them as in figure 8 (Hanneman and Riddle 2005, 8-9; Marsden 2002; Nooy et al. 2005, 145; Wasserman and Faust 1994, 41-43). This approach is particularly useful in situations where it is not possible to track down the full network (Hanneman and Riddle 2005, 8), because the data are just not available or because the full network is not relevant to answering specific research questions. In fact, the ego-network is a representation of the idea that individuals only have local knowledge of the social networks they are part of (Kleinberg 2000; Watts et al. 2002). By focusing on a single person, his or her direct relationships and the relationships among them, we can begin to explore how the direct social environment influences one from an individual's point of view. However, in no way does this ego-approach attribute an inherent simplicity to the process of influence and the evolution of entire social networks. Ego-networks can be seen as attributes of individuals, representing one of the many reflections of social contexts they are embedded in (Granovetter 1985; Knox et al. 2006, 118).

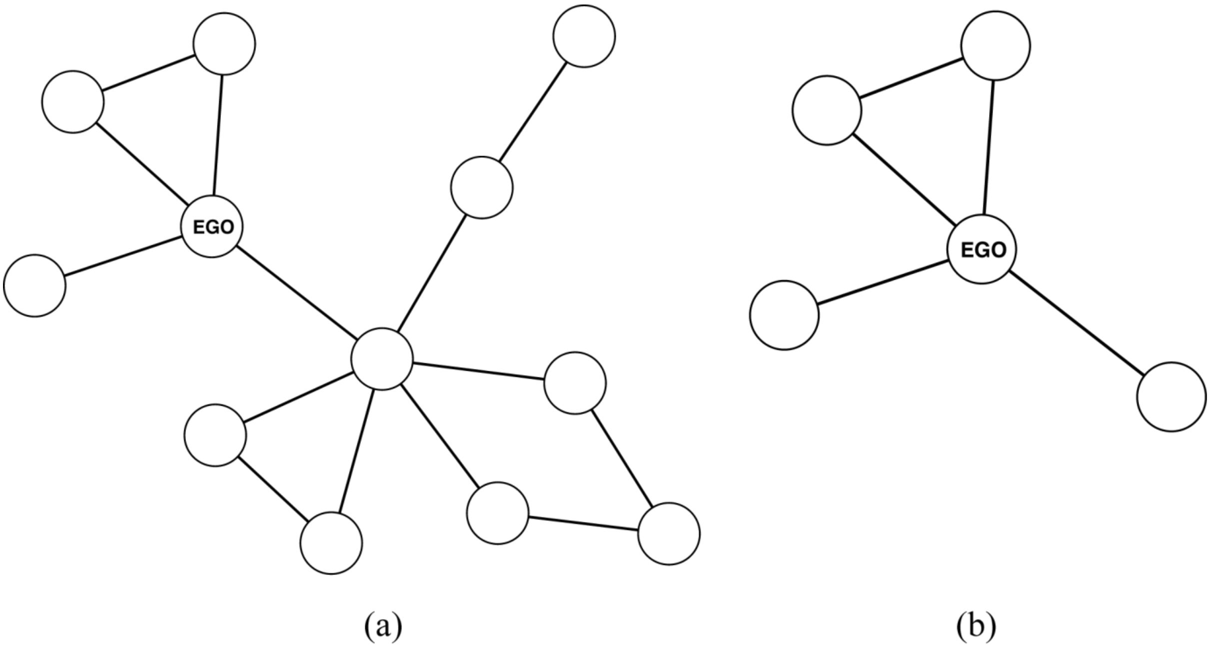


Fig. 8. Example of an undirected network (a) from which an ego-network (b), consisting of the ego, its direct neighbors and the edges between them, was extracted.

Published examples of the archaeological use of ego-networks are few (see Mol et al. 2014b for a rare example). I would suggest, however, that an ego-network approach is particularly promising for interpreting specific patterns or hypothetical processes from the bottom-up. Like scholars in many other disciplines archaeologists explore social relationships indirectly through the traces people leave behind. For historians these traces might be found in textual data and sociologists might use modern media like email correspondence. As archaeologists, our data are typically the material residues of individuals' actions influenced by local knowledge of the social networks they were embedded in. It can only inform us of parts of social networks and exploring entire social networks as a patchwork of “local knowledge” therefore becomes problematic. This essentially boils down to sampling issues that exist both on the whole network and the ego-network analytical scales. Network techniques allow archaeologists to traverse these different scales within the same methodological framework. I believe that the results of an ego-network study in particular can aid an interpretation of patterns identified in top-down network approaches, specifically in those cases where sampling issues on the whole network scale are considerable. For example, the structure of many complex social networks has been empirically proven to have a non-random degree distribution (Albert and Barabási 2002), a pattern that becomes clear on the level of the whole network. It has been shown, however, that this has a significant impact on the structure of ego-networks (Newman 2003a; Roberts et al. 2009). Exploring differing structures of ego-networks within a larger network can therefore provide an indication of whole-network patterning or help explain it. The analytical potential the multi-scalar nature of network approaches allow for has been recognized by archaeologists (e.g. Coward 2010) and features prominently in Carl Knappett's (2011) work. On his micro-networks scale of analysis Knappett traced hypothetical relationships between individuals and artefacts using affiliation networks. A similar approach could be used to focus explicitly on key ego's (or each ego in turn).

2.4. Discussion: beyond SNA

In this section I have illustrated how SNA tools and techniques have been used by archaeologists. Their applications are largely restricted to network visualization and exploring the static structure of archaeological datasets or social hypotheses. However, one can observe some general trends in the use of formal SNA methods. A number of methods, like centrality measures, seem to have been more popular than others, including affiliation and ego approaches. The reason for this might lie in their prominence within SNA itself as well as their clear aptness for exploring issues related to transport, power and control. Whereas early SNA applications (e.g. Irwin-Williams 1977) were strongly influenced by network methods in geography (Haggett 1965; Haggett and Chorley 1969), more recent ones drew upon a number of key SNA reference works (Carrington et al. 2005; Hanneman and Riddle 2005; Nooy et al. 2005; Wasserman and Faust 1994) as well as some well-known SNA theories and applications (Granovetter 1973; Hage and Harary 1996). This shift is at least in part a result of technological factors thanks to the more general availability of potent computing power and large digital datasets, which makes the application of many SNA techniques more worthwhile, and through the use of popular SNA software that is strongly linked with these SNA reference works. Although a growing number of software packages can be used to perform SNA techniques (for an overview see Huisman and van Duijn 2005; 2011), most archaeologists used either Pajek (Nooy et al. 2005) or UCINET (Borgatti et al. 2002), arguably the two most popular programs in SNA that are frequently expanded with new SNA techniques.

The most difficult hurdle for archaeological network analysts to overcome is not technological, however, nor is it related to a critical application of formal techniques. Indeed, SNA measures are largely adopted as they are available in software packages and are rarely adjusted to specific archaeological networks. It is the interpretative jump from identifying patterns in static network structures using SNA to explaining them in terms of past social processes that proved to be difficult in many cases. This is an issue present in any archaeological method, yet it is worth pointing out its influence in the particular case of SNA. Many SNA techniques come pre-packaged with traditional social explanations and it is tempting to adopt these in archaeological studies. However, in many cases the units of analysis are not social entities which makes adopting traditional social explanations problematic. It should be recognized that identifying and explaining network patterns are two different things. Most archaeological network analysts are aware of this and stress that both are completely dependent on how network nodes, links and measures are defined. We therefore cannot adopt SNA techniques into our discipline without question, although many of the topics and their traditional approaches discussed might prove useful for answering complex social research questions in archaeology.

The nature of archaeological data makes the archaeological application of social network analysis as an interpretative tool problematic for a number of reasons. Firstly, the full complexity of past social interactions is not reflected in the archaeological record and social network analysis does not succeed in representing this complexity. Secondly, the use of social network analysis as an explanatory tool is limited and it poses the danger that the network as a social phenomenon and as an analytical tool are confused (Knox et al. 2006; Riles 2001). Thirdly, human actions are based on local knowledge of social networks, which makes the task of exploring past complex social systems through particular material remains problematic. These issues should not be considered unique to archaeology and archaeological data. An approach consisting of a number of aggregated SNA techniques could be suggested for understanding aspects of past social relationships. However, we will never be informed about the full complexity of past social relationships and, even if we were, SNA would not succeed in understanding this complexity.

The recently very popular research tradition of complex network simulation in physics seems more promising in this respect, although it is by no means perfect itself. Indeed, neither SNA nor complex networks techniques are designed to unravel the full complexity of social interactions and archaeologists should definitely not apply them as if they were. As I will argue below, it is a combination of SNA and complex network simulation techniques that seems to hold the true potential of networks for archaeology. We will now turn our attention to network perspectives developed to understand properties of both human and non-human complex systems.

2.5. Complex networks and physics

Many of the ideas underlying the network science work done by physicists are rooted in complexity theory. Melanie Mitchell recently defined a complex system as "a system in which large networks of components with no central control and simple rules of operation give rise to complex collective behavior, sophisticated information processing, and adaptation via learning or evolution" (Mitchell 2009, 13). A long list of very diverse real-world systems fit this definition, including the World Wide Web (Adamic and Huberman 2000a; Albert et al. 1999; Broder et al. 2000; Huberman and Adamic 1999), co-authorship of papers (Barabási et al. 2002; Newman 2001), the brain (Sporns 2002; Sporns et al. 2000; White et al. 1986), and even the web of human sexual contacts (Liljeros et al. 2001). Although such systems are quite different, they have some features in common on an abstract level, as summarized by Mitchell: complex collective behavior, signaling and information processing, and adaptation (Mitchell 2009, 12-13). Notice how the first two are similar to Wasserman and Faust's first two principles of social network analysis introduced above. Indeed, network thinking is a popular perspective in complexity science as it forces one to think explicitly about how things relate and how local interaction between individual entities might give rise to patterning on a system-wide scale. Network thinking in complexity science is in part indebted to SNA as they share a research perspective and some of the techniques used to analyze complex networks were originally developed by social network analysts. Although social network analysts recognize the importance of dynamically changing networks through adaptation and have developed some methods to confront this problem (Lusher et al. 2013; Snijders 2005), most SNA applications still focus on structural properties of static networks. Contrary to social network analysis, however, the adaptation and evolution of systems through learning or evolutionary processes is a key assumption in complexity science (Bentley and Maschner 2003b; Mitchell 2009).

Much of the work on complex systems aims to identify and explain self-organizing emergent properties. Such properties are called self-organizing because they are patterns visible at the scale of the system, but emerge without any internal or external planning or control. They are called emergent because they arise out of the relatively simple interactions between individual entities or actors, who collectively form more complex behavior (Mitchell 2009, 13). Examples include the large and immensely variable mounds constructed by termites, or the way cities and even slums emerge without any top-down planning but merely through the needs and actions of (groups of) individuals. Identifying such properties is but a first step to understanding the complexity of systems, both past and present, and network models have been developed to do just that. I will not elaborate more on complex systems (for an overview of complex systems in archaeology: Bentey and Maschner 2003a; Bentley and Maschner 2007; Bintliff 2004; Garnsey and McGlade 2006; Kohler 2012; Lane et al. 2009; McGlade 2005) and focus entirely on some of these complex network models.

A few very popular models have been developed to identify and simulate particular processes that lead to the emergence of properties that turn out to be extremely common in diverse real-world networks. Although these models are by no means the only techniques for understanding properties of complex systems (for a few examples of other approaches to complex systems see Bak et al. 1987; Buldyrev et al. 2010; Turcotte 1999; West et al. 1999), they have dominated research in complex networks for the past decade. Two complex network models have been particularly influential to archaeologists: the small-world model and the scale-free model.

2.5.1. Small-world networks

In 1998 Duncan Watts and Steven Strogatz developed a simple model capturing a feature of complex networks that has puzzled sociologists for decades: the small-world problem (for the original paper: Watts and Strogatz 1998; some very readable overviews of the model and its implications followed: Watts 2003; Watts 2004; for an overview of pioneering work on the small-world problem see: Garfield 1979; Milgram 1967; Pool and Kochen 1978). The small-world problem was originally examined by Stanley Milgram (1967; 1992; Korte and Milgram 1970) in his experiments of how letters are passed on between two individuals who do not know each other. Milgram concluded that any one individual on the planet can be reached by any other individual in an average of six inter-personal steps, giving rise to the concept of *six degrees of separation*. The reason for this, Watts and Strogatz discovered, lies in the fact that "real-world networks are neither completely ordered nor completely random, but rather exhibit important properties of both" (Watts 2004, 244; Watts and Strogatz 1998). They identified a broad region in between both states where networks are highly clustered whilst the average path length is as small as possible (Fig. 9). Clustering is measured by the clustering coefficient which calculates the average probability that two neighbors of a vertex are themselves neighbors, as a ratio of the number of edges between the neighbors of a given node and the maximum number of edges that could possibly exist between these neighbors (Albert and Barabási 2002, 49; Newman 2010, 262-266; Watts and Strogatz 1998, 441). They adopted the name *small-world networks* to refer to this class of networks, a term first used by Eugene Garfield in 1979. The specific structure of small-world networks has direct consequences for the way networks evolve and how information, objects and people move through them. A crucial aspect of social networks this model does not address, however, is that human actions are limited by a strictly local knowledge of the networks they belong to and influenced by a general ignorance of the social system as a whole. Network analysts are aware of this (Kleinberg 2000; Watts et al. 2002). This issue does not undo the relevance and structural consequences of a small-world pattern however. Long-distance relationships between social clusters existed in the past and they did influence the lives of individuals, even if those people were not aware of their existence (e.g. Malkin 2011). A most striking example of the importance of long-distance relationships is reflected in the way infectious diseases spread through human networks. Bacteria do not care about whether people know of their own relatedness, because mere physical proximity suffices to jump between individuals. As such, long-distance relationships play a crucial role in transmitting diseases between largely independent communities (Newman 2003b; Watts 2003, 162-194).

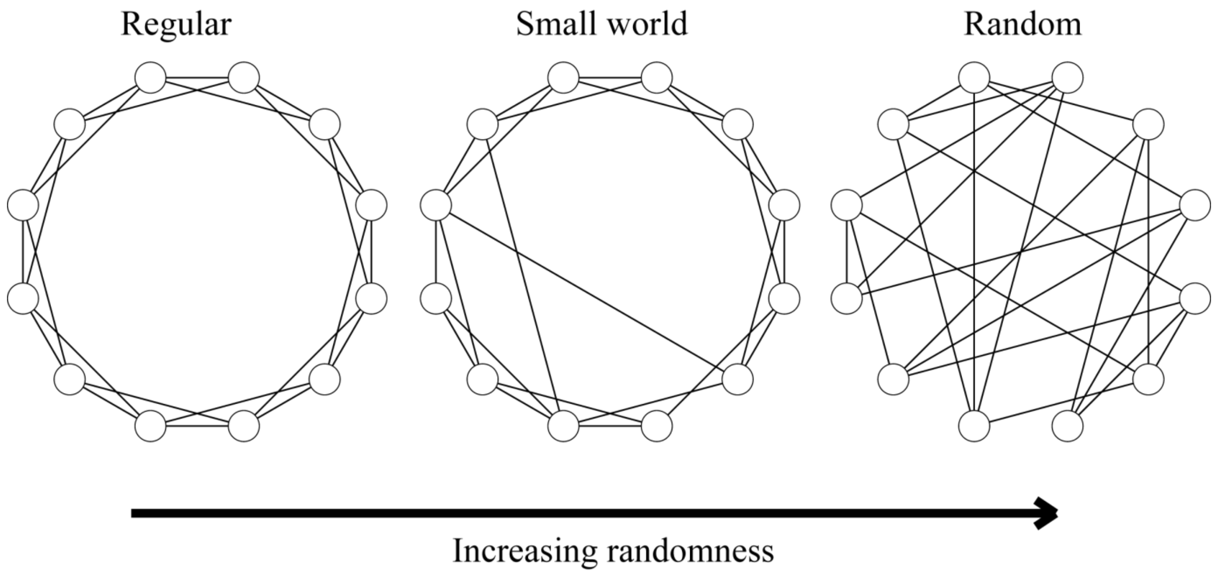


Fig. 9. Through a random rewiring procedure a small-world network structure emerges as a state between regular and random networks (after Watts and Strogatz 1998, Figure 1).

2.5.2. Scale-free networks and power-laws

A second popular model was published shortly after Watts and Strogatz's work and was in fact developed using the same real-world network datasets to address a fundamental assumption of the former model. Albert-László Barabási and his student Réka Albert concluded in their ground-breaking paper published in *Science* in 1999 that in real-world networks degree distribution (the fraction of nodes in a network with a certain number of relationships; Albert and Barabási 2002, 49; Newman 2010, 243-247) is not normal as Watts and Strogatz assumed, but is in fact highly skewed and follows the pattern of a power-law distribution, as in figure 10 (for the original paper: Barabási and Albert 1999; some very readable overviews of the model and its implications followed: Albert and Barabási 2002; Barabási 2002). The majority of vertices typically have less than the average number of relationships whilst a small fraction of hubs are much better connected than on average. Barabási and Albert made a simple mathematical model where nodes are continually added and attach preferentially to those nodes that are already well connected, effectively giving rise to a rich-get-richer effect (Barabási 2002, 79-92). Many real-world networks turn out to exhibit a scale-free structure. This realization had a significant impact on the way complex networks are approached, because not only does it imply a dramatic change in perspective away from random graphs (at least this is what Barabási claims), it also exhibits specific properties such as vulnerability to failures and attack that help us understand their functioning (Watts 2003, 109).

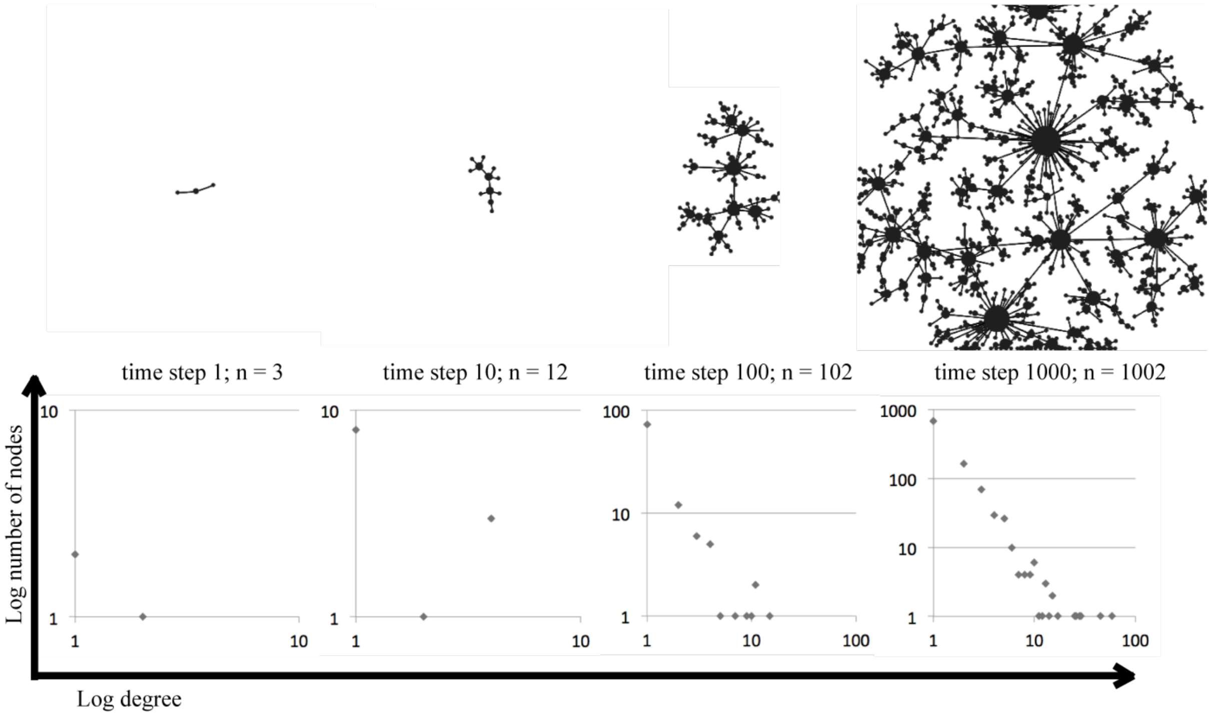


Fig. 10. Example of scale-free network growth (using a modified version of the model by Wilensky (2005) using Netlogo (Wilensky 1999)). In the Barabási-Albert model a new node is added to the network at every time step and attaches preferentially to an already well connected other node. The degree distribution when plotted on a chart with two logarithmic axes shows an approximate power-law for growing network size.

2.5.3. Small-world and scale-free networks in archaeology

The two complex network models described in the previous two sections are by far the most widely applied in archaeology as well as in other disciplines. Social network analysts often criticize the claims to novelty by physicists and others using these models, arguing that they ignore the advances made by the SNA communities over the years (e.g. Scott 2011). I will illustrate below that physicists and social network analysts are increasingly stressing the compatibility of their approaches and interests, a trend that I believe will allow for interesting archaeological applications as well. These two models and the multi-disciplinary work they triggered did, however, give rise to a strong increase of network studies in archaeology. I will briefly discuss a number of archaeological applications of these two popular models in this section.

In her analysis of the complex social interactions between Ancient Near Eastern sites in the Epipalaeolithic and Early Neolithic, Fiona Coward (2010; 2013) combined a traditional social network analysis with a description of small-world network structure. Discrete ^{14}C dated levels of sites formed the nodes in her networks and these were linked by the co-occurrence of particular forms of material culture which were considered “a material reflection of some form of social relationship (in its widest sense) between those sites” (Coward 2010, 464). These networks were explored in 1,000 year time-slices and their density, centralization, average degree and average path length was compared. The increases in network centralization, density and average degree were considered highly significant whilst the declining trend in average shortest path length was not. Coward interpreted these results as a structure of dense kin- and proximity-based groups linked by *weak ties* (as described by Granovetter 1973; 1983), which appears to develop towards a small-world phenomenon. Based on her description of the results, however, the identification of weak-links and a small-world structure is problematic. Coward gives no measure of clustering but rather assumes a high degree of clustering from the increasing network density and average degree, and the decreasing average shortest path length. This was one of the issues Coward raised with the application of network measures to a patchy archaeological record.

Søren Sindbæk (2007b) developed a networks approach for the study of artifact type distributions using small-world and scale-free networks. In his networks Early Viking Age sites in South Scandinavia formed the nodes and a connection between them was made when selected artefact types were co-present. Sindbæk stressed that “A shared artifact type does not show actual communication between sites, rather it indicates the existence of a group within which every site was connected to at least one other site" (Sindbæk 2007b, 66). Since Sindbæk was interested in the organization and dynamics of communication between these sites he interpreted the archaeological network as an impression of the underlying complex network by using network models. He suggested that the existence of a small number of hubs with a far more than average degree indicate that the settlements might have communicated as a scale-free network. Sindbæk also identified a high level of geographical clustering with each geographical region having some sites closely associated with the dense core of the network, which was interpreted as a probable small-world structure. Interpreting the archaeological evidence through these models suggests that "Communications across long distances were achieved through a spindly combination of hubs and weak ties" (Sindbæk 2007b, 70) and that this structure was also very vulnerable to targeted node removal. However, Sindbæk was very much aware of the issues involved in assigning the known behavior of dynamic models to a static structure visible in his archaeological network: "The fact that only a single phase is analyzed means that the dynamics involved are only hinted at" (Sindbæk 2007b, 70). Sindbæk shares this concern with almost all archaeological network analysts. Their general response to this issue involved the strict definition of network nodes, edges, measures, models and how to interpret the results, as well as a sense that one should not expect to understand more than a hint of past dynamic processes outside the static picture offered by the network. Complementing descriptions of static archaeological network structure with actual simulations of network models might help archaeological network analysts to move from the static to the dynamic, as I will argue below and will illustrate through case studies 2 and 3 (chapters 4 and 5). A number of archaeologists used chronologically subsequent networks of archaeological data to explore changing network structure (e.g. Golitko et al. 2012; Collar, 2007; 2008; Graham, 2006b), some of which will be discussed in more detail below.

A far more quantitative use of complex network models dominated Bentley and Maschner's (2003a) edited volume entitled '*Complex Systems and Archaeology*' (reviewed by Janssen 2005), a collection of papers with a particular focus on scale-free networks, punctuated change and agency that grew from presentations at the *Theoretical Archaeology Group* meeting in 2000. In the first part of the volume the editors provided an introduction to complex systems, discussing the potential use of the small-world and scale-free network models for archaeology (Bentley 2003a). They further illustrated how the emergence of social inequality in prehistoric societies can be explored through scale-free network growth (Bentley 2003b) in their analysis of house sizes on the North Pacific (Maschner and Bentley 2003). The authors argued that house size variability is indicative of the size of groups and the status of their headmen and can therefore be used to explore the emergence of social inequality. They argued that given “the nature of status striving among North Pacific hunter-gatherers” competition and growth should feature prominently in their approach, which is why the scale-free network model was considered particularly well suited for their aims (Maschner and Bentley 2003, 52-53). Trends in the frequency distributions of house sizes were interpreted in light of the rich-get-richer effect of scale-free networks: “a few households grow huge, with the remainder staying at levels similar to, or only slightly larger than, those in egalitarian societies” (Maschner and Bentley 2003, 57). In this example (and similar to the work of Coward and Sindbæk discussed above) Maschner and Bentley consider static patterns in the archaeological record as an indication for the structure of complex network models, allowing them to attribute the dynamic behavior of these models to past processes. Contrary to all the studies discussed above, however, the authors also explored dynamic simulation approaches in this edited volume.

2.5.4. Discussion: complex networks beyond popular models

Although they have received a disproportionate amount of attention, the models introduced above and the tsunami of papers they triggered in a wide range of disciplines are subject to some fundamental critiques, which should be acknowledged by archaeologists (for a brief overview see Mitchell 2009, 253-255). The mere identification of emergent self-organizing properties does not explain how this behavior came about and what it meant for the individuals creating it. Indeed, Kohler stressed that “characterizing a property as emergent is at best a general description and never an explanation” (Kohler 2012). Although these models imply that changes arise through interactions at every scale (individuals, communities, system-wide), the properties they allow us to identify do not tell us anything about specific human actions on a local scale. This shortcoming is particularly crucial in archaeology, in that we are typically confronted with the material reflections of isolated actions by individuals or small groups of individuals. This makes summing up our evidence to reveal system-wide patterns problematic and forces archaeologists to explore local actions. Similarly, Bentley (2003a, 15) raised the crucial point that merely identifying a power-law distribution in archaeological data is close to meaningless (at least in part because these patterns are common in nature (Frank 2009)). One has to understand what mechanisms created the power-law and what it means. To be specific, preferential attachment should not be seen as the only cause for power-law degree distribution in nature (Bentley and Shennan 2005; Mitchell 2009, 254; Shalizi 2011).

Archaeologists should also not forget that these models are minimizing abstractions of real-world networks, “they are overly simplified and based on unrealistic assumptions” (Mitchell 2009, 254). George Box captured the unrealistic nature of models perfectly by stressing that “all models are wrong, but some are useful” (Box and Draper 1987, 424). Mitchell mentioned that complex network models will never be able to represent the full complexity of real social systems as all nodes are often assumed to be identical except for their degree, and all links are the same type and have the same strength (Mitchell 2009, 255; although many models no longer share this assumption, e.g. Evans et al. 2009). Physicists are indeed aware of these shortcomings (Watts 2003; Watts 2004) and the “unimaginative” deterministic assumptions that follow from this form the favorite stick of social network analysts to beat them with (Carrington et al. 2005, 2; Scott 2011). For example, one key aspect of real-world systems that is not taken into account in these models is geographical space. Archaeologists seem more aware of this shortcoming than physicists, given the high number of archaeological network analyses that aim to explore spatial networks (e.g. Allen 1990; Branting 2007; Coward 2013; Earl and Keay 2007; Knappett et al. 2008; Pouncett and Lock 2007; Terrell 2010b; Zubrow 1990), although it must be said that recently the spatial nature of real-world complex systems has come to the attention of physicists (e.g. Barthélemy 2010; Gastner and Newman 2006) as well as the SNA community (Adams et al. 2012; see the special issue of *Social Networks* (34:1) dedicated to spatial networks). The graph theoretical techniques developed in space syntax (Hillier and Hanson 1984) also often incorporate distance as an edge attribute. Verhagen and colleagues (2013) have argued that these techniques are preferable to their non-spatial equivalents when applied to least-cost-path networks. Space syntax techniques are commonly used in archaeology to study the structure of well-preserved architectural features. Good reviews of archaeological applications of space syntax are published by Marion Cutting (2003; 2006) and will not be discussed in more detail in this thesis.

Archaeologists should not expect complex network models to capture the full complexity of systems, nor should they attribute the most popular explanations to descriptions of emergent properties. In light of these issues I would argue that archaeologists’ focus on a few popular complex network models severely limits the potential descriptive power of network modeling and holds the danger of introducing a routinized explanatory process. The examples from the previous section clearly illustrate the potential for complex network models to describe and explore archaeological hypotheses. A wealth of alternative network models exists, however, simulating a range of different behaviors that might be used in different archaeological research contexts (for overviews see: Costa et al. 2007; Newman 2010). The inability of many of these models to address more than one property of a complex network makes approaches that critically compare different models' behaviors with the archaeological record particularly promising (as in Bentley and Shennan's (2003) work described below). In fact, it is striking that random graph models (Erdős and Rényi 1959; 1960; 1961), that underlie much of the scale-free and small-world research, have hardly been used in archaeology as comparative models (for an exception see Graham 2006b), nor have spatial network models (Barthélemy 2010; Gastner and Newman 2006; for an exception see Bevan and Wilson 2013; Rihll and Wilson 1987; 1991). Archaeological problems might even drive the development of original complex network models (e.g. Knappett et al. 2008; 2011; Evans et al. 2009 described below), as I will illustrate in case studies 2 and 3 (chapters 4 and 5). Alternative network models allow for a wide range of applications, which again illustrates the need for archaeologists to clearly define network elements, contextualize results and, where possible, validate these with empirical data (Graham 2006a). Some pioneering archaeological applications of alternative complex network models, discussed in the next section, illustrate that this is a research direction worth pursuing. The simulations used in these applications all require some advanced mathematical and computational knowledge, which might explain the relative lack of uptake in archaeology, but also give them the great advantage of being able to explore the processes driving network change.

2.5.5. Dynamic network models in archaeology

Bentley and Shennan's (2003) study of different processes of cultural transmission is a particularly good example of how the distinct behavior of slightly different network models can be compared in an archaeological context. The authors suggested three quantifiable types of cultural transmission with a testable difference: independent decisions in a highly simplified model show an exponential decay in variant frequencies (an artifact variant was assumed to have a discrete nature, existed for some finite time and can be copied (Bentley and Shennan 2003, 461)); unbiased cultural transmission is characterized by a power-law or log-normal distribution; biased cultural transmission deviates significantly from a null-model of unbiased cultural transmission. Bentley and Shennan argued that Adamic and Huberman's (2000b) model of stochastic network growth is particularly applicable to unbiased cultural transmission. This model generates scale-free networks through a slightly different process than the Barabási and Albert (1999) model and was preferred since the process of preferential attachment can also occur in a network that is not growing. This model was adapted to represent biased cultural transmission by making preferential attachment proportional to an exponent of the number of connections of a node. Bentley and Shennan subsequently used three variations of this model with different values for the exponent in their case-study to model change in Linear Bandkeramik pottery motif frequencies from the Merzbach valley (Germany). In these networks individual motifs (variants) formed the nodes and each copy of a motif was connected with an arc to the source motif. The authors concluded that they found a good fit between the stochastic network model for unbiased cultural transmission and the later-phase Merzbach pottery data. The earlier phase motifs, however, were suspected to have known a pro-novelty biased cultural transmission. As far as their use of networks is concerned, Bentley and Shennan have pioneered how complex network models can be adopted, critically modified to represent archaeological hypotheses and how their resulting behavior can be compared with an archaeological dataset.

Together with Mark Lake, Bentley and Shennan also studied evolving networks using agent-based modeling (Bentley et al. 2005) to explore how an exchange network coevolves with the changing specializations of the agents within it. This model is of particular interest for this PhD project since it inspired the model of Roman tableware distribution presented in case study 3 (chapter 5). The authors argued that power law wealth distributions "are ubiquitous for a wide range of economic scales" and that this behavior might be linked to the benefits of specialization in exchange networks (Bentley et al. 2005, 1346-1347). Their model therefore aimed "to test whether specialization and wealth inequalities are natural, self-organizing qualities of a small-scale economy" (Bentley et al. 2005, 1347). They modified a simple model to simulate exchange developed by Jin, Girvan and Newman (2001) by adding variables for agents' possession of two different products (A and B) and a *strategy* variable, which determines the relative amount of A vs. B an agent produces per time step. A significant difference in the wealth distributions was identified for two scenarios where, on the one hand, two agents either trade when both possess sufficiently different amounts of a certain commodity (resulting in normal wealth distributions), and on the other, where agents only trade if they both like the price of a commodity (resulting in highly skewed wealth distributions). The authors argued this suggests “a basic analogy to the profound ideological differences that likely existed between certain indigenous populations and incoming agricultural colonists” (Bentley et al. 2005, 1353). Although the authors do not apply this model to an archaeological case study, it has clear potential for testing archaeological hypotheses concerning wealth distribution and exchange networks. In this agent-based approach as well as in Bentley and Shennan’s (2003) work described above the adopted network models are therefore not modified to represent attested static patterns in the archaeological record, rather their value lies in thinking through a range of hypothetical dynamic networks as possible processes underlying the creation of the archaeological record.

Another combination of agent-based modeling and networks is Shawn Graham’s (2006a) study of Roman itineraries discussed above. Graham aimed at exploring the diffusion of information on the Antonine Itineraries by populating a map of the itineraries’ static network structure with digital agents who could interact and share a piece of knowledge. Contrary to Bentley and colleagues’ (2005) model, the network used in Graham’s model was a pixelated map of the spatial representation of places connected by routes. Nodes represented places rather than the agents and played a far smaller role than the network edges in this model since agents were allowed to interact at any point on the network map. The model therefore did not allow Graham to make statements about the role of nodes. Rather, the author was interested in how the different structures of provinces as a whole affected the diffusion of information. The model simulates agents moving along the paths of the itinerary and passing on a message to agents who have not heard it. Graham concluded that there are distinctive regional differences in the fashion and speed of information diffusion. It is these simulated processes of diffusion that form the dynamic aspect of Graham’s work. Rather than network evolution through time, Graham explored the static structure of a conception of Roman space as presented in the Antonine Itineraries through hypothetical dynamic processes (for alternative network models of diffusion see e.g. Cowan and Jonard 2004; Guardiola et al. 2002; Valente 2005; Zhuang et al. 2011).

A unique example of a complex network model developed for a specific archaeological research context is Knappett, Evans and Rivers’ (2008; 2011; Evans et al. 2009) *Ariadne*[2] model for maritime interaction in the Aegean Bronze Age, which emerged as a fruitful

[2] The *Ariadne* model can be downloaded here: http://figshare.com/articles/ariadne/97746 (accessed 27-05-2014).

collaboration between one archaeologist (Carl Knappett) and two theoretical physicists (Tim Evans and Ray Rivers). The model was formulated as a reaction to more geographically deterministic network methods (Rihll and Wilson 1987; 1991) and the work of Cyprian Broodbank (2000) in particular. In his study of the Early Bronze Age Cyclades Broodbank examined networks consisting of archaeologically attested sites as well as hypothetical sites added to islands based on population estimates derived from site surveys. He used the PPA method (introduced above) to link each site to its three geographically closest neighbors. In this case links were considered to be equal which Knappett and colleagues argued was not the case for interactions in the Middle Bronze Age Aegean. The authors set about developing a complex network model that specifically addressed their assumptions about maritime interactions in the Middle Bronze Age Aegean. 39 archaeologically attested sites that formed the *centre of mass* for their immediate areas (for technical details see Evans et al. 2009). These were represented as vertices that were assigned a fixed carrying capacity reflecting their local resources, and a variable indicating each site's relative importance. The links between two sites were given a measure of the physical distance between them and a variable representing the effort one site puts into the interaction with the other (for technical details and the cost/benefit optimization function in this model see Evans et al. 2009). Attributing such values to links is a key difference with the other archaeological applications of complex network models described above. The authors used this model to explore the effect on Late Minoan civilization of the catastrophic destruction of Akrotiri on Thera (Santorini) by volcanic eruption (Knappett et al. 2011). For the pre-eruption period the model revealed a high level of clustering with a number of weak and stronger links connecting clusters. The immediate post-eruption period was modeled by the removal of Akrotiri. The results indicated that the removal of a key node in this network had little immediate effect on overall activity. The authors suggested that the removal of a key node might not lead to big changes initially, but will inevitably increase exchange costs. To test this hypothesis they increased exchange costs. At first, total activity was not reduced substantially, but eventually this hypothetical change led to fewer strong links, causing major sites to focus on maintaining fewer links. This was considered an unsustainable situation that might lead to collapse, represented by regional clusters in the network becoming disconnected. One could argue that the authors did not succeed in testing their hypothesis since it is the increase in exchange cost imposed by the authors that caused the network to disintegrate and not the removal of Akrotiri. It could be argued that this scenario should have been compared with one where the exchange costs are increased without removing Akrotiri. The work by Knappett, Evans and Rivers also raises the issue of the role of archaeological data in complex network modeling. The model inputs are only based on the archaeological record to a very limited extent and the results of their case study were not validated against empirical data. It should be clear, however, that the value of this model lies largely in its ability to make an archaeological hypothesis of interaction explicit by modeling it as a process on a network. The results should therefore not be interpreted as predictions of past processes, but rather as stressing the potential evolution of hypothetical structures that require subsequent validation. Knappett, Evans and Rivers have illustrated that complex network models developed specifically to explore archaeological hypotheses can lead to innovative and useful ways of thinking about past processes that help guide future research efforts.

2.6. Network science in Roman Archaeology

The use of network science in Roman archaeology is limited. Despite a relatively recent and limited adoption, some Roman archaeologists have nevertheless developed some highly original network science studies. This is at least in part thanks to the availability of more diverse data types that are not available in prehistoric archaeology, such as brick stamps, epigraphy, and other written sources. The combination of archaeological and historical data types in Roman archaeology means that more complex questions can be asked and sometimes answered, such as the structure and functioning of the Roman economy (e.g. Scheidel et al. 2007), of maritime transport (e.g. Arnaud 2005), or of Roman social systems (e.g. Bang 2008). Roman archaeologists using network science have focused on these particular questions, no doubt because, as I have shown above through examples and definitions, network science is particularly suited for addressing questions involving the flow of goods, information, and the structure of social networks. Indeed, some of the studies I will discuss in this section include the most explicit examples in archaeology of the use of network science techniques for the analysis of past social networks (for example, the individuals named on brick stamps (Graham 2006b; 2009), on inscriptions (Broekaert 2013a), or in Cicero’s writings (Alexander and Danowski 1990)).

A number of studies have used network science techniques for the analysis of Roman road networks, including that by Graham (2006a) mentioned above. Another example is the use of closeness and betweenness centrality measures by Leif Isaksen (2007; 2008) in his study of the transport system in Roman Southern Spain. He combined data from the Antonine

Itineraries with the river network based on the Vicarello Goblets, and with the Ravenna Cosmography. Isaksen proceeds with a critical description of the centrality results for each data source and the significant issues involved in interpreting them, before he draws them all together in an attempt to explore aspects of the Roman transport framework of the region.

In her study of religious innovation in the Roman Empire, Anna Collar (2007; 2008; 2013) made a more descriptive use of complex network models as concepts to think with than some of the studies mentioned above. Collar's work aimed to explore why some religious movements succeed and spread, while others ultimately fail. Rather than seeing this process as a direct consequence of the religions' inherent properties, she adopted a bottom-up approach by exploring the social networks that linked the individuals who drove religious change. Through three case-studies, Collar examined the power to communicate religious ideas of three types of social networks: the military networks that she argues were instrumental to the diffusion of the cult of Jupiter Dolichenus, the ethnic network of the Jewish Diaspora, and the religious network of the cult of Theos Hypsistos. Collar explored these social networks through a combination of a close reading of an exhaustive epigraphic dataset and archaeological networks generated from the spatial distribution of inscriptions. The PPA technique previously used by Broodbank (2000; discussed above) and developed by Terrell (1976; 1977a; 2010a), was used to create these networks. For all three case studies Collar created a simple PPA network where sites with evidence of inscriptions were connected to their three geographically closest neighbors. The discussions of the resulting networks focus on the description of clusters, isolated communities and centers. Collar repeatedly stressed that these networks are not a reflection of the actual connections that existed between sites. Rather, through PPA she visualized and explored the assumption that communities interact most intensely with their closest geographical neighbors. Only in her final attempt to interpret the processes of diffusion that the discussed social networks hint at did Collar refer to concepts derived from small-world, scale-free, and complex systems research. She referred to the hypothesized roles of hubs, weak- and strong-links, information cascade, self-organized criticality and stochastic network growth in the spread of religious innovation. In short, Collar illustrated how networks can be used to explore the distribution of archaeological data and think explicitly in terms of social networks without driving their interpretation. In addition, she used the vocabulary of complex networks to describe hypothetical processes.

Few archaeologists have combined both SNA techniques and complex network modeling in a single approach. To some extent Fiona Coward's (2010) work discussed above is an example of this. However, a particularly successful example is provided by Roman archaeologist Shawn Graham (2006b; 2009) in his study of the individuals active in the Roman brick industry in the Tiber valley. Graham created two types of networks. First he looked at patronage networks in the brick industry, evidenced in the names of individuals appearing on brick stamps. Then he examined brick manufacturing networks, where brick makers were connected if they shared the same clay sources (as derived from an archaeometrical analysis). These networks were considered to represent social relationships (of the kind petrified in bricks) between individuals and were analyzed in four chronologically subsequent periods (Julio-Claudian, Flavian, Nerva-Antonines, Severans). Graham combined SNA and complex network models in a method that switches between the local and the global scales. On a local level, Graham used the degree and Bonacich centrality measures to identify social *hubs* and *bridges* (Graham 2006b, 103). To explore patterns on a global level he calculated the networks' average path lengths, degree distributions and clustering coefficients. The results led Graham to conclude that the patronage and manufacturing networks exhibited what he called *egalitarian* (characterized by nodes roughly having the same degree) and *hierarchic* (characterized by the presence of a power-law degree distribution, i.e. a scale-free network) small-world structures at different periods. He went on to argue that the known behavior of small-world and scale-free networks can be attributed to the structure of different social networks in the Tiber valley brick industry at different times. Graham stressed that although his networks are at best static snapshots of an evolving industry, a comparison with the structure of complex network models still allows one "to explore how (and why) the industry assumed these different shapes at different times; it allows us to move from the static to the dynamic" (Graham 2006b, 97). By drawing together network techniques developed in different disciplines, Shawn Graham provided a real multi-scalar network perspective that allows one to explore social structure both from the bottom-up and top-down.

I would like to argue, however, that Graham's interpretations of network structure are sensitive to the sampling issues typically surrounding archaeological data. Although network measures might indicate that archaeological data networks have a small-world and/or scale-free structure, interpreting this structure is quite another thing. The work by Stumpf and colleagues (2005) on sampling properties of networks revealed the extent of this issue:

> "Only if the degree distributions of the network and randomly sampled subnets belong to the same family of probability distributions is it possible to extrapolate from subnet data to properties of the global network. We show that this condition is indeed satisfied for some important classes of networks, notably classical random graphs and exponential random graphs. For scale-free degree distributions, however, this is not the case. Thus, inferences about the scale-free nature of a network may have to be treated with some caution" (Stumpf et al. 2005, 4221).

Graham's networks are relatively small, making their structure very likely to change dramatically through the addition of nodes and edges. In fact, they might turn out not to have a small-world or scale-free structure at all. Graham did argue and illustrate that his hypotheses of preferential attachment, wealth condensation (Bouchaud and Mézard 2000) and cascading failures (Albert et al. 2000) are supported by historical events. As I argued above for SNA, however, it should be clear that the known behavior of complex networks cannot be extrapolated to the structure of archaeological data without question. Indeed, one could argue that this issue is present in all exploratory archaeological network analyses and that it favors network simulation approaches for testing hypothetical whole networks. As I will repeat quite often throughout this thesis: identifying network structure and explaining it are two different things and requires one to move beyond a direct application of these popular models.

Graham's work illustrates the advantage Roman archaeology offers of combining archaeological and historical sources in a single approach. He enhances the use of SNA for the study of social networks of named individuals, such as in Alexander and Danowski's (1990) SNA study of the Romans named in Cicero's letters, by adding archaeometrical information. Similarly, the study of prosopographies or lists of ancient individuals named on a variety of sources (including inscriptions and ceramics) can be enhanced by incorporating both written and material evidence in a network science approach (the prosopographical SNA work by Broekaert (2013a) illustrates this potential).

2.7. Conclusion: the need to evaluate network science in archaeology

In the recently published SAGE handbook of social network analysis John Scott (2011) severely criticized the claims to novelty made by social physicists (and Albert-Lázló Barabási in particular. See also Bentley and Shennan 2005) given their "almost total ignorance shown concerning the vast amount of prior work in social network analysis" (Scott 2011, 55). Scott illustrated that there is a long history of social physics in sociology of which the new social physicists seem to "know little or nothing" (Scott 2011, 55). Indeed, citations to previous work in SNA by social physicists are largely limited to popular textbooks (Degenne and Forsé, 1994; Scott, 1991; Wasserman and Faust, 1994). Scott seems particularly unhappy with Barabási's claim that the work in social physics triggered by Watts and Strogatz (1998) posed "the first serious challenge to the view that real networks are fundamentally random" (Barabási 2002, 15). Scott replied "that no sociologists, to the best of my knowledge, have ever thought that complex social networks are purely random phenomena" (Scott 2011, 58). It should be said in Barabási's defense that Scott only cites Barabási's popular science book *Linked* (2002). In fact, Scott himself seems quite unaware of the scope of recent work in physics by Barabási and his followers. Scott concludes on a hopeful note, stressing that the work by some social physicists like Duncan Watts shows potentially valuable contributions to the social sciences. Scott considered much of the work by social network analysts and physicists to be complementary (complexity theory perspectives and agent-based dynamical modeling used by physicists in particular hold great potential for SNA) and argues strongly for collaborations between the two.

In this literature review I have argued (although worded less strongly than John Scott) that a similar process is emerging in the archaeological use of formal network methods. A number of general trends can be identified that attest of a general unawareness of the historicity and potential diversity of existing network-based approaches or of suitable archaeological applications of known models and techniques and of the issues related to them (a similar development was recognized by Claire Lemercier (2012) for the discipline of history); which lead me to believe that, as far as formal network methods are concerned, there is a clear need for multi-disciplinary collaboration.

Firstly, archaeological applications of graph theory, which have been around since the 1960's, were not influential at all on more recent archaeological network studies. This is peculiar since many of the SNA techniques that only in the past decade have become more popular with archaeologists are rooted in graph theory. In fact, the archaeological applications of graph theory clearly illustrated the potential of the graph as a technique for visualization and analysis in research contexts that showed strong similarities with studies in social network analysis. The introduction of graph theory and social network analysis into archaeology therefore happened largely independently.

Secondly, the potential of social network analysis techniques was explored (largely theoretically) through Cynthia Irwin-Williams' (1977) network model. Many of the techniques she described were not applied in archaeological research until the last ten years, which might at least in part be due to technological factors. Like so many other quantitative methods in archaeology the earliest network analyses in archaeology were also strongly influenced by the New Geography and the work of Chorley and Haggett (1970; Haggett and Chorley 1969) in particular.

Thirdly, social network analysis and physics have been the most influential research traditions, yet only a very limited range of models and techniques have been explored by archaeologists so far. Centrality measures in SNA and the small-world and scale-free complex network models are among the most popular approaches, whilst a wealth of alternative complex network models (like the ones described in the previous section) and SNA techniques (including affiliation- and ego-networks) that I have argued show great potential for being applied to address archaeological research questions remain largely unexplored.

Lastly and most crucially, there is a danger of falling towards a standardized explanation of attested network structure. It is as if every network method comes with a social interpretation that needs only be moulded to fit the specific archaeological research context in question. This is definitely not the case as different processes can explain the emergence of the same structure. This issue might at least in part be caused by an unawareness or a limited discussion of the network science research process (Brandes et al. 2013), i.e. the abstraction of past phenomena into network concepts, the representation of network concepts as network data, and the formulation of clear motivations for making each of these steps.

These general trends are the result of two critical issues that will need to be addressed in future archaeological network analysis and that further emphasise the need for the evaluation of network science in archaeology as I have claimed in the introduction to this PhD: (1) a general unawareness of the historicity and diversity of formal network methods both within and outside the archaeological discipline or of suitable archaeological applications of known models and techniques has resulted in a very limited methodological scope; (2) the adoption or development of network methods has very rarely been driven by specific archaeological research questions and is dominated by a few popular models and techniques, which has in some cases resulted in a routinized explanatory process.

These issues should not necessarily be seen as a critique towards existing archaeological applications of network science. If anything, they stress that network science has far greater potential than has already been explored by archaeologists. Moreover, the increasing number of archaeological applications in the last decade, largely triggered by the popularity of complex networks research in the early 2000's, seems to indicate that there is a genuine interest in formal network methods for archaeology. In order to channel this interest towards more diverse and more explicitly archaeological future applications, however, the above two issues will need to be confronted. To do this I would argue that three things are crucial: taking a broad multi-disciplinary scope, letting the specific archaeological research context dominate the application, and explicitly work through each step in the network science research process. The first two conclusions might sound contradictory at first but in light of this review of archaeological applications their relevance and complementary nature becomes clear. Carl Knappett recently argued that "for new network approaches to be successful in archaeology they have to be as profoundly transdisciplinary as we can possibly make them" (Knappett 2011, 37). Indeed, we have seen that studies addressing a specific archaeological problem through a multi-disciplinary collaboration (e.g. Knappett et al. 2008; 2011) or through a combination of different network techniques (e.g. Bentley and Shennan 2003; Graham 2006b) have often been the most fruitful ones. The combination of SNA techniques and complex network modeling (e.g. Coward 2010; Graham 2006b; 2006a), which is considered to have great potential according to John Scott (as well as other social network analysts, e.g. Borgatti et al. 2009), is particularly promising for archaeology as it allows for a top-down as well as a bottom-up perspective to explore the multi-scalar nature of network thinking (Knappett 2011). The adoption or development of network methods should, however, always be motivated by specific archaeological research questions. The building blocks of every network (nodes, edges, and their parameters) as well as the techniques used should always be clearly defined from the outset, as they will dominate the interpretation of the results. These results in turn require a re-contextualization within their wider archaeological framework before one can make the jump from the identification of network structure to its explanation in terms of past dynamic social processes.

2.8. Implications for developing case studies

The main conclusions of this literature review for enabling an evaluation of the potential of network science in archaeology are:

- take a broad multi-disciplinary scope;
- let the specific archaeological research context dominate the application;

- explicitly work through each step in the network science research process.

These three conclusions drawn from the literature review will guide my use and evaluation of network science techniques through the three archaeological case studies in the next three chapters. All three case studies include an exploratory network analysis, whilst case studies 2 and 3 evaluate how this can be combined with different types of confirmatory network modelling approaches. The first case study concerns a multi-scalar exploratory network analysis, in the sense that different sample sizes drawn from a large citation dataset are considered to illustrate processes of academic influence taking place on different scales. Case study 3 on the other hand, explores a different kind of multi-scalar network approach, where the actions of individual software agents (for which there is little evidence in the archaeological record) give rise to large scale patterns that can be compared to general trends in the archaeological record. Every case study will be developed in light of an archaeological research question and the network science method most suitable to address this question will be selected, rather than adopting innovative methods with limited potential in archaeology just for innovation’s sake. In each case study I will explicitly work through the network science research process to evaluate how this influences the selection of network methods, how this fascilitates the definition of network data and dependence assumptions, and eventually how this approach allows archaeologists to better evaluate the potential contribution network science can make in their research. Finally, the implications and relevance of the quantitative results obtained from the network science techniques for evaluating the archaeological research question will always be discussed by re-contextualising them within a wider archaeological framework.

A final conclusion should be drawn from the use of network science in Roman archaeology. Although section 2.6 illustrated some original applications and the potential to address complex research questions thanks to a wealth and diversity of data, these kinds of studies are rare in Roman archaeology. Indeed, Roman archaeology has neglected the adoption and development of computational techniques to leverage this wealth of data and allow for testing the many existing conflicting hypotheses of Roman complex systems (Morris et al. 2007). Moreover, due to the availability of different data types and the focus of academic discussions in Roman archaeology on complex phenomena like urbanization, transport, economy, and social structure, the research themes discussed and methods more commonly used in prehistoric archaeology are neglected in Roman archaeology. Case study 2 (chapter 4) offers an example of this, where it is argued that the potential role played by lines of sight between Roman urban settlements in Southern Spain in structuring human behavior is neglected, even though it is a common research theme in the study of the immediately preceding Late Iron Age period. Both case studies 2 and 3 are developed in order to address the lack of formal methods for testing complex network-related research questions in Roman archaeology, by adopting and adapting network techniques which have proven useful in other archaeological subdisciplines or outside of archaeology.

Box 3. Summary chapter 2

Issues:

- A general unawareness of the historicity and diversity of formal network methods both within and outside the archaeological discipline or of suitable archaeological applications of known models and techniques has resulted in a very limited methodological scope.
- The adoption or development of network methods has very rarely been driven by specific archaeological research questions and is dominated by a few popular models and techniques, which has in some cases resulted in a routinized explanatory process.
- The use of network science in Roman archaeology is limited but highly creative, thanks to the diversity of the available data and the complexity of research questions.

Future archaeological network science approaches need to:

- take a broad multi-disciplinary scope;
- let the specific archaeological research context dominate the application;
- explicitly work through each step in the network science research process.

CHAPTER 3

CASE STUDY 1:

A CITATION NETWORK ANALYSIS OF THE ADOPTION, USE, AND ADAPTATION OF FORMAL NETWORK TECHNIQUES IN ARCHAEOLOGY

3. Case study 1: a citation network analysis of the adoption, use, and adaptation of formal network techniques in archaeology

3.1. Introduction[3]

This first case study aims to illustrate how the publications discussed in the literature review[4] could be explored in a different way, by focusing on their citations.

Formal methods, like network science, seem to constantly fade in and out of scientific practice. Some are lucky enough to find a large audience, such as the archaeological use of GIS, whilst others are quickly forgotten, such as the archaeological use of fractals. But how does a new technique emerge in a discipline, where did it originate and how does it evolve in a new research context? A well-established set of bibliometric methods has been developed to quantitatively address these sorts of questions using scientific literature and this case study is concerned with citation network analysis in particular (Börner et al. 2004; Garfield et al. 2003; Hummon and Doreian 1989). Recently, a wider availability of powerful computational resources, bibliometric software (e.g. HISTCITE[5]; PAJEK[6]; PUBLISH OR PERISH[7]) and large bibliographic datasets in the sciences as well as the humanities resulted in unprecedented progress in the analysis of citation networks (Börner et al. 2004). This case study will build on these recent advances.

The archaeological application of formal network methods forms a particularly suitable case to explore academic processes of adoption and adaptation through the related citation behaviour. Through a close reading of archaeological applications of formal network methods the literature review presented in chapter 2 provides an intuitive understanding of the general academic traditions, techniques, models and individuals influential to these applications. But

[3] Please see the electronic supplementary material for this case study, which includes the citation network files.

[4] I choose to explore the use of citation network techniques in relation to the literature review rather than for the published examples of the use of network science in Roman archaeology. This because the literature review revealed the use of network science in Roman archaeology is very limited and it would not allow me to explore the potential of exploratory citation network techniques. However, these techniques could also be usefully applied to study the citation behaviour of Roman archaeologists and to map out how their citation practices differ from those of other archaeologists and of historians.

[5] http://thomsonreuters.com/products_services/science/science_products/a-z/histcite/ (accessed 26.06.2014).

[6] http://pajek.imfm.si/doku.php (accessed 28.04.2013); and WoS2Pajek in particular http://vlado.fmf.uni-lj.si/pub/networks/pajek/WoS2Pajek/default.htm (accessed 26.06.2014).

[7] http://www.harzing.com/pop.htm (accessed 26.06.2014).

do the combined bibliographies of these applications tell a similar story? It is the aim of this case study to explore how an exploratory citation network analysis can usefully contribute to an archaeological literature review. The phenomenon studied in this case study through a network science research process is therefore an academic one: the adoption and adaptation of network science techniques by archaeologists. The abstraction into network concepts and the subsequent representation as and analysis of network data will be guided by an evaluation of the following research questions:

- Can the intuitive results drawn from a close reading be quantitatively expressed and validated using citation network analysis?
- Can an exploratory network analysis provide new insights into the citation behaviour within the limited bibliographic corpus of the close reading?
- Does a citation network analysis allow one to usefully contextualise a close reading within a larger and multi-disciplinary corpus of literature? And does it provide new insights into the adoption and adaptation of network science in archaeology?
- Do archaeologists publish and cite differently from scholars in other disciplines? If so, does this require the development of specific citation network analysis techniques or strategies?

Many of the answers to these research questions will sound familiar, since they aim to reproduce the conclusions drawn from the literature review. However, this case study will advance the general aim of this PhD by evaluating in practice the potential of network science techniques as exploration and communication tools (third research question of this PhD). Much of this potential must lie in the ability of this approach to ask and answer different questions than those of the literature review, and I will therefore focus the discussion at the end of this case study on the added value of exploratory network analysis as compared to a close reading of the same body of literature. Moreover, throughout this first case study I will introduce in practice some of the basic network science techniques that will also be used in the other two case studies.

3.2. Citation network analysis and archaeology

The foundations of citation network analysis were laid by Garfield and colleagues (1964) and the application of graph theory for citation network analysis was explored by Garner (1967). Despite this long tradition, and although citation analysis is more commonly used in the field of anthropology (e.g. Choi 1988; Clark and Clark 1982; Robinson and Posten 2005), its use in an archaeological context has not been thoroughly explored yet. A number of studies use simple counts of citations or other bibliometric data to track trends in the archaeological sciences and compare the impact and evolution of archaeological journals (e.g. Butzer 2009; Marriner 2009; Rehren et al. 2008; Rosenswig 2005; Sterud 1978), or to evaluate the impact of gender differentiation in archaeology (e.g. Beaudry and White 1994; Hutson 2002; 2006; Victor and Beaudry 1992). Of particular interest is the work by Schich and colleagues (Schich et al. 2009; Schich and Coscia 2011) who analysed the co-occurrence of classifications in the *Archäologische Bibliographie* (a library database consisting of over 450,000 titles, 45,000 classifications, and 670,000 classification links). Their approach of working on three analytical levels is followed in this case study. Schich and Coscia (2011) also point out that in classical archaeology and in the arts and humanities in general "citation indices are of limited use and literature is still not available in digital form". The available citation indices for the Arts and Humanities (and the Institute for Scientific Information's Arts and Humanities Citation Index in particular) have significant limitations (Nederhof 2006) and for this reason citation network analyses in the Arts and Humanities are rare (Leydesdorff et al. 2011). Indeed, a study by Knievel and Kellsey (2005) has shown that disciplines within the Humanities have different citation patterns and should be considered separately, and that monographs (rather than peer-reviewed journal articles) are often the dominant format of cited sources (which is a problem since monographs are traditionally not included in citation indices).

It is often claimed that the citation and publication behaviour of archaeologists is in fact different from other disciplines. Xia emphasised the nature of archaeological data as one of the reasons for this and described the traditional practice of archaeological publishing as of 'broad scope, slow speed, limited priority, and selected distribution' (Xia 2011, 234). Work of relevance to archaeologists is published in science, humanities and history journals. To capture the multi-disciplinary nature of archaeological publishing and citation, the strategy of this case study will consist of exploring how a select number of archaeological publications is embedded within a wider multi-disciplinary web of publications. As such, this strategy does not allow to evaluate the citation and publishing behaviour of the entire archaeological community. Instead, it is interested in the degree of multi-disciplinary engagement formally expressed through citation.

Citation analyses in the arts and humanities should not be discarded out of hand, however, since they can still provide an alternative look at scientific practice through large aggregated datasets, as long as the nature of the datasets and their limitations are thoroughly understood. This case study can be considered a step towards evaluating the issues involved in using citation network analysis for the archaeological discipline by aiming to reveal archaeological citation patterns and how these interlink with citation practices in other disciplines.

3.3. Data collection, citation network creation and sample size issues[8]

In a citation network (e.g. Fig. 11) vertices represent publications and a directed edge (or arc) between two vertices indicates a citation (Eom and Fortunato 2011). The seeds of the citation network analysed in this case study are 69 published examples of the archaeological use of formal network techniques (these will be referred to as egos, for a full list see Appendix II), all of which were mentioned in the literature review in chapter 2. The decision was made to include rather than exclude publications where formal network techniques were only mentioned and not applied or where merely the vocabulary of formal network techniques was used, in the hope that this will reveal a wider variety of influences and archaeological uses. The selection was strongly influenced by my own knowledge of the topic, a bias that will be evaluated in the analysis.

Bibliographic data for these 69 egos were extracted from Thomson Reuters Web of Knowledge (WOK)[9], which provides access to the most comprehensive citation databases and is the most commonly used resource for citation network analysis (Newman 2010, 68). Many publications are missing from WOK, however, which introduces the second bias in this dataset (see Table 1). Although it includes over 12,000 international and regional journals from a wide range of disciplines, a high number of journals are not indexed and their citation data are therefore not available[10]. The citation data in books is also not included in this study since until recently it was not covered by WOK (the launch of the Book Citation Index for WOK was announced a few months after data collection for the current case study was

[8] The network data used in this analysis is included as an electronic supplement to the PhD.

[9] http://wokinfo.com/about/whatitis/ (accessed 26.06.2014).

[10] These missing journals include journals that might still have a wide readership like *Journal of Roman Archaeology* or *Archaeological Review from Cambridge*.

completed)[11]. Most crucially, WOK assumes that English is the language of science and its coverage of journals that accept publications in languages other than English is therefore limited. These are significant issues for the discipline of archaeology where publications in books (in particular for site excavation reports) and in languages other than English are extremely common. Nevertheless, WOK is still the largest available body of citations with a clear focus on high-impact journals. The journal selection criteria are clearly described and are subject specific[12], which allows for the influence of this selection bias on the dataset to be evaluated.

Table 1. Count of egos and neighbouring publications that could be found in WOK and those whose bibliographies had to be extracted manually. The last row “missing publications” refers to the number of publications that were cited by the egos but that could not be found on WOK.

	Egos	Cited by egos
Total publications	69	1859
Extracted from WOK	28	449
Extracted manually	41	0
Missing bibliographies	0	1410

Only 28 of the 69 egos could be found on WOK, which meant the bibliographies of the remaining 41 had to be extracted manually (Table 1). To better represent the citation context of the 69 egos a so-called ego-network was created, consisting of the egos, their direct neighbours and all citations between them (Marsden 2002; Wasserman and Faust 1994, 41–43). Collectively, these 69 core publications cite 1,859 other publications whose bibliographic data were in turn extracted from WOK to allow for the creation of an ego-network. Only 449 of these could be found, leaving 1,410 publications for which the bibliographic data are not available. It was decided not to extract this data manually given time constraints and the potential of introducing significant inconsistencies into the dataset.

By combining these 518 bibliographies a citation network could be created consisting of 33,556 publications, up to two steps away from the 69 egos, and linked by 42,993 citations. The aims of this case study and the large number of missing bibliographies were the main

[11] About the Book Citation Index: http://thomsonreuters.com/content/press_room/science/book-citation-index-launches (accessed 26.06.2014).

[12] http://thomsonreuters.com/products_services/science/free/essays/journal_selection_process/ (accessed 26.06.2014).

motivation for my decision to perform a citation network analysis on three analytical scales (Fig. 11). The *local scale* concerns the 69 archaeological applications and the citations between them (Table 2; discussed in section 3.5.1). The *meso scale* consists of the 69 egos, their direct neighbours and the citations between them, thus forming the ego-network of the 69 archaeological applications (discussed in section 3.5.2). Two different networks will be analysed on this scale to evaluate the influence of missing data: on the one hand the ego-network of 518 publications for which we have the complete dataset (referred to in the text as *WOK data network*), and on the other hand this ego-network with the 1,410 publications added for which the bibliography was not available on WOK (referred to in the text as *all data network*). The *global scale* includes the entire bibliographic dataset in its raw form, consisting of the archaeological ego-network and the publications cited by the direct neighbours of the egos (where this data are available; discussed in section 3.5.3).

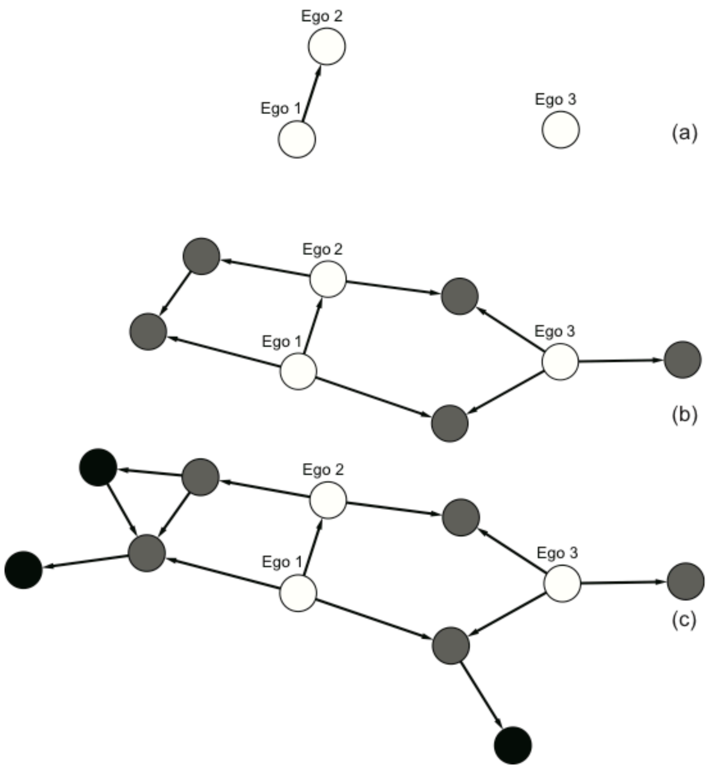


Fig. 11. The three scales of analysis used in this case study. (a) the local scale consisting of the egos and the citations between them (white nodes); (b) the meso scale consisting of the egos, their direct neighbours (grey nodes) and the citations between them; (c) the global scale consisting of the egos, their direct neighbours, the neighbours' neighbours (black nodes) and all available citations between them.

Table 2. Simple network measures (described in section 3.4) for all networks in the case study.

	Local scale Archaeological applications	Meso scale Ego-network		Global scale Combined dataset
		WOK data network	All data network	

Number of nodes	69	518	1928	33556
Number of arcs	128	1489	3987	42993
Missing bibliographies	0	0	1342	33032
Connected components	9	6	1	2
Isolated nodes	6	3	0	0
Network diameter	6	7	7	8
Network density	0.055	0.011	0.002	<0.001
Average shortest path length	1.689	2.377	2.355	3.191
Clustering coefficient	0.185	0.197	0.175	0.083
Average degree	3.71	5.75	4.14	2.562

3.4. Methods

As a result of their chronological evolution citation networks are *acyclic*, which means that there can be no closed loops of directed edges (Newman 2010, 69) (e.g. Fig. 12). The acyclic nature was enforced for the networks in this case study before any analysis took place. This allows for a number of techniques that can only be applied to acyclic networks and dramatically impacts the use and interpretation of many other techniques. The analyses and visualisations were executed with the software packages PAJEK (Nooy et al. 2005) and CYTOSCAPE (Smoot et al. 2011).

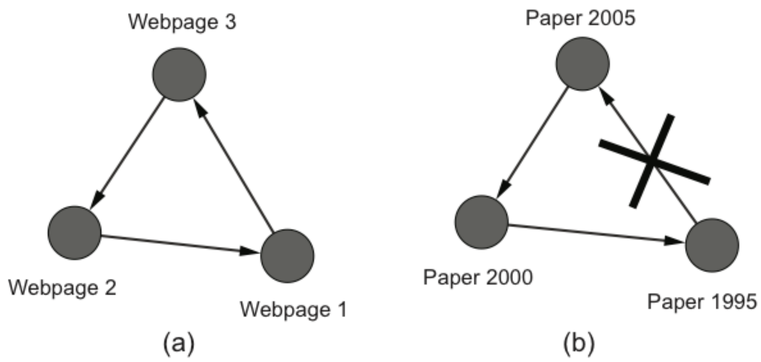


Fig. 12. Citation networks are acyclic like in (b), which shows that a paper published in 2005 can cite a paper published in 2000, which in turn can cite a paper published in 1995. The latter, however, cannot have a paper published in 2005 in its bibliography, thus preventing the creation of closed loops of arcs. This is not the case for many other networks like, for example, the World Wide Web (a).

A large number of exploratory network analysis measures are used in this paper, some are commonly applied to citation networks and others are not. In this section I will first describe general network measures that are often used to compare all sorts of networks (and which will also feature in the exploratory network analyses of case studies 2 and 3) and are therefore not specifically designed for analysing citation networks. Secondly I will introduce network measures describing aspects of the structural position of nodes, with reference to how these are interpreted within the context of citation networks. Lastly, I will introduce network projections and measures that can be used to analyse acyclic networks in particular.

A number of general network measures are used in Table 2 to compare networks. The *degree* of a node is the number of edges connected to it and the *average degree* is therefore the average of all degree scores in a single network (Newman 2010, 133-136; Nooy et al. 2005, 63-64). A *shortest path* (or geodesic path) in network terms is the shortest route over the network that runs from one vertex to another along the edges of the network and the *average shortest path length* is the average of all shortest path scores between all possible pairs of vertices in the network (Newman 2010, 136-140; Nooy et al. 2005, 127). The Watts-Strogatz *clustering coefficient* measures the average probability that two neighbours of a vertex are themselves neighbours as a ratio of the number of edges between the neighbours of a given node and the maximum number of edges that could possibly exist between these neighbours (Newman 2010, 262-266; Watts and Strogatz 1998, 441). The *degree distribution* represents the fraction of nodes in a network with a certain degree (Albert and Barabási 2002, 49; Newman 2010, 243-247). The network *diameter* is the length of the longest geodesic path between any pair of nodes in the network (Newman 2010, 140; Wasserman and Faust 1994, 111-112) and the number of *components* is the number of unconnected subgroups in networks (Newman 2010, 142; Nooy et al. 2005, 68). Network *density* is the fraction of all possible edges that are actually present (Newman 2010, 134-135).

The network analysis measures used in Tables 3-6 list node specific results for a number of simple network measures. I consider the *input domain* (which represents the number of all other vertices that are connected to a given vertex by a path (Nooy et al. 2005, 193)) of particular interest for analysing citation networks since older publications generally receive a higher score. It can therefore be used to evaluate the citation behaviour between the old and more recent archaeological applications of network methods. The *outdegree* is the number of arcs a given node sends and in a citation network represents the number of items in a publication’s bibliography. The *indegree* is the number of arcs a given node receives, which in a citation network equals the number of citations a publication receives. The indegree scores representing only citations within these relatively small networks are compared to the number of citations these publications received in total according to *Google Scholar* (which includes books but is less controlled and includes duplicates) and *Web of Knowledge* (which is manually edited, is more consistent but suffers from the selection criteria discussed above).

Hubs are publications that cite many other publications and many good authorities in particular. *Authorities* are publications that are cited by many other publications and by good hubs in particular (Kleinberg 1999). Hubs can be regarded organisational leaders whilst authorities are intellectual leaders (White 2011, 276). The *citation weights* measure is used to identify the path traversed by the largest number of other paths (Hummon and Doreian 1989). This path can therefore be considered the most important path of influence (as represented through citation) (Hummon et al. 1990, 464). The Simple Path Count (SPC) approach to citation weights was chosen since all three methods discussed by Hummon and Doreian (1989) provided the same results and SPC should be considered a first choice (Batagelj 2003, 24). For the citation weights the networks were transposed to represent the flow of influence, changing the *X cites Y* citation networks into *Y is cited by X* (Bataglej 2003).

Acyclic networks can also be transformed into undirected networks and analysed through the cocitation and bibliographic coupling network projections (Fig. 13) (Newman 2010, 115-118). In a *cocitation network* edges are drawn between a pair of vertices if they are both cited by the same paper(s). In a *bibliographic coupling network*, on the other hand, edges are drawn between a pair of vertices if they both cite the same paper(s). Bibliographic coupling networks do not change over time since they are derived from published bibliographies whilst cocitation networks evolve as citations to papers emerge (often only months or years after publication). Both networks are considered a good indicator for clusters of papers that deal with related topics (Newman 2010, 116-117). A number of techniques are used in this case study to analyse these two network types. *M-slices* include vertices linked by lines with a value equal to or greater than *m* (Nooy et al. 2005, 109), which will be used to explore nested subnetworks with a minimum cocitation or bibliographic coupling edge weight. *Betweenness centrality* is the proportion of all shortest paths between pairs of other vertices that include this vertex (Nooy et al. 2005, 131), used in Figures 17 and 18 to identify nodes that bridge clusters of different topics. *Hierarchical clustering* through Ward’s clustering method was used to identify ‘communities’ with similar citation behaviour. The ‘corrected Euclidean distance’ algorithm used in PAJEK was applied to compute dissimilarity between vertices since it takes the value of lines into account (Batagelj et al. 1992; Nooy et al. 2005, 265–73).

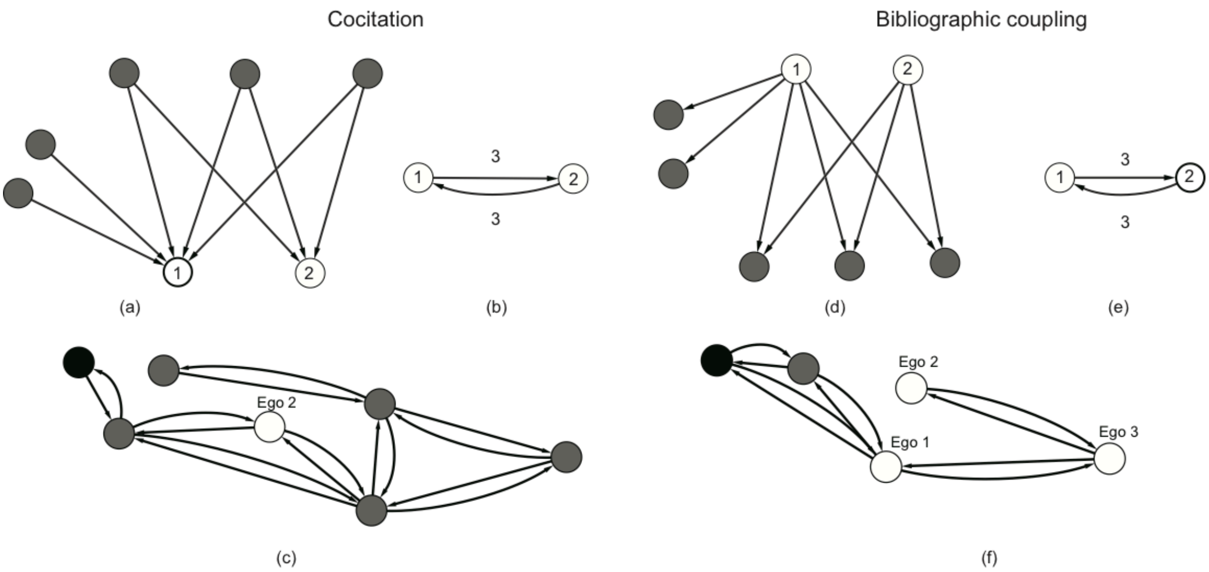


Fig. 13. Network (b) is the cocitation projection of network (a), e.g. network (c) is the cocitation projection of network (c) in Fig. 1. Network (e) is the bibliographic coupling projection of network (d), e.g. network (f) is the bibliographic coupling projection of network (c) in Fig. 11.

3.5. Results

In this case study a citation is taken to represent influence and relatedness of subject matter (Newman 2010, 67-70). The citation networks used do not reveal the nature of the citation (e.g. positive or negative; for an experimental approach to include this see Alen Sula and Miller 2014). The absence of citations is considered equally important as their presence, although purposefully not citing is not directly identifiable. Box 4 provides an overview of how the discussion of aims, research questions, data collection, and methods above reveal my thinking through a network science research process. In this section I will describe the results that this research process have led me to, the significance of which will be discussed in the next section.

Box 4. Network science research process case study 1

Phenomenon studied:

- The adoption and adaptation of network science techniques by archaeologists.

Abstraction as network concepts:

- Only one formal representation of adoption and adaptation is studied here: citation behaviour.
- Publications are individual entities of research interest.
- One publication can affect another, this process of influence is formalised as citations.
- Citations always describe a directed relationship: the flow of influence runs from an older to a younger publication. However, a citation represents a reference to the source of influence and therefore runs from the younger to the older publication.

Representation as network data:

- Publications are represented as nodes.

- Citations are represented as arcs going from the younger to the older publication.
- Citation networks are acyclic.

Dependence assumptions:

- Frequently cited publications tend to attract more citations over time (authorities).
- Publications with similar subject matter tend to cite each other more than other publications (community formation).

Network science techniques used:

- Network-based exploratory network measures: average degree, average shortest path length, clustering coefficient, degree distribution, diameter, connected components, density.
- Node-based exploratory network measures: input domain, outdegree, indegree, m-slices, betweenness centrality, hierarchical clustering.
- Techniques for acyclic networks: hubs, authorities, citation weights, cocitation, bibliographic coupling.

3.5.1. Local scale: network of archaeological applications

A visual exploration and the simple network measures in Table 2 reveal that the bulk of publications are part of a single large connected component, whilst only ten publications are either isolated or tied to only one other vertex. These isolated nodes include studies of network analysis in GIS (e.g. Lock and Pouncett 2007; Pouncett and Lock 2007; Zubrow 1990), and the bibliometric studies by Schich and colleagues (Schich et al. 2009; Schich and Coscia 2011). The network diameter and average shortest path length are both low, which seems to be the result of a few nodes being far better connected than other nodes in the network, given the low average degree and clustering coefficient.

One of the patterns observed in the literature review is immediately apparent from the chronological visualisation (Fig. 14): archaeological applications of formal network methods are rare before the late 1990s, and especially in the last decade there seems to have been a dramatic increase in the number of published archaeological applications (see also Fig. 5). Something that was less clear in the literature review but obvious from this citation network visualisation is the sheer density of citations between publications in the last decade, in contrast to the relative scarcity of citations between the older publications themselves and between the older and more recent ones. The citation behaviour seems to indicate a divide between an early group of adopters and a more recent one. Some network measures can help explore this pattern (see Table 3).

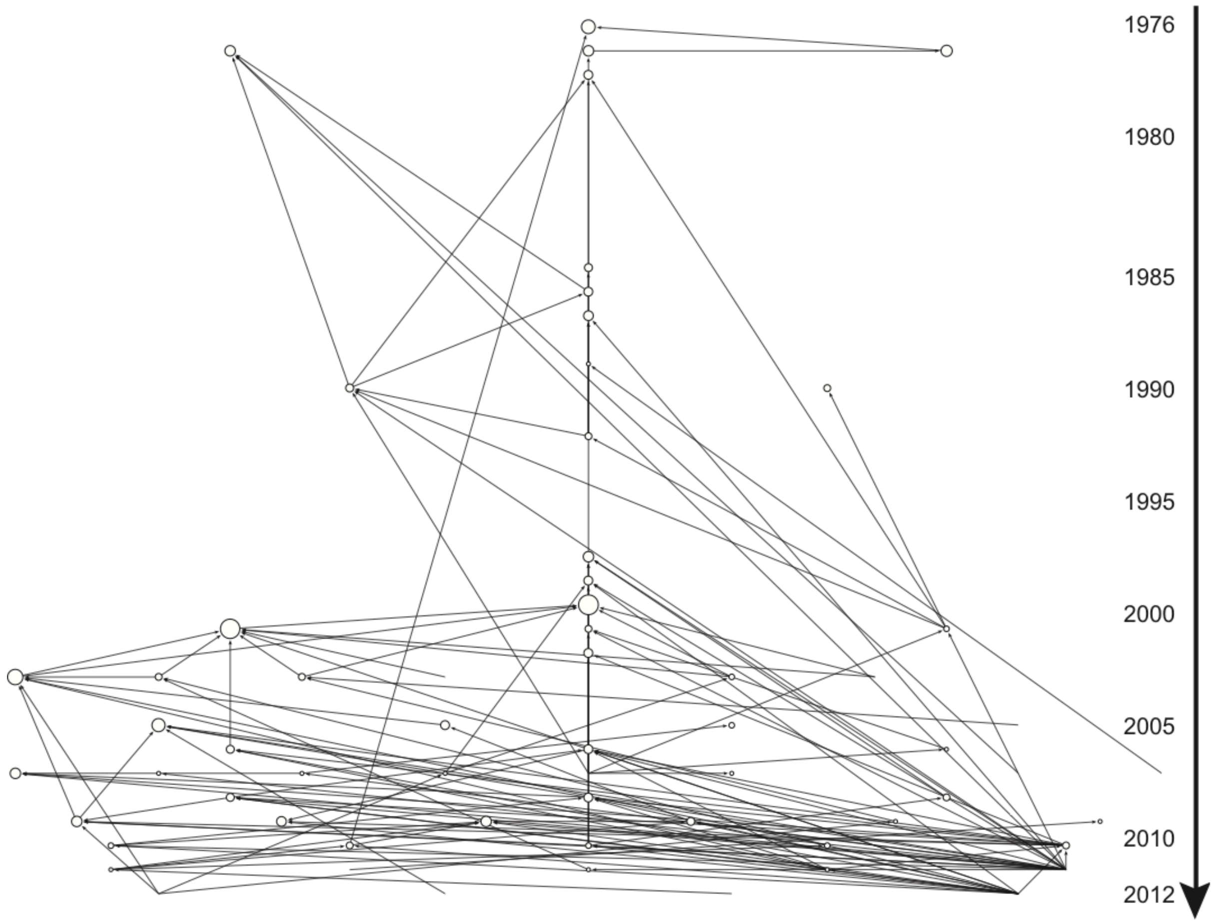


Fig. 14. Chronological visualisation of the local scale network of the 69 archaeological applications of network methods before 2012. Nodes represent publications and arcs represent citations. The size of nodes represents the relative input domain scores. Nodes are positioned along the Y-axis according to their date of publication. The positioning of nodes along the X-axis does not reflect any logic other than avoiding the overlap of nodes and arcs to make the patterns more visually 'readable'.

Table 3. Node specific measures for the local scale network. See Appendix II section 10.1 for full references.

Publication	Input domain	Indegree	Outdegree	Citations WOK	Citations Google Scholar
Bentley and Maschner, 2000	25	5	0	7	17
Bentley and Maschner, 2001	24	8	1	10	17
Bentley and Maschner, 2003a	3	2	2		
Bentley and Maschner, 2003b	0	0	0		
Bentley and Shennan, 2003	15	6	2	42	97
Bentley and Shennan, 2005	5	1	1	11	30
Bentley et al., 2005	0	0	1	2	5
Bentley, 2003a	3	2	2		10
Bentley, 2003b	0	0	2		8
Bintliff, 2003	0	0	0		4
Branting, 2007	0	0	1		
Brughmans, 2010	3	3	9	0	4
Brughmans, 2012b	0	0	15		0
Cochrane and Lipo, 2010	2	1	1	1	4
Collar, 2007	1	1	0	1	6
Coward and Gamble, 2008	3	1	1	8	10
Coward, 2010	2	2	5		1

Earl and Keay, 2007	0	0	6		0
Earl et al., 2011	1	1	6		0
Evans et al., 2009	7	5	2		15
Gamble, 1998	7	4	0	38	69
Gamble, 1999	5	4	0		718
Graham, 2006a	4	3	1		5
Graham, 2006b	5	5	1		13
Graham, 2009	6	3	0		4
Hage et al., 1986	4	1	0	4	7
Hamilton et al., 2007	1	1	2	30	46
Hart and Engelbrecht, 2011	1	1	4		1
Hart, 2012	0	0	1	0	0
Hill, 2009	1	1	1		1
Hunt, 1988	6	3	1		17
Irwin-Williams, 1977	7	4	0		23
Irwin, 1978	5	3	0	17	30
Isaksen, 2007	1	1	0		2
Isaksen, 2008	4	4	2		4
Jenkins, 2001	2	2	4	7	7
Jiménez and Chapman, 2002	5	2	0		5
Keay and Earl, 2006	1	1	1		2
Knappett and Nikolakopoulou, 2005	11	5	0		6
Knappett et al., 2008	5	5	2	9	17
Knappett, 2011a	1	1	1		
Knappett, 2011b	0	0	21		2
Knappett, 2012	0	0	1		
Kohler, 2012	0	0	4		1
Lock and Pouncett, 2007	0	0	1		
Mackie, 2001	3	3	1		7
Maschner and Bentley, 2003	2	1	2		11
McGlade, 2003	0	0	1		9
Milicic, 1993	3	1	4	5	8
Mizoguchi, 2009	7	6	0	4	9
Munson and Macri, 2009	4	3	2	3	4
Peregrine, 1991	4	4	3	8	18
Pouncett and Lock, 2007	0	0	0		
Rothman, 1987	5	1	1		7
Santley, 1991	3	1	0		
Schich and Coscia, 2011	0	0	1		
Schich et al., 2009	1	1	0		3
Shuchat, 1984	0	0	0	6	13
Sindbæk, 2007a	7	4	0	7	18
Sindbæk, 2007b	1	1	2	3	5
Swanson, 2003	0	0	0	6	11
Terrell, 1976	12	2	1	11	28
Terrell, 1977a	8	2	1	8	22
Terrell, 1977b	7	1	1		31
Terrell, 2010a	2	2	2		6
Terrell, 2010b	3	2	2	4	8
Welinder and Griffin, 1984	0	0	0	1	1
Wells, 2005	2	1	0	2	7
Zubrow, 1990	1	1	0		16

Three publications (Brughmans 2010; Brughmans 2012b; Knappett 2011a) have a far higher outdegree score than others and are to a large extent responsible for the higher density of citations in the group of more recent applications. These are all reviews of network science in archaeology and two of them are my own, emphasizing the impact of my own knowledge of existing applications on the creation of this dataset. This bias will be evaluated below by removing these three publications from the network. With the exception of these three all archaeologists cite six or less other archaeological applications, whilst the vast majority cites only one or none at all. More than half of all nodes have an indegree score of either zero or one. A few publications are cited more frequently than others, although in some cases this seems to be at least in part due to multiple papers by the same authors citing their earlier work on the same topic. The publications cited five or more times within this network were all published after 1999. The external citation counts generally show a similar pattern to the indegree scores in this network, although there are some interesting discrepancies at the lower and upper ends of the indegree scores. A number of publications with only one citation within this network have a significantly higher score outside of the network. Most of the publications that received many citations within the network, however, do not rank among the highest external citation scores. This can only in part be explained by their relatively recent publishing date and seems more related to the influence of these authors to other archaeological network analysts as well as citation of one’s own work in some cases.

Although the results in Table 3 indicate that many publications published before 1998 do have higher than average input domain scores (e.g. Terrell 1976; Terrell 1977b; Irwin-Williams 1977) it is striking that the top three highest scores are publications from the early 2000s that pioneered the application of complex network models in archaeology (Bentley and Maschner 2000; 2001; Bentley and Shennan 2003). The influence of early archaeological adopters is also visible in the citation weights SPC measure, which again distinguishes between an older and a more recent group of applications by indicating that ties between recent publications are traversed more frequently by all possible paths over this network (Fig. 15). The absence of older publications within the list of hubs and authorities can be seen as a reflection of the higher internal density of citations within the group of recent applications.

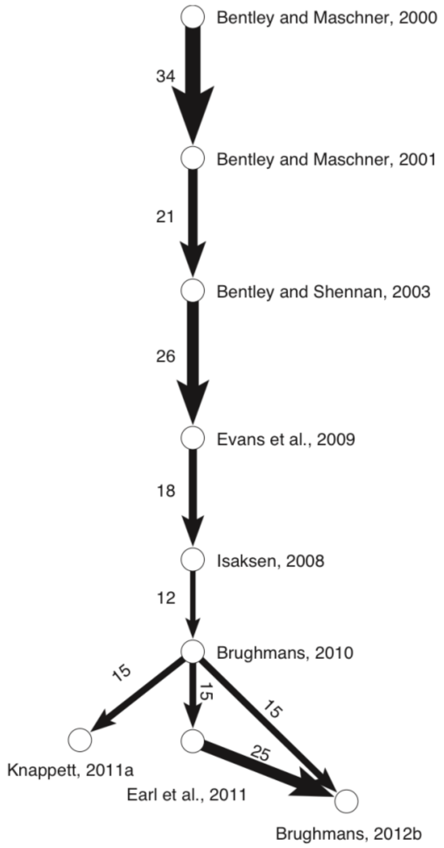


Fig. 15. SPC citation weight path of the local scale network. Nodes represent publications, arcs represent an inversed citation (i.e. *X is cited by Y*) and arc weights represent the number of paths passing through it.

Many of the more recently published works are isolated nodes in the cocitation network (Fig. 16a) since they had less time to accumulate citations. Most interestingly, however, the older applications are also isolated in the cocitation network because they are rarely cited together. The only connected component in this network represents recent publications being cited together. It would be wrong to argue that all recent publications cite many other recent applications. By removing my own review publications (Brughmans 2010; 2012b) and the review by Knappett (2011a), which were considered hubs in the network and have the highest outdegree, it becomes clear how limited cocitation of archaeological applications really is (Fig. 16b). The network fragments into two smaller components with no more than one cocitation per pair of publications. The bibliographic coupling network shows a similar pattern (Fig. 17). The older publications are again isolated, this time largely because there are hardly any older publications in this network that could create a tie between a pair. The connected component of this network therefore represents recent publications that have similar bibliographies. The core of this component consists of the three review publications that have a similar bibliography. When these three are extracted from the bibliographic coupling network it falls apart into smaller components (Fig. 17b). The bibliographies of the remaining publications are quite different and the majority of clusters in this network can be attributed to authors citing their own previously published work, or to topical, national and institutional contacts between authors.

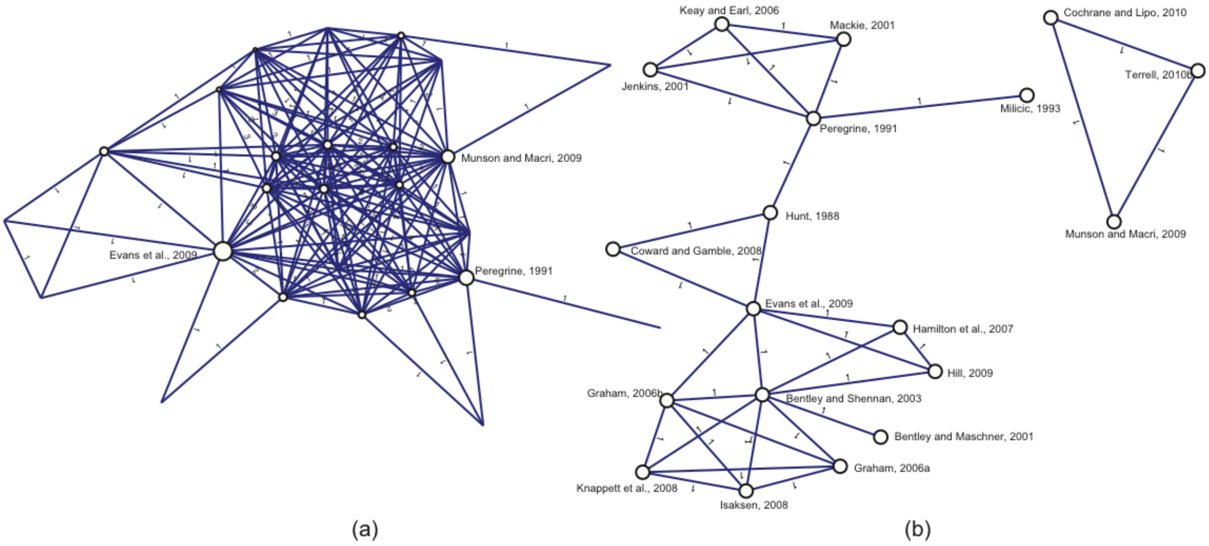


Fig. 16. Cocitation network of the local scale network with (a) and without (b) three recent review publications (Brughmans 2010; 2012b; Knappett 2011). Size of nodes in network (a) represent betweenness centrality: the proportion of all shortest paths between pairs of other nodes that include this node (Nooy et al. 2005, 131).

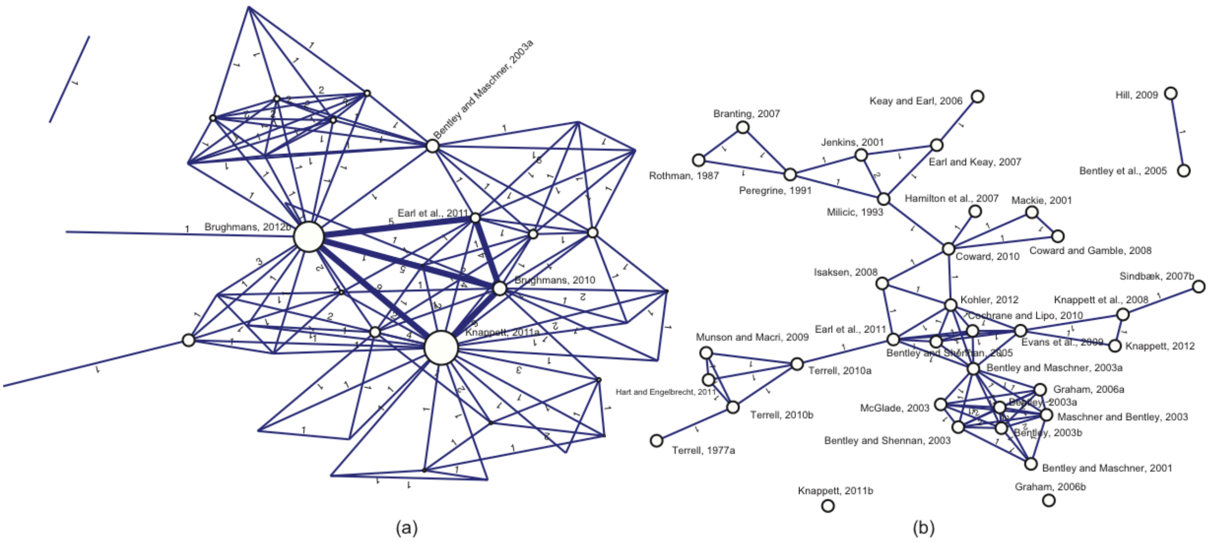


Fig. 17. Bibliographical coupling network of the local scale network with (a) and without (b) three review publications (Brughmans 2010; 2012b; Knappett 2011a). Size of nodes in network (a) represent betweenness centrality: the proportion of all shortest paths between pairs of other nodes that include this node (Nooy et al. 2005, 131).

3.5.2. Meso scale: ego-network of archaeological applications

WOK data network

This network includes the 518 publications for which the full bibliographies are available. The few smaller components and isolated nodes are all archaeological applications of network techniques that have strongly different citation behaviour from all other archaeological applications. Although this network is denser and nodes have a higher average degree than the local scale network, the diameter and average shortest path length remain low.

A chronological visualisation of this network (Fig. 18) reveals that publications dating back to 1964 are included, although the citation pattern is most dense from the late 90s onwards. Influential SNA publications are included among the older ones whilst influential studies in physics are only present since the late 90s. The network measures in Table 4 indicate, however, that these physics papers seem to be most prominent in this network. The highest number of citations both within this network and according to external citation indices are for the extremely influential models by Watts and Strogatz (1998) and Barabási and Albert (1999), as well as other influential publications in social physics. Most of these are even considered authorities within the network, whilst the hubs include a number of archaeological publications that cite many of these influential papers in physics. Granovetter's (1973) work '*The strength of weak ties*' is arguably the only SNA publication in this indegree top ten list, since many of the most prominent SNA publications are books and therefore not included in the current network. Most interestingly, the input domain measure picks out those physics papers from the late 1990s that were responsible for the emergence of what is sometimes referred to as '*The new science of networks*', which I have argued in the literature review in chapter 2 was key to the renewed interest in network techniques by archaeologists. Only a few older important publications are drawn out, notably Korte and Milgram's (1970) review of the '*Small World*' experiments and Bak, Tang and Wiesenfeld's (1987) original paper on Self-organized Criticality. The citation weight SPC measure again emphasizes the prominence of many of these physics papers (Fig. 19).

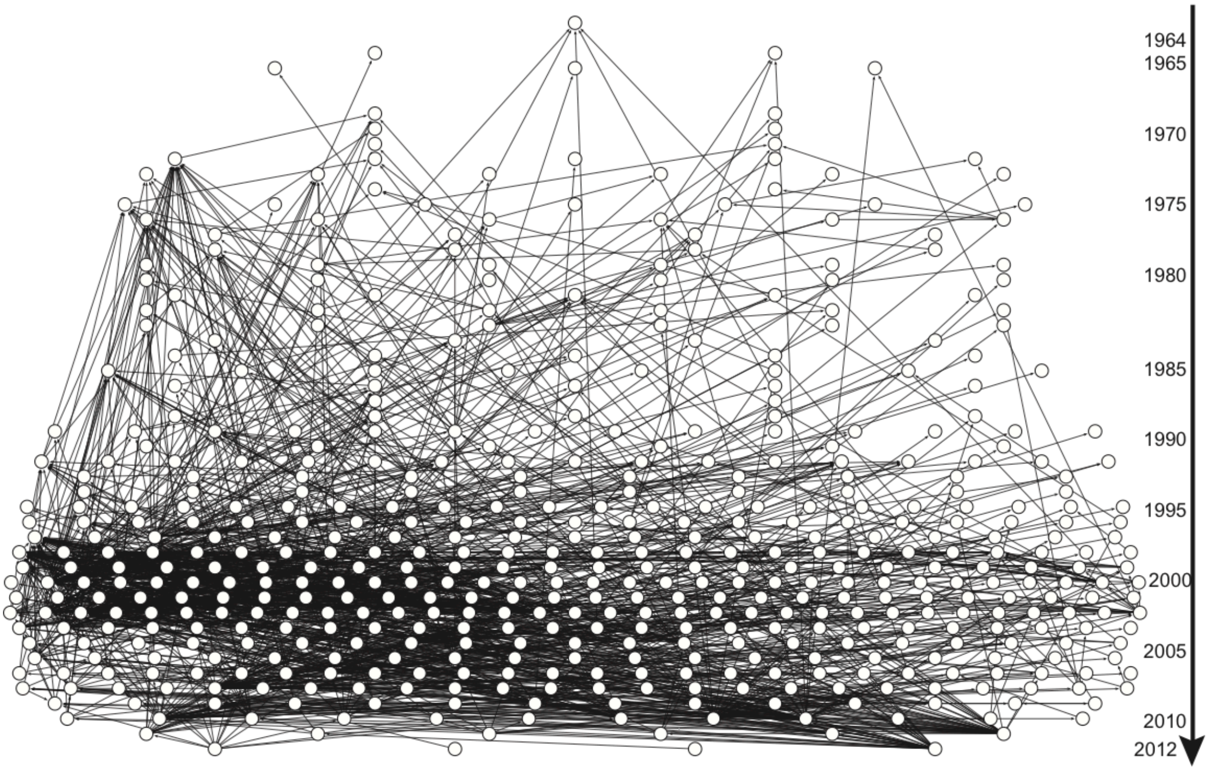


Fig. 18. Chronological visualisation of the meso scale WOK data network discussed in Section 3.5.2. Nodes represent publications and arcs represent citations.

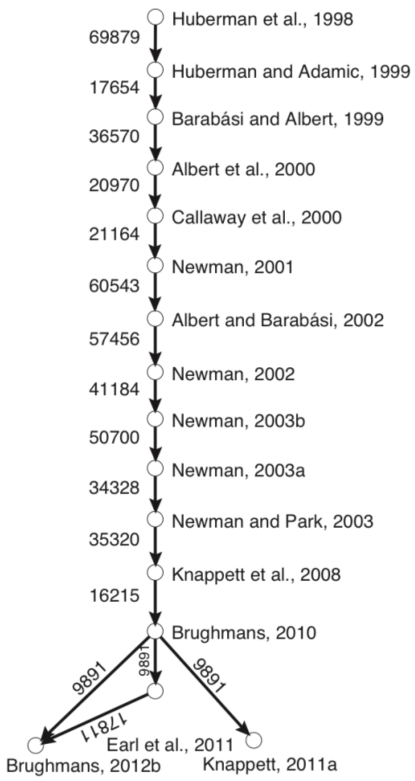


Fig. 19. SPC citation weight path of the meso scale WOK data network discussed in Section 3.5.2. Nodes represent publications, arcs represent an inversed citation (i.e. *X is cited by Y*) and arc weights represent the number of paths passing through it.

At the core of the cocitation network lie the publications with the highest indegree (Table 4) that are also most frequently cited together (m-slices 9-20). A hierarchical clustering of the dissimilarity matrix of this network further reveals groups of publications that are considered similar because they are often cited together. The frequently cited physics papers are most dissimilar to all other publications in the network. The bulk of the egos are grouped in one cluster that also includes influential SNA publications. Other clear clusters include works in archaeology/anthropology, biology and less frequently cited physics papers. Publications from the social and behavioural sciences are surprisingly peripheral or indeed absent from this network. Results of the bibliographic coupling network indicate that bibliographies of many egos are most similar (in many cases this can be attributed to multiple publications on a similar topic by the same authors), along with those of influential and review papers in physics (m-sclices 6-34). A hierarchical clustering of the dissimilarity matrix of this network also picks out these egos and physics publications as being dissimilar to the rest of the network, whilst other topical clusters are less immediately obvious than for the cocitation network. In fact, clusters seem to be comprised of a few egos and their direct influences from SNA, physics and archaeology/anthropology. Extracting the three archaeological review papers (Brughmans 2010; 2012b; Knappett 2011a) does not have a significant impact on these networks as it had on the local scale of analysis.

Table 4. Node specific measures for the top ten scoring nodes in the meso scale WOK data network discussed in Section 3.5.2. See Appendix II section 10.2 for full references.

Input domain		Outdegree		Indegree		WOK citations	Google Scholar citations	Hubs	Authorities
Watts and Strogatz, 1998	115	Coward and Gamble, 2008	50	Barabási and Albert, 1999	45	7192	13619	Evans, 2004	Barabási and Albert, 1999
Albert et al., 1999	112	Knappett, 2011b	47	Watts and Strogatz, 1998	44	6997	15811	Bentley and Shennan, 2003	Watts and Strogatz, 1998
Huberman et al., 1998	112	Bentley and Maschner, 2001	41	Albert et al., 1999	34	1542	3104	Newman, 2003	Albert et al., 1999
Huberman and Adamic, 1999	111	Terrell, 2010b	40	Albert and Barabási, 2002	28	5687	10223	Newman, 2001	Amaral et al., 2000
West et al., 1997	106	Bentley and Shennan, 2003	39	Granovetter, 1973	24	4965	20156	Watts, 2004	Newman, 2001
Banavar et al., 1999	103	Brughmans, 2010	36	Amaral et al., 2000	20	1032	1856	Ravasz and Barabási, 2003	Pastor-Satorras and Vespignani, 2001
Arthur, 1999	100	Brughmans, 2012b	36	Newman, 2001	19	489	1082	Bentley, 2003a	Albert et al., 2000
Barabási and Albert, 1999	99	Cochrane and Lipo, 2010	35	Pastor-Satorras and Vespignani, 2001	19	1230	2090	Fortunato, 2010	Faloutsos et al., 1999
Bak et al., 1987	97	Bentley, 2003a	34	Newman, 2003	19	4005	7829	Bentley, 2003b	Liljeros et al., 2001
Korte and Milgram, 1970	92	Newman, 2003	33	Albert et al., 2000	19	1963	3453	Albert and Barabási, 2002	Albert and Barabási, 2002

Table 5. Node specific measures for the top ten scoring nodes in the meso scale All data network discussed in Section 3.5.2. See Appendix II section 10.2 for full references.

Input domain		Outdegree		Indegree		WOK citations	Google Scholar citations	Hubs	Authorities
Wasserman and Faust, 1994	121	Wells, 2005	345	Barabási and Albert, 1999	45	7192	13619	Bentley and Shennan, 2003	Barabási and Albert, 1999
Kochen, 1989	117	Coward and Gamble, 2008	135	Watts and Strogatz, 1998	44	6997	15811	Newman, 2003	Barabási, 2002
Zipf, 1949	116	Gamble, 1998	122	Wasserman and Faust, 1994	39	na	12335	Watts, 2004	Watts and Strogatz, 1998
Watts and Strogatz, 1998	114	Brughmans, 2010	112	Albert et al., 1999	34	1542	3104	Bentley, 2003a	Albert et al., 1999
Erdós and Rényi, 1960	114	Jenkins, 2001	107	Boyd and Richerson, 1985	34	na	4140	Fortunato, 2010	Amaral et al., 2000
Mitchell, 1969	112	Harary and Norman, 1953	100	Albert and Barabási, 2002	28	5687	10223	Bentley, 2003b	Newman, 2001
Albert et al., 1999	111	Munson and Macri, 2009	92	Granovetter, 1973	24	4965	20156	Brughmans, 2010	Pastor-Satorras and Vespignani, 2001
Huberman et al., 1998	111	Bentley and Shennan, 2003	80	Amaral et al., 2000	20	1032	1856	Brughmans, 2012b	Albert et al., 2000
Crow and Shimizu, 1988	111	Cochrane and Lipo, 2010	79	Newman, 2001	19	489	1082	Albert and Barabási, 2002	Albert and Barabási, 2002
Huberman and Adamic, 1999	110	Bentley and Maschner, 2001	79	Pastor-Satorras and Vespignani, 2001	19	1230	2090	Knappett, 2011a	Wasserman and Faust, 1994

Table 6. Node specific measures for the top ten scoring nodes in the global scale network discussed in section 3.5.3. See Appendix II section 10.2 for full references.

Input domain		Outdegree		Indegree		WOK citations	Google Scholar citations	Hubs	Authorities
Milgram, 1967	184	Cross and Hohenberg, 1993	1237	Barabási and Albert, 1999	45	7192	13619	Evans, 2004	Barabási and Albert, 1999
Mitchell, 1969	153	McBrearty and Brooks, 2000	946	Watts and Strogatz, 1998	44	6997	15811	Newman and Park, 2003	Watts and Strogatz, 1998
Rogers, 1962	143	Fonagy et al., 2007	459	Wasserman and Faust, 1994	39	na	12335	Albert and Barabási, 2002	Albert and Barabási, 2002
Wasserman and Faust, 1994	142	Turcotte, 1999	457	Albert et al., 1999	34	1542	3104	Newman, 2003	Albert et al., 1999
Zipf, 1949	141	Newman, 2003	428	Boyd and Richerson, 1985	34	na	4140	Newman, 2000	Amaral et al., 2000
Harvey and Pagel, 1991	138	Flinn, 1997	427	Albert and Barabási, 2002	28	5687	10223	Newman, 2001	Newman, 2001
Travers and Milgram, 1969	138	Fortunato, 2010	401	Granovetter, 1973	24	4965	20156	Watts, 2004	Pastor-Satorras and Vespignani, 2001
Homans, 1950	137	Dietler, 1997	375	Amaral et al., 2000	20	1032	1856	Ravasz and Barabási, 2003	Amaral et al., 2000
Peters, 1986	137	Gronenborn, 1999	372	Newman, 2001	19	489	1082	Fortunato, 2010	Faloutsos et al., 1999
Kochen, 1989	136	Terrell et al., 1997	355	Pastor-Satorras and Vespignani, 2001	19	1230	2090	Jin et al., 2001	Wasserman and Faust, 1994

All data network

This network consists of the 69 egos and all of their first neighbours. The network consists of one large connected component and a preliminary visual exploration reveals that this component shows a topical or disciplinary clustering with a far larger number of archaeological publications than the previous network. The chronological visualisation of this network shows that the period between 1970 and 2012 is particularly dense, whilst the oldest publication dates back to 1854 (Fig. 20). Although this pattern is of course caused by the sampling strategy of selecting the first neighbours of the egos and the scientific tendency to cite recent papers, I believe this is at least in part a reflection of the existence of communities of academics using network methods in an increasingly formalised way since the late 1960s (Freeman 2004). Although this network is much larger, the diameter and average shortest path length of both meso scale networks are almost the same. Degree scores are very different, however, due to the fact that the vast majority (1,533) of publications in this network have no outgoing citations and the network is less dense as a result of this.

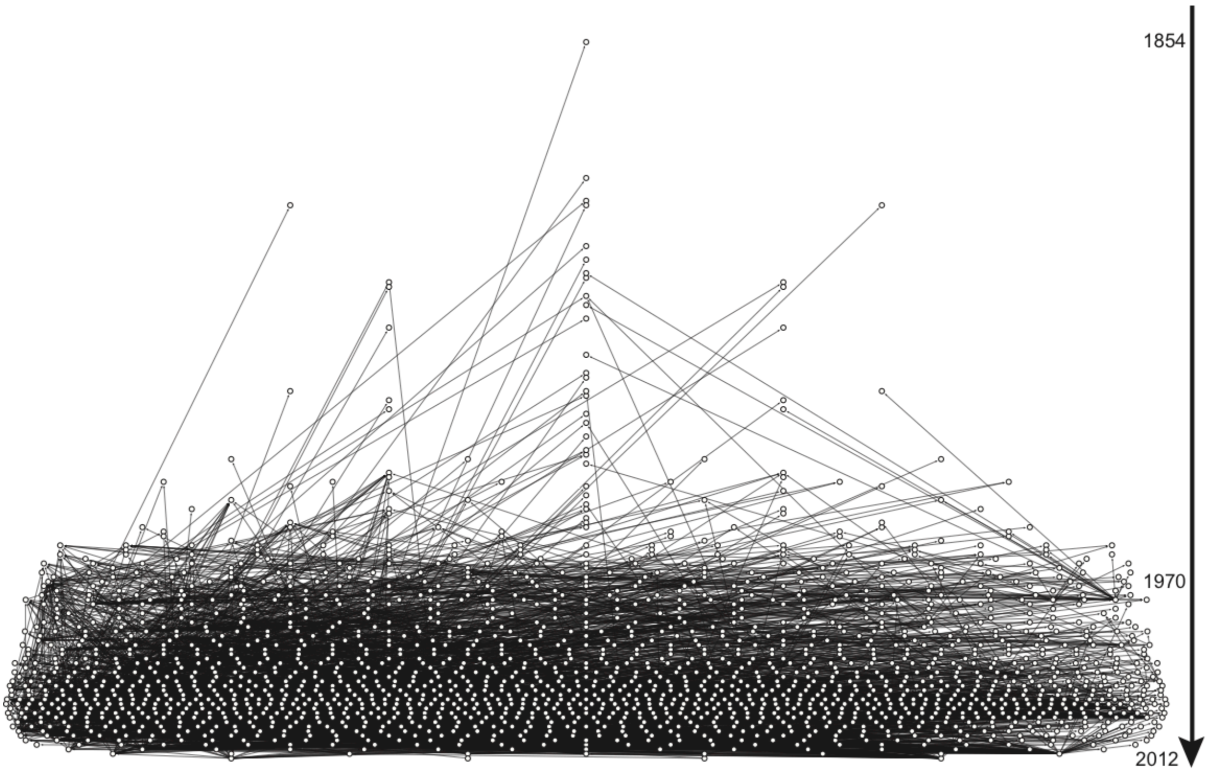


Fig. 20. Chronological visualisation of the meso scale All data network discussed in Section 3.5.2. Nodes represent publications and arcs represent citations.

The difference between both meso scale networks is most striking in light of the network analysis results shown in Table 5. Most crucially, influential SNA books now occupy prominent positions even though their bibliographies are not included in the network. The input domain measure this time emphasises older publications that are frequently cited by physicists (e.g. Erdós and Rényi 1960; Zipf 1949) as well as publications at the end of the 1990s that stimulated the renewed interest in complex networks. The SNA textbook by Wasserman and Faust (1994) has the highest input domain score, however, and is also cited very frequently within and outside of this network. Another newcomer is *Culture and the evolutionary process* by Boyd and Richerson (1985), which is frequently cited by archaeologists/anthropologists. With the exception of these two, the top ten list of indegree scores is still dominated by the same influential physics papers. The full bibliographies of the egos are included in this network and the outdegree score (345) of Wells (2005) in particular stands out. This publication is in fact peripheral in the network, since his use of network techniques is limited to mentioning the ubiquity of relationships within complex systems and referring to Barabási's (2002) popular science book *Linked*, but the sheer size of its bibliography caused it to dominate the hubs and authorities measures. I decided to remove Wells (2005) from the network and recalculate the hubs and authorities (Table 5). The authorities hardly changed compared to the WOK data network, although Wasserman and Faust (1994) is now included. The citation weight SPC measure (Fig. 21) reveals that the prominent position of Wasserman and Faust (1994) within this network might be due to its influence on Watts and Strogatz (1998). Indeed, the cocitation network shows Wasserman and Faust (1994) along with a few other SNA publications (e.g. Scott 2000; White et al. 1976) at its core (m-slices 11-34), whilst most other SNA works are clustered together in the periphery. The list of hubs is also similar to the previous network although it now includes three reviews of archaeological network analysis (Brughmans 2010; 2012b; Knappett 2011a) because these cite many of the authorities in this network. The bibliographic coupling network is the same as for the previous network because no new bibliographies are added.

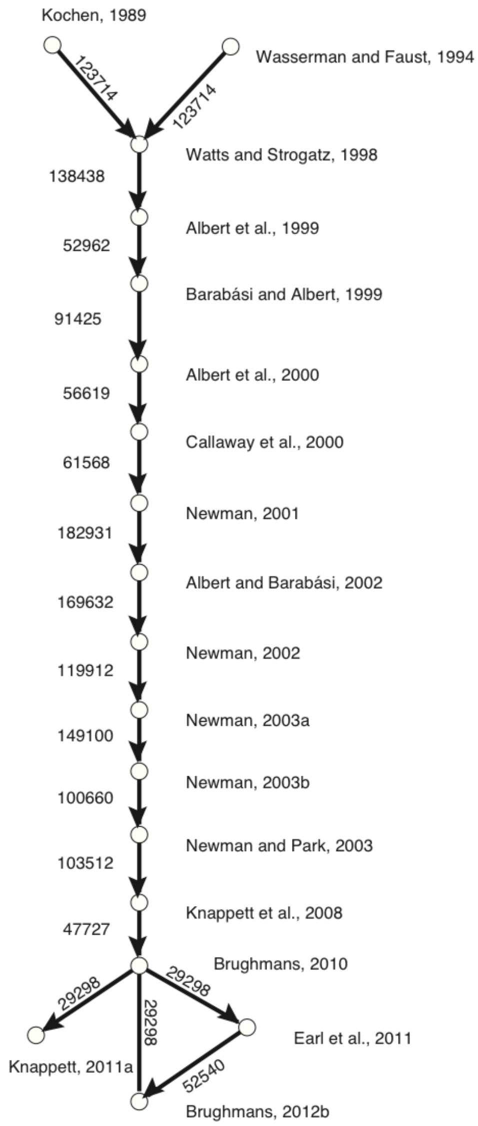


Fig. 21. SPC citation weight path of the meso scale All data network discussed in Section 3.5.2. Nodes represent publications, arcs represent an inversed citation (i.e. *X is cited by Y*) and arc weights represent the number of paths passing through it.

3.5.3. Global scale: combined dataset

Although this network is less dense than any of the preceding ones and has far more nodes, the average shortest path length and diameter are still low. The input domain (Table 6) picks up old publications on popular topics like small-worlds (Milgram 1967; Travers and Milgram 1969; Kochen 1989), the diffusion of innovations (Rogers 1962), Zipf's law (Zipf 1949) as well as a few sociological books (Homans 1950; Wasserman and Faust 1994). This seems to indicate that compared to the previous networks books are more prominent at this global scale of analysis. The outdegree results show the exceptionally large bibliographies of Cross and Hohenberg (1993) and McBrearty and Brooks (2000). These two publications significantly affected the hubs and authorities (the first emphasizing publications on self-organised criticality and fractals, the second on evolutionary anthropology) and have been removed before the hubs and authorities listed in Table 6 were calculated. These are very similar to the results of the networks described in Section 3.5.2 (Tables 4-5) without the archaeological publications. The citation weight SPC measure (Fig. 22) still focuses on popular physics publications, although there are now two branches influencing recent archaeological applications.

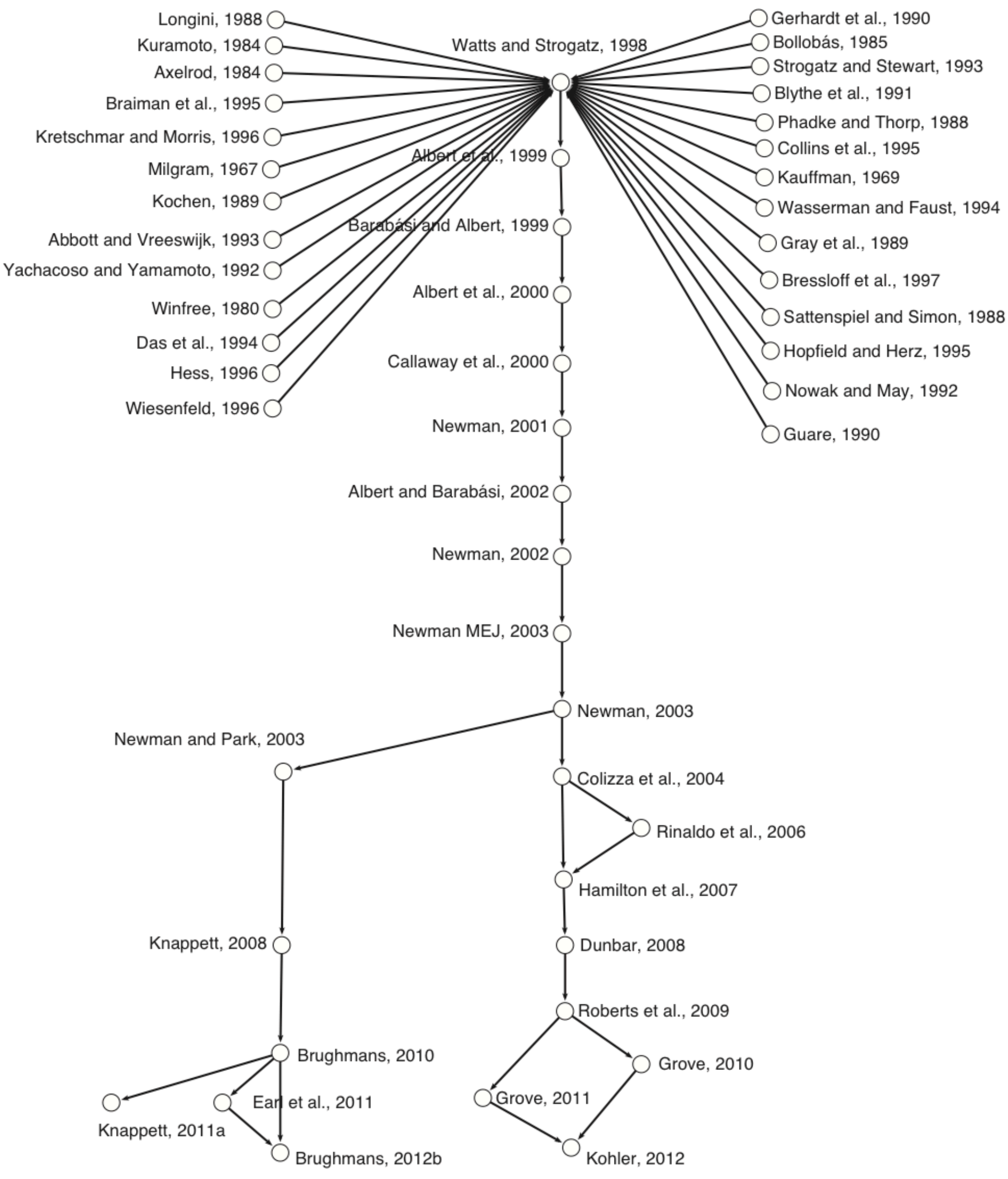


Fig. 22. SPC citation weight path of the global scale network. Nodes represent publications, arcs represent an inversed citation (i.e. *X is cited by Y*) and arc weights represent the number of paths passing through it.

Listing the top ten high scores as in Table 6 is almost senseless for such a large network. Moreover, the top ranking publications are very similar to those of the previous smaller networks, whilst the sheer size of the current network holds the potential of revealing publications not listed in the top ten with possibly interesting and novel archaeological applications. Some of these might be revealed through a close scrutiny of the ‘communities’ of similar citation practice illustrated by the cocitation and bibliographic coupling networks. This is sadly not within the scope of the current case study. It suffices to mention that the latter two network projections show a similar disciplinary clustering as was observed on the meso scale of analysis. The core of the cocitation network (m-slices 8-33) is similar to the All data network described in Section 3.5.2.

3.6. Discussion

A very large number of network science techniques was used in this case study, and the implications of their results for this PhD project should be made clear. In this section I will discuss how this long technical description of the exploratory citation network analysis results allows me to answer the research questions set out at the start of this case study.

> Question 1: can the intuitive results drawn from a close reading be quantitatively expressed and validated using citation network analysis?
>
> Question 2: can an exploratory network analysis provide new insights into the citation behaviour within the limited bibliographic corpus of the close reading?

On the local scale of analysis the citation networks confirmed some of the intuitive results of the literature review: archaeological applications of formal network methods before the 1990s are rare, the last decade has seen a strong increase of applications and these recent applications are rarely influenced by the older ones. Two research traditions have been most influential to archaeological network analysts: social network analysis (SNA) and complex network studies in physics. The citation networks clearly showed that some methods and models in both SNA and physics are significantly more often applied (or at least cited) than others, and archaeologists have initially focused on these few popular approaches that have proven their worth in other disciplines (e.g. small-worlds, centrality measures). The results also provided a new perspective, however, by identifying the strong differences in citation densities between the groups of old and more recent archaeological applications. The input domain results are the best example of this. The input domain on the meso and global scale picked out the first publications of particular methods and theories that have been extremely influential (especially in physics), whilst on the local scale more recent archaeological applications receive the highest scores. The higher citation density caused the recent applications to be most prominent according to most of the network measures used. It is tempting to interpret this in light of the observation that many of these recent applications use a limited set of similar network techniques. This methodological similarity does not explain the higher citation density, however, since popular methods were mainly adopted from a few non-archaeological publications. In fact, the high citation density was largely the result of a few recent review publications, which were also almost exclusively responsible for the few citation links bridging the recent and the older applications. Recent archaeological applications, therefore, seem largely uninfluenced by other archaeological applications (or did not make possible influence explicit through citation).

> Question 3: does a citation network analysis allow one to usefully contextualise a close reading within a larger and multi-disciplinary corpus of literature? And does it provide new insights into the adoption and adaptation of network science in archaeology?

The meso and global scales of analysis did indeed reveal that key methodological influences came from outside the archaeological discipline. A few network models and methodological textbooks by physicists and social network analysts were cited by most of the archaeological network scientists. In fact, the literature review revealed that only in a few cases formal network techniques were adopted from publications other than these highly influential ones. The citation network analysis revealed that these publications were not only influential to archaeologists but were in fact key authorities in a wide range of disciplines. Using citation weights, a number of other influential publications (mainly in physics) were identified that elaborated on these authorities but have been largely ignored by archaeologists. Moreover, the bibliographical coupling and cocitation networks revealed disciplinary and topical clusters of publications that include research areas tangential to the egos and not revealed through a close reading that might well have innovative and useful archaeological applications. In short, this large citation network re-contextualises the archaeological use of formal network methods within wider research communities and holds the potential to guide future discovery and adoption of published formal methods that have proven useful in different fields of research.

Compared to the close reading the citation network reveals a more explicitly fragmented picture of archaeological network scientists seemingly working in isolation (as far as the archaeological use of their network methods is concerned) despite clear methodological similarities, whilst frequently engaging with developments in other disciplines. This idea of isolation sheds an interesting new light on the processes of adoption and (published) use of network science by archaeologists. The early adopters knew of each other's network-related work but their methodology was peripheral to the key topics being discussed in their sub-disciplines (Irwin pers. comm; Terrell pers. comm.). Therefore, in addition to an obvious technological barrier, these sub-disciplinary silos and in some cases purposeful non-citation might have led to a limited exposure of network-related work (Cochrane, pers. comm.; Terrell, pers. comm.). The newer generation of archaeological network analysts on the other hand did manage to achieve a wider exposure, at least in part because now the technologies (software, big data, hardware) were readily available, but there was also a strong academic influence. Indeed, the more recent applications often use techniques published by physicists, and the adoption of new techniques follows the trends in physics (but also SNA) with a short delay.

This fragmentary picture and over-emphasis on popular techniques is changing rapidly, however. It must be said that the citation network only reveals the situation up to 2011 when the data collection for this case study was finished (only two publications from 2012 are included). Since then a good number of innovative and critical archaeological network studies have been published, as mentioned in the literature review, which might considerably shake up citation patterns as suggested by Carl Knappett (2014, 182) in his response to these results: "Were Brughmans to repeat his citation analysis in just another year or two, I believe it would look very different".

The citation network approach also has some clear limitations, however. Although the networks analysed are quite large, they still exclude a vast body of literature with potential for archaeological applications. When the aim of a citation network analysis is to discover new literature, the sampling strategy of the egos needs to be given more critical thought. Secondly, techniques used to identify the most influential publications often provide obvious results that were also clear from the literature review. By focusing on these key publications there is a danger of over-emphasizing the similarities between the archaeological applications. The literature review, on the other hand, showed that the most promising archaeological applications were those that used network techniques or models that were not used by others (e.g. the model by Jin et al. 2001 used by Bentley et al. 2005). Although citation network techniques can be used to focus on these meaningful differences, these are not as easy to derive given the high number of peripheral nodes in citation networks. It seems that citation network analysis is most useful for tracing the main ‘paths of influence’ whilst the close reading is invaluable for evaluating how and why certain publications were influential. This issue is most critical for the large cocitation and bibliographic coupling networks. Their clear potential for revealing the differences as well as the interface between disciplinary and thematic clusters requires closer scrutiny for which network community detection methods can be applied (Fortunato 2010; Rosvall and Bergstrom 2008).

> Question 4: do archaeologists publish and cite differently from scholars in other disciplines? If so, does this require the development of specific citation network analysis techniques or strategies?

This case study showed that the limited availability of citation data for the humanities and social sciences is the most significant limitation to using citation network analysis in archaeology (Alen Sula and Miller 2014; Linmans 2010). By tracing the adoption of techniques developed in a variety of disciplines the relative prominence of each discipline within the citation network could be compared. This revealed that the meso and global scales of analysis were dominated by publications in physics whilst publications from other disciplines (including archaeology) were grouped together in peripheral clusters. The absence of influential SNA publications is particularly striking, with the exception of textbooks (e.g. Wasserman and Faust 1994) and popular theories (e.g. Granovetter 1973) often cited by physicists. It is interesting to note that this pattern was more clearly present in the meso scale network with only complete bibliographies (first half of section 3.5.2) than in the meso scale network with all available data. This suggests that the incompleteness of the available citation datasets does not make them useless. Citation network analysis arguably provides the most consistent results for its most complete segments, yet it reveals incomplete clusters grouped per discipline or research topic that can still provide new insights by exploring them independently (as suggested by Batagelj (2003) for the citation weights method).

This case study focused on the inter-disciplinary nature of archaeological publishing and citing, but it revealed very little about these practices within the archaeological discipline itself. In fact, the disciplinary clusters identified in the cocitation and bibliographic coupling networks suggest that the bulk of the archaeological publications included on the meso and global scales of analysis were not as multi-disciplinary as is often claimed. Archaeology formed a distinct cluster and only a few archaeological publications (the more methodologically focused egos in particular) acted as a bridge to other disciplines. An alternative bibliographic resource, the cocitation network of the Scopus database (accessible through the SCImago Netviewer[13]), also does not attribute a particularly multi-disciplinary role to archaeology. It suggests that archaeology documents are most often cited together with history publications, which in some years are cited most often with the social sciences and in other years with biological sciences. The inter-disciplinary nature is not the only feature that is often claimed to make archaeological practices stand out. Of particular interest is the so-called ‘grey literature’, the reports of archaeological activities that are not easily searchable and rarely cited (although see Richards et al. 2011; May et al. 2012). Indeed, the nature of archaeological data, which vary in type and format, are collected in large quantities and require significant investments in both time and money to process (Xia 2011, 234), often put limits on the frequency and format of publication as well as influence the nature of citations to analysis- and data-based resources (Boast and Biehl 2011, 120). I believe that the nature of publishing and citing in archaeology is rather similar to some of the findings of Robinson and Posten (2005) for anthropology as a whole: books are frequently cited; government documents, museum publications and grey literature or site reports are rarely cited; working papers are not common; and older materials are commonly cited. Moreover, the importance of the context of the citation (its frequency, its location in the document, and its positive/negative nature) which has been argued to be particularly crucial in bibliometrics in the humanities as a whole (Alen Sula and Miller 2014) is clearly also key in archaeological bibliometrics. However, one should not over-emphasize these differences: archaeology is not special; each discipline has its idiosyncrasies. What is crucial, however, is an awareness of these differences and a modification of the strategies for data collection and citation analysis. Future work should therefore build on this case study, necessarily through an alternative strategy where the structural position of the egos is embedded within the entire archaeological citation network, including grey literature where possible. Cocitation and bibliographic clustering projections that tend to over-emphasize differences between disciplines should be complemented with techniques that focus on the cross-fertilization between disciplines. The discrepancies between the indegree and external citation counts of papers (Table 3) indicate differences between the role of papers within the archaeological citation network, a pattern that can be explored through this alternative strategy.

[13] The SCImago map generator interface was used to explore cocitation networks of subject categories (as recorded in the Scopus database) for publications published in the United States and the United Kingdom for the years between 2005-2011. http://www.scimagojr.com/mapgen.php (accessed 28.04.2013).

It has been argued that in the humanities the same bibliometric methods can be used as for the sciences if a broader range of publications is included in citation databases (Nederhof 2006). The current case study (and the complete local scale network in particular) seems to confirm this conclusion. It emphasised that a critical awareness of the nature and limitations of one's dataset and the development of a purpose-built strategy of analysis (e.g. multi-scalar) are crucial. This issue should be further explored in future work, by extracting disciplinary citation networks and confronting these with networks of book/journal titles, subject categories and co-authorship which is, sadly, outside of the scope of the current case study.

It can be concluded that citation network data can enhance an archaeological literature review in cases where high quality datasets are available and where dependence assumptions inherent in citation network data are of research interst, and that network science offers techniques that are useful for the exploration of large network datasets. However, I believe the approach taken in this case study should be considered one of a few interesting ways of exploring such a rich dataset, one that does not conclusively illustrate or prove the potential of network science for archaeology. Working through the network science research process was enlightening, but in addition I believe an evaluation and example of how archaeologists formulate dependence assumptions is needed. The second case study provides such an example. Some of the exploratory network techniques introduced in this first case study will be combined in the second case study with confirmatory network techniques aimed at representing and testing archaeological dependence assumption.

Box 5. Summary chapter 3

Research questions:

- Can the intuitive results drawn from a close reading be quantitatively expressed and validated using citation network analysis?
- Can an exploratory network analysis provide new insights into the citation behaviour within the limited bibliographic corpus of the close reading?
- Does a citation network analysis allow one to usefully contextualise a close reading within a larger and multi-disciplinary corpus of literature? And does it provide new insights into the adoption and adaptation of network science in archaeology?
- Do archaeologists publish and cite differently from scholars in other disciplines? If so, does this require the development of specific citation network analysis techniques or strategies?

Conclusions:

- Citation network techniques succeed in identifying results obtained through a literature review, but some were time-consuming and led to ‘obvious’ results.
- The most enlightening citation network techniques were those where dependence assumptions about the ‘flow of influence’ were formulated. These led to new insights about the archaeological use of network techniques, but also about their multi-disciplinary influences.
- Publication and citation behaviour of archaeologists is similar to that in anthropology and to a large extent to that in the humanities as a whole. The same bibliometric techniques as used in the sciences can be applied in the humanities. A more detailed evaluation of the idiosyncracies of archaeological citation behaviour requires an alternative data sampling strategy than the one adopted here.

Implications for this PhD project:

- Citation data are an example of network data archaeologists are commonly confronted with and where exploratory network analysis and visualisation techniques offer an alternative to other types of exploratory data analysis.
- Formulating dependence assumption about network data aids archaeologists when interpreting the data in light of the processes they are interested in.
- The completeness and composition of a network dataset determine the selection of network techniques and how they should be interpreted. The available citation datasets in the Humanities are still particularly incomplete.

CHAPTER 4

CASE STUDY 2:

UNDERSTANDING INTER-SETTLEMENT VISIBILITY IN IRON AGE AND ROMAN SOUTHERN SPAIN WITH EXPONENTIAL RANDOM GRAPH MODELS FOR VISIBILITY NETWORKS

4. Case Study 2: understanding inter-settlement visibility in Iron Age and Roman Southern Spain with exponential random graph models for visibility networks

4.1. Introduction[14]

The citation network analysis presented in the first case study offered a practical example of how exploratory network techniques could aid a study of processes of academic influence. However, it did not analyse such processes directly. It merely suggested certain network structures as being indicative of certain processes. In the second case study I will make the jump from exploratory network science techniques to confirmatory techniques, by suggesting a method that formalises our assumption of the kinds of network structures that are the outcomes of certain processes and allows one to test these processes. This second case study therefore wishes to advance the research aims of this PhD project by illustrating how archaeological research questions can lead to the formulation of dependence assumptions, and the selection of exploratory and confirmatory network science techniques as the best method for understanding the implications of these assumptions. In doing this it will introduce a new network method into the archaeological discipline: Exponential Random Graph Models (ERGMs).

Traditional approaches to the archaeology of Roman Southern Spain have neglected the study of inter-site connections (Keay 1998a; Keay and Earl 2006). The transition from the Iron Age II (ca. 5th c. BC to 3rd c. BC, here referred to as Iberian) to the Roman period (ca. 3rd c. BC to 5th c. AD) is a most striking example of this. Iron Age II and Roman settlements and towns are often investigated independently, which is necessary for a critical understanding of the excavated materials, but it also sidelines the study of ways in which past communities might have interacted and of long-term continuity or discontinuity of occupation. This case study illustrates how a long-term and large-scale multi-site analysis allows for traditional research themes concerning inter-site connections in Iberian and Roman archaeology to be confronted. It focuses in particular on long-term changes in visibility patterns between urban settlements, a factor considered important for understanding Iberian settlement locations but largely ignored in Roman studies. It further compares these visibility patterns with the location of

[14] Please see the electronic supplementary material for this case study, which includes the visibility network and the location of all sites.

towns along transport routes, and the Early Imperial urban status of settlements, which are considered key factors for explaining locations of Roman settlements. An exploratory network analysis will reveal similarities and differences in the patterns of visibility networks. Hypotheses of the emergence and long-term change of visibility networks will then be tested using ERGM. In this case study I argue that visibility might still have structured interactions between communities in Roman times and should not be dismissed out of hand. However, the way in which it affected human behaviour might have been different in Roman times as compared to the Iron Age. I argue that simulating archaeologists' hypotheses of the emergence of inter-site visibility is a promising way of understanding such differences.

4.2. Visibility networks

4.2.1. The study area

The study area can be more or less equated with the Guadalquivir basin, primarily lying within the modern province of Seville, but also the edges of the adjacent provinces of Córdoba, Huelva and Cádiz, corresponding to the central and western portions of the Roman province of *Baetica* (Fig. 23). The Guadalquivir river (known as the *Baetis* in antiquity) cuts through the heart of this area, leaving a landscape of rolling hills behind in Córdoba to meander its way to the Atlantic through an increasingly wide and flat plain (Mayoral Herrera 2004). In antiquity, however, the river mouth was not situated in its current location at the Atlantic coast. Instead it fed into a large inland sea known as the *lacus ligustinus* south of Seville (Roos et al. 1995), and today referred to as the Marismas. The climate in this part of Southern Spain is hot and dry and can be described as semi-arid. The broadest stretch of the valley on the left bank of the river between Seville and Córdoba, known as the vega and campiña, is a particularly fertile part of the area and was suitable for growing a diversity of crops including grain and olives. The Guadalquivir is flanked by two mountain ranges: the Sierra Morena in the north and the Sistema Sub-bético in the south. The study area and the Sierra Morena in particular are rich in metals such as iron, copper, silver, and lead (Domergue 1990). The direction of the river and the alignment of the mountains give the basin its orientation towards the Atlantic, indicating a greater ease of access to the Atlantic than to the Mediterranean (Cunliffe 1995). This basin with its wide valley flanked by mountains has a funnelling effect on visibility in the landscape: vantage points on the plateau flanks and foothills offer great views throughout much of the lower lying areas.

It has been argued that the study of Roman *Hispania Ulterior Baetica* is dominated on the one hand by a largely ancient historical framework and on the other by highly detailed studies of individual sites or assemblages (Keay 1998a, 22). The many historical and epigraphic sources related to key Roman towns like *Italica, Hispalis, Gades,* and *Corduba* have dominated scholarship of Roman *Baetica*. This research tradition has resulted in a picture where most Roman towns are considered similar, while the transition from the Iron Age II period seems to have been an unproblematic break. This is reflected in key publications on the subject: although the province is included in a number of syntheses of Roman Spain (e.g. Le Roux 1995; Richardson 1996), in-depth studies and archaeological syntheses of the province itself are rare (e.g. Keay 1998c). One regional study of particular relevance for this project is the work by Michel Ponsich, who performed large-scale surveys of rural settlements in the Guadalquivir valley, the results of which were published in a number of detailed reports (Ponsich 1974; 1979; 1987; 1991).

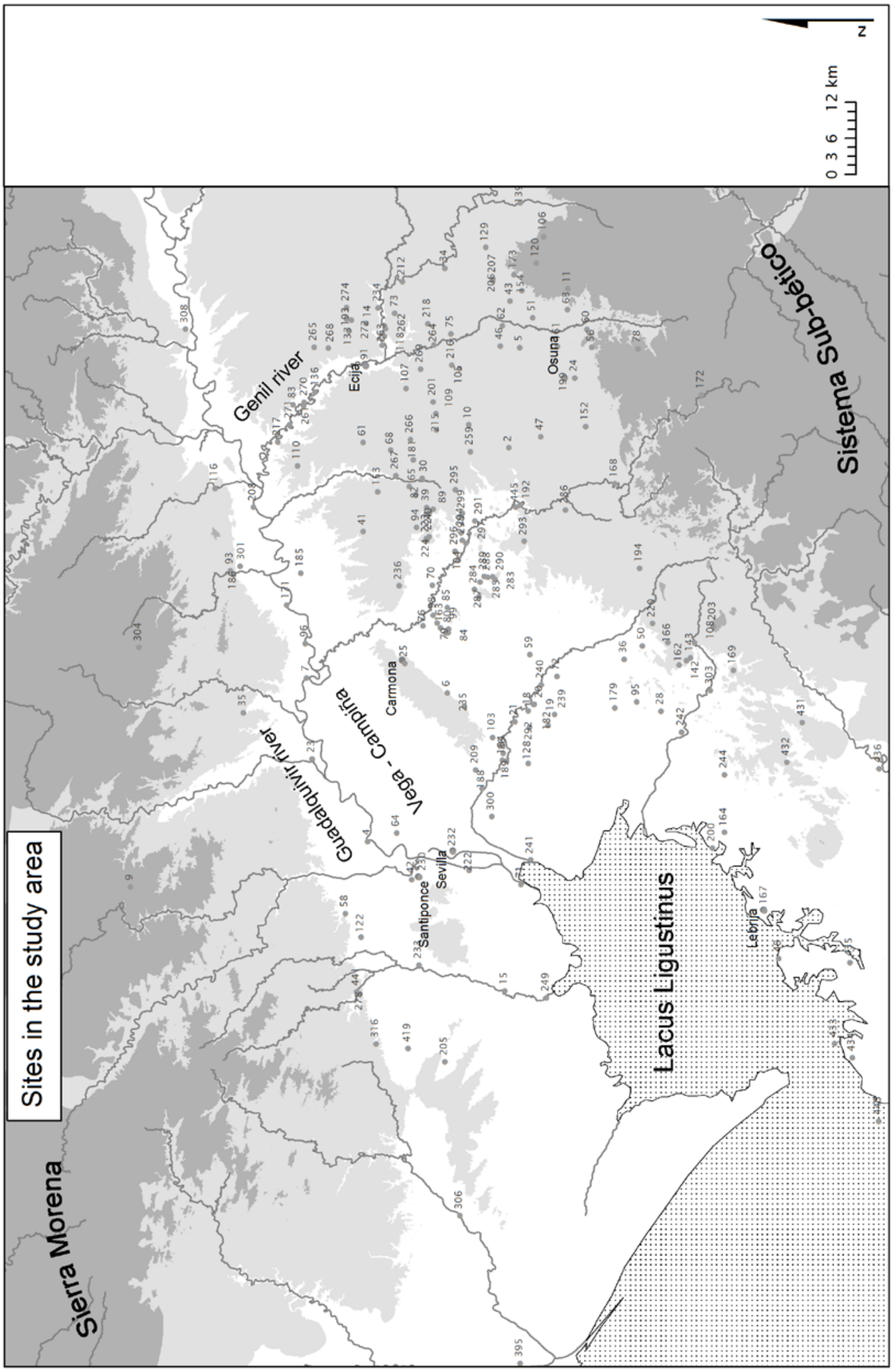


Fig. 23. Sites in the study area and geographical context. A full list of all site names mentioned on this map is included in Appendix III, section 11.1.

The Iberian period

The urbanization of Southern Spain has its roots in the Late Bronze Age, when the *Campiña* knew settlements of considerable size that grew into cities in the Iron Age (Ruiz Rodríguez 1997). During the Late Bronze Age the *Campiña* and the adjoining areas rich in metals were home to communities known to us as the Tartessos. Literary sources suggest the Tartessos were tied into wider Mediterranean exchange systems and traded with the Phoenician colonies that dotted the Atlantic and Mediterranean coasts of Southern Spain, although this picture should be nuanced as suggested by archaeological evidence (Chamorro 1987). The settlements of the Tartessos had in fact emerged before the rise of Phoenician colonies and continued to develop into the fifth and fourth centuries BC when the area formed part of a wider system of Iberian states (Cunliffe 1995).

Immediately before the establishment of Roman provinces in Southern Spain an Iberian people known as the *Turdetani* inhabited the Guadalquivir valley. Ancient authors write of the fertility and richness of *Turdetania* and of the civility and urbanization of its people (Pliny the Elder, *Naturalis Historia* 3.7; Strabo, 3.2.1-3.2.8). The archaeology of the region seems to support these authors' statements that this was by no means a barren land. Large and often fortified towns showing the architectural, social and economic features of urbanization were recorded all along the Guadalquivir river (Downs 1998; Keay 1998b). The settlement pattern in the *Campiña* was focused on these large settlements, located on elevated land near water sources. These settlements (sometimes called *oppida* although this is a very loaded term) were regularly spaced, housed substantial populations dependent on agriculture and formed the nuclei for surrounding settlements (Escacena and Belén 1998; García Fernández 2003; Ruiz Rodríguez 1997). The role of rural settlements is less well understood. However, the identification of the Turdetani as a distinct and homogeneous cultural entity (and even as an 'Iberian' people) inhabiting these lands between the end of the Bronze Age (6^{th} c. BC) and the beginning of the Roman period is problematic (Escacena and Belén 1998). Rather, the Iron Age and Early Roman period of this area was marked by a degree of continuity (Downs 1998), most notably in settlement occupation.

Roman Baetica

The study area became the Roman province of *Hispania Ulterior* in 197 BC (Richardson 1986), only a few years after it formed the stage for the final chapter of the second Punic war and the victory of the Romans. This made it one of the earliest provinces established by the

Roman Empire, and its past settlement pattern made it one of the most densely urbanized places in the empire. The urban landscape changed little during the Roman Republican period (Keay 1998b). Of note are the establishment of Roman settlements at *Italica* (modern Santiponce) in 206 BC and *Corduba* (modern Córdoba, 169 or 152 BC). *Italica* is a good example of the degree of continuity in Republican times. Although Roman veterans of the second Punic war were present at the site, it has been argued that *Turdetanian* cultural traditions nevertheless prevailed here (Keay 1997). *Corduba* shows a similar degree of continuity, where a pre-Roman fortified settlement existed alongside the newly founded Roman settlement, which acted as a river port (Ventura et al. 1998).

More significant changes came during the middle of the first century BC and the first century AD. Under Caesar and Augustus a number of *coloniae* and *municipia* were founded or, more frequently, existing settlements were awarded this status. The foundation of such urban settlements that formed part of the Roman Imperial administrative system was a key Roman strategy to impose and maintain control over provincial populations. An urban status can be considered an indication of political integration in the Roman Empire and is sometimes considered an expression of political urban hierarchies (Keay 1998b). The study of the diversity of interactions between Roman urban settlements and in particular their relationships with pre-Roman towns is therefore of great interest (Keay 1998a). In some towns (in particular the *coloniae* and *municipia*) during the Augustan and Julio-Claudian period public Roman buildings were constructed, which have been interpreted as the need of the provincial élite to publicly express their loyalty to the emperor (Keay 1998b; Trillmich and Zanker 1990). However, it should be stressed that many settlements remained essentially native in character, indicating that this was very much a period of gradual transformation (Keay 1998b). Indeed, it was not until the second half of the first century AD that rural Roman towns became a significant part of the landscape (Keay 2003, 190-191).

Under Augustus the province as a whole also became a more defined territorial entity as the new province of *Hispania Ulterior Baetica*, with the *colonia* of *Corduba* as its capital. After extensive administrative re-organization between 12 and 2 BC part of the province was absorbed by *Hispania Tarraconensis*. The remaining area of *Baetica* consisted of a number of districts: the *conventus hispalensis* with its capital in the *colonia Hispalis* (modern Seville), the *conventus cordubensis* with *Corduba* as a capital, the *conventus astigitanus* with its capital in the *colonia Astigi* (modern Écija), and the *conventus gaditanus* with the *municipium*

Gades (modern Cádiz) as its capital (Cortijo Cerezo 1993). The area of this case study roughly corresponds to the *conventus hispalensis*, and neighbouring sectors of the *conventus astigitanus* and *cordubensis*. Even more towns were given the status of *municipium* during the Flavian period, with the extension of Latin rights to a number of Baetican communities by Vespasian. Many of these Flavian *municipia* were previously native towns and knew a radical transformation of urban space and monumentalisation during the Flavian period (Keay 1998b). It has been argued that this transformation was made possible thanks to financial benefits enjoyed in this period and that it was politically motivated (Mayer 1977, 5-6). Finally, of particular note is the monumental enlargement of Italica in the second century AD during the reign of Emperor Hadrian (117-138 AD) who, like his predecessor Trajan, had close links with *Italica*.

Ancient authors (Pliny the Elder, *Naturalis Historia* 3.1, 3.3; Strabo 3.2.1-8) and archaeological evidence also stress the economic importance of *Baetica* to Rome. The best archaeological example of this is probably the large-scale production of Dressel 20 amphorae in kiln sites along the Guadalquivir and Genil rivers. These were filled with Baetican olive oil and transported in particularly large quantities to Rome as well as in more modest volumes all over the empire (Remesal Rodríguez 1998). A transport network existed that physically connected urban settlements and allowed for this produce to be efficiently transported out of the province. The most important components of this network were the *Via Augusta* (running through the *Campiña* from Córdoba to Seville through Écija and Carmona, and finally terminating in Cádiz (Corzo and Toscano 1992)), the Genil river and the Guadalquivir river itself, which in antiquity was navigable up to Córdoba (Strabo 3.2.3; Ponsich 1991, Fig. 6). One should also not forget the role of the many transhumance routes with origins in antiquity, connecting towns throughout the entire study area (Ponsich 1991, Fig. 10). Previous studies have explored the structural position of urban settlements on this transport network (Isaksen 2008). Sites that form part of this transport network can be considered to be integrated within established Roman economic networks (Ponsich 1991). In this case study the location of settlements with Iberian origins on this transport network will receive particular attention.

It has been argued that at the end of the second century AD a number of events (possibly including the invasion of *Baetica* by the Mauri who crossed the straits of Gibraltar from Mauretania) marked the start of a period of change in *Baetica* under the Severan dynasty (193-235 AD). The sites in this period seem to suggest a degeneration of the urban phenomenon, evidenced by the infrequent restoration of public buildings, and personal benefactions, and the exercising of public magistracies by citizens became more rare. However, the state's involvement in building projects as well as an increase in private buildings can be considered evidence for a degree of continuity. This has been interpreted as an adaptation to new needs and realities rather than decay (Rodà 1997).

At the end of the third century AD the study area saw some significant changes, in particular the replacement of the organizational system established under Augustus for a new administrative framework as part of Diocletian's (284-305 AD) reforms. Towns in the area continued to be occupied in the fourth century AD, and the first Christian religious buildings started to appear. The invasion of the Germanic peoples in 409 AD signalled the fragmentation of Roman provinces in the Iberian Peninsula. The Visigoths first entered the Peninsula in 415 AD and at the end of the century they would finally incorporate the area into their kingdom of *Tolosa*, which was previously limited to southwestern Gaul (Rodà 1997).

Bridging periods

The larger Roman sites have invariably received more attention from archaeologists than smaller ones, in particular those settlements that played a key role in the Roman administrative system as provincial capitals or by carrying the status of *colonia* or *municipium*. Although the province knew a fair number of these large sites like *Urso*, *Corduba*, *Italica* and *Carmo*, the vast majority of sites nevertheless had more modest proportions. Most crucially, a number of large Iberian settlements continued to play an important role in the Roman province in the Early Imperial period (Keay 1998b). The degree of continuity between the Late Iron Age and the Early Imperial period suggests that the Roman settlement patterns cannot be understood without exploring the preceding Iberian settlement patterns. This study will explore these long-term urban transformations through one particular aspect: their inter-visibility.

4.2.2. Why study visibility networks in Iron Age II and Roman Southern Spain?

Inter-visibility between settlements, both now and in the past, is high in Southern Spain (Keay 1998b; Keay and Earl 2011). This is a direct result of the landscape and the locations of settlements, positioned mainly along the large rivers or on the many hills surrounding the river valleys. In the Iron Age II period, the settlement pattern is generally understood to have consisted of large and often fortified nuclear settlements. These are sometimes referred to by archaeologists as *oppida* and are frequently surrounded by smaller settlements (Ruiz

Rodríguez 1997). While there is considerable debate as to the extent to which these settlements may have acted as political centres and, thus, a town in the Graeco-Roman sense, they are considered as such for the purposes of this research. Many Iron Age settlements of all kinds across Southern Spain as a whole are located on hilltops, terraces or at the edges of plateaus. At some of these there is evidence of defensive features and architecture, indicating that these locations were purposefully selected for their defendable nature and the ability to visually control the surrounding landscape, and maybe even for the ability to observe other settlements (Grau Mira 2004; Ruiz Rodríguez and Molinos 1993). It is possible that the patterns of inter-visibility between urban settlements were partly intentionally created and that these patterns had a role to play in structuring the interactions between Iron Age II communities: in other words through the visual control of urban settlements over surrounding rural settlements or the inter-visibility between urban settlements required to spread information in a signalling network as has been suggested for the south-east of Spain (Grau Mira 2003; 2005).

There seems to be no reason to believe that inter-visibility was a significant feature between towns in the Roman period (Keay and Earl 2011). Major Roman administrative centres were often located in low-lying areas along the main rivers and roads, to ensure economic and political integration within the wider Roman Empire (Keay 1998b). Due to the continuity in occupation of Iberian settlements, however, the percentage of them to be found on hills and plateau sides is only slightly lower in Roman times than in the Iron Age II period (Fig. 24, Table 7; Keay 1998b), which suggests that visually prominent site locations were not exclusively an Iron Age II phenomenon. This raises a number of questions:

- The trend in Figure 24 suggests that the slight decrease in settlements on hills and plateau sides reflects the degree to which Iron Age II sites continue to be occupied in the Roman period. Does the degree of inter-visibility between sites follow the same gentle downward trend through time as a result of this?
- To what extent were visually prominent settlements with Iron Age II origins integrated as towns into the political and economic structure of the Roman Empire? Was this truer than with less visually prominent settlements with Iron Age II origins?
- Is there evidence suggesting that inter-visibility of urban settlements was considered less important under the Roman empire? To what degree does the establishment

(whether intentionally or not) of inter-visibility patterns differ between the Iron Age II and Roman periods?

The past phenomenon being studied in this case study is therefore the ways in which lines of sight between Iron Age II and Roman settlements structured human behaviour, and the processes through which they emerged. In order to address these research questions, a method is needed to analyze both the changing patterns of observed inter-visibility as well as address hypotheses about how these patterns emerged. Very few contemporary accounts exist that could be informative. In *De Bello Hispaniensi* we read that “Most of the towns in this province are pretty well protected by the mountains, and are situated on natural eminences, so that one has to climb up to reach them, and the approach is thereby made difficult” (10.3) and “Pompeius pitched his camp on the hills, within sight of both towns, but did not venture to come to the help of his own side” (6.2). A quantitative visibility analysis might offer a way forward to shedding light on the significance of these comments.

Archaeologists have often used GIS-based visibility analysis methods for this purpose (Conolly and Lake 2006, 225-233; Wheatley and Gillings 2002, 201-216). They explore the visibility (or not) of landscapes, sites, features and objects to evaluate its possible impact on human behaviour and the probability that visibility patterns were intentionally created. Existing formal approaches, however, rarely analyze hypothetical processes that might have given rise to the observed visibility patterns. A reason for this is no doubt the complex mix of socio-cultural and environmental factors that influenced the establishment of settlement locations: new settlements emerge within an existing cultural landscape; the factors considered important at the emergence of the settlement can rarely be optimised; and the role of settlements and the important factors that accompany them change through time. To claim that inter-visibility of urban settlements was the only factor that determined site location would clearly be risible, but so would the opposite statement that inter-visibility played no role at all. In order to test the degree to which the emergence of inter-visibility patterns differed through time and make interpretations about the intentional creation of these patterns, an analysis of the structure of observed patterns should be combined with hypothetical models of how these patterns emerged.

In this case study I will apply a novel approach based on exploratory network analysis combined with ERGMs. This method allows one to explore the structure of patterns of inter-visibility between sites, the commonalities and differences in the inter-visibility patterns of individual sites, as well as to evaluate hypotheses of how patterns of high or low inter-visibility could have emerged on their own or in relation to a few external attributes. Such a combined approach will allow to suggest which types of inter-visibility pattern were most or least prominent in each period, whether processes that simulate the emergence (or not) of these particularly common or less common patterns are likely to have led to the observed pattern of each period, and to evaluate the possible importance of inter-visibility for different periods.

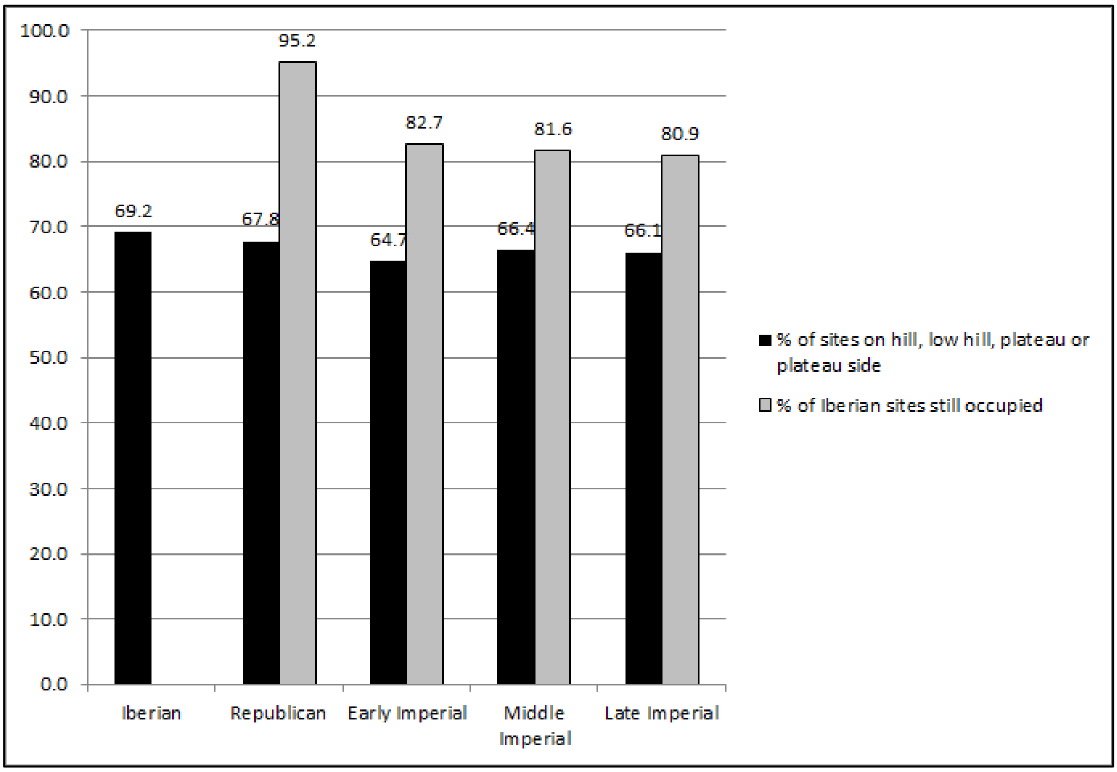


Fig. 24. Proportion of sites per period with a prominent location and with Iberian origins.

Table 7. Number of sites per period and location type.

	Total number of sites	Iberian sites still occupied	Sites on hill, low hill, plateau, or plateau side	Sites on hill	% of sites on hill, low hill, plateau or plateau side	% of sites on hill	Newly founded sites	Sites no longer occupied
Iberian	159	159	110	74	69.18	46.54		
Republican	146	139	99	69	67.81	47.26	7	20
Early Imperial	150	124	97	66	64.67	44.00	20	16
Middle Imperial	125	102	83	57	66.40	45.60	2	27
Late Imperial	115	93	76	51	66.09	44.35	1	11

4.2.3. Why processes of emerging inter-visibility?

Many formal studies of visibility have similar assumptions: they focus on analysing the structure of visibility patterns to understand their roles and evaluate whether these patterns were intentionally created. In such approaches, the types of inter-visibility patterns that are particularly suitable for a certain purpose such as visual control over a landscape or communication through visible signals, are considered most likely to have been purposefully established. These approaches imply a sequence of events based on the emergence of these particular types of patterns that resulted in the observed structure, a process of changing inter-visibility patterns. In other words, previous studies have assumed that the observed patterns could not have emerged through a random process. A good example of this is Tilley's (1994) study of a network of inter-visibility between barrows on Cranborne Chase, in which an observed network pattern is interpreted as the intentionally established outcome of an untested process: "One explanation for this pattern might be that sites that were particularly important in the prehistoric landscape and highly visible 'attracted' other barrows through time, and sites built later elsewhere were deliberately sited so as to be intervisible with one or more other barrows. In this manner the construction of barrows on Cranborne Chase gradually created a series of visual pathways and nodal points in the landscape" (Tilley 1994, 159).

Very few visibility studies have explored hypotheses about such processes explicitly (see Swanson 2003 for a notable exception). In this study, however, the decisions to establish certain patterns of visibility among urban settlements are the focus of attention. Most crucially, I will try to evaluate to what degree this changed through time, and how this was affected by factors like location along major transport routes or the Early Imperial urban status gained by settlements. The approach taken here is experimental. It will initially focus exclusively on the patterns of inter-visibility between settlements, exploring their observed structure as a static snapshot, and then addressing the following hypothetical question: if the visibility patterning that we have observed was the only reason for selecting the locations of sites, what then would be the process that is most likely to have led to the observed patterning? This question will be evaluated through an ERGM approach that models the creation of visibility patterns in abstract space (i.e. by simulating the creation of points and lines without taking the landscape's topography into account as a constraint). These processes will be simulated a second time, incorporating a number of non-visibility factors, such as the urban status of settlements or their location on Roman transport networks, to evaluate whether these factors are important for understanding visibility network creation processes. Finally, the results of this exploratory network analysis and ERGM approach will be re-contextualised within a wider archaeological discussion to shed light on aspects of the changing interactions between urban settlements in the study area through time, as reflected through visibility patterns.

Not including topography as a constraint in my ERGMs is a limitation of this study, because the approach presented here does not allow to evaluate how common or rare the observed patterns are in this particular landscape. This is only partly overcome in this case study by including the elevation of sites as an attribute in the ERGMs. Instead, the ERGM results show how likely it is that the observed patterns are the outcome of processes that are abstract expressions of archaeologists' hypotheses of visibility network creation (e.g. a tendency to visually control surrounding settlements, or a tendency for inter-visibility). However, within the context of the current case study I argue that this abstraction is justified and useful for the following reasons. Firstly, it allows one to express a very wide range of hypothetical processes of network creation in a (relatively) non-computationally-intensive way. It also gives one an idea of which hypothetical processes are less likely to lead to the observed networks, and I believe that it follows that if a process is unlikely to give rise to the observed network in abstract space with no geographical constraints, it will be even less likely to give rise to it in geographical space (although this should be proven in future research). The results will allow to focus more computationally-intensive modelling efforts with geographical constraints in the future on a more narrow range of processes. Moreover, a method for ERGMs with geographical constraints does not exist yet and will need to be developed in close collaboration with statisticians. Some progress is being made in this direction on which future work could be based, for example by including the distance between pairs of nodes (e.g. Daraganova et al. 2012), or by considering pairs of sites which are not inter-visible in the observed network as 'structural zeros' (pairs of nodes that cannot be connected). The approach advocated here is of course a highly simplistic abstraction of complex phenomena. But it has proven an insightful thinking process to evaluate the importance of particular patterns of visibility, and to discard possible but highly unlikely hypotheses.

4.2.4. Hypotheses of visibility network creation processes

The formal method I use requires for the past phenomena I am interested in to be abstracted and represented as network data. In this study the entities of research interest are individual settlements, where observer locations are used as an abstraction of the ability of members of the settlement communities to see other settlements. Relationships between settlements are used as an abstraction of lines of sight which could structure past human behaviour. Iberian settlements and Roman towns are represented as nodes, whilst arcs (directed edges) represent lines of sight from an observer on one site to an observed point at another site (Fig. 27). I can subsequently use these nodes and arcs to construct network data representations as the outcomes of the hypotheses I am interested in testing. In what follows I will formulate three groups of hypotheses tested in this case study, and the network data representations used to identify their expected outcomes (referred to as configurations from now on. The configurations used in the models presented here are shown in figures 25 and 26).

Firstly, the simplest hypothesis, which is the assumption that lines of sight appear and disappear independently of each other. This reflects a null-hypothesis where no specific configurations of lines of sight have a higher probability of emerging. Such a random process can be simulated with Bernoulli random graph models (Erdős and Rényi 1959). Although such an assumption is unrealistic for visibility networks, it is nevertheless commonly used as a baseline for comparison with the other two groups of models (Koskinen and Daraganova 2013, 56).

Secondly, one can formulate a number of hypotheses in which settlements are established where the visibility of other settlements was considered important, but where no other factors are taken into account:

1. If communication or signalling between settlements needed to occur then these need to be inter-visible. This is represented by reciprocal arcs.
2. In order to perform some sort of visual control over surrounding settlements, these need to be visible from a given settlement. This is represented by many incoming arcs.
3. If a settlement is purposefully visually prominent, it needs to be visible from surrounding settlements. This is represented by many outgoing arcs.
4. If visual isolation is considered important then settlements will be expected to be invisible from surrounding settlements. This is represented by isolated nodes.

Thirdly, a set of hypotheses can be formulated where we take into account the possible importance of other factors (here referred to as site attributes) in the establishment of visibility patterns. These hypotheses also imply certain assumptions about network patterns, this time not just about the configurations of nodes and arcs but also about nodes with certain attributes in interaction. Figure 26 represents different configurations of sites with the four attributes considered here:

1. Iron Age II settlements that continue to be occupied in Roman times could be inter-visible, visually prominent or visually active (i.e. having many outgoing lines of sight).
2. Roman towns with urban status could be inter-visible, visually prominent, or visually active.
3. Roman towns on Roman roads and navigable rivers could be inter-visible, visually prominent, or visually active.
4. Sites on hilltops could be inter-visible, visually prominent, or visually active.

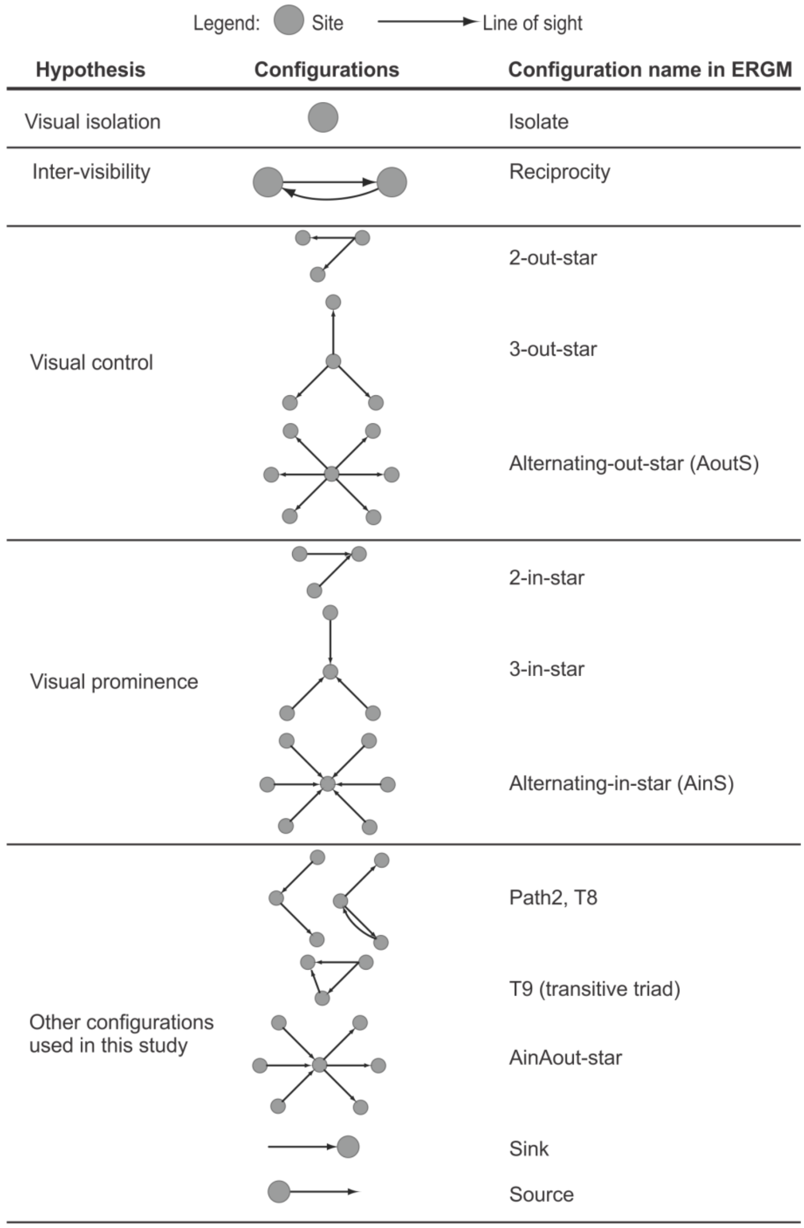


Fig. 25. Configurations (network building blocks) used in the models in this study.

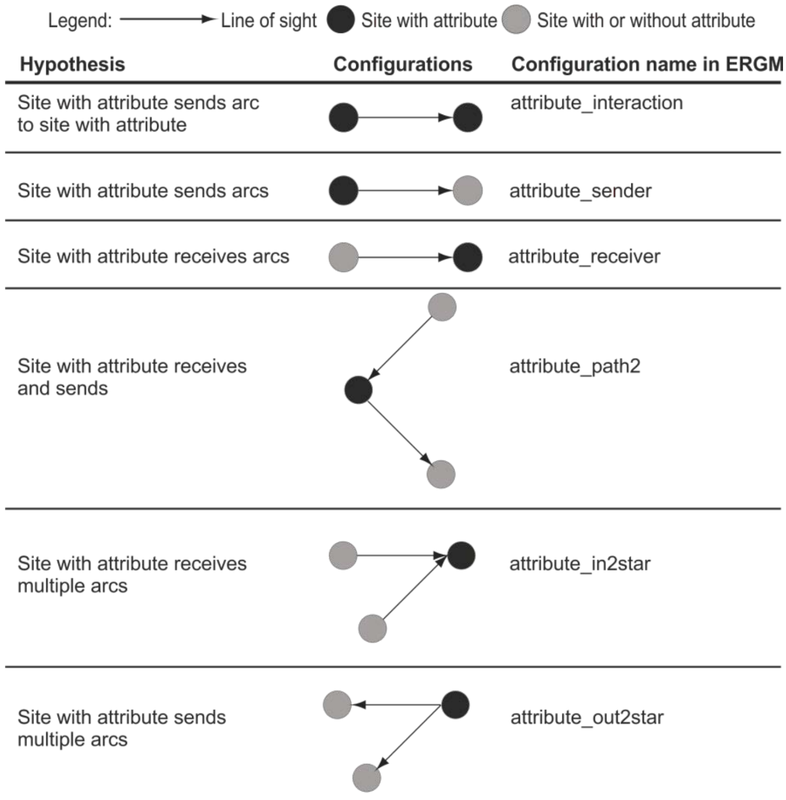


Fig. 26. Configurations with attribute effects for binary attributes. In black nodes the attribute is present, and in grey nodes the attribute is either present or not present.

4.2.5. Previous archaeological applications of visibility networks

Although visibility is generally considered an important feature in the study area, little work has been done to formally analyze visibility in Southern Spain (e.g. Garrido González 2011; Mayoral Herrera 2004). Visibility network analyses in archaeology in general are even rarer (early examples include Davidson 1979; Fraser 1983; Tilley 1994) and most of these concern communication networks (e.g. De Montis and Caschili 2012; Shemming and Briggs 2013). Of particular interest for this paper is the work on inter-visibility of Iron Age hillforts in Catalunya by Ruestes Bitrià (2008), who derives an inter-visibility network based on a probable viewshed analysis. Only mutual visibility between hillforts is included and network lines are either present or absent (no reference is made to how the probability of the viewshed results could be included in the network). The inter-visibility network was subsequently analyzed visually. The work of Grau Mira (2003; 2004; 2005) on the continuity and change between Iberian and Roman settlement patterns of Eastern Iberia is also of particular interest. He identified a strong degree of inter-visibility of *oppida* and argues for the emergence from the sixth century BC onwards of a hierarchical settlement system where *oppida* visually control surrounding rural settlements and access routes. Grau Mira also identified a more dispersed settlement pattern in the late Iron Age as compared to the Classical Iron Age, resulting in a decrease of the number of sites being located in areas that are highly visible from other sites.

The performance of a hypothetical communication network is examined more explicitly by Swanson (2003) in his study of hilltop features around the site of Paquimé (Mexico) that have been interpreted as forming a fire-signalling network. An undirected network of inter-visibility is created by considering only mutually visible sites derived from binary viewsheds. The resulting network is subsequently compared with randomly generated networks by calculating inter-visibility from randomly selected hilltops. The results show a clear difference between the observed network and the random network. The former has no isolated nodes, which is to be expected in an efficient communication network, while the latter networks regularly include isolates. By using such a method Swanson therefore aims to test the likelihood that the observed degree of inter-visibility between features could have arisen by chance.

The current case study aims to advocate an alternative approach, one that follows the different steps of the network science research process introduced in chapter 1 explicitly and allows researchers to test the dependence assumptions underlying network data. The case study will illustrate how directed networks can be derived from probable viewsheds, which allows for a new range of node-based network techniques to be applied. Swanson's (2003) use of randomly generated networks as a benchmark will be adopted here and pushed further by applying ERGMs specifically for the study of small-scale patterns.

4.2.6. What does inter-visibility mean?

An incredibly useful formal approach to analysing inter-visibility between settelements would be merely to consider it a direct result of site locations and the elevations of their surrounding landscapes. This straightforward combination of location and topography is often implied by simple binary viewsheds. But when we aim to understand what it means to be visually prominent this sum can be easily criticised through a range of arguments (for a more in-depth discussion see Wheatley and Gillings 2000; Llobera 2007):

- Spatial experience in a landscape is a complex mix of multi-sensorial information gathering and movement (Tilley 1994).

- A visibility study implies dominance of vision over the other senses in acquiring knowledge about one’s surroundings. This is sometimes considered an imposition of Western biases on the past (Thomas 1993; 2004).
- Geomorphological processes and human actions might cause significant differences between past topographies and the topographical models used. These models themselves are inaccurate abstractions of the modern topography generated through interpolation algorithms selected by the researcher (Fisher 1994; Wheatley and Gillings 2000; Wood 1996).
- Atmospheric conditions and past vegetation affect visibility. Although these often follow cycles they remain very hard to model.
- The ability to identify visible features is dependent on the closeness of the observer to the feature, the object-background clarity, and the observer.

Although these are all valid arguments, by no means should they cause archaeologists to abandon their pursuit of understanding past visibility patterns through formal methods altogether. Wheatley and Gillings (2000) argue that these critiques need to be confronted and embraced into formal archaeological visibility studies. The most common answer to these critiques has been to include some form of variability into formal visibility analyses, for example through Higuchi viewsheds (Wheatley and Gillings 2000), fuzzy viewsheds (Fisher 1992), probable viewsheds (Fisher 1994; used here and discussed in more detail in section 4.4.1 below), the selection of multiple viewing points on a single site (e.g. Mitcham 2002), or cumulative viewsheds (Wheatley 1995). Archaeologists have combined many of these different methods to achieve more accurate viewsheds (e.g. Llobera 2007). A particularly relevant example of such a combined approach is the study of visual control through inter-visibility of Iron Age hillforts in Catalunya discussed above (Ruestes Bitrià 2008).

Llobera (2007) rightly states, however, that introducing any type of variability imposes limitations on our interpretations and that it becomes crucial to define clearly what we can actually say with our particular combination of data and method. It is therefore necessary first to define how visibility will be approached and interpreted in this study, after which a method will be discussed that fits within this approach. Wheatley and Gillings (2000, 3) defined the term visibility as “*past* cognitive/perceptual acts that served to not only inform, structure and organise the location and form of cultural features, but also to choreograph practice within and around them.” Llobera (2003; 2007) similarly emphasises the role of visibility patterns in structuring space through the intentional positioning of physical features in the landscape. It is up to the archaeologist to decipher how this structuring was achieved in order to identify exactly which patterns were intentionally created, and most importantly to try to understand the role these visual patterns played in the past.

In this case study I assume that the presence as well as the absence of a line of sight from one settlement to another reflects the possibility that (i) this was intentional, (ii) it structured the surrounding space, and (iii) that the way in which it structured space might reveal aspects of the roles ascribed to the line of sight in the past.

4.3. Data: topography, sites, chronology, and attributes

4.3.1. Digital Elevation Model

The digital elevation model (DEM) was created from two sources: points and contour lines (source: ICA, Junta de Andalucía; contour interval 10m). The point heights are unevenly distributed over the landscape, with a higher density of observation in the mountainous Sierra Morena and the foothills of the Sistema Sub-Bético, as well as the more densely urbanised area around modern Seville, and a lower density in the low-lying Guadalquivir valley. The DEM was interpolated with the ‘Topo to Raster’ tool in ArcGIS 9.3, which was selected because it recreates a more correct representation of ridges from input point and contour data, features that have a significant impact on the results of visibility analyses (Wheatley and Gillings 2000, 10). This interpolation technique is specifically designed for creating hydrologically correct DEMs, drawing on both the points and contour datasets (the point heights were selected as the dominant source in the tool). A resolution of 35m was chosen for the resulting DEM and it has a root mean square error (RMSE) of 3.37m (reflecting the degree of inexactness of the interpolation method used). The RMSE was calculated in Microsoft Excel by comparing the observed point heights with the heights predicted by the DEM at the same points.

4.3.2. Settlement data

A dataset of 190 sites assembled by the ‘Urban Connectivity in Iron Age and Roman Southern Spain’ project[15] is considered in this study (see Appendix III section 11.1). The

[15] Directed by Prof. Simon Keay and Dr. Graeme Earl was funded by the UK *Arts and Humanities Research Council* (AHRC) between 2002 and 2005 with subsequent support by the University of Southampton and institutions in Seville, notably the Departamento de Prehistoria I Arqueología de la Universidad de Sevilla and the Delegación de Cultura de la Junta de Andalucía.

amount and reliability of information available for each site differs greatly due to a variable research history. All 190 sites are settlements of some kind, including major Iberian settlement sites and rural settlements, Roman urban settlements and semi-urban/semi-rural agglomerations: it excludes villas and other kinds of rural settlement. Since the reliability of settlement type identification varies I do not distinguish between types of settlement in the formal analysis, although these will be taken into account when interpreting the results. I believe this dataset to be representative of the Roman period towns and agglomerations as well as the larger Iberian settlements. Our knowledge of Iberian rural settlement is less complete; this will be taken into account when interpreting the results. For the majority of sites little is known other than their location and broad periods of occupation. For a further 253 sites no reliable dating information was available, and for most of these identification as past settlements is highly uncertain; these possible sites were not included in this analysis. The dataset reflects our total available knowledge of Iron Age II and Roman settlements in the study area up to the end of the project in 2005. The project believes this represents a settlement pattern that allows one to draw meaningful conclusions with the approach presented here. Although subsequent discoveries can certainly be expected, in particular of Iberian rural settlements, I believe they will not substantially change the conclusions drawn concerning the hypothesis of the inter-visibility of Iberian urban settlements but might change the conclusions concerning the visual prominence of Iberian urban centres as compared to Iberian rural settlements. For this reason, the interpretation of the results will focus in particular on the most densely surveyed areas at the centre of the study area (Vega and Campiña). Given the sparcity of sites along the boundaries of the study area (i.e. the included sectors of the *conventus astigitanus* and *cordubensis*), few visibility patterns of interest could be indentified in the results and these will be interpreted with caution.

The dataset was assembled primarily from the following sources:

- Archival: sites listed in regional sites and monuments catalogues held by the Delegacion Provincial de Cultura de la Junta de Andalucía (specifically the ARQUEOS and its later replacement the SIPHA). These draw upon information recorded by archaeologists in the 20^{th}c., as well as works published earlier. The work of Ponsich (1974; 1979; 1987; 1991) forms a fundamental part of this, particularly for the Roman settlements.

- Administrative: these are specifically surveys of Iberian and Roman sites undertaken for administrative reasons by the Delegación Provincial de Cultura as part of the Junta de Andalucía's ongoing strategy of updating its regional sites and monuments registers.
- Research: these are sites that were found, investigated or excavated in the course of research projects undertaken by archaeologists based at the Universidad de Sevilla or elsewhere.
- Accidental: sites found or investigated as a result of rescue work ('urgencias') undertaken by archaeologists on behalf of the Junta de Andalucía, whether by excavation or survey.
- Project: sites visited in the course of the 'Urban Connectivity' project for the purposes of checking location and analysis of surface materials for indications of chronology.

4.3.3. Chronology

The sites included in this analysis were all occupied for differing lengths of time within the period of the ten centuries this study is concerned with: from the early fifth century BC until the late fifth century AD. Dates of individual sites were derived from excavation and survey reports as well as from their archaeological record. The accuracy of these sites' chronologies therefore varies enormously. In order to explore long-term change it was decided to use five time-slices according to which sites were grouped together for analysis (for the sites dated to each time-slice see Appendix III sections 11.1.1 and 11.1.2):

- Iberian (Iron Age II): early 5th c. BC to late 3rd c. BC
- Roman Republican: late 3rd c. BC to late 1st c. BC
- Early Imperial: late 1st c. BC to early 3rd c. AD
- Middle Imperial: early 3rd c. AD to early 4th c. AD
- Late Imperial: early 4th c. AD to late 5th c. AD

The transition periods used for these time-slices refer to periods of change that are defined by the chronology of different classes of ceramics that are commonly found in the area; the late 3rd c. BC is marked by the first Roman settlements (viz. *Italica*) and the appearance of imported Italic Black Gloss pottery and subsequently imported Dressel 1 and other varieties of Italic wine amphora; the late 1st c. BC marks the disappearance of the earlier kinds of pottery and the appearance of Terra Sigillata Italica and a range of other kinds of well defined locally produced and imported pottery, such as Terra Sigillata Clara A; the early 3rd c. AD coincides with the appearance of a distinctive class of imported north African pottery, Terra Sigillata Clara C; the early 4th c. AD onwards is represented by the appearance of yet another distinctive class of imported north African pottery, Terra Sigillata Clara D and other imported material. Sadly the coarse chronologies of the vast majority of sites (often defined with an accuracy of a century, especially for those sites known exclusively through survey and surface scatters) do not allow for a much more precise overall periodization. However, it was felt that this archaeologically defined periodization conformed much more closely to the archaeological veracity of these sites, and that was therefore more appropriate than imposing historically defined periods. Although this is a convenient approach for exploring long-term changes of patterning in a large dataset using formal methods, it is necessarily coarse and even enforces the traditional chronological boundaries this paper seeks to challenge. A more fuzzy or probabilistic approach to dealing with temporal uncertainty should be preferred (e.g. Crema et al. 2010) but lies outside the scope of this case study. Still, earlier in the project fuzzy dates were employed as a means to integrate data from various sources, and in particular to record perceived bias in use of terminology (Earl and Keay 2007). We believe the current approach will still allow for the observation of large-scale long-term change that is the aim of this case study, although with a low degree of chronological accuracy.

4.3.4. Site attributes

In this case study I will evaluate the role of three attributes in giving rise to the observed inter-urban connections: urban status, prominent elevated locations, and location on road or river networks. In addition to these three I will consider the potential Iberian origins of sites as an attribute to evaluate interactions between Iberian sites that continue in occupation. Lists of sites for each of these attributes are included in Appendix III section 11.1.

Urban status: as mentioned above, many sites in the study area were attributed an urban status by Rome in the Early Imperial period. Thus, epigraphic and historical sources indicate that *Hispalis* (Seville), *Astigi* (Écija) and *Urso* (Osuna) were established as *coloniae* under Caesar and Augustus, while many others gained the status of *municipium* in the course of the 1st c. AD, particularly after the later 1st c. AD (Keay 1998b, 85). The acquisition of the legal privileges implicit in these statuses is a good measure of the political integration of these communities into the Roman Empire and, therefore, can be read as a regional expression of politically-based urban hierarchies. However, there are many difficulties surrounding the identification of an archaeological site with a town mentioned in ancient written sources, an issue that is particularly acute for the many possible *municipia* in the region. Here I decided to follow the list of *coloniae* and *municipia* compiled by Simon Keay (1998b, Appendix II) as a prime reference. This list was itself compiled from a range of secondary sources (Fear 1996; Knapp 1983; Sillières 1991; Stylow pers. comm.; Tovar 1974).

Prominent elevated locations: a qualitative assessment of prominent locations is included as a site attribute in this analysis to evaluate the effect of the DEM (which provides a quantitative assessment of prominent locations). Project members visited all of the sites in the study area in the framework of the 'Urban Connectivity in Iron Age and Roman Southern Spain' Project. They described the physical location of sites using a range of qualitative categories. Sites that were described with the following attributes were considered to have a prominent elevated location: hill, low hill, plateau, plateau side.

Road and river networks: this attribute represents sites located on the main transport routes in the study area: the *Via Augusta*, the Guadalquivir river, and the Genil river. Transhumance routes were not taken into account because their exact course and dating is often unknown. The course of the *Via Augusta* in this area is well known and has been thoroughly studied (Sillières 1991; Corzo and Toscano 1992). In antiquity the Guadalquivir and Genil rivers were navigable up to Córdoba (*Corduba*) and Écija (*Astigi*) respectively (Strabo 3.2.3; Ponsich 1991, Fig. 6). All sites along these rivers up to these points were considered to be part of the river network. It is notable that almost all towns along these transport routes were either *coloniae* or *municipia* and for most of them we know the ancient name (see Appendix III section 11.1.4 and 11.1.5 for full lists).

4.4. Method: an approach for exploring changing visibility networks

To address this study's research questions I developed an approach that allows one to explore the structure of visibility networks as well as evaluate assumptions about the factors driving the emergence of these changing structures. This combined method consists of a few steps that will be introduced here: generating probable viewsheds, creating visibility networks, exploring network structure, and creating statistical models of emerging networks (with ERGMs).

4.4.1. Probable viewsheds

As mentioned above, in formal visibility studies a measure of variability is often added to address methodological and theoretical shortcomings of the approach. In this study I will add variability in two ways: through probable viewsheds and by analysing the results in different bands of distance. A probable viewshed uses a Monte Carlo simulation approach to identify the probability that a location is visible to an observer from a specific viewing point. The probable viewshed was first developed by Fisher (1992; 1995) to compensate for the imperfections in the DEM and the variability between algorithms used to create viewsheds. As such the error introduced in a probable viewshed reflects the RMSE of the DEM. In this study I use a 5m error, which is slightly higher than the RMSE of 3.37m and enhances the effect of decreasing probability of visibility with distance away from the observer. Since the differences between the observed and simulated point heights used in the calculation of the RMSE are normally distributed, I decided to implement the error in the probable viewshed creation as follows: in each iteration, the elevation of each cell in the DEM is increased by a float from a normally distributed probability distribution. The probable viewshed was iterated 100 times for each of the 190 sites. The curvature of the earth was taken into account when calculating lines of sight, given the large study area. The observer height was fixed at 1.7m since I assume visibility by a human observer. An elevated target height was considered uncritical since for many sites we have no knowledge of the architecture and a target height of 1.7m was discarded since inter-visibility between observers was not the aim of the analysis. It was decided to consistently use a target height of 0m (Fig. 27). Although these assumed observer and target heights are selected in light of the above theoretical considerations, they also enforce a higher degree of asymmetry in the resulting visibility networks than would be expected for equal values for target and observer heights (as illustrated below in section 4.4.3). I argue this is defendable for two reasons: (1) the probable viewshed approach taken here inevitably leads to higher asymmetry of visibility networks than would be expected for equal values for target and observer heights (since both are increased by different amounts when introducing a random error into the DEM at each iteration); (2) the hypotheses suggested by Iron Age archaeologists concern asymmetric visibility networks and need to be evaluated as such.

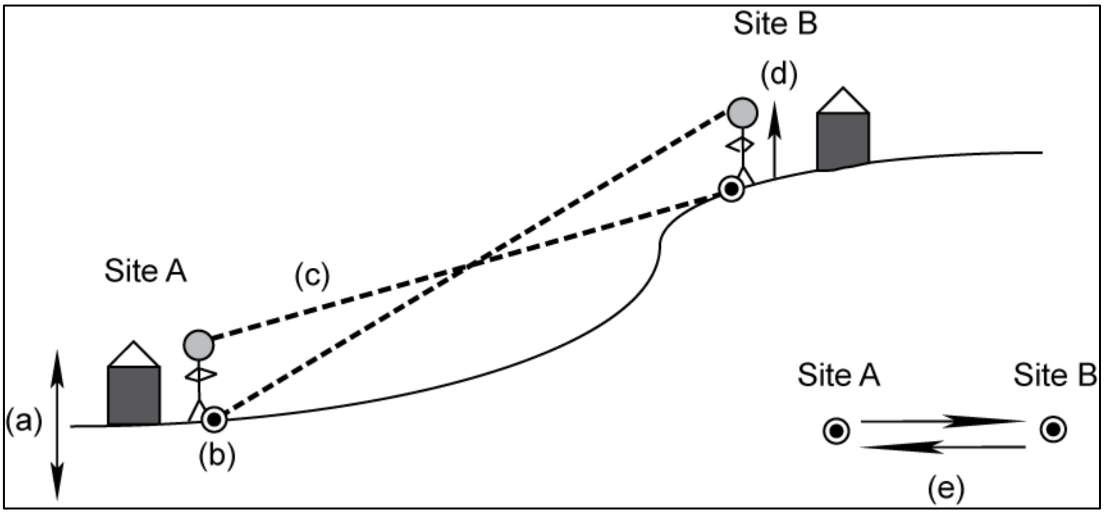


Fig. 27. (a) maximum random error in probable viewshed = 5m; (b) observer point site A and target point for observer site B; (c) line-of-sight with probability p_a from observer A to target point site B; (d) observer height = 1.7m; (e) inter-visibility network where site A is connected to site B with probability p_a and site B is connected to site A with probability p_b.

4.4.2. Observation point locations

A single observer point was created for each of the 190 sites in the analysis. In general the observer point was either positioned on the site location or at the centre of the site area (if the area is known, which is not the case for the vast majority of sites). Observer points were positioned on prominent features when known (for example the Roman gate at Carmona, see below). These point locations are more dense in the campiña and around the Genil river, areas that have traditionally always been more densely occupied (Ponsich 1991), and they are less dense around the Guadalquivir, the Sierra Morena, and the *lacus ligustinus* (Fig. 23). This distribution has a significant impact on the results of this study, as revealed through a cumulative viewshed analysis (results described in Appendix III section 11.2; Figs. 73-77), and it will be taken into account when interpreting the results. The cumulative viewsheds also reveal a gradual decrease over time of overall visibility in the study area, the increase over time of the visual prominence of one area (northwest of Osuna), the to-be-expected consistent visual prominence of sites located on elevations, and the consistently low degree of visibility of the major roads and rivers (with few exceptions).

However, past urban settlements were not just point locations but extended over sometimes considerable areas. One could argue that inter-visibility should therefore be derived from multiple observer points per settlement. However, for 23% of sites in this study the extent is unknown and for most of the others the minimum area is an estimate for which the exact geographical extent is unknown (Fig. 28; see also the rank-size analysis in Keay 1998b, Fig. 4). Deciding where the multiple observer points per site should be located therefore becomes impossible for most sites. I decided to remain consistent and only select a single observation point per site, even though this involves a significant limitation on the area visible from each site. At an earlier stage of the project some preliminary analyses were undertaken assigning ten viewer locations at random within the site polygons. However, the considerable uncertainty regarding the accuracy of these polygons, and in particular their sentivity to period of analysis, meant that this was not explored further. A detailed study of the past occupied area of each site was not possible for this study and should form the focus of future work, allowing for re-analysis of inter-visibility with multiple viewer points.

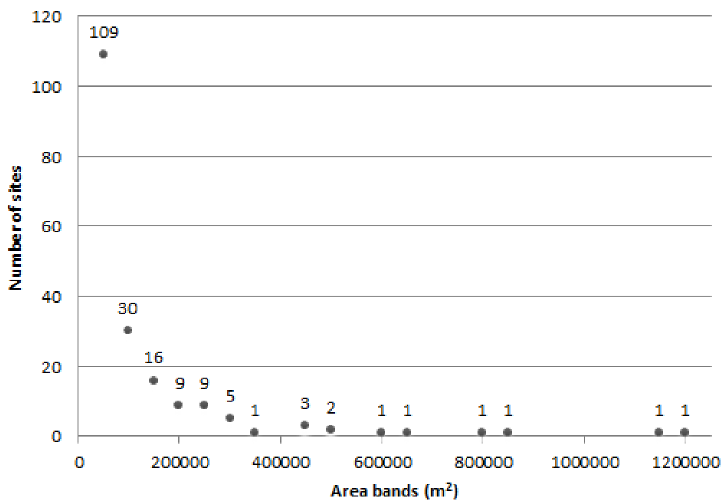


Fig. 28. Frequency distribution of the number of sites per area band of 50,000 m^2. All areas are minimum estimates. Of the 109 sites in the lowest band 43 sites have an area that is unknown or lower than 1,000 m^2.

4.4.3. Visibility network creation

From the probable viewsheds of all 190 sites a visibility network is created in which nodes represent sites (the observer point locations) and the arcs represent the presence of a line-of-sight from a site to another site. These arcs have a probability attributed to them derived from the number of times out of 100 that the cell in site A's probable viewshed on which the observer point of site B is located is visible. This results in a directed network in which lines-of-sight from and to sites can be distinguished, and where all sites have a differing probability of being visible from every other site (Fig. 27). The network can subsequently be explored per period by only including those sites occupied during a certain period. I will also explore the differences between network patterns with high probability arcs and those with low probabilities. Finally, I decided to focus the analysis on lines of sight at a distance of up to 20km, because at such distances large architectural features, and presumably communication signals, are still visible. However, in the discussion of the results I will also make reference to some patterns of interest over distances greater than 20km, since observations in the study area confirm that long-distance inter-visibility of sites is not impossible (Keay and Earl 2011). The inter-visibility of Carmona in the vega and Castillo de Mulva in the foothills of the Sierra Morena over a distance of ca. 30km is a clear example of this.

The coarseness of the DEM and the use of a single observer point per site might shed doubt on the use of directed networks. However, I believe their use is justified in this study for two reasons. Firstly, the probabilistic viewshed method gives rise to sometimes strong differences in the probability of lines of sight between pairs of sites: 31.6% of node-pairs have a difference in the probability of the arcs between them of more than 10%; whilst 4.7% of node-pairs have a difference in the probability of the arcs between them of more than 50%. These differences have a strong impact on the number of reciprocal ties and the network patterns when the networks are explored using thresholds on different probability values. Secondly, the archaeological hypotheses I aim to test require an abstraction and representation as directed networks (e.g. visual control or visual prominence).

4.4.4. Exploratory Network Analysis

The resulting networks are static and are not suitable to explore issues of movement through the landscape. Instead, this study will focus on how patterns of inter-visibility between sites change over time, how these patterns emerged, and how these might have structured interactions between urban settlements. This can in part be done by exploring these static networks using exploratory network measures.

The local node-based measures used include indegree, outdegree, and clustering coefficient. The global network-based measures include number of nodes, number of arcs, clustering coefficient, number of connected components, average degree, diameter, average shortest path length, and density (for definitions see Newman 2010; Wasserman and Faust 1994). The indegree of site A is the number of sites from which site A can be seen. The outdegree of site A is the number of sites that can be seen from site A. The diameter of a network is the longest path between a pair of nodes; if a network is disconnected the diameter of the largest connected component is given.

4.4.5. Exponential Random Graph Models

Exponential Random Graph Modelling is a statistical simulation approach that compares the structure of simulated networks which represent a researcher’s hypothesis of how a certain observed network emerged, with the known structure of this observed network. For a detailed ERGM reference work, see Lusher et al. (2013).

Social network analysts often use an archaeological analogy to explain the concept of an ERGM (e.g. Lusher and Robins 2013, 18). Past material remains are like static snapshots of dynamic processes in the past. Archaeologists explore the structure of these material residues to understand past dynamic processes. Such snapshots made up of archaeological traces are like static fragmentary cross-sections of a social process taken at a given moment. If one were to observe multiple cross-sections in sequence, changes in the structure of these fragmentary snapshots would become clear. This is exactly what an ERGM aims to do: to explore hypothetical processes that could give rise to observed network structure through the dynamic emergence of small network fragments or subnetworks (called configurations, e.g. Figs. 25-26). These configurations can be considered the building blocks of networks; indeed, LEGO blocks offer a good analogy for explaining ERGMs. To give an example, a network's topology can be compared to a LEGO castle boxed set, where a list of particular building blocks can be used to re-assemble a castle. But a LEGO castle boxed set does not assemble itself through a random process. Instead, a step by step guide needs to be followed, detailing how each block should be placed on top of the other in what order. By doing this we make certain assumptions about building blocks and their relationship to each other. We assume that in order to achieve structural integrity in our LEGO castle, a certain configuration of blocks needs to appear, and in order to make it look like a castle other configurations will preferentially appear creating ramparts, turrets, etc. ERGMs are similar: they are models that represent our assumptions of how certain network configurations affect each other, of how the presence of some ties will bring about the creation or the demise of others. This is where the real strength of ERGMs lies: the formulation and testing of assumptions about what a connection between a pair of nodes means and how it affects the evolution of the network, explicitly addressing the dynamic nature of our archaeological assumptions. The hypotheses introduced in section 4.2.4 above represent different assumptions archaeologists make about certain network patterns (for example, in order for a communication link to exist and function between two settlements they need to be inter-visible). These theoretical assumptions can be represented using network configurations (Figs. 25-26).

I believe this method is particularly promising as an approach for evaluating the central claim of this PhD introduced in section 1.3, that network science reveals its true potential for archaeology when we address dependence assumptions inherent in network data explicitly. In what follows I will therefore describe this method in full technical detail, using examples to illustrate this complicated approach.

Definition

ERGMs[16] belong to a family of statistical models originally developed for social networks (Anderson et al. 1999; Wasserman and Pattison 1996) that aim to investigate the *dependence assumptions* underpinning hypotheses of network formation by comparing the frequency of particular *configurations* in *observed networks* with their frequency in *stochastic models*. The terms in this definition in italics have a specific meaning in network science and will be defined below.

As mentioned in my definition of network science, *dependence assumptions* are theoretical assumptions grounded in the idea that pairs of nodes do not just become connected independent of what happens in the rest of the network: “The presence of some ties will encourage other ties to come into existence, to be maintained, or to be destroyed” (Robins 2011, 485). These assumptions, therefore, reflect the researcher’s theories of how ties emerge relative to their position in the network. An example of a dependence assumption and its representation using configurations is shown in figure 29.

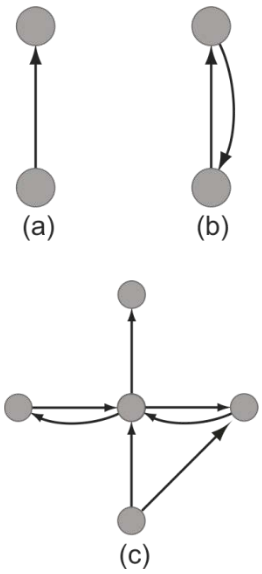


Fig. 29. (a-b) Example of a visibility network where nodes are settlements and arcs lines of sight between them. If an observer can see one settlement from another (a), it is likely that both settlements are inter-visible (b). (c) This network consists of 5 nodes, 7 arcs, and 2 reciprocity configurations.

I use the term *observed network* here to refer to the network created on the basis of data collected by archaeologists. The researcher using ERGM is interested in modelling the observed network (Robins et al. 2007a, 175). In visibility networks this would typically be a set of nodes representing the observation locations (e.g. settlements) connected by a set of

[16] These are sometimes called $p*$ models to distinguish them from the earlier p_1 (Holland and Leinhardt 1981) and p_2 (Lazega and Van Duijn 1997) model classes.

arcs representing lines of sight (Fig. 27). Visibility network data can either be collected by observations in the field, or more formally by using visibility analysis techniques in a GIS (Conolly and Lake 2005; Wheatley and Gillings 2002). An ERGM aims to study the range of processes that could give rise to such networks, and it is therefore crucial for the observed visibility network to be as complete as possible if the ERGM is to suggest realistic processes. Moreover, the selection of the boundaries of the visibility network will need to be clearly argued for, and the impact these boundaries have on the results of the ERGM will need to be assessed when interpreting the results. In archaeology the observed visibility networks are often a pattern that aggregates evidence over a long timespan. This is less of a problem than the issue of having a complete network, in particular for relatively slow-changing processes such as settlement patterns, as long as the researcher is confident about the contemporaneity of the set of nodes that make up the observed network.

Configurations[17] are small network patterns consisting of a few nodes and the arcs between them (e.g. Figs. 25-26). They play a number of roles in the ERGM procedure: representing dependence assumptions, describing observed network structure, comparison with simulated networks, and as *effects* in the ERGM. Dependence assumptions can be formally represented by particular configurations. For example, inter-visibility could be represented by arcs in two directions (referred to as reciprocity; Fig. 29b). One can also describe an observed visibility network by counting the frequency of each configuration in the network. This provides a way of describing a visibility network's structure, but also allows comparison with the number of configurations of simulated networks. For example, the network shown in figure 29c consists of five nodes, seven arcs, and two reciprocity configurations. This information is used to determine how similar the networks simulated by a certain ERGM are to the observed network. When creating an ERGM researchers select those configurations to be included as *effects* in the model which they believe to be representations of their dependence assumptions. This means that the model will not let these particular configurations that are of research interest emerge purely by chance, but rather it will estimate whether there is a positive or negative tendency for each configuration to appear throughout the simulation process. For example, in the results below I describe an ERGM which includes the

[17] The term *configurations* is used here following key publications in ERGM (e.g. Robins 2011) and first used by Moreno and Jennings (1938), instead of the term *motifs* which recently became popular through the work of Milo et al. (2002).

reciprocity configuration as an effect, and the results indicate that in this model there is a tendency for settlements to be inter-visible.

These configurations are assembled through a *stochastic process*: at each time step two randomly selected nodes are considered and an arc may be created or removed between them. The probability that an arc is created between these two nodes is determined by the effects in the model, and therefore by the presence or absence of other ties. To give the example of this case study, when in an ERGM with a strong inter-visibility effect a pair of nodes A and B is considered that already have one arc from A to B, then the probability will be high that another arc from B to A is created in that time step. This stochastic simulation process is an implementation of the idea that the observed network is only one particular outcome out of a wide range of possible networks. We do not know what process generated the observed network, this is what we are trying to find out. But we do know the dependence assumptions we can formulate based on our theories. The goal of an ERGM is to draw on these assumptions to propose a plausible theoretical hypothesis for the process that led to the observed network (Robins et al. 2007a, 175).

Creation process of ERGMs for visibility networks

Now that I have defined ERGM and its most important concepts, the next thing to do is explain step-by-step how an ERGM is created. In their general framework for ERGM construction Robins et al. (2007a) describe five steps followed when creating an ERGM. By going through these steps archaeologists can test their theoretical decisions about how lines of sight are created through statistical data analysis. I will discuss these five steps and give examples relevant to the analysis of visibility networks. Figure 30 offers a simplified overview of this design process.

1. **Each arc can either be present or absent:** we start with a fixed set of nodes that are unconnected, and assume that throughout the simulation every pair of nodes can either be connected or not. Although the arc is considered a random variable, some arcs will have a higher probability of appearing than others. For the visibility networks in this case study this means that the number of settlements remains the same throughout the simulation and that settlement A can either be seen from settlement B or not with a certain probability.
2. **A dependence assumption is proposed:** this is the most crucial step and concerns the explicit formulation of dependence assumptions representing the proposed

processes generating the network, i.e. one decides how arcs affect each other's presence or absence. The nine theoretical dependence assumptions about how lines of sight are established that will be tested in this case study have been introduced in section 4.2.4 above.

3. **The dependence hypothesis implies a particular form to the model:** the dependence assumptions formulated above can be represented by particular configurations. This means the theoretical assumptions of how lines of sight between settlements emerge have to be represented as network data, i.e. nodes and arcs. The researcher should select those configurations to be included in the model that are considered the best representation of the dependence assumptions (Figs. 25-26). Multiple configurations can be included.
4. **Simplification of models:** the hypotheses archaeologists formulate can often be a complex mix of the assumptions introduced above. Although multiple configurations can be included in an ERGM to represent this complexity from the start, this should be avoided since the more configurations included in the model, the harder it becomes to understand which configuration causes a good fit between the model and the observed visibility network. For this reason it is recommended that one starts with a simple model with few effects, and gradually build up the complexity of the model by adding more effects. The simplest assumption mentioned above is that lines of sight emerge independently of each other. Such an assumption could be represented by an ERGM with only one effect, which is the probability that an arc will be created (such models are called Bernoulli random graph models; Erdös and Rényi 1959). After evaluating and interpreting how well this model fits with the observed networks, one could then increase the complexity of the model by adding more effects (e.g. a reciprocity effect). Moreover, one should also consider whether some effects can be equated or related in some way, in order to limit the number of effects included in the model.
5. **Estimate and interpret model parameters:** the previous four steps are arguably the most important ones, since they determine the formulation of theoretical assumptions, their representation as network data, and the creation of a proposed model. The goal of an ERGM is to find a set of parameter values (representations of how important particular configurations are for generating the observed patterns) that best represent a single observed network. The observed network can then be interpreted in light of

these configurations and the dependence assumptions underlying them. The researcher runs the model and estimates parameter values for all of the configurations in the model, i.e. whether certain configurations have a positive or negative tendency of appearing in the simulated networks. This is an iterative process where parameter values are gradually refined until one ends up with a model with parameter values that give rise to networks with frequencies of the included configurations very similar to the observed network. One then performs a goodness of fit test to evaluate whether this model also gives rise to similar counts of configurations or aspects of the network's structure that were not explicitly included in the model (such as the degree distribution). This is done to confirm whether the model succeeds in generating non-modelled features of the observed networks. The final parameter values of the included configurations are subsequently interpreted. It is important not to over-interpret the results -a good rule of thumb is to only distinguish between positive or negative tendencies to creating a certain configuration, and to pay particular attention to statistically significant effects. For example, if in an ERGM for a visibility network the reciprocity effect is significant and positive, one can formulate the following interpretation "in the process that led to the creation of this visibility network, there was a tendency for settlements to become inter-visible, more so than expected by chance".

In the following three sections some of these steps are described in more technical detail (see also Lusher et al. 2013; Robins et al. 2007a; Robins et al. 2007b; Robins et al. 2009).

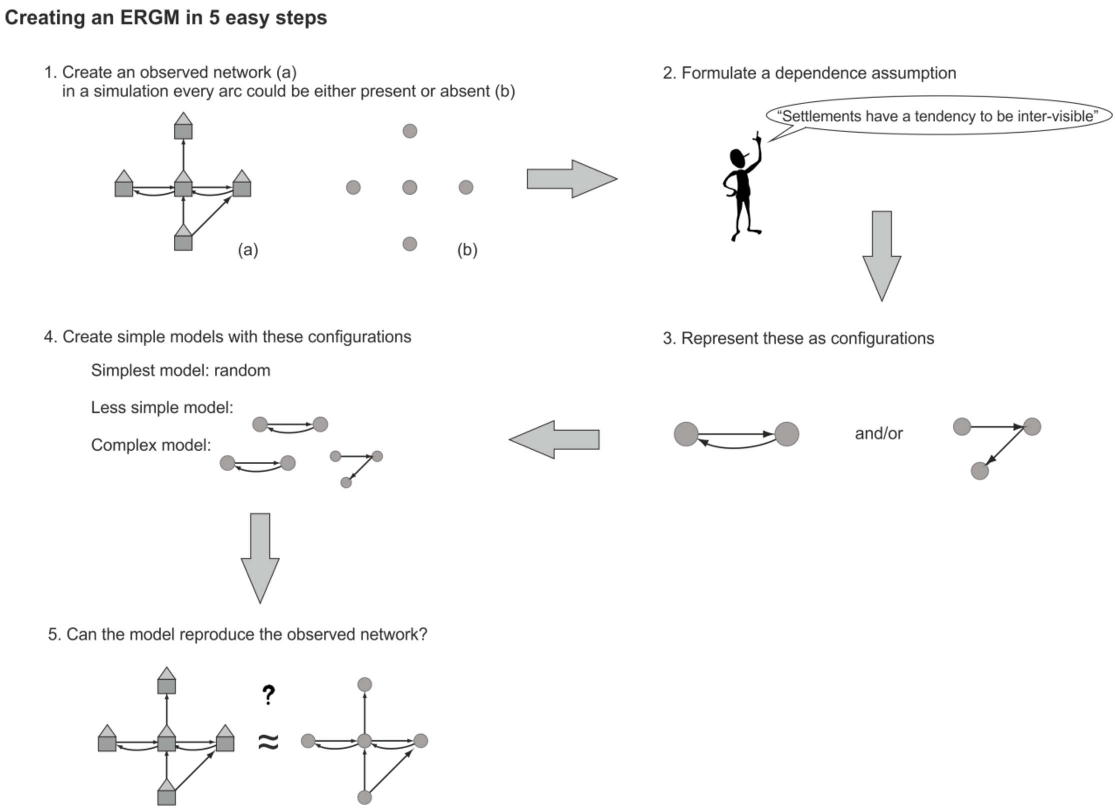


Fig. 30. A simplified representation of the creation process of an ERGM. (1a) an empirically observed network is considered; (1b) in a simulation we assume that every arc between every pair of nodes can be either present or absent; (2) dependence assumptions are formulated about how ties emerge relative to each other (e.g. the importance of inter-visibility for communication); (3) configurations or network building blocks are selected that best represent the dependence assumptions (e.g. reciprocity and 2-path); (4) different types of models are created (e.g. a model without dependence assumptions (Bernoulli random graph model) and one with the previously selected configurations) and the frequency of all configurations in the graphs simulated by these models is determined; (5) the number of configurations in the graphs simulated by the models are compared with those in the observed network and interpreted.

The form of an ERGM

The definition and creation process of an ERGM can be expressed more formally as equation 1. All ERGMs have the same general form. To this general form different dependence assumptions can be added depending on the hypotheses tested. Equation 1 describes a general probability distribution of networks, where the probability that a particular network will exist in this distribution (Pr) is dependent on the configuration parameter in the model (η_A) and the count of this configuration in the observed network ($z_A(\mathbf{x})$).

On the left hand side of this equation we distinguish between the randomly generated network ($\mathbf{X}$) and the observed network ($\mathbf{x}$), both have the same number of nodes (or settlements in the case study presented below). Between every pair of nodes there can either be an arc or not (i.e. settlements can either be connected by a line of sight or not). In the observed visibility network we know exactly which nodes are connected by a line of sight, but in randomly generated networks this line of sight will be created with a certain probability. This probability is determined by the effects one includes in the ERGM.

More formally, for every pair i and j that are distinct nodes of a set N of n nodes, a random variable X_{ij} exists, where $X_{ij} = 1$ if there is an arc from node i to node j, and $X_{ij} = 0$ if there is no arc. If X_{ij} is an arc random variable that can have a value 1 or 0 with a certain *probability*, then let x_{ij} be the *observed* value (the arc that is part of our observed visibility network). Similarly, we define **X** as the matrix of all *variables* and **x** the matrix of *observed* ties. Since nodes are not supposed to have self-loops the diagonal of these matrices are empty cells. In the case of our visibility networks, X is *directed* which means that X_{ij} is different from X_{ji} (Robins 2011; Robins et al. 2007a).

The general form of an ERGM is:

$$\Pr(\mathbf{X} = \mathbf{x}) = \left(\frac{1}{\kappa}\right) \exp\left\{\sum_{A} \eta_A \, z_A(\mathbf{x})\right\}$$

Eq. 1: general form of ERGM.

where the summation is over all configuration types A; η_A is a parameter corresponding to configuration type A, it reflects the dependence assumption of a particular configuration (i.e. η_A cannot be zero if the frequency of configuration type A is considered to be dependent on the rest of the network); $z_A(\mathbf{x})$ is a count of the number of configurations A observed in **x**; κ is a normalizing quantity which ensures that equation 1 is a proper probability distribution (Robins 2011; Robins et al. 2007a).

Estimation

An ERGM goes through a process of estimation before it can be fitted to the observed networks. The process of estimation described here and applied in this case study is called the Monte Carlo Markov Chain Maximum Likelihood Estimation (MCMCMLE) approach (Koskinen and Snijders 2013).

The estimation process is aimed at refining the parameter values (the weight attributed to the configurations) by comparing the frequency of modelled configurations in the observed network against that in a distribution of random networks generated by a stochastic simulation using the approximate parameter values. These parameter values are adjusted through iterating the simulation so that the means of the values of the configuration in question can get as close as possible to the observed values. With “as close as possible” we mean: a t-ratio for the estimate of every configuration derived at every simulation; the t-ratio is calculated as follows:

$$\frac{(observed\ frequency\ of\ configuration - mean\ of\ simulated\ frequencies\ of\ configuration)}{standard\ deviation\ of\ simulated\ frequencies}$$

Eq. 2. Calculating the t-ratio of an estimated configuration.

The t-ratio indicates how well the estimate has converged with the observed data; a good convergence is indicated by t-ratios for parameter estimates of less than 0.1 in absolute value. These final parameter values are called the maximum likelihood values. Statistically significant effects (here indicated by *) have a parameter estimate in absolute value more than twice the standard error. Table 8 offers an example of this for a visibility network. We see that the t-ratio for the reciprocity effect is less than 0.1, indicating that the estimated parameter value for reciprocity produces a similar frequency of this configuration as has been observed in the visibility network. Moreover, the estimate is positive and more than twice the standard error, indicating there is a significant tendency towards the creation of inter-visible arcs, more than expected by chance.

Table 8. Example estimate of one effect in an ERGM for a visibility network.

Effects	Estimates	Standard error	T-ratio	
reciprocity	8.00	0.79	-0.07	*

Goodness of fit and interpretation

Once maximum likelihood values are obtained for the configurations included in the model the ERGM needs to be fitted to the observed network. This is done to evaluate whether the frequency of observed configurations included in the model are well reproduced by the model, as well as to check if all the other features of the observed network that are not explicitly modelled are replicated (e.g. degree distribution). The rationale behind this "goodness of fit" test is that the plausibility of an ERGM is higher if it can replicate all or most of the features of an observed network (Robins et al. 2007b, 206). The guidelines set out in Harrigan (2007) for determining whether the goodness of fit results suggest that the model is plausible are commonly used:

1. "If the parameter was estimated and specified in your model … then the t-statistic needs to be below 0.1 (as it was in the estimation)."

2. "If the parameter was not estimated and specified in your model … then the t-statistic should be below 2 for the model to not be a bad fit."

To give an example, table 9 offers a fragment of the goodness of fit results of an ERGM for a visibility network, where the reciprocity effect was included in the model whilst the 2-in-star effect was not. It shows that the mean number of these configurations in the simulated networks is very close to their frequency in the observed network. Moreover, this table offers the information needed to calculate the t-ratio as described in equation 2 above. For the reciprocity effect this is (32-31.966)/2.612=0.013 and for the 2-in-star effect this is (78-79.134)/21.836=-0.052. Since both are lower than 0.1 in absolute value they indicate a good fit between the model and the observed network.

Table 9. Example of the goodness of fit results of an ERGM for a visibility network.

Effects	Observed	Mean	Standard deviation	T-ratio
reciprocity	32	31.97	2.61	0.01
2-in-star	78	79.13	21.84	-0.05

The results of the estimated configurations are subsequently interpreted. A positive parameter estimate indicates a tendency to form this particular configuration higher than purely by chance and a negative parameter estimate indicates that the configuration appears less often then expected purely by chance. Robins and colleagues (2009, Table 1) provide a useful key for technical interpretation of effects in ERGMs of directed networks, such as visibility networks. Although a technical interpretation of ERGM results is a necessary first step, this should always be followed by a discussion of the importance and implications of the results within the archaeological research context.

4.4.6. Software

The probable viewsheds were created in ArcGIS 10.1 using a python script written by David Wheatley (University of Southampton). Networks were created using UCINET, PAJEK and CYTOSCAPE. Exploratory network measures were derived with PAJEK. ERGMs were estimated and fitted using PNET.

4.5. Results of the combined approach

I decided to include this long and technical description of the functioning of the ERGM method because it forces a researcher to explicitly work through the network science research process, and because I believe this approach shows particular promise for the evaluation of the potential of network science in archaeology. In the above I have discussed a particular archaeological research problem and I have developed an approach to this problem that draws on a diversity of exploratory and confirmatory network techniques. Box 6 offers a summary of the research process presented by this case study. In what follows I will present the results of this innovative method, discuss their implications for this case study's research questions, and finally discuss how this case study advances the aims of this PhD project.

Box 6. Network science research process case study 2

Phenomenon studied:

- The ways in which lines of sight between Iberian and Roman urban setltements structured human behaviour and the processes through which they were established.

Abstraction as network concepts:

- The entities of research interest are individual settlements, where observer locations are used as an abstraction of the ability of members of the settlement communities to see other settlements.
- Relationships between settlements are used as an abstraction of lines of sight which could structure past human behaviour.

Representation as network data:

- Iberian settlements and Roman towns are represented as nodes.
- Arcs represent lines of sight from an observer on one site to an observed point at another site.

Dependence assumptions:

- Lines of sight appear and disappear independently of each other (random).
- If communication or signalling between settlements needed to occur then these need to be inter-visible (represented by reciprocal arcs).
- In order to perform some sort of visual control over surrounding settlements, these need to be visible from a given settlement (represented by many incoming arcs).
- If a settlement is purposefully visually prominent, it needs to be visible from surrounding settlements (represented by many outgoing arcs).
- If visual isolation is considered important then settlements will be expected to be invisible from surrounding settlements (represented by isolated nodes).
- Iron Age II settlements that continue to be occupied in Roman times could be inter-visible, visually prominent or visually active.
- Roman towns with urban status could be inter-visible, visually prominent, or visually active.
- Roman towns on Roman roads and navigable rivers could be inter-visible, visually prominent, or visually active.
- Sites on hilltops could be inter-visible, visually prominent, or visually active.

Network science techniques used:

- Node-based exploratory network analysis: indegree, outdegree, node clustering coefficient.
- Network-based exploratory network analysis: number of nodes, number of arcs, clustering coefficient, number of connected components, average degree, diameter, average shortest path length, density.

- Exponential random graph modelling: Bernoulli random graph model; social circuit model; attributed model.

4.5.1. Exploratory network analysis[18]

Sensitivity analysis of exploratory network metrics

In the following exploratory network analysis and the ERGMs I will largely focus on lines of sight with a probability over 50%. However, the results of network measures can vary significantly when a network is reduced to subnetworks using certain thresholds (Peeples and Roberts 2013). I performed a sensitivity analysis to explore the effects of different thresholds on the probability of arcs on the overall results of the network measures and the site rankings. This analysis was performed for the Early Imperial network only, although similar results can be expected for networks of other periods given their general structural similarity. The networks used are directed, contain isolated nodes, and do not contain loops (arcs from one node to itself).

The results suggest that for many measures (global density, all degree centralization, local clustering coefficient) the networks with probability >50% give stable results, indicating that a focus on these high-probability subnetworks in the exploratory analysis is representative (Fig. 31). However, the exploratory network analysis should also incorporate the global clustering coefficient, and the indegree and outdegree rankings of low probability networks, since these show significant sensitivity to changing thresholds.

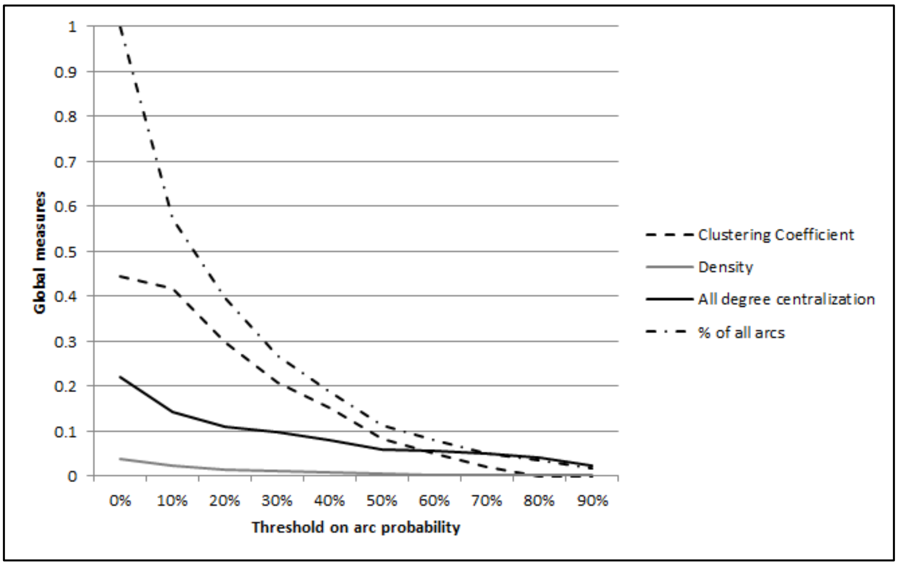


Fig. 31 Results of global network measures for ten networks with a different minimum probability of arcs (in %).

[18] Appendix III section 11.3 presents the results of the local exploratory network measures.

***Global network measures*[19]**

The global network measures in table 10 confirm that the number of nodes and arcs decrease significantly through time. To some extent this can be explained through the general decrease in the number of occupied settlements (see Table 7). However, in this exploratory network analysis I will try to address the extent to which the network patterns cannot be explained simply with reference to a decrease in the number of sites. The number of sites and lines of sight per period are nevertheless important since they determine the minimum and maximum values of many network measures, and they are therefore always used as a benchmark for interpreting the analytical results and comparing across periods.

The networks limited to a 20km radius of all periods are very fragmented and sparse (Table 10). The number of visibility links per site remains low throughout all periods, although there is a slight increase in the normalized average degree. The normalized number of connected components and the density show a similar slight increase, indicating an increasing fragmentation through time while the largest components become denser.

It is notable that the percentage of highly probable arcs (>=50%) generally decreases through time, with the Iberian and Republican periods showing a significantly higher proportion of high-probability lines of sight than the Imperial periods (Fig. 32). A more nuanced picture emerges when one compares lines of sight of different lengths. The drop in the proportion of high probability lines of sight in the Early Imperial period shown in figure 32 is the result of a decrease of high-probability arcs longer than 50km. When taking arcs of all probabilities (1-100%) into account we see that the proportion of shorter-distance arcs increases through time, while the proportion of arcs with lengths between 20 and 50km decreases (Fig. 33). The graph in figure 34 with the number of high probability arcs is quite different, however. A significantly higher proportion of arcs in the Iberian and Republican periods have a length larger than 50km, indicating that a considerable proportion of the shorter distance links in these periods have a low probability. Moreover, the majority of arcs with high probability in the Iberian period fall within the 20 to 50km range. Although this proportion does not change very much, the subsequent periods do show an increase in the proportion of arcs shorter than

[19] The results presented in this section only take into account pairs of nodes that are connected and do not include the many isolated nodes dated to these periods (see Fig. 35). This affects the normalized and averaged results as well as the density and clustering coefficient (note that network density including isolates is also included in the tables). This decision was made to enhance the differences between these measures and to focus this first step of the analysis explicitly on the lines of sight that are present. Isolates are included in the local exploratory network analsysis and the ERGMs.

20km at the expense of arcs longer than 50km. The low probability arcs show a similar trend towards an increasing proportion of short distance lines of sight. All this indicates that long-distance lines of sight become extremely rare in the Imperial periods. This might be considered a result of the decrease in the number of settlements on elevations. The lines of sight shorter than 20km on the other hand become more prominent in the Imperial periods.

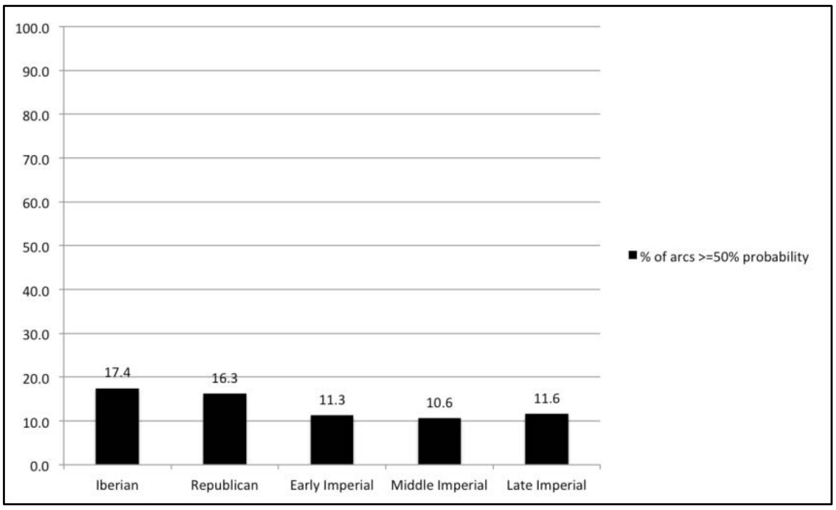


Fig. 32. Percentage of arcs per period with a probability higher than or equal to 50%.

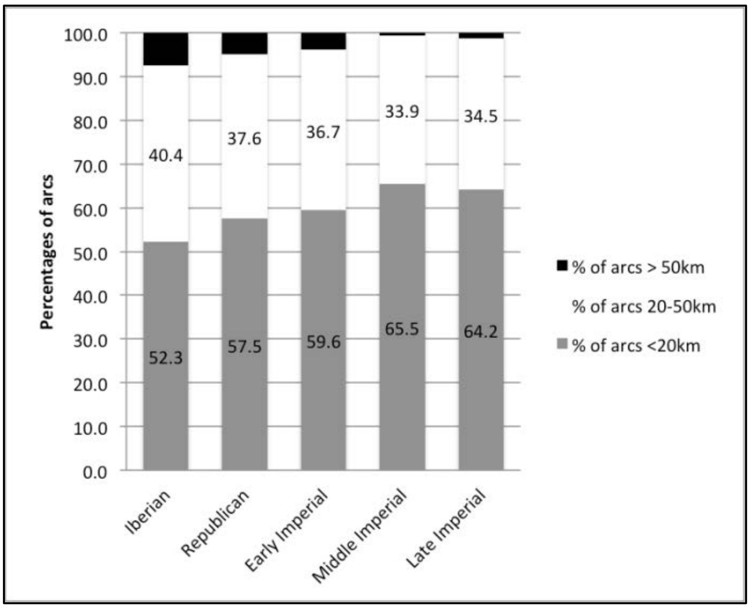


Fig. 33. Percentages of arcs shorter than 20km and 50km and longer than 50km for each period. All arcs with probabilities between 1 and 100% were included in this graph.

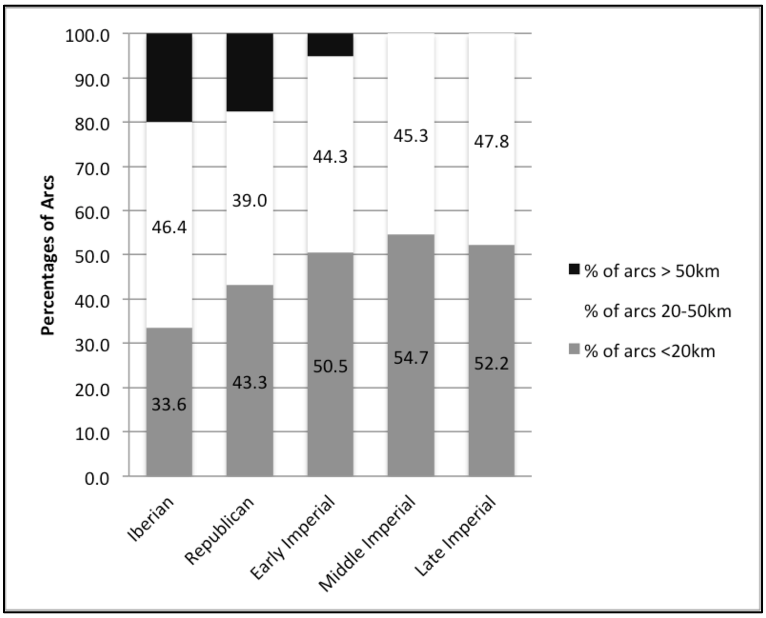


Fig. 34. Percentages of arcs shorter than 20km and 50km and longer than 50km for each period. Only arcs with probabilities over 50% were included in this graph.

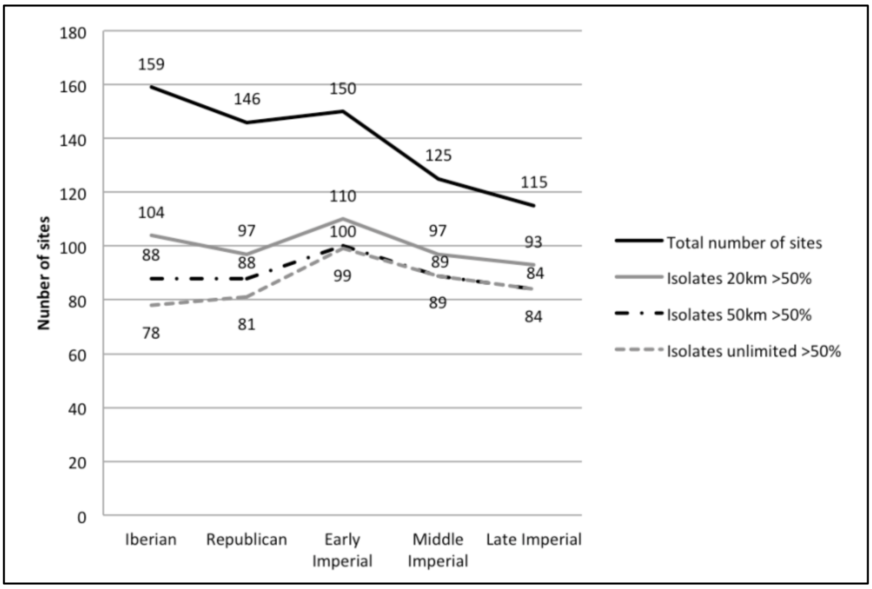


Fig. 35. Total number of sites and number of unconnected nodes per period and network type.

Table 10: global measures networks limited to 20km radius with probability >50%.

	Iberian	Republican	Early Imperial	Middle Imperial	Late Imperial
Clustering coefficient	0.107	0.108	0.056	0.000	0.000
Connected components	13	12	11	9	8
Con. Comp. normalised	0.241	0.245	0.282	0.321	0.364
Diameter	5	6	4	4	4
Node pairs	204	181	98	51	40
Av. Shortest path length	1.917	2.044	1.694	1.608	1.500
Av. Degree	1.815	1.837	1.641	1.429	1.364
Av. Degree normalised	0.034	0.037	0.042	0.051	0.062
Nodes	54	49	39	28	22

Density	0.028	0.030	0.033	0.038	0.054
Density (incl. isolates)	0.0031	0.0034	0.0022	0.0019	0.0019
Arcs	79	71	49	29	25

Network analysis of arcs <20km

As mentioned above, these networks of lines of sight up to 20km are very fragmentary and the vast majority of nodes are unconnected (Table 10, Fig. 36-37). Only a few areas have a higher number of sites that are inter-visible and these areas make up the main components of the networks: the middle- and upper-Genil valley (including the sites of La Alcuza, Mochales, El Mocho, and Castillo de Alhonoz), the area around Pancorvo (including the sites of Pancorvo, Cerro del Bollo, Las Aguzaderas, and Las Mazmorras), and the middle Corbones valley (including the sites of San Pedro I, La Torre II, and Porcún I). The latter cluster disappears in the Middle Imperial period when these sites cease to be occupied. As one would expect, the sites within the denser components have higher clustering coefficients and a higher indegree and outdegree.

The networks including arcs of lower probability show the same core areas (Fig. 37). What the networks of higher probability do not show, however, are the two clusters of low probability inter-visibility along the lower Guadalquivir valley (the area of Santiponce and Seville) and the lower Genil valley. The former emerges in the Republican period and is no longer present in the Late Imperial period, while the latter is a fully connected cluster present in all periods (the connected cluster includes Doña Mencia, Cortijo Nuevo, La Saetilla, Las Valbuenas, Isla del Castillo, and Las Animas).

There are more sites with exclusively incoming arcs of high probability and no outgoing arcs (i.e. sites from which no other sites can be seen but that are visible from other sites) than the other way around. An example of this in the Iberian and Republican periods is Pozo del Carretero, a hilltop site located in the densely urbanised middle Corbones valley. Morón de la Frontera serves as an example for the Early and Middle Imperial periods (although this site was occupied throughout the other periods as well, when it was not visible over short distances from other sites). Morón de la Frontera is located on a hilltop in an area with a low density of sites. An exception to this trend is Tejada la Vieja in the Iberian period, which cannot be seen but from which two sites can be seen. The latter occupies a prominent fortified location on a plateau in the foothills of the Sierra Morena on the western extent of the study area, but given the low density of sites in this part of the study area the site does not occupy as prominent a position on the networks as its location would suggest. The ERGMs discussed below suggest that these types of configurations are indeed quite key to understanding these networks' structures.

The sites with a Roman urban status are not very prominent in these 20km networks. The *coloniae* of Seville, Écija, and Osuna are not inter-visible with any other sites with high probability, and low probability arcs connected to these sites are few as well. Seville can be seen from Piesolo I with 1% probability, Écija can be seen from La Alcuza with 29% probability, and Osuna can be seen from Cerro del Calvario with 10% probability. Very few of the *municipia* have high probability inter-visibility links. Las Cabezas de San Juan and Palmilla I are inter-visible (84% probability), Pancorvo can be seen from the *municipium* of Torre del Aguila (66%), the riverside *municipia* of Cantillana and Alcolea del Rio are inter-visible (100-96%), and Mesa del Almendro can be seen from the *municipium* Peñaflor (56%). Many of the *municipia* do have low probability inter-visibility links, however, including El Casar, El Gandul, Torre del Aguila, Torres de Alocaz, Lebrija, Castillo de la Monclova, Isla del Castillo, Alcala del Rio, Santiponce, Gerena, and San Juan de Aznalfarache. Carmona does not have any lines of sight to close-by sites.

One can conclude that these networks limited to a 20km radius around settlements represent local patterns of inter-visibility and are sensitive to the distribution of sites. All key clusters identified in these networks are areas with a high density of sites. Many visually prominent sites that have a key position in the networks are occupied in the Iberian and Republican periods but cease to be so in the Imperial periods (e.g. Pancorvo, Tejada la Vieja, Tablada, San Pedro II, Pozo del Carretero).

Although I consider lines of sight longer than 20km of minor importance for enabling communication and control, a few striking patterns of the lines of sight with a length between 20 and 50km should be mentioned. These lines create links between areas of dense local inter-visibility, and a number of sites (e.g. Morón de la Frontera) can only be considered visually prominent at these longer distances. As the number of occupied sites decreases through time, however, so does the degree of local clustering, the indegree and outdegree. The longer distance networks therefore confirm the trend evidenced by the shorter distance networks: the networks increasingly fragment and with a few exceptions (e.g. Morón de la Frontera) the proportion of local visibility ties becomes increasingly dominant. Few sites with an urban status are inter-visible with high probability over such long distances. The *colonia* Osuna is connected to Cagancha with 57% probability. Few *municipia* have arcs of high probability. Noteworthy are the strong inter-visibility links between Carmona and Cantillana, and Cantillana and Alcolea del Rio. El Casar is inter-visible with Palmilla I and Dehesa de las Majadilas.

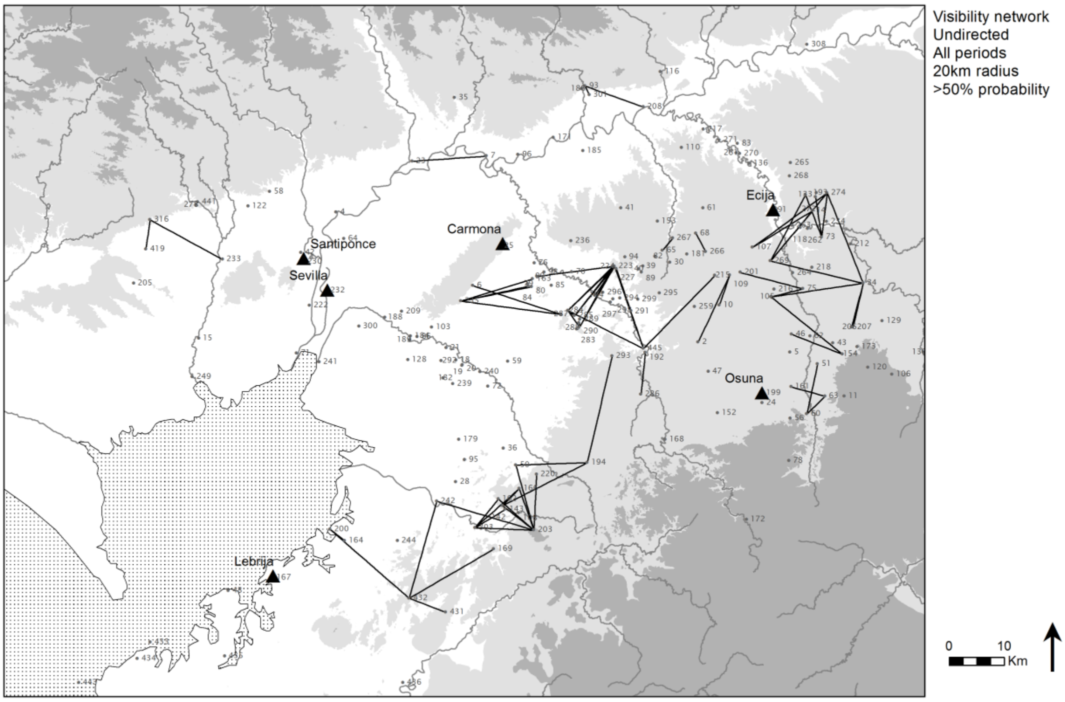


Fig. 36. Undirected visibility network for all periods combined limited to a 20km radius around observers and ties >50% probability.

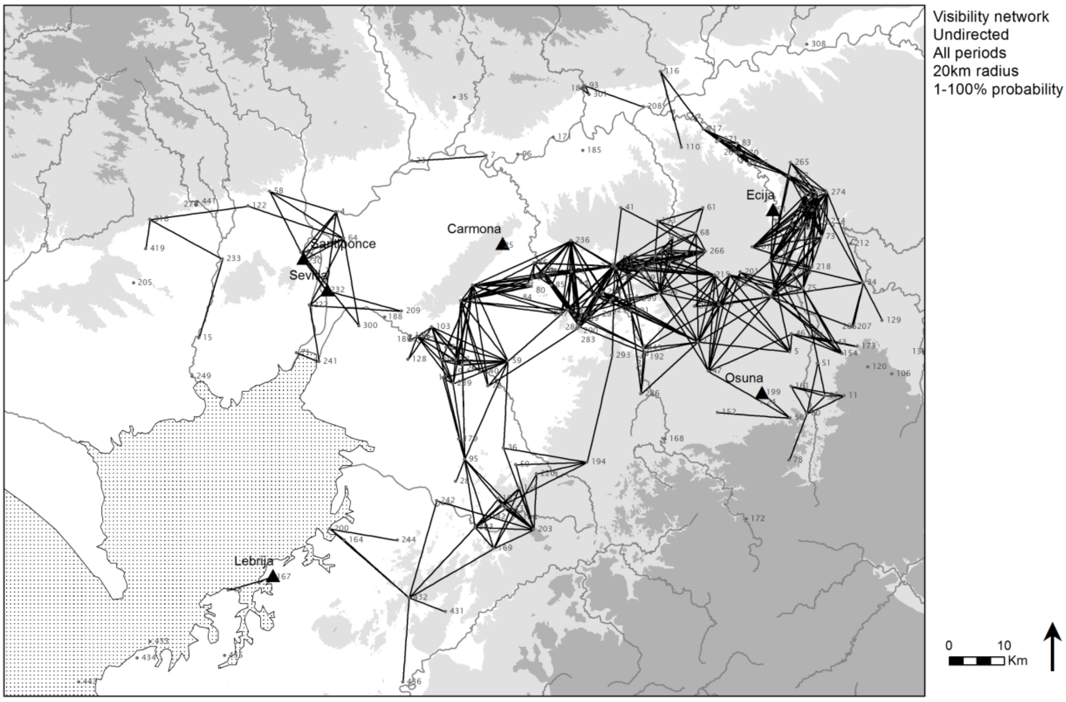


Fig. 37. Undirected visibility network for all periods combined limited to a 20km radius around observers and ties >0% probability.

4.5.2. Exponential Random Graph Models

This section will explore what kinds of processes were likely or unlikely to have led to the observed visibility networks. The assumptions that go into these models are introduced from the bottom-up, starting with random processes, followed by simple models focusing just on the emergence of network structural features, and finally models that include site attributes. Each of these assumptions represents different hypotheses, as discussed in section 4.2.4. Unlike for the global exploratory network analysis, isolates were included in the ERGMs since the tendency for sites to be visually isolated is of interest in this study. All models discussed here only take arcs with a probability >50% into account. Estimate and goodness of fit results are available in appendix III section 11.4.

Bernoulli random graph models

In the first instance models with only one dependence assumption, the probability of arc creation, were formulated for networks of each period limited to 20km arcs. For each model 50 million networks were simulated[20] from which 1000 samples were selected.[21] The

[20] By using the directed density results of the observed networks it was ensured that all randomly simulated networks have the same directed density, and therefore the same number of nodes and arcs, as the observed networks (although small variations in the number of arcs were allowed due to rounding the density score to four decimal places).

frequency of each configuration was counted for the 1000 sample networks of each model and were compared with the configuration counts obtained from the observed networks in a goodness of fit test (see Appendix III section 11.4.1).

Table 11 shows that none of the Bernoulli random graph models manages to produce networks which are structurally similar to the observed networks. This indicates that these observed networks are significantly different from randomly generated networks with the same directed density. These results suggest that non-random processes gave rise to the observed visibility networks. Although this conclusion might sound obvious, it is by no means trivial. The processes of site selection that give rise to these visibility patterns are governed by a complex mix of social decisions that can hardly be considered random. Yet our information about these decisions (in this case represented by the observed networks) is fragmentary and might therefore appear to us to have been the result of random processes. In such cases we cannot suggest processes other than seemingly random processes that are more likely to have given rise to these networks, since the information available to us does not allow validation of such alternative processes. It is worth pointing out, however, that the Middle and Late Imperial Bernoulli random graph models manage to reproduce a number of structural features of the observed networks for these periods. Indeed, as I shall argue below, the ERGM method suggests that the data for these periods might not allow us to draw many conclusions about the emergence of these networks.

Having established that for all observed networks random processes do not provide a good solution, we can start to formulate new models suggesting alternative processes that might turn out to be more likely to have happened in the past.

Table 11. Proportion of configurations of the simulated networks which show a good fit with the observed networks (networks limited to a 20km radius for each period, no attributes). 1000 samples were taken from 50 million randomly generated networks. None of the Bernoulli random graph models succeed in reproducing the structure of the observed networks, although those for the Middle and Late Imperial periods manage to reproduce some structural features.

20km networks	Iberian	Republican	Early Imperial	Middle Imperial	Late Imperial
Proportion configurations with good fit	0/52	0/52	0/52	26/52	21/52

[21] The number of simulations was taken as high as 50 million to have a higher diversity in the distribution of results, which led in most cases to better fits than when only 1 million simulations were performed. However, the number of samples taken from these simulations (and on the basis of which the significance tests were performed) was not increased beyond the default 1000 to avoid enforcing significant results.

ERGMs of arcs <20km

ERGMs were estimated for the observed network of each period. The ERGMs presented here all show a good fit with the observed networks, but are by no means the only possible ERGMs and should be considered the outcome of a process of trial-and-error guided by our theoretical assumptions of what processes gave rise to these networks (for an example of this trial-and-error process see Brughmans et al. 2014). All estimated models and goodness of fit results are presented in Appendix III section 11.4.2. When interpreting these results I focus on the significant effects and on whether these have a positive or negative value. A positive value indicates a tendency towards the creation of this type of configuration, while a negative value indicates a tendency against this configuration.

The ERGMs show some similarities throughout all periods (Table 12). Most striking is the significant and positive reciprocity effect of all periods, i.e. we see more inter-visibility than we might expect to emerge purely by chance given the other effects in the model. In the probable viewshed approach used for this case study a high degree of inter-visibility is to be expected, although there are sometimes strong differences in the probability of incoming and outgoing lines of sight (as explained in section 4.4.3 above). All periods also show a positive alternating-in-star effect, which suggests that the indegree distribution is spread and includes sites with a high number of incoming lines of sight.

However, the similarities seem to end there. In fact, the models suggest that the role of indegree hubs is very different for each period. Indeed, this alternating-in-star effect is only significant for Iberian, Republican, and Early Imperial periods, making the indegree hubs a less important structural feature for the later periods. Most interesting is the significant and negative 2-in-star effect for the Republican and Early Imperial periods, which indicates that in these periods the network tends to be dominated by indegree hubs and not sites with a lower indegree. The distribution of outgoing lines of sight on the other hand suggests that only for the Iberian period can one argue for a tendency towards the emergence of outdegree hubs with a high degree of visual control: the alternating-out-star is only significant and positive for the Iberian period, while the other periods do not seem to have evidence for processes giving rise to sites with high visual control (for most periods it is in fact negative).

Another difference between the Iberian and later period networks is its significant and negative 2-path effect. This suggests that settlements that are visually prominent tend not to be visually controlling. It also suggests that paths through the network necessary for passing on information through signalling might not have been purposefully established. Moreover, this effect is combined with a positive and significant transitive triad (T9) in the Iberian, Republican and Early Imperial periods. This indicates that if a 2-path occurs it tends to be closed. I believe this process to be a better representation of one leading to clusters of inter-visible settlements around a visually controlling settlement, than of one leading to a communication network.

Finally, one needs to acknowledge the significant and negative isolates effect in the Iberian and Republican periods: sites will tend to have at least one line of sight connected to them. Moreover, the Republican period model has a significant negative source effect and so there are fewer peripheral sites with only one incoming line of sight than one would expect to see purely by chance. This can be partly explained through the higher number of sites and higher site density in this period. The Middle Imperial period on the other hand sees a significant positive isolates effect, which could be explained by the lower site density, but also the lower number of sites located on visually prominent elevations.

Although all of these models have a good fit to the observed networks and succeed in reproducing structural features that are not explicitly modelled (like the path length- and degree distributions), one cannot ignore the low number of significant effects for the Middle and Late Imperial periods. These results seem to reflect the results of the Bernoulli random graph models for these periods, and suggest that the available data might not allow us to deduce more than the most obvious effects (like reciprocity).

Table 12. Final ERGMs for networks limited to 20km and with arcs >50% probability. For all circuit configurations λ was set to 2. Asterisks indicate significant effects for which absolute value of estimates are more than twice the standard error. A positive value indicates a tendency towards the creation of this type of configuration, while a negative value indicates a tendency against this configuration.

Configurations	**Iberian**		**Republican**		**Early Imperial**		**Middle Imperial**		**Late Imperial**	
reciprocity	8.00	*	6.81	*	8.69	*	6.98	*	8.59	*
2-in-star			-3.66	*	-4.63	*	-3.30			
2-out-star			0.15		-0.05		-0.25			
path2	-0.52	*	0.12		0.90		0.86		-0.74	
T8									0.08	
T9	0.40	*	0.54	*	0.58	*				
Sink	-2.29		-1.16							
Source	-1.28		-5.87	*						
Isolates	-3.23	*	-6.20	*	-0.74		0.87		1.07	
AinS	2.35	*	10.02	*	5.75	*	3.92		1.85	
AoutS	2.61	*	0.21		-0.03		-1.07		0.56	

AinAout-star	0.26	-0.33

ERGMs of Early Imperial attributed networks

The previous sections described models with purely structural effects, addressing the extreme question: "what if lines of sight were the only factor influencing site location." Addressing this question was necessary to be able to explore what processes are suggested by the visibility networks themselves, before adding other explanations to the mix. The current section will add another layer of complexity to the processes suggested by the visibility networks, by exploring to what extent a few features or attributes of sites might have shaped these processes. This can only be tested for the Early Imperial period due to the nature of the attributes (discussed above). I should stress again that ERGMs are not designed to take the landscape's topography into account as a constraint in these processes (as explained in section 4.2.3). This section therefore merely illustrates another way in which ERGMs can be used to explore factors influencing site locations, one that should be complemented in future research with approaches that do take topography into account.

The results of the attributed models are presented in Table 13. They are based on the 20km Early Imperial ERGM shown in table 12. What is important, however, is that some of these models now include significant attribute related effects. Indeed, these new models show more significant effects than the ERGMs without attributes, indicating that these attributes are of importance and are likely to have influenced the structure and evolution of the visibility networks. These results seem to suggest that the observed network structure cannot be fully understood without reference to three of these attributes: sites on elevations, with Iberian origins, and with an urban status. The only attributed model that does not show any significant attribute effects is the transport network model. This indicates that the location of sites along the main road or river networks did not influence (neither positively nor negatively) the emergence of the observed visibility network.

The urban status model has a positive and significant attribute_out-2-star effect, indicating that sites with an urban status have a tendency to visually control surrounding sites, but not as hubs (i.e. their number of outgoing lines of sight is limited). There is also a negative and significant 2-path attribute effect, indicating that sites with an urban status that are visually prominent do not tend to visually control. This suggests a discrepancy between the indegree and outdegree scores for sites with an urban status. Finally, it is particularly interesting to note the significant negative 2-path effect of the Iberian 20km model, which might be an indication that the role of sites with an urban status was more similar to that of Iberian-period hub sites than that of contemporary sites without an urban status. These results are quite different to those of the exploratory network analysis, which discerned no particularly prominent position for sites with an urban status.

Similar results were obtained for the model with the Iberian origins attribute: sites with Iberian origins that are visually prominent do not tend to have many outgoing lines of sight, while sites with Iberian origins do have a strong tendency to visually control surrounding sites, far more so than all sites in general. These sites with Iberian origins also do not tend to be particularly visually prominent. This is interesting because it indicates that the more recently founded Roman sites and the abandonment of Iberian sites changed the network structure significantly, suggesting different roles for lines of sight in structuring inter-urban interaction. A model including a qualitative assessment of sites located on elevations was also created. This was done largely as a control of the method, since one would expect such a model to deliver significant results given the dominant role attributed to sites on hills and plateau sides as a result of the viewshed approach. The model suggests similar processes to the previous two models: sites on elevations have a very stong tendency towards visual control, there is no real evidence of an exceptional tendency towards visual prominence, and sites on elevations that are visually prominent do not tend to visually control. Finally, there seems to be no indication that sites with any one attribute have a tendency towards reciprocating lines of sight, i.e. sites with an urban status, Iberian origins or located on elevations are not inter-visible more frequently than one would expect in random processes.

The results of these ERGMs are highly suspicious given that the exploratory network analysis results were interpreted quite differently, especially for sites with an urban status. Further analysis of these factors is therefore needed. Indeed, our interpretation of how these attributes might have affected the creation of the observed visibility patterns changes dramatically when we create models that combine these attributes. Table 14 shows a model that combines the attributes that have shown significant effects: elevation, Iberian origins and the urban status of particular communities. What is immediately striking is the absence of any significant attribute effects save one: sites on elevations have a tendency of being visually prominent (in-2-star). This leads to the not so surprising conclusion that our qualitative observation of sites on elevations better succeeds in explaining the emergence of the Early Imperial visibility patterns than any of the other attributes. I created three additional models in which two of the three attributes are included in turn: status and elevation, status and occupation, elevation and occupation (these models did not show a perfect fit of certain non-modeled features [clustering and triangles], they are therefore not published here and any conclusions drawn from them should be expressed with extreme caution). These again suggest the significance of effects related to sites on elevations and no significant effects were found in the model without the elevation attribute. We can conclude that none of the attributes other than the elevation of sites should be considered of particular interest for explaining the observed visibility patterns.

Table 13. Final ERGMs for the Early Imperial network limited to 20km and with arcs >50% probability. Each model includes one of four attribute effects: sites on elevation, Iberian origins of site, urban status, and location on river or road network (transport networks). For all circuit configurations λ was set to 2. Asterisks indicate significant effects for which absolute value of estimates are more than twice the standard error. A positive value indicates a tendency towards the creation of this type of configuration, while a negative value indicates a tendency against this configuration.

	Early Imperial period							
Configurations	Sites on elevation		Iberian origins		Urban status		Transport networks	
reciprocity	11.88	*	11.42	*	10.73	*	10.06	*
2-in-star	-6.44	*	-7.39	*	-6.98	*	-5.71	*
2-out-star	-6.49	*	-9.22	*	-3.18	*	-2.00	
path2	4.32	*	4.82	*	2.67	*	1.85	*
T9(030T)	0.60	*	0.66	*	0.59	*	0.58	*
Isolates	-1.66		-1.92	*	-1.67		-1.34	
AinS	6.79	*	7.68	*	7.65	*	6.79	*
AoutS	0.76		1.78		1.98		1.28	
AinAout-star	0.18		-0.04		-0.38		-0.12	
attribute_interaction	0.08		-0.64		-0.46		-0.03	
attribute_sender	0.83		0.04		0.30		0.56	
attribute_receiver	1.77		1.37		0.53		-0.30	
attribute_in2star	0.95		1.89		1.52		-1.32	
attribute_path2	-3.39	*	-4.30	*	-2.63	*	-0.29	
attribute_out2star	6.25	*	8.98	*	3.28	*	1.26	

Table 14. Final ERGM for the Early Imperial network limited to 20km and with arcs >50% probability. The model includes three attribute effects: sites on elevation, Iberian origins of site and urban status. For all circuit configurations λ was set to 2. Asterisks indicate significant effects for which absolute value of estimates are more than twice the standard error. A positive value indicates a tendency towards the creation of this type of configuration, while a negative value indicates a tendency against this configuration.

Configurations	Estimate	
reciprocity	9.99	*
2-in-star	-7.70	*
2-out-star	-1.09	
path2	2.05	
030T	0.59	

isolates	-1.25	
AinS	8.24	*
AoutS	0.55	
AinAout-star	-0.03	
Status_interaction	0.33	
Elevation_interaction	-0.30	
Occupation_interaction	0.29	
Status_sender	0.38	
Elevation_sender	0.96	
Occupation_sender	-0.62	
Status_receiver	-0.56	
Elevation_receiver	-0.07	
Occupation_receiver	0.05	
Status_in2star	-0.13	
Elevation_in2star	3.21	*
Occupation_in2star	-2.12	
Status_path2	0.24	
Elevation_path2	-1.95	
Occupation_path2	0.90	
Status_out2star	0.12	
Elevation_out2star	0.99	
Occupation_out2star	-0.52	

4.6. Discussion: towards a better understanding of urban connectivity in Iberian and Roman Southern Spain

How does this long and rather technical description of results allow me to gain new insights into the structure of and processes governing urban connectivity Iron Age and Roman Southern Spain? I will focus the discussion of these results and therefore my answer to this question on the research questions of this case study:

> Question 1: does the degree of inter-visibility between sites decrease slightly as a result of the gradual discontinuity in occupation of Iron Age sites?
> Question2: to what extent were visually prominent settlements with Iron Age II origins integrated as towns into the political and economic structure of the Roman Empire? Was this truer than with less visually prominent settlements with Iron Age II origins?
> Question 3: is there evidence suggesting that inter-visibility of urban settlements was considered less important under the Roman empire? To what degree does the

establishment (whether intentionally or not) of inter-visibility patterns differ between the Iron Age II and Roman periods?

All of these questions concern a comparison between the Iron Age and the Roman periods. Therefore, a good starting point would be to evaluate how the results compare with our current understanding of Iberian settlement patterns, which can subsequently be compared with that of the Roman periods. It has been argued that this pattern consisted of fortified hilltop settlements or *oppida* around which rural settlements were positioned, where the inter-visibility between *oppida* and the visual control of the *oppidum* over rural settlements in its territory can be considered to have been important (Grau Mira 2003; 2005; Ruestes Bitrià 2008). When only considering lines of sight with high probability, the results only allow one to confirm such a pattern in a few cases. Pancorvo is probably the best example of a fortified urban settlement with strong local (<20km) and long-distance (20-50km) inter-visibility links with rural settlements in its direct vicinity. It is only visible (not inter-visible with high probability) from one other major Iberian settlement site, Torre del Aguila. Another good example is that of the fortified urban settlement of Castillo de Alhonoz, inter-visible with multiple smaller sites, but mainly over distances less than 20km. Castillo de Alhonoz is not inter-visible with any other major Iberian sites. Finally the major settlement of Vico (probably related to nearby Montemolín) is inter-visible with three smaller sites over short distances and with five over more than 20km distance. It is important to note that all three of these sites are located in parts of the study area with high densities of sites; these dense clusters might well give us a clue about the territories of which these settlements may have acted as some kind of political or aristocratic focus. However, the pattern of inter-visible major settlement sites is even less obvious from the results. There is just one example when one only considers high probability lines of sight shorter than 20km: El Nuño – Alamillo. There are no additional examples when considering lines of sight between 20 and 50km.

However, it would be wrong to argue against inter-visibility as an important feature of the Iberian settlement pattern. When considering the position of sites other than the major settlement sites it is clear that overall only a limited proportion of them are not part of the visibility network (Fig. 35) and the ERGMs indicate a tendency for sites to be integrated within the network (Table 12). A rather high proportion of lines of sight has a probability of 50% or over (Fig. 32) and most exploratory network measures, both global and local, are indicative of an integrated and dense visibility network with a high number of hubs (in particular Porcún I, Cerro del Manzano, Morón de la Frontera and San Pedro I: very little is known of the Iberian period layout for most of these). I therefore argue against an exclusive focus on the inter-visibility patterns of the major settlement sites, a research bias that might obscure the potentially interesting role that is played by lines of sight connecting rural settlements.

One could interpret this limited inter-visibility between the larger settlements and the rare occurrence of these sites exercising visual control over surrounding rural settlements in light of Grau Mira's (2005) observation for Late Eastern Iberia (3rd c. BC). In comparing these to the Classical Iberian period (4th c. BC) he observed a more dispersed settlement pattern and the disintegration of inter-visibility between settlements. Sadly the dataset and coarse chronology used here do not allow a comparison between these different Iberian periods. The ERGMs, however, do suggest that the appearance of highly visible settlements exercising visual control over surrounding settlements, and a tendency for sites to be integrated in the visibility network are significant features of the possible processes leading to the Iberian settlement pattern. I argue that the observed Iberian visibility networks have a structure that facilitates mainly local visual control and communication, as well as occasional signaling and control over greater distance. These features are only possible thanks to the existence of hubs that are inter-visible with a high number of sites and hold the potential of sharing information or exercising control between the dense local clusters of sites. As Grau Mira argues, the role of *oppida* (akin to our major settlement sites) might have diminished although in some cases the typically hierarchical settlement pattern of *oppida* surrounded by rural settlements linked by inter-visibility is still clearly present. Moreover, since the settlement dataset used is more complete for the larger settlements than for smaller rural settlements, I believe the identification of additional rural settlements through future fieldwork will emphasise this structure of inter-visible clusters even more.

This is the settlement pattern that confronted the Romans when they arrived in this part of Iberia in the late 3rd c BC. The more dispersed settlements and decrease in inter-visibility observed by Grau Mira for Eastern Iberia might imply a change in the structure and function served by lines of sight between urban settlements. This study has aimed at understanding this structure and its function in more detail in the rather different cultural context of south-western Iberia, and a number of general conclusions can be drawn. Firstly, in no way does the changing structure of visibility patterns in Roman times indicate a clean break with the preceding Iron Age. This is most clear in the structural similarity of Republican and Early Imperial period networks to those of the Iberian period. I also observed some very similar tendencies towards certain configurations in the <20km ERGMs of these periods. Secondly, there are also some clear differences between all periods and it would be equally wrong not to emphasise these. Only the Iberian and Republican ERGMs of short-distance arcs show a tendency against isolated settlements and a strong tendency towards hubs with a high number of lines of sight. Moreover, the exploratory network analysis shows a gradual disintegration of that visibility network through time, resulting in a highly fragmentary and low-density network in the Middle and Late Imperial periods. The communication function of the Roman inter-visibility network was strongly affected by the discontinuous occupation of the major settlement hubs in the Iberian network (e.g. Pancorvo and Tejada la Vieja). Thirdly, I noticed that through time an increasing proportion of lines of sight were shorter than 20km. This suggests that the importance of long-distance inter-visibility diminished in favour of short-distance inter-visibility. Fourthly, a few factors that were no doubt crucial for understanding Roman inter-urban connectivity, such as the location of towns on transport networks or their urban status, cannot be considered fundamental in understanding the processes giving rise to Early Imperial visibility networks. Moreover, the ERGMs suggested that the role of towns with Iberian origins in shaping these Early Imperial visibility patterns (for example as visual focal points for new settlements) should not be over-emphasised, no doubt due to the discontinuity in occupation of Iberian-period hubs.

These general statements do tend to tip the balance more in the direction of a disintegration of the visibility networks in Roman times. But rather than direct disintegration in post-Iberian periods, I believe these results should be interpreted initially as gradual changes in its role in structuring inter-settlement interactions, possibly followed by disintegration. Both the ERGMs and the exploratory network analysis identified the importance of hubs of a lower degree, which could have taken over the pivotal role played by Iberian settlement hubs. Indeed, many of these low degree Iberian hubs were probably rural settlements that continued in occupation throughout the Roman period, such as Cerro del Bollo; Morón de la Frontera, however, was a major settlement that continued as a Roman town. An example of an Iberian fortified settlement that retains its hub-like function until the Early Imperial period is Mesa del Almendro. The increased proportion of short-distance lines of sight from the Early Imperial period onwards (largely caused by the disappearance of Iberian settlements visible over long distances) creates a pattern of small pockets of local inter-visibility, mainly in the areas where I identified a high level of inter-visibility in the Iberian period, allowing for the possibility that these lines of sight continued to function as media for local control and communication. While key Roman settlements along the navigable part of the Guadalquivir are hardly inter-visible (although in some cases I did identify chains of high probability inter-visibility along the river: e.g. Alcolea del Rio - Cantillana), the *Via Augusta* does cross through a few areas between Écija and Carmona that are highly visible from many sites (as identified by cumulative viewsheds, see Appendix III section 11.2). However, the degree of disintegration is much more pronounced for the Middle and Late Imperial periods.

Grau Mira (2005, 332) states that “The Iberian oppida were the spatial manifestation of an economic and socio-political model of control maintained by an aristocratic elite over an extensive peasant base living in rural sites.” The changes in the settlement pattern in the study area throughout the Roman period could be considered to reflect a similar phenomenon, in the rather different cultural context of south-western Spain, of a move away form this Iberian model. However, this was a gradual process and the Republican period visibility network in particular should be considered a reflection of a very slow transition to a different model. Such new socio-political models should probably be understood in terms of urban settlement location close to transport networks, agrarian products, and minerals, or their integration within the wider Roman administrative system, possibly reflected by the urban status of some communities. The ERGMs allowed me to confront different aspects of settlement location, but the results indicated these factors did not necessarily influence each other.

This case study has not revealed alternative factors that might help explain the Roman period inter-visibility structure among the few tested factors, but it did identify changes in its structure and indications of its changing role. If anything, this case study has shown that the core of the Iberian inter-visibility network initially persisted but possibly played a different role within a different socio-political model. I believe this generalizing picture can be refined by focusing future work on smaller parts of the study area and by considering evidence for other influential factors in the Roman period, for which exploratory network analysis and ERGMs are not necessarily the best approaches. Future more local case studies should take the known area of occupation of sites into account and use multiple viewer locations (e.g. Garrido González 2011) preferably with a higher resolution DEM, to avoid missing key inter-visibility patterns. A good example of this is the line of sight observed in visits to the study area between Castillo de Mulva and Carmona, which was not picked up by the probable viewsheds as a result of my decision to select one observer location on the southwestern side of Carmona. This also emphasizes the need in future local studies to compare these computationally derived lines of sight with those that project members observed and recorded in the study area (which was sadly not possible within the limits of this PhD project). One way in which a network approach can still offer new information in such future local studies is through representing major Iberian settlements and key Roman urban settlements as ego-networks (Marsden 2002), where particular local patterns of inter-urban interaction can be represented and compared. Comparative studies exploring the differences in visibility patterns observed in different study areas also offer a particularly promising way forward (e.g. Moreno Escobar 2014). In future work I should also move away from the abstract space represented by the ERGMs and aim to simulate the same hypotheses within this particular landscape (i.e. including topography as a constraint in the ERGMs). This will allow me to answer different research questions than those addressed in this case study: for example, how likely is it for the observed visibility network to emerge by chance in this particular landscape. The models presented here will allow one to focus this future work on the identified range of models with significant effects. However, such future work is dependent on methodological developments in the study of ERGMs. Finally, I could compare the changes in visibility patterns with changes in the distribution of material culture in order to better place these results in a large-scale framework for the study of urban connectivity in Iron Age and Roman Southern Spain.

In this case study I have discussed in detail a critical and innovative statistical network approach that was developed to answer a particular archaeological research question. I believe this approach has allowed me to gain a critical understanding of the nature and limitations of the network dataset used, to be aware of the processes it might suggest as well as explore the implications of such processes. I concluded that data quality and theory are crucial for formulating statistical models of past phenomena, and that therefore ERGMs can only be used in research contexts where complete datasets and clearly formulated dependence assumptions are available. The next case study will illustrate how confirmatory network approaches can allow one to test Roman archaeology hypotheses through falsification, even in cases where the available data are considered a far less direct reflection of the past human behaviour of research interest and statistical modelling approaches are therefore not an option.

Box 7. Summary of chapter 4

Research questions:

- Does the degree of inter-visibility between sites decrease slightly as a result of the gradual discontinuity in occupation of Iron Age sites?
- To what extent were visually prominent settlements with Iron Age II origins integrated as towns into the political and economic structure of the Roman Empire? Was this truer than with less visually prominent settlements with Iron Age II origins?
- Is there evidence suggesting that inter-visibility of urban settlements was considered less important under the Roman empire? To what degree does the establishment (whether intentionally or not) of inter-visibility patterns differ between the Iron Age II and Roman periods?

Conclusions:

- The degree of inter-visibility between sites decreases gradually throughout the Roman periods. This is only partly caused by the discontinuity in occupation of Iron Age sites. A proportion of Iron Age sites continues to be occupied throughout the Roman periods, but occupy a different position in the visibility network.
- The way in which lines of sight might have structured human behaviour was different in Roman periods as compared to the Iron Age. Key features include low degree hubs and an emphasis on short-distance lines of sight.
- There is no correlation between the position of Roman towns on the visibility networks and their urban status, and location on river or road networks. Visibility does not seem to have been an important factor for those settlements that are integrated in the administrative and economic structure of the Roman empire.

Implications for this PhD project:

- The ability of an exploratory network analysis to enable a better understanding of past phenomena is significantly enhanced when dependence assumptions are explicitly formulated.
- I introduced exponential random graph modelling (ERGM) as a method for bridging static and dynamic approaches to interpreting visibility networks.
- ERGMs allow archaeologists to use network data to evaluate existing hypotheses and formulate new hypotheses of the past processes that drove the phenomena they are interested in. New hypotheses can be focused on the more narrow range of processes which the ERGMs suggest can lead to the datasets available to us.
- ERGMs are only as reliable as the network datasets they are based on. The observed network needs to be complete.
- Common research themes from other subdisciplines can still be of interest in Roman archaeology if good quality datasets are available to address them.

CHAPTER 5

CASE STUDY 3:

BANG'S ROMAN BAZAAR AND/OR TEMIN'S MARKET ECONOMY? AN AGENT-BASED NETWORK MODEL OF TABLEWARE DISTRIBUTION IN THE ROMAN EAST

5. Case study 3: Bang's Roman Bazaar and/or Temin's Market Economy? An agent-based network model of tableware distribution in the Roman East

5.1. Introduction[22]

In the previous case study I assumed that the visibility network dataset was complete and that the observed visibility network patterns were the direct outcomes of (at least in part) network evolution processes. This allowed me to apply a statistical modelling approach to test hypotheses of the processes driving change in visibility networks. However, what if we cannot make such assumptions? What if our datasets are far less complete, if the hypotheses that aim to explain them do not concern network evolution but rather processes taking place on networks, and if network representations of the archaeological dataset in question cannot be compared with simulated networks? In this third case study I will evaluate another confirmatory network technique which is more promising in such cases: agent-based network modelling (ABM).

Ceramic tableware is arguably the most common find on sites in the Roman Eastern Mediterranean and therefore lends itself particularly well to quantification. Such quantitative studies have contributed to a better understanding of tableware morphology, manufacture, distribution and consumption. However, quantifications of tableware in the Roman East have also raised new questions. A particularly challenging issue is posed by the significant differences in distribution patterns of different wares. Some wares like Eastern Sigillata A (ESA) were distributed on a supra-regional scale for centuries, others were more of regional importance, whilst yet other wares were purely produced for local consumption. Moreover, this pattern was not static but changed through time: from 100AD onwards the width of ESA distribution declined significantly, whilst Eastern Sigillata B (ESB) knew a supra-regional distribution in the period 30-75AD. What were the mechanisms that led to these different patterns? Admittedly, many contributing factors can be easily formulated (e.g. state involvement, redistributive centres, consumption "pulling forces", commercial "piggy-back" trade, closeness to large-scale agricultural production). Many potential answers have already been published suggesting particular combinations of factors or shifting more weight in

[22] Please see the electronic supplementary material for this case study, which includes the ABM presented here, and the ceramic dataset used for the exploratory data and network analyses as a matrix per 25-year period.

favour of some factors (e.g. Abadie-Reynal 1989; Bes 2007; Lewit 2011; Reynolds 1995). However, the challenging aspect of this question does not lie in the formulation of convincing contributing factors but in the interplay of factors. Indeed, extreme hypotheses surrounding the distribution mechanisms of tableware have rarely been very popular. Instead, most scholars seem to agree that a complex mix of mechanisms working on multiple levels was responsible for the considerable differences in tableware distribution patterns (e.g. Bes 2007, 203). Since there is evidently no lack of hypotheses and contributing factors, the key research question becomes “what mix of factors is best supported by the available evidence?” and the main research challenge of this case study then becomes the search for an approach that allows one to distinguish between the impact and archaeological “signatures” of different hypothetical scenarios.

This case study aims to contribute to this on-going discussion by evaluating published hypotheses through a combination of an exploratory analysis of the collected tableware evidence (using the ICRATES database of tablewares) with computational modelling of hypothetical trade mechanisms. Tableware trade in the Roman East will be considered to function as a complex system, where the particular small-scale actions of agents making decisions to interact based on their limited knowledge of their social network, gives rise to large-scale patterns that can be compared to the combined archaeological record. In doing this it will further this PhD project’s aims by:

(1) evaluating the impact of abstracting hypotheses as network concepts;
(2) by illustrating how network science techniques allow one to perform confirmatory analyses through falsification in research environments where statistical approaches to the network data would not allow for this;
(3) by addressing the need in Roman studies for a workable method for the evaluation of a complex mix of hypothetical processes taking place on multiple levels to better understand archaeologically attested large-scale distribution patterns (Davies 2005; Morris et al. 2007).

A number of Roman historical and archaeological contexts could serve to explore this question, including for example the distribution of African Red Slip Ware during the Middle and Late Roman Empire. In this case study, however, I will focus on an example that poses particular problems: the large-scale distribution of four ceramic tablewares produced and circulated in the Eastern Mediterranean between ca. 150BC and 200AD, with a particular focus on the period 25BC-150AD when all four tablewares circulated in the Eastern Mediterranean.[23] A large number of wares of mainly local or regional importance must have existed, but only one ware maintained a supra-regional distribution for centuries: ESA. Moreover, it reached this wide distribution before most of the Eastern Mediterranean was incorporated into the Roman Empire. This fact forces one to look for a complex combination of contributing factors rather than accepting a political framework as the sole explanation for the observed distribution patterns. In this case study I will focus my efforts on exploring the potential role of social networks as a driving force, a concept that can take many shapes. The reason for placing this focus on social networks is that many of the archaeological hypotheses suggested to explain the distribution pattern under scrutiny in this case study concern large-scale processes that imply the agency of individual traders but rarely address their functioning explicitly, let alone test this functioning with the available evidence. Social networks are here considered to represent the commercial opportunities of traders, acting as a medium for the flow of information and the trade in resources between traders. Much like one could consider space as the medium for physical transport, social networks can be considered the medium on which the distribution of tableware took place. As such, under different conditions they could allow for direct state-controlled trade in luxuries over vast distances as well as merchants transporting some “pots and pans” on pack animals for trade in a rural village. This definition also means that I make the assumption that in the Roman East social relationships between individuals are a prerequisite for the flow of goods. This assumption could refer to long-established contacts between traders as well as more casual one-off encounters between two merchants in a market. The assumption also does not merely imply a restriction to the direct flow of goods between pairs of individuals. Rather, thanks to acting as a medium for the flow of information, social networks allow for pairs of individuals to act as representatives of a pair of commercial partners to enable the indirect flow of goods.

Considering such a concept as the key to my approach is convenient, yet in order to be useful the abstraction of tableware trade in the Roman East in terms of network concepts needs to be well motivated. Moreover, an approach needs to be developed that allows for hypotheses of trade through social networks to be quantitatively expressed and compared with the observed

[23] Rather than including the many maps showing the spatial distribution of the pattern studied in this case study, I thought it more useful to present them online, see: http://icrates.arts.kuleuven.be/icrates/network-analysis/webpages/icrates_maps.html (accessed 26.06.2014).

archaeological distribution patterns. Many such hypotheses for explaining the key driving forces behind the Roman economy have been published. In this case study I will consider two very different hypotheses as a starting point: Peter Bang's '*The Roman Bazaar*' (2008) and Peter Temin's '*The Roman Market Economy*' (2013). In his model, Bang considers three factors key to understanding trade and markets in Roman Imperial times: bazaar-style markets, the tributary nature of the Roman Empire, and the agrarianate nature of ancient societies. The engine of the model, however, is clearly the concept of the bazaar: local markets distinguished by high uncertainty of information, relative unpredictability of supply and demand, leading to poorly integrated markets throughout the empire. Temin agrees with Bang that the information available to individuals was limited and that local markets are key structuring factors. However, contrary to Bang he believes that the Roman economy was a well-functioning integrated market where prices are determined by supply and demand. These conflicting models and the issues surrounding their implementation in this case study will be discussed in more detail below. For now, it suffices to mention that Bang claims that the differences in tableware distribution patterns can be better understood when thinking through this model. He argues that different circuits for the flow of goods could emerge as the result of different circuits for the flow of information. In other words, the observed distribution patterns of wares and different workshops' products (when these can be identified) are a reflection of the functioning of past social networks as I have defined them above. Temin's model can be considered to offer an alternative approach, where the structure of social networks as a channel for the flow of information must have allowed for integrated markets.

The concept of social networks in trade will be pursued in more detail throughout this case study. A number of research questions will be addressed that will guide the implementation of hypothetical scenarios, the exploration of the collected tableware evidence and the confrontation between computational model and empirical data:

- What differences can be observed in the distribution patterns of different tablewares and forms (here considered modern analytical constructs)? Are forms' distributions always more similar to those of the same ware? Were similar distribution processes responsible for the distribution of forms of the same ware?
- Can similarities and differences between wares', and forms' distribution patterns be better understood from the perspective of individual agents active in ancient trade,

assuming different roles, having limited information available to them determining their economic opportunities? Can Bang or Temin's hypotheses be falsified through computational modelling and comparison with the observed tableware distributions?

These research questions require a combined method that moves from the archaeological data to hypothetical models and back again: an approach that combines the examination of empirical data, the exploration of different archaeological hypotheses, and the ability to test these hypotheses with archaeological data. The method suggested here consists of two parts: an exploratory analysis of the distribution patterns of different tableware forms, and agent-based network modelling (ABM) of selected hypotheses. By developing and evaluating this innovative approach this case study therefore builds on and extends the exploration of formal network methods for archaeology performed in the rest of this PhD project. In particular, it pushes my research agenda further by evaluating how exploratory data analysis can be used to evaluate the simulated output of a more abstract agent-based network model representing an archaeological/historical hypothesis.

Box 8. Research outline of case study 3

Observed pattern: for centuries ESA tableware had a significantly wider distribution in the Roman East than any other eastern-made tableware, whilst many wares of purely local importance must have existed.

Archaeological aim: evaluate hypotheses of social networks acting as a driving force in tableware trade in the Roman East, in order to better understand the differences in the tablewares' distribution patterns.

Hypotheses:

- The Roman market systems consisted of weakly integrated markets due to community structure of social networks which served to protect commercial interests and opportunism whilst disadvantaging outsiders (Bang).
- The Roman economy was a well-functioning integrated market where prices are determined by supply and demand (Temin).

Methodological aim: explore how archaeological hypotheses of individuals in interaction can be expressed as simple ABMs, and how exploratory data analysis can help to compare the observed patterns with the simulated outputs of the ABMs.

Method:

- Exploratory analysis of the dataset in order to understand the structure of the pattern of interest and the limitations of the data and to identify similarities and differences in the distribution patterns of wares, and forms.
- Simulation of hypotheses through ABMs and comparison of model outputs with results of exploratory network analysis.

5.2. Tableware studies, distribution patterns, and processes

5.2.1. Introduction

What makes the study of differences in tableware distribution patterns an interesting research topic? The easy answer would be that a large number of archaeologists that came before certainly considered this to be the case, and in this section I will discuss many of the mechanisms and factors they argued contributed to the creation of these patterns. But a more meaningful answer would be based on our assumption that the vast amounts of Roman pottery are the archaeologist's (almost only) way towards a better understanding of the ancient economy. Moreover, it has been argued that the differences in the distribution pattern observed in the Roman East could be explained by the model of Roman trade suggested by Peter Bang. So far, the development of formal approaches to evaluate such models of the ancient economy with archaeological data has been limited. In this section I will work towards a description of what such an approach should look like, by discussing the types of questions Roman pottery allows us to ask, the archaeological explanations of the distribution patterns, and models of Roman trade.

5.2.2. Studying Roman pottery distributions

Any attempt at understanding Roman pottery distributions will need to address the following two issues: “The study of ceramic distribution involves, first, identifying the source of the vessel, either the workshop or the region in which it was produced, and second, examining mechanisms by which it may have reached its ultimate destination” (Sinopoli 1991, 103). Indeed, the study of distributions is concerned with places of production, places of deposition, and absolutely everything that happened in between. This last aspect is often the most elusive archaeologically, and it is exactly at this aspect that I will aim the full force of network science in an attempt to obtain new insights into the mechanisms underlying the observed patterns. This case study therefore does not aim to test hypotheses surrounding tableware production and consumption but will focus on distribution.

Roman pottery must have been used in every household alongside objects in other materials (e.g. glas, metal, wood), but thanks to its durability and unrecyclable nature pottery is often the only material found on sites around the Mediterranean in astronomical quantities. This overwhelming amount of pots holds advantages and disadvantages for the study of roman antiquity (Poblome et al. 2013). The sheer number allows one to quantify pottery and make statistical inferences about the ancient economy. Indeed, despite the traditional (and necessary) focus of ceramologists on the morphological properties of vessels, roman pottery can be considered as an index or proxy for the flows of other (possibly more perishable) goods (Greene 2005; Peacock 1982; Orton et al. 1993). It can be used to make inferences about the frequency with which different routes were used, to indicate the main direction and intensity of exchange, and ultimately to make claims about the nature and size of economic growth in antiquity (Morris et al. 2007). Pottery is the archaeologist’s means to exploring the underlying principles, mechanisms and processes driving the ancient economy (Peacock 1982, 4).

Vast amounts of pottery also come with disadvantages, however. The larger the amount of pottery on a site, the less likely a full study and quantification of all available material becomes (Poblome et al. 2013). Indeed, the study of pottery distributions (like most everything else in archaeology) is the study of fragmentary samples drawn from an unknown whole. The issue is most critical when, as in the current study, one works on an empire-wide level. Usually the samples we obtain are determined by modern-day socio-political considerations rather than by concerns for acquiring a representative sample. It are the regions and sites described most elaborately in ancient written sources, with visible evidence of monumental architecture, or in countries with a long-term Western European imperialist presence that have traditionally received more attention, whilst their hinterlands and the smaller settlements have largely escaped archaeologists' interest (although this is changing to some extent). Moreover, many institutions, archaeologists, and ceramologists are involved which leads to a differential availability, quantity and quality of data. Empire-wide studies of pottery distributions are therefore particularly vulnerable to the issue of charting the intensity of archaeological work, something that was clearly born out in previous network analyses of the ICRATES dataset (Brughmans 2010; Brughmans and Poblome *in press*).

This case study aims to take advantage of the large available datasets of roman tablewares to make inferences about processes of exchange in the Eastern Mediterranean. However, this will require a thorough understanding of the sampling bias present in the datasets and a critical method that acknowledges and addresses this bias. The method applied below includes sensitivity analyses where the impact of changing the size of the dataset on the results of the network measures is evaluated, which will influence the use of certain network techniques (see appendix IV section 12.2). But first the hypotheses and issues surrounding the study of tableware exchange mechanisms will need to be addressed.

5.2.3. Distribution processes

Many factors undoubtedly contributed to the observed tableware distribution patterns. This section provides a focused summary of some hypotheses most relevant to this case study.

The life cycle of Roman tableware

> "Une question de fond reste en effet posée à quiconque s'intéresse à la navigation et au commerce antiques: par où les navires passaient-ils réellement lorsqu'ils acheminaient, à partir d'un point A un objet produit en un point B jusqu'à un port C, à partir duquel il était écoulé vers un lieu de consommation D, sachant que B et D sont généralement les seules certitudes?" (Arnaud 2005, 7-8).

This question posed by Pascal Arnaud reveals the challenge we are faced with when studying ancient commerce and transport through pottery evidence: what happened to an object between place of production and place of deposition is generally unknown. Developing hypothetical models of the life cycle of Roman tableware is one research avenue scholars have walked down to address this challenge, and such studies therefore offer a good starting point for the discussion of distribution processes.

A simple model for artefact life cycles was proposed by Schiffer (1972) and included phases of manufacture, use, maintenance and discard. However, the particular nature of tableware data and our knowledge of the complex mechanisms surrounding its production and exchange allow one to formulate more detailed models. Theodore Peña (2007, 328-329) suggests a formal model for the life cycle of Roman tableware, expanding Schiffer's basic model with processes of distribution and distinguishing between prime use, reuse, recycling and reclamation (Fig. 38). Moreover, Peña suggests hypothetical proportions of tableware affected by certain processes and hypothetical durations of prime use (for similar models for the Eastern Mediterranean, see contributions in Lawall and Lund 2011). Although this model is a useful guideline, it is obviously a generalising representation of a complex mix of processes. It does not explicitly represent particular mechanisms of manufacture, distribution and consumption. Nor does it reflect how individuals' social contexts or roles/professions affect their involvement in such mechanisms. It is also a deterministic model, it does not allow for variability or particular unexpected events. It should be clear that these arguments are not criticisms. In order for Peña's model to be a useful representation of a general past phenomenon (e.g. the "life cycle of tablewares distributed beyond locale of manufacture" Peña 2008, 329) it needs to be generalising. However, in this case study I aim to explore a particular distribution pattern, framed by particular archaeological research questions and contexts. For the purposes of this case study one could therefore consider it useful to expand Peña's model even more, by paying more attention to the distribution mechanisms. And so we enter the academic battlefield of models and hypotheses that is the study of Roman trade.

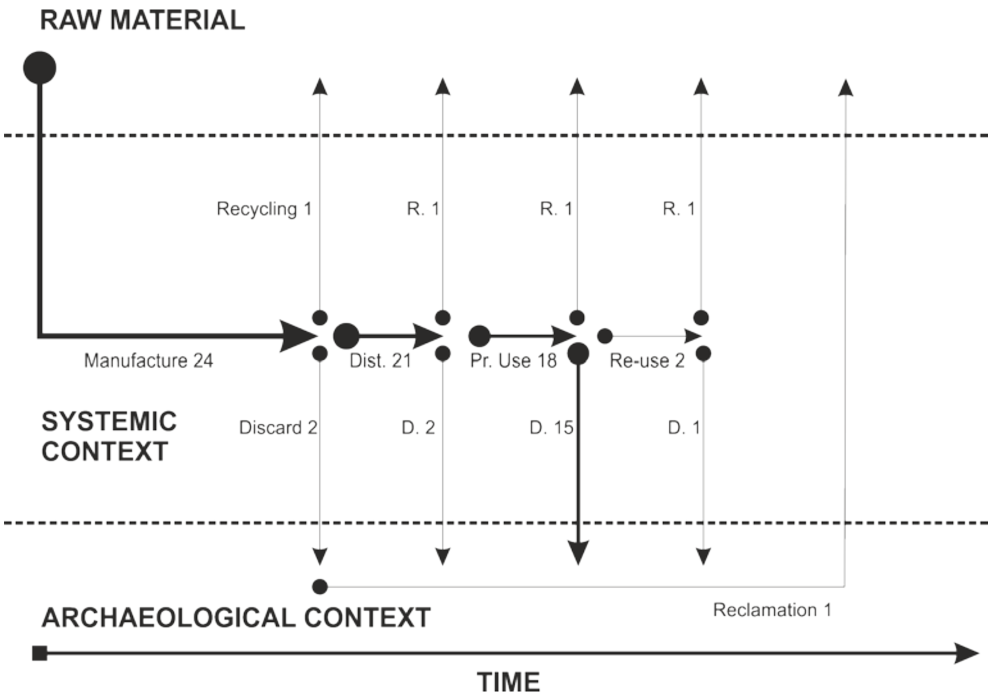


Fig. 38. Simplified version of flow diagram representing the life cycle of tablewares distributed beyond locale of manufacture, based on Peña 2007, Fig. 11.5. Each arrow represents a process, the numbers represent the proportion of all production that undergoes this process, and the width of the arrows reflects this number. A hypothetical total of 24 vessels is considered.

One archaeological pattern, many hypotheses

Clearly an in-depth discussion of the study of the ancient economy is neither feasible nor desirable in this context (for a recent overview see Scheidel et al. 2007). It is more useful to be guided through this massive research topic by the hypotheses and mechanisms surrounding the phenomenon I aim to explore in this case study: the differences in the geographical distribution patterns of different tablewares.

This research topic is immersed in a wide diversity of ideas and opinions about the functioning and performance of the larger, more visible flows (for an overview see Willet 2012, 443-449). The discussion can be suitably opened with the treatment of the topic by Philip Bes in his study of tableware distributions in the Roman East (using the ICRATES database). He argued that "four mutually dependent factors may underpin the supra-regional distribution pattern of … *sigillatas* and *red slip wares*" (2007, 204-205. Bes' italics):

1. The symbiosis between an active urban hub and a productive countryside (see also Lund 2003; Poblome 2006; Poblome et al. 2007).
2. The existence and creation of pulling forces (political-administrative, economic, religious, military or a combination of these), e.g. Delos, Corinth, Alexandria.

3. Interconnectivity of the people and places involved through lines of communication.
4. A political or other system that 'encapsulates' these factors.

By formulating these mutually dependent factors Bes clearly argues against simplistic one-dimensional and purely exogenous explanations of tableware distribution patterns. These factors imply a long list of variables to be considered: physical transport, geographical and ecological context, agrarian productivity, urbanization, administrative and political system, military presence, inter-regional connectivity, social networks. Rather than a single factor dominating the explanation of the much wider distribution of some wares, each of these factors might have had a proportionally stronger influence in the complex mix of factors at any point in time and space. Bes therefore concludes that "more diversified socio-economic reconstruction is required, which will need to take the message of local and regional levels of production into account" (Bes 2007, 203).

Similar contributing factors are suggested by Tamara Lewit (2011), although the absence of the important third point about connectivity makes Lewit's arguments more generic than Bes'. Nevertheless, Lewit poses a question that is strongly related to my current aims and her arguments should therefore be compared to those of others. Lewit argues that archaeologists have used Roman tablewares as dating tools whilst "less attention has been paid to a key economic question they pose: what exceptional factors provided the impetus for some groups of producers to break the typical pattern of local/regional distribution and become large-scale, supra-regional exporters, sometimes over a period of centuries, when similar fine pottery was made and could have been obtained more locally almost anywhere in the Roman world?" (Lewit 2011, 318). Although Lewit considers the case of the distribution of ARSW and Phocaean Red Slip Ware (PRSW), her arguments raised to answer this question are of interest to my current focus on ESA as well. She considers four crucial factors:

1. Plentiful availability of resources, including large amounts of fuel for firing kilns. Lewit argues fuel might have been got from oil production, and suggests tableware might have been transported together with oil amphorae.
2. The Annona state and military supply system, implying an involvement (directly or indirectly) of the Roman state.
3. Commercial exchanges, both long- and short-distance (see also Peacock 1982, 159).
4. Closeness to other commercially valuable resources such as minerals.

She concludes that supra-regional tableware exports should be considered the result of exceptional stimuli in combination generating a not entirely predictable dynamic. These are all rather specific factors that Lewit argues occurred in combination and were of changing importance. Her focus on exceptional external stimuli largely leaves potentially influential endogenous circumstances out of the equation. Bes' third point about connectivity suggests such endogenous circumstances, although they are admittedly rather vaguely described.

Towards the more deterministic side of the spectrum as compared to Lewit and Bes we can position the argumentation by Reynolds (1995, 126-139). Tableware production and distribution are entirely understood in light of the flows of other goods, possible local variability relatively untouched by external stimuli is completely absent in the following arguments:

1. It is considered unlikely that tablewares were carried as sole cargoes (see Parker 1984).
2. Tablewares are considered secondary cargoes, whilst foodstuffs and raw materials are primary cargoes (this is sometimes referred to as 'Piggy-back' trade, Greene 1992, 58-59).
3. "the development for export of secondary industries (e.g. pottery, fish-sauce) is directly related to the success of primary industries … and the successful exploitation by secondary industries of markets and shipping routes developed for primary industries" (see also Fulford 1987) and "the greater the frequency of contact between major agricultural producers and consumers, the greater the range and quantities of pottery exported" (Reynolds 1995, 128). The latter argument is considered particularly important for annona trade.
4. "The distribution of secondary and tertiary goods must always reflect the distributional paths, i.e. shipping routes, of the primary cargoes, whether the latter are identifiable … or not" (Reynolds 1995, 129).
5. The distance travelled by secondary goods depends on the value of the primary cargo as well as on the frequency of exports of primary goods.
6. "Various, inter-related, models, it can be argued, account for the regional patterns in the distribution of Eastern Mediterranean products in the West: a) strong, frequent shipping routes; b) the distinct composition of cargoes … ; c) distinct, separate, shipping routes (hence, contract or economic ties) supplying specific ports (cf. b); d)

> direct distribution in Eastern ships; e) redistribution in Western … ships" (Reynolds 1995, 132).

Reynolds provides us with a model emphasizing the importance of "piggy-back" trade, where tableware is not traded over long-distances for its own sake but rather makes up a relatively small proportion of cargoes (c. 10% as suggested from one shipwreck by Millett 1993). Shippers in both commercial and official movements of goods were often able to include such secondary (and tertiary) products in addition to their main cargo. Although these arguments are perfectly logical and extremely useful, they do tend to imply that an understanding of the mechanisms governing tableware distributions is 100% dependent on primary goods. They do not allow answering questions such as whether a wide distribution of tablewares could have been obtained without direct regular trade in primary cargoes between redistributive centres.

Together with Bes and Lewit, however, Reynolds provides us with a wealth of factors and hypotheses that no doubt all played some role or other. Indeed, it is not the arguments themselves that I aim to challenge here, but the approach to evaluating particular combinations of factors that could explain tableware distributions. Whereas the abovementioned authors relied on exploratory data analysis and assumptions about what patterns are to be expected as outcomes of certain mechanisms, I aim to additionally simulate the expected outcomes of specific mechanisms and compare them with the archaeological record. Moreover, although many of these arguments imply local and particular commercial actions, their functioning and roles are not explicitly stated. Indeed, Willet (2012, 443-449) argues that a lack of detail in the data prevents us from being able to explain "the complexities of smaller scale trade" (the inability to trace the production and distribution of individual potters due to the lack of stamps or other markers on many Eastern wares is discussed as a specific limiting factor). I believe this statement might be usefully explored and challenged by considering the admittedly crude distributional data the potential outcome of aggregated small- and large-scale processes. It is clear that an approach is needed that allows for attributing differing importance to different factors, active on multiple geographical scales as well as allowing for some degree of variability. Below I will argue that computational modelling might hold some potential in this respect. First, however, I will briefly elaborate on two additional factors that add to the complexity of the archaeological explanatory efforts introduced here, but the functioning of which is not always made specific in these efforts: transport mechanisms and the roles of intermediaries.

What about transport processes?

Distribution is a rather generic term that implies both transport and trade: in the models of the life cycle of pottery, distribution involves all flows between production and deposition. Arguably, then, the question of what processes were responsible for tableware distribution cannot be addressed without considering different transport processes, or can it? Many of the archaeological hypotheses mentioned above imply or explicitly concern transport processes (especially those suggested by Reynolds), and we will see below that intermediaries are also often concerned with transport to differing degrees. Both land-based and maritime transport should be considered, although the costs, transported volumes, and people involved in both types of transport differ greatly. For maritime transport, Pascal Arnaud (2005, 9-10) argues that the limitations of our data and the research questions we are typically interested in leads to a tendency to limit ancient navigation to two opposite but equally reductionist patterns: direct navigation from point of departure to destination, and *cabotage* (commercial coastal navigation and hopping between close-by ports). These two patterns are very much a product of the limitations of our data: our general ignorance of what happened between production and deposition of an object suggests direct connections between these two points, and the common pattern of decreasing quantity of a certain object as one moves away from its place of production suggests *cabotage*. However, here I will follow Arnaud in making the important distinction between the flows of objects and the routes taken by the vessels that transported them:

> “Une confusion certaine s’attache en fait à la notion de route maritime, que l’on distingue insuffisamment du flux maritime. La route maritime est l’itinéraire effectivement suivi entre A et B quand le flux s’attache au constat qu’un produit réputé élaboré en un point A a été transporté jusqu’à un point B. Les conditions naturelles de la navigation n’ont jamais interdit les flux lorsque l’économie ou la politique l’imposaient. Elles imposaient en revanche un nombre limité de routes déterminées et favorisent certains itinéraires. La plupart des cartes des échanges établissent des relations directes, comme tirées à la règle sur la carte, jusqu’à évoquer parfois un pont aérien, entre le lieu de production et le lieu de consommation d’une série cohérente d’objets. Certes, le produit concerné est bien allé d’un point A à un point B, et l’on peut ainsi déterminer un flux commercial. Il convient néanmoins de prendre garde de ne jamais confondre le flux commercial ainsi identifié avec le parcours réellement effectué par un navire. Si les flux commerciaux peuvent, et doivent constituer des indices pour établir des hypothèses de route, ils ne sauraient à eux seuls les révéler” (Arnaud 2005, 11).

When making this distinction it becomes clear that some of the archaeological hypotheses mentioned above are concerned more with trade and others with navigation. When exploring these hypotheses, different factors should be considered key driving forces: in navigation we should consider ship technology, winds, currents, seasonality, visibility and other environmental factors; in trade we consider the integration of markets and the available information, the intermediaries mediating transactions, the diversity in supplied products, the demand, and the price setting mechanism. It should also be clear that this case study does not aim at reconstructing particular sailing routes in the Mediterranean through a comprehensive computational model. The approach taken here is rather different: in order to be usefully explored and tested, specific hypotheses surrounding tableware distribution will be expressed as simple models that will include a limited number of key parameters representing transport or trade mechanisms, with a stronger focus on trade mechanisms. The resulting models are generalizing abstractions and do not aim at capturing the full complexity of ancient trade or navigation. However, they might prove useful 'tools to think with' when evaluating the role of different factors in giving rise to the tableware distribution pattern under study in this chapter.

The importance of intermediaries

There is a need to elaborate on another factor that deserves particular attention, the importance of which becomes clear from its absence in archaeological explanatory efforts that focus on the larger flows: the role of intermediaries, the agents that made trade happen. This is a factor that features rather generically in most models (e.g. Abadie-Reynal 1989; Lewit 2011; Peacock 1982; Reynolds 1995) yet its impact on tableware distribution patterns is hard to tie down since they are the archaeologically least visible components of Roman trade. However, given the emphasis on understanding tableware trade as a multi-scalar phenomenon I believe the different roles of individuals active in tableware trade need to be evaluated (as well as those in tableware production and consumption, although that is not the focus of this study). The exact titles of these intermediaries is therefore of less importance in this case study. Rather, we should have a general idea of the types of interactions that could have taken place between agents in trade processes.

The titles and roles of merchants were variable, but two of the most common titles are those of the *negotiatores* and *mercatores*. From the Late Republic onwards, the title of *negotiator* was used to refer to individuals engaged in a diversity of activities: commerce, financial business, and the exploitation of agricultural estates. However, from the Early Imperial period onwards the title became more associated with individuals engaged in trade only (the meaning of the title then became more similar to that of the *mercator*). Individuals engaged in trading a wide range of goods could be referred to as *negotiatores,* but they typically perform their activities abroad and on a private basis (Broekaert 2013b, 15-17). They could be involved in overland or maritime transport of goods. The title of *mercator* refers to "basically anyone whose main profession is to organize the sale or resale of merchandize, with the intention to make a profit" (Broekaert 2013b, 151). The mobility of *mercatores* is considered their most defining feature: "A mercator is constantly travelling around, searching for the most profitable trading opportunities and transporting merchandize from one place to another" (Broekaert 2013b, 152) and "As their entrepreneurship and, more importantly, desire for profit forces them to inspect unknown regions for commercial opportunities, they are frequently mentioned as an ideal source for information" (Broekaert 2013b, 153). *Mercatores* could be engaged in both land and sea transport and could possess their own vessels, although they would rely on the transport services of others (*nautae* or *navicularii*) if this were not the case. Both *mercatores* and *negotiatores* could be specialised in trading particular types of goods (such as tableware) and some focused on specific trade routes.

The main difference between *mercatores* and *negotiatores* seems to be one of scale. Where the *mercator* would probably be mainly concerned with providing parts of a ship's cargo, the *negotiator* would charter an entire ship, although this differed through time (Peacock 1982, 158; Rougé 1966, 290). The *negotiator*'s activities were more international, capital intensive, and profitable than those of the *mercator* (Broekaert 2013b, 20). Although *negotiatores* were mobile, they nevertheless were often doing business in a shop or booth and can be considered to have been less mobile than *mercatores*. The nature of their mobility is well summarised by Wim Broekaert (2013b, 18-19): "Commercial activities evidently forced the majority of *negotiatores* to travel around, always searching for the most profitable combination of supply and demand. Literary texts abound with descriptions of *negotiatores* transporting their merchandize all over the Mediterranean and beyond. They are invariably present in the most isolated corners of the empire, risking their lives for profit and often being the sole connections between the 'civilized' empire and the 'barbarians' living on the edges. Yet, contrary to the semantics of *mercator,* mobility is not a central aspect of the *negotiator*'s nature. Various inscriptions show, especially in the city of Rome, how *negotiatores* were constantly doing business in a shop or booth, located near a famous statue, a temple etc."

However, trade cannot be considered without discussing transportation, despite the focus of the current case study on trade processes. Some merchants might have been shipowners whilst some shippers might have been active in trade. The types and roles of shippers are equally variable as those of traders. Rougé (1966, 244) suggests the *nauclerus* and *magister navis* might have been agents of shipowners and in charge of the sale of the ship's cargo. *navicularii* are considered financiers and owners in shipping operations (Peacock 1982, 158). Wim Broekaert (2013b, 221-222) disagrees with this distinction between these titles and argues *navicularius* and *nauclerus* can be more usefully considered as synonyms. Three meanings can be attributed to these titles: ship-owners, using ships for commercial purposes, or being a representative of a ship-owner. These shippers could have been engaged in commercial activities themselves, and therefore their role could overlap with that of *mercatores* and *negotiatores*, but the emphasis of their activity would always lie on transport. Although many shippers might also have been ship-owners, the evidence does not allow for assuming this was always the case (Broekaert 2013b, 217-220). The *navicularius* and *nauclerus* were mainly concerned with sea transport, while the different title of *nautae* refers to those concerned with "the organisation of transport on rivers, lakes and swamps" (Broekaert 2013b, 177). *Nautae* could also organise land transport themselves, rather than outsourcing it, and could be engaged in trade as *negotiatores* (Broekaert 2013b, 175-177).

A wide range of intermediaries can be considered to have been active in trade and transport, both inter-regionally or locally, often combining different roles, making small or large profits, and collectively producing the distribution patterns of tableware we are now confronted with. In this case study it is less important to know exactly which titles were associated with which roles. Rather, the existence of different roles intermediaries could take up, that these could be combined and, indeed, that tableware trade in the Roman East did not always take place directly between producer and consumer, is what matters and should influence the method developed to address this case study's research questions. Moreover, an additional complication is the difficulty in evaluating the contribution of arguably the largest group of merchants active in tableware trade: those who traded small amounts of pottery locally (in the place of production, redistributive centres or in markets close to the place of consumption). Wim Broekaert (2013b, 251-252) suggests that the title of *propolae* might be used to refer to those engaged in small-scale local distribution of merchandise. However, it is difficult to quantify their contribution to trade due to the limited mentioning of small-scale merchants in literary sources. As I will argue below, I believe the key to understanding their contribution to creating the tableware distribution patterns is an approach that takes into account connectivity and, more specifically, social networks connecting individuals over short and long distances. In the next few sections I will introduce two models of the ancient economy where such connectivity is a major driving force. For now, it suffices to conclude that intermediaries with different roles, active on different scales should be an integral part of the approach developed in this case study.

Bang's hypothesis: The Roman bazaar in a tributary empire

Bang (2008) argues that markets in Roman Imperial times functioned very differently from our current conception of large-scale, integrated entities where very informed specialists trade, facilitated by extensive and efficient communication networks. He draws on a comparative history approach, focusing mainly on the Indian Mughal empire, to explore the functioning of Roman trade in more detail. Bang suggests the concept of Bazaar-style markets, "distinguished by high uncertainty of information and relative unpredictability of supply and demand. This makes the prices of commodities in the bazaar fairly volatile. As a consequence, the integration of markets is often low and fragile; it is simply difficult for traders to obtain sufficiently reliable and stable information on which effectively to respond to developments in other markets. Considerable fragmentation of markets prevails" (Bang 2008, 4). This hypothesis sees the Roman market as a fragmentary system with low standardization, with traders making do to the best of their knowledge of the system, which on average is very limited. The agents braving this irregular trade landscape were faced with a variety of challenges. Due to variable consumer demands, producer supplies, environmental uncertainties and transport challenges, the market knew huge irregularities and low transparency. Agents' responses to these challenges were twofold: (1) instead of market integration, merchants would aim to benefit from opportunism and speculation; (2) a social network of personal trusted relations and communal ties was maintained that organized protection (both commercial and physical) and determined to a large extent the information available to the agent and their economic opportunities (Bang 2008, 200-201). This social network was not just a dense ball of random ties connecting individuals in towns either: "two key concepts of the social fabric of the bazaar were communal associations and the household" (Bang 2008, 241). More than that, the social network allowed for inter-regional trade to take place, through an integration of political and commercial spheres, as well as the specialization of intermediaries. However far away from home merchants roamed, the tendency to form and use communities structured around native identities is illustrative of the local emergence of social networks, often with a preference for native connections, but giving rise to supra-regional distributions of goods (Bang 2008, 249-250). The community structure of social networks served to protect community interests and opportunism whilst disadvantaging outsiders, thus reinforcing the fragmentation of the Roman market system. One can see that the key concepts in this model are traders as intermediaries and the social networks connecting them, concepts that are responsible for the shape and functioning of the Roman trade system.

Of course these are not just two factors working in isolation. Two other major components of Bang’s model are the tributary nature of the Roman Empire and the agrarianate nature of its societies. The former refers to the taxes provinces owe to the Roman imperial state. With the concept of an agrarianate society, on the other hand, Bang refers to the fact that “literary high-culture and cities co-existed with peasant agriculture in the pre-industrial world” (Bang 2008, 11). These two concepts draw attention to the importance of redistributive measures and state influence through tribute extraction and peasant surplus production. The increasing legal uniformity must no doubt have correlated markets to some extent (despite the huge uncertainties that came with the local variability of its implementation and enforcement), but it is unlikely they ever “broke the pattern of strong local variations in custom and practice” (Bang 2008, 188). Rather than seeing state influence and market trade as diametrically opposed spheres where one always needs to dominate the other, Bang argues that the dichotomy between state and market redistribution is a modernist construct. Although imperial surplus extraction is seen as a key stimulus of inter-regional economic flows, Bang also argues the state simply did not have the organizational capacity to structure these flows on its own or within a single integrated sphere: “markets were a necessary intermediary” (Bang 2008, 119-121). This dichotomy has also dominated discussions of the more physical aspects of roman trade: cargo, ship and route sizes, and itineraries. Big ships carrying a cargo destined to supply major cities and following popular routes between producing and consuming regions no doubt existed. Yet focusing on these bigger flows means the more casual flows of small amounts of goods within a region and their impact escape our attention, even though these must have been the norm. Large cities existed in the Roman East (e.g. Alexandria, Antioch, Ephesos) and their markets might well have acted as major redistributive centres, yet the vast majority of markets were small and maintaining close connections with these markets to obtain accurate consumer information must have been a formidable challenge. Indeed, logistical transport difficulties are intertwined with the problem of limited information through the importance of intermediaries. Accurate and up-to-date information as well as the right social contacts were a necessary requirement to make these large and very risky undertakings happen and worthwhile. Transport at all scales can therefore not be seen independent of considerations about intermediaries and their economic opportunities.

Critiques of Bang's hypothesis

To say that some scholars disagree with Bang's model would be an understatement. His work has attracted a very mixed reception, with some arguing Bang signals a return to or defence of the 'primitivist' paradigm in the study of the ancient economy (Shaw 2010). By far the most elaborate and extreme of commentaries on Bang's 'Roman Bazaar' model is Morris Silver's (2009). It is worth here to follow Silver's lead in identifying the key issues surrounding Bang's work. This will lead me to the conclusion that Bang' model is wrong, just like that of every other scholar of the ancient economy, including Silver's, but that the key aspects of each model are extreme hypotheses that need to be evaluated with non-anecdotal evidence.

Silver's main critique is that the key elements in Bang's model "do not necessarily approximate closely to the Roman facts" and that "Bang underestimates the integration of the Roman economy". He further develops this critique through a number of counter-arguments:

- "There is ample evidence of large, non-household-based enterprises in Rome. Bang presents no evidence indicating that larger firms were relatively less important in the 'Roman bazaar' than in early modern Europe, or that the size distribution of Roman firms more closely resembled that in the East" (Silver 2009, 423).
- Silver presents plentiful evidence for élite participation in the Roman economy, cashless transfers and the existence of deposit banks.
- "Roman commercial law undoubtedly had difficulties in terms of both content and degree of proactivity. However, Bang has not demonstrated that in either respect it resembled the eastern bazaar more than early modern Europe" (Silver 2009, 428).
- Bang argues that a capitalist economy is not a suitable model for the Roman economy, but Silver disagrees. He claims that Bang minimises the evidence that Rome "compared favourably with other pre-industrial regions", and that the arguments Bang raises do not preclude a classification as a market economy. Silver

> further argues that the impact of the *pax romana* implied a moderate to revolutionary reduction in transaction costs, and that the use of the concept of “peasant” economy is not an argument against capitalism (the concept of a “peasant” economy refers to the claim that the bulk of the labour force were peasants who were largely self-sufficient, reflecting a big chunk of produce that always stood outside the money economy (Bang 2007, 25)).

Many of Silver’s critiques are arguments that Bang’s model is not perfect since it attributes a minor role to some evidenced phenomena (e.g. participation of élites in commerce, cashless transfers, the existence of large firms). Possibly most important is the claim that Bang purposefully downplays any evidence that does not fit well within his model. However, I believe Bang probably does this in the belief that these phenomena can take place within his model but are not its defining features or driving forces. Although the claims that Bang misinterprets or excludes evidence might be true in some cases, the debate between Silver and Bang seems an inconclusive one where each author uses (sometimes anecdotal) evidence to emphasise different aspects of a complex past system, assuming (without testing these assumptions) that their model can account for the evidence mentioned by the other author. In some sense, then, the debate can be considered to miss the point since both authors have evidence to support their claims and neither is in the position to conclusively disprove key aspects of the other’s model. What is really needed is an approach based on hypothesis testing that is able to evaluate whether Bang’s claims can be substantiated: can the patterns or sources Bang quotes result from the processes he mentions? However, this is easier said than done. This statement seemingly ignores the limitations of the available evidence, which is the reason why Bang and Silver necessarily adopt their style of argumentation. The required approaches should never ignore the complexities of the hypotheses to be tested, and it is most likely that they will succeed in only testing a few of the most well-defined aspects of Bang’s model. The approach developed for this case study, the specific questions it allows one to address, and how this approach might help one to evaluate some of the claims made by Silver and Bang will be discussed below.

However, this elaboration on Silver’s criticisms begs the question what other models are out there, and in particular models based on the concept of a market economy. Before introducing the approach developed for this case study I will discuss an alternative to Bang’s model: Peter Temin’s *The Roman Market Economy* (2013).

Temin's hypothesis: The Roman market economy

In some ways, Silver's critique of Bang's work is not very helpful because no workable alternative model is proposed. Peter Temin's work on the Roman economy, however, can be considered an admirably specific model that reflects many of Silver's arguments. Temin agrees with Bang that government involvement in (wheat) trade must have been rather limited and that private enterprises must have dominated trade (Temin 2013, 32). The way both scholars present a model that is based on this assumption, however, could not be more different. Temin believes that Roman markets were integrated and strongly interconnected:

> "I argue that the economy of the early Roman Empire was primarily a market economy. The parts of this economy located far from each other were not tied together as tightly as markets often are today, but they still functioned as part of a comprehensive Mediterranean market. There are two reasons why this conclusion is important. First, it brings the description of the Roman economy as a whole into accord with the fragmentary evidence we have about individual market transactions. Second, this synthetic view provides a platform on which to investigate further questions about the origins and eventual demise of the Roman economy and about conditions for the formation and preservation of markets in general." (Temin 2013, 4).

Providing a workable research platform is as crucial to Temin's efforts as the hypotheses he derives from it. Temin argues that simple concepts from modern economics can offer useful insights into the ancient economy. For example, the concepts of supply and demand are considered tools for understanding price-setting mechanisms and exchanges of individual commodities or services; the *New Institutional Economics* (North 1981; 1990) is considered an approach that helps to evaluate the operation of markets by focusing on the role of institutions; the concept of comparative advantage is suggested as a way to understand the economic interactions of regions. In addition to these, a few other concepts are of particular relevance to this case study. Similarly to Bang, Temin states that ancient people are considered to have had access to far less information than people in the modern world, and that institutions were crucial in mediating information. However, Temin suggests more specifically defined concepts than Bang to explore this limitation: expensive information (there were costs involved in maintaining long-distance communication, costs which were reduced by being part of economic and social institutions) and asymmetric information ("when one party to a transaction knows more than the other" Temin 2013, 13, 112-113).

These concepts should be used in simple models that are necessarily abstractions of a complex reality, where a good model is distinguished from a bad model because it fits the available data better (Temin 2013, 5). In providing this research platform, Temin does something very few other scholars of the Roman economy do: his key statement (that the Roman economy was a well-functioning integrated market where prices are determined by supply and demand) is not taken for granted but is considered a null-hypothesis that needs to be tested (Temin 2013, 5-6). Temin even suggests two ways in which this hypothesis can be tested:

1. "we can infer from the existence of prices that market exchange more closely describes the interaction containing the prices than reciprocity or redistribution."
2. "people will behave instrumentally in market exchanges" so by looking at the incentives people have to continue their particular behaviour one can evaluate the existence of market exchange (Temin 2013, 8.)

Temin performs the first type of test for the late Republican and Early Imperial period wheat trade: "I use wheat prices … to test the proposition that many wheat markets across the Mediterranean were interconnected and interdependent" (Temin 2013, 29). A series of logical statements is proposed to identify what experiment is necessary in order to evaluate his proposition:

- Private enterprises are assumed to have dominated wheat trade
- "If there had been a unified wheat market, the main market would have been in Rome". This is where the largest supplies and demands came together and where the price for wheat would have been set (Temin 2013, 36).
- "Under these circumstances, wheat outside of Rome would be valued by what it was worth in Rome" (Temin 2013, 36).
- Therefore the price outside Rome equals the price in Rome minus the costs involved in transporting it to Rome
- One would therefore expect a correlation between the price of wheat at a certain market and the distance to Rome. Alternatively, in the absence of an integrated market there would be only local prices determined by local conditions. If markets were not integrated then we would not expect a relationship between location and price.

Temin collects six prices (which are argued to be all the known prices that can be used in this experiment) from different markets around the Mediterranean and with different dates. He performs a regression analysis which shows a strong correlation between as-the-crow-flies distance from Rome and price, seemingly confirming his hypothesis that the Roman wheat markets in the late Republic and early Empire were integrated (Temin 2013, 29-52). “the regressions confirm with very high probability that there was a unified wheat market that extended from one end to the other of the Mediterranean Sea. Transport costs were roughly proportional to distance, and the effects of distance were larger than the idiosyncratic influences of particular markets and places” (Temin 2013, 46).

Critiques of Temin’s hypothesis

There can be no doubt that the economic concepts used by Temin and his approach based on hypothesis testing are useful and extremely promising for the study of the ancient economy. However, as Temin himself acknowledges (2013, 48-52), there are some issues with his conclusions. The experiment itself is misleading because its results offer less proof to Temin’s hypothesis than he is willing to admit. From a descriptive statistics point of view it seems that there is a correlation between price and location. However, with only six data points there is no way that one can statistically prove or reject Temin's null-hypothesis. Temin acknowledges this point (2013, 49), and the absence of relevant data to support his hypothesis leads him to turn his attention to the larger dataset available for Hellenistic Babylon, even though this is not relevant temporally and spatially (Temin 2013, 53-69). Nevertheless, the correlation between price and distance from Rome is a key argument in Temin's defense of the hypothesis that the Roman wheat market was integrated. The limited data leaves this hypothesis unproven for the late Republic and early Empire, despite Temin's claims to the contrary.

Temin’s hypothesis disagrees with the findings of Paul Erdkamp (2005), who has argued against a strong integration of the wheat market and attributes a bigger role to government involvement in wheat trade. It was Erdkamp’s hypothesis Temin set out to disprove with his experiment. It is therefore no surprise that Erdkamp produced an extensive review of the strengths and shortcomings of Temin’s approach. Most crucially, Erdkamp argues that it is unnecessary to downplay state involvement in an argument for the existence of a market economy, since both state and private enterprises can function side-by-side. Temin interprets evidence for distribution and transportation as trade, which Erdkamp disagrees with. Moreover, it is argued that the annona is not the only non-market supply channel, Erdkamp mentions in particular the large-landowners who bring part of their produce to Rome to support their urban households and clientele: “the private market’s contribution to provisioning the empire’s capital was much smaller than Temin implies” (Erdkamp 2013, 2).

Both Erdkamp (2013, 7) and Temin agree that Temin’s models are abstract simplifications of a complex past phenomenon, and that the nature, quantity and critique of the available data largely determines their success. The available historical data are clearly insufficient to evaluate the integrated nature of the Roman market economy. The modelling approach is nevertheless promising, although not exclusively in the way Temin used it. His approach can be considered rather macroeconomic and deterministic in nature, where prices are considered to be in equilibrium and the particular interactions of individuals each with their own motives does not have a place. Temin constructed the model like this on purpose and it is clear that interesting insights can be gained from it. However, such a model will never succeed in capturing the interesting variation and distinctive traits of Roman society that fascinates scholars of the ancient economy. Erdkamp argues that what is needed to solve the debate surrounding the Roman economy is “a more balanced and nuanced approach, allowing for the shades of grey that characterize all historic reality” (Erdkamp 2013, 7). Maybe computational network modelling and drawing on the vast amounts of pottery data offer a way forward?

5.2.4. Being wrong can be useful: questions, methods, and issues for a new approach

I can now come back to my statement in the introduction to this case study: there is no lack of hypotheses explaining tableware distributions in the Roman East; most scholars agree the Roman trade system was a complex thing involving multiple factors; what is needed is an approach that allows one to evaluate the relative importance of these factors and how they are reflected in our datasets. This case study suggests such an approach. I believe that the very diverse arguments made by the archaeologists and historians mentioned above can be considered useful models of tableware distribution in the Roman East or of the Roman economy at large, despite many criticisms to the contrary. Each author places an emphasis on different key concepts, acknowledges the simplifying nature of their model, and argues for the complexity of the phenomenon under study. In this case study, I aim to suggest a novel (for Roman archaeology) approach to exploring the implications of different hypotheses suggested by authors and evaluate their potential contribution in giving rise to the observed tableware distributions. This approach therefore does not aim at being ‘right’, or at capturing the full complexity of such past phenomena. Instead, it tries to map out the grey-zone between these hypotheses; the kind of approach that has been curiously absent in the study of Roman tableware distributions and the Roman economy at large (Davies 2005; Morris et al. 2007). An outline of the research aims and process of this case study is provided in box 8.

To address the research questions formulated in the introduction I will combine a number of methods and techniques. Firstly, I will perform an exploratory analysis of the distributions evidenced in a large tableware dataset (section 5.4). This will allow me to critically evaluate the nature of the data and its many limitations. Secondly, similarity matrices of forms and wares will be produced and analysed using exploratory network analysis to address the first research question: the extent to which forms and wares show similarities in their distribution patterns (section 5.4 and appendix IV section 12.2). This will also highlight some features of the known distribution patterns that could suggest types of distribution mechanisms. The third aspect of my methodology will (initially) take a step away from the data and its limitations (section 5.5). It concerns a more abstract agent-based model representing individuals trading tableware in the Eastern Mediterranean, a formal representation of many of the hypotheses mentioned above. This simple model will be simulated under various different settings to explore the role of different combinations of factors introduced above. The aim of this last method is to evaluate to what extent current hypotheses surrounding trade in the ancient world could have given rise to the observed distribution patterns. The latter step concerns a confrontation of the computational simulation output (i.e. hypothetical distributions of goods) with the observed tableware distributions (i.e. the ICRATES dataset).

But what place do models of the Roman economy have in an archaeological network science PhD project? Their importance should be clear by now: our archaeological research questions reflect our interest in certain past phenomena, and models are necessary to abstract these phenomena in terms of network concepts and data. It could be argued that Bang's hypothesis offers a suitable framework or starting point for exploring differences in Roman tableware distributions: specialized intermediaries, social networks, a tributary empire, and agrarianate societies. These concepts combined offer a model for exchange that works from the bottom-up through individuals in interaction, yet does not disregard the importance of large-scale state influence and the kinds of trade systems possible in agrarianate societies. Yet this does not necessarily mean Bang's hypothesis is useful for exploring the different factors responsible for tableware distribution introduced above. Bang himself argues quite explicitly for his model as a possible explanation for differences in tableware distribution patterns.

Tableware was found in large quantities and knew a wide distribution, yet some wares, forms or stamps travelled significantly farther and are much more common than others. This is interpreted as pointing to the existence and importance of different networks for the distribution of different potters' or production centres' produce (Bang 2008, 288; although see also Willet's (2012, 443-449) argument on the limitations of Eastern wares for performing this kind of analysis due to the general absence of informative stamps). Rather than taking this hypothesis at face value, this case study aims to evaluate it. I believe it allows for many of the factors argued for by archaeologists to take place and have an impact on multiple geographical scales without being deterministic. It is not a model of a world in chaos, rather it is a concept that unites tableware trade in the Eastern Mediterranean with that of luxury goods from India and staple goods from Southern Spain or Africa, and can accommodate for both small-scale trade at local markets as well as large-scale long-distance distributions of agrarian products for state or military redistribution (Bang 2008, 304-305). But only a formal implementation of this hypothesis as a computational model can teach us what its implications are and what kinds of patterns we can expect if we believe the hypothesis, and only a confrontation with our collected datasets can lead to an evaluation of this hypothesis.

A similar role in this case study can be attributed to Temin's model: it suggests macroeconomic mechanisms that are considered to govern Roman trade and provides economic concepts through which these can be explored. What I will not adopt here, however, is Temin's deterministic modelling approach. Instead, by adding a measure of stochasticity and allowing individual agents to be the driving forces behind the price-setting mechanism, I believe the concepts introduced by Temin can also serve as a framework for testing archaeological hypotheses. As should be clear from the critical discussion above, Temin and Bang's work is not considered mutually exclusive: they still agree on the limited availability of information, the importance of markets and limited state involvement. The grey-zone between these two models themselves will therefore also need to be explored.

The next section will introduce the ICRATES dataset and our current understanding of the tableware distribution pattern in the Roman East. The two sections that follow this describe the methods employed in this approach and their results.

5.3. Tableware distribution patterns: data and the state of the art

5.3.1. Tableware fabric-groups in the Roman East

This case study is concerned with the distribution of tablewares in the Eastern Mediterranean. A wide range of wares must have existed and three levels of production can be usefully considered to classify wares (Bes 2007, 202; see also Peacock 1982). Firstly, there are the wares produced either in households or on estates with a purely local distribution; secondly, wares produced in professional workshops that knew a regional distribution; and thirdly, wares produced in large manufactories which were of supra-regional importance. The underlying mechanisms of exchange leading up to the observed distributions might have been very different for these three levels of production, yet they must have been strongly intertwined. Wares of all types and distributions therefore need to be studied together if we wish to understand the processes driving their distribution. However, our knowledge of the first two levels is rather limited, since archaeologists have traditionally focused on imported wares which offer a number of advantages: they were better studied, are indicative of inter-regional interactions and can act as chronological markers. Often large proportions of site assemblages all over the Eastern Mediterranean are made up of locally-produced wares or wares of regional importance (a number of wares that fall into this category are mentioned by Bes 2007, 202). It is believed that such wares which provided for the demand of a settlement and its hinterland or for a small region must have been a common phenomenon (Bes 2007, 109). However, our knowledge of these wares is limited and our dataset is therefore not representative for these levels of production. These wares will not feature prominently in the formal analyses of this case study, since here I am mainly concerned in the different distributions of wares of regional and supra-regional importance. I believe future work which focuses more on differences in production processes would be better placed to study such wares.

The exploratory data analysis will focus on wares produced in large manufactories, with a distribution of either regional (ESB, ESC, ESD) or of supra-regional (ESA) importance. The wares this case study is concerned with are four major sigillatas produced in the Eastern Mediterranean. The term ‘Eastern Sigillata’ was coined by Kathleen Kenyon through the publications of the 1931-1933 Samaria-Sebaste excavations (Crowfoot et al. 1957; Kenyon 1957). A brief description of these wares will have to suffice for this case study, more in depth discussions of each ware can be found in Bes (2007) and Hayes (1972; 1980; 1985; 2008).

Table 15: typo-chronological references and (possible) region of production for major eastern tablewares.

Ware	Abbreviation	Typo-chronological reference	Region of production
Eastern Sigillata A	ESA	Hayes 1985	Coast between Tarsos (TUR) and Latakia (SYR)
Eastern Sigillata B	ESB	Hayes 1985	Maeander Valley in western Asia Minor (TUR). Possibly Aydin (ancient Tralleis)
Eastern Sigillata C	ESC	Meyer-Schlichtmann 1988 and Hayes 1972, 1985	Pergamon and surrounding region
Eastern Sigillata D	ESD	Hayes 1985	Cyprus (probably the western part)

Eastern Sigillata A (ESA)

Although no ESA kilns have been found yet, archaeometrical analyses and the distribution of early, late and rare forms suggest the ware must have been produced “somewhere in the coastal area between Latakia in Syria and Tarsus in Turkey” (Schneider 1995, 415-416). The core area of ESA distribution consisted of the Levant, Cilicia and (Eastern) Cyprus, although it was distributed all over the Central and Eastern Mediterranean (with sporadic finds further west, see references in Hayes 2008, 18). The earliest forms appeared in the mid-second century BC and had distinctly Hellenistic shapes and decorations. The morphology of ESA was subsequently influenced by that of Italian Sigillata (ITS) with the appearance of ITS around 40-30BC. ESA was distributed most widely between 125BC and AD20, and disappeared (or continued on a much reduced scale) at the end of the second century AD. The standard typology used here and in the ICRATES database is that by Hayes (1985).

Eastern Sigillata B (ESB)

ESB was produced in the Maeander Valley in western Asia Minor, possibly in Aydin (ancient Tralleis). A number of potters’ names on ESB stamps shared with Arretine produced ITS suggests Italian potters were involved in establishing the ESB manufactories, or are at the very least indicative of deliberate copying (Hayes 2008, 31). ESB also shows a strong morphological influence from ITS. An early version of the ware called ESBI appeared around 25BC but was superseded by ESBII around AD40, which introduced a new range of shapes influenced by metal vessels. In this case study, ESBI, ESBII and the transitional ESBI-II are referred to collectively as ESB. The standard typology used here and in the ICRATES database is that by Hayes (1985), who attributes a date of no later than AD150 for the last ESB forms. The core area of ESB distribution is southern Asia Minor, although it is quite common throughout the Aegean.

Eastern Sigillata C (ESC)

ESC was produced in Pergamon and its surrounding region including the kiln sites of modern Çandarli (ancient Pitane). One can distinguish between two wares: one produced at Pergamon and the early material at Çandarli, and the second only produced at Çandarli (Hayes 2008, 49). Both wares will be referred to as ESC in this case study. The core area of ESC distribution is Northern Asia Minor and the Northern Aegean. The standard typology used here and in the ICRATES database is that by Meyer-Schlichtmann (1988) and Hayes (1972; 1985). ESC first emerged around the late second century BC and continued to be produced until the end of the third century AD.

Eastern Sigillata D (ESD)

No ESD kiln sites have been found yet, but a provenance in (Western) Cyprus is most likely. The core area of ESD distribution is Cyprus and the south coast of modern Turkey, and at times it is present on the Levantine coast in large quantities and diversity. The standard typology used here and in the ICRATES database is that by Hayes (1985). ESD was likely produced between the end of the second century BC and the second half of the second century AD.

5.3.2. The ICRATES database

The ICRATES project (‘*Inventory of Crafts and Trade in the Roman East*’) aims to come to a better understanding of mechanisms of production and exchange in the Roman East between the second century BC and the seventh century AD through the study of multiple material data types.[24] Two PhD projects have been completed under the umbrella of this project (Bes 2007; Willet 2012). The work by Bes in particular, as well as that of many other project members has led to the creation of a database of over 33,000 individually recorded tableware

[24] The ICRATES project was supported by the Belgian Programme on Interuniversity Poles of Attraction (IAP 07/09), the Research Fund of the University of Leuven (BOF-GOA 13/04), Project G.0562.11 of the Research Council-Flanders (FWO) and the Hercules Foundation (AKUL/09/16).; http://icrates.arts.kuleuven.be/Icrates/Default.aspx (accessed 29-08-2013).

sherds, a great resource from which the dataset for the current case study is extracted. The ICRATES database is compiled mainly from published sources, supplemented with unpublished data from the Boeotia Survey in Greece and from the site of Amata in Jordan. Each entry in the database concerns a single tableware sherd, the site it was found on, the publication in which it was mentioned, the standardized fabric and form attributed by the authors (sometimes corrected or standardized by project members), and more information if this is available (see Bes 2007).

I decided to use the same version of the ICRATES database as was used for a previous exploratory network analysis (Brughmans and Poblome *in press*).[25] The last modifications to this version of the database were made in August 2010 and the total number of records or individual pieces of tableware is 28,501.[26] The total number of sherds in the database of the major wares discussed in this case study are: ESA 7,458; ESB 2,100; ESC 2,601; ESD 772.[27] Due to some sherds of unknown forms, and forms with undefined standard chronologies or chronologies falling outside the chronological limits of this case study, the number of sherds actually used in this analysis is further reduced to: ESA 4,543; ESB 1,032; ESC 1,896; ESD 602. An exploratory data analysis of the dataset used is presented in section 5.4.

5.3.3. Issues with the data

The limitations of the ICRATES dataset have already been discussed in depth by Bes (2007, 10, 100-101) and Willet (2012, 43-58). It suffices here to list some of the key issues and how these will be addressed in this case study:

- Although the ICRATES dataset is large by archaeological standards, it is still a sample and includes only a fraction of the total amount of tableware that circulated in the past (Abadie-Reynal 1989, 143). An eye-opener is Willet's (2012, 43-49) hypothetical calculation of a total minimum number of 75 million tableware pieces ever to have circulated in the Roman East and a total maximum of 33 billion.

[25] For the results and networks of this analysis see http://icrates.arts.kuleuven.be/icrates/network-analysis/webpages/icrates_mainframe.html (accessed 01-06-2014).

[26] Please note this is an earlier version than that used by Willet (2012) which includes 33,587 records. The decision to use the older version was motivated by the ability to compare the different network methods used in Brughmans and Poblome (*in press*) with the methods used here on a single dataset

[27] These counts are slightly higher in the later version of the database used by Willet (2012): ESA 7,649; ESB 2,195; ESC 1,705; ESD 798.

- Some areas and sites around the Eastern Mediterranean have seen more archaeological activity than others, leading to an overrepresentation of these in the dataset. Moreover, the published evidence is skewed towards urban settlements whilst rural settlements are underrepresented.
- The organisation and place of production of many wares is poorly understood. Most problematic for this study is our limited knowledge of the ESA production process and centre(s), as well as an almost complete ignorance of the many wares with a very limited distribution in the direct vicinity of the production centre.
- On the other hand, for some wares a much higher number of sites around the production centre are included. This issue is most crucial for the suggested production area of ESA, for which a very large number of sites with often incredibly high numbers of ESA sherds are included in the database.
- The quality and method of quantification differs per site and ceramologist. They are the results of the author's or excavator's objectives. Discarding non-diagnostic sherds was quite common until the first half of the 20th century. Not all publications mention whether the published sample is representative of the material excavated. Some forms or wares might be misidentified by the authors. The latter issue was corrected as much as possible by those entering data into the ICRATES database and estimated by Bes (2007, 10-11) to be no more than 1-2% of the total database.

These issues should be acknowledged and one should find a productive way of addressing them in one's method. For even with all its flaws, to use the data we have is still to be preferred to ignoring it completely. I will confront these issues in a variety of ways, both quantitative and qualitative. These approaches do not aim at completely overcoming the limitations of the collected data, but rather at understanding the nature of the limitations better and incorporating these findings in the development of a critical quantitative method.

- This case study will only concern the largest possible patterns, not small particular changes. It is believed that collecting a large dataset such as the ICRATES database allows one to make statements of large-scale patterns: "certain basic patterns are undeniably present in the data" (Bes 2007, 10). Bes argues that by collecting the data *en masse* we can still overcome some of these issues (Bes 2007, 10, 100).

- Rather than considering the actual amounts of sherds published per site as representative, I will focus in the exploratory data analysis on the diversity of pottery forms attested (or not) at sites.
- The proportion of types per ware in site assemblages will be the focus of analysis, rather than the proportion of sherds published.
- The dataset is only considered representative for the major wares distributed on a regional (e.g. ESB, ESC, ESD) or supra-regional (e.g. ESA) level. It is not representative for the many locally-produced wares that undoubtedly make up big chunks of site assemblages, and these will therefore not be included in the quantitative analyses.
- The exploratory network analysis includes a sensitivity analysis where the impact on the quantitative results of focusing the analysis on the most robust patterns will be explicitly evaluated.
- The overrepresentation of urban settlements should be taken into account when suggesting a realistic estimate for the demand distribution in the ABM, since this is dependent on the population distribution. I therefore consider the dataset representative for the urban centres with a high population and assume an exponential distribution of population size in the ABM, which reflects the idea that there are a low number of settlements with a high population and a large number of settlements (largely not included in the dataset) with a low population.
- The results from the quantitative analyses will be recontextualised within their socio-political and archaeological frameworks in the discussion in section 5.6.

5.3.4. Chronological overview tableware distribution in the Roman East

Bes (2007) uses a chronological framework describing major developments in the main Eastern wares, more or less dictated by the collected evidence in the ICRATES dataset. The first four of these are particularly relevant for this study:

Table 16. First four phases of the chronological framework used by Bes (2007, 12).

Phase	Lower Date	Upper Date	Criteria
1	150BC	30BC	Initiation and development of ESA, ESC and ESD
2	30BC	25/30AD	The inception of ITS and ESB, typological change in the repertoire of mainly ESA
3	30AD	60/70AD	Contraction distribution of ESA, flourishing of ITS, development of ESB and ESD
4	70/75AD	200AD	Contracting distribution of ITS, disappearance of ESA, ESB and ESD, first ARSW

Phase 1 (150-30BC, the pre-Augustan period) is marked by the inception of ESA and its distribution on a supra-regional scale. ESA was the first of the eastern fully red-slipped tablewares to have a large-scale production. The earliest ESA forms (Bes' phase 1a, c. pre-75BC) have a more limited geographical distribution compared to those in the later part of this phase. These early forms are attested as far away as Greece, although they are only present in larger quantities on the Levantine coast and Cyprus. Bes (2007, 103-104) mentions a number of possible reasons for the relatively limited distribution of these early ESA forms: (a) ESA circulated predominantly in the Seleucid kingdom and less so in the Ptolemaic kingdom; (b) possible economic interactions with the west and the Roman republic might have played a role, although the effect of this is difficult to evaluate; (c) ESA production may have been limited or aimed at a certain quantity, which was in part kept low by the production and local or regional consumption of other tablewares in many parts of the Eastern Mediterranean; (d) the early ESA forms were largely of Hellenistic style and were probably not considered a novelty by consumers. The first ESC and ESD forms also emerged in phase 1a but only in more limited quantities. The distribution of ESC was largely restricted to the west coast of Asia Minor, whilst ESD was mainly present on Cyprus. In the later part of this phase (Bes' phase 1b, c. 125/100–30/25BC) ESA becomes more common throughout the whole Eastern Mediterranean. A number of factors might have contributed to this increased distribution (Bes 2007, 105-109): (a) the innovative character of the new ESA forms; (b) some ESA forms have a particularly wide distribution (forms 3, 4A and 22A) and it has been argued that they may have been sold as sets, which might imply that merchants purposefully decided to include those sets in their cargoes which they knew they could sell widely (Bes 2007, 108); (c) the growing influence and involvement of the Romans on the

Eastern Mediterranean stage and the decreasing power of the Seleucids; (d) thanks to the decrease in piracy, the creation of the provinces of Syria and Cyprus and the end of the Mithridatic wars, the Eastern Mediterranean might have become more peaceful, stimulating supra-regional trade. In phase 1b ESC knew a marginal growth in the Aegean, ESD on the other hand remained largely limited to Cyprus. Both ESC and ESD were of regional importance throughout this phase, whilst ESA already knew a supra-regional distribution. It has been suggested that the manufacturers of ESC did not use the establishment of the *provincia* of Asia to their benefit (Poblome et al. 2007, 224), and neither did the ESD manufacturers benefit from the creation of the *provincia* of Cyprus (Lund 2002, 206). Moreover, Cyprus was firmly in the core distribution area of ESA: the locally produced ESD is only the second ware in most Cypriot site assemblages. Bes emphasises the importance of possible redistributive centres such as Rhodes and (after 167BC) Delos. These two centres might have formed pivots for trade between east and west. Alexandria on the other hand might have served as a redistributive centre in the east and even as a commercial pulling force, given the important export of Egyptian grain.

Phase 2 (c. 30BC-25/30AD, the Augustan-Tiberian period) is marked by the introduction of ITS and ESB. The earliest ITS forms are mainly found in larger urban centres, with harbours or direct access to the Mediterranean, and some of which played a role in provincial administration. Bes (2007, 111-112) therefore argues that early ITS might have travelled mainly along the major lines of communication between east and west. Corinth might have played an important role as a redistribution centre, from which ITS reached other eastern sites. Throughout this phase ITS becomes the main ware at those sites where it arrived earliest. Moreover, ITS had a strong morphological influence on ESA and ESB (as indicated by ‘italicized’ forms and the use of stamps, which is considered a western phenomenon), but less so on the other two eastern wares, ESC and ESD. Indeed, it is argued that Italian potters or owners of manufactories were involved in the inception of ESB (Bes 2007, 85, 109; Zabehlicky-Scheffenegger 1995, 1996), which in this phase knew a rather limited distribution focused on the Aegean. ESA continued its wide distribution in the Eastern Mediterranean, and the quantity and distribution of its production was unmatched by the other contemporary eastern tablewares. The distribution and quantity of ESC increased, perhaps related to the start of production at coastal Çandarli (Loeschke 1912), dominating the coastal area around Pergamon. Its presence beyond the Aegean remains small. ESD continued its expansion on Cyprus, as well as on a more modest scale on the Levantine coast neighbouring Cyprus and on some sites in the Southern Aegean. The increased distribution led to increasingly diverse site assemblages of wares and forms, although each ware has a clear core area of distribution. All of this took place at a time when the Mediterranean and the lands beyond were for the first time part of a single political, administrative and military system, with a tax system applied throughout the area. These systems drew on resources from all over the empire, a fact that might well have stimulated inter-regional exchange. The capital of Rome in particular started to act as a pulling-force of resources. The tableware distributions do evidence of an increase in geographical scope and quantity in the east, an increase in exchange with the west, and a more uniform shape of terra sigillata during this period, which has been termed by some as the “tableware boom” (Poblome et al. 2001, 144; Poblome and Zelle 2002, 277). The role of Alexandria as a redistributive centre in the east has often been emphasised, due to its role in supplying the west with Egyptian grain in the framework of the *annona*. Yet Bes (2007, 115) rightly points out that this has led to a focus on official exchange only, whilst “A category of tableware such as ESA must also have been traded within strictly commercial patterns”. Other centres, such as Corinth, Knossos, Gortyn and Ephesos must also have played an important part as redistributive centres in the east and connecting exchanges between east and west.

Phase 3 (c. 30AD-60/70AD, the Claudian-Neronian period) signals a decrease in the quantity as well as geographical spread of ESA distribution. Although ESA becomes less common throughout the Eastern Mediterranean, its decrease is most striking in the Aegean. ESB distribution increased substantially during this period, especially on sites in the Aegean, Asia Minor and in Paphos on Cyprus, making ESB a ware of supra-regional importance during this phase. ESC is attested at fewer (especially Middle-Eastern) sites in this phase, although its presence in the area around Pergamon and the west coast of Asia Minor remained strong. ESD continued its geographically restricted distribution although it now takes the lead from ESA as the most common ware in Western Cyprus. It also makes up a large proportion of site assemblages in Cyprus as a whole, the Southern Levantine coastal region and Cilicia. ITS continued its wide geographical distribution but quantities generally decrease, although ITS knew a sharp increase in the Southern Aegean and Cyprus compared to phase 2. When considering the total number of tablewares attested for this period we notice a decrease, which might suggest the intensity of tableware distribution shrank. Although ESD and ESB increased in importance, they did not completely fill the holes left by the decrease of ITS and ESA. Bes (2007, 116-117) suggests that locally produced wares might have gained importance compared to imported wares during this phase and the next (as evidenced by the enormous proportions of locally produced wares in the assemblages of Corinth, Knossos and Tanagra, Bes 2007, 116-117). However, this hypothesis and its implications (e.g. a decrease of exchange between east and west, the impoverishment of regional economies, or a change in taste or fashion?) are difficult to test due to our poor knowledge of wares of local importance. The collected tableware evidence also suggests a continuation of the wider exchange patterns between east and west as discussed for phase 2. The large quantities of ITS in the Aegean, Crete and Berenice still indicate a strong link with Italy. However, the more mixed assemblages in Cyprus and the Southern Levant might suggest a stronger focus on exchange routes connecting Alexandria via the south of Cyprus, rather than the north along the south coast of Asia Minor (Bes 2007, 117).

During phase 4 (c. 70/75AD-200AD) most wares maintained a wide distribution. ESA remained the major ware at sites in the Levant and Eastern Cyprus, although its presence in the Aegean and Africa decreased significantly. The presence of ESA at sites with a military character along the Euphrates also suggests ESA moved farther inland during this period. The distribution of ESA forms dated after 100AD in particular is limited to the Levantine coast, which was the core area of ESA production and consumption. The distribution of ESB in the later first century AD in the Eastern Aegean decreased whilst elsewhere it largely retained its distribution of phase 3. ESB therefore remained of regional importance in the Aegean, but was far less common beyond. The distribution of ESC expanded significantly in this phase, retaining its regional importance in the area around Pergamon, and being more prominently present at sites everywhere in the Aegean, Crete, and Asia Minor. However, only small amounts of ESC are attested farther east of the Aegean. Bes (2007, 97) suggests that the co-presence of ESB and ESC on some sites in this phase might be indicative of both wares being distributed together. ESD maintains a similar distribution to that of phase 3, with a strong focus on Cyprus, Cilicia and the Southern Levant. The presence of ESA on Cyprus and of ESD on the Levantine coast suggests continued direct exchange between the two areas. The distribution of ITS decreased significantly, focusing on a few regions: the South-western Aegean, Crete and the Cyrenaica. As was argued for phase 3, the evidence seems to suggest that ESB, ESC and ESD filled the gaps left by ESA and ITS only partly (Bes 2007, 118). ITS stamps point to two major commercial routes: a first route linking the Cyrenaica with Italy (as indicated by Pisan ITS stamps), and a second route connecting Italy with Corinth, which continued to act as a redistributive centre supplying the rest of the Aegean (as indicated by central Italian ITS stamps). Based on the quantities of ESB and ESC that reached the Cyrenaica, Bes (2007, 119) further suggests a possible route between the Cyrenaica and Knossos. It is likely that Alexandria continued its redistributive role, forming a link between Italy and the east in the wake of the grain supply to the west and the opening of stone quarries in this period (*Mons Porphyrites* and *Mons Claudianus*). However, the tableware evidence in Egypt (showing a limited presence of imported wares in site assemblages) does not allow to conclusively confirm this hypothesis, and Bes (2007, 120) argues it is likely that locally-produced wares played an important role in supplying for Egyptian households in this period. The end of phase 4 saw the disappearance of the supra-regional distribution of ESA, ESB and ESD. These wares either stopped being produced or their scale of production was very much reduced. After the second century AD ESC continued being widely distributed, alongside the Tunisia-produced ARSW that first appears in this phase in modest quantities and at a limited number of sites.

5.4. Exploratory data analysis

5.4.1. Introduction

Before proceeding with the network methods I first need to obtain a better understanding of the structure of the ICRATES dataset, its limitations and how the distribution pattern that is the focus of this case study is reflected in this dataset. This will be done through an exploratory data analysis, the results of which will inform methodological decisions when designing the network approach and will also prove invaluable in comparing this archaeological dataset with the simulated outputs of the ABM. I will first address the issue of classification systems and argue for the use of wares, forms and functional categories as useful complementary analytical constructs in this case study. This will be followed by a description of the method used to divide the combined pottery dataset into 25-year time-slices. This manipulated data will then be described by focusing on the main chronological trends and exploring different types of frequency distribution. This will be followed by the results of an exploratory network analysis of the dataset, which is presented in full detail in appendix IV section 12.2.

5.4.2. Typologies and functional categories

First of all, it should be clear that in this case study I consider the tableware forms derived from the standard typologies (discussed per ware in the previous section) as modern constructs. They are conventions that were constructed by individual ceramicists or groups of scholars, and reflect their decisions and assumptions, but mainly their need to reduce a mass of information into analytically useful categories. The patterns emerging from the distributions of these forms will therefore, at least in part, reflect such academic decisions and assumptions (as revealed in previous network analyses of the ICRATES dataset, see Brughmans 2010; Brughmans and Poblome *in press*). Tableware forms are not considered to represent categories that were considered as distinct by people in the past.

This reveals an issue familiar to any archaeologist: if in our analyses we merely reveal patterns imposed by our classification systems, patterns of distinctions we in the present consider important or useful, how then will we ever reveal the differences that were perceived by people in the past? This issue can never be overcome. But what we can do is come to a better understanding of the impact of our classification systems on the creation of the patterns through our analyses, by analysing the same dataset using different classification systems representing different archaeological assumptions. In this case study I will use two different classification systems for grouping tableware together and exploring the resulting patterns. At the lowest level, or at least the most extensive system, I will group sherds together based on the standard typologies (i.e. tableware forms); a second grouping concerns all sherds of the same ware. A reassessment of the standard typologies is not a feasible task for this case study. A third classification was planned to be included in this case study, which groups forms together based on the shape (e.g. plate, bowl, cup) and possible function (e.g. beverage, food, serving) attributed to them by the authors of the publications in which they were described. However, this was not completed due to time limitations and will be included in future work.

5.4.3. Gaussian distribution

In order to explore the changing distribution patterns over time of tableware forms with different chronological ranges, the dataset will need to be chopped up into different time-slices. There are different ways of doing this, discussed in detail by Willet (2012) based on the method first suggested by Fentress and Perkins (1988). For this case study I decided to assume a Gaussian (or normal) distribution for the popularity and circulation of tableware forms, applied as described by Willet (2012, 35-38). This assumption entails that the majority of sherds of a certain form are more likely to have circulated around the middle of the typo-chronological date attributed to it, rather than around the earliest or latest dates. This is a theoretical assumption and a probability distribution I consider likely. However, in most cases the chronological accuracy needed to identify the popularity and circulation of Eastern forms or wares is simply not available, making uniform or skewed probability distributions equally probable (exceptions are those cases where certain wares or forms are particularly common or rare in well-dated contexts). A similar assumption about pottery probability distribution was made in the exploratory network analysis by Barbara Mills and colleagues (Mills et al. 2013; Roberts et al. 2012). Gaussian probabilities for each 25-year time-slice were calculated per form and the observed number of sherds per form was multiplied by this probability for each time-slice (Fig. 39). This results in a 'probable' number of sherds per form per 25-year time-slice, allowing for differences in distribution patterns per form (and per ware) to be explored through time (Fig. 40). It is important to note that the distribution obtained with this method and dataset reproduces the general trends observed by other studies using different versions of the ICRATES dataset and different methods of generating time-slices (Bes 2007; Brughmans and Poblome *in press*; Willet 2012). In previous network analyses of the same dataset we assumed a uniform distribution (Brughmans and Poblome *in press*), and it is therefore of methodological interest in this case study to compare the similarities and differences in particular patterns generated by these two approaches.

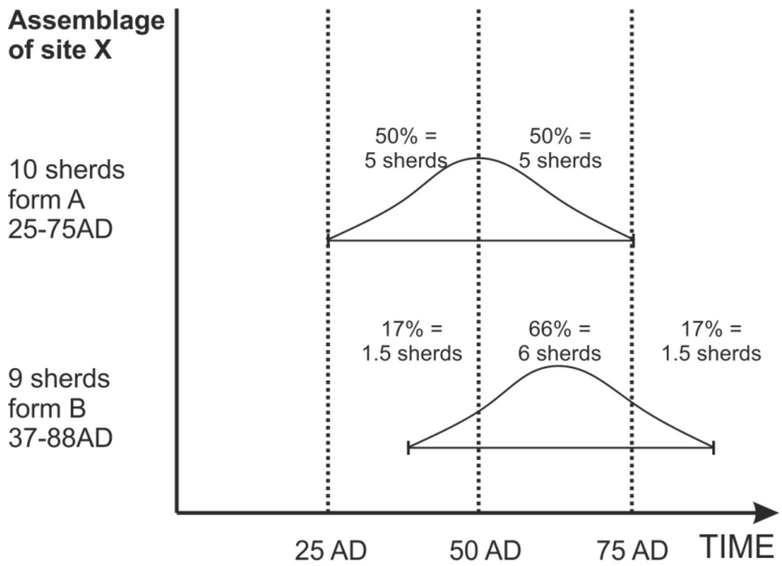


Fig. 39. Schematic example of the gausian probability distribution method.

5.4.4. Number of sherds per ware

By chopping up the collected tableware evidence into 25-year periods it becomes possible to identify some general patterns. The chart displaying the number of sherds per period (Fig. 40) shows that the dataset includes a very low number of sherds at the chronological boundaries of this case study, which is to be expected since the lower boundary reflects the first production of the Eastern sigillatas under study and the upper boundary marks the transition to the increasing popularity of other wares. One also notices that the bulk of the dataset until at least 75AD consists of ESA sherds. After that, ESB becomes slightly more common than ESA, whilst ESC sees a slight increase in the period after 125AD. However, this chart shows us only part of the available information. The higher quantity of ESA alone does not indicate a wide distribution since it could be concentrated on a limited number of sites, or many sites in only particular regions of the Eastern Mediterranean, or it can be the result of the typology used (if a ware has more forms in its typology then the process of only publishing diagnostic sherds will lead to an over-emphasis on this ware). Exploring the trends in the number of forms, the number of sites and the frequency distributions below will therefore be necessary to pin down the particularity of the ESA distribution.

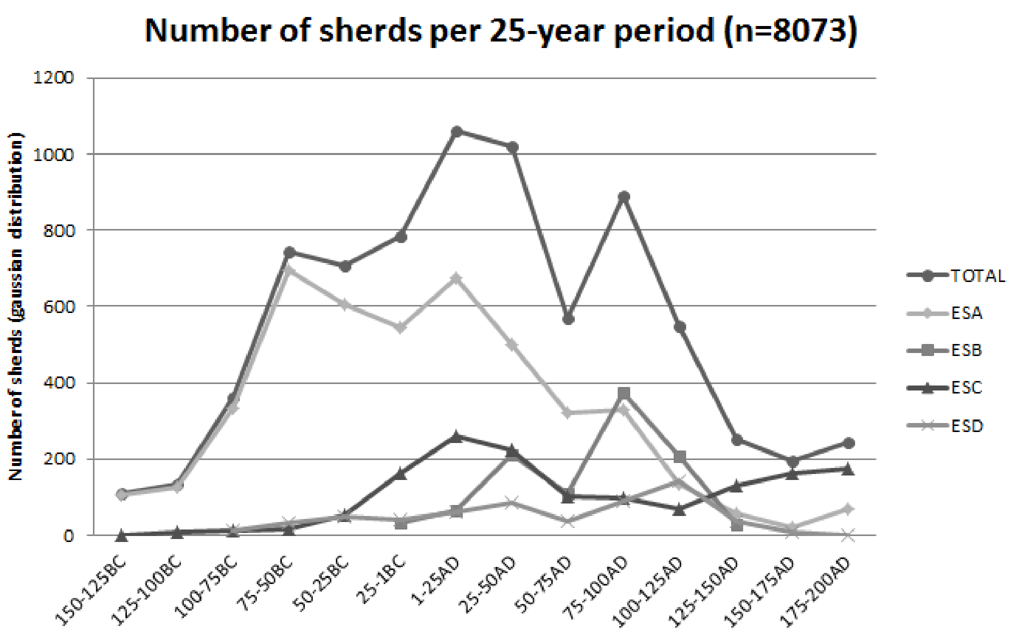


Fig. 40. Number of sherds per ware, per 25-year period, calculated using the Gaussian distribution method (n=8073).

5.4.5. Number of forms per ware

Figure 41 represents the number of forms in each wares' standard typology that is dated to a particular 25-year period. Again, we notice the highest diversity of forms between 50BC and 150AD, which is to be expected since these wares were most common in this period. However, it is interesting to note that ESA does not have the highest number of forms throughout all periods. The diversity of ESC is larger from 75BC onwards, even though its distribution pattern is much more limited than that of ESA.

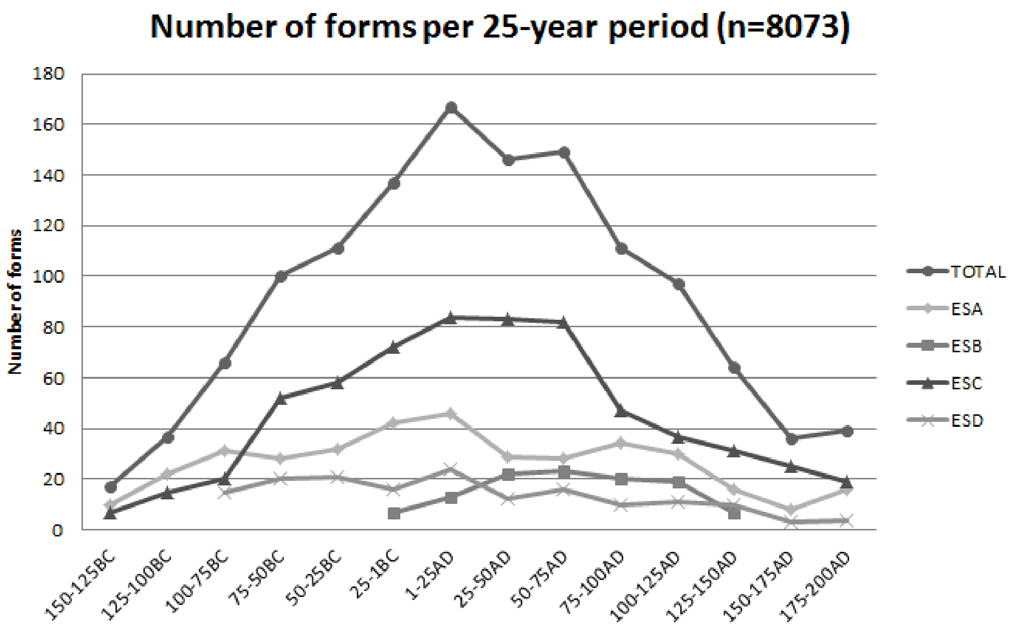


Fig. 41. Number of forms per ware, per 25-year period (n=8073).

5.4.6. Number of sites per ware[28]

The pattern under investigation in this case study is most obvious when we plot a chart of the number of sites per period per ware (Fig. 42). The boxplot in figure 43 presents an alternative representation of this pattern, which will be used to compare the simulated distribution of wares of the agent-based model presented in section 5.5 with the distribution observed in the ICRATES database. This figure presents the same data but normalised to a scale of 100 sites, in order to facilitate comparison with the simulations which include only 100 sites. It also only includes those periods where four wares circulated in the Eastern Mediterranean, which is the scenario simulated in the model.

It is immediately clear that ESA has by far the widest distribution until at least 75AD. After that its distribution becomes more limited like that of the other wares. Between 100 and 150AD ESD has a wider distribution than ESA, but not on such a scale as before. It is clear that only ESA attained a supra-regional distribution for centuries whilst all other wares were mainly of regional importance (although we should not forget the significant distribution of ESB and ESD in the period 50-125AD). It is also worth noting that there must have been an

[28] For maps showing the spatial distribution of the dataset per period and per ware, see: http://icrates.arts.kuleuven.be/icrates/network-analysis/webpages/icrates_maps.html (accessed 01/06/2014).

even wider diversity of wares, mainly of local importance, than those represented here. In light of this fact, the supra-regional distribution of ESA becomes even more striking.

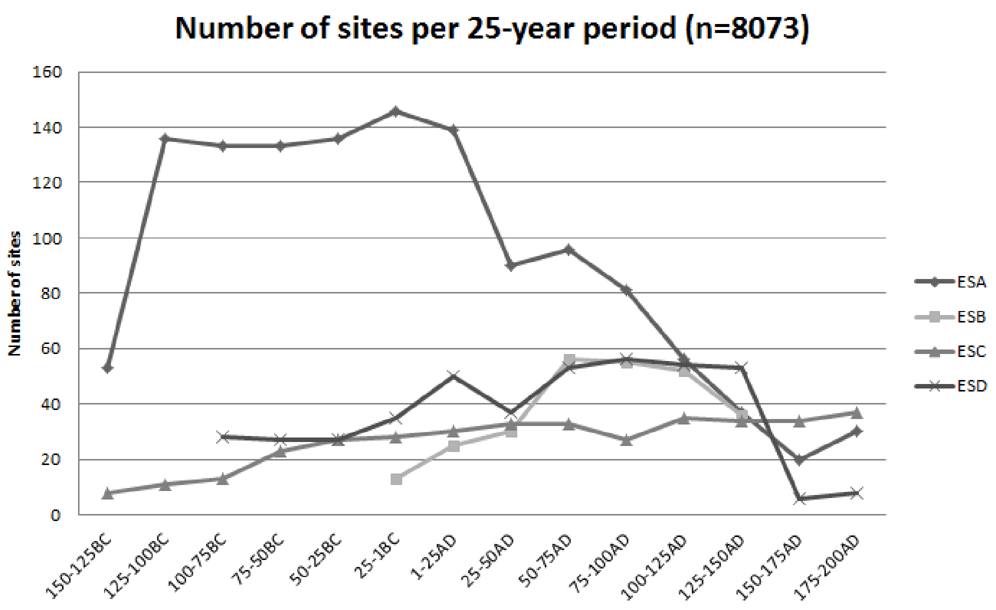


Fig. 42. Number of sites a certain ware is attested at, per 25-year period (n=8073).

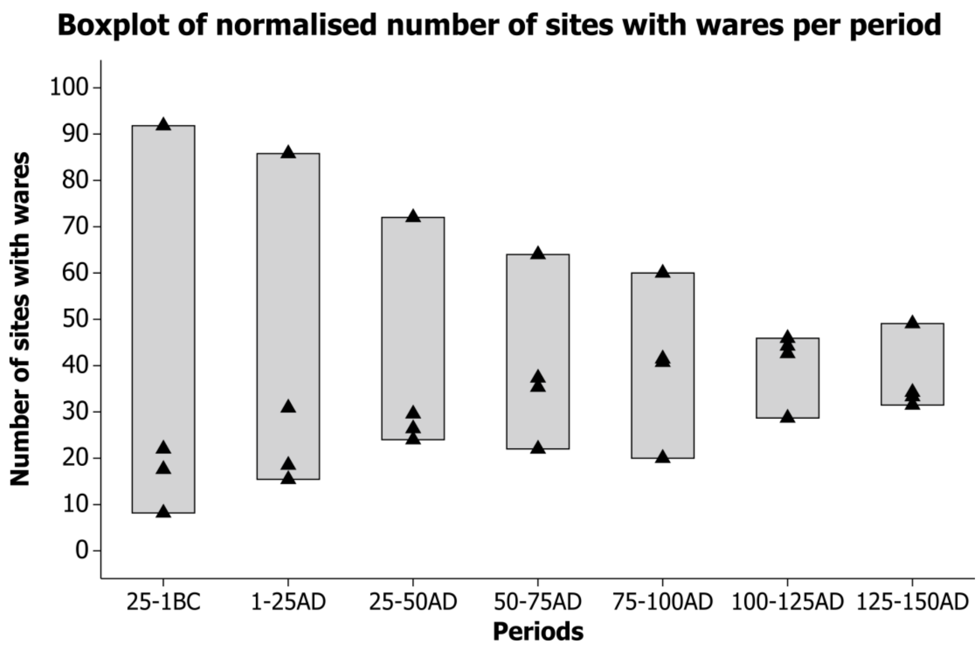


Fig. 43. Boxplot of number of sites with wares per period. Triangles represent individual wares, grey boxes indicate the range (wideness) of distribution. Values were normalized to a scale of 100 to facilitate comparison with the simulated distributions.

5.4.7. Frequency distributions

In this section I will explore the general patterns identified in the charts above in a bit more detail through three frequency distributions, which compare the number and proportion of sherds, forms, and sites with each other per period. This offers the advantage of exploring how the pattern of interest in this case study is made up of particular site assemblages with different configurations and volumes of wares and forms. These distributions will be compared to the simulation results in section 5.5 and I will therefore only describe them for the period from 25BC to 150AD in which four tablewares circulated on a large scale in the Eastern Mediterranean.

Wares per site

Figure 44 shows an example of a frequency distribution of the number of wares attested at sites (according to the Gaussian probability distribution). The frequency distributions for each 25-year period between 25BC and 150AD show the same general trend: the vast majority of sites have evidence of only one ware, whilst a low number of sites have evidence of two, three or four wares.

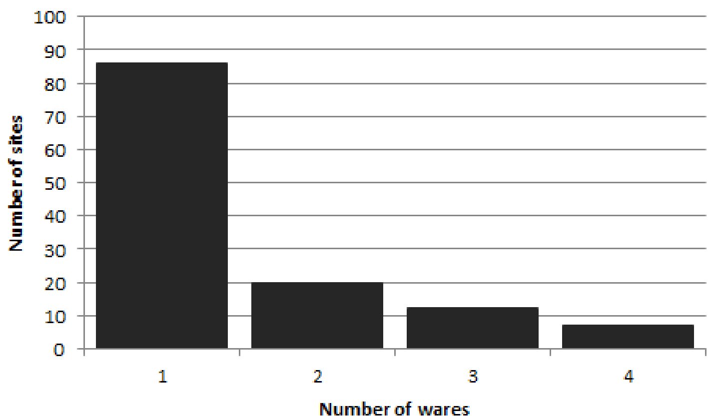


Fig. 44. Frequency distribution of the period 25-50AD showing the number of sites (Y-axis) on which a certain number of wares (X-axis) is attested. The frequency distributions of all other periods between 25BC and 150AD show the same general trend.

Sherds per site

The histograms in figure 45 show the proportion of sherds per ware that make up sites' assemblages. The frequency distributions therefore reflect the number of sites per 25-year period where certain wares are the majority or minority according to the numbers of sherds recorded in the ICRATES database. One pattern is immediately obvious: ESA has a significantly wider distribution than the other wares and is the dominant ware at a very high number of sites until at least 75AD. After that, the geographical distribution of ESA is matched by that of other wares. One might argue that the number of sherds is sensitive to sampling issues and that the proportion of types of each ware found on sites is a better indicator of assemblage diversity and wares' distributions.

Types per site

The frequency distributions of types per site (Fig. 46) show a largely similar pattern, which is what we would expect given that many publications included in the ICRATES database only describe diagnostic sherds (i.e. not a full quantification but rather providing evidence of the diversity of attested types). However, these frequency distributions of types also show a higher diversity of assemblages (see the 26-50% categories), which was less clear from the frequency distributions of sherds. On a large number of sites ESA is the only ware, indicating a wide distribution unmatched by the other wares until at least 75AD. Until that same period, a high number of sites also have more diversity of ESA forms than of any other ware. It is interesting to note that although ESC has a much higher number of forms than ESA and the other wares from 75BC onwards (see Fig. 41), the number of sites these were found on is actually quite limited (see Fig. 42). This is the reason why ESC is the majority ware in relatively few sites throughout most periods, and is a minority ware in more sites than ESA. One could point out that many of the sites where ESA is the only ware are sites with only one sherd of one form in the database, which is especially the case for the sites close to the production area of ESA covered by the Tell Rifa'at survey and published by Kenrick (1981). But even if we disregard sites where one ware makes up 100% of the assemblage, we still notice that ESA is the majority ware on more sites than any other ware for most of the period under study.

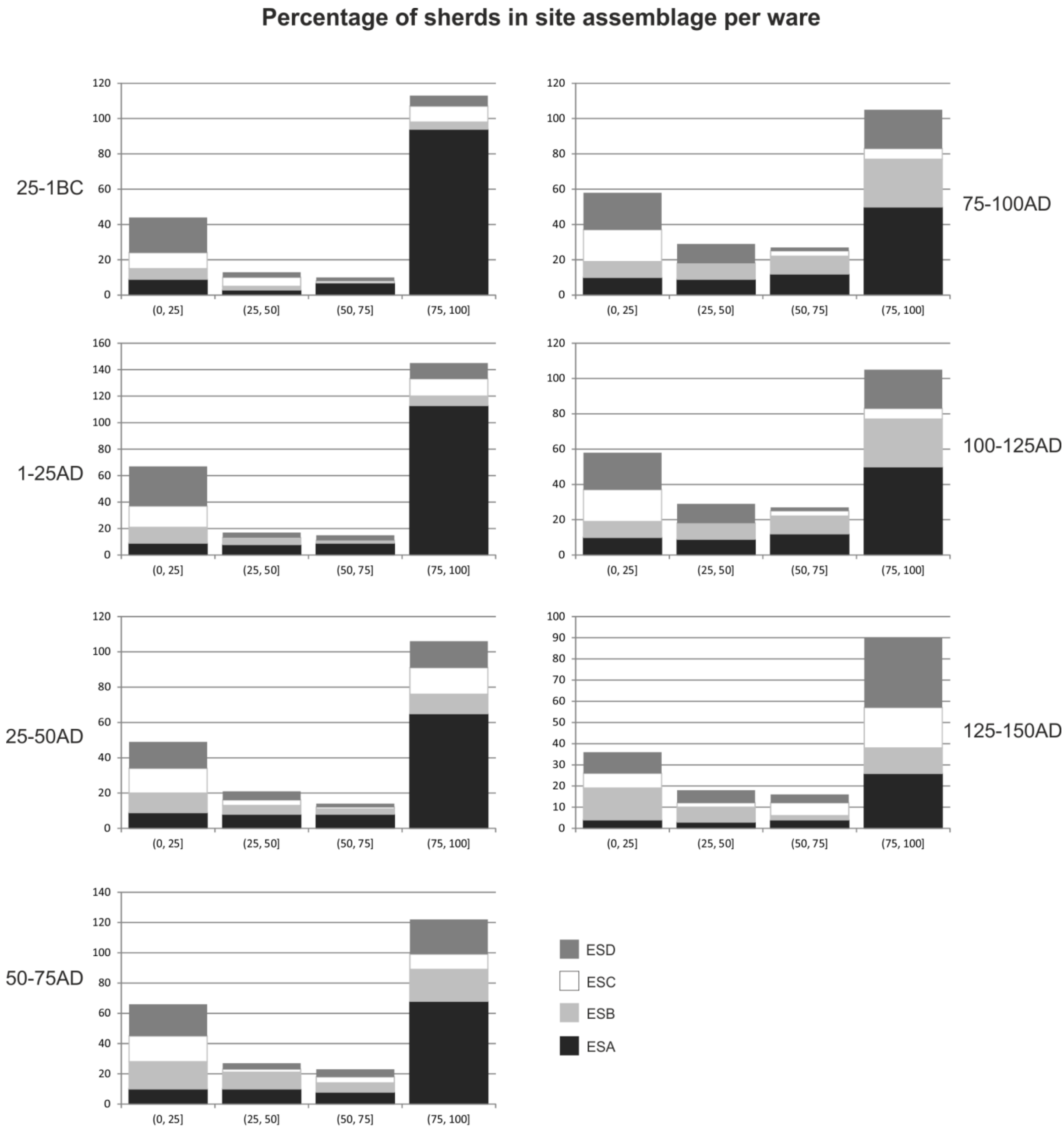


Fig. 45: Frequency distribution per 25-year period showing the number of sites (Y-axis) for which a certain percentage of sherds of a certain ware (X-axis) is attested. The X-axis shows four percentage categories: 1-25; 26-50; 51-75; 76-100. For example, in the period 25-1BC ESA makes up 100% of all sherds on 94 sites, whilst ESC makes up 100% of all sherds on only 9 sites.

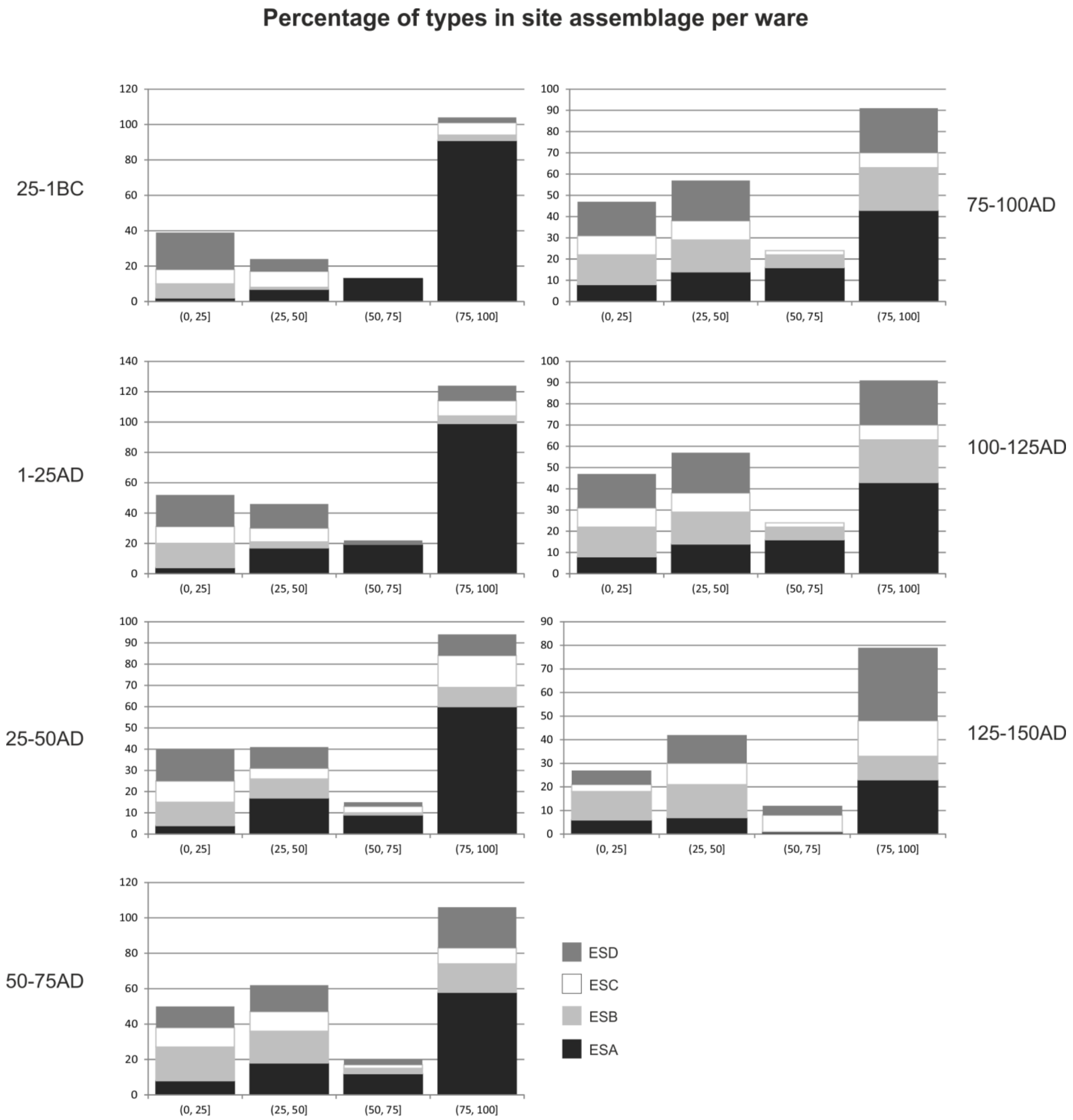


Fig. 46: Frequency distribution per 25-year period showing the number of sites (Y-axis) for which a certain proportion of forms per ware (X-axis) is attested. The X-axis shows four percentage categories: 1-25; 26-50; 51-75; 76-100. For example, in the period 25-1BC there are 91 sites where 100% of the total number of forms is ESA.

5.4.8. Exploratory network analysis results

An exploratory network analysis of the dataset was performed to further explore the degree to which forms of the same ware have a similar distribution. To this aim the analysis uses an inversed Brainerd-Robinson similarity index, which results in a numerical indication of how similar the distribution of a pair of forms is. Using these Brainerd-Robinson (BR) similarity values a similarity network can be created and analysed. Through a sensitivity analysis I evaluated the impact on the exploratory network measures' results of focusing the exploratory network analysis on subnetworks with a different minimum similarity value, i.e. an evaluation of the impact of the common practice of placing an arbitrary threshold on the edge value in order to focus analysis on the most robust patterns. I decided to include this exploratory network analysis and sensitivity analysis as an appendix (see appendix IV section 12.2) since it does not introduce innovative network measures compared to the previous two case studies (indeed the kind of analysis presented here is by far the most common way archaeologists have been applying network science techniques) and since its results are not directly comparable to the simulation output of the ABM. The main contribution of this exploratory network analysis and sensitivity analysis to this case study is therefore as an additional exploratory data analysis which helps answer the first research question of this case study concerning the general trends in the similarity of forms' and wares' distribution patterns. Here I will provide the results of this analysis and their implications for the aims of the case study are given in section 5.4.8.

The distributions of tableware forms are not very similar in general, although there is great variation in the BR values. The sensitivity analysis suggests that overall the exploratory network measures show the same general trends for the complete networks as for the networks with a threshold applied: decreasing density, high clustering coefficient, peak in heterogeneity, similar trend in the number of nodes. This suggests these trends are quite robust and should inform the discussion of the exploratory network analysis results. However, not surprisingly the effect of applying a threshold on the 'mean + standard deviation' BR value had the strongest impact on the results, for example concerning the decrease in the number of ESA forms included. The proportional change of nodes in rankings as a result of changing thresholds is very common but never dramatic. The change for the clustering coefficient is slightly higher on average, between 0 and 45%. Change for the degree measure is between 15 and 35%. The period 150-125BC shows strong sensitivity, especially for the degree measure. Overall this sensitivity analysis indicates that the clustering coefficient and degree rankings of all periods are sensitive to changing thresholds, but not dramatically, suggesting that the most persistent patterns will be reflected in all thresholds. This analysis also indicates that the outliers and networks with a high degree of change should be addressed in more detail, which was done in the exploratory network analysis.

It should be clear that the sensitivity analysis performed here is by no means conclusive about the robustness of the exploratory network analysis results. One could argue that the sampling bias in the ICRATES dataset is not directly addressed (something that could be done by removing randomly selected chunks of data, performing the analysis again and comparing results, or by comparing the observed networks with simulated random networks with the same number of nodes and edges). This sensitivity analysis merely served to explore the dataset in its network form and calibrate any possible interpretative statements about pairs or groups of forms that show particularly high similarity. A more in-depth evaluation of the sampling bias in the ICRATES dataset was not within the scope of this project which focuses mainly on the network side of things.

The technical description of similarity networks provided in Appendix IV section 12.2 can be summarised by focusing on the general trends. From 25BC onwards ESA distribution becomes less internally homogeneous. Although ITS becomes widely distributed around this time, this is nevertheless also a local Eastern phenomenon. ESD, ESB and ESA show increasing similarities in their distribution patterns, especially between 25BC and 25AD. However, one should not forget that during this period ESA was still present on a much larger number of sites than any of the other wares. The higher similarity between wares in this period should also be seen in light of ESA being the dominant ware on many sites with a diverse assemblage (see section 5.4.7). After that the similarity of ESB and ESD to ESA forms' distributions decreases, whilst ESA maintained an unrivalled wideness of distribution. There is another increase of similarity between these wares' distributions from 75AD until 125AD. As well as a higher similarity to ESB and ESD, it is important to note that the period between 75AD and 125AD also marks the decrease of the wideness of ESA distribution. ESC on the other hand retains its internal homogeneity until its corpus of forms decreases from 75AD onwards. From then on it becomes increasingly similar to ESB and to some extent to ESA, although this trend stops around 125AD. ESC remains very dissimilar to ESD throughout all periods.

5.4.9. Conclusion

The exploratory data and network analysis results together provide an answer to the first research question of this case study: "What differences can be observed in the distribution patterns of different tablewares, forms (here considered modern analytical constructs)? Are forms' distributions always more similar to those of the same ware?". We have seen that the answer to this question is a complex one: during some periods ESA, ESB and ESD knew a very similar distribution whilst ESA was still most widely distributed, during other periods all wares' distributions were dissimilar. However, the detailed technical descriptions also highlighted those forms whose distribution is more similar to that of other wares. It becomes clear that we cannot merely assume that all forms of the same ware were similarly distributed. Moreover, when we bring the assumption used here to mind, that similarity in distribution patterns of forms might indicate similarity in the processes driving distribution, this suggests that tableware distribution processes were largely ware-specific but with some exceptions. However, we need to acknowledge that forms are a modern construct and might not be the best analytical category to make such statements about. Future work should focus on similarities in distributions of functional categories as well. Indeed, the exploratory network analysis also emphasised the importance of clearly stating the theoretical assumptions underlying the selection of a certain similarity measure, i.e. archaeologists should not motivate their decision to use the Brainerd-Robinson method because it is commonly used by archaeologists but because the assumptions underlying the way it captures similarity of artefact assemblages or forms' distributions is desirable within their research context.

The exploratory data and network analyses presented here made the exceptional distribution of ESA clear. It is observed both on the level of the aggregated data per 25-year time-slice, as well as when looking at differences in the distributions per time-slice in more detail. To this exploratory data analysis we can add the geographical wideness of ESA distribution identified in previous work (Bes 2007; Brughmans and Poblome *in press*; Willet 2012). It can be concluded that this selection of the ICRATES dataset mirrors the results of previous studies of the topics, suggesting that ESA was for centuries of supra-regional importance, whilst other Eastern-produced wares were more of regional importance for most of their production-life. It also became clear that the smaller trends identified are often created by the standard typo-chronological frameworks used, by the different densities and number of sites excavated by archaeologists in different regions around the Mediterranean, or by the common practice to only publish diagnostic sherds. It should therefore be emphasised that only the more general trends identified here should be considered as the archaeological pattern I aim to study as the outcome of distribution mechanisms that are the research focus of this case study. This has the further implication that our knowledge of this distribution pattern is not as accurate as we might need it to be for scientific hypothesis testing, since it is not possible to describe this distribution pattern in anything other than the most general terms with a considerable error range. However, this condition should not be considered an argument against hypothesis testing, since we can still falsify hypotheses by identifying those distribution mechanisms that do not even succeed in producing outcomes that fit this very general pattern. These distribution patterns will therefore serve as a reference when exploring the results of the agent-based network models in particular (section 5.5), where I will evaluate which simulated scenarios are and are not able to reproduce the patterns identified here.

This section has emphasised the need to only compare the observed tableware distribution with simulated distributions through descriptive statistics, and focus hypothesis testing on falsification rather than cherry picking those experiments that show a perfect fit with the data. However, these exploratory data and network analyses still did not bring me any closer to an indication of what the distribution processes that gave rise to the tableware distribution patterns might be. This work was nevertheless crucial in order to be able to critically compare the archaeologically observed distribution pattern with simulated distribution patterns. In the next section I will evaluate two hypotheses about the processes that could have given rise to the differences in tableware distribution patterns, in light of the complex picture of the data painted in this section.

5.5. Agent-based model

In section 5.2 above I argued that agent-based network modelling might form a useful part of an innovative method for expressing hypotheses about the workings of the Roman economy, and to compare their simulated outcomes with the archaeological record. However, I also highlighted the wide range of hypotheses argued by archaeologists and historians alike to account for the archaeologically observed distribution patterns, and I emphasised that these cannot all be evaluated through one single approach. Rather, I see more use in Roman archaeology for approaches that allow for evaluating the impact of individual contributing factors. Such a research process forces archaeologists to explicitly formulate their hypotheses, their assumptions about how exactly these factors functioned, and how the archaeological record can allow us to answer the questions we pose. I decided to focus on one key contributing factor which, as I have argued above, could be considered a concept holding together many of the hypotheses: the structure of social networks.

In this section I will first describe the model and, most crucially, explain and motivate my abstraction of the phenomenon under study (tableware trade in the Roman East) in terms of network concepts. This abstraction will then lead to an implementation of my ideas as a computational model, whose main procedures will be discussed in detail (the model is available as an electronic supplement to this PhD; the model code is provided in Appendix IV section 12.3). A summary of the network science research process applied in the creation and analysis of this ABM is given in Box 9. Finally, the results of the model are presented and compared with the archaeological record.

5.5.1. Model description

An agent-based model was created that represents the structure of social networks between traders that act as the channels for the flow of commercial information and goods. When the model is initialised a social network is created between traders, who are distributed among sites. Four of these sites are production centres of four different wares, and traders located at these sites obtain a number of items of this locally produced ware in each turn. At each time step traders will determine the local demand for tableware they want to satisfy, and will estimate the price they believe an item of tableware is worth based on their knowledge of the supply and demand of the traders they are connected to. Every item of tableware is then put up for sale, and pairs of traders who are connected in the network can buy or sell an item. When an item is successfully traded, the buyer will decide to either sell it to a local consumer to lower the demand (in which case the item is taken out of the trade system and is deposited at that site), or to store it for redistribution in the following turn in case this promises a higher profit. Over time, this model therefore gives rise to distributions of four tablewares.

The following section explains how the hypotheses of Bang and Temin can be tested in this model by changing the network structure in two specific ways.

Conceptualisation and representation

The social networks described by Bang consist of a strong *community structure* within *markets*, *limited availability of commercial information* between communities, and *weakly integrated markets*. The social networks described by Temin consist of *limited availability of commercial information*, and *well integrated markets* where prices are determined by *supply and demand*. The terms in italics are concepts these authors use to abstract their ideas about the phenomenon of Roman trade, and will need to be clearly defined before implementation in a computational model. I argue that the hypotheses of Bang and Temin can be usefully conceptualised and represented as follows:

- Commercial actors have agency (these actors are here referred to as traders, but they could represent any type or number of commercial actors that enable the flow of goods and commercial information). In this model traders are software agents represented as nodes.
- Markets are places where traders come together, where different communities of traders exist, and where tableware and commercial information are available. In this model all traders are distributed among a set number of markets.
- Traders have the ability to share commercial information and goods with other traders. This ability is represented as a link between a pair of nodes.
- The availability of information is limited when a proportion of the traders a trader is able to trade with does not share commercial information with the trader. This proportion of available commercial information is represented as the proportion of neighbouring nodes a node knows the supply (volume of tableware in possession) and demand of. This proportion is tested by varying the variable *local-knowledge* in experiments. A low value of this variable represents limited availability of accurate commercial information (≈Bang and Temin).
- Communities consist of traders who are more likely to trade and share commercial information with each other. This is represented by a social network structure with a high clustering coefficient within markets, and with a lower number of links between clusters than within clusters.
- The integration of markets is limited if the ability to share commercial information and goods between markets is limited. This is represented by the proportion of all possible links between nodes that connect nodes on different sites. This proportion is tested by varying the variable *proportion-inter-site-links* in experiments. A high value for this variable represents highly integrated markets (≈Temin), a low value represents weakly integrated markets (≈Bang).
- The supply of a trader is the amount of tableware this trader owns and is willing to sell. The demand is here used as the demand of consumers a trader is aware of and is willing to supply for. In this model supply and demand of traders constitute commercial information. Traders obtain this information from a proportion of the traders they are connected to and use it to estimate the price they believe tableware is worth within their part of the social network. A seller will only agree to sell tableware for more or the exact amount he estimates it is worth. A buyer will only buy tableware

> for less or the exact amount he estimates it is worth. (this means that there is no "haggling" and negotiation of prices in this model, a decision made to simplify the model and to be better able to evaluate the impact of differential availability of commercial information).

This conceptualisation and representation is implemented in the ABM presented here, which allows me to explore the hypotheses by Bang and Temin in two ways: by changing the proportion of network links between traders located on different sites (*proportion-inter-site-links* variable), and by changing the proportion of accurate commercial information a trader has access to (*local-knowledge* variable).

The structure of the social network does not change within a single experiment, a decision made in order to be able to assess the impact the social network structure has on giving rise to the archaeological pattern of interest. This means that this model does not represent traders moving from one town to another and create new trade contacts over time. Instead, the model represents the availability of commercial information and trade partners in different towns. In this model, therefore, an edge between a pair of traders located on different markets represents the possibility of obtaining commercial information from a different market and the ability to trade with someone located in a different market. Whether those traders actually physically moved between markets is not the main interest in this study and it is therefore not implemented; what matters is their ability to enable the flow of goods and information. For this reason, I decided against a geographically more realistic modelling of the Eastern Mediterranean and to focus on these flows in an abstract topological space. The connectivity of commercial actors is a key feature of both Bang's and Temin's hypotheses, and I believe a representation as network data is preferable as an initial comparison of these hypotheses. This model, therefore, aims to express these two hypotheses as different network data representation, and it is the effect of changing the structure of the hypothesised network that will be studied here.

A final important factor deserves our attention before describing the technical details of the model: the conceptualisation of time. The accuracy of dates is always problematic in archaeology, and no less so for Roman tableware as shown by Theodore Peña (2007) in his model of the use-life of this type of pottery, and in Rinse Willet's (2012, 44-46) hypothetical calculation of tableware volumes in circulation in the Eastern Mediterranean. One could adopt Peña's estimate for the use-life of tableware of between one and three years to suggest a realistic setting for new demand caused by the deposition of previously owned tableware. However, this would depend on realistic estimates of the ancient population and its distribution among towns in the Eastern Mediterranean. Estimates for this are available (e.g. Parkin and Pomeroy 2007; Scheidel et al. 2007) but they are by no means unproblematic. Moreover, this model does not aim to recreate the changing distribution pattern observed in the archaeological record through time, and its production and consumption procedures are purposefully kept as simple as possible. It merely wishes to evaluate whether different structures of social networks can give rise to differences in the wideness of wares' distribution. In this model I therefore decided to work with a relative 'transaction time' rather than an absolute timeframe: the time of each time step is the time it takes for all tableware available for trade to be considered in a transaction, and the demand to increase by one item per trader if it is not at its maximum. In future work this model should be given a more realistic time conceptualisation by modifying settings for tableware volumes and the increase in demand.

Variables and outputs

I have already introduced the two variables whose values I change in order to test the hypotheses. The ABM described in the next few sections includes a large number of variables, although most of them are simply used to store a certain value (e.g. the fact that a site is a tableware production site, or the amount of tableware deposited on a site). A number of other variables have a specific default value, which are motivated in the next few sections. Tables 17-18 list all variables included in this ABM, with a short description, and a default value for some. Before I describe the model in more detail it is important to distinguish here between independent and dependent variables, and my definition of these. Independent variables can be defined as the inputs, causes, or the variables tested to evaluate whether they cause variation in the dependent variables. Dependent variables are the outputs or effects, or are tested to see if they are the effects. The independent variables listed in table 17 are therefore hypothesised to cause the differences in the amount of products deposited at sites (dependent variables), and they will not change during a simulation. The dependent variables listed below will change throughout the simulation as a result of the trade procedures. The outputs of the model are the values of the dependent variables at the end of an experiment, i.e. the simulated volume of tableware at sites, diversity of site assemblages, and wideness of wares' distributions. What I aim to do in this model is to evaluate the effects of two specific independent variables (*proportion-inter-site-links* and *local-knowledge*) on the variation observed in the dependent variables, because these two variables are my abstractions of the hypotheses I wish to express and test.

Table 17. Independent variables: the inputs, causes, or the variables tested to evaluate whether they cause variation in the dependent variables.

Independent variables		
Variable	**Description**	**Default value**
Global variables		
Num-traders	The total number of traders to be distributed among all sites	2000
Num-sites	The total number of sites	100
Traders-distribution	Determines how the traders are distributed among the sites, this can be either “uniform”, “normal”, or “exponential” frequency distributions	exponential
Network-structure	When set to “hypothesis” this connects traders to create a small-world structure that represents the hypothesised social network structure; when set to “random” this connects traders to create a random structure with the same number of nodes and edges as would be expected in a “hypothesis” network with the same global variable settings. (not tested here)	hypothesis
Maximum-degree	The maximum number of connections any single trader can have	5
Proportion-inter-site-links	The proportion of all pairs of traders that are connected in step two of the network creation procedure by inter-site links	tested (0, 0.00005, 0.0001, 0.0003, 0.0006, 0.001)
Proportion-intra-site-links	The proportion of all pairs of traders that are considered for becoming connected in step three of the network creation procedure by intra-site links	0.0005
Proportion-mutual-neighbors	The proportion of all pairs of traders with a mutual-neighbor that are considered for becoming connected in step four of the network creation procedure by intra-site-links	2
Site-specific variables		
Producer-A	Set to "true" if the site is the production centre of product-A	
Producer-B	Set to "true" if the site is the production centre of product-B	
Producer-C	Set to "true" if the site is the production centre of product-C	
Producer-D	Set to "true" if the site is the production centre of product-D	
Trader-specific variables		
Demand	The proportion of the demand at the market the trader is located at that he aims to satisfy by obtaining products through trade	Constant increase of 1 per turn; maximum = the number of traders at the site
Local-knowledge	The proportion of all traders a trader is connected to that he receives commercial information (supply and demand) of in each turn	tested (0.1, 0.2, 0.5, 1)
Transport-cost	The fee which is subtracted from the profit a trader expects to make from a transaction with a trader located at a different market (not tested here)	0

Table 18. Dependent variables: the outputs or effects, or are tested to see if they are the effects.

Dependent variables	
Variable	**Description**
Site-specific variables	
Volume-A	The number of items of product A deposited on this site as a result of a successful transaction
Volume-B	The number of items of product B deposited on this site as a result of a successful transaction
Volume-C	The number of items of product C deposited on this site as a result of a successful transaction
Volume-D	The number of items of product D deposited on this site as a result of a successful transaction
Trader-specific variables	
Product-A	The number of items of product A a trader owns and can trade or store in this turn
Product-B	The number of items of product B a trader owns and can trade or store in this turn
Product-C	The number of items of product C a trader owns and can trade or store in this turn
Product-D	The number of items of product D a trader owns and can trade or store in this turn
Stock-A	The number of items of product A a trader puts in his stock in this turn as a result of an unsuccessful transaction or for redistribution in the next turn
Stock-B	The number of items of product B a trader puts in his stock in this turn as a result of an unsuccessful transaction or for redistribution in the next turn
Stock-C	The number of items of product C a trader puts in his stock in this turn as a result of an unsuccessful transaction or for redistribution in the next turn
Stock-D	The number of items of product D a trader puts in his stock in this turn as a result of an unsuccessful transaction or for redistribution in the next turn
Maximum-stock-size	The number of items a trader is willing to obtain through trade this turn in addition to his own demand if the average demand is higher than his demand
Price	The price a trader believes an item is worth based on his knowledge of supply and demand on the market

Setup procedures

The model is initialised by creating 100 sites and 2000 traders, distributing the traders among the sites, and connecting traders to construct a social network. Sites are positioned along a circle, which is convenient for setting up the social network and for visualisation, but it does not represent sites that are geographically close. Traders are then distributed on these sites following an exponential frequency distribution with the mean equal to the mean number of traders per site. This exponential frequency distribution will result in strong differences between the number of traders per site, where a few sites have a very high number of traders (between 58 and 135 traders in the experiments) and most sites have a much lower number of traders. Since the demand of a site in this model is determined by the number of traders present at a site, this exponential distribution is considered to reflect the differing demands of markets throughout the Mediterranean: markets with an extremely high demand are relatively rare (e.g. Antioch) whilst most markets will have a much more modest demand.

Traders are subsequently connected to each other to form a social network with a structure that represents the hypotheses tested by setting the variable *network-structure* to the value "hypothesis". This structure is considered to have a high clustering coefficient within markets relatively few links between clusters, and a modifiable proportion of links between markets. It was decided to use the 'small-world' network structure as a baseline for creating this hypothesised network. The procedure to create a network with a 'small-world' structure is inspired by the simplified model[29] for the growth of social networks by Jin, Girvan and Newman (2001), which has previously been applied in an archaeological model of exchange by Bentley, Lake and Shennan (2005). The simplified version of the model was selected because it gives rise to the network structure of interest (maximum degree, low average shortest-path length, high clustering coefficient) with relatively few parameters. In particular, the procedure to initialise this model is used here, it repeats two steps until "all or most vertices have degree z^*" (z^* is the maximum degree, 5 by default): (step1) a proportion of all node pairs is randomly selected and connected if they are not connected yet and if neither node has the maximum degree; (step2) a proportion of all pairs of nodes with a mutual neighbour is connected if they are not connected yet and if neither node has the maximum degree. This procedure was modified to represent the hypothesis being tested here. The model's network creation procedure therefore consists of five steps:

- Firstly, one pair of randomly selected traders located on neighbouring sites in the circular layout is connected between each pair of neighbouring sites. This ensures commercial information can flow from each site to its two neighbouring sites whilst offering a minimum of connectivity between sites that allows for goods to still be distributed to all sites. In scenarios where no other inter-site links are added, information and goods will therefore need to travel from site to site along the circular layout.
- Secondly, a variable number of inter-site links is created. A proportion s_0 (determined by the variable *proportion-inter-site-links*) of all trader pairs n_p is connected if a pair is not located on the same site and is not connected yet. The total number of trader pairs n_p is calculated as:

$$\frac{1}{2} N (N - 1)$$

Eq. 3. Calculation of total number of trader pairs.

[29] Referred to in Jin et al. 2001 as Model II.

where N is the total number of traders. The variable s_0, which determines the proportion of all node pairs to be connected in this way, is used in the experiments to represent the availability of commercial information (together with the *local-knowledge* variable) and the ability to trade with traders in different sites. s_0 is a float, in the experiments I use a value of between 0 and 0.001 (representing a proportion of 0% to 40.5% of all links being inter-site links).

- Thirdly, a proportion r_0 (determined by the variable *proportion-intra-site-links*) of all trader pairs n_p is connected if they meet the following requirements: the pair is not connected yet, neither of the traders has the maximum number of connections z^*, and they are located at the same site. The maximum degree z^* (determined by the variable *maximum-degree*) is an integer and is 5 in the default setup, whilst the proportion r_0 of trader pairs selected in this way is a float and is 0.0005 in the default setup.
- Fourthly, $n_m r_1$ traders are selected at random (determined by the variable *proportion-mutual-neighbors*). If these traders are connected to a pair of traders on the same site that are not connected yet and do not have the maximum number of connections z^*, then this pair of traders of whom the randomly selected trader is a mutual neighbour will be connected. r_1 is a proportion of all trader pairs with a mutual neighbour n_m, the latter is calculated as:

$$\frac{1}{2}\sum_i z_i(z_i - 1)$$

Eq. 4. Calculation of all trader pairs with a mutual neighbour.

where z_i is the degree of the i^{th} trader. This step is responsible for the high level of clustering and is a process common in social networks called transitivity, which stands for the idea that a pair of individuals who have a mutual friend have a high probability of becoming friends themselves in the future.

Steps three and four of this network creation procedure are subsequently repeated until the average degree of the network approximates the maximum degree of each trader, i.e. many traders will have a degree close to the maximum, those who do not have a maximum degree cannot create any further links without violating the rules of steps three and four. The default values for the variables in steps three and four were adopted from Jin et al. (2001) and result in a ‘small-world’ network structure of traders on each site.

- Fifthly, at this stage the network can still have a few isolated traders and consist of multiple components. Therefore, a minimum number of edges are added between pairs of traders on the same site which are not in the same connected component, to ensure all traders become part of a single component. This last step was enforced to make it theoretically possible for each trader to receive each of the products. It generally results in very few (between 0 and 3) extra links being created and does not affect the ‘small-world’ structure much.
- For every experiment the number of edges created in each of these five steps is recorded.

The result is a network structure where neighbouring sites are connected by at least one link between a pair of traders, where traders within the same site are connected in clusters with few connections crossing clusters, and a variable number of inter-site links depending on the hypothesis being tested (Fig. 47). Traders on each site are therefore connected following a ‘small-world’ network structure (Watts and Strogatz 1998), with a high clustering coefficient and a low average shortest path length. However, the overall network structure of all traders on all sites combined will not always show the characteristics of a ‘small-world’ network, since the number of inter-site links added in addition to the links connecting traders on neighbouring sites is determined by the variable *proportion-inter-site-links* used to represent different degrees of integration of markets (Table 19).

Table 19. Examples of the number of links added in each step of the setup procedure for different settings of the proportion-inter-site-links variable. Values were derived from experiments with the same randomization seed, due to the stochasticity in the model minor differences in the numbers for steps 3 to 5 appear for different randomization seeds.

Variable: proportion-inter-site-links	**0**	**0.00005**	**0.0001**	**0.0003**	**0.0006**	**0.001**
Step1: links to neighboring sites	100	100	100	100	100	100
Step2: proportion inter-site links	0	100	200	600	1200	1999
Step3: randomly created intra-site links	2942	2895	2889	2779	2541	2251
Step4: intra-site links created between pairs with mutual neighbors	1787	1738	1642	1390	1029	580
Step5: links created between components	3	2	1	0	0	1
Total number of links	4832	4835	4832	4869	4870	4931

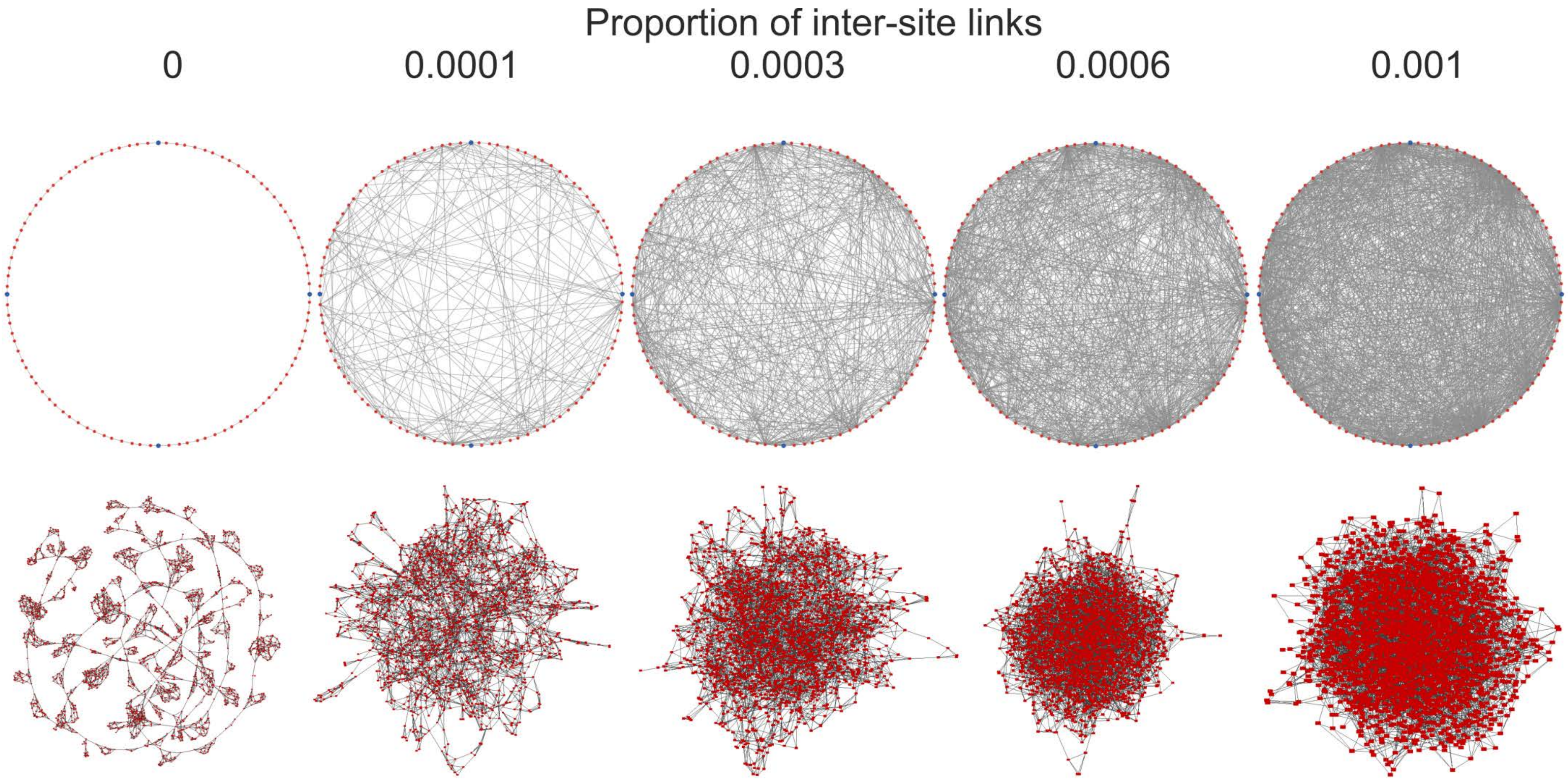


Fig. 47. Example of the network structure generated in the setup procedure of the ABM for different values of the proportion-inter-site-links variable. At the top sites are layed out along a circle and traders are positioned at sites. At the botton sites are no longer included and the traders' social network is layed out using a force-directed layout algorithm (yFiles Organic layout in Cytoscape) to display its structure. Note the existence of clusters of traders on sites connected to few other clusters, a pattern which gradually disappears as traders receive more inter-site links and the sites become more integrated.

Trade procedures

In every time step of the model traders determine their demand, they discard part of their stock (due to broken or unfashionable items), traders on tableware production sites obtain new items, traders obtain commercial information, they determine what they believe to be the current price of an item, and then all items owned by all traders in that turn are considered for trade.

Determining demand: *demand* in this model is a variable of the traders and represents the demand of consumers they are aware of and they believe they can supply for. Demand can only be satisfied by obtaining an item of any product through a successful transaction between a pair of traders. The demand of each trader is 0 at the start of the simulation. To represent the unequal demand for products of sites the maximum demand of a single trader equals the total number of traders at the site he is located at, and the maximum demand of traders at that site is therefore the number of traders located there squared (i.e. the maximum demand at a site increases by the number of traders squared, following a quadratic function $maxDemand = NumberOfTraders^2$, see Fig. 48). Each time step, a traders' demand is increased by 1 if his demand is lower than the maximum. Demand of consumers therefore has a maximum (i.e. at some point the inhabitants of a site do not require any more tableware), it is satisfied for a while (i.e. the demand is not topped up to the maximum every turn: when a consumer obtains an item they do not immediately require a new one), and gradually increases to the maximum (i.e. new demand is created by the breaking, or becoming unfashionable of old items, or by renewed need for the item).

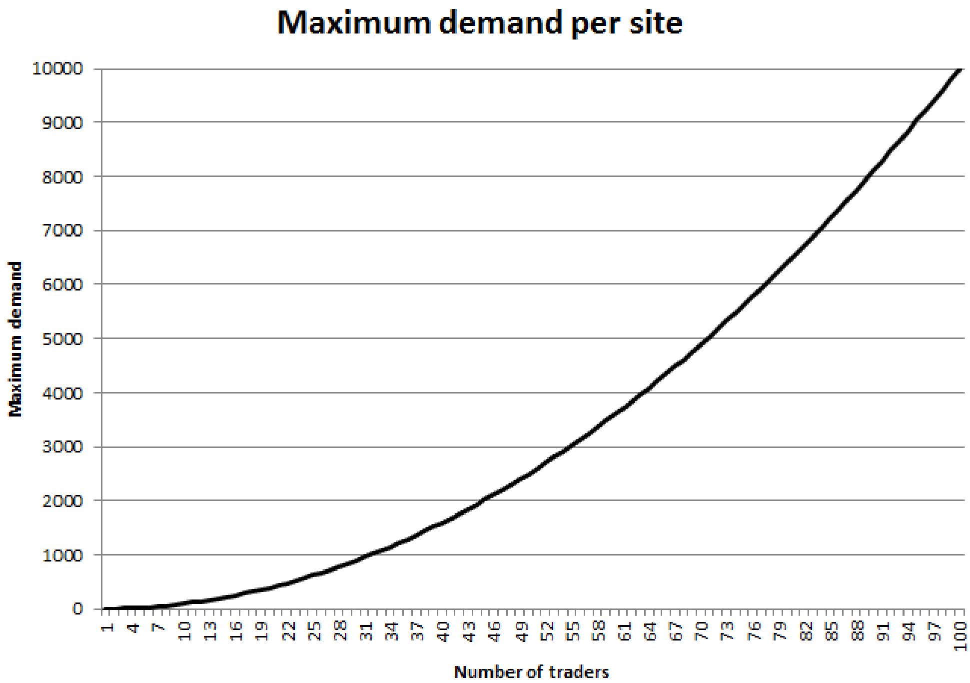


Fig. 48. The maximum demand at a site equals the number of traders squared.

Stock: in a previous turn items can be put in stock as a result of either a failed transaction or a deliberate storing of the item for redistribution in the next turn. At the start of a turn traders who have stock will need to discard a fixed proportion of it, this penalty reflects the risks involved in not immediately selling an item on to a consumer but storing it for redistribution and represents broken and unfashionable items. The proportion of discarded items is set to 14% in this model, as suggested by Peña's model (2007, Fig. 11.5) to have been either recycled or discarded.

Production: traders on the production site obtain newly produced items each time step, only if their total current possession of all products is less than their demand. If this is the case then they obtain items of the product being produced at that site equal to their demand minus the sum of their possessed products, i.e. they obtain the number of items needed to satisfy their demand.

Obtaining commercial information and price-setting: every time step each trader will only have commercial information available from a proportion of its link neighbours. This proportion is determined by the variable *local-knowledge* and is used in the experiments to test scenarios with differing availability of information (together with the *proportion-inter-site-links* variable). Commercial information is gathered once in each time step before trade happens, i.e. the traders whose stock and supply becomes known to the trader in question remain the same during a time step, but could be different in subsequent time steps. The trader then calculates the average demand and average supply of this proportion of neighbours, including his own supply and demand. Using this commercial information available to him he then determines what he believes is the price of one item of any product as follows:

$$\boldsymbol{price} = \frac{\boldsymbol{average{-}demand}}{\boldsymbol{average{-}supply{+}average{-}demand}}.$$

Eq. 5. Price estimate by agent

This results in a normalised float value between 0 and 1 following the logic of supply and demand: if the average demand is equal to the average supply then the price will be 0.5; if the average demand is higher than the average supply then the price will be between 0.5 and 1; if the average demand is lower than the average supply then the price will be between 0 and 0.5.

Determine maximum stock: the traders will subsequently determine how many items they are happy to store for redistribution in the next turn. The maximum stock is only higher than 0 when the average demand the trader is aware of is higher than his own demand and his total stock, i.e. when the trader believes there is a high demand which promises higher profits (because his own demand is lower) he will be willing to store the number of items necessary to supply for the average demand.

Each item is considered for trade once per time step. An item is put in a trader’s stock if he cannot make a profit or if none of his neighbours in the network require an item (i.e. their demand equals 0). Items in stock can be redistributed in the next time step. An item is sold to a buyer if the buyer’s price promises a profit or break-even for the seller. The buyer either places the obtained item in stock for redistribution if the average-demand is higher than his demand (i.e. redistribution holds the promise of a higher profit), or he sells it to a consumer. This is the scenario when deposition of tableware takes place. A buyer decides not to redistribute an item and therefore sells it to a consumer: the buyer’s demand is decreased by 1 because some of the local demand is satisfied; the item is taken out of the trade system because the consumer does not redistribute it; and it is deposited on the buyer’s site. The volume and diversity of tableware deposited on sites can subsequently be used to compare simulated tableware distributions with observed ones.

Experiments and variable settings

A total of 24 experiments were performed in order to test my implementation of the hypotheses by Bang and Temin. Each experiment has a different combination of the values for the two tested variables: *proportion-inter-site-link* and *local-knowledge*. I selected a larger number of low variable values in order to see the impact of a limited variable effect on the results. The values chosen for *local-knowledge* are 0.1, 0.2, 0.5, and 1. The values chosen for *proportion-inter-site-links* are 0, 0.00005, 0.0001, 0.0003, 0.0006, and 0.001.

Because these experiments were only designed to test two variables, all other variables were given the default value in all experiments, as shown in tables 17-18. I used 100 sites because this is close to the maximum distribution of ESA in the database (146 sites), the number is convenient to scale the results, and it was close to the maximum of what was computationally possible within the framework of this project. A total number of 2000 traders was used and were exponentially distributed among the sites to reflect the unequal sizes and demands of ancient towns. In the experiment this results in sites having a minimum of 1 trader and a maximum of 135 traders: the vast majority of sites have between 1 and 20 traders (in most cases enough to give rise to local clusters at sites given the maximum degree of 5) whilst just a few sites have a much higher number of traders (50-135). The default variable settings for the three network creation variables (*maximum-degree, proportion-intra-site-links, proportion-mutual-neighbors*) were adopted from the Jin et al. (2001) model from which these variables were derived, because these default values give rise to networks on sites with a high clustering coefficient and a low average shortest path length (features of a social network which we believe to reflect well the hypotheses tested).

The model was implemented in the Netlogo language, which is easy to learn and read, offers a user-friendly interface for visualising the workings of an agent-based model and run experiments with different variable settings. Moreover, it has a large community of users and many example network models are freely available which can be drawn on for inspiration. The most significant drawback of Netlogo is that it can be extremely slow for complicated models with many agents, as the one presented here. Indeed, due to time limitations each experiment was run only ten times (ten iterations). This is insufficient to expect robust results given the stochasticity in the model, and the number of iterations should be higher in future work. However, the results still allow for a descriptive statistics approach to the results and will help understand the behaviour of a model constructed within the available time limits.

These results will be crucial in shaping future versions of this model, and maybe even future attempts at testing Bang’s and Temin’s hypotheses. Each iteration uses a different randomisation seed, but the same ten seeds were used across all experiments. This means that results between experiments are comparable, because the random factors involved in setting up the social network and running the distribution processes are the same, the only difference are the values for the two tested variables.

Each experiment was run for 10,000 time steps (or ticks in Netlogo jargon). This is arguably an arbitrary end to the simulation, but was considered acceptable since preliminary simulations of the model showed that the maximum width of a ware’s distributions rarely changes after time step 5,000. The final results confirm this observation.

The output obtained from these experiments is rich and will be studied in two different ways. Firstly, we obtain the values for all dependent variables (Table 18) at the end of the experiment (i.e. at time step 10,000). Most crucial of these for testing the hypotheses and comparing simulations with the observed distributions of tablewares is the volume of each ware deposited at each site. These simulated results allow me to identify how many sites each ware was deposited on and what proportion each ware takes up in a site’s assemblage, which will be compared with the combined archaeological data. Although one could also study differences in the price of goods between different sites, a key aspect of Temin’s hypothesis, this is not done here due to time limitations. Secondly, in each simulation the social network will have a different structure and we can use network measures to describe these differences. In particular, I will compare the closeness centrality, betweennes centrality, number of inter-site links, and number of traders at a site with the output of the dependent variables: the total volume deposited at a site at the end of the simulation. The correlation between the social network structure and these results from the simulation will provide a better insight into the workings of this complex model, and will guide the modification of the model in future work.

5.5.2. Results

The previous section discussed how the archaeological research question of this case study, the hypotheses being tested, and the dataset used can inform a network science research process of conceptualisation and representation as network data (see Box 9). An ABM was created by following this research process, and this section provides a technical description of the results. This will be followed by a detailed discussion of the relevance of these results for advancing our understanding of the Roman trade system.

Box 9. Network science research process ABM case study 3

Phenomenon studied:

- The functioning of social networks of traders, and their ability to drive the distribution of tableware in the Roman East.

Abstraction as network concepts:

- Commercial actors (traders) with agency are the entities of research interest.
- Traders can trade on markets.
- Through a social network traders can share commercial information and goods with others.
- The social network consists of communities of traders which trade more frequently with each other.
- The social network allows a limited number of trader-pairs to trade between markets.

Representation as network data:

- Commercial actors are represented as nodes.
- The ability to share information and goods is represented as an edge.
- Community structure is represented by a high clustering coefficient.
- The limited contacts between communities is represented by edges between a low proportion of randomly selected pairs of traders.
- The tested degree of market integration is represented by edges between a variable proportion of randomly selected traders on different sites.
- The tested degree of the availability of reliable commercial information is represented by the ability of a node to obtain information and goods from a variable proportion of its neighbours.

Dependence assumptions:

- Trader pairs with a mutual contact will be more likely to become commercial partners themselves.
- Traders with contacts on different markets will have a better knowledge of prices and will have the opportunity to distribute tableware to different markets.

Network science techniques used:

- Exploratory network analysis: nodes, connected components, average degree, clustering coefficient, density, heterogeneity, node clustering coefficient, code degree.
- Confirmatory network analysis: ABM; Jin et al.’s (2001) social network model.

Range of distribution per ware: simulations

The boxplots in figures 49-58 present the simulated ranges of wares’ distributions of all experiments and compare these with the observed ranges per period. A boxplot summarises a lot of information and it is worth explaining here what they represent: stars represent outliers (i.e. exceptionally low or high values); the vertical line represents the total range of the distribution excluding outliers (i.e. the area between the minimum and maximum number of sites wares are found on); the grey box shows the interquartile range (i.e. the middle 50% of values); the horizontal line in this box represents the mean value. Each boxplot in these figures shows the combined results of 10 iterations of each experiment. This means I identified on how many sites each of the four wares was deposited on for each iteration, and then combined all fourty values per experiment to produce this boxplot. Each group of experiments with a different value for the *proportion-inter-site-links* variable is separated by a vertical dashed line.

Figure 49 presents only the simulated results of all experiments, not the observed distributions. We notice a clear trend in the wideness of wares' distributions between experiments. Low values for the *proportion-inter-site-links* variable (0; 0.00005; 0.0001) will give rise to an extremely limited distribution of wares, whilst high values (0.0003; 0.0006; 0.001) result in very widely distributed wares. This suggests that this variable strongly affects the wideness of wares' distributions. In contrast, the *local-knowledge* variable affects the wideness of wares' distributions very little. We notice only slight increases in the range of distributions as this variable is increased. However, we also notice very peculiar results for each experiment with this variable set to 1: the range is always lower than that of other experiments with the same value for *proportion-inter-site-links*. To ensure this is a robust trend and not merely a result of using the round number 1, I performed an additional experiment with *local-knowledge* set to 0.99. The results of this experiment shown in figure 50 demonstrate that a high *local-knowledge* does indeed result in a more limited distribution (only shown for *proportion-inter-site-links* = 0.0006, although similar results are to be expected for other settings of this variable).

These combined results of all experiments suggest that the *proportion-inter-site-links* variable has a stronger effect on the wideness of wares' distributions (as suggested by Temin's hypothesis), and that the *local-knowledge* variable is unlikely to give rise to strong differences in the distribution of wares (contra Bang's hypothesis).

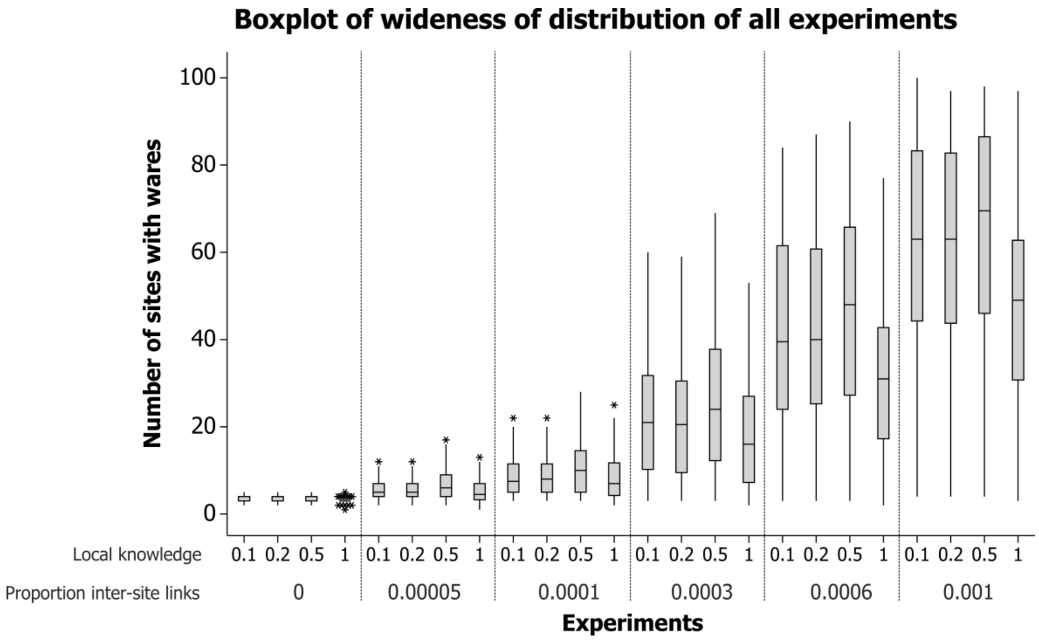


Fig. 49. Boxplots show the range of the number of sites simulated wares in different experiments were deposited at (each boxplot shows the combined results of 10 iterations of the experiment). The vertical lines divide the results into experiments with different values for proportion-inter-site-links.

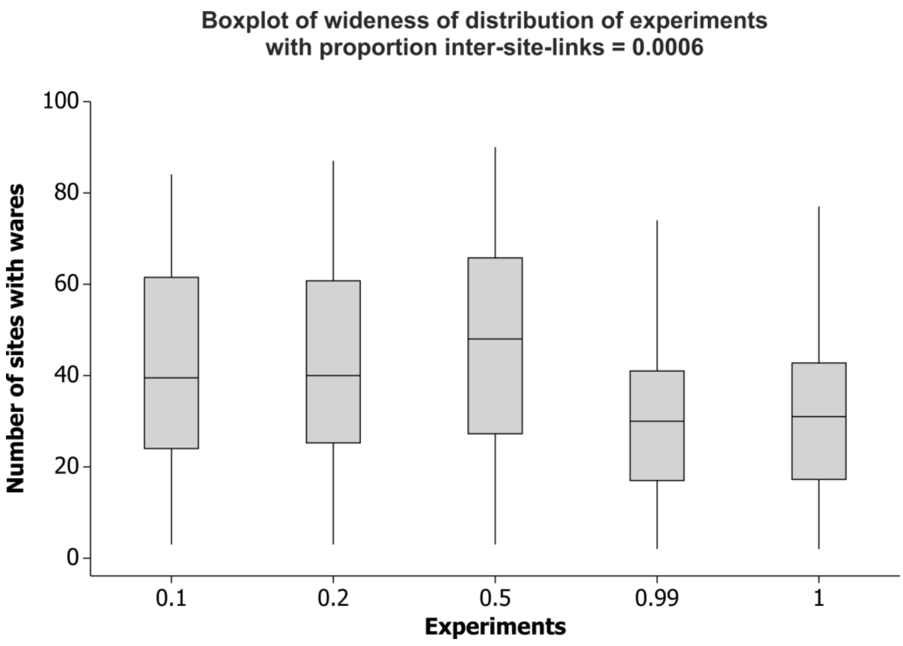


Fig. 50. Boxplot of the experiments with *proportion-inter-site-links* set to 0.0006. The additional experiment with local-knowledge = 0.99 indicates that the downward trend with high values for this variable is real and not merely a product of the value 1. Boxplots show the range of the number of sites simulated wares in different experiments were deposited at (each boxplot shows the combined results of 10 iterations of the experiment).

However, when interpreting these results it is crucial to keep one important limitation in mind: these are combined results of only ten iterations. Indeed, we notice an extremely high variance between different iterations of the same experiment, in particular for experiments with a value for *proportion inter-site links* higher than 0.0001. Figure 51 illustrates this for one experiment: we see that iteration 8 has a range of only 15, whilst iteration 5 has a range of 83. Moreover, the ranges of some iterations do not even overlap (e.g. iterations 6 and 7). This suggests that the wideness of wares distributions' varies greatly depending on the setup of the social network, which in this model is different for every iteration. Since the same ten randomization seeds were used for the ten iterations of each experiment, the results of the experiments are nevertheless comparable. It can be concluded that the trend described in the previous paragraph seems robust but the variability between iterations needs to be better understood. Future work should therefore perform a much higher number of iterations to check whether the results are normally distributed, and it should evaluate what factors give rise to such strong differences between iterations. The latter is done at the end of this section.

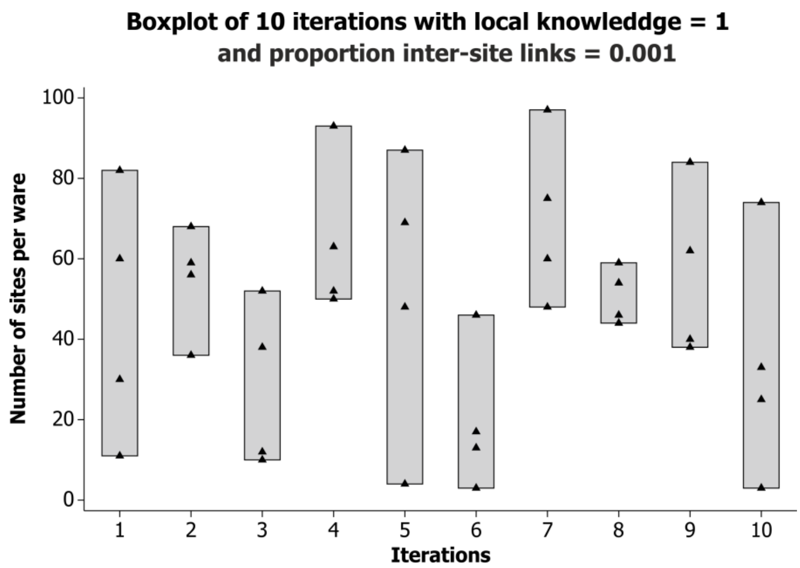


Fig. 51. Boxplot showing the results of ten iterations of a single experiment with local-knowledge = 1 and proportion-inter-site-links = 0.001. Grey boxes represent the range of simulated distributions, whilst each ware (data point) is also shown as a black triangle. Please note the strong differences in ranges within this one experiment, suggesting that a higher number of iterations is necessary.

Range of distribution per ware: simulations VS observations

In this section I will compare the simulated distribution of wares discussed above with the distribution observed in the ICRATES database per 25-year period. I will only do this for the seven periods between 25BC and 150AD during which four Eastern produced wares circulated in the Eastern Mediterranean, since these are comparable with the four simulated wares. These periods show very different trends: from very strong differences due to the ubiquity of ESA, to smaller differences due to the demise of ESA and the increased circulation of other wares. These trends will be compared with the simulation results in a descriptive way, because the small number of iterations per experiment, the strong variance of the results, and the fact that the observed range consists of only four data points (one for each ware) suggest a statistical goodness of fit test would be inappropriate.

The difference between the number of sites wares are present on is greatest in the periods 25BC-25AD (Figs. 52-53). The interquartile ranges of experiments with *proportion-inter-site-links* values of 0.0003, 0.0006, and 0.001 lie completely within this range (for 0.0003 only in the period 25-1BC). In the period 25-50AD the difference in the wideness of wares' distribution is still large, but much more limited than for the previous two periods, and now the experiments with *proportion-inter-site-links* = 0.0006 seem to fit best within this range. In the subsequent four periods the difference in the wideness of distribution becomes increasingly limited. The periods 50-100AD still show a good fit with the experiments with *proportion-inter-site-links* = 0.0006. None of the experiments provide comparable results with those of the periods 100-150AD. It is important to note that experiments with a low value for the *proportion-inter-site-links* variable show results very dissimilar to the observed distribution of tablewares. Even for periods where the range of distribution is very limited (100-150AD), the number of sites tablewares are found on is much higher and therefore not comparable with the simulations.

It can be concluded that a high proportion of inter-site links is needed to give rise to the differences in the wideness of wares' distributions as observed in the ICRATES database, that the model results shows a favourable fit of some experiments with the observations for the period 25BC-100AD, and that the model does not succeed in reproducing the distributions observed for the period 100-150AD.

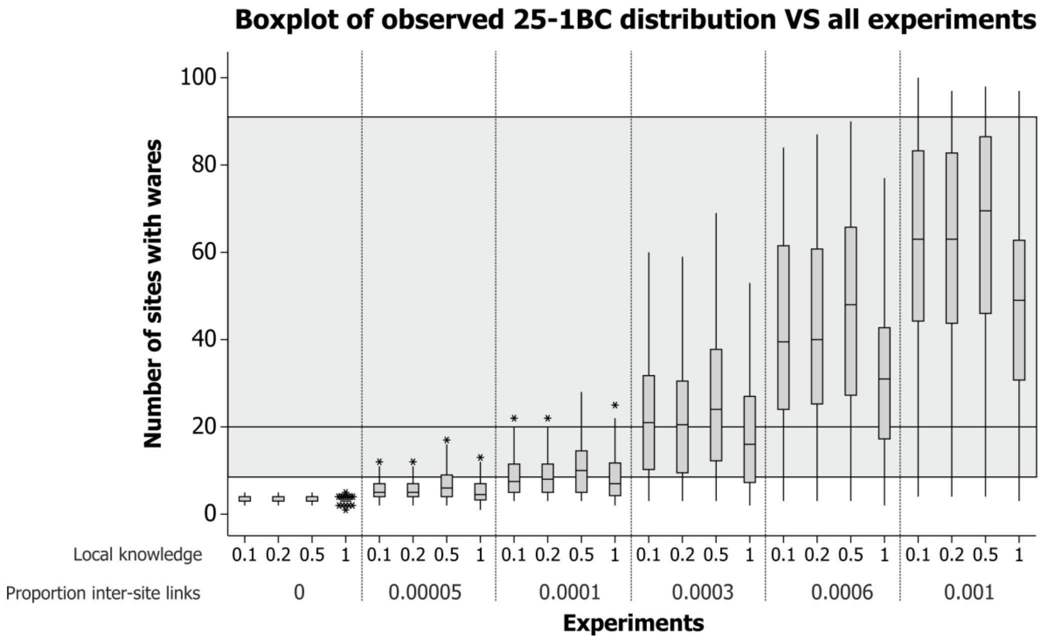


Fig. 52. Observed distribution of wares for the period 25-1BC (grey box in background) compared to the simulated distributions (boxplots). Boxplots show the range of the number of sites simulated wares in different experiments were deposited at (each boxplot shows the combined results of 10 iterations of the experiment). The grey box represents the observed minimum, maximum and average distribution of four wares for the period 25-1BC.

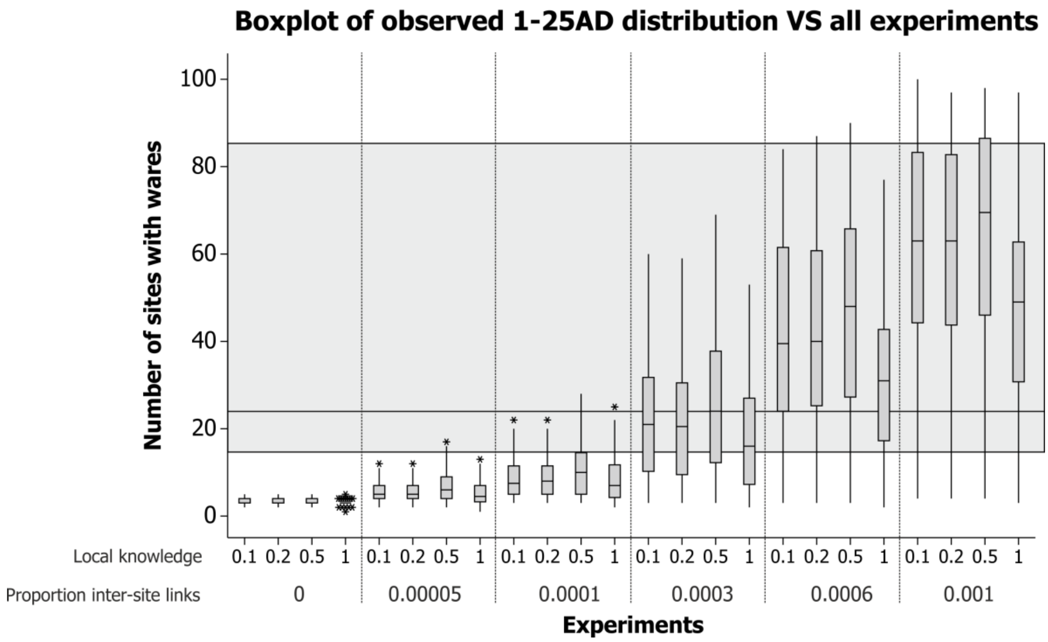


Fig. 53. Observed distribution of wares for the period 1-25AD (grey box in background) compared to the simulated distributions (boxplots). Boxplots show the range of the number of sites simulated wares in different experiments were deposited at (each boxplot shows the combined results of 10 iterations of the experiment). The grey box represents the observed minimum, maximum and average distribution of four wares for the period 1-25AD.

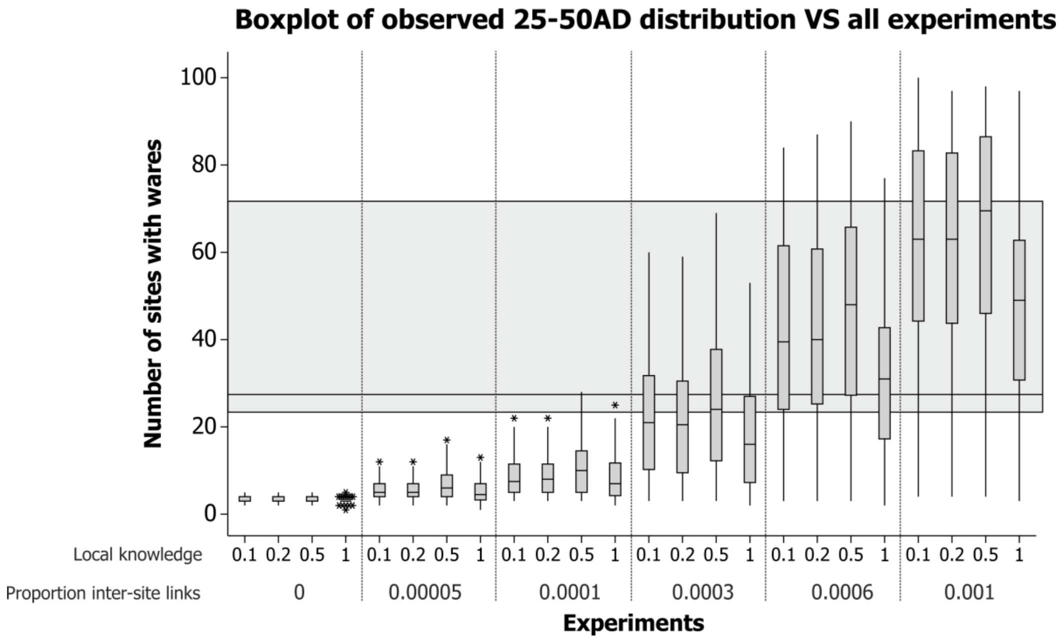


Fig. 54. Observed distribution of wares for the period 25-50AD (grey box in background) compared to the simulated distributions (boxplots). Boxplots show the range of the number of sites simulated wares in different experiments were deposited at (each boxplot shows the combined results of 10 iterations of the experiment). The grey box represents the observed minimum, maximum and average distribution of four wares for the period 25-50AD.

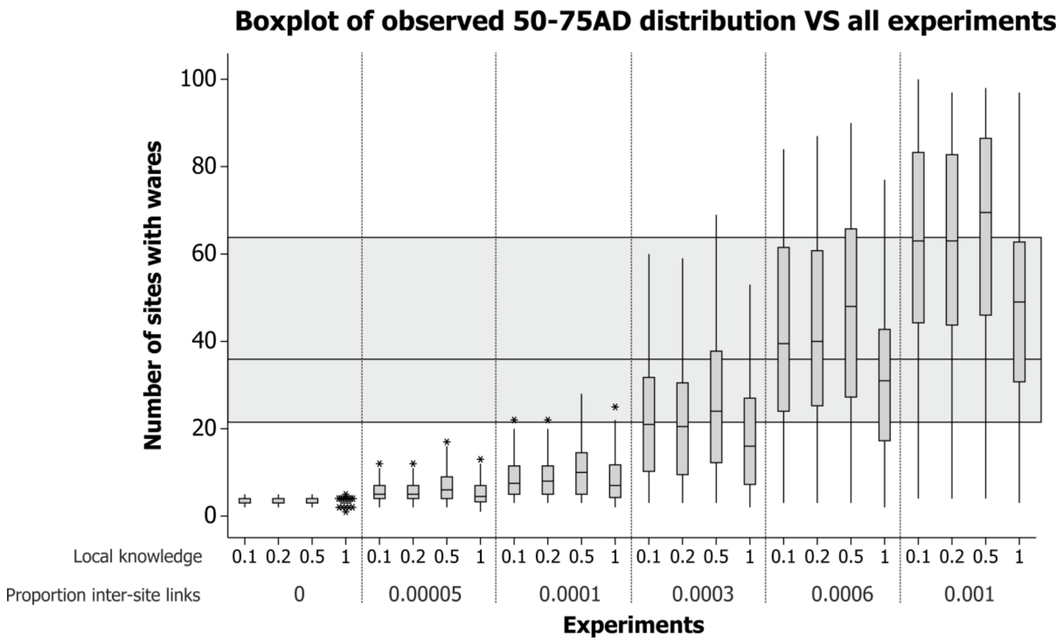


Fig. 55. Observed distribution of wares for the period 50-75AD (grey box in background) compared to the simulated distributions (boxplots). Boxplots show the range of the number of sites simulated wares in different experiments were deposited at (each boxplot shows the combined results of 10 iterations of the experiment). The grey box represents the observed minimum, maximum and average distribution of four wares for the period 50-75AD.

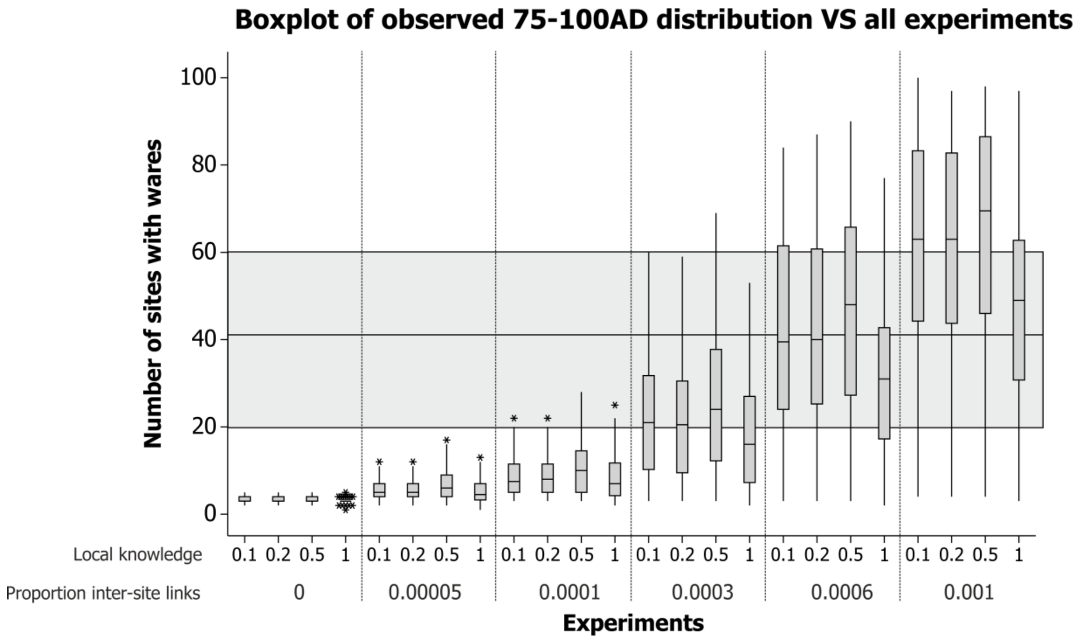


Fig. 56. Observed distribution of wares for the period 75-100AD (grey box in background) compared to the simulated distributions (boxplots). Boxplots show the range of the number of sites simulated wares in different experiments were deposited at (each boxplot shows the combined results of 10 iterations of the experiment). The grey box represents the observed minimum, maximum and average distribution of four wares for the period 75-100AD.

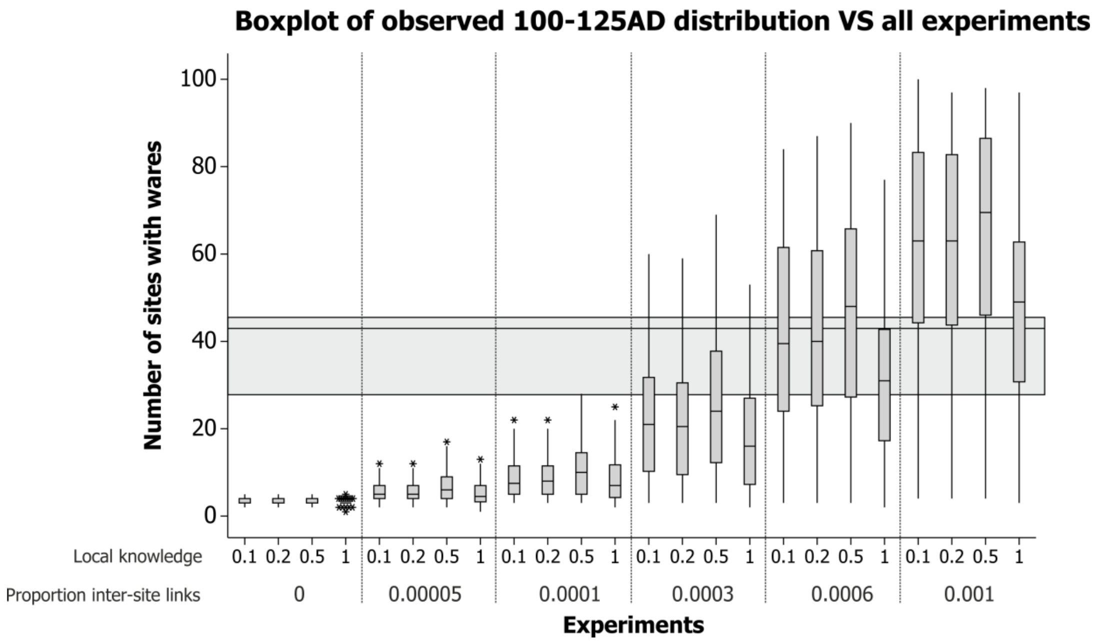


Fig. 57. Observed distribution of wares for the period 100-125AD (grey box in background) compared to the simulated distributions (boxplots). Boxplots show the range of the number of sites simulated wares in different experiments were deposited at (each boxplot shows the combined results of 10 iterations of the experiment). The grey box represents the observed minimum, maximum and average distribution of four wares for the period 100-125AD.

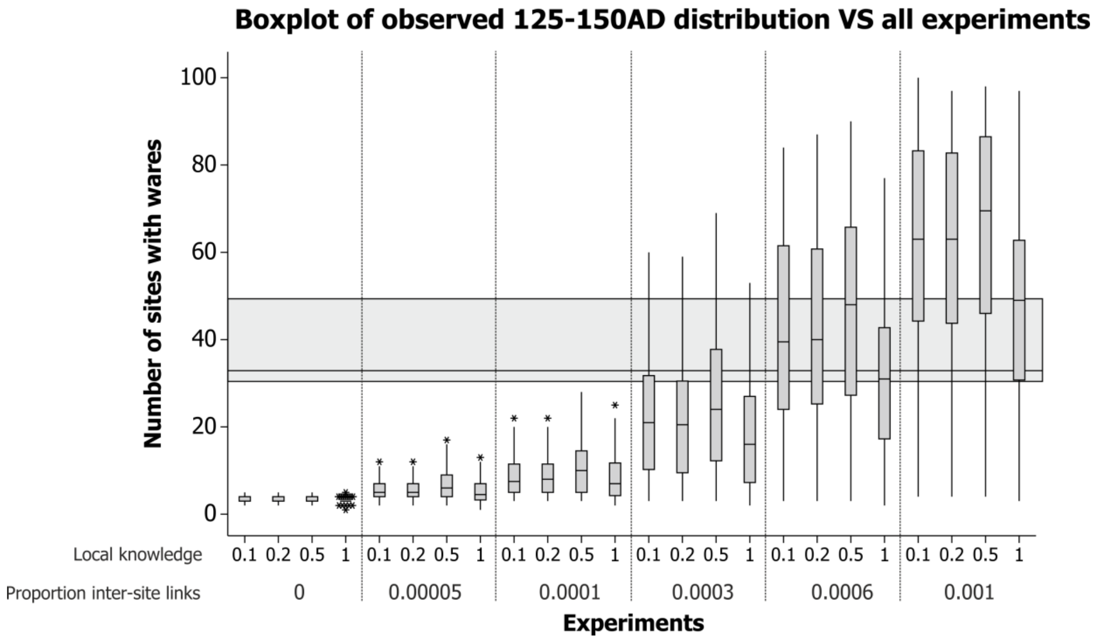


Fig. 58. Observed distribution of wares for the period 125-150AD (grey box in background) compared to the simulated distributions (boxplots). Boxplots show the range of the number of sites simulated wares in different experiments were deposited at (each boxplot shows the combined results of 10 iterations of the experiment). The grey box represents the observed minimum, maximum and average distribution of four wares for the period 125-150AD.

Frequency distributions: simulated assemblages VS observed assemblages

In section 5.4.7 above I described three frequency distributions of the observed tableware data: the number of wares per site, the proportion of sherds per ware in site assemblages, and the proportion of types per ware in site assemblages. These will now be compared with simulated frequency distributions of the experiments which showed some similarity with the observed distributions, i.e. experiments with *proportion-inter-site-links* of 0.0003 or higher. The results of experiments with a lower value for this variable are very dissimilar to the observed frequency distributions and are therefore not presented here. Given the strong variation between individual iterations this comparison is done descriptively.

The results suggest that both the frequency distributions of the number of wares (e.g. Fig. 59) and the frequency distribution of the percentage of wares in assemblages are very different from the observations of all periods. In most experiments the number of sites with only one ware is greater than those with more wares, but the differences are not as striking as with the observed distributions (see Fig. 44 in section 5.4.7). The simulated percentages of wares on sites differ greatly between individual iterations. There is no identifiable tendency for one ware to be the dominant ware in the majority of sites, as was seen in the observed distributions (see Figs. 45-46 in section 5.4.7), and in future work a higher number of iterations should be performed to allow for comparison with the percentages of wares in site assemblages. It can be concluded that although the model might approximate the strong discrepancy in the wideness of wares' distributions in certain experiments, it does not succeed in reproducing the observer trends in the number of wares per site and the percentages of wares in site assemblages.

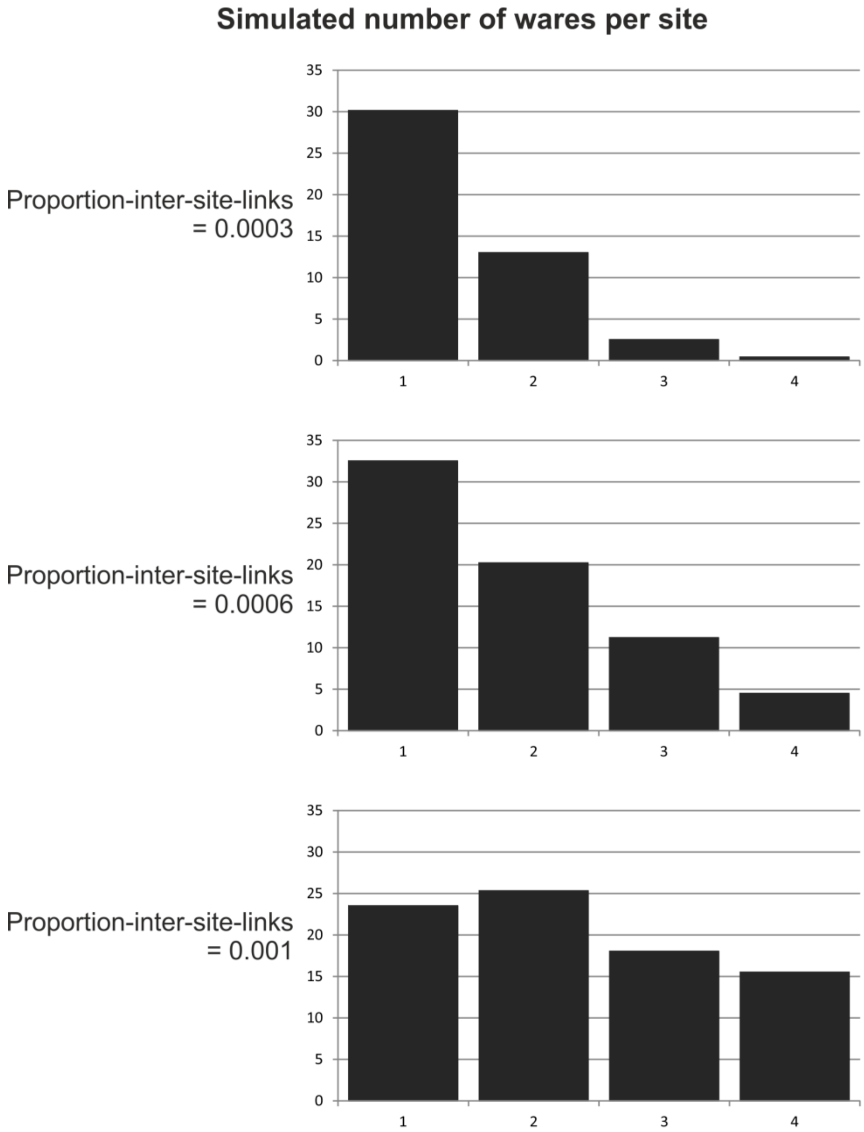


Fig. 59. Example frequency distributions showing the number of sites (Y-axis) on which a certain number of wares (X-axis) is attested, for the combined results of three experiments (local-knowledge = 1) with a proportion-inter-site-links value of 0.0003, 0.0006, 0.001.

Simulated volume of wares at sites VS social networks

In this final results section I will compare the simulated total volume of wares deposited on sites with some feature of the simulated social networks, for two reasons. Firstly, the approach to testing Bang's and Temin's hypotheses taken here strongly depends on my abstraction and representation of their hypotheses as social networks. The model includes a series of procedures to create social networks with a structure that I believe to be a suitable representation of these hypotheses. However, these procedures are complicated and the structural features of the resulting social networks should be studied to understand exactly what features give rise to the outcomes of research interest. Secondly, the previously discussed results indicate a high degree of variation that cannot merely be explained by the two tested variables. Due to the high number of procedures in the model, many of which include a measure of stochasticity, it is hard to identify exactly what causes this variation. However, in order to improve this model in future work its functioning will need to be better understood. In this section I will explore to what extent variation in one output, the total volume of tableware at a site, can be explained by features of the simulated social networks. Due to the time limitations of this project I was not able to perform similar tests for other outputs.

I decided to perform a multiple regression analysis since I wanted to evaluate the influence of a number of features of the social network on generating the variation in the total volume of tablewares per site. A multiple regression analysis produces a model that describes the statistical relationship between multiple independent variables (in this case the features of the social networks) and one dependent variable (in this case the simulated volume of tablewares on sites). The expected outcome of this model is a description of how the volume of tablewares changes as a result of a change in any one of the independent variables, and an indication of how much of the variation in the volume of tablewares on sites is explained by this model.

The following features of the simulated social networks were used as independent variables: average closeness, maximum closeness, average betweenness, maximum betweenness,[30] number of inter-site links, and number of traders at a site. The centrality measures were used because I believe it might be the case that if traders are closer to all other traders in the network or are important intermediaries, then they will obtain more tableware. Moreover, the number of traders at a site might correlate with the accumulated volume, as well as the total number of inter-site links of all traders per site.

[30] Please note that I decided to use the average and maximum of these centrality values rather than individual values because the centralities are not a feature of the sites themselves (which are the entities of analysis here), but rather of the traders located at sites.

A multiple regression analysis was performed for each experiment, using the combined results of ten iterations per experiment. In order to create a model one first needs to evaluate whether the independent variables themselves do not have a high correlation, because if they do then it will be unclear exactly which independent variable explains the variance of the dependent variable. To do this, a correlation matrix was created for each experiment where each cell shows the Paerson correlation coefficient for a pair of independent variables (e.g. Table 20). The correlation matrices were very similar for every experiment: they showed a high correlation among centralities, and a very high correlation between the number of traders and the number of inter-site links. I therefore decided to create multiple regression models with two independent variables: inter-site links and average closeness. Inter-site links was chosen over the number of traders, since this feature is key to understanding the hypotheses tested. Closeness was chosen over betweenness since I believe the distance from a trader to every other trader in the network might be more indicative for the volume of tableware accumulated at a site.

Table 20. Example of a correlation matrix for the experiment with proportion-inter-site-links = 0.001 and local-knowledge = 0.1. Rows and column headers show the independent variables, cells show Pearson's correlation coefficient. Very high correlation values are underlined and in grey boxes. Results were very similar for all experiments. To avoid these high correlations between independent variables it was decided to perform the multiple regression analysis with the average closeness and inter-site links variables.

	AvCloseness	MaxCloseness	AvBetweenness	MaxBetweenness	InterSiteLinks
MaxCloseness	0.665				
AvBetweenness	0.76	0.634			
MaxBetweenness	0.419	0.831	0.581		
InterSiteLinks	0.394	0.501	0.242	0.423	
NumberTraders	0.331	0.457	0.18	0.376	0.987

The multiple regression could not be performed for the experiments with values for the *proportion-inter-site-links* variable of 0 and 0.00005, since these did not satisfy the normal distribution and independence of residuals assumptions of the technique. In other words, multiple regression is not a good technique to evaluate the results for these experiments because they do not fulfil the necessary assumptions. The results for all the other experiments are presented in table 21, we notice that the inter-site links variable is significant for all experiments, and average closeness only in a few cases (experiments with *local-knowledge* = 1). Although the same model could be created for each experiment, there is a very strong discrepancy in the proportion of variance explained. This discrepancy is greatest between experiments with different values for *proportion-inter-site-links*: for low values of this variable very little variance is explained by the model suggesting that there must be other untested independent variables that give rise to this variance, and for high values up to and more than 50% of variance is explained suggesting that other untested independent variables might explain a minor part of the variance. There is also a small difference between experiments with a different value for *local-knowledge*: a higher proportion of the variance is explained with high values for this variable, suggesting that a better knowledge of possible trade contacts on other sites will give rise to higher volumes of accumulated tablewares. In all cases the coefficient for inter-site links is positive, which means that if the number of inter-site links increases then the volume of tablewares at a site will increase. However, where it is significant the average closeness coefficient is negative, indicating that if the average closeness increases then the volume of tablewares at a site will decrease.

This multiple regression analysis confirmed the importance of the number of inter-site links for explaining the variation in the total volume of tablewares at sites, in particular for experiments with a high *proportion-inter-site-links*. Moreover, higher values for *local-knowledge* further increase the importance of the number of inter-site links. Combined with the analysis of the wideness of wares' distribution, this seems to suggest that the *local-knowledge* variable affects both the wideness and volume of distribution where a better knowledge of inter- and intra-site links will lead to a slightly more limited spread of a ware but higher volumes on sites. However, the variable determining the number of inter-site links, *proportion-inter-site-links* succeeds far better in explaining variation in both wideness and volume of distribution. All results also indicate that these tested features of the social network do not suffice for explaining the variation observed in the simulated output. In future work other features of the models should be scrutinized to better understand its workings.

Table 21. Results of the multiple regression per experiment, where the aim is to explain the variation in the total volume of tableware deposited at a site at the end of the simulation (response variable) by comparing it to features of the social network (predictor variables). The first two columns provide the variable settings of the experiment, the third column gives the independent variables included in the multiple regression model for that experiment (* indicates a significant effect), and the last column indicates what proportion of the variation in the total volume is explained by the social network variables (R-Sq).

proportion-inter-site-links	local-knowledge	Variables included in multiple regression model	Proportion of variance explained (R-Sq)
0.0001	0.1	inter-site-links*; av-closeness	7.5
	0.2	inter-site-links*; av-closeness	7.6
	0.5	inter-site-links*; av-closeness	10.5
	1	inter-site-links*; av-closeness	13.9
0.0003	0.1	inter-site-links*; av-closeness	21.8
	0.2	inter-site-links*; av-closeness	21.4
	0.5	inter-site-links*; av-closeness	24.9
	1	inter-site-links*; av-closeness*	34.6
0.0006	0.1	inter-site-links*; av-closeness	40.9
	0.2	inter-site-links*; av-closeness	43
	0.5	inter-site-links*; av-closeness	48
	1	inter-site-links*; av-closeness*	54.8
0.001	0.1	inter-site-links*; av-closeness	57
	0.2	inter-site-links*; av-closeness	57
	0.5	inter-site-links*; av-closeness	59.8
	1	inter-site-links*; av-closeness*	63.6

5.5.3. Conclusions

A number of conclusions can be drawn from these results, which will be discussed in more detail in the next section:

- The *local-knowledge* variable has a limited effect on the wideness of tableware distribution.
- The *proportion-inter-site-links* variable has a strong effect on the wideness of tableware distribution.
- Limited commercial knowledge can still give rise to wide differences in distributions, but only in systems with highly integrated markets. This means that the *local-knowledge* variable is not instrumental in giving rise to the pattern of interest, whilst the *proportion-inter-site-links* variable is.
- Limited availability and high uncertainty of information, and a weak integration of different markets in an economy governed by supply and demand, is unlikely to give rise to large differences in the distribution patterns of tableware.

- A strong integration of different markets and the availability of commercial information from different markets, irrespective of the uncertainty and availability of information, is likely to give rise to large differences in the distribution patterns of tablewares.
- The model succeeds in simulating differences in the wideness of tableware distributions similar to the observed ones.
- The model does not succeed in reproducing the observed frequency distributions of the number of wares on sites and the proportion of wares in site assemblages.
- This model will therefore need to be reworked, and the analysis of the behaviour of this model has revealed the following aspects deserve more attention in future versions: the number of iterations need to be much higher; the conceptualisations of demand and time need to be more realistic; other aspects of Bang's and Temin's hypotheses need to be compared with the model's workings (e.g. correlation of prices with distance).

5.6. Discussion

In this case study I have illustrated that a method combining exploratory data and network analysis, and agent-based network modelling can aid our understanding of the processes giving rise to an archaeologically observed distribution pattern of tablewares. A large number of very diverse results were obtained, and their relevance to answering the research questions set out at the start of this case study now needs to be made clear.

> Question 1: what differences can be observed in the distribution patterns of different tablewares and forms (here considered modern analytical constructs)?

The first research question was addressed through exploratory data and network analysis in section 5.4. The distribution patterns of the four tablewares studied in this case study changed through time, but a number of general trends can be identified. ESA knew an exceptionally wide distribution until at least 75AD. Moreover, ESD and ESB knew a similar but more limited distribution to ESA in the period 25BC-25AD. This similarity decreases after 25AD, although ESA remains more widely distributed until 75AD. It is only after this period that ESA, ESB and ESD distributions again show stronger similarities. ESC retains a specific regional distribution throughout this period, its distribution being dissimilar and more limited than that of the other wares, although its distribution shows some similarity to that of ESA and ESB in the period between 75AD and 125AD. In addition to these general trends, a detailed description of the similarity of distribution patterns of different forms identified those forms whose distributions are more similar to those of other wares. This suggests that the processes that gave rise to these distributions should not be seen as exclusively ware-specific. It is true that one persistent pattern is the higher volume and diversity of a ware close to its production area, and the hypothesis that the distribution of a wide range of a ware's production in the region around its production area was distributed through similar processes is not unlikely. However, focusing exclusively on the core areas of a ware's distribution will not allow one to understand such exceptional cases that suggest similar distribution processes for different wares, and to what extent different wares were considered different products in the past.

Sadly but importantly, the exploratory analysis performed here also added a sobering note to our ability to answer such interesting questions using our current knowledge of tableware distributions. Since forms were considered modern analytical constructs, the differences in their distributions can hardly be considered a direct reflection of differences in past human behaviour. Moreover, the strong variability in the number of forms and sherds per form recorded at different sites did not allow me to use this evidence to perform a statistical goodness of fit test to evaluate hypotheses of distribution processes, unless the sampling bias is more thoroughly quantified (an exercise that was sadly not possible within the framework of this project). However, this was by no means considered a reason to abandon hypothesis testing for Roman archaeology altogether. It merely required me to focus my efforts on that general trend in the data that was considered a robust pattern: the strong differences in the wideness of tablewares' distributions. This pattern can be compared descriptively with the outcomes of simulations and allows for the falsification of hypotheses that cannot possibly give rise to this general trend.

> Question 2: can Bang or Temin's hypotheses be falsified through computational modelling and comparison with the observed tableware distributions?

This close scrutiny of the available data influenced the creation of a computational model to express and possibly falsify two hypotheses surrounding the importance of the structure of social networks between traders in giving rise to differences in the distribution patterns of tablewares. I believe the results of this model allow me to reject Peter Bang's claim that weak integration of markets and limited availability of reliable information can give rise to differences in the distribution patterns of products (in this case tablewares). My conceptualisation of the availability of commercial information gave rise to only minor differences in the wideness of wares' distributions. Conversely, differences in my conceptualisation of the integration of markets did give rise to such differences. The experiments with a high proportion of inter-site links in particular showed results similar to the observed wideness of tablewares' distributions in the period 25BC-25AD. This leads me to conclude that limited availability of commercial information can still give rise to wide differences in distributions, but only in systems with highly integrated markets. This means that the integration of markets is more instrumental in generating the pattern studied here: strong integration of different markets and the availability of commercial information from different markets, irrespective of the uncertainty and availability of information, is likely to give rise to large differences in the distribution patterns of products. I believe these results add more credibility to the tested aspects of Temin's hypothesis, and succeed in falsifying the tested aspects of Bang's hypothesis.

The results also shed some light on the specific ways in which my implementation of Temin's and Bang's hypotheses function. The results suggest that only a very strong integration of the markets will give rise to differences in the wideness of tableware distributions. Since the variable settings have so far only been discussed as abstract proportions, it is worth to clearly state what these proportions of market integration mean in real terms. The setup procedure used here creates links between 0.25% of all node pairs, with some variation due to the stochasticity in the model. The tested values for the *proportion-inter-site-links* variable will create links between 0%, 0.005%, 0.01%, 0.03%, 0.06%, and 0.1% of all node pairs. This means that the links added to enforce the integration of markets is a very high proportion of the actual number of links created in some experiments: 0%, 2.1%, 4.1%, 12.3%, 24.6%, and 40.5% respectively for *proportion-inter-site-links* values of 0, 0.00005, 0.0001, 0.0003, 0.0006, and 0.001. So in the most extreme scenario this means that 40.5% of all transactions could potentially be cross-market transactions. Moreover, the results suggest that only a very high integration of markets can give rise to strong differences in the wideness of wares' distributions, and could explain the deposited volumes at sites (see multiple regression in section 5.5.2). The experiments with 12.3% to 40.5% of all links being inter-site links showed the most similarity with the observed distributions of the periods 25BC-25AD, which suggest a very strong integration of markets and a large proportion of all transactions in the Roman trade system taking place between markets. However, this does not mean that transactions in local markets did not matter. The beauty of this approach is that it allows for the existence of local trade within individual markets as well as trade between markets, thus avoiding one-dimensional explanations of complex phenomena and identifying the grey-zone between hypothesised contributing factors.

The results also shed light on the impact of the availability of reliable commercial information. It would be wrong to state this factor has no role to play in a model of Roman trade, indeed the results of my experiments suggest that it gave rise to minor changes in the wideness of distribution of tablewares. However, an interesting trend is seen in the differences between the settings of the *local-knowledge* variable of 0.1 and 0.5, which result in more or less similarly wide distributions, and a setting of 1 which results in a drop in the wideness of distribution. It seems that the way in which the availability of information affects the wideness of distribution is as follows: the more a trader knows the commercial situation in his direct surroundings in the social network, the less he is inclined to take risks that could lead to wider distributions of products. However, this trend is minor compared to that seen for the increasing integration of markets. More importantly, the necessity for this strong degree of market integration and the limited impact of availability of information suggest to me that future modelling work should focus on the factors that allow for or drive market integration, some of which have been discussed in section 5.5.2 above: larger commercial entities with representatives on multiple markets, transport mechanisms and infrastructure, the gravitational pull of large consumption centres or institutions (such as the army), and regulation and stimulus provided by a political system. The results presented here reveal very little about these important factors, other than: the ability to share commercial information and goods between markets in relatively few steps is more important for explaining variance in tableware distributions than the proportion of available information. Again, none of these alternative factors should be considered as working in isolation, but should always function within systems where a high proportion of transactions are intra-market transactions.

It should be clear that this model had a very tight focus and therefore is necessarily simplifying. This analysis has focused on, to use the distinction made by Arnaud (2005, 11) mentioned above, the flux of goods rather than on their transportation. The specific route the object travelled was not tested, rather how it got from place of production to place of deposition through transactions. These two factors are of course interlinked and the ability to address them depends on the conceptualisation used. Minor modifications to the model might allow for the introduction of transport costs and making distinctions between markets that are geographically distant.

The model also did not distinguish between types of traders, like those defined by Broekaert (2013b) introduced in section 2. That is to say, only a few "types" of traders existed in the model: traders were either able or not able to trade across markets (depending on their position in the social network), whilst transactions invariably concerned one item at a time. Nevertheless, the results obtained do allow a few minor statements about the roles and types of traders, and their proportional contributions to the pattern of interest. Firstly, a very large proportion of all trade must have taken place between traders within markets: in this model between 88 and 60%, depending on the degree of market integration which was considered realistic. This suggests that we cannot just focus on those big entrepreneurs active in inter-regional exchange on a large scale, such as the *negotiatores* and some *mercatores*, but rather the more casual day to day transactions between the 'little fish' within a market are what makes up the bulk of all trade, e.g. possibly by those merchants referred to as *propolae*. Secondly, between 12 and 40% of all transactions are inter-market trade, suggesting that a large proportion of intermediaries needed to be able (financially and logistically) to perform such transactions. However, the model does not allow for realistic estimates of the number or proportion of such traders. It might be that similarly wide distributions would be achieved in systems with a much larger number of traders, a lower proportion of inter-site links and the same number of sites. This would indicate that it is not the proportion of all links between traders that matters but the proportion of sites connected through one or more inter-site links. Such a scenario would fit better with Bang's claim that whatever inter-regional trade and weak integration of markets did exist was caused by the specialisation of intermediaries. However, what this model does show, through the results of the multiple regression analysis, is that the *number* of inter-site links is strongly positively correlated with the volume of tableware on sites, i.e. the more pairs of traders are able to trade between a pair of sites, the higher the volume. Although this suggests that it is indeed the proportion of trader pairs that matters and not the proportion of site pairs which are connected, allowing traders to trade more than just one item in a transaction in the latter scenario might change this interpretation, i.e. when a trader sees the promise of a high profit due to higher demands on another market and high local supply, he can decide to buy a lot locally and sell it on the other market. Moreover, the simulated volumes were not compared with the volumes observed in the

ICRATES database due to sampling biases. However, the inability of the model to reproduce the observed frequency distributions of number of wares on sites might have been caused by omitting this factor: the simulated distributions showed a higher proportion of sites with a diverse assemblage than that observed. Could it be that experiments with a much lower number of trade links between a similar proportion of site pairs would show a better fit with frequency distributions of number of wares on sites? More experiments using a slightly modified conceptualisation are needed to critically address this possible alternative scenario.

Although it is not explicitly tested, this model does illustrate a statement that both Temin and Bang would agree with: markets are necessary intermediaries, small-scale transactions within markets are important, and state involvement cannot explain everything. However, it might well have been state stimulus and regulation that allowed for the high proportion of market integration. This would nevertheless only make up for a maximum of 40% of all transactions as suggested by this model: local intra-market transactions are necessary and more frequent.

But it would be wrong to argue that the results of the ABM succeed in burying one of the two tested hypotheses in favour of the other. The model presented here is a work in progress, and the main contribution of the results is their ability to guide future work on improving the model. It needs to be emphasised that the model only succeeds in reproducing the wideness of wares' distributions, and not the frequency distributions of the number of wares per site or the proportions of wares in sites' assemblages. Moreover, a number of procedures in this model are considered unrealistic, most crucial among them are my treatment of demand and time. To some extent this is a result of my attempt to keep the model simple as well as focused on answering a specific question. It was decided to simplify the production and consumption procedures, which are strongly linked to the demand. In future work these procedures need to be given more thought, in particular by only increasing demand after one to three years (as suggested by Peña's (2007) model of the use-life of tableware). A realistic implementation of demand will also allow for a more realistic implementation of time in the model. In its current state the model is not designed to recreate the changes observed in the archaeological record over time. Indeed, the results can only be argued to show similarity for the periods between 25BC and 25AD, when the differences in the wideness of tableware distributions are strongest. Moreover, the motivation of an ending point for experiments, which is currently rather arbitrarily set at 10,000 time steps, will benefit from a more realistic implementation of time. However, our ability to do so is plagued by our complete ignorance on the increase in demand and the time tableware remained in use; every implementation, even one based on Peña's estimates, will necessarily be based on an untested assumption. Although I argued no more than a descriptive comparison of the simulated output with the archaeological record could be performed, future work on quantifying the sampling bias in the ICRATES database will also allow for modifying the output of the model to incorporate this bias. Finally, the preliminary results presented here allowed me to identify the many ways in which the stochasticity in the model affects the outcomes, and will allow me in future work to limit some of this stochasticity or to further test its impact by setting random variables to constant values and observing difference in the output.

Another interesting result which could not be properly tested due to the narrow focus of the model is the strong positive correlation observed between the number of traders at sites, the number of inter-site links, and the total volume of tableware deposited at sites. It was decided to explore the ability for the number of inter-site links to explain the variation in the deposited volume, since this was the aim of the model. However, the correlation with the deposited volume and the number of traders at a site seems to lend credibility to the hypothesis that larger towns exercise a 'gravitational pull' on resources (an archaeological hypothesis introduced in section 5.2.3). Although this sounds obvious, it is nevertheless interesting to observe that this effect is very much visible in the model. Sadly, I cannot test this hypothesis because, unlike for the proportion of inter-site links, I did not modify the variable which determines the distribution of traders on sites in my experiments. An additional complication is the need to formulate experiments that allow me to distinguish between the effects of the inter-site links and the number of traders at sites. Further experiments are needed in future work in order to be able to claim that the 'gravitational pull' hypothesis is likely.

A number of suggestions can be made to slightly widen the scope of this model and to make its ability to falsify the tested hypotheses more robust. Firstly, a much higher number of iterations per experiment is needed in order to see whether the results, which currently show a high variance, become normally distributed or not. Secondly, currently items of different wares are valued similarly by agents. This can be modified to explore whether the same differences in distributions could arise when wares are considered as different products by agents. Thirdly, the trade procedures implemented in this model are very complex and it would be interesting to evaluate whether the same results could be obtained through a more simple process of diffusion (e.g. as implemented in the SIR network model of the spread of a virus, Newman 2010, 627-675). Fourthly, the effect of the currently implemented social network structure on the creation of the results would be even better understood by comparing the results with those of the same trade procedures taking place on a random social network with the same number of nodes and edges. Fifthly, would the same results be obtained when a transport cost is charged for transactions between traders located at different sites? Preliminary results of experiments that implement this suggest the distribution would be even more limited, even with a high proportion of inter-site links. Finally, a key aspect of Temin’s hypothesis is the correlation between the price of goods at a market and the distance from its place of production. This issue is connected to the previous one, since the price of a good is considered to increase with distance as a result of transport costs, which could be tested through new experiments. Most of these suggestions have already been built into the model, but time limitations did not allow for the necessary experiments to be performed and analysed.

A more drastic reconceptualisation could be suggested, one in which the integration of markets is not enforced but emerges through trade processes. This would require a more complex model in which traders can move between markets (without having prior knowledge of the commercial information at this market), to obtain new commercial contacts, and to end others. Such a model would be more similar to that of Bentley, Lake and Shennan (2005). The advantage of this would be that we would be able to test Bang's and Temin's claims that the integration of markets is an emergent property of the Roman trade system, rather than a given. However, I believe the current model was a necessary first step in order to evaluate the effects of the hypothetical trade procedures on social networks with a specific structure. Moreover, both Temin’s and Bang’s hypotheses require the social network of traders to be based on strong and stable connections. Due to the limited availability of reliable information and the sheer time it takes for messages to be sent across the empire, trusted long-term commercial contacts will need to have existed. I therefore consider the current conceptualisation a good representation of this, and suggest that future work should evaluate the effect of slow rates of network change on the results in an experiment.

One could criticise the approach presented here of cherry-picking comparable aspects of Bang’s and Temin’s hypotheses to allow for quantitative hypothesis testing. I strongly believe this is a good thing to do, but only if the results of a quantitative approach subsequently feed back into a richer culture-historical discussion. This case study has shown that the conceptualisation of a key aspect of Bang's hypothesis presented here is unlikely to give rise to the phenomena he claims they give rise to. Conversely, the conceptualisation of a key aspect of Temin's hypothesis presented here lends more credibility to this scholar's claims. However, I have also argued that these are preliminary conclusions and that they can only be considered robust through a refinement of the model in future work, as well as through comparison with alternative models that use different conceptualisations to express Bang's and Temin's hypotheses (e.g. Graham and Weingart 2014).

Box 10. Summary of chapter 5

Research questions:

- What differences can be observed in the distribution patterns of different tablewares and forms?
- Can Bang or Temin's hypotheses be falsified through computational modelling and comparison with the observed tableware distributions?

Conclusions:

- Exploratory data analysis revealed significant differences in the wideness of tableware distributions.
- Exploratory network analysis allowed for the identification of differences in the distribution patterns of individual forms, which confirmed the results of previous studies using the ICRATES dataset.
- Contrary to Bang's hypothesis, limited availability of reliable commercial information from different markets is unlikely to give rise to the large differences in the wideness of product's distributions observed in the archaeological record.
- The ABM does not succeed in reproducing all variability in the ICRATES dataset identified through the exploratory data analysis.
- The model will need to be reworked significantly by reducing stochasticity, implement more realistic conceptualisations of demand and time, and re-run experiments with a much higher number of runs.

Implications for this PhD project:

- The importance of clearly formulating how a past phenomenon of research interest was abstracted into network concepts became clear. This allows for formally representing one's hypothesis, as well as comparing different conceptualisations of the same phenomenon.
- In cases where dependence assumptions cannot be clearly defined, the potential for exploratory network techniques to lead to insights other exploratory data analysis cannot offer is more limited.
- When the archaeological datasets necessary to prove a hypothesis are not available, Roman archaeology hypotheses can still be falsified if the authors offer specifications of the kinds of data they would expect to see if their hypothesis was right.

CHAPTER 6

DISCUSSION:

ADDRESSING CHALLENGES IN ARCHAEOLOGICAL NETWORK SCIENCE

6. Discussion: addressing challenges in archaeological network science

6.1. Lessons learned from the case studies

The literature review and case studies have revealed a large number of challenges that archaeologists are confronted with when evaluating the potential contribution network science can make to achieving their research aims. Through three case studies I have explored some possible responses to these challenges, whilst being aware that many others remain unaddressed. In this section I will discuss this first set of challenges and the contributions this PhD made to overcoming some of them. In the first three parts of this discussion I will evaluate how the three case studies have allowed me to answer the PhDs methodological research questions, while the fourth section will focus on the implications for Roman archaeology specifically. The wider significance of the lessons learnt for the future use of network science in archaeology will then be discussed in the last few sections of this chapter.

6.1.1. Case study 1: methodological implications

In the first case study I have followed the network science research process explicitly to investigate the academic phenomenon of the adoption and adaptation of network science in archaeology. This led to specifications of how such phenomenon could be abstracted as network concepts, how these could be represented as network data, what sorts of dependence assumptions were driving change in this network data, and what network science techniques were suitable for exploring this data (Box 4). The results revealed a number of advantages and limitations to this approach. Many conclusions drawn from citation analysis could be identified through the literature review in chapter 2. However, a number of additional conclusions could also be obtained. These new insights are largely restricted to the limited influence of old archaeological applications on new ones, and the way in which archaeological publications cite literature outside their own discipline. The incompleteness of the available bibliometric data was considered a key limitation. Moreover, some results that were immediately obvious from the literature review were obtained through a very time-consuming network method that requires significant technical skills and, therefore, could lead one to question the usefulness of this approach. However, even though not all of the applied techniques necessarily revealed new insights, they did offer an alternative way of exploring a dataset, a phenomenon of interest, and my research questions. Such an exploratory process which forces one to think explicitly about the effect of relationships on the process of academic influence can lead one to ask new research questions, to validate the intuitive conclusions drawn from a close reading through a reproducible method, or to refocus analysis on previously neglected aspects of a dataset.

Interestingly, this leads me to conclude that network science can still be useful as a perspective that forces one to see a research problem in a different light even in cases where network science techniques do not lead to new insights and where techniques more commonly used in archaeology might be preferable. However, in line with the key claim of this PhD project stated in section 1.3, I believe that the real potential of network science techniques is revealed when it allows us to do things that no other approaches can do: address dependence assumptions through the analysis of network data. To give a concrete example, the results of the four techniques employed in the first case study, input domain, citation weights, cocitation and bibliographic coupling, enabled the evaluation of different types of dependence assumptions and allowed me to focus explicitly on the main ‘flows of influence’. In answer to the third research question of this PhD project, I believe the usefulness of network visualisations for communicating research results was clear in those cases with a limited number of nodes, and in particular those where arcs are an intuitively understandable representation of flow processes such as in the citation weights results (Figs. 15, 19, 21-22). Figures that include too many nodes did not allow me to make more than the most obvious statements, such as the density of citations and higher number of publications in the second half of the 20th century shown in figure 20.

6.1.2. Case study 2: methodological implications

I have introduced exponential random graph modelling as a method that has great potential for enhancing the study of archaeological network data, and of visibility network data in particular. I have argued that archaeological network scientists have tended to focus on exploring the outcomes of past phenomena, rather than on the processes that give rise to them. I believe that the dependence assumptions archaeologists formulate when representing archaeological data as networks is key to overcoming this problem, and ERGM offers a formal method for doing so. ERGMs allow archaeologists to use network data to evaluate existing hypotheses and formulate new hypotheses of the past processes that drove the phenomena they are interested in. New hypotheses can be focused on the more narrow range of processes, which the ERGMs suggest can lead to the creation of archaeological datasets.

The case study has also taught me that data quality and theory are crucial for formulating models of past phenomena. It is clear that ERGMs are only as reliable as the network datasets they are based on, suggesting that for some datasets (e.g. long-distance visibility networks) it is difficult to make statistical inferences about processes with much certainty. One needs to be confident that the observed network is as close as possible to complete, or one needs to make theoretical arguments why the created network can be used for testing one's hypotheses. Experiments of modelling a network with lines of sight up to 50km illustrated this most clearly: such network data did not allow for the evaluation of hypotheses concerning inter-visiblity or visual control among settlements, therefore the effects included in ERGMs were not suitable for explaining the network's creation. The impact of so-called 'edge-effects', i.e. determining the geographical boundary of the network of interest, and the impact of thresholding the network on the probability or distance of its lines of sight need to be evaluated more thoroughly in future work through sensitivity analyses (as suggested by Peeples and Roberts 2013). Another valuable experiment in future work would be to evaluate the use of ERGM for visibility networks by comparing the process suggested by an ERGM with an observed process of settlements established in a known order, within a research area where such detailed information is available. I also expressed caution with pushing the interpretations of the statistical models too far. As I have shown, different combinations of attributes in the models can lead to wildly different results. Moreover, every aspect of this combined network approach always requires a re-contextualisation within a cultural and historical context before the significance of its results become clear, as the discussion in section 4.6 illustrates.

The results of this case study have a number of implications for the methodological research questions of this PhD project. Firstly, some of the datasets archaeologists are confronted with translate more intuitively into networks than others, visibility networks are a particularly good example of this although their study is less common than that of road and river networks. This case study's exploratory network analysis can be considered similar to how previous archaeological studies have approached visibility networks, although a much wider range of network techniques was used here. However, I believe, this case study illustrates how the ability of an exploratory network analysis to enable a better understanding of past phenomena is significantly enhanced when dependence assumptions are explicitly formulated. Largely because when doing so the results of exploratory network science techniques cease to be mere numbers but carry the baggage of one's theoretical assumptions:

the results have specific implications for the inference of social processes governing the studied dataset. Secondly, this case study illustrated how an exploratory network analysis allows one to understand the structure of a network dataset, but such an approach reveals little about how it emerged. A confirmatory statistical network modelling approach shows greater potential for understanding such processes of network creation, although its ability to do so is intertwined with the quality and completeness of the dataset, as well as with one's arguments of the relationship between a dataset and a certain past phenomenon, i.e. how the observations in the dataset mirror past human behaviour.

6.1.3. Case study 3: methodological implications

This case study illustrated even more than the other two case studies the importance and impact of clearly formulating and arguing for a selected abstraction of the studied system in terms of network concepts. A different conceptualisation and representation of network data could have given rise to different results (e.g. see the network model of Bang's hypothesis by Graham and Weingart 2014). Indeed, the many archaeologists and historians who have joined the debate surrounding the Roman trade systems and whose diverse hypotheses have been introduced in section 5.2.3 use a wide range of concepts to abstract past trade. Hence, I believe that the ABM and network techniques used should be less the focus of future academic scrutiny of this model than the conceptualisation adopted here. The implementation I adopted can be easily criticised and I offer a number of methodological critiques that should be addressed in future work in section 5.6. Indeed, my model would have to be discarded if the Roman archaeology/history community disagrees with the conceptualisation used and offers evidence that an alternative conceptualisation is preferable. Regardless of how much time and effort I invested in this model, if it triggered discussion about the conceptualisations used in different hypotheses it would have served its purpose very well. This case study has illustrated that confirmatory network approaches aimed at falsifying hypotheses rather than making statistical inferences from archaeological datasets stand or fall with the argumentation surrounding the conceptualisation used.

A wide range of quantitative techniques were used in this case study, and it became clear that in some cases network techniques did not lead to new insights. This was most clear when performing both an exploratory data analysis and network analysis on the same dataset. The former is more commonly applied in archaeology. It leads to a critical understanding of the dataset and its limitations, and allowed me to answer in part the first research question of this case study concerning similarities of forms' distributions. The latter was based on a rather complicated technique that reproduced some of the results of the exploratory data analysis and added more detailed insights on the similarities of forms' distributions. However, this added knowledge was almost exclusively the result of a close scrutiny of the Brainerd-Robinson similarity matrices, whilst the network visualisations and the network measures revealed very little new information. I believe this again emphasizes the importance of dependence assumptions in archaeological network science. The assumption in the exploratory network analysis performed in this case study was that forms that have a similar distribution are more likely to have been distributed through similar processes. However, this is of little use as a dependence assumption unless the processes in question are more explicitly defined, which was not possible or even desirable in this case. My interpretation of the exploratory network results was therefore limited to an additional type of exploratory data analysis. Through this case study it became clear that the selection of a methodological approach for addressing a Roman archaeology research question needs to be well motivated, and that network science techniques are only worth applying if they add something unique to the mix of techniques, which is more likely to happen in cases where dependence assumptions can be clearly formulated.

6.1.4. Implications for Roman archaeology

The literature review indicated the use of network science in Roman archaeology was limited but highly creative, and it attributed this creativity largely to the availability of both material and textual datasets, which allowed archaeologists to address a wider range of complex research questions. Nevertheless, it was also argued that Roman archaeologists have neglected the development or adoption of methods that allow them to compare and falsify the complex hypotheses they formulate. Moreover, because of the availability of more data types as compared to the prehistoric periods, some common research themes in prehistoric archaeology have received limited attention, which in some cases creates the impression of a clean break between prehistoric and Roman periods caused by different research traditions. The second and third case studies confirmed these impressions and were specifically designed to make a few methodological contributions to help overcome them. They introduced network science methods that are rarely or never used in Roman archaeology, that are aimed at addressing complex hypotheses in Roman archaeology, and that take the nature of the available data into account in order to evaluate our ability to prove or falsify hypotheses.

The second case study argued that the potential role of lines of sight in structuring past human behaviour was an example of such a neglected research theme. It illustrated how the transition from the Iron Age to the Roman periods in Southern Spain could be studied through a long-term regional analysis of visibility networks. Moreover, hypotheses representing theories of how lines of sight potentially structured inter-urban interactions in the Iron Age were tested against the Roman settlement patterns. The results suggested that the structure of the visibility networks connecting urban settlements disintegrated from the Iron Age to the Roman period. However, this was a very gradual process of disintegration which only takes place from the Early Imperial period onwards. The ERGMs nevertheless suggested that the processes that gave rise to the Iberian and Republican period visibility network were very different from those that resulted in the Roman Imperial periods.

Two key implications of this case study for Roman archaeology should be mentioned. Firstly, the network approach used here proved fruitful to study long-term and large-scale past phenomena concerning visibility that are rarely addressed in Roman archaeology. It, therefore, showed that common research themes from other subdisciplines can still be of interest in Roman archaeology if good quality datasets are available to address them, and that such research themes allow for bridging subdisciplinary divides, and highlight the processes driving change between academically defined periods. Moreover, it emphasised the gradual nature of change and suggested a range of abstract processes that could have driven this change. Secondly, network science offers a good approach for addressing some of these themes and it can even be used for quantitative hypothesis testing, a practice which has been particularly rare in Roman archaeology. Moreover, the case study has shown that hypothesis testing in Roman archaeology does not necessarily have to draw scholars away from a close scrutiny of their datasets. The statistical modelling technique introduced here stands or falls with the motivations of data specialists of why they believe a given dataset allows them to better understand a certain past phenomenon.

The third case study offered an example of the kinds of complex hypotheses that Roman archaeologists and historians can formulate thanks to a wealth and diversity of data available to them. The different approaches, datasets used, and research traditions of archaeologists and historians resulted in a large number of sometimes conflicting hypotheses about the functioning of the Roman trade systems. But despite the wealth of data that allow for such a diversity of hypotheses to be voiced, few of them can be tested. In this case study I have illustrated how falsification of hypotheses through agent-based network modelling might offer a way forward. The conflicting hypotheses of Peter Bang and Peter Temin formed the focus of my analysis, one aspect of their models in particular: the degree of integration of markets in the Roman world. Bang argues that differences in the wideness of tableware distributions can be caused by trade systems with weakly integrated markets, whilst Temin argued the Roman trade systems knew more strongly integrated markets. These hypotheses were abstracted as network concepts, which allowed me to formulate specifications of the kinds of network data I believe best represents both hypotheses. The results of this case study allowed me to reject Bang's claim, thus falsifying one factor in his complex model of the Roman trade systems.

The implications of this case study for Roman archaeology are twofold. Firstly, the importance of making the abstraction of past phenomena as network concepts a transparent process, as mentioned already in section 6.1.3 above. Alternative models implementing alternative network conceptualisations could be suggested to prove me wrong, and I believe such practice should become more common in Roman archaeology in order to shed highly unlikely hypotheses and work towards stronger hypotheses (the high number of untested hypotheses surrounding the Roman economy discussed in section 5.2 is but one example of a research context which would benefit greatly from this approach). Secondly, the case study showed that even though the archaeological datasets necessary to prove a hypothesis are not available, Roman archaeology hypotheses can still be falsified if the authors offer specifications of the kind of data they would expect to see if their hypothesis was right. Peter Temin and to a lesser extent Peter Bang are very specific at times about these specifications, which makes parts of their hypotheses testable. If network science is to be more usefully applied to Roman archaeology research questions this has to become more common, since the true potential of network science seems to lie in its ability to deal with dependence assumptions requiring clear network data specifications. Moreover, I believe that such a formal hypothesis falsification approach holds great potential for Roman archaeology, in particular because of the availability of archaeological and historical datasets that allow for the formulation of complex hypotheses.

6.2. 'We shall overcome': challenges and suggested solutions

The case studies offer specific practical examples of the use of network science in different archaeological research contexts, and their contributions to answering this PhDs research questions are therefore most relevant to these particular research contexts. However, the three case studies were purposefully selected because they offer particularly interesting methodological challenges, which allowed innovative network science techniques to be applied to address some methodological lacunae in Roman archaeology. Although they are not representative for the full diversity of practice in the archaeological discipline as a whole, I believe it is possible to extrapolate the lessons learned through these case studies in order to approach the aim of my PhD: to gain a better understanding of the advantages and limitations of network science in archaeology, and of its position in the archaeological research process, as well as to suggest specifications for the future archaeological use of network science. This discussion of the implications of the case studies for archaeological network science will be structured around five topics: challenges related to method, data, space, process, and a brief note on past social networks.

6.2.1. Method

Like any other formal techniques in the archaeologist's toolbox (e.g. GIS, radiocarbon dating, statistics), network science techniques are methodological tools that work according to a set of known rules (the algorithms underlying them). These allow one to answer certain questions, and have clear limitations (what the algorithms are not designed to answer). This means that their formal use is fundamentally limited by what they are designed to do, and that they can only be critically applied in an archaeological context when serving this particular purpose. In most cases, however, formal network science results are not the aim of one's research; archaeologists do not use network methods just because they can. Instead one thinks through a networks research process about the past phenomena one is actually interested in. The literature review has shown that this is often done implicitly in archaeological network research, although in the case studies I have argued it is crucial to let the network science research process guide the development of a suitable method. When this research process is only implied it holds the danger of an epistemological issue that all archaeological tools struggle with: there is a danger that formal network techniques are equated with the past networks we are trying to understand (Isaksen 2013; Knox et al. 2006; Riles 2001).

In other cases, however, formal analysis is avoided altogether and concepts adopted from formal network methods are used to describe hypothetical past structures or processes (e.g. Malkin 2011). Although this sort of network thinking can lead to innovative hypotheses, it is not formal network analysis (see reviews of Malkin (2011) by Ruffini (2012) and Brughmans (2013b)). Good examples of this are Bang's and Temin's hypotheses addressed in the third case study: neither scholar worked explicitly from a networks perspective but concepts of connectivity between markets and traders were clearly key to both hypotheses. Indeed, a large number of hypotheses of the structure and functioning of the Roman trade systems have been proposed by archaeologists and historians, using a range of different conceptualisations, some of which translate easier into network concepts than others. This case study has shown that such alternative conceptualisations of the same phenomenon cannot be easily compared, unless specifications are formulated of how the concepts are represented in data. In cases where hypotheses are best abstracted in terms of network concepts (such as the tested factors of Bang's and Temin's hypotheses), formulating such specifications allows for the use of confirmatory network approaches to compare and falsify hypotheses.

However, when scholars argue that concepts were adopted from formal network science (as in the case of Malkin (2011)), they must recognise that these concepts have a very specific meaning to network scientists and are associated with specified data requirements in order to represent them. Most crucially, when the concepts one uses to explain a hypothesis cannot be demonstrated through data (not even hypothetically through simulation), there is a real danger that these concepts become devalued since they are not more probable than any other hypotheses. Moreover, the interpretation of past social systems runs the risk of becoming mechanised when researchers adopt the typical interpretation of network concepts from the SNA or physics literature without validating their use with archaeological data or without modifying their interpretation to a particular archaeological research context. This criticism is addressed at the adoption of formal network concepts only. It should be clear that other theoretical concepts could well use a similar vocabulary whilst not sharing the same purpose or data requirements, in which case I would argue to refrain from using the same word to refer to different concepts or explicitly address the difference between these concepts in order to avoid confusion. The archaeological use of actor-network theory (ANT) is a good example of a highly confusing use of metaphors because, just as I argue is the case for archaeological network science, archaeologists have not yet given ANT a clear place in the archaeological research process. Archaeologists thinking through an ANT perspective commonly use network concepts, which creates the impression that ANT and network science are somehow similar, or that they offer a compatible method and theory for a homogeneous research perspective. Van Oyen (*in press*) provides an archaeological comparison of the two approaches and clearly shows how both are designed to do very different things and should be treated as such by archaeologists. In summary, I argue that the metaphorical use of network concepts in archaeology is useful only when one refers to existing specifications of how these concepts are represented by data if one adopts existing concepts, or if one develops new network concepts then one should formulate such specifications. Without these specifications hypotheses cannot be falsified and archaeologists cannot discuss their hypotheses on equal terms. The implication of this statement for the archaeological use of network science is that confirmatory network science techniques can only be usefully applied when specifications are formulated of how the network concepts used should be represented as network data.

Although it is easy to claim that the rules underlying network science techniques are known, it is less straightforward to assume that the traditional education of archaeologists allows them to decipher these algorithms. Archaeologists are not always sufficiently equipped to critique the mathematical underpinnings of network science techniques, let alone to develop novel techniques tailor-made to address an archaeological question. For many archaeologists this means a real barrier or at least a very steep learning curve. Sadly, it also does not suffice to focus one's efforts on the most common techniques (as the literature review revealed was the case until recently in archaeological network science) or on learning graph theory. Like GIS, network science is not a single homogeneous method: it incorporates every formal technique that manages, visualises, or analyses the interactions between nodes (either hypothetical or observed), and it is only the particular nature of the network as a data type that holds these techniques together (Brandes et al. 2013). For this purpose it draws on graph theory, statistical and probability theory, algebraic models, but also agent-based modelling and GIS. The review of network science techniques developed in different disciplines performed for this PhD project revealed that archaeologists should not become specialists in all of these different approaches. Rather, archaeologists should be able to understand what kinds of questions different techniques are designed to answer and to evaluate whether it allows them to achieve their research aims. For this, hardly any technical skills are needed, merely a willingness to evaluate a scientific method. Archaeologists should be supported in this by 'guides to good practice' and critical archaeological case studies, which are appearing ever more frequently. The specifications formulated at the end of this chapter might be one of the steps to enhancing the ability of archaeologists to do this. However, the subsequent application of the selected technique does require a thorough knowledge of its functioning, and I have argued that multi-disciplinary engagement or collaboration significantly aids this process.

A thorough understanding of the technical underpinnings of particular applied network techniques is not an option but a prerequisite for a critical interpretation of the results. The case studies illustrated how even the most commonly used network science techniques, like degree or betweenness centrality, mean very different things depending on the research context in which they are applied, and that in some contexts theoretical assumptions (rather than methodological motivations) might prevent some techniques from being used. For example, because I assumed in the exploratory network analysis in case study 3 that the similarity of forms' distributions does not indicate or allow for any kind of flow between forms, I argued path-based measures (such as betweenness and closeness centrality) should not be used. Moreover, many exploratory network science techniques are sensitive to the sample size or boundary selection of the network used, as illustrated through the sensitivity analyses in case studies 2 and 3. It is not sufficient to obtain the results of the most commonly applied exploratory network measures to a network dataset: the meaning of nodes, links, measures, and the network dataset boundaries need to be discussed explicitly and will determine the interpretation of the results.

Another good example of this challenge is network visualisation. Many archaeologists consider the visualisation of networks as graphs a useful exploratory technique to understand the nature of their data, in particular when combined with geographical visualisations (e.g. Golitko et al. 2012; Mol 2014a). However, there are many different graph layout algorithms, and all of them are designed for a particular purpose: to communicate a certain structural feature most efficiently (Conway 2012; Freeman 2005). These days, user-friendly network analysis software is freely available and most of it includes a limited set of layouts, often not offering the option of modifying the impact of variables in the layout algorithms. Not understanding the underlying 'graph drawing aesthetics' or limiting one's exploration to a single layout will result in routinized interpretations focusing on a limited set of the network's structural features. The first case study offered some examples of this, where some network visualisations (the citation weights in particular) were considered highly effective at exploring and communicating a certain pattern, whilst others included too many nodes and arcs to allow for more than the most general of interpretative statements. Moreover, the use of network visualisations in the exploratory network analysis of case study 3 was considered to add very little that could not be gained through alternative approaches. The issues surrounding network visualisation and their implications for archaeological research were not addressed in more detail in this PhD project and should form the focus of future work.

However, the case studies made clear that network visualisations offer an interesting alternative approach for exploring a dataset that forces one to think through relationships and one's assumptions of what these mean. For communication purposes the case studies revealed that only in cases with a small number of nodes and where dependence assumptions gave rise to specific patterns (e.g. the citation weights visualisations in case study 1) were network visualisations preferable over other types of data representation (e.g. the geographical distribution maps[31] or the matrix representation of BR results in case study 3).

6.2.2. Data

It should be clear by now that network science is by no means a method devoid of any theoretical considerations. This PhD project has focused on how archaeological theoretical assumptions influence the network science research process and the selection of techniques. Most interestingly, however, in many other archaeological network studies theoretical critiques are often triggered by issues concerning the role of archaeological data. This is usually a result of the material nature of archaeological data serving as proxy evidence for past human behaviour, which poses a number of challenges.

Firstly, imposing categories and sometimes hierarchical relationships on data is a prerequisite for any network science. This results in the assumption that categories can actually be defined with any certainty (Butts 2009), and from the need to establish data categories ahead of the analysis, rather than letting them emerge from the analysis (Isaksen 2013). Indeed, the definition of nodes, ties and the network as a whole are a crucial part of the network science research process. However straightforward such definitions seem, doing so in a critical manner is not as easy as it sounds as the case studies have shown. For example, in the third case study I chose to follow a formal ceramic typology, where each node represents a distinct form of tableware. When doing so I had to acknowledge that such typologies are modern constructs and that alternative categorisations can easily be developed. This led me to argue for the importance of re-doing the exploratory network analysis using alternative functional categories. This in turn raises the issue that the network data we analyse is not necessarily identical to the past networks we are trying to understand. For example, although in some cases it can be proven that particular ceramic types were used for particular purposes and in certain contexts, their meaning can nevertheless change through time, requiring a

[31] For maps showing the spatial distribution of the dataset per period and per ware, see: http://icrates.arts.kuleuven.be/icrates/network-analysis/webpages/icrates_maps.html (accessed 01/06/2014).

modification of our categorisation (van Oyen *in press*). The case studies have shown that network science techniques require the establishment of a categorisation of data before analysis, but they have also shown that the categorisation is not imposed by the network techniques themselves but rather by the archaeologist's motivations to abstract the phenomenon of research interest into network concepts. Categorisation is common practice in archaeological research and is a requirement of many formal methods. The issues surrounding categorisation are therefore not specific to network science but rather stem from the scholar's theoretical perspective. If a method is needed where the boundaries of entities are ill-defined and fluid, and where one argues these can not under any circumstances be tied down for analytical purposes, then network science does not offer the solution.

Secondly, unlike network analysts in many other disciplines, archaeologists work with primary data sources of a material nature. Social network analysts often only consider inter-personal interactions, whilst archaeological network analysts are forced to consider *object-person* and *object-object* interactions. A range of interactionist theoretical perspectives exist to confront materiality, for example the ANT theoretical perspective mentioned above places a strong focus on the analysis of *object-person* interactions. Archaeological network scientists are faced with finding a workable framework that combines both network theories and methods. Indeed, this has been the focus of Carl Knappett's interest in networks. In *An Archaeology of Interaction* Knappett (2011) discusses a range of interactionist theories and how formal network techniques such as affiliation networks could be used to operationalise these. I find that the diversity of network methodological and theoretical network perspectives illustrated in Knappett's book and in the literature review in chapter 2 leads to the false impression that different network perspectives need to be able to be applied together within a single framework. I believe there is no need for a great unifying theory or method in archaeology, not even for one that just focuses on questions of connectivity. The multi-vocality achieved by contrasting results derived from different perspectives has already proven vastly more fruitful for the study of the past. Rather, I consider the different network perspectives as tools that function according to certain rules, and once these rules are known the tools have a potential to make small but crucial contributions to our knowledge of the past. I believe that if we are to ever achieve the full potential of these exciting new approaches for archaeology we will need to first critically explore them in isolation. This thesis has aimed to do this for formal network science methods only.

In summary, the decisions archaeological network scientists make when defining nodes and edges, when selecting or modifying analytical techniques and when interpreting the outcomes, are fundamentally influenced by their theoretical preconceptions, as determined in the network science research process. There is not a single right way to incorporate and interpret archaeological data in network approaches. Network science is not a method that solves archaeological challenges related to the use of fragmentary material data used as proxy evidence for past human behaviour. Therefore, network science (as is the case for any other formal method) in our discipline can never be separated from the archaeological theoretical motivations of how and why certain archaeological evidence allows one to better understand a past phenomenon.

6.2.3. Space

The representation of network data is not only dependent on data type categorisation but also necessarily reflects the research questions being asked, revealing an issue of spatial scales. Do the past phenomena we are interested in concern interactions between regions, sites or individuals? How can this be represented in node, tie and network data definitions? The ability of network approaches to work on multiple scales is often mentioned as one of the advantages of using formal network methods (Knappett 2011). In practice, however, archaeological network analysts have traditionally focused on inter-regional or macro-scales of analysis. Knappett (2011) argues that it is on the macro-scale that network analysis comes into its own and a recently published edited volume reveals this regional emphasis (Knappett 2013b). The second and third case studies of this PhD provide yet another example of regional network approaches. This insistence to work on large scales becomes quite unique in light of social network analysts' traditional focus on individual social entities in interaction. SNA provides a multitude of good examples of how network methods could be usefully applied on a micro- or local scale of analysis (such as triad census, see Blake 2014, or ego-networks, see Mol et al. 2014b). However, the nature of archaeological data, which rarely allows for individuals and their interactions to be identified with any certainty, should not be considered the only reason for this focus on the macro-scale. Arguably, networks lend themselves very well to exploring inter-regional interaction, and archaeologists have always had a particular interest in the movements and flows of people, resources and information across large areas. Moreover, many of the early applications of network methods in archaeology, which in some cases might have served as an example to more recent applications, concerned inter-regional interaction (e.g. Terrell 1976).

However, one should acknowledge the importance of exploring how local actions give rise to larger-scale patterns if we are to benefit from the multi-scalar advantage of formal network methods. The third case study offered an example of how this can be done, through simulating hypotheses of how commercial agents interacting on and between markets are the driving force behind the Roman trade systems. The ABM approach focused on falsifying these hypotheses, and was therefore not reliant on detailed archaeological evidence of individuals in interaction. Arguably, such data requirements will remain a strong limitation to the kinds of questions archaeologists can answer. But it should be clear that confirmatory network methods (both statistical and ABM) offer archaeologists an additional approach to understanding how large-scale patterns emerge through the particular interactions of individual agents.

It is not surprising that many archaeological network analysts are interested in exploring the dynamics between relational and geographical space, given the importance of spatial factors in understanding archaeological data and archaeologists' traditional interest in geographical methods (e.g. Hodder and Orton 1976). Despite early work by archaeologists on geographical networks (for an overview see chapter 2 in Knappett 2011), geographical space has been almost completely ignored by sociologists and physicists, resulting in a very limited geographical network analysis toolset for archaeologists to draw on (although see a recent special issue of the journal Social Networks (issue 34(1), 2012) and the review work by Barthelemy (2011), as well as techniques used in Space Syntax (Hillier and Hanson, 1984)). I believe this is one of the areas in which archaeologists can make valuable contributions to network science, and a number of archaeologists have already done so by suggesting or applying spatial network methods (e.g. Bevan and Wilson 2013; Bikoulis 2012; Knappett et al. 2008; Menze and Ur 2012; Wernke 2012). Indeed, a review paper of spatial and social network techniques applied to historical and archaeological data was recently published by Marc Barthelemy (2014) in *Nouvelles de l'Archéologie*. The second case study of this PhD presents a method that draws on techniques developed in SNA and GIS to make such a methodological and archaeological contribution to network science.

6.2.4. Process

Many archaeological network studies treat networks as static snapshots. This is at least in part a result of the nature of archaeological data and our inability to observe past processes directly. Graph visualisations and many network science techniques further enforce this idea of a static network by exploring structural features of particular networks in isolation.

However, the past systems we study were governed by dynamic phenomena and the network approach used to understand these phenomena should reflect their changeable nature. In fact, one could argue that no network is truly static since our assumptions underlying network data imply flows of resources and network change as a result of this. Sander van der Leeuw makes a similar point in his discussion of the collection of archaeological network science articles edited by Knappett (2013a): in most of the archaeological literature on networks "… the network itself is viewed independent of the temporal dimensions that affect different interactions differently" (van der Leeuw 2013, 341). He suggests that primacy, the advantages of nodes that are established in the network before others, should be given more explicit attention, and that "… we should not so much look for the origins of phenomena that we observe, but instead *look for their emergence*" (italics by van der Leeuw 2013, 337).

Archaeological data often does not have the chronological accuracy to reconstruct an exact sequence of events which is required to empirically identify primacy: which ties and nodes appeared and disappeared in what order? The second case study presented an example of this problem, where very coarse chronological time-slices were used to observe changes in a settlement pattern. However, a number of network modelling approaches exist that can help one deal with this issue, including agent based modelling (e.g. Graham 2006a; case study 3), algebraic modelling (e.g. Menze and Ur 2012; not applied here), and statistical modelling (e.g. Lusher et al. 2013; case study 2). Underlying all of these modelling approaches are clearly formulated assumptions of what a relationship means and what types of flows it allows for. They therefore require one to explicitly acknowledge the dynamic nature of past processes and the dynamic assumptions underlying the definition of ties. Because of this, I believe that these confirmatory network science techniques reveal the potential contribution of network science for archaeology far more than the results of the exploratory network techniques used in this PhD. For example, although the dependence assumptions underlying the citation network data in case study 1 implied dynamic processes such as the flow of 'influence', the exploratory network analysis applied there did not allow me to identify these processes. Rather, they allowed me to study the structure of the outcome of such processes. I nevertheless believe that exploratory network methods that represent these dependence assumptions most explicitly (e.g. the citation weights method in case study 1) offer an interesting way of considering dynamic processes in exploratory data analyses. In the second case study, a technique was presented that relied strongly on the available data in order to make statistical inferences of which kinds of dynamic processes could give rise to this data.

The evaluation of this technique revealed that the data requirements are high and that it can only be applied in research contexts where such data are available. It also showed how strongly theoretical assumptions weigh on the kinds of results one can expect in network science. For example, the ERGMs of networks with lines of sight up to 50km, which make little theoretical sense since few features are inter-visible over such distances, did not reveal any processes based on our archaeological assumptions that could give rise to such networks.

But which model is best? Many models, representing different hypothetical processes, can be created that could all give rise to the same observed network. Since archaeologists cannot directly observe past processes, and given that our data are incomplete and are merely indirect proxies, how then can we ever claim that one model is more probable than any other? The problem of equifinality (the idea that multiple processes can have the same end result) is equally critical for network analysis as for any other technique in the archaeologist's toolbox. There are a few ways in which formal network methods can help us address this issue. Firstly, archaeological data (however flawed) used in statistical models can help us to identify very general processes that are more probable than others, as I have illustrated in the second case study. This method allows one to narrow down the range of hypothetical processes that could give rise to an observed network. However, the suggested processes are very abstract and proved hard to interpret. Moreover, as stressed above, this ERGM method can only be used in cases where good quality datasets and clear dependence assumptions make it a suitable approach. Secondly, these models can help us to formally express otherwise ill-defined hypotheses and evaluate their likeliness given certain archaeological assumptions. Thirdly, they might not be able to prove certain processes, but models can definitely be used to negatively test or falsify hypotheses, or at least identify which processes are less likely than others (given our current knowledge). In this way, such models serve as experimental laboratories (Premo 2006). This was the approach used in the third case study, where an ABM was developed to falsify existing archaeological/historical hypotheses. In such approaches a focus on falsification is preferable over one that aims to prove hypotheses, since the issue of equifinality is most critical here: alternative models of the Roman trade systems could easily be formulated and could potentially provide similar results. In such cases it becomes difficult to determine which model is more likely in the absence of good data, but it is nevertheless of interest that two alternative conceptualisations of the same hypothesis succeed in falsifying it. One has to acknowledge, however, that some past processes are unknowable given our current techniques and datasets. All archaeological approaches suffer from this disadvantage and network science is no exception.

6.2.5. Past social networks

Finally, I would like to briefly discuss a persistent issue surrounding the use of network science in archaeology for studying past social networks: that combining more network data and network science techniques allows one to better understand past social phenomena. Although I have argued the use of network science in archaeology should not be restricted to the study of past social networks (as a conceptualisation of structure of different ways in which past individuals were related), it is still the research focus of many archaeological applications. Yet the archaeological literature reveals a false expectation that all archaeological network science should concern the study of past social networks, regardless of our ability to do so. I would like to illustrate this statement by van der Leeuw’s (2013) critical evaluation of archaeological network science. He takes a broader social science perspective in which he distinguishes between three fundamental commodities that constitue the metabolism of human societies (matter, energy, and information), to formulate his critique that:

> “… in most of the archaeological literature on networks I am aware of, we seem to be dealing with ‘flat’ networks, or at least flat projections of networks, even though in human societies hierarchy is a very important structuring characteristic. In such networks, only the position of any particular node in the network is taken into account, and the network itself is viewed independent of the temporal dimension that affect different interactions differently” (van der Leeuw 2013, 341).

> “there are two kinds of networks involved in any case study – one for organization (information processing), and one for resources (matter and energy) – and they should not be conflated or confused. In real life, these two kinds of flows have a very complicated relationship, which we may not always be able to disentangle” (van der Leeuw 2013, 339).

He emphasises that past social phenomena are complex and that archaeological network scientists should confront this complexity if they are to come to a better understanding of these phenomena. Van der Leeuw suggests two ways in which this can be done, drawing on the discussion of dynamic networks by White (2009). Firstly, he argues that hierarchical network sctructures represented by acyclic networks[32], introduced in case study 1, offer a

[32] In his discussion Van der Leeuw creates the wrong impression that all directed asymmetric graphs have no directed cycles and are therefore acyclic. Given his focus on hierarchic networks I assume van der Leeuw meant to refer to acyclic networks only, and this is how I will interpret his suggestions.

useful conceptualisation. These have the advantage of not being 'flat' but having a pyramidal structure with few nodes 'on top', from which information, materials, or energy flows and is channelled to peripheral nodes. Secondly, he suggests that one should study multiple networks representing different aspects of complex past social phenomena (e.g. spatial, business, kinship relationships), the way in which these multiple networks interconnect, and how they emerged.

I believe van der Leeuw's perspective is important and that his methodological suggestions are promising. Moreover, this thesis shares van der Leeuws argument that a shift from static to dynamic network approaches is needed, and has illustrated ways of putting this into practice. However, in light of this PhD project's aims I would like to emphasise that the promise in the approach he suggests does not derive from an ability inherent in network science to offer a solution for the study of complex social phenomena. A research perspective that considers the interlock of multiple factors of complex phenomena conceptualised as networks is interesting, but its ability to produce non-trivial analytical results will depend more on whether the network conceptualisation of each factor offers advantages over other conceptualisations, and whether comparable datasets are available. Moreover, the insistence on acyclic networks seems to disregard the assumption that all relationships between social entities allow for flows, and the need for network conceptualisations to be phenomenon- and research-context specific. In short, one cannot assume that all archaeological network science should strive to tackle as much as possible of the complexity of the studied past social phenomena in one network approach. Rather, network science offers techniques which allow for different aspects of past social phenomena to be studied if a network conceptualisation is considered best and if the data requirements are met. Aspects of social phenomena which are best conceptualised differently and studied using alternative tools should not be forced into a network science method with the assumption that the results will be comparable and non-trivial.

6.3. A 'guide to good practice': specifications for future archaeological network science

This chapter introduced a large number of advantages and limitations of the archaeological use of network science, and discussed how I believe network science best fits in the archaeological research process. These lessons hold the potential to guide future archaeological network science in a very specific practical way: as a set of suggested specifications. I agree with Isaksen (2013) that archaeologists should be provided with the means to evaluate how and whether network science is a tool that can lead to innovative insights in their own research contexts. This is a task for the more mathematically- and computer-literate archaeologists, who should provide critical applied case studies as examples and work towards a 'guide to good practice'[33] accessible to all archaeologists. This PhD offered three such case studies, and to conclude I would like to summarise the lessons I learned as a set of suggestions which could become part of a more elaborate 'guide to good practice' in the future. I purposefully write these suggestions in a very generic manner rather than provide particular examples, in line with the aim of this PhD to evaluate network science for the general use in archaeology:

- Network science techniques are methodological tools with clear rules and limitations.
- Archaeologists could be provided with guides to good practice and archaeological examples, making them able to understand what kinds of questions different network science techniques are designed to answer and to evaluate whether it allows them to achieve their research aims. To do this hardly any familiarity with mathematical and computational techniques is required, only a willingness to explore the potential of a scientific method. This thesis along with other recent publications (e.g. contributions in Knappett 2013a) can help archaeologists by providing specifications on whether network science techniques are desirable for addressing their research questions, and by offering examples of archaeological research questions that benefit from a network science treatment.
- An evaluation of the potential contribution of network methods to addressing a particular research problem might be enhanced by working explicitly through the network science research process, which again does not require much technical skills.
- However, once archaeologists have decided to apply a specific network science technique, then a thorough understanding of the technical underpinnings of this technique is not an option but a prerequisite for a critical interpretation of its results. Archaeologists could be aided in this process by multi-disciplinary engagement and collaboration where possible.

[33] A 'guide to good practice for archaeological network science' could be inspired by the guides to good practice provided by the Archaeology Data Service. Indeed, it would fit perfectly within this series among other data analysis and visualisation techniques like GIS, CAD, and VR. http://guides.archaeologydataservice.ac.uk/ (accessed 06-09-2014).

- Network concepts developed in network science are associated with specified data requirements, which should be acknowledged by the archaeologists adopting them. If the data requirements cannot be identified in empirical or simulated data then the network concepts loose all explanatory value.
- When developing new network concepts, one should formulate network data specifications such that it becomes clear how the concept differs from exisiting concepts.
- Formulating specifications of how network concepts are represented in network data allows for different conceptualisations of the same past phenomenon to be compared and possibly falsified.
- A shift in perspective from the study of static structures to the emergence of empirical observations and past phenomena might be needed (van der Leeuw 2013).
- Confirmatory network science techniques offer archaeologists an approach to understanding how large-scale patterns emerge through the particular interactions of individual agents or relationships.
- Confirmatory network science techniques can only be usefully applied when specifications are formulated of how the network concepts used should be represented as network data.
- Confirmatory network science techniques require one to explicitly acknowledge the dynamic nature of past processes and the dynamic assumptions underlying the definition of ties. Because of this, I believe these techniques reveal the potential contribution of network science for archaeology far more than the exploratory network techniques used in this thesis.
- The past systems we study were governed by dynamic phenomena and the network approach used to understand these phenomena should reflect their changeable nature.
- The case studies revealed that only in cases with a small number of nodes and where dependence assumptions gave rise to specific patterns, were network visualisations preferable over other types of data representation for communication purposes.
- Even in cases where network science techniques do not offer additional functionality compared to other more common archaeological techniques, it could still lead to interesting insights by forcing one to explore a dataset or hypothesis through the lens of one’s assumptions about why and which relationships matter.

- If a method is needed where the boundaries of entities are ill-defined and fluid, and where one argues these can not under any circumstances be tied down for analytical purposes, then network science does not offer the solution.
- Network science can never be separated from the archaeological theoretical motivations of how and why certain archaeological evidence allows one to better understand a past phenomenon.
- Some past processes are unknowable, due to our current techniques and datasets. All archaeological approaches suffer from this disadvantage and network science is no exception.

CHAPTER 7

CONCLUSION:

THE POTENTIAL OF NETWORK SCIENCE IN ARCHAEOLOGY

7. Conclusion: the potential of network science in archaeology

7.1. The contribution of this PhD

This PhD project aimed to evaluate the potential of network science as a method in archaeology. My approach to testing this potential has been to argue that network science offers methodological advantages that no other method in the archaeologist's toolbox shares and that enhances archaeological research. This was done by adopting the definition of network science by Brandes and colleagues (2013), which claims that network science is the study of the management, representation, and analysis of network data, and network data represents our theoretical statements about why relationships matter; which was reformulated as a central claim in this PhD:

> Archaeologists are confronted with network data; network science offers the suite of techniques necessary to deal with network data.

In order for this claim to be confirmed and not to be trivial I needed to address a number of research questions:

1. Do archaeologists commonly ask research questions that are best addressed through the analysis of network data? What kinds of archaeological research questions are well suited to explore from a networks perspective? Does network science allow archaeologists to ask new questions?
2. Are archaeologists commonly confronted with network data? What data are typically well suited to submit to a network science approach?
3. Are network techniques and visualisations useful exploration and communication tools in archaeology?
4. What are the limitations of formal network methods in archaeology?
5. Do the typical research questions and datasets in Roman archaeology pose particular advantages or disadvantages?

These questions were first addressed through a review of previous archaeological applications of network science (chapter 2), which led me to two key conclusions

1. A general unawareness of the historicity and diversity of formal network methods both within and outside the archaeological discipline or of suitable archaeological

applications of known models and techniques has resulted in a very limited methodological scope.

2. The adoption or development of network methods has very rarely been driven by specific archaeological research questions and is dominated by a few popular models and techniques, which has in some cases resulted in a routinized explanatory process.

These conclusions suggested that in order to evaluate the potentially innovative contribution of network science to archaeology, future studies should take a broad multi-disciplinary scope, let the specific archaeological research context dominate the application, and explicitly work through each step in the network science research process. I have argued that doing this would force one to think through the definition of network science adopted in this PhD in order to evaluate whether network science would allow for a useful approach to a specific archaeological research question, and which network technique would be best suited for this purpose. In order to test these suggestions they were put into practice through three case studies, to illustrate how this would indeed allow me to answer this project's research questions (chapters 3-5).

The three case studies offered examples of archaeological research questions that imply dependence assumptions about how relationships between entities matter, as well as of archaeological datasets, which are commonly treated as network data by archaeologists. However, I have argued that the ways in which archaeologists have dealt with these dependence assumptions did not allow them to test their hypotheses. A range of formal network and non-network techniques were used to address this issue in each case study. This led to the adoption of innovative methods (in particular the use of citation network techniques, the application of ERGM to visibility networks, and the use of ABM for falsifying hypotheses surrounding Roman tableware distributions), which allowed me to test past processes which were so far only hinted at, but which also allowed me to re-focus existing research questions or address completely new research questions. In each case study I have provided examples of results that could not have been obtained through non-network science methods. However, it also provided examples of cases where network science techniques did not lead to new insights. Most interesting were the disappointing results of some of the exploratory network measures used in case study 1 and of most of the exploratory network analysis of case study 3. Although the latter offered new insights and more detailed results in a few cases, it would be hard to argue why in this case study an exploratory network analysis would be preferable over a thorough exploratory data analysis or a statistical treatment of the dataset. However, it is worth noting that in case study 1 the exploratory network techniques which offered the most interesting results were those for which the dependence assumptions were most clearly formulated and relevant to answering the case study's research questions.

I therefore conclude that the literature review and case studies presented here suggest that archaeologists commonly formulate (implicitly or explicitly) dependence assumptions, that they are confronted with network data, and that in such cases network science techniques can offer methodological advantages over other methodological tools used by archaeologists. This conclusion firmly establishes network science in the archaeologist's methodological toolbox. The main contributions of this PhD project are, therefore, i) to argue why network science offers archaeologists unique and necessary methodological advantages, ii) to provide three critical practical examples of how it can be applied, iii) to offer a number of suggestions which could guide the future archaeological use of network science (section 6.3), iv) to introduce innovative methods of citation analysis and statistical network modelling to our discipline, and v) to emphasise the importance and possibility of falsifying and abstracting Roman archaeology hypotheses as transparent and comparable conceptualisations.

However, these statements should be qualified. Although this thesis has a very clear aim and comes to the above strongly formulated conclusions, one must recognize it is the result of a four-year process of trial and error. It should be clear from the case studies and the discussion in chapter 6 that I was confronted with a lot of interesting challenges, many of which are still unadressed, and surfacing these issues is as much a result of this PhD project as my argument that network science offers methodological advantages for some archaeological research. In order to overcome all these challenges and uncover new and useful archaeological uses of network science it is important that archaeological network science should be considered more than just a hype: archaeological network scientists should be cautious to make grand claims about the novelty of their methods and instead enable a larger body of archaeologists to critique their approaches.

7.2. Where to go from here? Archaeological network science beyond the "hype"

A number of suggestions can be formulated to guide future archaeological research that aims to overcome the remaining challenges. Firstly, although I have formulated my methodological conclusions by addressing the archaeological discipline at large, it must be stated that this is an extrapolation of results obtained through two case studies in Roman archaeology as well as an archaeological case study and literature review. Other archaeological subdisciplines where less complete datasets are available, such as palaeolithic archaeology, or less intensively investigated geographical areas might confront one with particular methodological challenges. These should be evaluated on a case-by-case basis. Secondly, the strong focus on network science methodology, which was necessary for this project, meant their relation to more theoretical network perspectives was given little attention. Relational theoretical perspectives are becoming increasingly popular in archaeology (Hodder 2012; Knappett 2011; Malafouris 2013) and some have argued they might be compatible with network science methods (e.g. Knappett 2011). I would like to reiteratre from the introduction to this PhD that there can be no network “doing” without network “thinking” and that the possible combination of these methodological and theoretical perspectives should be critically scrutinized in future work. Thirdly, archaeologists might make valuable methodological contributions to network science thanks to their focus on material and spatial data. Many of the theoretical relational perspectives concern *object-person* interactions, which have received very little attention in the SNA and physics communities that have driven the development of network science techniques. Archaeologists should argue and illustrate why such interactions require the development of different methodological tools. Archaeologists also have a strong tradition in working with spatial data whilst network science has only recently started to develop spatial network techniques. The combination of spatial analysis techniques that have proven valuable in archaeology (such as visibility or least-cost path analyses) with network science techniques might offer a fruitful area of collaboration between archaeologists and network scientists. One particular example of this is the need to incorporate spatial constraints in ERGMs as argued in case study 2. Finally, archaeological network scientists should enable a wider archaeological audience to critique their methods. More published case studies with clear and transparent explanations of the network methods used aimed at an archaeological audience are needed. Moreover, archaeological network scientists should collaborate to create an accessible ‘guide to good practice’ in a similar format as those published by the Archaeology Data Service.

I am optimistic that these remaining challenges will be addressed in the future thanks to the current speed of developments in archaeological network science. When I started my PhD research in 2010, the use of network science in archaeology was very different to what it was when finishing this project in 2014. This change is best illustrated by the trend observed when counting the number of archaeological publications which apply network science techniques over time (Fig. 5): the published use of network techniques became slightly more common between 2000 and 2010, but this increase is dwarfed by the number of publications between 2012 and 2014.[34] In addition to these publications, the last few years has known a strong increase in the number of conferences, sessions and workshops on archaeological network science, and most importantly an increase in the number of people attending these events[35]. One could argue these are indications that both the application and visibility of network science in our discipline has recently increased significantly.

A useful tool to interpret this is Gartner's "hype cycle" (Fenn and Raskino 2008), which offers a model for the life-cycle of emerging technologies (Fig. 60). It is a representation of how a technological innovation is surrounded by inflated speculations of its prospects and overenthusiasm soon after its emergence, followed by a period of disillusionment where it does not seem to live up to the expectations, until finally its place or role in a domain is understood allowing it to be used to its full potential. We have seen that network methods have been used by archaeologists at least since the 1960's, but only in the last decade or so have they been more widely applied. Does this mean we are heading towards a peak of inflated expectations? Or have we already reached that peak and are we racing down the slope towards a trough of disillusionment?

[34] This very recent surge can be attributed to the publication of a number of key monographs and collected volumes in archaeological network science: Knappett's (2011) '*An archaeology of interaction*'; a volume of case studies edited by Knappett (2013a) published as '*Network analysis in archaeology*'; a special issue of the *Archaeological Review from Cambridge* edited by Evans and Felder (2014) titled '*Social network perspectives in archaeology'*; a special issue of *Nouvelles de l'archéologie* edited by Knappett (2014) published as '*Analyser les réseaux du passé en archéologie et en histoire*'; and a special issue of the *Journal of Archaeological Method and Theory* edited by Brughmans, Collar, Coward and Mills (2014) titled '*The connected past: critical and innovative approaches to networks in archaeology'*.

[35] See for example the series of events organised by the '*Arts, Humanities, and Complex Networks*'(), '*The Connected Past*' (http://connectedpast.soton.ac.uk/ accessed 2-6-2014), '*Réseaux et Histoire*' (http://reshist.hypotheses.org/ accessed 2-6-2014), and '*Historical Network Research*' communities (http://historicalnetworkresearch.org/ accessed 2-6-2014), and the start of yearly archaeological and historical network science sessions at the '*Sunbelt SNA*' (http://www.insna.org/ accessed 2-6-2014), '*Computer Applications and Quantitative Techniques in Archaeology*' (http://caa-international.org/ accessed 2-6-2014), and '*Digital Humanities*' conferences (http://adho.org/conference accessed 2-6-2014).

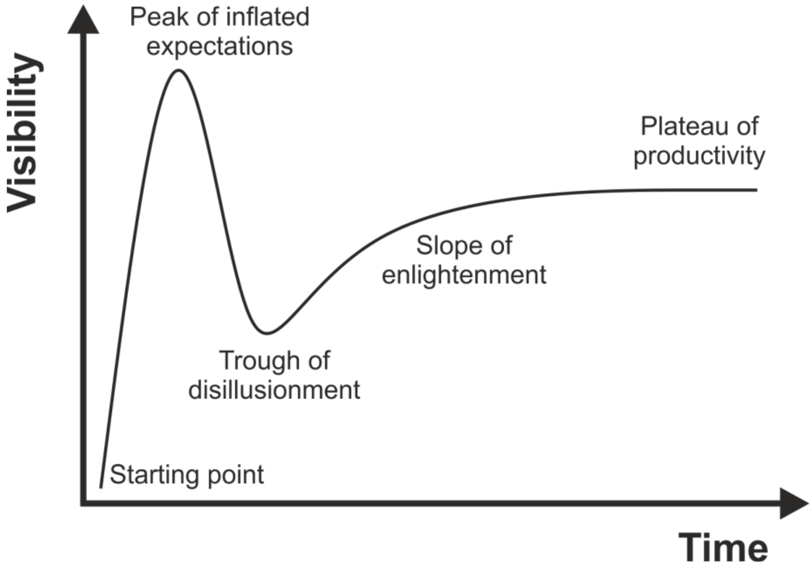


Fig. 60. Gartner's hype cycle for emerging technologies: the moment of technological innovation is the starting point; expectations rise rapidly leading to a peak in visibility; the technology does not live up to overexpectations creating a negative hype; inevitably the technology will get through these ups and downs and mature, its potential is better understood. Note that this curve is a model, and that it does not represent the full life-cycle of technologies, which could still fail or increase after the initial patterns.

An increasing number of scholars argue for the innovative contributions network science can make to archaeology, whilst emphasising the many challenges archaeologists are faced with when using formal network methods (Brughmans et al. *in press*; Fulminante 2014; Isaksen 2013; Knappett 2011; *in press*). But can we really claim to understand the role network science can play in the archaeological research process when its use is surrounded by inflated expectations and challenges? How far along are we in understanding the potentially productive contribution of network science to archaeology? As a passionate advocate of archaeological network science I would like to think we have already reached this consolidation phase, but this PhD shows that most likely we are still in a phase of optimism. Regardless of whether we are now in a positive or negative hype, I believe that the Gartner hype cycle teaches us an important lesson: innovations will inevitably overcome the disillusionment and mature. Going through these ups and downs is a necessary process, and the judgement of adopters should therefore not be clouded by positive or negative jumps and drops. Instead, adopters should focus on whether and what the technological innovation really contributes to a domain.

I believe the literature review and case studies presented in this PhD show that now it is possible for archaeological network scientists to work towards the "plateau of productivity". This PhD project has illustrated that network science enriches the methodological toolkit of our domain. By recognising the particular nature of network data, by maintaining a multi-disciplinary perspective, and by confronting the many methodological challenges ahead in practice, archaeologists are well on their way to harnessing the full potential network science has to offer to our discipline. Moreover, archaeologists' focus on the relationships between material culture and people, their ability to explore cultural change over the long-term, and their challenge of dealing with a diversity of fragmentary textual and material data types that serve as proxy evidence for past human behaviour might hold the promise of valuable and unique contributions from archaeology to network science in general. This might allow us to further discard some of the remaining expectations and challenges, in order to better position network science in archaeology, with the "plateau of productivity" as our goal.

8. Bibliography

Abadie-Reynal, C. (1989). Céramique et commerce dans le bassin Egéen du IVe au VIIe siecle. In V. Kravari, J. Lefort, & C. Morrisson (Eds.), *Hommes et richesses dans l'empire byzantin. Tome I, IVe-VIIe siècle* (pp. 143–159). Paris.

Adamic, A. L., & Huberman, B. A. (2000a). Power-Law Distribution of the World Wide Web. *Science*, *287*(5461), 2115. doi:10.1126/science.287.5461.2115a

Adamic, A. L., & Huberman, B. A. (2000b). The Nature of Markets on the World Wide Web. *Quarterly Journal of Electronic Commerce*, *1*, 5–12.

Adams, J., Faust, K., & Lovasi, G. S. (2012). Capturing context: Integrating spatial and social network analyses. *Social Networks*, *34*(1), 1–5. doi:10.1016/j.socnet.2011.10.007

Albert, R., & Barabási, A.-L. (2002). Statistical Mechanics of Complex Networks. *Reviews of Modern Physics*, *74*, 47–97.

Albert, R., Jeong, H., & Barabási, A.-L. (1999). Internet: Diameter of the World-Wide Web. *Nature*, *401*, 130.

Albert, R., Jeong, H., & Barabási, A. (2000). Error and attack tolerance of complex networks. *Nature*, *406*(6794), 378–82. doi:10.1038/35019019

Alen Sula, C., & Miller, M. (2014). Citations, contexts, and humanistic discourse: Toward automatic extraction and classification. *Literary and Linguistic Computing*. doi:10.1093/llc/fqu019

Alexander, M. C., & Danowski, J. A. (1990). Analysis of an ancient network: personal communication and the study of social structure in a past society. *Social networks*, *12*, 313–335.

Allen, K. M. S. (1990). Modelling early historic trade in the eastern Great Lakes using geographic information systems. In K. M. S. Allen, S. W. Green, & E. B. W. Zubrow (Eds.), *Interpreting space: GIS and archaeology* (pp. 319–329). London - New York - Philadelphia: Taylor & Francis.

Anderson, C. J., Wasserman, S., & Crouch, B. (1999). A p* primer: logit models for social networks. *Social Networks*, *21*(1), 37–66. doi:10.1016/S0378-8733(98)00012-4

Arnaud, P. (2005). *Les routes de la navigation antique. Itinéraires en Méditerranée*. Paris: Editions errance.

Bak, P., Tang, C., & Weisenfeld, K. (1987). Self-organized-criticality: an explanation of 1/F noise. *Physical Review Letters*, *59*, 381–384.

Bang, P. F. (2008). *The Roman bazaar, a comparative study of trade and markets in a tributary empire*. Cambridge: Cambridge university press.

Barabási, A.-L. (2002). *Linked: The New Science of Networks*. Cambridge, Massachusetts: Perseus.

Barabási, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, *286*(5439), 509–512. doi:10.1126/science.286.5439.509

Barabási, A., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, *311*(3-4), 590–614. doi:10.1016/S0378-4371(02)00736-7

Barnes, J. (1972). Social networks. *Addison-Wesley module in anthropology*, *26*, 1–29.

Barnes, J., & Harary, F. (1983). Graph theory in network analysis. *Social Networks*, *5*(2), 235–244. doi:10.1016/0378-8733(83)90026-6

Barthelemy, M. (2011). Spatial networks. *Physics Reports*, *499*(1-3), 1–101. doi:10.1016/j.physrep.2010.11.002

Barthelemy, M., 2014. Discussion: social and spatial networks. Nouv. l'archéologie 135, 51–61.

Batagelj, V. (2003). Efficient Algorithms for Citation Network Analysis. *Arxiv preprint cs/0309023*, 1–27.

Batagelj, V., Ferligoj, A., & Doreian, P. (1992). Direct and Indirect Methods for Structural Equivalence. *Social networks*, *14*, 63–90.

Beaudry, M., & White, J. (1994). Cowgirls with the Blues? A Study of Women's Publication and the Citation of Women's Work in Historical Archaeology. In C. Claassen (Ed.), *Women in Archaeology* (pp. 138–158). Philadelphia: University of Pennsylvania Press.

Bentley, R. A. (2003a). An introduction to complex systems. In R. A. Bentley & H. D. G. Maschner (Eds.), *Complex systems and archaeology* (pp. 9–23). Salt Lake City: University of Utah Press.

Bentley, R. A. (2003b). Scale-free network growth and social inequality. In R. A. Bentley & H. D. G. Maschner (Eds.), *Complex systems and archaeology* (pp. 27–46). Salt Lake City: University of Utah Press.

Bentley, R., Lake, M., & Shennan, S. (2005). Specialisation and wealth inequality in a model of a clustered economic network. *Journal of Archaeological Science*, *32*(9), 1346–1356. doi:10.1016/j.jas.2005.03.008

Bentley, R. A., & Maschner, H. D. G. (2003a). *Complex systems and archaeology*. Salt Lake City: University of Utah Press.

Bentley, R. A., & Maschner, H. D. G. (2003b). Preface: considering complexity theory in archaeology. In R. A. Bentley & H. D. G. Maschner (Eds.), *Complex systems and archaeology* (pp. 1–8). Salt Lake City: University of Utah Press.

Bentley, R. A., & Maschner, H. D. G. (2007). Complexity Theory. In R. A. Bentley, H. D. G. Maschner, & C. Chippendale (Eds.), *Handbook of archaeological theories* (Vol. 48, pp. 245–270). AltaMira Press. doi:10.1016/j.ijnurstu.2010.09.012

Bentley, R. A., & Shennan, S. J. (2003). Cultural Transmission and Stochastic Network Growth. *American Antiquity*, *68*(3), 459–485.

Bentley, R. A., & Shennan, S. J. (2005). Random Copying and Cultural Evolution. *Science*, *309*, 877–878. doi:10.1126/science.309.5736.874b

Bernardini, W. (2007). Jeddito yellow ware and Hopi social networks. *Kiva*, *72*(3), 295–328.

Bes, P. (2007). *A geographical and chronological study of the distribution and consumption of tablewares in the Roman East. Unpublished doctoral thesis*. Leuven: K.U.Leuven.

Bevan, A., & Wilson, A. (2013). Models of settlement hierarchy based on partial evidence. *Journal of Archaeological Science*, *40*(5), 2415–2427.

Bikoulis, P. (2012). Revisiting prehistoric sites in the Göksu valley: a GIS and social network approach. *Anatolian Studies*, 62: 35–59. DOI: 10.1017/S0066154612000026

Bintliff, J. (2004). Time, structure, and agency: the Annales, emergent complexity, and archaeology. In J. Bintliff (Ed.), *A companion to archaeology* (pp. 174–194). Oxford: Blackwell.

Blake, E., 2014. Dyads and triads in community detection: a view from the Italian Bronze Age. Nouv. l'archéologie 135, 28–31.

Boast, R., & Biehl, P. (2011). Archaeological knowledge production and dissemination in the digital age. In E. C. Kansa, S. W. Kansa, & E. Watrall (Eds.), *Archaeology 2.0. New approaches to communication and collaboration* (pp. 119–156). Los Angeles: UCLA Press.

Boissevain, J. (1973). An exploration of two first-order zones. In J. Boissevain & J. C. Mitchell (Eds.), *Network Analysis: Studies in Human Interaction*. The Hague: Mouton.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of mathematical sociology*, *2*, 113–120.

Börner, K., Maru, J. T., & Goldstone, R. L. (2004). The simultaneous evolution of author and paper networks. *PNAS*, *101*, 5266–527.

Boruvka, O. (1926). O jistem problemu minimalnim. *Acta Societatis Scientarium Naturalium Moravicae*, *3*, 37–58.

Borgatti, S. P. (2005). Centrality and network flows. *Social Networks*, *27*(1), 55–71.

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *Ucinet 6 for Windows: Software for Social Network Analysis, Analytic Technologies*. Harvard University.

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, *323*(April), 892–896.

Bott, E. (1957). *Family and social network*. London: Tavistock.

Bouchaud, J., & Mézard, M. (2000). Wealth condensation in a simple model of economy. *Physica A: Statistical Mechanics and its Applications*, *282*(3-4), 536–545. doi:10.1016/S0378-4371(00)00205-3

Box, G. E. P., & Draper, N. R. (1987). *Empirical Model-Building and Response Surfaces*. Oxford: Wiley.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.

Brainerd, G. W. (1951). The Place of Chronological Ordering in Archaeological Analysis, *16*(4), 301–313.

Brandes, U., Robins, G., McCranie, A., & Wasserman, S. (2013). What is network science? *Network Science*, *1*(01), 1–15. doi:10.1017/nws.2013.2

Branting, S. (2007). Using an Urban Street Network and a PGIS-T Approach to Analyze Ancient Movement. In E. M. Clark, J. T. Hagenmeister (Ed.), *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 87–96). Budapest.

Breiger, R. L. (1981). The social class structure of occupational mobility. *American Journal of Sociology*, *87*, 578–611.

Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., … Wiener, J. (2000). Graph structure in the Web. *Computer Networks*, *33*(1-6), 309–320. doi:10.1016/S1389-1286(00)00083-9

Broekaert, W. (2013a). Financial experts in a spider web. A social network analysis of the archives of Caecilius Iucundus and the Sulpicii. *Klio*, *95*(2), 471–510.

Broekaert, W. (2013b). Navicularii et negotiantes: a prosopographical study of Roman merchants and shippers. *Pharos*, *1*.

Broodbank, C. (2000). *An island archaeology of the early Cyclades*. Cambridge: Cambridge University Press.

Brown, C. T., Witschey, W. R. T., & Liebovitch, L. S. (2005). The Broken Past: Fractals in Archaeology. *Journal of Archaeological Method and Theory*, *12*(1), 37–78. doi:10.1007/s10816-005-2396-6

Brughmans, T. (2010). Connecting the dots: towards archaeological network analysis. *Oxford Journal of Archaeology*, *29*(3), 277–303.

Brughmans, T. (2012). Facebooking the Past : a critical social network analysis approach for archaeology. In A. Chrysanthi, M. P. Flores, & C. Papadopoulos (Eds.), *Thinking beyond the Tool: Archaeological Computing and the Interpretative Process. British Archaeological Reports International Series* (pp. 191–203). Oxford: Archaeopress.

Brughmans, T. (2013a). Networks of networks: a citation network analysis of the adoption, use and adaptation of formal network techniques in archaeology. *Literary and Linguistic Computing, The Journal of Digital Scholarship in the Humanities*, *28*(4), 538–562. doi:10.1093/llc/fqt048

Brughmans, T. (2013b). Review of I. Malkin 2011. A Small Greek World. Networks in the Ancient Mediterranean. *The Classical Review*, *63*(01), 146–148. doi:10.1017/S0009840X12002776

Brughmans, T. (2013c). Thinking Through Networks: A Review of Formal Network Methods in Archaeology. *Journal of Archaeological Method and Theory*, *20*, 623–662. doi:10.1007/s10816-012-9133-8

Brughmans, T. (2014). The roots and shoots of archaeological network analysis: A citation analysis and review of the archaeological use of formal network methods. *Archaeological Review from Cambridge*, *29*(1), 18–41.

Brughmans, T., Collar, A., & Coward, F. (in press). Introduction: challenging network perspectives on the past. In T. Brughmans, A. Collar, & F. Coward (Eds.), *The Connected Past: challenges to network studies of the past*. Oxford: Oxford University Press.

Brughmans, T., Collar, A., Coward, F., & Mills, B. J. (2014). The connected past: critical and innovative approaches to networks in archaeology. Special issue of Journal of archaeological method and theory.

Brughmans, T., Collar, A., Coward, F., & Mills, B. J. (2014). Introduction: positioning networks in archaeology. *Journal of Archaeological Method and Theory*.

Brughmans, T., Keay, S., & Earl, G. (2012). Complex networks in archaeology: urban connectivity in Iron Age and Roman Southern Spain. *Leonardo*, *45*(3).

Brughmans, T., Keay, S., & Earl, G. P. (2014). Introducing exponential random graph models for visibility networks. *Journal of Archaeological Science*. doi:10.1016/j.jas.2014.05.027

Brughmans, T., & Poblome, J. (in press). Pots in space: understanding Roman pottery distribution from confronting exploratory and geographical network analyses. In E. Barker, S. Bouzarovski, C. Pelling, & L. Isaksen (Eds.), *New worlds out of old texts: developing techniques for the spatial analysis of ancient narratives*. Oxford: Oxford University Press.

Buldyrev, S. V, Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature*, *464*, 1025–1028.

Burt, R. S. (2011). *Neighbor Networks: Competitive Advantage Local and Personal*. Oxford: Oxford University Press.

Butts, C. (2006). Exact bounds for degree centralization. *Social Networks*, *28*(4), 283–296. doi:10.1016/j.socnet.2005.07.003

Butts, C. T. (2009). Revisiting the foundations of network analysis. *Science, 325*(5939), 414–6. doi:10.1126/science.1171022

Butzer, K. W. (2009). Evolution of an interdisciplinary enterprise: the Journal of Archaeological Science at 35years. *Journal of Archaeological Science*, *36*(9), 1842–1846. doi:10.1016/j.jas.2009.04.011

Carrington, P. J., Scott, J., & Wasserman, S. (2005). *Models and methods in social network analysis* (pp. xiv, 328). Cambridge ; New York: Cambridge University Press.

Chamorro, J. G. (1987). Survey of Archaeological Research on Tartessos. *American Journal of Archaeology*, *91*(2), 197–232.

Choi, J. M. (1988). Citation Analysis of Intra- and Interdisciplinary Communication Patterns of Anthropology in the U.S.A. *Behavioral & Social Sciences Librarian*, *6*(3-4), 65–84. doi:10.1300/J103v06n03_04

Chorley, R. J., & Haggett, P. (1967). *Models in geography*. London: Methuen.

Chorley, R. J., & Haggett, P. (1970). *Integrated models in geography*. London: Methuen.

Clark, B. M., & Clark, S. E. (1982). Core journals in anthropology. *Behavioral & Social Sciences Librarian*, *2*(2-3), 95–110. doi:10.1300/J103v02n02_06

Collar, A. (2007). Network theory and religious innovation. *Mediterranean Historical Review*, *22*(1), 149–162.

Collar, A. C. F. (2008). *Networks and religious innovation in the Roman empire*. Unpublished PhD thesis. University of Exeter.

Collar, A. (2013). *Religious networks in the Roman Empire. The spread of new ideas*. Cambridge: Cambridge University Press.

Conolly, J., & Lake, M. (2006). *Geographical information systems in archaeology*. Cambridge - New York: Cambridge University Press.

Conway, S. (2012). A Cautionary Note on Data Inputs and Visual Outputs in Social Network Analysis. *British Journal of Management*. doi:10.1111/j.1467-8551.2012.00835.x

Cortijo Cerezo, M. L. (1993). *La división territorial de la Baetica romana*. Córdoba.

Corzo, R. S., & Toscano, M. S. G. (1992). *Las Vias Romanas de Andalucía*. Sevilla: Consejería de Obras Públicas y Transportes, Junta de Andalucía.

Costa, L. D. F., Rodrigues, F. A., Travieso, G., & Villas Boas, P. R. (2007). Characterization of complex networks: A survey of measurements. *Advances in Physics*, *56*(1), 167–242. doi:10.1080/00018730601170527

Cowan, R., & Jonard, N. (2004). Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control*, *28*(8), 1557–1575. doi:10.1016/j.jedc.2003.04.002

Coward, F. (2010). Small worlds, material culture and ancient Near Eastern social networks. *Proceedings of the British Academy*, *158*, 449–479.

Coward, F. (2013). Grounding the net: social networks, material culture and geography in the Epipalaeolithic and early Neolithic of the Near East (~21 – 6,000 cal BCE). In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 247–280). Oxford: Oxford University Press.

Cowgill, G. L. (1990). Why Pearson's r is Not a Good Similarity Coefficient for Comparing Collections. *American Antiquity*, *55*(3), 512–521.

Crema, E. R., Bevan, A., & Lake, M. W. (2010). A probabilistic framework for assessing spatio-temporal point patterns in the archaeological record. *Journal of Archaeological Science*, *37*(5), 1118–1130.

Cross, M. C., & Hohenberg, P. C. (1993). Pattern-Formation Outside of Equilibrium. *Reviews of modern physics*, *65*, 851–1112.

Crowfoot, J. W., Crowfoot, G. M., & Kenyon, K. M. (1957). *The objects from Samaria (Samaria-Sebaste III)*. London.

Cunliffe, B. (1995). Diversity in the landscape: the geographical background to urbanism in Iberia. In B. Cunliffe & S. Keay (Eds.), *Social complexity and the development of towns in Iberia. From the Copper Age to the second century AD* (pp. 5–28). Oxford: The British Academy.

Cutting, M. (2003). The use of spatial analysis to study prehistoric settlement architecture. *Oxford Journal of Archaeology*, *22*, 1–21.

Cutting, M. (2006). More than one way to study a building: Approaches to prehistoric household and settlement space. *Oxford Journal of Archaeology*, *25*, 225–246.

Daraganova, G., Pattison, P., Koskinen, J., Mitchell, B., Bill, A., Watts, M., & Baum, S. (2012). Networks and geography: Modelling community network structures as the outcome of both spatial and network processes. *Social Networks*, *34*(1), 6–17. doi:10.1016/j.socnet.2010.12.001

Davidson, D. A. (1979). The Orcadian environment and cairn location. In C. Renfrew (Ed.), *Investigations in Orkney* (pp. 7–20). London: Thames and Hudson.

Davies, J. K. (2005). Linear and nonlinear flow models for ancient economies. In J. G. Manning & I. Morris (Eds.), *The ancient economy: evidence and models* (pp. 127–156). Stanford.

Degenne, A., & Forsé, M. (1994). *Les réseaux sociaux. Une approche structurale en sociologie*. Paris.

De Montis, A., & Caschili, S. (2012). Nuraghes and landscape planning: Coupling viewshed with complex network analysis. *Landscape and Urban Planning*, *105*(3), 315–324. doi:10.1016/j.landurbplan.2012.01.005

Domergue, C. (1990). *Les mines de la Péninsule Ibérique dans l´antiquité romaine*. Rome: École Française de Rome.

Doran, J. E., & Hodson, F. R. (1975). *Mathematics and Computers in Archaeology*. Cambridge (MA): Harvard University Press.

Downs, M. E. (1998). Turdetani and Bastetani: cultural identity in Iberian and early Roman Baetica. In S. Keay (Ed.), *The archaeology of Early Roman Baetica. Journal of Roman Archaeology Supplement Series 29* (pp. 39–53). Portsmouth - Rhode Island: Journal of Roman Archaeology.

Earl, G., & Keay, S. (2007). Urban Connectivity of Iberian and Roman Towns in Southern Spain : A Network Analysis Approach. In *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 77–86). Clark, J. T. Hagenmeister, E. M.

Eom, Y.-H., & Fortunato, S. (2011). Characterizing and Modeling Citation Dynamics. *PLoS ONE*, *6*(9), e24926. doi:10.1371/journal.pone.0024926

Epstein, A. L. (1969). Gossip, norms and social network. In J. C. Mitchell (Ed.), *Social networks in urban situations: analyses of personal relationships in Central African towns*. Manchester: Published for the Institute for Social Research University of Zambia by Manchester University press.

Erdkamp, P. (2013). How modern was the market economy of the Roman world? *Oeconomia. History, Methodology, Philosophy*, *3*(4).

Erdős, P., & Rényi, A. (1959). On random graphs. *Publicationes Mathematicae*, *6*, 290–297.

Erdős, P., & Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, *5*, 17–61.

Erdős, P., & Rényi, A. (1961). On the strength of connectedness of a random graph. *Acta Mathematica Scientia Hungarica*, *12*, 261–267.

Erickson, B. (1988). The relational basis of attitudes. In B. Wellman & S. D. Berkowitz (Eds.), *Social Structures: A Network Approach* (pp. 99–121). Cambridge: Cambridge University Press.

Escacena, J. L., & Belén, M. (1998). Pre-Roman Turdetania. In S. Keay (Ed.), *The archaeology of Early Roman Baetica. Journal of Roman Archaeology Supplement Series 29* (pp. 23–37). Portsmouth - Rhode Island: Journal of Roman Archaeology.

Evans, S., & Felder, K. (2014). Social network perspectives in archaeology. A special issue of Archaeological Review from Cambridge.

Evans, T., Knappett, C., & Rivers, R. (2009). Using statistical physics to understand relational space: a case study from Mediterranean prehistory. In D. Lane, D. Pumain, S. Van Der Leeuw, & G. West (Eds.), *Complexity Perspectives in Innovation* (Vol. d, pp. 451–479).

Everett, M. G., & Borgatti, S. P. (2005). Extending centrality. In P. J. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (pp. 57–76). Cambridge: Cambridge University Press.

Faust, K. (2005). Using correspondence analysis for joint displays of affiliation networks. In P. J. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (pp. 117–147). Cambridge: Cambridge University Press.

Fear, A. T. (1996). *Rome and Baetica: urbanization in Southern Spain c. 50 BC-AD 150*. Oxford: Oxford University Press.

Fenn, J., & Raskino, M. (2008). Mastering the hype cycle. How to choose the right innovation at the right time. Boston: Harvard Business Press.

Fentress, E., & Perkins, P. (1988). Counting African Red Slip Ware. In A. Mastino (Ed.), *L'Africa Romana: Atti del V Convegno di studio Sassari, 11-13 dicembre 1987* (pp. 205–214). Sassari.

Fisher, P. F. (1992). First experiments in viewshed uncertainty: simulating fuzzy viewsheds. *Photogrammetric Engineering and Remote Sensing*, *58*(3), 345–352. Retrieved from http://www.colorado.edu/geography/leyk/geog_5113/readings/fisher_1992_pers.pdf

Fisher, P. F. (1994). Probable and fuzzy models of the viewshed operation. In M. F. Worboys (Ed.), *Innovations in GIS: selected papers from the First national Conference on GIS Research UK* (pp. 161–175). London: Taylor & Francis.

Fisher, P. F. (1995). An exploration of probable viewsheds in landscape planning. *Environment and Planning B: Planning and Design*, *22*, 527–546.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*. Retrieved from http://www.sciencedirect.com/science/article/pii/S0370157309002841

Frank, S. a. (2009). The common patterns of nature. *Journal of evolutionary biology*, *22*(8), 1563–85. doi:10.1111/j.1420-9101.2009.01775.x

Fraser, D. (1983). *Land and society in Neolithic Orkney. BAR British Series 117*. Oxford: Archaeopress.

Freeman, L. (1979). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239. doi:10.1016/0378-8733(78)90021-7

Freeman, L. (2004). *The development of social network analysis*. Vancouver: Empirical Press.

Freeman, L. C. (2005). Graphic techniques for exploring social network data. In P. J. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (Vol. 5, pp. 248–268). Cambridge: Cambridge University Press. doi:10.3917/enje.005.0059

Fulford, M. G. (1987). Economic interdependence among urban communities of the Roman Mediterranean. *World archaeology*, *19*, 58–75.

Fulminante, F. (2014). The network approach: Tool or paradigm? *Archaeological Review from Cambridge*, *29*(1), 167–178.

García Fernández, F. J. (2003). *El poblamiento Turdetano en el Bajo Guadalquivir*. Universidad de Sevilla, Sevilla.

Garfield, E. (1979). It's a small world after all. In E. Garfield (Ed.), *Essays of an information scientist* (pp. 299–304). Philadelphia, PA: ISI.

Garfield, E., Irving, H. S., & Richard, J. T. (1964). *The use of citation data in writing the history of science. Director*. Philadelphia: Institute for scientific information.

Garfield, E., Pudovkin, A. I., & Istomin, V. S. (2003). Why do we need algorithmic historiography? *Journal of the American Society for Information Science and Technology*, *54*(5), 400–412.

Garner, R. (1967). A computer-oriented graph theoretic analysis of citation index structures. In B. Flood (Ed.), *Three drexel information science research studies* (pp. 3–46). Philadelphia: Drexel press.

Garnsey, E., & McGlade, J. (2006). *Complexity and Co-evolution: Continuity and Change in Socio-economic Systems*. Cheltenham: Elgar.

Garrido González, P. (2011). *La Ocupación Romana del Valle del Guadiamar y la conexión Minera. Unpublished PhD thesis. ISBN 9788469478226*. Sevilla, Universidad de Sevilla. Retrieved from http://fondosdigitales.us.es/tesis/tesis/1564/la-ocupacion-romana-del-valle-del-guadiamar-y-la-conexion-minera/

Gastner, M. T., & Newman, M. E. J. (2006). The spatial structure of networks. *The European Physical Journal B*, *49*(2), 247–252. doi:10.1140/epjb/e2006-00046-8

Golitko, M., Meierhoff, J., Feinman, G. M., & Williams, P. R. (2012). Complexities of collapse : the evidence of Maya obsidian as revealed by social network graphical analysis. *Antiquity*, *86*, 507–523.

Graham, S. (2006a). Networks, agent-based models and the Antonine Itineraries: implications for Roman archaeology. *Journal of Mediterranean Archaeology*, *19*(1), 45–64.

Graham, S. (2006b). *EX FIGLINIS, the network dynamics of the Tiber valley brick industry in the hinterland of Rome, BAR international series 1486*. Oxford: Archaeopress.

Graham, S. (2009). The space between: the geography of social networks in the Tiber valley. In F. Coarelli & H. Patterson (Eds.), *Mercator Placidissimus: The Tiber Valley in Antiquity New research in the upper and middle river valley*. Rome: Edizioni Quasar.

Graham, S., & Weingart, S. (2014). The Equifinality of Archaeological Networks: An Agent Based Exploratory Lab Approach. *Journal of Archaeological Method and Theory*.

Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, *78*(5), 1360–1380.

Granovetter, M. (1983). The strength of weak ties: A network theory revisited. *Sociological theory*, *1*(1), 201–233. doi:10.2307/202051

Granovetter, M. (1985). Economic Action and Social Structure: The Problem of Embeddedness. *The American Journal of Sociology*, *91*(3), 481–510.

Grau Mira, I. (2003). Settlement Dynamics and Social Organization in Eastern Iberia during the Iron Age (Eighth-Second Centuries BC). *Oxford Journal of Archaeology*, *22*(3), 261–279. doi:10.1111/1468-0092.00187

Grau Mira, I. (2004). La construcción del paisaje ibérico: aproximación SIG al territorio protohistórico de la Marina Alta. *SAGVNTVN (P.L.A.V.)*, *36*, 61–75.

Grau Mira, I. (2005). Romanization in Eastern Spain: a GIS approach to Late Iberian Iron Age landscape. In J.-F. Berger, F. Bertoncello, F. Braemer, D. Gourguen, & M. Gazenbeek (Eds.), *Temps et espaces de l'homme en société, analyses et modèles spatiaux en archéologie. XXVième rencontres internatioales d'archéologie et d'histoire d'Antibes* (pp. 325–334). Antibes: Éditions APDCA.

Greene, K. (1992). *Roman pottery*. London: Published for the Trustees of the British Museum by British Museum Press.

Greene, K. (2005). Roman pottery: models, proxies and economic interpretation. *Journal of Roman Archaeology*, *18*, 34–56.

Guardiola, X., Díaz-Guilera, A., Pérez, C., Arenas, A., & Llas, M. (2002). Modeling diffusion of innovations in a social network. *Physical Review E*, *66*(2). doi:10.1103/PhysRevE.66.026121

Hage, P., & Harary, F. (1983). *Structural models in anthropology*. Cambridge: Cambridge University Press.

Hage, P., & Harary, F. (1991). *Exchange in Oceania: a graph theoretic analysis*. Oxford: Clarendon Press.

Hage, P., & Harary, F. (1996). *Island networks: communication, kinship and classification structures in Oceania*. Cambridge: Cambridge University Press.

Haggett, P. (1965). *Locational analysis in human geography*. London: Wiley.

Haggett, P. (1970). Network models in geography. In R. J. Chorley & P. Haggett (Eds.), *Integrated models in geography* (pp. 609–668). London: Methuen.

Haggett, P., & Chorley, R. J. (1969). *Network Analysis in Geography*. London: Arnold.

Hanneman, R. A., & Riddle, M. (2005). *Introduction to social network methods*. *Network*. University of California Riverside, CA. Retrieved from http://wiki.gonzaga.edu/dpls707/images/6/6e/Introduction_to_Social_Network_Methods.pdf

Harary, F., Norman, R. Z., & Cartwright, D. (1965). *Structural models : an introduction to the theory of directed graphs*. New York: Wiley.

Hart, J. P., & Engelbrecht, W. (2011). Northern Iroquoian Ethnic Evolution: A Social Network Analysis. *Journal of Archaeological Method and Theory*. doi:10.1007/s10816-011-9116-1

Harrigan, N. (2007). PNet for dummies: An introduction to estimating exponential random graph (p*) models with PNet. Retrieved from http://ink.library.smu.edu.sg/soss_research/810/

Hayes, J. W. (1972). *Late Roman Pottery*. London.

Hayes, J. W., & British School at Rome. (1980). *A supplement to Late Roman pottery*. London: British School at Rome.

Hayes, J. W. (1985). Sigillate orientale. In *Enciclopedia dell'arte antica. classica e orientale. Atlante della forme ceramiche II: ceramica fine romana nel bacino mediterraneo (tardo ellenismo e primo imperio)* (pp. 1–96). Rome.

Hayes, J. W. (2008). *Roman pottery, fine-ware imports. The Athenian Agora, results of excavations conducted by the American School of Classical Studies at Athens, volume XXXII*. Princeton: The American School of Classical Studies at Athens.

Hillier, B., & Hanson, J. (1984). *The social logic of space*. Cambridge: Cambridge University Press.

Hiorns, R. W. (1971). Statistical studies in migration. In F. R. Hodson, D. G. Kendall, & P. Tăutu (Eds.), *Mathematics in the archaeological and historical sciences, Proceedings of the Anglo-Romanian conference, Mamaia 1970* (pp. 291–302). Edinburgh: Edinburgh University Press.

Hodder, I. (2012). *Entangled: an archaeology of the relationships between humans and things*. Malden, Ma.: Wiley-Blackwell.

Hodder, I., & Orton, C. (1976). *Spatial analysis in archaeology*. Cambridge: Cambridge University Press.

Homans, G. C. (1950). *The Human Group*. New York: Harcourt, Brace.

Huberman, B. A., & Adamic, L. A. (1999). Growth dynamics of the World-Wide Web. *Nature*, *401*, 131.

Huisman, M., & Duijn, M. A. J. van. (2005). Software for social network analysis. In P. J. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (pp. 270–316). Cambridge: Cambridge University Press.

Huisman, M., & van Duijn, M. A. J. (2011). A reader's guide to SNA software. In J. Scott & P. J. Carrington (Eds.), *The SAGE handbook of social network analysis* (pp. 578–600). London: Sage.

Hummon, N. P., & Doreian, P. (1989). Connectivity in a citation network: The development of DNA theory. *Social Networks*, *11*(1), 39–63.

Hummon, N. P., Doreian, P., & Freeman, L. C. (1990). Analyzing the Structure of the Centrality-Productivity Literature Created Between 1948 and 1979. *Science Communication*, *11*(4), 459–480. doi:10.1177/107554709001100405

Hunt, T. L. (1988). Graph theoretic network models for Lapita exchange: a trial application. In P. V. Kirch & T. L. Hunt (Eds.), *Archaeology of the Lapita cultural complex: a critical review* (pp. 135–155). Seattle: Thomas Burke Memorial Washington State Museum Research Reports no. 5.

Hutson, S. (2002). Gendered citation practices in American Antiquity and other archaeology journals. *American antiquity*, *67*(2), 331–342.

Hutson, S. R. (2006). Self-Citation in Archaeology: Age, Gender, Prestige, and the Self. *Journal of Archaeological Method and Theory*, *13*(1), 1–18. doi:10.1007/s10816-006-9001-5

Irwin, G. (1978). Pots and entrepots: a study of settlement , trade and the development of economic specialization in Papuan prehistory. *World Archaeology*, *9*(3), 299–319.

Irwin-Williams, C. (1977). A network model for the analysis of Prehistoric trade. In T. K. Earle & J. E. Ericson (Eds.), *Exchange systems in Prehistory* (pp. 141–151). New York: Academic Press.

Isaksen, L. (2007). Network analysis of transport vectors in Roman Baetica. In J. T. Clark & E. M. Hagenmeister (Eds.), *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 76–87). Budapest: Archaeolingua.

Isaksen, L. (2008). The application of network analysis to ancient transport geography : A case study of Roman Baetica. *Digital Medievalist*, *4*, http://www.digitalmedievalist.org/journal/4/isakse.

Isaksen, L. (2013). "O What A Tangled Web We Weave" – Towards a Practice That Does Not Deceive. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 43–70). Oxford: Oxford University Press.

Janssen, M. A. (2005). Review : Complex Systems and Archaeology : Empirical and Theoretical Applications. *American Journal of Archaeology*, *109*(3), 568–569.

Jenkins, D. (2001). A Network Analysis of Inka Roads, Administrative Centers, and Storage Facilities. *Ethnohistory*, *48*(4), 655.

Jenkins, D. (2008). Anthropology, mathematics, kinship: a tribute to the anthropologist Per Hage and his work with the mathematician Frank Harary. *Mathematical Anthropology and Cultural Theory*, *2*(3), http://mathematicalanthropology.org/pdf/Jenkins070.

Jin, E. M., Girvan, M., & Newman, M. E. (2001). Structure of growing social networks. *Physical review. E, Statistical, nonlinear, and soft matter physics*, *64*(4 Pt 2), 046132.

Johnson, J. C. (1994). Anthropological contributions to the study of social networks: a review. In S. Wasserman & J. Galaskiewicz (Eds.), *Advances in social network analysis* (pp. 113–151). London: Sage.

Kapferer, B. (1969). Norms and the manipulation of relationships in a work context. In J. C. Mitchell (Ed.), *Social networks in urban situations: analyses of personal relationships in Central African towns*. Manchester: Published for the Institute for Social Research University of Zambia by Manchester University press.

Kenrick, P. M. (1981). Fine wares of the Hellenistic and Roman periods. In J. Matthers (Ed.), *The river Qoueiq, northern Syria, and its catchment. Studies arising from the Tell Rifa'at survey 1977-79 Part ii, BAR International Series 98(ii)* (Vol. ii, pp. 439–458). Oxford: Archaeopress.

Keay, S. (1997). Early Roman Italica and the romanisation of western Baetica. In A. Caballos & P. León (Eds.), *Italica MMCC. Actas de la fundacíon de Italica* (pp. 21–47). Sevilla: Consejería de Cultura de la Junta de Andalucía.

Keay, S. (1998a). Introduction: the archaeology of Early Roman Baetica. In S. Keay (Ed.), *The archaeology of Early Roman Baetica. Journal of Roman Archaeology Supplement Series 29* (pp. 11–22). Portsmouth - Rhode Island: Journal of Roman Archaeology.

Keay, S. (1998b). The development of towns in Early Roman Baetica. In S. Keay (Ed.), *The archaeology of Early Roman Baetica. Journal of Roman Archaeology Supplement Series 29* (pp. 55–86). Portsmouth - Rhode Island.

Keay, S. (1998c). *The archaeology of Early Roman Baetica*. Portsmouth - Rhode Island: Journal of Roman Archaeology.

Keay, S., & Earl, G. (2006). Inscriptions and Social Networks in western Baetica. In A. Sartori & A. Valvo (Eds.), *Hiberia Italia, Italia-Hiberia. Hiberia-Italia, Italia-Hiberia: Convegno Internazionale di Epigrafia e Storia Antica Milan, Italy, Cisalpino (Acta et Studia 2)* (pp. 269–290). Milan: Cisalpino.

Keay, S., & Earl, G. P. (2011). Towns and territories in Roman Baetica. In A. Bowman & A. Wilson (Eds.), *Settlement, urbanization, and population* (pp. 276–316). Oxford: Oxford University Press.

Kendall, D. (1969). Incidence matrices, interval graphs and seriation in archeology. *Pacific Journal of mathematics*, *28*(3), 565–570.

Kendall, D. G. (1971a). Seriation from abundance matrices. In F. R. Hodson, D. G. Kendall, & P. Tăutu (Eds.), *Mathematics in the archaeological and historical sciences, Proceedings of the Anglo-Romanian conference, Mamaia 1970* (pp. 215–252). Edinburgh: Edinburgh University Press.

Kendall, D. G. (1971b). Maps from marriages: an application of non-metric multi-dimensional scaling to parish register data. In F. R. Hodson, D. G. Kendall, & P. Tăutu (Eds.), *Mathematics in the archaeological and historical sciences, Proceedings of the Anglo-Romanian conference, Mamaia 1970* (pp. 303–318). Edinburgh: Edinburgh University Press.

Kenyon, K. M. (1957). Roman and later wares 1. Terra Sigiata, 2. stratified groups. In J. W. Crowfoot, G. M. Crowfoot, & K. M. Kenyon (Eds.), *The objects from Samaria (Samaria-Sebaste III)* (pp. 281–306). London.

Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, *46*(5), 604–632.

Kleinberg, J. M. (2000). Navigation in a small world. *Nature*, *406*(August), 845.

Knapp, R. (1983). *Roman Cordoba. History and archaeology*. Berkeley: University of California publications.

Knappett, C. (2011). *An archaeology of interaction: network perspectives on material culture and society*. Oxford: Oxford University Press.

Knappett, C. (2013a). *Network analysis in archaeology. New approaches to regional interaction*. Oxford: Oxford University Press.

Knappett, C. (2013b). Introduction: why networks? Carl Knappett. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 3–16). Oxford: Oxford University Press.

Knappett, C. (2014). What are social network perspectives in archaeology? *Archaeological Review from Cambridge*, *29*(1), 179–184.

Knappett, C. (2014). Analyser les réseaux du passé en archéologie et en histoire. Special issue of Nouvelles de l'archéologie.

Knappett, C. (in press). Networks in archaeology: between scientific method and humanistic metaphor. In T. Brughmans, A. Collar, & F. Coward (Eds.), *The Connected Past: challenges to network studies of the past*. Oxford: Oxford University Press.

Knappett, C., Evans, T., & Rivers, R. (2008). Modelling maritime interaction in the Aegean Bronze Age. *Antiquity*, *82*(318), 1009–1024.

Knappett, C., Evans, T., & Rivers, R. (2011). The Theran eruption and Minoan palatial collapse: new interpretations gained from modelling the maritime network. *Antiquity*, *85*(329), 1008–1023.

Knievel, J. E., & Kellsey, C. (2005). Citation analysis for collection development: a comparative study of eight humanities fields. *The Library Quarterly*, *75*(2), 142–168. doi:10.1086/431331

Knox, H., Savage, M., & Harvey, P. (2006). Social networks and the study of relations: networks as method, metaphor and form. *Economy and Society*, *35*(1), 113–140. doi:10.1080/03085140500465899

Kochen, M. (1989). *The Small World*. Norwood, New Jersey: Ablex.

Kohler, T. A. (2012). Complex systems and archaeology. In I. Hodder (Ed.), *Archaeological Theory Today II*. Cambridge: Polity Press.

Korte, C., & Milgram, S. (1970). Acquaintance networks between racial groups - application of the small world method. *Journal of personality and social psychology*, *15*(2), 101.

Koskinen, J., & Daraganova, G. (2013). Exponential random graph model fundamentals. In D. Lusher, J. Koskinen, & G. Robins (Eds.), *Exponential Random Graph Models for Social Networks* (pp. 49–76). Cambridge: Cambridge university press.

Koskinen, J., & Snijders, T. A. B. (2013). Simulation, estimation, and goodness of fit. In D. Lusher, J. Koskinen, & G. Robins (Eds.), *Exponential Random Graph Models for Social Networks* (pp. 141–166). Cambridge: Cambridge university press.

Lane, D., Van der Leeuw, S. E., Pumain, D., & West, G. (2009). *Complexity Perspectives in Innovation and Social Change*. Dordrecht: Springer.

Lawall, M. L., & Lund, J. (2011). *Pottery in the Archaeological Record: A View from the Greek World (Gösta Enbom Monograph Series 1)*. Copenhagen.

Lazega, E., & van Duijn, M. A. J. (1997). Position in formal structure, personal characteristics and choices of advisors in a law firm: A logistic regression model for dyadic network data. *Social Networks*, *19*, 375–397.

Lemercier, C. (2012). Formal network methods in history: why and how? *Österreichische Zeitschrift für Geschichtswissenschaften*, *1*.

Le Roux, P. (1995). *Romains d'Espagne. Cités et politique dans les provinces*. Paris.

Leydesdorff, L., Hammarfelt, B., & Salah, A. (2011). The structure of the Arts & Humanities Citation Index: A mapping on the basis of aggregated citations among 1,157 journals. *Journal of the American Society for Information Science and Technology*, *62*(12), 2414–2426. doi:10.1002/asi

Lewit, T. (2011). Dynamics of fineware production and trade: the puzzle of supra-regional exporters. *Journal of Roman Archaeology*, *24*, 313–332.

Liljeros, F., Edling, C. R., Amaral, L. a, Stanley, H. E., & Aberg, Y. (2001). The web of human sexual contacts. *Nature*, *411*(6840), 907–8. doi:10.1038/35082140

Linmans, A. J. M. (2010). Why with bibliometrics the Humanities does not need to be the weakest link: Indicators for research evaluation based on citations, library holdings, and productivity measures. *Scientometrics*, *83*, 337–354.

Llobera, M. (2003). Extending GIS-based visual analysis: the concept of visualscapes. *International Journal of Geographical Information Science*, *17*(1), 25–48. doi:10.1080/13658810210157732

Llobera, M. (2007). Reconstructing visual landscapes. *World Archaeology*, *39*(1), 51–69. doi:10.1080/00438240601136496

Lock, G., & Pouncett, J. (2007). Network analysis in archaeology session introduction: and introduction to network analysis. In J. T. Clark & E. M. Hagenmeister (Eds.), *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 71–73). Budapest: Archaeolingua.

Loeschke, S. (1912). Sigillata-töpfereien in Tschandarli. *Mitteilungen des deutschen Instituts, Athenische Abteilung*, *37*, 342–407.

Lund, J. (2002). The onthogenesis of Cypriot Sigillata. In A. Rathje, M. Nielsen, & B. Bundgaard Rasmussen (Eds.), *Pots for the living, pots for the dead, Acta Hyperborea 9* (pp. 185–223). Copenhagen.

Lund, J. (2003). Eastern Sigillata B. A ceramic fine ware industry in the political and commercial landscape of the Eastern Mediterranean. In C. Abadie-Reynal (Ed.), *Les céramiques en Anatolie aux époques hellenistiques et romaine, Varia Anatolica 15* (pp. 125–136). Paris.

Lusher, D., Koskinen, J., & Robins, G. (2013). *Exponential Random Graph Models for Social Networks*. Cambridge: Cambridge university press.

Lusher, D., & Robins, G. (2013). Formation of social network structure. In D. Lusher, J. Koskinen, & G. Robins (Eds.), *Exponential Random Graph Models for Social Networks* (pp. 16–28). Cambridge: Cambridge University Press.

Malafouris, L. (2013). *How things shape the mind: a theory of material engagement*. Cambridge (MA): MIT Press.

Malkin, I. (2011). *A small Greek world: networks in the Ancient Mediterranean*. Oxford - New York: Oxford University Press.

Markovsky, B., Willer, D., & Patton, T. (1988). Power relations in exchange networks. *American Sociological Review*, *53*, 220–236.

Marriner, N. (2009). Currents and trends in the archaeological sciences. *Journal of Archaeological Science*, *36*(12), 2811–2815. doi:10.1016/j.jas.2009.09.009

Marsden, P. V. (2002). Egocentric and sociocentric measures of network centrality. *Social Networks*, *24*(4), 407–422.

Maschner, H. D. G., & Bentley, R. A. (2003). The power law of rank and household on the North Pacific. In R. A. Bentley & H. D. G. Maschner (Eds.), *Complex systems and archaeology* (pp. 47–60). Salt Lake City: University of Utah Press.

May, K., Binding, C., Tudhope, D., & Jeffrey, S. (2012). Semantic Technologies Enhancing Links and Linked Data for Archaeological Resources. In M. Zhou, I. Romanowska, Z. Wu, P. Xu, & P. Verhagen (Eds.), *Revive the past: Proceedings of Computer Applications and Quantitative Techniques in Archaeology conference 2011, Beijing* (pp. 261–272). Amsterdam: Amsterdam University Press.

Mayer, M. (1977). *Aspectes de la Catalunya del segle II*. Barcelona: Institut d'Estudis Catalans, Secciò Històrico-Arqueològica.

Mayoral Herrera, V. (2004). *Paisajes agrarios y cambio social en Andalucía oriental entre los periodos ibérico y romano. Anejos de Aespa XXXI*. Madrid: CSIC.

McBrearty, S., & Brooks, A. S. (2000). The Revolution that Wasn't: a new interpretation of the origin of modern human behavior. *Journal of Human Evolution*, *39*, 453–563.

McGlade, J. (2005). Systems and simulacra: modeling, simulation, and archaeological interpretation. In H. D. G. Maschner & C. Chippindale (Eds.), *Handbook of archaeological methods* (pp. 554–602). Lanham - New York - Toronto - Oxford: AltaMira Press.

Menze, B. H., & Ur, J. a. (2012). Mapping patterns of long-term settlement in Northern Mesopotamia at a large scale. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(14), E778–87. doi:10.1073/pnas.1115472109

Meyer-Schlichtmann, C. (1988). *Die pergamenischen Sigillata aus der Stadtgrabung von Pergamon. Mitte 2. JH v. Chr. – Mitte 2. JH n. Chr.* Berlin/New York.

Michael, J. H., & Massey, J. G. (1997). Modelling the communication network in a sawmill. *Forest products journal*, *47*, 25–30.

Milgram, S. (1967). The Small World Problem. *Psychology Today*, *2*, 60–67.

Milgram, S. (1992). *The individual in a social world: essays and experiments, 2nd ed.* New York: McGraw-Hill.

Mills, B. J., Clark, J. J., Peeples, M. a, Haas, W. R., Roberts, J. M., Hill, J. B., … Shackley, M. S. (2013). Transformation of social networks in the late pre-Hispanic US Southwest. *Proceedings of the National Academy of Sciences of the United States of America*, 1–6. doi:10.1073/pnas.1219966110

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science*, *298*, 824–27.

Mitcham, J. (2002). In search of a defensible site: a GIS analysis of Hampshire hillforts. In D. Wheatley, G. Earl, & S. Poppy (Eds.), *Contemporary Themes in Archaeological Computing* (pp. 73–79). Oxford: Oxbow.

Mitchell, J. D. (1969). *Social networks in urban situations: analyses of personal relationships in Central African towns*. Manchester: Published for the Institute for Social Research University of Zambia by Manchester University press.

Mitchell, J. C. (1974). Social networks. *Annual Review of Anthropology*, *3*, 279–299.

Mitchell, M. (2009). *Complexity: a guided tour*. New York: Oxford University Press.

Mizoguchi, K. (2009). Nodes and edges: A network approach to hierarchisation and state formation in Japan. *Journal of Anthropological Archaeology*, *28*(1), 14–26. doi:10.1016/j.jaa.2008.12.001

Mol, A. A. A. (2014a). The connected caribbean. A socio-material network approach to patterns of homogeneity and diversity in the pre-colonial period. Leiden: Sidestone press.

Mol, A. A. A., Hofman, C. L., & Hoogland, M. (2014b). Remotely Local: An ego-network model of the 14th century site of Kelbey's Ridge, Saba. *Journal of Archaeological Method and Theory*.

Moreno, J. L. (1934). *Who shall survive?* Washington DC: Nervous and Mental Disease Publishing Company.

Moreno, J. L. (1946). Sociogram and sociomatrix: a note to the paper by Forsyth and Katz. *Sociometry*, *9*, 348–349.

Moreno, J. L. (1960). *The sociometry reader*. Glencoe Ill.: The Free Press.

Moreno, J. L., & Jennings, H. H. (1938). Statistics of social configurations. *Sociometry*, *1*, 342–374.

Moreno Escobar, M. C. (2014). *Patrones de asentamiento en la Bética romana. Un estudio del proceso de romanización desde el análisis arqueológico espacial. Unpublished Doctoral thesis. Universidad Pablo de Olavide*.

Morris, I., Saller, R. P., & Scheidel, W. (2007). Introduction. In W. Scheidel, I. Morris, & R. P. Saller (Eds.), *The Cambridge economic history of the Greco-Roman world* (pp. 1–12). Cambridge: Cambridge University Press.

Munson, J. L., & Macri, M. J. (2009). Sociopolitical network interactions: A case study of the Classic Maya. *Journal of Anthropological Archaeology*, *28*(4), 424–438. doi:10.1016/j.jaa.2009.08.002

Nederhof, A. (2006). Bibliometric monitoring of research performance in the Social Sciences and the Humanities : a review. *Scientometrics*, *66*(1), 81–100.

Newman, M. (2001). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E*, *64*(1), 1–8. doi:10.1103/PhysRevE.64.016131

Newman, M. (2003a). Ego-centered networks and the ripple effect. *Social Networks*, *25*(1), 83–95. doi:10.1016/S0378-8733(02)00039-4

Newman, M. E. J. (2003b). The structure and function of complex networks. *SIAM review*, *45*, 167–256.

Newman, M. E. J. (2010). *Networks: an introduction*. Oxford: Oxford University Press.

Nooy, W. de, Mrvar, A., & Batagelj, V. (2005). *Exploratory social network analysis with Pajek*. Cambridge ; New York: Cambridge University Press.

Orton, C. (1980). *Mathematics in archaeology*. London: Collins.

Orton, C., Tyers, P., & Vince, A. (1993). *Pottery in Archaeology*. Cambridge: Cambridge University Press.

Osa, M. (2003). *Solidarity and Contention: Networks of Polish Opposition*. Minneapolis, MN: University of Minnesota Press.

Östborn, P., & Gerding, H. (2014). Network Analysis of Archaeological Data: a Systematic Approach. *Journal of Archaeological Science*. doi:10.1016/j.jas.2014.03.015

Parkin, T. G., & Pomeroy, A. J. (2007). *Roman Social History. A sourcebook*. London: Paulissen.

Peacock, D. P. S. (1982). *Pottery in the Roman world: an ethnoarchaeological approach*. London - New York: Longman.

Peeples, M. A. (2011a). *Identity and Social Transformation in the Prehispanic Cibola World: A.D. 1150-1325*. *Archaeology*. Arizona State University.

Peeples, M. A. (2011b). R Script for Calculating the Brainerd-Robinson Coefficient of Similarity and Assessing Sampling Error. http://www.mattpeeples.net/br.html.

Peeples, M. a., & Roberts, J. M. (2013). To binarize or not to binarize: relational data and the construction of archaeological networks. *Journal of Archaeological Science*, *40*(7), 3001–3010. doi:10.1016/j.jas.2013.03.014

Peña, T. (2007). *Roman pottery in the archaeological record*. Cambridge: Cambridge University Press.

Peregrine, P. (1991). A Graph-Theoretic Approach to the Evolution of Cahokia. *American Antiquity*, *56*(1), 66–75. Retrieved from http://www.jstor.org/stable/280973

Phillips, S. C. (2011). *Networked glass: lithic raw material consumption and social networks in the Kuril islands, Far Eastern Russia. Unpublished PhD dissertation*. University of Washington, Seattle.

Pitts, F. R. (1965). A graph theoretic approach to historical geography. *The Professional Geographer*, *17*(5), 15–20. doi:10.1111/j.0033-0124.1965.015

Pitts, F. (1979). The medieval river trade network of Russia revisited. *Social Networks*, *1*(3), 285–292. doi:10.1016/0378-8733(78)90025-4

Poblome, J. (2006). Made in Sagalassos: modelling regional potential and constraints. In S. Menchelli & M. Pasquinucci (Eds.), *Territorio e produzioni ceramiche: paesaggi, economia e società in età romana. Convegno internazionale. Pisa, 20-22 ottobre 2005* (pp. 355–363). Pisa.

Poblome, J., Bes, P., & Lauwers, V. (2007). Winning hearts, minds and stomachs? Artefactual or artificial evidence for romanization. In M. Meyer (Ed.), *Neue Zeiten - Neue Sitten. Zu Rezeption und Integration römischen und italischen Kulturguts in Kleinasien (Wiener Forschungen zur Archäologie 12)* (pp. 221–232). Vienna.

Poblome, J., Bounegru, O., Degryse, P., Viaene, W., Waelkens, M., & Erdemgil, S. (2001). The sigillata manufactories of Pergamon and Sagalassos. *Journal of Roman Archaeology*, *14*, 143–165.

Poblome, J., Malfitana, D., & Lund, J. (2013). Pottery, Roman Empire. In R. S. Bagnall, K. Brodersen, C. B. Champion, A. Erskine, & S. R. Huebner (Eds.), *The Encyclopedia of Ancient History* (pp. 5482–5489). Chichester: Wiley-Blackwell. doi:10.1002/9781444338386.wbeah18104

Poblome, J., & Zelle, M. (2002). The table ware boom. A socio-economic perspective from western Asia Minor, Patris und Imperium. Kulturelle und 248 politische Identität in den Städt der römischen Provinzen Kleinasiens in. In C. Berns, H. von Hesberg, L. Vandeput, & M. Waelkens (Eds.), *Patris und Imperium. Kulturelle und politische Identität in den Städt der römischen Provinzen Kleinasiens in der frühen Kaiserzeit. Kolloquium Köln, November 1998* (pp. 275–287). Leuven/Paris/Dudley.

Ponsich, M. (1974). *Implantation rurale antique sur le Bas-Guadalquivir. T. 1*. Madrid: Laboratoire d'archéologie de la Casa de Velázquez.

Ponsich, M. (1979). *Implantation rurale antique sur le Bas-Guadalquivir. T. 2, La Campana, Palma del Rio, Posadas*. Paris: E. de Boccard.

Ponsich, M. (1987). *Implantation rurale antique sur le bas-Guadalquivir T. 3*. Madrid.

Ponsich, M. (1991). *Implantation rurale antique sur le bas-Guadalquivir T. 4*. Madrid: Laboratoire d'archéologie de la Casa de Velázquez.

Pool, I. de S., & Kochen, M. (1978). Contacts and influence. *Social Networks, 1*, 1–48. Retrieved from http://deepblue.lib.umich.edu/handle/2027.42/22688

Pouncett, J., & Lock, G. (2007). A Vector-based Approach to the Integration of Geophysical and Test-pitting Data : Phasing the South Cadbury Environs Project. In E. M. Clark, J. T. Hagenmeister (Ed.), *Digital* *discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 97–106). Budapest.

Premo, L. S. (2006). Agent-based models as behavioral laboratories for evolutionary anthropological research. *Arizona Anthropologist*, 91–113.

Rehren, T., Grattan, J., & Klein, R. (2008). Going strong, and growing. *Journal of Archaeological Science*, *35*, 94305. doi:10.1016/j.jas.2007.11.026

Remesal Rodríguez, J. (1998). Baetican olive oil and the Roman economy. In S. Keay (Ed.), *The archaeology of Early Roman Baetica. Journal of Roman Archaeology Supplement Series 29* (pp. 183–199). Portsmouth - Rhode Island.

Renfrew, C., & Sterud, G. (1969). Close-proximity Analysis: A rapid method for the ordering of archaeological materials. *American antiquity*, *34*, 265–277.

Reynolds, P. (1995). *Trade in the Western Mediterranean, AD 400-700: the ceramic evidence. BAR international series 604* (Vol. 604). Oxford: Archaeopress.

Richards, J., Jeffrey, S., Waller, S., Ciravegna, F., Chapman, S., & Zhang, Z. (2011). The Archaeology Data Service and the Archaeotools project: faceted classification and natural language processing. In E. C. Kansa, S. W. Kansa, & E. Watrall (Eds.), *Archaeology 2.0. New approaches to communication and collaboration* (pp. 31–56). Los Angeles: UCLA Press.

Richardson, J. S. (1986). *Hispaniae and the development of Roman imperialism 218-82 B.C.* Cambridge: Cambridge University Press.

Richardson, J. S. (1996). *The Romans in Spain*. Oxford: Blackwell.

Rihll, T. E., & Wilson, A. G. (1987). Spatial interaction and structural models in historical analysis: some possibilities and an example. *Histoire & Mesure*, *2*, 5–32.

Rihll, T. E., & Wilson, A. G. (1991). Modelling settlement structures in ancient Greece : new approaches to the polis. In J. Rich & A. Wallace-Hadrill (Eds.), *City and country in the ancient world* (pp. 59–96). London.

Riles, A. (2001). *The Network inside Out*. Ann Arbor, MI: University of Michigan Press.

Roberts, J. M., Mills, B. J., Clark, J. J., Haas, W. R., Huntley, D. L., & Trowbridge, M. a. (2012). A method for chronological apportioning of ceramic assemblages. *Journal of Archaeological Science*, *39*(5), 1513–1520. doi:10.1016/j.jas.2011.12.022

Roberts, S. G. B., Dunbar, R. I. M., Pollet, T. V., & Kuppens, T. (2009). Exploring variation in active network size: Constraints and ego characteristics. *Social Networks*, *31*(2), 138–146. doi:10.1016/j.socnet.2008.12.002

Robins, G. (2011). Exponential random graph models for social networks. In J. Scott & P. J. Carrington (Eds.), *The SAGE handbook of social network analysis* (pp. 484–500). London: Sage.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007a). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, *29*(2), 173–191. doi:10.1016/j.socnet.2006.08.002

Robins, G., Snijders, T., Wang, P., Handcock, M., & Pattison, P. (2007b). Recent developments in exponential random graph (p*) models for social networks. *Social Networks*, *29*(2), 192–215. doi:10.1016/j.socnet.2006.08.003

Robins, G., Pattison, P., & Wang, P. (2009). Closure, connectivity and degree distributions: Exponential random graph (p*) models for directed social networks. *Social Networks, 31*(2), 105–117. doi:10.1016/j.socnet.2008.10.006

Robinson, W. S. (1951). A method for chronologically ordering archaeological deposits. *American antiquity, 16*(4), 293–301.

Robinson, W., & Posten, P. (2005). Literature use of scholars publishing in leading anthropology periodicals. *Behavioral & Social Sciences Librarian*, (April 2013), 37–41.

Rodà, I. (1997). Hispania: from the second century AD to Late Antiquity. In M. Díaz-Andreu & S. Keay (Eds.), *The Archaeology of Iberia. The dynamics of change* (pp. 211–234). London - New York: Routledge.

Rogers, E. (1962). *Diffusion of Innovations*. New York: Free Press.

Rogers, E. M. (1979). Network analysis of the diffusion of innovations. In P. W. Holland & S. Leinhardt (Eds.), *Perspectives on social network research* (pp. 137–164). New York: Academic Press.

Rogers, E. M. (1995). *Diffusion of innovations, 4th edition*. New York: The Free Press.

Roos, A. M., Schulz, H. D., Arteaga Matute, O., & Schulz, H. (1995). El problema del "Lacus Ligustinus". Investigaciones geoarqueológicas en torno a las marismas del bajo Guadalquivir. In *Actas del Congreso Conmemorativo del V Symposium Internacional de Prehistoria Peninsular. Tartessos: 25 años después, 1968-1993* (pp. 99–135). Jerez de la Frontera.

Rosenswig, R. (2005). A tale of two antiquities: Evolving editorial policies of the SAA journals. *The SAA Archeological Record, 5*(1), 15–21.

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences, 105*, 1118–1123.

Rothman, M. (1987). Graph theory and the interpretation of regional survey data. *Paléorient, 13*(2), 73–91. doi:10.3406/paleo.1987.4430

Rougé, J. (1966). *Recherches sur l'organisation du commerce maritime en Méditerranée sous l'empire romain*. Paris.

Ruestes Bitrià, C. (2008). A Multi-technique GIS Visibility Analysis for Studying Visual Control of an Iron Age Landscape. *Internet Archaeology, 23*, http://intarch.ac.uk/journal/issue23/4/index.html.

Ruffini, G. (2012). Review of Malkin, I. 2011 A Small Greek World: Networks in the Ancient Mediterranean. *American Historical Review*, 1643–1644.

Ruiz Rodríguez, A. (1997). The Iron Age Iberian peoples of the upper Guadalquivir valley. In M. Díaz-Andreu & S. Keay (Eds.), *The Archaeology of Iberia. The dynamics of change* (pp. 175–191). London - New York: Routledge.

Ruiz Rodríguez, A., & Molinos, M. (1993). *Iberos. Análisis arqueológico de un proceso histórico*. Barcelona: Crítica.

Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika, 31*, 581–603.

Santley, R. S. (1991). The structure of the Aztec transport network. In C. D. Trombold (Ed.), *Acient road networks and settlement hierarchies in the New World* (pp. 198–210). Cambridge: Cambridge University Press.

Scheidel, W., Morris, I., & Saller, R. P. (2007). *The Cambridge economic history of the Greco-Roman world*. Cambridge: Cambridge University Press.

Schich, M., & Coscia, M. (2011). Exploring Co-Occurrence on a Meso and Global Level Using Network Analysis and Rule Mining. In *Proceedings of the ninth workshop on mining and Learning with Graphs (MLG '11)*. San Diego: ACM.

Schich, M., Hidalgo, C. A. ., Lehmann, S. J. ., & Park, J. (2009). The Network of Subject Co-Popularity in Classical Archaeology. *Bollettino di archeologia on line*, 49–57.

Schiffer, M. B. (1972). Archaeological context and systemic context. *American antiquity*, *37*(2), 156–165.

Schneider, G. (1995). Roman red and black slipped pottery from NE-Syria and Jordan. First results of chemical analysis. In *Hellenistic and Roman pottery in the Eastern Mediterranean - Advances in scientific studies. Acts of the II Nieborów pottery workshop, Nieborów, 18-20 December 1993* (pp. 415–422). Warsaw.

Scott, J. (1991). *Social network analysis: a handbook*. London.

Scott, J. (2011). Social physics and social networks. In J. Scott & P. J. Carrington (Eds.), *The SAGE handbook of social network analysis* (pp. 55–66). London: SAGE.

Scott, J., & Carrington, P. J. (2011). *The SAGE handbook of social network analysis*. London: Sage.

Shalizi, C. (2011). Power Law Distributions, 1/f Noise, Long-Memory Time Series. Retrieved from http://cscs.umich.edu/~crshalizi/notebooks/power-laws.html

Shaw, B. D. (2010). Review of: The Roman Bazaar: a comparative study of trade and markets in a tributary empire. By Peter F. Bang. *Journal of interdisciplinary history*, *41*(1), 126–127.

Shemming, J., & Briggs, K. (2014). Anglo-saxon communication networks. Retrieved March 31, 2014, from http://keithbriggs.info/AS_networks.html

Shennan, S. (1997). *Quantifying archaeology. 2nd edition* (2nd ed.). Edinburgh: Edinburgh University Press.

Shuchat, A. (1984). Matrix and network models in Archaeology. *Mathematics Magazine*, *57*(1), 3–14.

Sillières, P. (1991). *Les voies de communication de l'Hispanie méridionale*. Paris: Diffusion De Boccard.

Silver, M. (2009). Historical otherness, the Roman Bazaar and primitivism. *Journal of Roman Archaeology*, *22*, 421–440.

Sindbæk, S. M. (2007a). Networks and nodal points: the emergence of towns in Early Viking Age Scandinavia. *Antiquity*, *81*(311), 119–132.

Sindbæk, S. M. (2007b). The Small World of the Vikings : Networks in Early Medieval Communication and Exchange. *Norwegian Archaeological Review*, *40*, 59–74.

Sindbæk, S. M. (2013). Broken links and black boxes: material affiliations and contextual network synthesis in the Viking world. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 71–94). Oxford: Oxford University Press.

Sinopoli, C. M. (1991). *Approaches to archaeological ceramics*. New York - London: Plenum.

Smoot, M., Ono, K., Ruscheinski, J., Wang, P.-L., & Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, *27*(3), 431–432.

Snijders, T. A. B. (2005). Models for longitudinal network data. In P. J. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (pp. 215–247). Cambridge: Cambridge University Press.

Snyder, D., & Kick, E. L. (1979). Structural position in the world system and economic growth, 1955-1970: A multiple-network analysis of transnational interactions. *American Journal of Sociology*, *84*(5), 1096–1126. doi:10.1086/226902

Sporns, O. (2002). Network analysis, complexity, and brain function. *Complexity*, *8*(1), 56–60. doi:10.1002/cplx.10047

Sporns, O., Tononi, G., & Edelman, G. M. (2000). Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cerebral cortex*, *10*(2), 127–41.

Sterud, E. (1978). Changing Aims of Americanist Archaeology: A Citations Analysis of American Antiquity. 1946-1975. *American Antiquity*, *43*(2), 294–302.

Stumpf, M. P. H., Wiuf, C., & May, R. M. (2005). Subnets of scale-free networks are not scale-free: sampling properties of networks. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(12), 4221–4224. doi:10.1073/pnas.0501179102

Swanson, S. (2003). Documenting prehistoric communication networks: A case study in the Paquimé polity. *American antiquity*, *68*(4), 753–767. Retrieved from http://www.jstor.org/stable/3557071

Temin, P. (2013). *The Roman Market Economy*. Princeton: Princeton University Press.

Terrell, J. E. (1976). Island biogeography and man in Melanesia. *Archaeology and physical anthropology in Oceania*, *11*(1), 1–17.

Terrell, J. E. (1977a). Geographic systems and human diversity in the North Solomons. *World archaeology*, *9*(1), 62–81.

Terrell, J. E. (1977b). Human biogeography in the Solomon Islands. *Fieldiana Anthropology*, *68*(1), 1–47.

Terrell, J. E. (2010a). Language and Material Culture on the Sepik Coast of Papua New Guinea: Using Social Network Analysis to Simulate, Graph, Identify, and Analyze Social and Cultural Boundaries Between Communities. *The Journal of Island and Coastal Archaeology*, *5*(1), 3–32.

Terrell, J. E. (2010b). Social network analysis of the genetic structure of Pacific islanders. *Annals of human genetics*, *74*(3), 211–232. doi:10.1111/j.1469-1809.2010.00575.x

Thomas, J. (1993). The politics of vision and the archaeologies of landscape. In B. Bender (Ed.), *Landscapes: Politics and Perspectives* (pp. 1–19). Oxford: Berg.

Thomas, J. (2004). *Archaeology and Modernity*. London: Routledge.

Thompson, L. (1940). *Southern Lau, Fiji: An ethnography. B. P. Bishop Museum Bulletin 162.* Honolulu.

Tilley, C. Y. (1994). *A Phenomenology of Landscape*. Oxford: Berg.

Tovar, A. (1974). *Iberische Landeskunde. 2. Teil: Die Völkerunddie Städte des antiken Hispanien, Band I: Baetica*. Baden-baden: Valentin Koerner.

Travers, J., & Milgram, S. (1969). An Experimental Study of the Small World Problem. *Sociometry*, *32*(4), 425–443.

Trillmich, W., & Zanker, P. (1990). *Stadtbild und Ideologie. Die Monumentalisierung hispanischer Städte zwischen Republik und Kaiserzeit*. Munich: Verlag der Bayerischen Akademie der Wissenschaften.

Turcotte, D. L. (1999). Self-organized criticality. *Reports on progress in physics*, *62*, 1377–1430.

Valente, T. W. (1995). *Network models of the diffusion of innovations*. Creskill, NJ: Hampton Press.

Valente, T. W. (2005). Network models and methods for studying the diffusion of innovations. In P. J. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (pp. 98–116). Cambridge: Cambridge University Press.

Van der Leeuw, S. E. (2013). Archaeology, networks, information processing, and beyond. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 335–348). Oxford: Oxford University Press.

Van Oyen, A. (in press). Networks or work-nets? Actor-Network Theory and multiple social topologies in the production of Roman terra sigillata. In T. Brughmans, A. Collar, & F. Coward (Eds.), *The Connected Past: challenging networks in archaeology and history*. Oxford: Oxford Univeristy Press.

Ventura, A., León, P., & Márquez, C. (1998). Roman Córdoba in the light of recent archaeological research. In S. Keay (Ed.), *The archaeology of Early Roman Baetica. Journal of Roman Archaeology Supplement Series 29* (pp. 87–107). Portsmouth - Rhode Island: Journal of Roman Archaeology.

Verhagen, P., Brughmans, T., Nuninger, L., & Bertoncello, F. (2013). The long and winding road: combining least cost paths and network analysis techniques for settlement location analysis and predictive modelling. *Proceedings of Computer Applications and Quantitative Techniques in Archaeology conference 2012, Southampton*, pp. 357–66. Amsterdam University Press: Amsterdam.

Victor, K., & Beaudry, M. (1992). Women's Participation in American Prehistoric and Historic Archaeology: A Comparative Look at the Journals American Antiquity and Historical Archaeology. In C. Claassen (Ed.), *Exploring Gender through Archaeology* (pp. 11–22). Madison, Wisconsin: Prehistory Press.

Wasserman, S., & Faust, K. (1994). *Social network analysis : methods and applications*. Cambridge: Cambridge University Press.

Wasserman, S., & Pattison, P. (1996). Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and p*. *Psychometrika*, *61*(3), 401–425.

Watts, D. J. (2003). *Six Degrees: The Science of a Connected Age*. London: Vintage.

Watts, D. J. (2004). The "new" science of networks. *Annual Review of Sociology*, *30*, 243–270.

Watts, D. J., Dodds, P. S., & Newman, M. E. J. (2002). Identity and search in social networks. *Science*, *296*(5571), 1302–5. doi:10.1126/science.1070120

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*(6684), 440–2. doi:10.1038/30918

Wells, P. S. (2005). Creating an Imperial Frontier: Archaeology of the Formation of Rome's Danube Borderland. *Journal of Archaeological Research*, *13*(1), 49–88. doi:10.1007/s10814-005-0810-y

Wernke, S. a. (2012). Spatial network analysis of a terminal prehispanic and early colonial settlement in highland Peru. *Journal of Archaeological Science*, *39*(4), 1111–1122. doi:10.1016/j.jas.2011.12.014

West, G. B., Brown, J. H., & Enquist, B. J. (1999). The Fourth Dimension of Life: Fractal Geometry and Allometric Scaling of Organisms. *Science*, *284*(5420), 1677–1679. doi:10.1126/science.284.5420.1677

Wheatley, D. W. (1995). Cumulative Viewshed Analysis: a GIS-based method for investigating intervisibility, and its archaeological application. In G. Lock & Z. Stančič (Eds.), *Archaeology and Geographical Information Systems: a European Perspective* (pp. 171–186). London: Taylor & Francis.

Wheatley, D., & Gillings, M. (2000). Vision, perception and GIS: developing enriched approaches to the study of archaeological visibility. In G. R. Lock (Ed.), (pp. 1–27). Amsterdam: IOS Press.

Wheatley, D., & Gillings, M. (2002). *Spatial technology and archaeology : the archaeological applications of GIS*. London - New York: Taylor & Francis.

White, D. (2009). Innovation in the context of networks, hierarchies, and cohesion. In D. Lane, D. Pumain, S. E. van der Leeuw, & G. West (Eds.), *Complexity Perspectives in Innovation and Social Change* (pp. 153–193). Berlin: Springer.

White, H. D. (2011). Scientific and scholarly networks. In J. Scott & P. J. Carrington (Eds.), *The SAGE handbook of social network analysis* (pp. 272–285). London: Sage.

White, H. C. (1981). Where do markets come from? *American Journal of Sociology*, *87*, 517–547.

White, H. C., Boorman, S. A., & Breiger, R. L. (1976). Social structure from multiple networks. I. Blockmodels of roles and positions. *The American Journal of Sociology*, *81*(4), 730–780.

White, J. G., Southgate, E., Thompson, J. N., & Brenner, S. (1986). The structure of the nervous system of the nematode C. Elegans. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *314*, 1–340.

Wilensky, U. (1999). NetLogo. http://ccl.northwestern.edu/netlogo/. *Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.*

Wilensky, U. (2005). NetLogo Preferential Attachment model. http://ccl.northwestern.edu/netlogo/models/PreferentialAttachment. *Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL*.

Willet, R. (2012). *Red slipped complexity. The socio-cultural context of the concept and use of tableware in the Roman East (second century BC - seventh century AD). Unpublished PhD thesis*. Katholieke Universiteit Leuven.

Wolfe, A. (1978). The rise of network thinking in anthropology. *Social Networks*, *1*, 53–64. Retrieved from http://www.sciencedirect.com/science/article/pii/0378873378900126

Wolfe, A. W. (2011). Anthropologist view of social network analysis and data mining. *SOCNET*, *1*, 3–19. doi:10.1007/s13278-010-0014-4

Wood, J. (1996). *The Geomorphological Characterisation of Digital Elevation Models. Unpublished PhD Thesis*. Leicester.

Xia, J. (2011). Open access for archaeological literature: a manager's perspective. In E. C. Kansa, S. W. Kansa, & E. Watrall (Eds.), *Archaeology 2.0. New approaches to communication and collaboration* (pp. 233–249). Los Angeles: UCLA Press.

Zabehlicky-Scheffenegger, S. (1995). Subsidiary factories of Italian sigillata potters: the Ephesian evidence. In H. Koester (Ed.), *Ephesos. Metropolis of Asia. An interdisciplinary approach to its archaeology, religion and culture (Harvard Theological Studies 41)* (pp. 217–228). Valley Forge: Trinity Press International.

Zabehlicky-Scheffenegger, S. (1996). Der Italiener in Ephesos. *Rei Cretariae Romanae Fautores*, *34*, 253–271.

Zhuang, E., Chen, G., & Feng, G. (2011). A network model of knowledge accumulation through diffusion and upgrade. *Physica A: Statistical Mechanics and its Applications*, *390*(13), 2582–2592. doi:10.1016/j.physa.2011.02.043

Zipf, G. K. (1949). *Human Behaviour and the Principle of Least Effort*. Reading (Mass): Addison-Wesley.

Zubrow, E. B. W. (1990). Modelling and prediction with geographic information systems: a demographic example from prehistoric and historic New York. In K. M. S. Allen, S. W. Green, & E. B. W. Zubrow (Eds.), *Interpreting space: GIS and archaeology* (pp. 307–318). London - New York - Philadelphia: Taylor & Francis.

FACULTY OF HUMANITIES

DEPARTMENT OF ARCHAEOLOGY

Evaluating network science in archaeology.

A Roman Archaeology perspective

by

Tom Brughmans

VOLUME II: APPENDICES

Thesis for the degree of Doctor of Philosophy

June 2014

9. Appendix I: literature review

9.1. Publications used to compile figure 5

Alexander, C. (2008). The Bedolina map: An Exploratory Network Analysis. In A. Posluschny, K. Lambers, & I. Herzog (Eds.), *Layers of perception: proceedings of the 35th international conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, Germany, April 2-6, 2007. (Kolloquien zur Vor- und Frühgeschichte, Vol. 10)* (pp. 366–371). Bonn: Dr. Rudolf Habelt GmbH.

Allen, K. M. S. (1990). Modelling early historic trade in the eastern Great Lakes using geographic information systems. In K. M. S. Allen, S. W. Green, & E. B. W. Zubrow (Eds.), *Interpreting space: GIS and archaeology* (pp. 319–329). London - New York - Philadelphia: Taylor & Francis.

Barthelemy, M. (2014). Discussion: social and spatial networks. *Nouvelles de l'archéologie, 135*, 51–61.

Bentley, R. A., & Maschner, H. D. G. (2003). *Complex systems and archaeology*. Salt Lake City: University of Utah Press.

Bentley, R. A., & Maschner, H. D. G. (2007). Complexity Theory. In R. A. Bentley, H. D. G. Maschner, & C. Chippendale (Eds.), *Handbook of archaeological theories* (Vol. 48, pp. 245–270). AltaMira Press. doi:10.1016/j.ijnurstu.2010.09.012

Bentley, R. A., & Shennan, S. J. (2003). Cultural Transmission and Stochastic Network Growth. *American Antiquity*, *68*(3), 459–485.

Bentley, R. A., & Shennan, S. J. (2005). Random Copying and Cultural Evolution. *Science*, *309*, 877–878. doi:10.1126/science.309.5736.874b

Bentley, R., Lake, M., & Shennan, S. (2005). Specialisation and wealth inequality in a model of a clustered economic network. *Journal of Archaeological Science*, *32*(9), 1346–1356. doi:10.1016/j.jas.2005.03.008

Bernardini, W. (2007). Jeddito yellow ware and Hopi social networks. *Kiva*, *72*(3), 295–328.

Bevan, A., & Crema, E. R. (2014). Une modélisation géographiquement explicite d'interaction culturelle. Dialectes crétois modernes, archéologie de l'âge du Bronze. *Nouvelles de l'archéologie*, *135*, 45–50.

Bevan, A., & Wilson, A. (2013). Models of settlement hierarchy based on partial evidence. *Journal of Archaeological Science*, *40*(5), 2415–2427.

Bintliff, J. (2004). Time, structure, and agency: the Annales, emergent complexity, and archaeology. In J. Bintliff (Ed.), *A companion to archaeology* (pp. 174–194). Oxford: Blackwell.

Blake, E. (2013). Social networks, path dependence, and the rise of ethnic groups in pre-Roman Italy. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 203–222). Oxford: Oxford University Press.

Blake, E. (2014). Dyads and triads in community detection: a view from the Italian Bronze Age. *Nouvelles de l'archéologie*, *135*, 28–31.

Branting, S. (2007). Using an Urban Street Network and a PGIS-T Approach to Analyze Ancient Movement. In E. M. Clark, J. T. Hagenmeister (Ed.), *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 87–96). Budapest.

Broodbank, C. (2000). *An island archaeology of the early Cyclades*. Cambridge: Cambridge University Press.

Brughmans, T. (2010). Connecting the dots: towards archaeological network analysis. *Oxford Journal of Archaeology*, *29*(3), 277–303.

Brughmans, T. (2012). Facebooking the Past : a critical social network analysis approach for archaeology. In A. Chrysanthi, M. P. Flores, & C. Papadopoulos (Eds.), *Thinking beyond the Tool: Archaeological Computing and the Interpretative Process. British Archaeological Reports International Series* (pp. 191–203). Oxford: Archaeopress.

Brughmans, T. (2013a). Networks of networks: a citation network analysis of the adoption, use and adaptation of formal network techniques in archaeology. *Literary and Linguistic Computing, The Journal of Digital Scholarship in the Humanities*, *28*(4), 538–562. doi:10.1093/llc/fqt048

Brughmans, T. (2013b). Thinking Through Networks: A Review of Formal Network Methods in Archaeology. *Journal of Archaeological Method and Theory*, *20*, 623–662. doi:10.1007/s10816-012-9133-8

Brughmans, T. (2014). The roots and shoots of archaeological network analysis: A citation analysis and review of the archaeological use of formal network methods. *Archaeological Review from Cambridge*, *29*(1), 18–41.

Brughmans, T., Isaksen, L., & Earl, G. (2012). Connecting the dots: an introduction to critical approaches in archaeological network analysis. In M. Zhou, I. Romanowska, Z. Wu, P. Xu, & P. Verhagen (Eds.), *Proceedings of Computer Applications and Quantitative Techniques in Archaeology conference 2011, Beijing*. Amsterdam: Amsterdam University Press.

Brughmans, T., Keay, S., & Earl, G. (2012). Complex networks in archaeology: urban connectivity in Iron Age and Roman Southern Spain. *Leonardo*, *45*(3).

Clarke, D. L. (1968). *Analytical archaeology*. London: Methuen.

Classen, E. (2004). Verfahren der “Sozialen Netzwerkanalyse” und ihre Anwendung in der Archäologie. *Archäologische Informationen*, *27*, 219–226.

Classen, E. (2008). Early Neolithic Social Networks in Western Germany. In A. Posluschny, K. Lambers, & I. Herzog (Eds.), *Layers of perception. Proceedings of the 35th international conference on computer applications and quantitative methods in archaeology (CAA), Berlin April 2-6 2007. Kolloquien zur Vor- und Frühgeschichte, Vol. 10.* Bonn: Dr. Rudolf Habelt GmbH.

Classen, E., & Zimmerman, A. (2004). Tessellation and triangulations. understanding Early Neolithic Social Networks. In K. F. Ausserer (Ed.), *Enter the Past. Proceedings of the 30th CAA conference held in Vienna, Austria, April 2003*. Oxford.

Cochrane, E. E., & Lipo, C. (2010). Phylogenetic analyses of Lapita decoration do not support branching evolution or regional population structure during colonization of Remote Oceania. *Proceedings of the Royal Society B*, *365*, 3889–3902.

Collar, A. (2007). Network Theory and Religious Innovation. *Mediterranean Historical Review*, *22*(1), 149–162. doi:10.1080/09518960701539372

Collar, A. (2013a). Re-thinking Jewish ethnicity through social network analysis. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 223–246). Oxford: Oxford University Press.

Collar, A. (2013b). *Religious networks in the Roman Empire. The spread of new ideas*. Cambridge: Cambridge University Press.

Collar, A., Brughmans, T., Coward, F., & Lemercier, C. (2014). Analyser les réseaux du passé en archéologie et en histoire. *Nouvelles de l’archéologie*, *135*, 9–13.

Collar, A. C. F. (2008). *Networks and religious innovation in the Roman empire*. University of Exeter.

Coward, F. (2010). Small worlds, material culture and ancient Near Eastern social networks. *Proceedings of the British Academy*, *158*, 449–479.

Coward, F. (2013). Grounding the net: social networks, material culture and geography in the Epipalaeolithic and early Neolithic of the Near East (~21-6,000 cal BCE). In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 247–280). Oxford: Oxford University Press.

Coward, F., & Gamble, C. (2008). Big brains, small worlds: material culture and the evolution of the mind. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *363*(1499), 1969–79. doi:10.1098/rstb.2008.0004

Doran, J. E., & Hodson, F. R. (1975). *Mathematics and computers in archaeology*. Edinburgh: Edinburgh University Press.

Dunn, H. (2014). Population genetics and the investigation of past human interactions. *Archaeological Review from Cambridge*, *29*(1), 42–66.

Earl, G., & Keay, S. (2007). Urban Connectivity of Iberian and Roman Towns in Southern Spain : A Network Analysis Approach. In *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 77–86). Clark, J. T. Hagenmeister, E. M.

Evans, S., & Felder, K. (2014). Making the connection: changing perspectives on social networks. *Archaeological Review from Cambridge*, *29*(1), 9–17.

Evans, T., Knappett, C., & Rivers, R. (2009). Using statistical physics to understand relational space: a case study from Mediterranean prehistory. In D. Lane, D. Pumain, S. Van Der Leeuw, & G. West (Eds.), *Complexity Perspectives in Innovation* (Vol. d, pp. 451–479).

Fulminante, F. (2012). Social Network Analysis and the Emergence of Central Places A Case Study from Central Italy (Latium Vetus). *BABESCH*, *87*, 27–53.

Fulminante, F. (2014). The network approach: Tool or paradigm? *Archaeological Review from Cambridge*, *29*(1), 167–178.

Gamble, C. (1999). *The Palaeolithic societies of Europe*. Cambridge: Cambridge University Press.

Gjesfjeld, E., & Phillips, S. C. (2013). Evaluating adaptive network strategies with geochemical sourcing data: a case study from the Kuril Islands. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 281–306). Oxford: Oxford University Press.

Golitko, M., Meierhoff, J., Feinman, G. M., & Williams, P. R. (2012). Complexities of collapse : the evidence of Maya obsidian as revealed by social network graphical analysis. *Antiquity*, *86*, 507–523.

Graham, S. (2005). Agent-based modelling, archaeology and social organisation: the robustness of Rome. *The Archaeological Computing Newsletter*, *63*, 1–6.

Graham, S. (2006a). Networks, agent-based models and the Antonine Itineraries: implications for Roman archaeology. *Journal of Mediterranean Archaeology*, *19*(1), 45–64.

Graham, S. (2006b). *EX FIGLINIS, the network dynamics of the Tiber valley brick industry in the hinterland of Rome, BAR international series 1486*. Oxford: Archaeopress.

Graham, S. (2006). Who's in Charge? Studying Social Networks in the roman Brick Industry in Central Italy. In C. Mattusch & A. Donohue (Eds.), *Common Ground: Archaeology, Art, Science, and Humanities - Proceedings of the XVI International Congress of Classical Archaeology, oxford* (pp. 359–362).

Graham, S. (2009). The space between: the geography of social networks in the Tiber valley. In F. Coarelli & H. Patterson (Eds.), *Mercator Placidissimus: The Tiber Valley in Antiquity New research in the upper and middle river valley*. Rome: Edizioni Quasar.

Graham, S. (2014). On connecting stamps - Network analysis and epigraphy. *Nouvelles de l'archéologie, 135,* 39–44.

Graham, S., & Steiner, J. (2007). TravellerSim: growing settlement structures and territories with agent-based modeling. In J. T. Clark & E. M. Hagenmeister (Eds.), *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 57–67). Budapest: Archaeolingua.

Hage, P., & Harary, F. (1983). *Structural models in anthropology*. Cambridge: Cambridge University Press.

Hage, P., & Harary, F. (1991). *Exchange in Oceania: a graph theoretic analysis*. Oxford: Clarendon Press.

Hage, P., & Harary, F. (1996). *Island networks: communication, kinship and classification structures in Oceania*. Cambridge: Cambridge University Press.

Hiorns, R. W. (1971). Statistical studies in migration. In F. R. Hodson, D. G. Kendall, & P. Tăutu (Eds.), *Mathematics in the archaeological and historical sciences, Proceedings of the Anglo-Romanian conference, Mamaia 1970* (pp. 291–302). Edinburgh: Edinburgh University Press.

Hunt, T. L. (1988). Graph theoretic network models for Lapita exchange: a trial application. In P. V. Kirch & T. L. Hunt (Eds.), *Archaeology of the Lapita cultural complex: a critical review* (pp. 135–155). Seattle: Thomas Burke Memorial Washington State Museum Research Reports no. 5.

Hutchinson, P. (1972). Networks and roman roads: A Further roman Network. *Area*, *4*(4), 279–280.

Irwin, G. (1978). Pots and entrepots: a study of settlement , trade and the development of economic specialization in Papuan prehistory. *World Archaeology*, *9*(3), 299–319.

Isaksen, L. (2007). Network analysis of transport vectors in Roman Baetica. In J. T. Clark & E. M. Hagenmeister (Eds.), *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 76–87). Budapest: Archaeolingua.

Isaksen, L. (2008). The application of network analysis to ancient transport geography : A case study of Roman Baetica. *Digital Medievalist*, *4*, http://www.digitalmedievalist.org/journal/4/isakse.

Isaksen, L. (2013). "O What A Tangled Web We Weave" – Towards a Practice That Does Not Deceive. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 43–70). Oxford: Oxford University Press.

Jenkins, D. (2001). A Network Analysis of Inka Roads, Administrative Centers, and Storage Facilities. *Ethnohistory*, *48*(4), 655.

Jiménez, D., & Chapman, D. (2002). An application of proximity graphs in archaeological spatial analysis. In D. Wheatley, G. Earl, & S. Poppy (Eds.), *Contemporary themes in archaeological computing* (pp. 90–99). Oxford: Oxbow Books.

Kendall, D. (1969). Incidence matrices, interval graphs and seriation in archeology. *Pacific Journal of mathematics*, *28*(3), 565–570.

Kendall, D. G. (1971a). Seriation from abundance matrices. In F. R. Hodson, D. G. Kendall, & P. Tăutu (Eds.), *Mathematics in the archaeological and historical sciences, Proceedings of the Anglo-Romanian conference, Mamaia 1970* (pp. 215–252). Edinburgh: Edinburgh University Press.

Kendall, D. G. (1971b). Maps from marriages: an application of non-metric multi-dimensional scaling to parish register data. In F. R. Hodson, D. G. Kendall, & P. Tăutu (Eds.), *Mathematics in the archaeological and historical sciences, Proceedings of the Anglo-Romanian conference, Mamaia 1970* (pp. 303–318). Edinburgh: Edinburgh University Press.

Knappett, C. (2011). *An archaeology of interaction: network perspectives on material culture and society*. Oxford: Oxford University Press.

Knappett, C. (2013a). *Network analysis in archaeology. New approaches to regional interaction*. Oxford: Oxford University Press.

Knappett, C. (2013b). Introduction: why networks? Carl Knappett. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 3–16). Oxford: Oxford University Press.

Knappett, C. (2014). Avant-propos. Dossier: Analyse des réseaux sociaux en archéologie. *Nouvelles de l'archéologie*, *135*, 5–8.

Knappett, C., Evans, T., & Rivers, R. (2008). Modelling maritime interaction in the Aegean Bronze Age. *Antiquity*, *82*(318), 1009–1024.

Knappett, C., Evans, T., & Rivers, R. (2011). The Theran eruption and Minoan palatial collapse: new interpretations gained from modelling the maritime network. *Antiquity*, *85*(329), 1008–1023.

Kohler, T. A. (2012). Complex systems and archaeology. In I. Hodder (Ed.), *Archaeological Theory Today II*. Cambridge: Polity Press.

Leidwanger, J. (2014). Maritime networks and economic regionalism in the Roman Eastern Mediterranean. *Nouvelles de l'archéologie*, *135*, 32–38.

Lock, G., & Pouncett, J. (2007). Network analysis in archaeology session introduction: and introduction to network analysis. In J. T. Clark & E. M. Hagenmeister (Eds.), *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 71–73). Budapest: Archaeolingua.

Mackie, Q. (2001). *Settlement archaeology in a fjordland archipelago : network analysis, social practice and the built environment of Western Vancouver Island, British Columbia, Canada since 2,000 BP. BAR international series 926*. Oxford: Archaeopress.

Menze, B. H., & Ur, J. a. (2012). Mapping patterns of long-term settlement in Northern Mesopotamia at a large scale. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(14), E778–87. doi:10.1073/pnas.1115472109

Mills, B. J., Clark, J. J., Peeples, M. a, Haas, W. R., Roberts, J. M., Hill, J. B., … Shackley, M. S. (2013). Transformation of social networks in the late pre-Hispanic US Southwest. *Proceedings of the National Academy of Sciences of the United States of America*, 1–6. doi:10.1073/pnas.1219966110

Mills, B. J., Roberts, J. M., Clark, J. J., Jr., W. R. H., Huntley, D., Peeples, M. A., … Breiger, R. L. (2013). The dynamics of social networks in the Late Prehispanic U.S. Southwest. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 181–202). Oxford: Oxford University Press.

Mizoguchi, K. (2009). Nodes and edges: A network approach to hierarchisation and state formation in Japan. *Journal of Anthropological Archaeology*, *28*(1), 14–26. doi:10.1016/j.jaa.2008.12.001

Mizoguchi, K. (2014). Evolution of prestige good systems: an application of network analysis to the transformation of communication systems and their media. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 151–179). Oxford: Oxford University Press.

Mol, A. A. A. (2014). Play-things and the origins of online networks: Virtual material culture in multiplayer games. *Archaeological Review from Cambridge*, *29*(1), 144–166.

Mol, A. A. A., & Mans, J. (2013). Old boy networks in the indigenous Caribbean. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 307–333). Oxford: Oxford University Press.

Müller, U. (2010). Zentrale Orte und Netzwerk Zwei Konzepte zur Beschreibung von Zentralität. In C. Dobiat, P. Ettel, & F. Fless (Eds.), *Zwischen Fjorden und Steppe. Festschrift fur Johan Callmer zum 65. Geburtstag* (pp. 57–67). Rahden: Verlag Marie Leidorf GmbH.

Müller, U. (2012). Networks of Towns – Networks of Periphery? Some Relations between the North European Medieval Town and its Hinterland. In S. Brather, U. Müller, & H. Steuer (Eds.), *Raumbildung durch Netzwerke? Der Ostseeraum zwischen Wikingerzeit un Spaetmittelalter aus archaeologischer und geschichtswissenschaftlicher Perspektive* (pp. 55–78). Bonn: Dr. Rudolf Habelt GmbH.

Munson, J. (2013). From Metaphors to Practice. Operationalizing network concepts for archaeological stratigraphy. *Journal of Archaeological Method and Theory*. doi:10.1007/s10816-013-9181-8

Munson, J. L., & Macri, M. J. (2009). Sociopolitical network interactions: A case study of the Classic Maya. *Journal of Anthropological Archaeology*, *28*(4), 424–438. doi:10.1016/j.jaa.2009.08.002

Nelson, E. S., & Kassabaum, M. C. (2014). Expanding social networks through ritual deposition: A case study from the Lower Mississippi Valley. *Archaeological Review from Cambridge*, *29*(1), 103–128.

Orser Jr., C. (2005). Network Theory and the Archaeology of Modern History. In *Global Archaeological Theory SE - 7* (pp. 77–95). Springer US. doi:10.1007/0-306-48652-0_7

Orton, C. (1980). *Mathematics in archaeology*. London: Collins.

Östborn, P., & Gerding, H. (2014). Network Analysis of Archaeological Data: a Systematic Approach. *Journal of Archaeological Science*. doi:10.1016/j.jas.2014.03.015

Peeples, M. A. (2011). *Identity and Social Transformation in the Prehispanic Cibola World: A.D. 1150-1325. Archaeology*. Arizona State University.

Peeples, M. a., & Roberts, J. M. (2013). To binarize or not to binarize: relational data and the construction of archaeological networks. *Journal of Archaeological Science*, *40*(7), 3001–3010. doi:10.1016/j.jas.2013.03.014

Peregrine, P. (1991). A Graph-Theoretic Approach to the Evolution of Cahokia. *American Antiquity*, *56*(1), 66–75.

Phillips, S. C. (2011). *Networked glass: lithic raw material consumption and social networks in the Kuril islands, Far Eastern Russia. Unpublished PhD dissertation*. University of Washington, Seattle.

Pouncett, J., & Lock, G. (2007). A Vector-based Approach to the Integration of Geophysical and Test-pitting Data : Phasing the South Cadbury Environs Project. In E. M. Clark, J. T. Hagenmeister (Ed.), *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 97–106). Budapest.

Riede, F. (2014). Eruptions and ruptures — a social network perspective on vulnerability and impact of the Laacher See eruption (c. 13,000 BP) on Late Glacial hunter-gatherers in northern Europe. *Archaeological Review from Cambridge*, *29*(1), 67–102.

Rivers, R., & Evans, T. S. (2013). What makes a site important? Centrality, gateways and gravity. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 125–150). Oxford: Oxford University Press.

Rivers, R., & Evans, T. S. (2014). New approaches to Archaic Greek settlement structure. *Nouvelles de l'archéologie*, *135*, 21–28.

Rothman, M. (1987). Graph theory and the interpretation of regional survey data. *Paléorient*, *13*(2), 73–91. doi:10.3406/paleo.1987.4430

Santley, R. S. (1991). The structure of the Aztec transport network. In C. D. Trombold (Ed.), *Acient road networks and settlement hierarchies in the New World* (pp. 198–210). Cambridge: Cambridge University Press.

Schich, M., & Coscia, M. (2011). Exploring Co-Occurrence on a Meso and Global Level Using Network Analysis and Rule Mining. In *Proceedings of the ninth workshop on mining and Learning with Graphs (MLG '11)*. San Diego: ACM.

Schich, M., Hidalgo, C. A. ., Lehmann, S. J. ., & Park, J. (2009). The Network of Subject Co-Popularity in Classical Archaeology. *Bollettino di archeologia on line*, 49–57.

Scholnick, J. B., Munson, J. L., & Macri, M. J. (2013). Positioning power in a multi-relational framework: a social network analysis of Classic Maya political rhetoric. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 95–124). Oxford: Oxford University Press.

Schweizer, T., Hage, P., Harary, F., Houseman, M., Kent, S., & Wolfe, A. W. (1997). Embeddedness of Ethnographic Cases: A Social Networks Perspective. *Current Anthropology*, *38*(5), 739–760.

Shuchat, A. (1984). Matrix and network models in Archaeology. *Mathematics Magazine*, *57*(1), 3–14.

Sindbæk, S. M. (2007a). Networks and nodal points: the emergence of towns in Early Viking Age Scandinavia. *Antiquity*, *81*(311), 119–132.

Sindbæk, S. M. (2007b). The Small World of the Vikings : Networks in Early Medieval Communication and Exchange. *Norwegian Archaeological Review*, *40*, 59–74.

Sindbæk, S. M. (2009). Open access, nodal points, and central places. Maritime communication and locational principles for coastal sites in south Scandinavia, c. AD 400–1200. *Estonian Journal of Archaeology*, *13*(2), 96. doi:10.3176/arch.2009.2.02

Sindbæk, S. M. (2013). Broken links and black boxes: material affiliations and contextual network synthesis in the Viking world. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 71–94). Oxford: Oxford University Press.

Stoner, J. (2014). "Extending the self" through material culture: Private letters and personal relationships in second-century AD Egypt. *Archaeological Review from Cambridge*, *29*(1), 129–143.

Swanson, S. (2003). Documenting prehistoric communication networks: A case study in the Paquimé polity. *American antiquity*, *68*(4), 753–767.

Terrell, J. E. (1976). Island biogeography and man in Melanesia. *Archaeology and physical anthropology in Oceania*, *11*(1), 1–17.

Terrell, J. E. (1977). Geographic systems and human diversity in the North Solomons. *World archaeology*, *9*(1), 62–81.

Terrell, J. E. (1977). Human biogeography in the Solomon Islands. *Fieldiana Anthropology*, *68*(1), 1–47.

Terrell, J. E. (2010). Language and Material Culture on the Sepik Coast of Papua New Guinea: Using Social Network Analysis to Simulate, Graph, Identify, and Analyze Social and Cultural Boundaries Between Communities. *The Journal of Island and Coastal Archaeology*, *5*(1), 3–32.

Terrell, J. E. (2010). Social network analysis of the genetic structure of Pacific islanders. *Annals of human genetics*, *74*(3), 211–232. doi:10.1111/j.1469-1809.2010.00575.x

Terrell, J. E. (2013). Social network analysis and the practice of history. In *Network analysis in archaeology. New approaches to regional interaction* (pp. 17–42).

Van der Leeuw, S. E. (2013). Archaeology, networks, information processing, and beyond. In C. Knappett (Ed.), *Network analysis in archaeology. New approaches to regional interaction* (pp. 335–348). Oxford: Oxford University Press.

Van Oyen, A. (2014). Les acteurs-réseaux en archéologie: état de la question et perspectives futures. *Nouvelles de l'archéologie*, *135*, 14–20.

Verhagen, P., Brughmans, T., Nuninger, L., & Bertoncello, F. (2013). The long and winding road: combining least cost paths and network analysis techniques for settlement location analysis and predictive modelling. In *Proceedings of Computer Applications and Quantitative Techniques in Archaeology conference 2012, Southampton* (pp. 357–366). Amsterdam: Amsterdam University Press.

Wernke, S. a. (2012). Spatial network analysis of a terminal prehispanic and early colonial settlement in highland Peru. *Journal of Archaeological Science*, *39*(4), 1111–1122. doi:10.1016/j.jas.2011.12.014

Zubrow, E. B. W. (1990). Modelling and prediction with geographic information systems: a demographic example from prehistoric and historic New York. In K. M. S. Allen, S. W. Green, & E. B. W. Zubrow (Eds.), *Interpreting space: GIS and archaeology* (pp. 307–318). London - New York - Philadelphia: Taylor & Francis.

10.Appendix II: case study 1

10.1. Bibliography local scale

This appendix includes full bibliographic references for all of the 69 publications that make up the local scale of this citation network analysis.

Bentley, R. A. (2003a). Scale-free network growth and social inequality. In R. A. Bentley & H. D. G. Maschner (Eds.), *Complex systems and archaeology* (pp. 27-46). Salt Lake City: University of Utah Press.

Bentley, R. A. (2003b). An introduction to complex systems. In R. A. Bentley & H. D. G. Maschner (Eds.), *Complex systems and archaeology* (pp. 9-23). Salt Lake City: University of Utah Press.

Bentley, R., Lake, M., & Shennan, S. (2005). Specialisation and wealth inequality in a model of a clustered economic network. *Journal of Archaeological Science*, *32*(9), 1346-1356. doi:10.1016/j.jas.2005.03.008

Bentley, R., & Maschner, H. (2000). A growing network of ideas. *Fractals*, *8*(3), 227-237.

Bentley, R. A., & Maschner, H. D. G. (2001). Stylistic change as a self-organized critical phenomenon: An archaeological study in complexity. *Journal of Archaeological Method and Theory*, *8*(1), 35-66.

Bentley, R. A., & Maschner, H. D. G. (2003a). Preface: considering complexity theory in archaeology. In R. A. Bentley & H. D. G. Maschner (Eds.), *Complex systems and archaeology* (pp. 1-8). Salt Lake City: University of Utah Press.

Bentley, R. A., & Maschner, H. D. G. (2003b). Punctuated agency and change in archaeology. In R. A. Bentley & H. D. G. Maschner (Eds.), *Complex Systems and Archaeology* (pp. 75-77). Salt Lake City: University of Utah Press.

Bentley, R. A., & Shennan, S. J. (2003). Cultural Transmission and Stochastic Network Growth. *American Antiquity*, *68*(3), 459-485.

Bentley, R. A., & Shennan, S. J. (2005). Random Copying and Cultural Evolution. *Science*, *309*, 877-878. doi:10.1126/science.309.5736.874b

Bintliff, J. (2003). Searching for structure in the past - or was it “one damn thing after another”? In A. Bentley & H. D. G. Maschner (Eds.), *Complex systems and archaeology* (pp. 79-83). Salt Lake City: University of Utah Press.

Branting, S. (2007). Using an Urban Street Network and a PGIS-T Approach to Analyze Ancient Movement. In E. M. Clark, J. T. Hagenmeister (Ed.), *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 87-96). Budapest.

Brughmans, T. (2010). Connecting the dots: towards archaeological network analysis. *Oxford Journal of Archaeology*, *29*(3), 277-303.

Brughmans, T. (2012b). Facebooking the Past : a critical social network analysis approach for archaeology. In A. Chrysanthi, M. P. Flores, & C. Papadopoulos (Eds.), *Thinking beyond the Tool: Archaeological Computing and the Interpretative Process. British Archaeological Reports International Series*. Oxford: Archaeopress.

Cochrane, E. E., & Lipo, C. (2010). Phylogenetic analyses of Lapita decoration do not support branching evolution or regional population structure during colonization of Remote Oceania. *Proceedings of the Royal Society B*, *365*, 3889–3902.

Collar, A. (2007). Network Theory and Religious Innovation. *Mediterranean Historical Review*, *22*(1), 149-162. doi:10.1080/09518960701539372

Coward, F. (2010). Small worlds, material culture and ancient Near Eastern social networks. *Proceedings of the British Academy*, *158*, 449-479.

Coward, F., & Gamble, C. (2008). Big brains, small worlds: material culture and the evolution of the mind. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *363*(1499), 1969-79. doi:10.1098/rstb.2008.0004

Earl, G. P., Isaksen, L., Keay, S., Brughmans, T., & Potts, D. (2011). Computational methods on the Roman Port Networks project. In S. Keay (Ed.), *Rome, Portus and the Mediterranean. Archaeological Monographs of the British School at Rome*. London.

Earl, G., & Keay, S. (2007). Urban Connectivity of Iberian and Roman Towns in Southern Spain : A Network Analysis Approach. *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 77-86). Clark, J. T. Hagenmeister, E. M.

Evans, T., Knappett, C., & Rivers, R. (2009). Using statistical physics to understand relational space: a case study from Mediterranean prehistory. In D. Lane, D. Pumain, S. Van Der Leeuw, & G. West (Eds.), *Complexity Perspectives in Innovation* (Vol. d, pp. 451-479).

Gamble, C. (1998). Palaeolithic society and the release from proximity: a network approach to intimate relations. *World Archaeology*, *29*(3), 426-449.

Gamble, C. (1999). *The Palaeolithic societies of Europe*. Cambridge: Cambridge University Press.

Graham, S. (2006). *EX FIGLINIS, the network dynamics of the Tiber valley brick industry in the hinterland of Rome, BAR international series 1486*. Oxford: Archaeopress.

Graham, S. (2006). Networks, agent-based models and the Antonine Itineraries: implications for Roman archaeology. *Journal of Mediterranean Archaeology*, *19*(1), 45-64.

Graham, S. (2009). The space between: the geography of social networks in the Tiber valley. In F. Coarelli & H. Patterson (Eds.), *Mercator Placidissimus: The Tiber Valley in Antiquity New research in the upper and middle river valley*. Rome: Edizioni Quasar.

Hage, P., Harary, F., & James, B. (1986). Wealth and Hierarchy in the Kula Ring. *American Anthropologist*, *88*(1), 108-115. doi:10.1525/aa.1986.88.1.02a00070

Hamilton, M. J., Milne, B. T., Walker, R. S., Burger, O., & Brown, J. H. (2007). The complex structure of hunter-gatherer social networks. *Proceedings. Biological sciences / The Royal Society*, *274*(1622), 2195-202. doi:10.1098/rspb.2007.0564

Hart, J. P., & Engelbrecht, W. (2011). Northern Iroquoian Ethnic Evolution: A Social Network Analysis. *Journal of Archaeological Method and Theory*. doi:10.1007/s10816-011-9116-1

Hart, J. P. (2012). The effects of geographical distances on pottery assemblage similarities: a case study from Northern Iroquoia. *Journal of Archaeological Science*, *39*(1), 128-134. Elsevier Ltd. doi:10.1016/j.jas.2011.09.010

Hill, M. A. (2009). *The Benefit of the Gift: Exchange and Social Interaction in the Late Archaic Western Great Lakes. Unpublished PhD thesis*. Washington State University.

Hunt, T. L. (1988). Graph theoretic network models for Lapita exchange: a trial application. In P. V. Kirch & T. L. Hunt (Eds.), *Archaeology of the Lapita cultural complex: a critical review* (pp. 135–155). Seattle: Thomas Burke Memorial Washington State Museum Research Reports no. 5.

Irwin, G. (1978). Pots and entrepots: a study of settlement , trade and the development of economic specialization in Papuan prehistory. *World Archaeology*, *9*(3), 299–319. JSTOR.

Irwin-Williams, C. (1977). A network model for the analysis of Prehistoric trade. In T. K. Earle & J. E. Ericson (Eds.), *Exchange systems in Prehistory* (pp. 141-151). New York: Academic Press.

Isaksen, L. (2007). Network analysis of transport vectors in Roman Baetica. In J. T. Clark & E. M. Hagenmeister (Eds.), *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 76-87). Budapest: Archaeolingua.

Isaksen, L. (2008). The application of network analysis to ancient transport geography : A case study of Roman Baetica. *Digital Medievalist*, *4*, http://www.digitalmedievalist.org/journal/4/isakse.

Jenkins, D. (2001). A Network Analysis of Inka Roads, Administrative Centers, and Storage Facilities. *Ethnohistory*, *48*(4), 655. Duke Univ Press.

Jiménez, D., & Chapman, D. (2002). An application of proximity graphs in archaeological spatial analysis. In D. Wheatley, G. Earl, & S. Poppy (Eds.), *Contemporary themes in archaeological computing* (pp. 90-99). Oxford: Oxbow Books.

Keay, S., & Earl, G. (2006). Inscriptions and Social Networks in western Baetica. In A. Sartori & A. Valvo (Eds.), *Hiberia Italia, Italia-Hiberia. Hiberia-Italia, Italia-Hiberia: Convegno Internazionale di Epigrafia e Storia Antica Milan, Italy, Cisalpino (Acta et Studia 2)* (pp. 269-290). Milan: Cisalpino.

Knappett, C. (2011a). *An archaeology of interaction. Network perspectives on material culture and society*. Oxford - New York: Oxford University Press.

Knappett, C. (2011b). Meaning in miniature: semiotic networks in material culture. In M. Jessen, N. Johannsen, & H. J. Jensen (Eds.), *Excavating the mind*. Aarhus: Aarhus University Press.

Knappett, C. (2012). A regional network approach to Protopalatial complexity. In I. Schoep, P. Tomkins, & J. Driessen (Eds.), *Back to the beginning. Reassessing social and political complexity on Crete during the Early and Middle Bronze Age* (pp. 384-402). Oxford: Oxbow Books.

Knappett, C., Evans, T., & Rivers, R. (2008). Modelling maritime interaction in the Aegean Bronze Age. *Antiquity*, *82*(318), 1009–1024. Portland Press.

Knappett, C., & Nikolakopoulou, I. (2005). Exchange and affiliation networks in MBA southern Aegean: Crete, Akrotiri and Miletus. In R. Laffineur & E. Greco (Eds.), *Emporia: Aegens in east and west mediterranean* (pp. 175-184). Liège: University of Liège.

Kohler, T. A. (2012). Complex systems and archaeology. In I. Hodder (Ed.), *Archaeological Theory Today II*. Cambridge: Polity Press.

Lock, G., & Pouncett, J. (2007). Network analysis in archaeology session introduction: and introduction to network analysis. In J. T. Clark & E. M. Hagenmeister (Eds.), *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 71-73). Budapest: Archaeolingua.

Mackie, Q. (2001). *Settlement archaeology in a fjordland archipelago : network analysis, social practice and the built environment of Western Vancouver Island, British Columbia, Canada since 2,000 BP. BAR international series 926*. Oxford: Archaeopress.

Maschner, H. D. G., & Bentley, R. A. (2003). The power law of rank and household on the North Pacific. In R. A. Bentley & H. D. G. Maschner (Eds.), *Complex systems and archaeology* (pp. 47-60). Salt Lake City: University of Utah Press.

McGlade, J. (2003). The map is not the territory: complexity, complication, and representation. In R. A. Bentley & H. D. G. Maschner (Eds.), *Complex systems and archaeology* (pp. 111-119). Salt Lake City: University of Utah Press.

Milicic, B. (1993). Exchange and Social Stratification in the Eastern Adriatic: A Graph Theoretic Model. *Ethnology*, *32*, 375-395.

Mizoguchi, K. (2009). Nodes and edges: A network approach to hierarchisation and state formation in Japan. *Journal of Anthropological Archaeology*, *28*(1), 14-26. Elsevier Inc. doi:10.1016/j.jaa.2008.12.001

Munson, J. L., & Macri, M. J. (2009). Sociopolitical network interactions: A case study of the Classic Maya. *Journal of Anthropological Archaeology*, *28*(4), 424-438. Elsevier Inc. doi:10.1016/j.jaa.2009.08.002

Peregrine, P. (1991). A Graph-Theoretic Approach to the Evolution of Cahokia. *American Antiquity*, *56*(1), 66-75.

Pouncett, J., & Lock, G. (2007). A Vector-based Approach to the Integration of Geophysical and Test-pitting Data : Phasing the South Cadbury Environs Project. In E. M. Clark, J. T. Hagenmeister (Ed.), *Digital discovery: exploring new frontiers in human heritage. Proceedings of the 34th CAA conference, Fargo, 2006* (pp. 97-106). Budapest.

Rothman, M. (1987). Graph theory and the interpretation of regional survey data. *Paléorient*, *13*(2), 73-91. doi:10.3406/paleo.1987.4430

Santley, R. S. (1991). The structure of the Aztec transport network. In C. D. Trombold (Ed.), *Acient road networks and settlement hierarchies in the New World* (pp. 198-210). Cambridge: Cambridge University Press.

Schich, M., & Coscia, M. (2011). Exploring Co-Occurrence on a Meso and Global Level Using Network Analysis and Rule Mining. *Proceedings of the ninth workshop on mining and Learning with Graphs (MLG ’11)*. San Diego: ACM.

Schich, M., Hidalgo, C. A. ., Lehmann, S. J. ., & Park, J. (2009). The Network of Subject Co-Popularity in Classical Archaeology. *Bollettino di archeologia on line,* 49-57.

Shuchat, A. (1984). Matrix and network models in Archaeology. *Mathematics Magazine*, *57*(1), 3–14. JSTOR.

Sindbæk, S. M. (2007a). Networks and nodal points: the emergence of towns in Early Viking Age Scandinavia. *Antiquity*, *81*(311), 119-132.

Sindbæk, S. M. (2007b). The Small World of the Vikings : Networks in Early Medieval Communication and Exchange. *Norwegian Archaeological Review*, *40*, 59-74.

Swanson, S. (2003). Documenting prehistoric communication networks: A case study in the Paquimé polity. *American antiquity*, *68*(4), 753–767. JSTOR.

Terrell, J. E. (1976). Island biogeography and man in Melanesia. *Archaeology and physical anthropology in Oceania*, *11*(1), 1-17.

Terrell, J. E. (1977a). Geographic systems and human diversity in the North Solomons. *World archaeology*, *9*(1), 62–81. Taylor & Francis.

Terrell, J. E. (1977b). Human biogeography in the Solomon Islands. *Fieldiana Anthropology*, *68*(1), 1-47.

Terrell, J. E. (2010a). Language and Material Culture on the Sepik Coast of Papua New Guinea: Using Social Network Analysis to Simulate, Graph, Identify, and Analyze Social and Cultural Boundaries Between Communities. *The Journal of Island and Coastal Archaeology*, *5*(1), 3-32. doi:10.1080/15564890903142891

Terrell, J. E. (2010b). Social network analysis of the genetic structure of Pacific islanders. *Annals of human genetics*, *74*(3), 211-232. doi:10.1111/j.1469-1809.2010.00575.x

Welinder, S., & Griffin, W. L. (1984). Raw material sources and an exchange network of the earliest farming society in central Sweden. *World archaeology*, *16*(2), 174–185. Taylor & Francis.

Wells, P. S. (2005). Creating an Imperial Frontier: Archaeology of the Formation of Rome’s Danube Borderland. *Journal of Archaeological Research*, *13*(1), 49-88. doi:10.1007/s10814-005-0810-y

Zubrow, E. B. W. (1990). Modelling and prediction with geographic information systems: a demographic example from prehistoric and historic New York. In K. M. S. Allen, S. W. Green, & E. B. W. Zubrow (Eds.), *Interpreting space: GIS and archaeology* (pp. 307-318). London - New York - Philadelphia: Taylor & Francis.

10.2. Bibliography tables 4-6

This appendix includes full bibliographic references for all publications mentioned in tables 4-6.

Albert, R., & Barabási, A. (2002). Statistical mechanics of complex networks. *Reviews of modern physics*, *74*(January), 47-97.

Albert, R., Jeong, H., & Barabási, A. L. (1999). Internet: Diameter of the World-Wide Web. *Nature*, *401*(6749), 130–131. Nature Publishing Group.

Albert, R., Jeong, H., & Barabási, A. (2000). Error and attack tolerance of complex networks. *Nature*, *406*(6794), 378-82. doi:10.1038/35019019

Amaral, L. A. N., Scala, A., Barthélémy, M., & Stanley, H. E. (2000). Classes of Small-World Networks. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(21), 11149-11152.

Arthur, W. B. (1999). Complexity and the Economy. *Science*, *284*(5411), 107-109. doi:10.1126/science.284.5411.107

Bak, P., Tang, C., & Weisenfeld, K. (1987). Self-organized-criticality: an explanation of 1/F noise. *Physical Review Letters*, *59*, 381-384.

Banavar, J. R., Maritan, A., & Rinaldo, A. (1999). Size and form in efficient transportation networks. *Nature*, *399*(6732), 130-2. doi:10.1038/20144

Barabási, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, *286*(5439), 509-512. doi:10.1126/science.286.5439.509

Barabási, A.-L. (2002). *Linked: The New Science of Networks*. Cambridge, Massachusetts: Perseus.

Bentley, R. A., & Maschner, H. D. G. (2001). Stylistic change as a self-organized critical phenomenon: An archaeological study in complexity. *Journal of Archaeological Method and Theory*, *8*(1), 35-66.

Bentley, R. A., & Shennan, S. J. (2003). Cultural Transmission and Stochastic Network Growth. *American Antiquity*, *68*(3), 459-485.

Bentley, R. A. (2003a). Scale-free network growth and social inequality. In R. A. Bentley & H. D. G. Maschner (Eds.), *Complex systems and archaeology* (pp. 27-46). Salt Lake City: University of Utah Press.

Bentley, R. A. (2003b). An introduction to complex systems. In R. A. Bentley & H. D. G. Maschner (Eds.), *Complex systems and archaeology* (pp. 9-23). Salt Lake City: University of Utah Press.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.

Brughmans, T. (2010). Connecting the dots: towards archaeological network analysis. *Oxford Journal of Archaeology*, *29*(3), 277-303.

Brughmans, T. (2012b). Facebooking the Past : a critical social network analysis approach for archaeology. In A. Chrysanthi, M. P. Flores, & C. Papadopoulos (Eds.), *Thinking beyond the Tool: Archaeological Computing and the Interpretative Process. British Archaeological Reports International Series*. Oxford: Archaeopress.

Cochrane, E. E., & Lipo, C. (2010). Phylogenetic analyses of Lapita decoration do not support branching evolution or regional population structure during colonization of Remote Oceania. *Proceedings of the Royal Society B*, *365*, 3889–3902.

Coward, F., & Gamble, C. (2008). Big brains, small worlds: material culture and the evolution of the mind. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *363*(1499), 1969-79. doi:10.1098/rstb.2008.0004

Cross, M. C., & Hohenberg, P. C. (1993). Pattern-Formation Outside of Equilibrium. *Reviews of modern physics*, *65*, 851–1112.

Crow, E. L., & Shimizu, K. (1988). *Lognormal Distributions: Theory and Applications*. New York: Dekker.

Dietler, M. (1997). The Iron Age in Mediterranean France: Colonial encounters, entanglements, and transformation. *Journal of World Prehistory*, *11*, 269-358.

Erdős, P., & Renyi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, *5*, 17-61.

Evans, T. S. (2004). Complex networks. *Contemporary Physics*, *45*(6), 455-474. doi:10.1080/00107510412331283531

Faloutsos, M., Faloutsos, P., & Faloutsos, C. (1999). On power-law relationships of the internet topology. *Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication* (Vol. 29, pp. 251–262). ACM. doi:10.1145/316194.316229

Flinn, M. (1997). Culture and evolution of Social Learning. *Evolution and Human Behavior, 18*(23-67).

Fonagy, P., Gergely, G., & Target, M. (2007). The parent–infant dyad and the construction of the subjective self. *Journal of Child Psychology and Psychiatry*, *48*, 288-328.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*(3-5), 75-174. Elsevier B.V. doi:10.1016/j.physrep.2009.11.002

Gamble, C. (1998). Palaeolithic society and the release from proximity: a network approach to intimate relations. *World Archaeology*, *29*(3), 426-449.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360-1380. doi:10.1037/a0018761

Gronenborn, D. (1999). A Variation on a Basic Theme : The Transition to Farming in Southern Central Europe. *Journal of World Prehistory*, *13*(2).

Harary, F., & Norman, R. Z. (1953). *Graph Theory as a Mathematical Model in Social Science*. Ann Arbor: University of Michigan Press.

Harvey, P. H., & Pagel, M. D. (1991). *The comparative method in evolutionary biology*. Oxford: Oxford Univeristy Press.

Homans, G. C. (1950). *The Human Group*. New York: Harcourt, Brace.

Huberman, B. A., & Adamic, L. A. (1999). Growth dynamics of the World-Wide Web. *Nature*, *401*, 131.

Huberman, B. A., Pirolli, P. L. T., Pitkow, J. E., & Lukose, R. M. (1998). Strong Regularities in World Wide Web Surfing. *Science*, *280*(5360), 95-97. doi:10.1126/science.280.5360.95

Jenkins, D. (2001). A Network Analysis of Inka Roads, Administrative Centers, and Storage Facilities. *Ethnohistory*, *48*(4), 655. Duke Univ Press.

Jin, E. M., Girvan, M., & Newman, M. E. (2001). Structure of growing social networks. *Physical review. E, Statistical, nonlinear, and soft matter physics*, *64*(4 Pt 2), 046132.

Knappett, C. (2011a). *An archaeology of interaction: network perspectives on material culture and society*. Oxford: Oxford University Press.

Knappett, C. (2011b). Meaning in miniature: semiotic networks in material culture. In M. Jessen, N. Johannsen, & H. J. Jensen (Eds.), *Excavating the mind*. Aarhus: Aarhus University Press.

Kochen, M. (1989). *The Small World*. Norwood, New Jersey: Ablex.

Korte, C., & Milgram, S. (1970). Acquaintance networks between racial groups - application of the small world method. *Journal of personality and social psychology*, *15*(2), 101.

Liljeros, F., Edling, C. R., Amaral, L. a, Stanley, H. E., & Aberg, Y. (2001). The web of human sexual contacts. *Nature*, *411*(6840), 907-8. doi:10.1038/35082140

McBrearty, S., & Brooks, A. S. (2000). The Revolution that Wasn't: a new interpretation of the origin of modern human behavior. *Journal of Human Evolution*, *39*, 453-563.

Milgram, S. (1967). The Small World Problem. *Psychology Today*, *2*, 60-67.

Mitchell, J. D. (1969). *Social networks in urban situations: analyses of personal relationships in Central African towns*. Manchester: Published for the Institute for Social Research University of Zambia by Manchester University press.

Munson, J. L., & Macri, M. J. (2009). Sociopolitical network interactions: A case study of the Classic Maya. *Journal of Anthropological Archaeology*, *28*(4), 424-438. Elsevier Inc. doi:10.1016/j.jaa.2009.08.002

Newman, M., & Park, J. (2003). Why social networks are different from other types of networks. *Physical Review E*, *68*(3), 1-8. doi:10.1103/PhysRevE.68.036122

Newman, M. E. J. (2000). Models of the small world. *Journal of Statistical Physics*, *101*(3), 819–841. Springer.

Newman, M. E. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 404-9. doi:10.1073/pnas.021544898

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM review*, *45*, 167-256.

Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic Spreading in Scale-Free Networks. *Physical Review Letters*, *86*(14), 3200-3203. doi:10.1103/PhysRevLett.86.3200

Peters, R. H. (1986). *The Ecological Implications of Body Size*. Cambridge: Cambridge University Press.

Ravasz, E., & Barabási, A.-L. (2003). Hierarchical organization in complex networks. *Physical Review E*, *67*(2), 1-7. doi:10.1103/PhysRevE.67.026112

Rogers, E. (1962). *Diffusion of Innovations*. New York: Free Press.

Terrell, J. E., Hunt, T. L., & Gosden, C. (1997). The Dimensions of Social Life in the Pacific Human: Diversity and the Myth of the primitive isolate. *Current Anthropology*, *38*(2), 155-195.

Terrell, J. E. (2010b). Social network analysis of the genetic structure of Pacific islanders. *Annals of human genetics*, *74*(3), 211-232. doi:10.1111/j.1469-1809.2010.00575.x

Travers, J., & Milgram, S. (1969). An Experimental Study of the Small World Problem. *Sociometry*, *32*(4), 425-443.

Turcotte, D. L. (1999). Self-organized criticality. *Reports on progress in physics*, *62*, 1377-1430. APS.

Wasserman, S., & Faust, K. (1994). *Social network analysis : methods and applications*. Cambridge: Cambridge University Press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*(6684), 440-2. doi:10.1038/30918

Watts, D. J. (2004). The "New" Science of Networks. *Annual Review of Sociology*, *30*(1), 243-270. doi:10.1146/annurev.soc.30.020404.104342

Wells, P. S. (2005). Creating an Imperial Frontier: Archaeology of the Formation of Rome's Danube Borderland. *Journal of Archaeological Research*, *13*(1), 49-88. doi:10.1007/s10814-005-0810-y

West, G. B., Brown, J. H., & Enquist, B. J. (1997). A general model for the origin of allometric scaling laws in biology. *Science (New York, N.Y.)*, *276*(5309), 122-6.

Zipf, G. K. (1949). *Human Behaviour and the Principle of Least Effort*. Reading (Mass): Addison-Wesley.

11. Appendix III: case study 2

11.1. Sites index: dates, urban status, location, road and river networks

11.1.1. All 190 sites included in this case study in alphabetical order

The project IDs mentioned here are used in Figure 2. The ‘Municipality’ column lists the modern municipalities in which the sites are located. The urban status and ancient names of sites are provided when known. This list of urban status and ancient names was derived from Keay 1998, Appendix II, with minor updates and changes. The last five columns show whether we have evidence to believe a site was occupied during a certain period (indicated by “1”) or not (indicated by “0”).

Site Name	Project ID	Municipality	Urban status	Ancient name	Iberian	Republican	Early Imperial	Middle Imperial	Late Imperial
Alamillo	2	Osuna			1	1	1	1	1
Alcalá del Rio	4	Alcalá del Rio	Municipium	Ilipa Magna	1	1	1	1	1
Alcalá Morisco	5	Osuna			0	1	1	1	1
Alcaudete	6	Carmona			1	1	0	0	0
Alcolea del Rio	7	Alcolea del Rio	Municipium Flavium	Canama	1	1	1	1	1
Almadén de la Plata	9	Almadén de la Plata			0	0	1	1	0
Almodovar del Rio	308	Almodovar del Rio		Carbula					
Aparicio el Grande	11	Gilena		Unknown	0	0	1	1	1
Arcos de la Frontera	436	Arcos de la Frontera		Saudo (temporary)	1	1	1	1	1
Atalaya de la Moranilla	14	Écija			1	1	1	1	1
Aznalcazar	15	Aznalcazar		Olontigi	1	1	1	1	1
Bonanza	433	Sanlúcar de Barrameda		Lux Dubia (temporary)	0	0	1	0	0
Cagancha	286	Marchena			1	1	0	0	0
Cañada Real de Morón (IX)	18	Alcala de Guadaira			1	0	0	0	0
Cañada Real de Morón (VII)	19	Alcala de Guadaira			1	0	0	0	0
Cañada Real de Morón (VIII)	20	Alcala de Guadaira			1	0	0	0	0
Cañada Real de Morón (XII)	21	Alcala de Guadaira			1	0	0	0	0
Cantillana	23	Cantillana	Municipium Flavium	Naeva	1	1	1	1	1
Capaparda	24	Osuna			1	1	1	1	1
Carmona	25	Carmona		Carmo	1	1	1	1	1

Casablanca	28	Utrera			1	1	1	1	1
Casilla de Barrera	30	Écija			1	1	1	0	0
Castillo de Alhonoz	34	Écija			1	1	0	0	0
Castillo de la Monclova	153	Fuentes de Andalucia	Municipium Flavium	Obulcula	1	1	1	1	1
Castillo de Lora	171	Lora del Rio	Municipium Flavium	Axati	1	1	1	1	1
Castillo de Luna	278	Aznalcollar			1	1	1	0	0
Castillo de Mulva	35	Villanueva del Rio y Minas	Municipium Flavium	Munigua	1	1	1	1	0
Casulillas	36	Arahal (el)			1	1	1	1	1
Cerro Barrero (I)	39	Fuentes de Andalucia			1	0	0	0	0
Cerro Barrero (II)	40	Fuentes de Andalucia			1	1	0	0	0
Cerro de Atalaya	41	Campana (la)			1	1	1	1	1
Cerro de la Cabeza	42	Santiponce			1	0	0	0	0
Cerro de la Higuera	43	Estepa			0	0	1	1	1
Cerro de las Cabezas	46	Osuna			1	1	1	1	1
Cerro de las Cabezas (Sobarbina)	233	Olivares		Laelia	1	0	1	1	1
Cerro de las Catorce	47	Osuna			1	1	0	0	0
Cerro de las Vacas	48	Lebrija		Conobaria	1	1	1	1	1
Cerro de los Ladrillos	50	Arahal (el)			0	0	1	1	1
Cerro de los Villares	51	Estepa			0	0	1	1	1
Cerro del Bollo	303	Utrera		Unknown	1	1	1	1	1
Cerro del Calvario	56	Osuna			1	0	0	0	0
Cerro del Castillo (Gerena)	58	Gerena			1	1	1	0	0
Cerro del Cincho	59	Carmona	Municipium Flavium	Basilippo	1	1	1	1	0
Cerro del Manzano	60	Osuna			1	0	0	0	0
Cerro del Orégano	293	Marchena			1	1	1	1	1
Cerro del Pascualejo	259	Écija			1	1	1	1	1
Cerro del Tesoro	61	Luisiana (la) or Ecija??			1	1	1	1	1
Cerro Duran	62	Rubio (el)			1	1	1	1	1
Cerro Esperilla	432	Espera		Cappa					
Cerro Gordo	63	Gilena			1	0	0	0	0
Cerro Macareno	64	Rinconada (la)			1	1	0	0	0
Cerro Rubio	65	Fuentes de Andalucia			1	1	1	1	1

Cerros de San Ignacio	287	Marchena			1	1	1	0	0
Chiclana (II)	68	Écija			1	1	1	1	0
Clarebout II	299	Marchena			1	1	1	0	0
Constantina	304	Constantina		Iporca	0	0	1	0	0
Consuegra	10	Osuna		Munda	1	1	1	1	0
Córdoba (II)	70	Carmona			1	1	1	1	1
Coria del Rio	71	Coria del Rio		Caura	1	1	1	1	1
Cortijo Cabeza Del Sordo	72	Alcala de Guadaira			1	1	1	0	0
Cortijo de Alcofría	73	Écija			0	0	1	0	0
Cortijo de Carija (Bornos)	431	Espera	Municipium	Caris(s)a Aurelia					
Cortijo de Casablanca	444	Arcos de la Frontera		Lacca (temporary)					
Cortijo de los Cosmes	75	Écija	Municipium Flavium	Carruca	1	1	1	1	0
Cortijo de los Olivos (VII)	76	Carmona			1	1	1	1	1
Cortijo de Repla	78	Corrales (los)	Municipium Flavium	Ilipula Minor	1	1	1	1	1
Cortijo de Torneji Viejo	79	Carmona			1	1	1	1	1
Cortijo del Cerro	80	Carmona			1	1	1	1	1
Cortijo del Membrillo	301	Lora del Rio							
Cortijo Escalera	82	Fuentes de Andalucia			1	1	1	1	1
Cortijo Nuevo	83	Écija			1	1	1	1	1
Cortijos del Cerro (II)	84	Carmona			1	1	1	0	0
Cuevalonga (III)	85	Carmona			1	1	1	1	1
Dehesa de las Majadilas	86	Alcala de Guadaira			1	0	0	0	0
Doña Mencia	261	Écija			1	1	1	1	1
Donadio (II)	89	Fuentes de Andalucia			1	1	1	1	1
Écija	91	Écija	Colonia	Colonia Augusta Firma Astigi	1	1	1	1	1
El Arenal (II)	94	Fuentes de Andalucia			1	0	0	0	0
El Batán	262	Écija			1	1	1	1	0
El Casar	95	Utrera	Municipium Flavium	Salpensa	1	1	1	1	1
El Castillejo (Arva)	96	Alcolea del Rio	Municipium Flavium	Arva	1	1	1	1	1
El Castillejo (Écija)	263	Écija			1	1	1	1	1
El Chiste (II)	98	Carmona			1	1	1	1	1
El Chiste (V)	99	Carmona			1	1	1	1	1
El Gandul	103	Alcala de Guadaira		Irippo	1	1	1	1	1
El Grullo	104	Marchena			0	0	1	1	1
El Guijo	105	Écija			1	1	1	1	1
El Hachillo	106	Lora de Estepa		Olaura	0	1	1	1	1

El Mocho	107	Écija			0	1	1	1	1
El Molino Pintado	108	Montellano		Callenses Aeneanici	1	1	1	1	1
El Nuño	109	Écija			1	1	1	1	1
El Palomarejo	201	Écija			1	1	1	1	1
El Picate	110	Palma del Rio			1	1	1	1	1
El Santo Siervo	264	Écija			1	1	1	1	1
El Turruñuelo	116	Peñaflor			1	0	0	0	0
El Villar	265	Écija			1	1	1	1	1
Ermita de San Antón	118	Écija or Carmona?			1	1	1	1	1
Estepa	120	Estepa	Municipium Flavium	Ostippo	1	1	1	1	1
Friíllas	266	Écija			1	1	1	1	1
Garrotal	267	Écija			1	1	1	1	1
Gerena	122	Gerena		Ilse	1	1	1	1	1
Hacienda de Quintos	300	Dos Hermanas			0	1	1	1	1
Haza de las Piedras	128	Alcala de Guadaira			1	1	1	1	1
Herrera	129	Herrera			0	0	0	1	1
Huelva	395	Huelva		Onuba	1	1	1	1	1
Huerta del Caño	133	Écija			1	1	1	1	1
Isla del Castillo	136	Écija	Municipium Flavium	Segovia	1	1	1	1	1
La Alberquilla	268	Écija			0	0	1	1	1
La Alcuza	269	Écija			1	1	1	1	1
La Atalaya Chica	139	Casariche	Municipium Flavium	Ventipo	1	1	1	1	1
La Esclavitud	142	Coronil (el)			1	0	0	0	0
La Foronguilla	143	Coronil (el)			1	1	1	1	1
La Lombriz Ib	295	Marchena			1	1	1	0	0
La Molina (I)	152	Osuna			0	1	1	1	1
La Platera	154	Estepa			0	1	1	1	0
La Saetilla	217	Palma del Rio	Municipium		1	1	1	1	1
La Torre II	290	Marchena			1	1	1	0	0
La Zorrilla I	288	Marchena			1	0	1	0	0
La Zorrilla II	289	Marchena			1	0	1	0	0
Las Aguilillas (I)	161	Osuna			0	0	1	1	1
Las Aguzaderas	162	Coronil (el)			1	1	1	1	1
Las Albaidas (I)	163	Carmona			1	1	0	0	0
Las Animas	270	Écija			1	1	1	1	1
Las Cabezas de San Juan	164	Las Cabezas de San Juan (las)	Municipium Flavium	Conobaria	1	1	1	1	1
Las Mazmorras	166	Morón de la Frontera			1	1	1	1	1
Las Valbuenas	271	Écija			1	1	1	1	1
Lebrija	167	Lebrija		Nabrissa Veneria	1	1	1	1	1
Lomas del Castillo	168	Puebla de Cazalla (la)							
Lopera (III)	169	Utrera			1	1	1	0	0
Los Abades	272	Écija			1	1	1	1	1
Los Baldios	172	Saucejo (el)	Municipium Flavium	Irni	1	1	1	1	1

Los Canterones	173	Estepa			1	0	1	1	1
Los Castrejones	441	Aznalcollar			1	1	1	0	0
Los Felipes III	298	Marchena			1	1	0	0	0
Los Galindos II	294	Marchena			1	1	1	0	0
Los Medianos III	291	Marchena			1	1	1	0	0
Los Rodeos	179	Molares (los)			1	1	1	0	0
Malaver	181	Écija			1	1	1	1	1
Marchamorón (I)	182	Alcala de Guadaira			0	0	1	1	1
Marchenilla (II)	184	Alcala de Guadaira			1	1	1	1	1
Mesa de Lora	185	Lora del Rio	Municipium Flavium	Oducia	1	1	1	1	1
Mesa de Setefilla	186	Lora del Rio			1	1	1	1	1
Mesa del Almendro	93	Lora del Rio			1	1	1	0	0
Mesas de Asta	435	Jerez de la Frontera	Colonia	Hasta Regia (temporary)	1	1	1	1	1
Mochales	274	Écija			1	1	1	1	1
Molino de Pelay Correa	188	Alcala de Guadaira			0	0	1	1	1
Molino Hundido (I)	189	Alcala de Guadaira			1	1	1	0	0
Montemolín	192	Marchena			1	1	1	0	0
Moranilla	193	Écija			1	1	1	1	1
Morón de la Frontera	194	Morón de la Frontera		Lucurgentum Genius Iulii	1	1	1	1	1
Niebla	306	Niebla		Ilipula	1	1	1	1	1
Osuna	199	Osuna	Colonia	Colonia Genetiva Iulia (urbanorum ?) Urso	1	1	1	1	1
Palmilla (I)	200	Cabezas de San Juan (las)			1	1	1	1	1
Pancorvo	203	Montellano			1	1	0	0	0
Pavia	205	Esacena del Campo or Ecija?			1	1	1	1	1
Pedro Cruzado (I)	206	Estepa			1	0	0	0	0
Pedro Cruzado (II)	207	Estepa			1	1	1	0	0
Peñaflor	208	Peñaflor		Celti	1	1	1	1	1
Piedra Salmedina	443	Rota		Caepionis Turris (temporary)	0	1	0	0	0
Piesolo (I)	209	Alcala de Guadaira			1	1	1	1	1
Porcún I	284	Marchena			1	1	1	0	0
Pozo del Carretero	285	Marchena			1	1	0	0	0
Qiñones	212	Écija			1	1	1	1	1
Rancho Pozo Blanquillo	283	Marchena			0	0	1	1	1
Reinoso	215	Écija			0	0	1	1	1
Ruiz Sánchez	216	Écija			0	0	0	1	1

Salinas de la Torre	218	Écija			1	1	1	1	1
San José	220	Morón de la Frontera			1	0	0	0	0
San Juan de Aznalfarache	222	San Juan de Aznalfarache		Osset Constantia Iulia	1	1	1	1	0
San Pedro (I)	223	Fuentes de Andalucia			1	1	1	0	0
San Pedro (II)	224	Fuentes de Andalucia			1	1	0	0	0
San Pedro (VII)	227	Fuentes de Andalucia			1	1	0	0	0
Sanlucar de Barrameda	434	Sanlúcar de Barrameda			1	1	1	1	1
Santa Ana	297	Marchena			1	1	0	0	0
Santiponce	230	Santiponce	Municipium	Colonia Aelia Augusta Itálica	1	1	1	1	1
Sevilla	232	Sevilla	Colonia	Colonia Iulia Romula Hispalis	1	1	1	1	1
Sotillo Gallego	234	Écija			1	1	1	1	1
Tablada	235	Viso del Alcor (el)			1	1	0	0	1
Tarajal	296	Marchena			1	1	1	1	1
Tejada la Nueva	419	Paterna del Campo		ancient name?	1	1	1	1	1
Tejada La Vieja	316	Escacena del Campo			1	0	0	0	0
Tinajuela	236	Carmona			1	1	0	0	0
Torre Abad (I)	239	Alcala de Guadaira			1	1	1	1	0
Torre de la Membrilla	240	Alcala de Guadaira			1	1	1	1	0
Torre de los Herberos	241	Dos Hermanas	Municipium	Orippo	1	1	1	1	1
Torre del Aguila	242	Utrera		Siarum	1	1	1	1	1
Torres de Alocaz	244	Cabezas de San Juan (las)	Municipium	Ugia Marti	1	1	1	1	1
Vado de Quema	249	Aznalcazar			0	0	1	1	1
Vico	445	Marchena			1	1	1	0	0
Vistalegre I	292	Alcala de Guadaira			1	1	1	0	0

11.1.2. All 190 sites included in this case study sorted according to their project ID

The project IDs mentioned here are used in Figure 2. The ‘Municipality’ column lists the modern municipalities in which the sites are located. The urban status and ancient names of sites are provided when known. This list of urban status and ancient names was derived from Keay 1998, Appendix II, with minor updates and changes. The last five columns show whether we have evidence to believe a site was occupied during a certain period (indicated by "1") or not (indicated by "0").

Site Name	Project ID	Municipality	Urban status	Ancient name	Iberian	Republican	Early Imperial	Middle Imperial	Late Imperial
Alamillo	2	Osuna			1	1	1	1	1
Alcalá del Rio	4	Alcalá del Rio	Municipium	Ilipa Magna	1	1	1	1	1
Alcalá Morisco	5	Osuna			0	1	1	1	1
Alcaudete	6	Carmona			1	1	0	0	0
Alcolea del Rio	7	Alcolea del Rio	Municipium Flavium	Canama	1	1	1	1	1
Almadén de la Plata	9	Almadén de la Plata			0	0	1	1	0
Consuegra	10	Osuna		Munda	1	1	1	1	0
Aparicio el Grande	11	Gilena		Unknown	0	0	1	1	1
Atalaya de la Moranilla	14	Écija			1	1	1	1	1
Aznalcazar	15	Aznalcazar		Olontigi	1	1	1	1	1
Cañada Real de Morón (IX)	18	Alcala de Guadaira			1	0	0	0	0
Cañada Real de Morón (VII)	19	Alcala de Guadaira			1	0	0	0	0
Cañada Real de Morón (VIII)	20	Alcala de Guadaira			1	0	0	0	0
Cañada Real de Morón (XII)	21	Alcala de Guadaira			1	0	0	0	0
Cantillana	23	Cantillana	Municipium Flavium	Naeva	1	1	1	1	1
Capaparda	24	Osuna			1	1	1	1	1
Carmona	25	Carmona		Carmo	1	1	1	1	1
Casablanca	28	Utrera			1	1	1	1	1
Casilla de	30	Écija			1	1	1	0	0

Barrera									
Castillo de Alhonoz	34	Écija			1	1	0	0	0
Castillo de Mulva	35	Villanueva del Rio y Minas	Municipium Flavium	Munigua	1	1	1	1	0
Casulillas	36	Arahal (el)			1	1	1	1	1
Cerro Barrero (I)	39	Fuentes de Andalucia			1	0	0	0	0
Cerro Barrero (II)	40	Fuentes de Andalucia			1	1	0	0	0
Cerro de Atalaya	41	Campana (la)			1	1	1	1	1
Cerro de la Cabeza	42	Santiponce			1	0	0	0	0
Cerro de la Higuera	43	Estepa			0	0	1	1	1
Cerro de las Cabezas	46	Osuna			1	1	1	1	1
Cerro de las Catorce	47	Osuna			1	1	0	0	0
Cerro de las Vacas	48	Lebrija		Conobaria	1	1	1	1	1
Cerro de los Ladrillos	50	Arahal (el)			0	0	1	1	1
Cerro de los Villares	51	Estepa			0	0	1	1	1
Cerro del Calvario	56	Osuna			1	0	0	0	0
Cerro del Castillo (Gerena)	58	Gerena			1	1	1	0	0
Cerro del Cincho	59	Carmona	Municipium Flavium	Basilippo	1	1	1	1	0
Cerro del Manzano	60	Osuna			1	0	0	0	0
Cerro del	61	Luisiana (la) or			1	1	1	1	1

Tesoro		Ecija??							
Cerro Duran	62	Rubio (el)			1	1	1	1	1
Cerro Gordo	63	Gilena			1	0	0	0	0
Cerro Macareno	64	Rinconada (la)			1	1	0	0	0
Cerro Rubio	65	Fuentes de Andalucia			1	1	1	1	1
Chiclana (II)	68	Écija			1	1	1	1	0
Córdoba (II)	70	Carmona			1	1	1	1	1
Coria del Rio	71	Coria del Rio		Caura	1	1	1	1	1
Cortijo Cabeza Del Sordo	72	Alcala de Guadaira			1	1	1	0	0
Cortijo de Alcofría	73	Écija			0	0	1	0	0
Cortijo de los Cosmes	75	Écija	Municipium Flavium	Carruca	1	1	1	1	0
Cortijo de los Olivos (VII)	76	Carmona			1	1	1	1	1
Cortijo de Repla	78	Corrales (los)	Municipium Flavium	Ilipula Minor	1	1	1	1	1
Cortijo de Torneji Viejo	79	Carmona			1	1	1	1	1
Cortijo del Cerro	80	Carmona			1	1	1	1	1
Cortijo Escalera	82	Fuentes de Andalucia			1	1	1	1	1
Cortijo Nuevo	83	Écija			1	1	1	1	1
Cortijos del Cerro (II)	84	Carmona			1	1	1	0	0
Cuevalonga (III)	85	Carmona			1	1	1	1	1
Dehesa de las Majadilas	86	Alcala de Guadaira			1	0	0	0	0
Donadio (II)	89	Fuentes de Andalucia			1	1	1	1	1

Écija	91	Écija	Colonia	Colonia Augusta Firma Astigi	1	1	1	1	1
Mesa del Almendro	93	Lora del Rio			1	1	1	0	0
El Arenal (II)	94	Fuentes de Andalucia			1	0	0	0	0
El Casar	95	Utrera	Municipium Flavium	Salpensa	1	1	1	1	1
El Castillejo (Arva)	96	Alcolea del Rio	Municipium Flavium	Arva	1	1	1	1	1
El Chiste (II)	98	Carmona			1	1	1	1	1
El Chiste (V)	99	Carmona			1	1	1	1	1
El Gandul	103	Alcala de Guadaira		Irippo	1	1	1	1	1
El Grullo	104	Marchena			0	0	1	1	1
El Guijo	105	Écija			1	1	1	1	1
El Hachillo	106	Lora de Estepa		Olaura	0	1	1	1	1
El Mocho	107	Écija			0	1	1	1	1
El Molino Pintado	108	Montellano		Callenses Aeneanici	1	1	1	1	1
El Nuño	109	Écija			1	1	1	1	1
El Picate	110	Palma del Rio			1	1	1	1	1
El Turruñuelo	116	Peñaflor			1	0	0	0	0
Ermita de San Antón	118	Écija or Carmona?			1	1	1	1	1
Estepa	120	Estepa	Municipium Flavium	Ostippo	1	1	1	1	1
Gerena	122	Gerena		Ilse	1	1	1	1	1
Haza de las Piedras	128	Alcala de Guadaira			1	1	1	1	1
Herrera	129	Herrera			0	0	0	1	1
Huerta del Caño	133	Écija			1	1	1	1	1

Isla del Castillo	136	Écija	Municipium Flavium	Segovia	1	1	1	1	1
La Atalaya Chica	139	Casariche	Municipium Flavium	Ventipo	1	1	1	1	1
La Esclavitud	142	Coronil (el)			1	0	0	0	0
La Foronguilla	143	Coronil (el)			1	1	1	1	1
La Molina (I)	152	Osuna			0	1	1	1	1
Castillo de la Monclova	153	Fuentes de Andalucia	Municipium Flavium	Obulcula	1	1	1	1	1
La Platera	154	Estepa			0	1	1	1	0
Las Aguilillas (I)	161	Osuna			0	0	1	1	1
Las Aguzaderas	162	Coronil (el)			1	1	1	1	1
Las Albaidas (I)	163	Carmona			1	1	0	0	0
Las Cabezas de San Juan	164	Las Cabezas de San Juan (las)	Municipium Flavium	Conobaria	1	1	1	1	1
Las Mazmorras	166	Morón de la Frontera			1	1	1	1	1
Lebrija	167	Lebrija		Nabrissa Veneria	1	1	1	1	1
Lomas del Castillo	168	Puebla de Cazalla (la)							
Lopera (III)	169	Utrera			1	1	1	0	0
Castillo de Lora	171	Lora del Rio	Municipium Flavium	Axati	1	1	1	1	1
Los Baldios	172	Saucejo (el)	Municipium Flavium	Irni	1	1	1	1	1
Los Canterones	173	Estepa			1	0	1	1	1
Los Rodeos	179	Molares (los)			1	1	1	0	0
Malaver	181	Écija			1	1	1	1	1
Marchamorón (I)	182	Alcala de Guadaira			0	0	1	1	1
Marchenilla (II)	184	Alcala de			1	1	1	1	1

		Guadaira							
Mesa de Lora	185	Lora del Rio	Municipium Flavium	Oducia	1	1	1	1	1
Mesa de Setefilla	186	Lora del Rio			1	1	1	1	1
Molino de Pelay Correa	188	Alcala de Guadaira			0	0	1	1	1
Molino Hundido (I)	189	Alcala de Guadaira			1	1	1	0	0
Montemolín	192	Marchena			1	1	1	0	0
Moranilla	193	Écija			1	1	1	1	1
Morón de la Frontera	194	Morón de la Frontera		Lucurgentum Genius Iulii	1	1	1	1	1
Osuna	199	Osuna	Colonia	Colonia Genetiva Iulia (urbanorum ?) Urso	1	1	1	1	1
Palmilla (I)	200	Cabezas de San Juan (las)			1	1	1	1	1
El Palomarejo	201	Écija			1	1	1	1	1
Pancorvo	203	Montellano			1	1	0	0	0
Pavia	205	Esacena del Campo or Ecija?			1	1	1	1	1
Pedro Cruzado (I)	206	Estepa			1	0	0	0	0
Pedro Cruzado (II)	207	Estepa			1	1	1	0	0
Peñaflor	208	Peñaflor		Celti	1	1	1	1	1
Piesolo (I)	209	Alcala de Guadaira			1	1	1	1	1
Qiñones	212	Écija			1	1	1	1	1
Reinoso	215	Écija			0	0	1	1	1
Ruiz Sánchez	216	Écija			0	0	0	1	1
La Saetilla	217	Palma del Rio	Municipium		1	1	1	1	1

Salinas de la Torre	218	Écija			1	1	1	1	1
San José	220	Morón de la Frontera			1	0	0	0	0
San Juan de Aznalfarache	222	San Juan de Aznalfarache		Osset Constantia Iulia	1	1	1	1	0
San Pedro (I)	223	Fuentes de Andalucia			1	1	1	0	0
San Pedro (II)	224	Fuentes de Andalucia			1	1	0	0	0
San Pedro (VII)	227	Fuentes de Andalucia			1	1	0	0	0
Santiponce	230	Santiponce	Municipium	Colonia Aelia Augusta Itálica	1	1	1	1	1
Sevilla	232	Sevilla	Colonia	Colonia Iulia Romula Hispalis	1	1	1	1	1
Cerro de las Cabezas (Sobarbina)	233	Olivares		Laelia	1	0	1	1	1
Sotillo Gallego	234	Écija			1	1	1	1	1
Tablada	235	Viso del Alcor (el)			1	1	0	0	1
Tinajuela	236	Carmona			1	1	0	0	0
Torre Abad (I)	239	Alcala de Guadaira			1	1	1	1	0
Torre de la Membrilla	240	Alcala de Guadaira			1	1	1	1	0
Torre de los Herberos	241	Dos Hermanas	Municipium	Orippo	1	1	1	1	1
Torre del Aguila	242	Utrera		Siarum	1	1	1	1	1
Torres de Alocaz	244	Cabezas de San Juan (las)	Municipium	Ugia Marti	1	1	1	1	1

Vado de Quema	249	Aznalcazar	0	0	1	1	1
Cerro del Pascualejo	259	Écija	1	1	1	1	1
Doña Mencia	261	Écija	1	1	1	1	1
El Batán	262	Écija	1	1	1	1	0
El Castillejo (Écija)	263	Écija	1	1	1	1	1
El Santo Siervo	264	Écija	1	1	1	1	1
El Villar	265	Écija	1	1	1	1	1
Friíllas	266	Écija	1	1	1	1	1
Garrotal	267	Écija	1	1	1	1	1
La Alberquilla	268	Écija	0	0	1	1	1
La Alcuza	269	Écija	1	1	1	1	1
Las Animas	270	Écija	1	1	1	1	1
Las Valbuenas	271	Écija	1	1	1	1	1
Los Abades	272	Écija	1	1	1	1	1
Mochales	274	Écija	1	1	1	1	1
Castillo de Luna	278	Aznalcollar	1	1	1	0	0
Rancho Pozo Blanquillo	283	Marchena	0	0	1	1	1
Porcún I	284	Marchena	1	1	1	0	0
Pozo del Carretero	285	Marchena	1	1	0	0	0
Cagancha	286	Marchena	1	1	0	0	0
Cerros de San Ignacio	287	Marchena	1	1	1	0	0
La Zorrilla I	288	Marchena	1	0	1	0	0
La Zorrilla II	289	Marchena	1	0	1	0	0
La Torre II	290	Marchena	1	1	1	0	0
Los Medianos	291	Marchena	1	1	1	0	0

III									
Vistalegre I	292	Alcala de Guadaira			1	1	1	0	0
Cerro del Orégano	293	Marchena			1	1	1	1	1
Los Galindos II	294	Marchena			1	1	1	0	0
La Lombriz Ib	295	Marchena			1	1	1	0	0
Tarajal	296	Marchena			1	1	1	1	1
Santa Ana	297	Marchena			1	1	0	0	0
Los Felipes III	298	Marchena			1	1	0	0	0
Clarebout II	299	Marchena			1	1	1	0	0
Hacienda de Quintos	300	Dos Hermanas			0	1	1	1	1
Cortijo del Membrillo	301	Lora del Rio							
Cerro del Bollo	303	Utrera		Unknown	1	1	1	1	1
Constantina	304	Constantina		Iporca	0	0	1	0	0
Niebla	306	Niebla		Ilipula	1	1	1	1	1
Almodovar del Rio	308	Almodovar del Rio		Carbula					
Tejada La Vieja	316	Escacena del Campo			1	0	0	0	0
Huelva	395	Huelva		Onuba	1	1	1	1	1
Tejada la Nueva	419	Paterna del Campo		ancient name?	1	1	1	1	1
Cortijo de Carija (Bornos)	431	Espera	Municipium	Caris(s)a Aurelia					
Cerro Esperilla	432	Espera		Cappa					
Bonanza	433	Sanlúcar de Barrameda		Lux Dubia (temporary)	0	0	1	0	0
Sanlucar de Barrameda	434	Sanlúcar de Barrameda			1	1	1	1	1
Mesas de Asta	435	Jerez de la	Colonia	Hasta Regia	1	1	1	1	1

		Frontera	(temporary)					
Arcos de la Frontera	436	Arcos de la Frontera	Saudo (temporary)	1	1	1	1	1
Los Castrejones	441	Aznalcollar		1	1	1	0	0
Piedra Salmedina	443	Rota	Caepionis Turris (temporary)	0	1	0	0	0
Cortijo de Casablanca	444	Arcos de la Frontera	Lacca (temporary)					
Vico	445	Marchena		1	1	1	0	0

11.1.3. All sites with an urban status

SiteID	site_name	Urban status	Ancient name
4	Alcalá del Rio	Municipium	Ilipa Magna
7	Alcolea del Rio	Municipium Flavium	Canama
23	Cantillana	Municipium Flavium	Naeva
35	Castillo de Mulva	Municipium Flavium	Munigua
59	Cerro del Cincho	Municipium Flavium	Basilippo
75	Cortijo de los Cosmes	Municipium Flavium	Carruca
78	Cortijo de Repla	Municipium Flavium	Ilipula Minor
91	Écija	Colonia	Colonia Augusta Firma Astigi
95	El Casar	Municipium Flavium	Salpensa
96	El Castillejo (Arva)	Municipium Flavium	Arva
120	Estepa	Municipium Flavium	Ostippo
136	Isla del Castillo	Municipium Flavium	Segovia
139	La Atalaya Chica	Municipium Flavium	Ventipo
153	Castillo de la Monclova	Municipium Flavium	Obulcula
164	Las Cabezas de San Juan	Municipium Flavium	Conobaria
171	Castillo de Lora	Municipium Flavium	Axati

172	Los Baldios	Municipium Flavium	Irni
185	Mesa de Lora	Municipium Flavium	Oducia
199	Osuna	Colonia	Colonia Genetiva Iulia (urbanorum ?) Urso
217	La Saetilla	Municipium	
230	Santiponce	Municipium	Colonia Aelia Augusta Itálica
232	Sevilla	Colonia	Colonia Iulia Romula Hispalis
241	Torre de los Herberos	Municipium	Orippo
244	Torres de Alocaz	Municipium	Ugia Marti
431	Cortijo de Carija (Bornos)	Municipium	Caris(s)a Aurelia
435	Mesas de Asta	Colonia	Hasta Regia (temporary)

11.1.4. Sites on the Via Augusta

Site Name	Project ID	Municipality	Urban status	Ancient name
Carmona	25	Carmona		Carmo
Écija	91	Écija	Colonia	Colonia Augusta Firma Astigi
Castillo de la Monclova	153	Fuentes de Andalucia	Municipium Flavium	Obulcula
Sevilla	232	Sevilla	Colonia	Colonia Iulia Romula Hispalis
Torre de los Herberos	241	Dos Hermanas	Municipium	Orippo
Torres de Alocaz	244	Cabezas de San Juan (las)	Municipium	Ugia Marti
Mesas de Asta	435	Jerez de la Frontera	Colonia	Hasta Regia (temporary)

11.1.5. Sites on the river network

Site Name	Project ID	Municipality	Urban status	Ancient name

Alcalá del Rio	4	Alcalá del Rio	Municipium	Ilipa Magna
Alcolea del Rio	7	Alcolea del Rio	Municipium Flavium	Canama
Cantillana	23	Cantillana	Municipium Flavium	Naeva
Coria del Rio	71	Coria del Rio		Caura
Cortijo Nuevo	83	Écija		
Écija	91	Écija	Colonia	Colonia Augusta Firma Astigi
El Castillejo (Arva)	96	Alcolea del Rio	Municipium Flavium	Arva
Isla del Castillo	136	Écija	Municipium Flavium	Segovia
Castillo de Lora	171	Lora del Rio	Municipium Flavium	Axati
Peñaflor	208	Peñaflor		Celti
La Saetilla	217	Palma del Rio	Municipium	
San Juan de Aznalfarache	222	San Juan de Aznalfarache		Osset Constantia Iulia
Santiponce	230	Santiponce	Municipium	Colonia Aelia Augusta Itálica
Sevilla	232	Sevilla	Colonia	Colonia Iulia Romula Hispalis
Torre de los Herberos	241	Dos Hermanas	Municipium	Orippo
Doña Mencia	261	Écija		
Las Animas	270	Écija		
Las Valbuenas	271	Écija		

11.2. Cumulative viewsheds

Most of the areas that are highly visible in the cumulative viewshed results (Figs. 73-77) are areas with a high site density (Fig. 23). This is not so much a surprising result as it is a useful one, since it suggests that in the exploratory networks one can expect clusters with a higher density of arcs to appear. The areas of high visibility do differ somewhat depending on the distance bands of the lines of sight. Here I decided to focus on a single band of lines of sight up to 20km, although a number of features visible over longer distances should be mentioned. The hill and ridge on which Carmona is located as well as the area between this ridge and the Genil are highly visible both in the 20km and the unlimited bands. The low lying area of the *Lacus Ligustinus* and its shores are also visible from many sites in both bands. The banks of the Guadalquivir and Genil are not highly visible, with the exception of the area around Santiponce and a number of hills and plateau sides around the Genil. The *Via Augusta* generally passes through areas of low local visibility, with the exception of Carmona and possibly the shores of the *Lacus Ligustinus*, although much of its trail is visible from many sites over long distances. A few areas are only highly visible over longer distances, however. In particular the foothills of the Sierra Morena, the hills west of Santiponce and the foothills of the Sistema Sub-bético. Most interesting are the large areas around Osuna that are highly visible over long distances but less so over short distances. Both these long-distance patterns at borders of the study area and the high degree of visibility at the centre of the study area can be considered to reflect edge effects. The centre of the study area will naturally have a higher probability of being highly visible (in particular given the large number of sites) and the edges will only be highly visible over long distances.

The cumulative viewsheds nevertheless reveal some interesting information that will be useful when interpreting the exploratory network analysis results, in particular when we focus on the changes through time in areas of high and low visibility. There is a general decrease over time in the size of the area that can be seen from a large number of sites. This is to be expected given the general decrease in the number of sites occupied (Fig. 24; Table 7). Especially from the Middle Imperial period there is a strong decrease in the results for the area around Seville and Santiponce, as well as the area around Carmona and the middle Corbones valley (one of the most highly visible areas in the periods before). On the other hand, the area northwest of Osuna becomes more visually prominent in the Early Imperial period. One consistent feature is the visual prominence of hilltops, ridges and plateau sides.

These cumulative viewshed results have shown that the impact of the site density and the observer locations is significant and is expected to strongly influence the structure of the exploratory networks. However, they also reveal a gradual decrease of overall visibility in the study area, the increase of the visual prominence of one area, the visual prominence of sites located on elevations, and the consistently low degree of visibility of the roads and rivers.

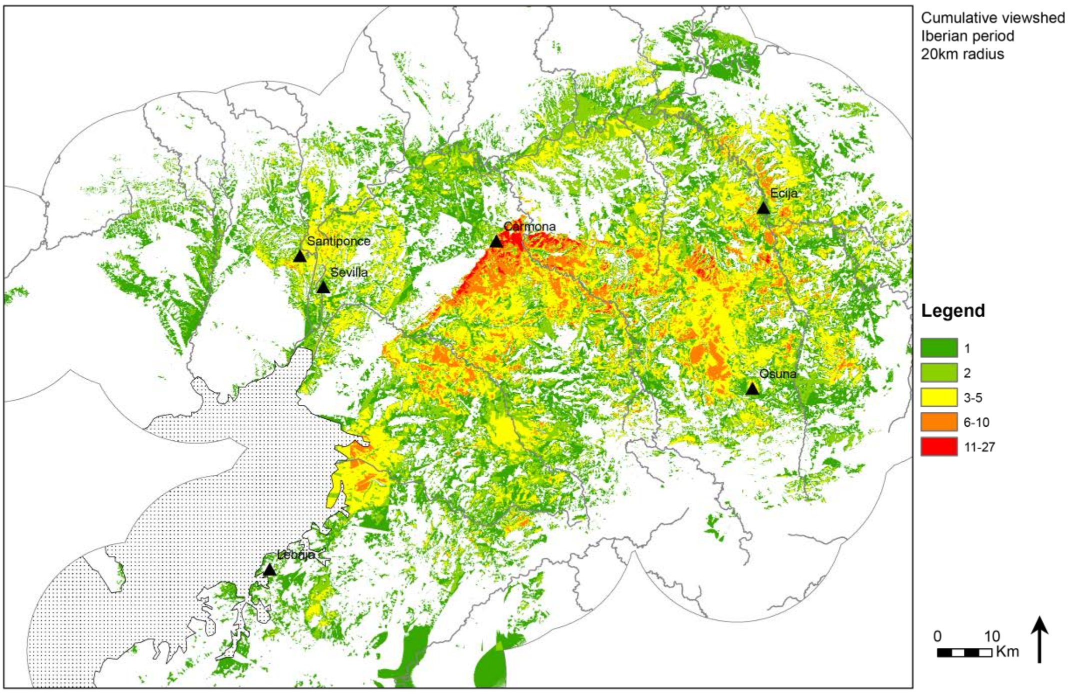


Fig. 61 Cumulative viewshed of sites dated to the Iberian period, lines of sight are limited to a 20km radius around observer locations. Colours represent the number of sites from which a particular part of the landscape is visible. The study area is cropped to a 20km area around Iberian sites.

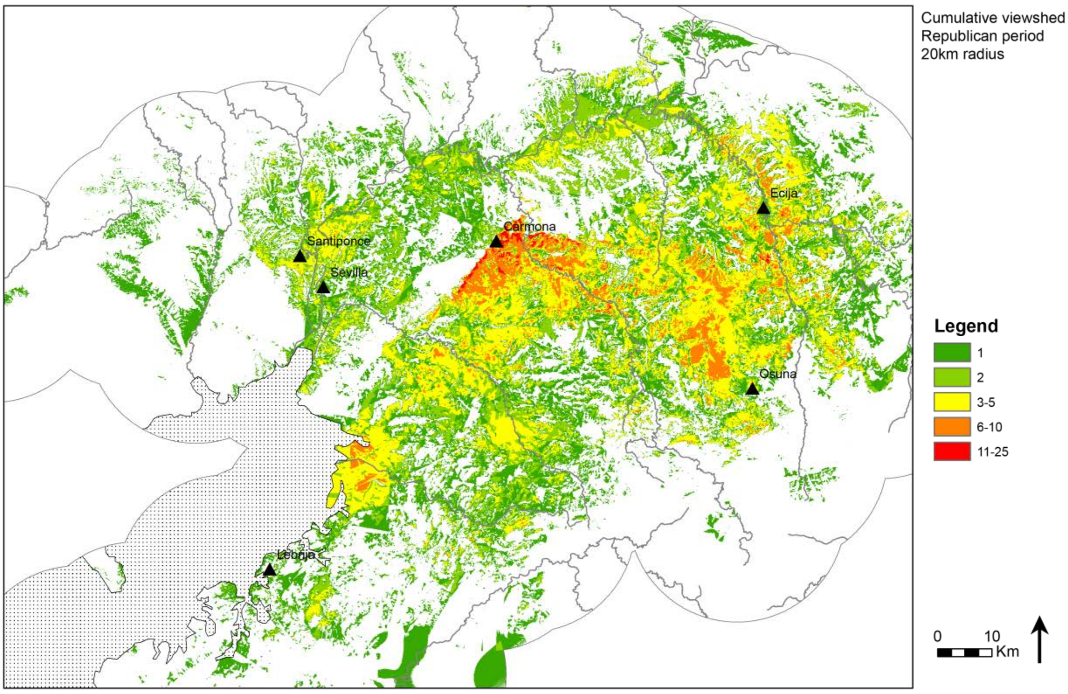


Fig. 62 Cumulative viewshed of sites dated to the Republican period, lines of sight are limited to a 20km radius around observer locations. Colours represent the number of sites from which a particular part of the landscape is visible. The study area is cropped to a 20km area around Republican sites.

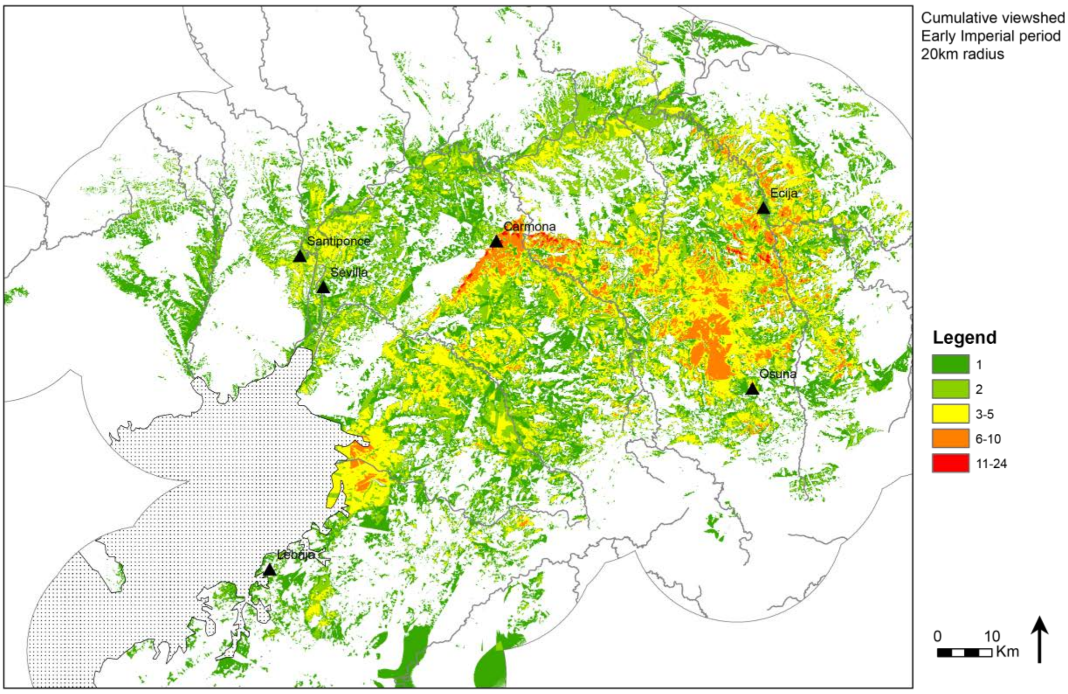


Fig. 63 Cumulative viewshed of sites dated to the Early Imperial period, lines of sight are limited to a 20km radius around observer locations. Colours represent the number of sites from which a particular part of the landscape is visible. The study area is cropped to a 20km area around Early Imperial sites.

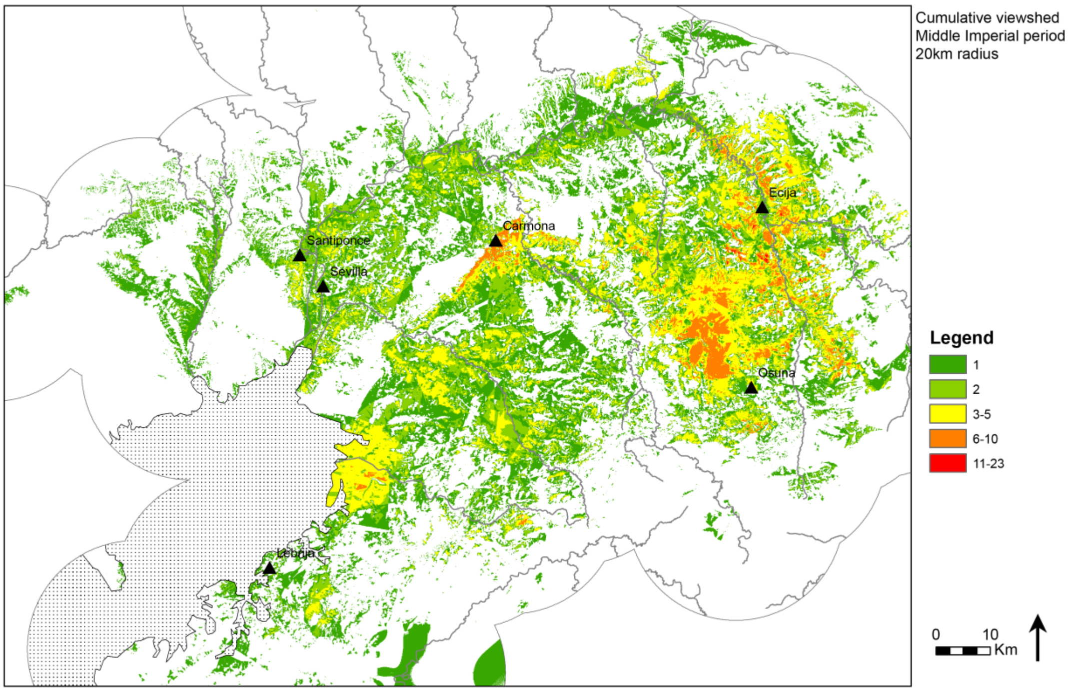


Fig. 64 Cumulative viewshed of sites dated to the Middle Imperial period, lines of sight are limited to a 20km radius around observer locations. Colours represent the number of sites from which a particular part of the landscape is visible. The study area is cropped to a 20km area around Middle Imperial sites.

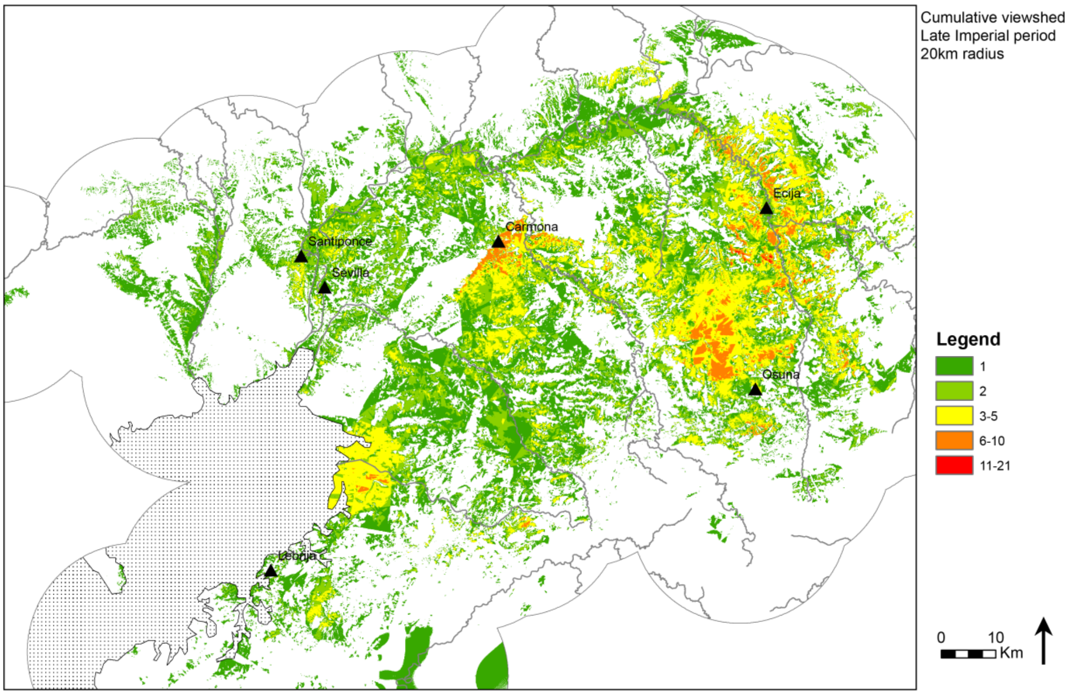


Fig. 65 Cumulative viewshed of sites dated to the Late Imperial period, lines of sight are limited to a 20km radius around observer locations. Colours represent the number of sites from which a particular part of the landscape is visible. The study area is cropped to a 20km area around Late Imperial sites.

11.3. Results of the node-based network measures

This appendix presents the results of the local exploratory network measures for each period's visibility network with arcs up to 20km and higher than 50% probability.

11.3.1. Iberian period network up to 20km

Site name	Project ID	Clustering Coefficient	Indegree	Outdegree
Alamillo	2	0	1	1
Alcaudete	6	0	0	1
Alcolea del Rio	7	0	1	1
Atalaya de la Moranilla	14	0	1	0
Cagancha	286	0	1	1
Cantillana	23	0	1	1
Castillo de Alhonoz	34	0.1	5	4
Cerro de las Cabezas (Sobarbina)	233	0	1	0
Cerro del Bollo	303	0.33333333	4	3
Cerro del Manzano	60	0	1	1
Cerro del Orégano	293	0	0	1
Cerro Gordo	63	0	1	1
Cerro Rubio	65	0	1	0
Cerros de San Ignacio	287	0	2	1
Chiclana (II)	68	0	0	1
Consuegra	10	0	0	1
Cortijo de los Cosmes	75	0	1	1
Cortijo del Cerro	80	0	1	1
Cortijos del Cerro (II)	84	0	0	1

El Chiste (V)	99	0	1	0
El Guijo	105	0	1	2
El Nuño	109	0	2	1
El Palomarejo	201	0	1	0
Friíllas	266	0	1	0
Garrotal	267	0	0	1
Huerta del Caño	133	0	1	1
La Alcuza	269	0.16666667	3	4
La Esclavitud	142	0.33333333	3	3
La Foronguilla	143	0	1	1
La Torre II	290	0.66666667	3	3
Las Aguzaderas	162	1	1	2
Las Albaidas (I)	163	0	0	1
Las Cabezas de San Juan	164	0	1	1
Las Mazmorras	166	0.66666667	3	3
Mesa del Almendro	93	0	1	0
Mochales	274	1	2	2
Montemolín	192	0	1	1
Morón de la Frontera	194	0	2	1
Palmilla (I)	200	0	1	1
Pancorvo	203	0.16666667	6	5
Pedro Cruzado (I)	206	0	1	1
Pedro Cruzado (II)	207	0	0	1
Peñaflor	208	0	0	1
Porcún I	284	0.5	4	4

Pozo del Carretero	285	0	2	0
San José	220	0	1	1
San Pedro (I)	223	0.2	3	5
San Pedro (II)	224	0.33333333	2	3
San Pedro (VII)	227	0	1	1
Tablada	235	0	4	2
Tejada la Nueva	419	0	1	0
Tejada La Vieja	316	0	0	2
Torre del Aguila	242	0	0	1
Vico	445	0.33333333	3	3

11.3.2. Republican period network up to 20km

Site name	Project ID	Clustering Coefficient	Indegree	Outdegree
Alamillo	2	0	1	1
Alcaudete	6	0	0	1
Alcolea del Rio	7	0	1	1
Atalaya de la Moranilla	14	0	2	1
Cagancha	286	0	1	1
Cantillana	23	0	1	1
Castillo de Alhonoz	34	0.16666667	4	3
Cerro de las Cabezas	46	0	0	1
Cerro del Bollo	303	0.33333333	4	3
Cerro del Orégano	293	0	0	1
Cerro Rubio	65	0	1	0

Cerros de San Ignacio	287	0	2	1
Chiclana (II)	68	0	0	1
Consuegra	10	0	0	1
Cortijo de los Cosmes	75	0	1	1
Cortijo del Cerro	80	0	1	1
Cortijos del Cerro (II)	84	0	0	1
El Chiste (V)	99	0	1	0
El Guijo	105	0	2	2
El Mocho	107	0	2	2
El Nuño	109	0	2	1
El Palomarejo	201	0	1	0
Friíllas	266	0	1	0
Garrotal	267	0	0	1
Huerta del Caño	133	0	1	1
La Alcuza	269	0.16666667	3	4
La Foronguilla	143	0	1	1
La Platera	154	0	1	1
La Torre II	290	0.66666667	3	3
Las Aguzaderas	162	1	1	2
Las Albaidas (I)	163	0	0	1
Las Cabezas de San Juan	164	0	1	1
Las Mazmorras	166	1	2	2
Mesa del Almendro	93	0	1	0
Mochales	274	0.33333333	3	3

Montemolín	192	0	1	1
Morón de la Frontera	194	0	1	0
Palmilla (I)	200	0	1	1
Pancorvo	203	0.25	4	3
Pedro Cruzado (II)	207	0	0	1
Peñaflor	208	0	0	1
Porcún I	284	0.5	4	4
Pozo del Carretero	285	0	2	0
San Pedro (I)	223	0.2	3	5
San Pedro (II)	224	0.33333333	2	3
San Pedro (VII)	227	0	1	1
Tablada	235	0	4	2
Torre del Aguila	242	0	0	1
Vico	445	0.33333333	3	3

11.3.3. Early Imperial period network up to 20km

Site name	**Project ID**	**Clustering Coefficient**	**Indegree**	**Outdegree**
Alamillo	2	0	1	1
Alcolea del Rio	7	0	1	1
Atalaya de la Moranilla	14	0	2	1
Cantillana	23	0	1	1
Cerro de las Cabezas	46	0	0	1
Cerro de los Ladrillos	50	0	0	1
Cerro del Bollo	303	0	3	2

Cerro del Orégano	293	0	0	1
Cerro Rubio	65	0	1	0
Chiclana (II)	68	0	0	1
Consuegra	10	0	0	1
Cortijo de Alcofría	73	0	2	3
Cortijo de los Cosmes	75	0	1	1
El Chiste (V)	99	0	1	0
El Grullo	104	0	1	0
El Guijo	105	0	1	1
El Mocho	107	0	2	2
El Nuño	109	0	2	1
El Palomarejo	201	0	1	0
Friíllas	266	0	1	0
Garrotal	267	0	0	1
Huerta del Caño	133	0	2	2
La Alcuza	269	0	2	3
La Foronguilla	143	0	1	1
La Platera	154	0	1	1
La Torre II	290	1	2	2
Las Aguzaderas	162	0	0	1
Las Cabezas de San Juan	164	0	1	1
Las Mazmorras	166	0	1	1
Los Abades	272	0	1	0
Mesa del Almendro	93	0	1	0
Mochales	274	0	3	3

Morón de la Frontera	194	0	2	0
Palmilla (I)	200	0	1	1
Peñaflor	208	0	0	1
Porcún I	284	0.66666667	3	3
Reinoso	215	0	1	1
San Pedro (I)	223	0.2	3	5
Vico	445	0.33333333	3	3

11.3.4. Middle Imperial period network up to 20km

Site name	Project ID	Clustering Coefficient	Indegree	Outdegree
Alamillo	2	0	1	1
Alcolea del Rio	7	0	1	1
Atalaya de la Moranilla	14	0	2	1
Cantillana	23	0	1	1
Cerro de las Cabezas	46	0	0	1
Cerro de los Ladrillos	50	0	0	1
Cerro del Bollo	303	0	3	2
Cerro del Orégano	293	0	0	1
Cerro Rubio	65	0	1	0
Chiclana (II)	68	0	0	1
Consuegra	10	0	0	1
Cortijo de los Cosmes	75	0	1	1
El Guijo	105	0	1	1
El Mocho	107	0	2	2

El Nuño	109	0	2	1
El Palomarejo	201	0	1	0
Friíllas	266	0	1	0
Garrotal	267	0	0	1
Huerta del Caño	133	0	1	1
La Alcuza	269	0	2	3
La Foronguilla	143	0	1	1
La Platera	154	0	1	1
Las Aguzaderas	162	0	0	1
Las Cabezas de San Juan	164	0	1	1
Las Mazmorras	166	0	1	1
Mochales	274	0	2	2
Morón de la Frontera	194	0	2	0
Palmilla (I)	200	0	1	1

11.3.5. Late Imperial period network up to 20km

Site name	Project ID	Clustering Coefficient	Indegree	Outdegree
Alamillo	2	0	1	1
Alcolea del Rio	7	0	1	1
Atalaya de la Moranilla	14	0	2	1
Cantillana	23	0	1	1
Cerro de los Ladrillos	50	0	0	1
Cerro del Bollo	303	0	3	2
Cerro del Orégano	293	0	0	1

Cerro Rubio	65	0	1	0
Cortijo del Cerro	80	0	1	1
El Mocho	107	0	2	2
El Nuño	109	0	1	1
Garrotal	267	0	0	1
Huerta del Caño	133	0	1	1
La Alcuza	269	0	2	3
La Foronguilla	143	0	1	1
Las Aguzaderas	162	0	0	1
Las Cabezas de San Juan	164	0	1	1
Las Mazmorras	166	0	1	1
Mochales	274	0	2	2
Morón de la Frontera	194	0	2	0
Palmilla (I)	200	0	1	1
Tablada	235	0	1	1

11.4. ERGM results

This appendix includes the results discussed in section 5.2 of the paper: the Bernoulli random graph models, the ERGMs without attributes, and the ERGMs with attributes.

11.4.1. Goodness of fit of Bernoulli random graph models

T-ratios express how good the model succeeds in reproducing the frequency of configurations in the observed network In order to be a good fit the t-ratio for these configurations should be lower than 2. Only the Bernoulli random graph models for the Middle and Late Imperial observed networks show a good fit for a small number of configurations.

Table 22. Iberian 20km network.

Effects	Observed	Mean	Standard deviation	T-ratio
arc	84			
reciprocity	32	0.126	0.35	91.163
2-in-star	78	21.658	4.54	12.411
2-out-star	73	21.64	4.706	10.914
3-in-star	59	3.523	2.717	20.417
3-out-star	57	3.658	3.001	17.773
path2	147	43.521	6.474	15.983
T1	6	0	0	
T2	41	0	0	
T3	46	0	0	
T4	23	0	0	
T5	23	0	0	
T6	53	0	0	
T7	129	0.113	0.459	280.959
T8	122	0.121	0.474	257.182

T9(030T)	52	0.148	0.405	127.938
T10(030C)	17	0.027	0.162	104.665
Sink	8	38.704	3.688	-8.325
Source	11	38.798	3.7	-7.512
Isolates	104	54.685	3.831	12.874
AinS(2.00)	54.563	19.986	3.627	9.534
AoutS(2.00)	50.563	19.918	3.765	8.14
Ain1out-star(2.00)	86.063	38.5	5.252	9.057
1inAout-star(2.00)	87.125	38.425	5.302	9.185
AinAout-star(2.00)	48.875	33.966	4.336	3.439
AT-T(2.00)	43.375	0.148	0.405	106.657
AT-C(2.00)	42.875	0.081	0.486	87.964
AT-D(2.00)	44	0.148	0.405	108.199
AT-U(2.00)	42.5	0.148	0.405	104.498
AT-TD(2.00)	43.688	0.148	0.405	107.428
AT-TU(2.00)	42.938	0.148	0.405	105.577
AT-DU(2.00)	43.25	0.148	0.405	106.348
AT-TDU(2.00)	43.292	0.148	0.405	106.451
A2P-T(2.00)	133.125	43.505	6.466	13.859
A2P-D(2.00)	65.25	21.633	4.706	9.268
A2P-U(2.00)	71.125	21.651	4.537	10.904
A2P-TD(2.00)	99.188	32.569	3.97	16.781
A2P-TU(2.00)	102.125	32.578	3.891	17.874
A2P-DU(2.00)	68.188	21.642	3.308	14.069
A2P-TDU(2.00)	89.833	28.929	3.034	20.076

Std Dev in-degree dist	1.113	0.723	0.039	9.881
Skew in-degree dist	2.614	1.288	0.247	5.367
Std Dev out-degree dist	1.084	0.723	0.041	8.816
Skew out-degree dist	2.698	1.297	0.264	5.309
CorrCoef in-out-degree dists	0.869	-0.007	0.078	11.3
Global Clustering Cto	0.356	0.003	0.01	36.937
Global Clustering Cti	0.333	0.003	0.009	35.94
Global Clustering Ctm	0.354	0.003	0.009	37.049
Global Clustering Ccm	0.347	0.002	0.011	32.655
Global Clustering AKC-T	0.326	0.003	0.009	34.08
Global Clustering AKC-D	0.337	0.003	0.01	34.942
Global Clustering AKC-U	0.299	0.003	0.009	32.163
Global Clustering AKC-C	0.322	0.002	0.011	30.302

Table 23. Republican 20km network.

Effects	Observed	Mean	Standard deviation	T-ratio
arc	74			
reciprocity	27	0.133	0.371	72.468
2-in-star	61	18.469	4.39	9.687
2-out-star	57	18.07	4.028	9.664
3-in-star	33	3.045	2.883	10.39
3-out-star	41	2.771	2.345	16.301
path2	117	36.495	5.754	13.991
T1	5	0	0	

T2	35	0	0	
T3	40	0	0	
T4	20	0	0	
T5	20	0	0	
T6	38	0	0	
T7	97	0.144	0.531	182.538
T8	92	0.129	0.503	182.79
T9(030T)	46	0.112	0.319	143.994
T10(030C)	15	0.039	0.204	73.423
Sink	6	35.192	3.515	-8.304
Source	11	35.396	3.443	-7.087
Isolates	97	52.375	3.673	12.148
AinS(2.00)	46.625	17.038	3.475	8.514
AoutS(2.00)	40.938	16.754	3.274	7.387
Ain1out-star(2.00)	76.875	32.441	4.742	9.37
1inAout-star(2.00)	76.938	32.437	4.737	9.395
AinAout-star(2.00)	48.438	28.827	3.916	5.008
AT-T(2.00)	38.375	0.112	0.319	120.067
AT-C(2.00)	37.875	0.117	0.611	61.767
AT-D(2.00)	39	0.112	0.319	122.028
AT-U(2.00)	37.5	0.112	0.319	117.321
AT-TD(2.00)	38.688	0.112	0.319	121.047
AT-TU(2.00)	37.938	0.112	0.319	118.694
AT-DU(2.00)	38.25	0.112	0.319	119.675
AT-TDU(2.00)	38.292	0.112	0.319	119.805

A2P-T(2.00)	104.125	36.483	5.752	11.76
A2P-D(2.00)	49.75	18.061	4.022	7.879
A2P-U(2.00)	54.625	18.459	4.388	8.242
A2P-TD(2.00)	76.938	27.272	3.486	14.246
A2P-TU(2.00)	79.375	27.471	3.571	14.533
A2P-DU(2.00)	52.188	18.26	2.973	11.41
A2P-TDU(2.00)	69.5	24.334	2.718	16.617
Std Dev in-degree dist	1.045	0.71	0.042	7.97
Skew in-degree dist	2.266	1.328	0.279	3.363
Std Dev out-degree dist	1.019	0.707	0.039	8.005
Skew out-degree dist	2.683	1.304	0.256	5.375
CorrCoef in-out-degree dists	0.858	-0.01	0.079	10.946
Global Clustering Cto	0.404	0.003	0.009	44.911
Global Clustering Cti	0.377	0.003	0.009	40.558
Global Clustering Ctm	0.393	0.003	0.009	43.653
Global Clustering Ccm	0.385	0.003	0.016	23.972
Global Clustering AKC-T	0.369	0.003	0.009	40.873
Global Clustering AKC-D	0.392	0.003	0.009	43.602
Global Clustering AKC-U	0.343	0.003	0.009	36.835
Global Clustering AKC-C	0.364	0.003	0.016	22.661

Table 24 Early Imperial 20km network.

Effects	Observed	Mean	Standard deviation	T-ratio
arc	51			

reciprocity	18	0.057	0.24	74.627
2-in-star	25	8.573	2.996	5.484
2-out-star	31	8.307	2.795	8.119
3-in-star	6	1.002	1.535	3.256
3-out-star	16	0.877	1.273	11.884
path2	57	16.652	4.004	10.078
T1	2	0	0	
T2	12	0	0	
T3	12	0	0	
T4	6	0	0	
T5	6	0	0	
T6	20	0.001	0.032	632.424
T7	44	0.038	0.273	160.923
T8	50	0.041	0.271	184.412
T9(030T)	12	0.034	0.181	65.994
T10(030C)	4	0.009	0.094	42.238
Sink	8	31.129	3.088	-7.491
Source	8	31.302	2.932	-7.947
Isolates	110	75.367	3.285	10.542
AinS(2.00)	22	8.097	2.548	5.455
AoutS(2.00)	24.125	7.885	2.419	6.713
Ain1out-star(2.00)	45	15.363	3.449	8.593
1inAout-star(2.00)	42.125	15.386	3.551	7.53
AinAout-star(2.00)	31.875	14.194	3.063	5.772
AT-T(2.00)	11	0.034	0.181	60.479

AT-C(2.00)	11	0.027	0.283	38.711
AT-D(2.00)	11	0.034	0.181	60.479
AT-U(2.00)	11	0.034	0.181	60.479
AT-TD(2.00)	11	0.034	0.181	60.479
AT-TU(2.00)	11	0.034	0.181	60.479
AT-DU(2.00)	11	0.034	0.181	60.479
AT-TDU(2.00)	11	0.034	0.181	60.479
A2P-T(2.00)	52	16.648	4.003	8.832
A2P-D(2.00)	28.5	8.306	2.795	7.224
A2P-U(2.00)	22.5	8.572	2.995	4.651
A2P-TD(2.00)	40.25	12.477	2.378	11.677
A2P-TU(2.00)	37.25	12.61	2.504	9.841
A2P-DU(2.00)	25.5	8.439	2.032	8.398
A2P-TDU(2.00)	34.333	11.175	1.868	12.396
Std Dev in-degree dist	0.749	0.583	0.034	4.885
Skew in-degree dist	2.31	1.639	0.328	2.045
Std Dev out-degree dist	0.801	0.58	0.032	6.918
Skew out-degree dist	2.975	1.614	0.296	4.592
CorrCoef in-out-degree dists	0.84	-0.011	0.08	10.664
Global Clustering Cto	0.194	0.002	0.012	16.618
Global Clustering Cti	0.24	0.002	0.011	20.988
Global Clustering Ctm	0.211	0.002	0.013	15.439
Global Clustering Ccm	0.211	0.002	0.017	12.526
Global Clustering AKC-T	0.212	0.002	0.013	15.514
Global Clustering AKC-D	0.193	0.002	0.012	16.569

Global Clustering AKC-U	0.244	0.002	0.011	21.379
Global Clustering AKC-C	0.212	0.002	0.017	12.587

Table 25. Middle Imperial 20km network.

Effects	Observed	Mean	Standard deviation	T-ratio
arc	29			
reciprocity	9	0.027	0.162	55.333
2-in-star	9	3.101	1.679	3.513
2-out-star	6	3.208	1.772	1.576
3-in-star	1	0.204	0.556	1.433
3-out-star	1	0.22	0.518	1.507
path2	17	6.518	2.484	4.22
T1	0	0	0	
T2	0	0	0	
T3	0	0	0	
T4	0	0	0	
T5	0	0	0	
T6	4	0	0	
T7	12	0.013	0.151	79.292
T8	10	0.015	0.157	63.405
T9(030T)	0	0.013	0.113	-0.115
T10(030C)	0	0.004	0.063	-0.063
Sink	4	20.711	2.238	-7.466
Source	7	20.625	2.22	-6.138

Isolates	97	78.282	2.512	7.452
AinS(2.00)	8.5	3.002	1.527	3.601
AoutS(2.00)	5.5	3.099	1.624	1.479
Ain1out-star(2.00)	15.5	6.199	2.272	4.093
1inAout-star(2.00)	15.5	6.205	2.296	4.047
AinAout-star(2.00)	14	5.902	2.1	3.857
AT-T(2.00)	0	0.013	0.113	-0.115
AT-C(2.00)	0	0.012	0.189	-0.063
AT-D(2.00)	0	0.013	0.113	-0.115
AT-U(2.00)	0	0.013	0.113	-0.115
AT-TD(2.00)	0	0.013	0.113	-0.115
AT-TU(2.00)	0	0.013	0.113	-0.115
AT-DU(2.00)	0	0.013	0.113	-0.115
AT-TDU(2.00)	0	0.013	0.113	-0.115
A2P-T(2.00)	16	6.518	2.484	3.817
A2P-D(2.00)	5.5	3.208	1.772	1.293
A2P-U(2.00)	8.5	3.101	1.679	3.215
A2P-TD(2.00)	10.75	4.863	1.503	3.917
A2P-TU(2.00)	12.25	4.809	1.527	4.871
A2P-DU(2.00)	7	3.155	1.222	3.146
A2P-TDU(2.00)	10	4.276	1.165	4.914
Std Dev in-degree dist	0.57	0.478	0.028	3.289
Skew in-degree dist	2.568	1.928	0.344	1.861
Std Dev out-degree dist	0.526	0.48	0.029	1.559
Skew out-degree dist	2.51	1.945	0.346	1.634

CorrCoef in-out-degree dists	0.755	-0.006	0.087	8.712
Global Clustering Cto	0	0.002	0.019	-0.107
Global Clustering Cti	0	0.002	0.026	-0.086
Global Clustering Ctm	0	0.002	0.02	-0.111
Global Clustering Ccm	0	0.001	0.017	-0.062
Global Clustering AKC-T	0	0.002	0.02	-0.111
Global Clustering AKC-D	0	0.002	0.019	-0.107
Global Clustering AKC-U	0	0.002	0.026	-0.086
Global Clustering AKC-C	0	0.001	0.017	-0.062

Table 26. Late Imperial 20km network.

Effects	Observed	Mean	Standard deviation	T-ratio
arc	25			
reciprocity	10	0.026	0.159	62.645
2-in-star	8	2.49	1.585	3.476
2-out-star	6	2.632	1.587	2.123
3-in-star	1	0.158	0.542	1.555
3-out-star	1	0.173	0.519	1.594
path2	13	5.053	2.238	3.551
T1	0	0	0	
T2	0	0	0	
T3	0	0	0	
T4	0	0	0	
T5	0	0	0	

T6	4	0	0	
T7	11	0.008	0.1	110.218
T8	10	0.01	0.118	84.692
T9(030T)	0	0.015	0.122	-0.123
T10(030C)	0	0.004	0.063	-0.063
Sink	2	18.474	1.946	-8.465
Source	4	18.346	2.009	-7.141
Isolates	93	73.994	2.185	8.7
AinS(2.00)	7.5	2.413	1.444	3.523
AoutS(2.00)	5.5	2.548	1.449	2.037
AinS(2.00)	7.5	2.413	1.444	3.523
AoutS(2.00)	5.5	2.548	1.449	2.037
Ain1out-star(2.00)	11.5	4.822	2.05	3.257
1inAout-star(2.00)	11.5	4.794	2.043	3.282
AinAout-star(2.00)	10	4.576	1.878	2.888
AT-T(2.00)	0	0.015	0.122	-0.123
AT-C(2.00)	0	0.012	0.189	-0.063
AT-D(2.00)	0	0.015	0.122	-0.123
AT-U(2.00)	0	0.015	0.122	-0.123
AT-TD(2.00)	0	0.015	0.122	-0.123
AT-TU(2.00)	0	0.015	0.122	-0.123
AT-DU(2.00)	0	0.015	0.122	-0.123
AT-TDU(2.00)	0	0.015	0.122	-0.123
A2P-T(2.00)	12	5.053	2.238	3.104
A2P-D(2.00)	5.5	2.632	1.586	1.809

A2P-U(2.00)	7.5	2.49	1.585	3.161
A2P-TD(2.00)	8.75	3.842	1.371	3.579
A2P-TU(2.00)	9.75	3.771	1.33	4.495
A2P-DU(2.00)	6.5	2.561	1.099	3.586
A2P-TDU(2.00)	8.333	3.391	1.021	4.838
Std Dev in-degree dist	0.559	0.463	0.029	3.26
Skew in-degree dist	2.726	1.981	0.37	2.01
Std Dev out-degree dist	0.526	0.466	0.029	2.058
Skew out-degree dist	2.7	2.012	0.374	1.838
CorrCoef in-out-degree dists	0.816	-0.013	0.091	9.068
Global Clustering Cto	0	0.004	0.036	-0.104
Global Clustering Cti	0	0.002	0.022	-0.112
Global Clustering Ctm	0	0.003	0.027	-0.108
Global Clustering Ccm	0	0.001	0.02	-0.063
Global Clustering AKC-T	0	0.003	0.027	-0.108
Global Clustering AKC-D	0	0.004	0.036	-0.104
Global Clustering AKC-U	0	0.002	0.022	-0.112
Global Clustering AKC-C	0	0.001	0.02	-0.063

11.4.2. ERGMs without attributes

Models of 20km networks

Significant configurations are indicated by *. Configurations are significant if the estimate is more than two times the standard error.

Table 27. Iberian 20km social circuit model.

Effects	Estimates	Standard error	T-ratio	
reciprocity	8.00	0.79	-0.07	*
path2	-0.52	0.17	-0.07	*
030T	0.40	0.05	-0.06	*
sink	-2.29	1.33	0.05	
source	-1.28	1.39	0.03	
isolates	-3.23	1.53	-0.10	*
AinS	2.35	0.92	-0.08	*
AoutS	2.61	0.96	-0.08	*

Table 28. Republican 20km circuit model.

Effects	Estimates	Standard error	T-ratio	
reciprocity	6.81	0.79	-0.05	*
2-in-star	-3.66	0.92	0.03	*
2-out-star	0.15	0.52	-0.06	
path2	0.12	0.35	-0.02	
030T	0.54	0.08	0.02	*
sink	-1.16	1.81	0.03	
source	-5.87	1.66	0.10	*

isolates	-6.20	1.97	-0.06	*
AinS(2.00)	10.02	2.08	0.01	*
AoutS(2.00)	0.21	1.88	-0.06	

Table 29. Early Imperial 20km circuit model.

Effects	Estimates	Standard deviation	T-ratio	
reciprocity	8.69	1.21	-0.01	*
2-in-star	-4.63	2.09	0.02	*
2-out-star	-0.05	0.83	0.04	
path2	0.90	0.66	0.03	
030T	0.58	0.16	0.09	*
isolates	-0.74	0.89	0.00	
AinS(2.00)	5.75	2.44	0.02	*
AoutS(2.00)	-0.03	1.28	0.02	
AinAout-star(2.00)	0.26	0.78	0.00	

Table 30. Middle Imperial 20km circuit model.

Effects	Estimates	Standard error	T-ratio	
reciprocity	6.98	1.41	-0.08	*
2-in-star	-3.30	2.01	0.02	
2-out-star	-0.25	2.19	0.09	
path2	0.86	0.75	0.05	
isolates	0.87	0.98	-0.07	
AinS(2.00)	3.92	2.30	0.02	
AoutS(2.00)	-1.07	2.52	0.02	

Table 31. Late Imperial 20km circuit model.

Effects	Estimates	Standard error	T-ration	
reciprocity	8.59	2.08	0.03	*
path2	-0.74	0.82	0.03	
111U	0.08	0.89	0.03	
isolates	1.07	1.35	0.01	
AinS(2.00)	1.85	1.04	-0.01	
AoutS(2.00)	0.56	1.61	0.01	
AinAout-star(2.00)	-0.33	0.79	0.04	

Goodness of fit 20km network models

Configurations included in the model are in bold. T-ratios express how good the model succeeds in reproducing the frequency of configurations in the observed network In order to be a good fit the t-ratio for these configurations should be lower than 0.1, t-ratios for all other configurations should be between 0.1 and 2.

Table 32. Goodness of fit Iberian 20km social circuit model. 100 million simulations, 1000 samples.

Effects	Observed	Mean	Standard deviation	T-ratio
arc	84	84	0	
reciprocity	**32**	**31.966**	**2.612**	**0.013**
2-in-star	78	79.134	21.836	-0.052
2-out-star	73	73.072	22.799	-0.003
3-in-star	59	66.401	40.74	-0.182
3-out-star	57	61.695	41.224	-0.114
path2	**147**	**143.565**	**45.57**	**0.075**
T1	6	7.57	8.386	-0.187
T2	41	46.554	50.843	-0.109
T3	46	47.717	51.418	-0.033
T4	23	23.947	25.677	-0.037
T5	23	23.935	25.7	-0.036
T6	53	55.73	24.342	-0.112
T7	129	128.75	46.8	0.005
T8	122	124.645	47.623	-0.056
T9(030T)	**52**	**49.275**	**51.95**	**0.052**
T10(030C)	17	16.303	17.347	0.04
Sink	**8**	**7.989**	**2.698**	**0.004**
Source	**11**	**11.108**	**3.324**	**-0.032**

Isolates	**104**	**103.666**	**5.94**	**0.056**
AinS(2.00)	**54.563**	**54.011**	**9.253**	**0.06**
AoutS(2.00)	**50.563**	**49.777**	**9.973**	**0.079**
AinS(2.00)	54.563	54.011	9.253	0.06
AoutS(2.00)	50.563	49.777	9.973	0.079
Ain1out-star(2.00)	86.063	80.832	12.799	0.409
1inAout-star(2.00)	87.125	82.022	12.624	0.404
AinAout-star(2.00)	48.875	46.65	4.647	0.479
AT-T(2.00)	43.375	30.296	22.206	0.589
AT-C(2.00)	42.875	29.966	22.308	0.579
AT-D(2.00)	44	30.295	22.215	0.617
AT-U(2.00)	42.5	30.26	22.203	0.551
AT-TD(2.00)	43.688	30.296	22.21	0.603
AT-TU(2.00)	42.938	30.278	22.203	0.57
AT-DU(2.00)	43.25	30.278	22.206	0.584
AT-TDU(2.00)	43.292	30.284	22.206	0.586
A2P-T(2.00)	133.125	121.068	21.351	0.565
A2P-D(2.00)	65.25	61.76	11.126	0.314
A2P-U(2.00)	71.125	67.829	11.08	0.297
A2P-TD(2.00)	99.188	91.414	15.75	0.494
A2P-TU(2.00)	102.125	94.449	15.539	0.494
A2P-DU(2.00)	68.188	64.795	10.092	0.336
A2P-TDU(2.00)	89.833	83.553	13.594	0.462
Std Dev in-degree dist	1.113	1.113	0.12	0
Skew in-degree dist	2.614	2.617	0.401	-0.009

Std Dev out-degree dist	1.084	1.077	0.128	0.056
Skew out-degree dist	2.698	2.64	0.419	0.14
CorrCoef in-out-degree dists	0.869	0.836	0.079	0.414
Global Clustering Cto	0.356	0.281	0.217	0.345
Global Clustering Cti	0.333	0.264	0.211	0.33
Global Clustering Ctm	0.354	0.284	0.216	0.321
Global Clustering Ccm	0.347	0.281	0.217	0.302
Global Clustering AKC-T	0.326	0.239	0.164	0.529
Global Clustering AKC-D	0.337	0.236	0.165	0.613
Global Clustering AKC-U	0.299	0.218	0.157	0.515
Global Clustering AKC-C	0.322	0.236	0.165	0.521

Table 33. Goodness of fit Republican 20km circuit model. 100milion simulations, 1000 samples.

Effects	Observed	Mean	Standard deviation	T-ratio
arc	74	74	0	
reciprocity	**27**	**26.814**	**2.233**	**0.083**
2-in-star	**61**	**60.689**	**11.074**	**0.028**
2-out-star	**57**	**56.406**	**16.189**	**0.037**
3-in-star	33	32.506	13.578	0.036
3-out-star	41	40.243	25.207	0.03
path2	**117**	**116.214**	**25.799**	**0.03**
T1	5	6.356	5.095	-0.266
T2	35	39.46	30.682	-0.145
T3	40	40.861	30.868	-0.028
T4	20	20.478	15.408	-0.031

T5	20	20.521	15.478	-0.034
T6	38	39.314	13.383	-0.098
T7	97	96.146	24.283	0.035
T8	92	93.843	28.798	-0.064
T9(030T)	**46**	**42.642**	**31.157**	**0.108**
T10(030C)	15	14.116	10.377	0.085
Sink	**6**	**6.07**	**2.33**	**-0.03**
Source	**11**	**11.215**	**3.116**	**-0.069**
Isolates	**97**	**96.717**	**5.176**	**0.055**
AinS(2.00)	**46.625**	**46.487**	**6.232**	**0.022**
AoutS(2.00)	**40.938**	**40.561**	**8.204**	**0.046**
AinS(2.00)	46.625	46.487	6.232	0.022
AoutS(2.00)	40.938	40.561	8.204	0.046
Ain1out-star(2.00)	76.875	74.496	10.445	0.228
1inAout-star(2.00)	76.938	74.506	9.182	0.265
AinAout-star(2.00)	**48.438**	**46.64**	**3.923**	**0.458**
AT-T(2.00)	38.375	30.125	16.575	0.498
AT-C(2.00)	37.875	29.853	16.565	0.484
AT-D(2.00)	39	30.233	16.619	0.528
AT-U(2.00)	37.5	29.915	16.455	0.461
AT-TD(2.00)	38.688	30.179	16.596	0.513
AT-TU(2.00)	37.938	30.02	16.512	0.479
AT-DU(2.00)	38.25	30.074	16.531	0.495
AT-TDU(2.00)	38.292	30.091	16.545	0.496
A2P-T(2.00)	104.125	101.605	15.021	0.168

A2P-D(2.00)	49.75	49.034	10.643	0.067
A2P-U(2.00)	54.625	53.4	6.594	0.186
A2P-TD(2.00)	76.938	75.32	12.548	0.129
A2P-TU(2.00)	79.375	77.503	10.304	0.182
A2P-DU(2.00)	52.188	51.217	7.606	0.128
A2P-TDU(2.00)	69.5	68.013	10.02	0.148
Std Dev in-degree dist	1.045	1.041	0.073	0.063
Skew in-degree dist	2.266	2.203	0.216	0.289
Std Dev out-degree dist	1.019	1.009	0.109	0.091
Skew out-degree dist	2.683	2.513	0.436	0.391
CorrCoef in-out-degree dists	0.858	0.853	0.048	0.121
Global Clustering Cto	0.404	0.344	0.182	0.328
Global Clustering Cti	0.377	0.326	0.191	0.269
Global Clustering Ctm	0.393	0.335	0.187	0.311
Global Clustering Ccm	0.385	0.332	0.187	0.278
Global Clustering AKC-T	0.369	0.289	0.148	0.537
Global Clustering AKC-D	0.392	0.299	0.146	0.638
Global Clustering AKC-U	0.343	0.278	0.151	0.431
Global Clustering AKC-C	0.364	0.286	0.148	0.522

Table 34. Goodness of fit Early Imperial 20km circuit model. 50milion simulations, 1000 samples.

Effects	Observed	Mean	Standard deviation	T-ratio
arc	51	51	0	
reciprocity	**18**	**18.039**	**1.705**	**-0.023**
2-in-star	**25**	**25.429**	**4.867**	**-0.088**

2-out-star	**31**	**31.467**	**9.448**	**-0.049**
3-in-star	6	6.595	3.65	-0.163
3-out-star	16	17.335	14.258	-0.094
path2	**57**	**57.907**	**12.689**	**-0.071**
T1	2	1.868	1.506	0.088
T2	12	11.733	9.093	0.029
T3	12	12.295	9.213	-0.032
T4	6	6.142	4.609	-0.031
T5	6	6.176	4.619	-0.038
T6	20	18.741	5.283	0.238
T7	44	43.449	9.841	0.056
T8	50	49.111	13.363	0.067
T9(030T)	**12**	**12.942**	**9.422**	**-0.1**
T10(030C)	4	4.302	3.132	-0.096
Sink	8	8.514	2.543	-0.202
Source	8	8.278	2.574	-0.108
Isolates	**110**	**110.211**	**4.075**	**-0.052**
AinS(2.00)	**22**	**22.297**	**3.48**	**-0.085**
AoutS(2.00)	**24.125**	**24.438**	**5.128**	**-0.061**
AinS(2.00)	22	22.297	3.48	-0.085
AoutS(2.00)	24.125	24.438	5.128	-0.061
Ain1out-star(2.00)	45	44.57	6.665	0.065
1inAout-star(2.00)	42.125	41.899	5.534	0.041
AinAout-star(2.00)	**31.875**	**32.093**	**3.229**	**-0.067**
AT-T(2.00)	11	11.829	7.661	-0.108

AT-C(2.00)	11	11.795	7.637	-0.104
AT-D(2.00)	11	11.889	7.717	-0.115
AT-U(2.00)	11	11.748	7.589	-0.099
AT-TD(2.00)	11	11.859	7.688	-0.112
AT-TU(2.00)	11	11.788	7.624	-0.103
AT-DU(2.00)	11	11.818	7.65	-0.107
AT-TDU(2.00)	11	11.822	7.653	-0.107
A2P-T(2.00)	52	56.175	11.169	-0.374
A2P-D(2.00)	28.5	30.593	8.691	-0.241
A2P-U(2.00)	22.5	24.567	4.213	-0.491
A2P-TD(2.00)	40.25	43.384	9.783	-0.32
A2P-TU(2.00)	37.25	40.371	7.447	-0.419
A2P-DU(2.00)	25.5	27.58	5.952	-0.349
A2P-TDU(2.00)	34.333	37.112	7.662	-0.363
Std Dev in-degree dist	0.749	0.752	0.043	-0.06
Skew in-degree dist	2.31	2.312	0.237	-0.011
Std Dev out-degree dist	0.801	0.801	0.077	-0.003
Skew out-degree dist	2.975	2.875	0.559	0.178
CorrCoef in-out-degree dists	0.84	0.841	0.043	-0.029
Global Clustering Cto	0.194	0.197	0.123	-0.032
Global Clustering Cti	0.24	0.241	0.151	-0.008
Global Clustering Ctm	0.211	0.211	0.131	-0.007
Global Clustering Ccm	0.211	0.211	0.131	-0.003
Global Clustering AKC-T	0.212	0.204	0.12	0.066
Global Clustering AKC-D	0.193	0.191	0.113	0.021

Global Clustering AKC-U	0.244	0.232	0.14	0.087
Global Clustering AKC-C	0.212	0.203	0.12	0.07

Table 35. Goodness of fit Middle Imperial 20km circuit model. 50milion simulations, 1000 samples.

Effects	Observed	Mean	Standard deviation	T-ratio
arc	29	29	0	
reciprocity	**9**	**8.878**	**1.474**	**0.083**
2-in-star	**9**	**9.028**	**2.406**	**-0.012**
2-out-star	**6**	**5.686**	**2.378**	**0.132**
3-in-star	1	0.975	1.071	0.023
3-out-star	1	0.701	1.144	0.261
path2	**17**	**16.756**	**4.327**	**0.056**
T1	0	0.012	0.109	-0.11
T2	0	0.09	0.667	-0.135
T3	0	0.124	0.716	-0.173
T4	0	0.055	0.355	-0.155
T5	0	0.057	0.358	-0.159
T6	4	3.038	1.828	0.526
T7	12	11.185	3.789	0.215
T8	10	8.383	3.894	0.415
T9(030T)	0	0.152	0.792	-0.192
T10(030C)	0	0.065	0.295	-0.221
Sink	4	4.211	1.88	-0.112
Source	7	7.247	2.153	-0.115
Isolates	**97**	**96.83**	**2.712**	**0.063**

AinS(2.00)	**8.5**	**8.547**	**2.019**	**-0.023**
AoutS(2.00)	**5.5**	**5.35**	**2.039**	**0.074**
AinS(2.00)	8.5	8.547	2.019	-0.023
AoutS(2.00)	5.5	5.35	2.039	0.074
Ain1out-star(2.00)	15.5	14.905	3.269	0.182
1inAout-star(2.00)	15.5	15.161	3.349	0.101
AinAout-star(2.00)	14	13.456	2.624	0.207
AT-T(2.00)	0	0.152	0.792	-0.192
AT-C(2.00)	0	0.195	0.884	-0.221
AT-D(2.00)	0	0.152	0.792	-0.192
AT-U(2.00)	0	0.152	0.792	-0.192
AT-TD(2.00)	0	0.152	0.792	-0.192
AT-TU(2.00)	0	0.152	0.792	-0.192
AT-DU(2.00)	0	0.152	0.792	-0.192
AT-TDU(2.00)	0	0.152	0.792	-0.192
A2P-T(2.00)	16	16.741	4.316	-0.172
A2P-D(2.00)	5.5	5.678	2.373	-0.075
A2P-U(2.00)	8.5	9.02	2.401	-0.216
A2P-TD(2.00)	10.75	11.209	3.162	-0.145
A2P-TU(2.00)	12.25	12.88	3.182	-0.198
A2P-DU(2.00)	7	7.348	2.05	-0.17
A2P-TDU(2.00)	10	10.479	2.734	-0.175
Std Dev in-degree dist	0.57	0.569	0.034	0.018
Skew in-degree dist	2.568	2.507	0.268	0.226
Std Dev out-degree dist	0.526	0.52	0.037	0.168

Skew out-degree dist	2.51	2.304	0.405	0.509
CorrCoef in-out-degree dists	0.755	0.746	0.079	0.115
Global Clustering Cto	0	0.011	0.059	-0.186
Global Clustering Cti	0	0.008	0.044	-0.181
Global Clustering Ctm	0	0.008	0.046	-0.183
Global Clustering Ccm	0	0.011	0.05	-0.211
Global Clustering AKC-T	0	0.008	0.046	-0.183
Global Clustering AKC-D	0	0.011	0.059	-0.186
Global Clustering AKC-U	0	0.008	0.044	-0.181
Global Clustering AKC-C	0	0.011	0.05	-0.211

Table 36. Goodness of fit Late Imperial 20km circuit model. 50milion simulations, 1000 samples.

Effects	Observed	Mean	Standard deviation	T-ratio
arc	25	25	0	
reciprocity	**10**	**10.038**	**1.145**	**-0.033**
2-in-star	8	8.362	3.109	-0.116
2-out-star	6	5.902	2.294	0.043
3-in-star	1	1.998	2.321	-0.43
3-out-star	1	0.791	1.095	0.191
path2	**13**	**12.894**	**4.563**	**0.023**
T1	0	0.009	0.094	-0.095
T2	0	0.065	0.575	-0.113
T3	0	0.077	0.603	-0.128
T4	0	0.042	0.307	-0.137
T5	0	0.04	0.304	-0.132

T6	4	4.442	2.221	-0.199
T7	11	11.679	4.601	-0.148
T8	**10**	**9.975**	**4.367**	**0.006**
T9(030T)	0	0.101	0.663	-0.152
T10(030C)	0	0.03	0.217	-0.138
Sink	2	2.116	1.456	-0.08
Source	4	3.615	1.911	0.202
Isolates	**93**	**93.044**	**2.235**	**-0.02**
AinS(2.00)	**7.5**	**7.441**	**2.325**	**0.026**
AoutS(2.00)	**5.5**	**5.519**	**1.94**	**-0.01**
AinS(2.00)	7.5	7.441	2.325	0.026
AoutS(2.00)	5.5	5.519	1.94	-0.01
Ain1out-star(2.00)	11.5	11.177	3.372	0.096
1inAout-star(2.00)	11.5	11.556	3.514	-0.016
AinAout-star(2.00)	**10**	**9.968**	**2.713**	**0.012**
AT-T(2.00)	0	0.101	0.663	-0.152
AT-C(2.00)	0	0.09	0.651	-0.138
AT-D(2.00)	0	0.101	0.663	-0.152
AT-U(2.00)	0	0.101	0.663	-0.152
AT-TD(2.00)	0	0.101	0.663	-0.152
AT-TU(2.00)	0	0.101	0.663	-0.152
AT-DU(2.00)	0	0.101	0.663	-0.152
AT-TDU(2.00)	0	0.101	0.663	-0.152
A2P-T(2.00)	12	12.878	4.547	-0.193
A2P-D(2.00)	5.5	5.892	2.285	-0.172

A2P-U(2.00)	7.5	8.352	3.101	-0.275
A2P-TD(2.00)	8.75	9.385	3.296	-0.193
A2P-TU(2.00)	9.75	10.615	3.554	-0.243
A2P-DU(2.00)	6.5	7.122	2.322	-0.268
A2P-TDU(2.00)	8.333	9.041	2.982	-0.237
Std Dev in-degree dist	0.559	0.562	0.048	-0.077
Skew in-degree dist	2.726	2.872	0.51	-0.287
Std Dev out-degree dist	0.526	0.523	0.038	0.079
Skew out-degree dist	2.7	2.526	0.4	0.435
CorrCoef in-out-degree dists	0.816	0.811	0.109	0.043
Global Clustering Cto	0	0.008	0.052	-0.149
Global Clustering Cti	0	0.005	0.036	-0.149
Global Clustering Ctm	0	0.007	0.043	-0.152
Global Clustering Ccm	0	0.006	0.042	-0.138
Global Clustering AKC-T	0	0.007	0.043	-0.152
Global Clustering AKC-D	0	0.008	0.052	-0.149
Global Clustering AKC-U	0	0.005	0.036	-0.149
Global Clustering AKC-C	0	0.006	0.042	-0.138

11.4.3. ERGMs of Early Imperial attributed networks

Estimations attribute models

Significant configurations are indicated by *. Configurations are significant if the estimate is more than two times the standard error.

Table 37. Early Imperial 20km circuit model with transport network attribute.

Effects	Estimates	Standard error	T-ratio	
reciprocity	10.055539	1.55015	-0.04154	*
2-in-star	-5.708482	1.90495	0.01498	*
2-out-star	-1.999473	1.38887	0.05745	
path2	1.847137	0.82768	0.04641	*
030T	0.578545	0.16138	0.05334	*
isolates	-1.34101	0.97791	-0.01551	
AinS(2.00)	6.79016	2.27348	-0.00888	*
AoutS(2.00)	1.284343	1.5239	0.04105	
AinAout-star(2.00)	-0.11767	0.81402	-0.06978	
attribute_interaction	-0.028675	0.71317	-0.00496	
attribute_sender	0.563154	0.70018	0.04241	
attribute_receiver	-0.295265	0.82344	0.06534	
attribute_in2star	-1.322136	2.06961	0.0622	
attribute_path2	-0.294563	1.35667	0.0645	
attribute_out2star	1.261559	1.29692	0.05888	

Table 38. Early Imperial 20km circuit model with ‘sites on elevation’ attribute.

Effects	Estimates	Standard error	T-ratio	
reciprocity	11.877478	1.5637	0.02118	*
2-in-star	-6.439163	1.98779	0.00043	*

2-out-star	-6.486679	1.67387	0.01959	*
path2	4.323526	0.92855	0.00294	*
030T	0.601893	0.15858	0.00768	*
isolates	-1.658075	0.94566	0.06074	
AinS(2.00)	6.792351	2.35062	0.01144	*
AoutS(2.00)	0.763858	1.27458	0.02395	
AinAout-star(2.00)	0.179412	0.80513	0.0326	
attribute_interaction	0.079965	1.25665	-0.02934	
attribute_sender	0.834882	1.16092	-0.0222	
attribute_receiver	1.765496	1.27449	-0.0633	
attribute_interaction_reciprocity	-2.219687	1.31417	-0.01667	
attribute_in2star	0.950885	1.40204	-0.0431	
attribute_path2	-3.386105	1.05128	-0.02311	*
attribute_out2star	6.253632	1.58822	0.00642	*

Table 39. Early Imperial 20km circuit model with urban status attribute.

Effects	**Estimates**	**Standard error**	**T-ratio**	
reciprocity	10.74847	1.46607	0.05017	*
2-in-star	-7.111442	1.95956	0.06285	*
2-out-star	-3.179045	1.28336	-0.00022	*
path2	2.668022	0.81965	0.04433	*
030T	0.584527	0.16368	0.08331	*
isolates	-1.711183	0.92062	0.03844	
AinS(2.00)	7.85798	2.40963	0.04644	*
AoutS(2.00)	2.051864	1.28003	0.01099	

AinAout-star(2.00)	-0.402732	0.76709	0.03331	
attribute_interaction	-0.426453	0.69449	0.0174	
attribute_sender	0.300056	0.67564	-0.06309	
attribute_receiver	0.526665	0.65393	-0.03843	
attribute_in2star	1.536405	1.44753	-0.01246	
attribute_path2	-2.642974	1.09304	-0.02586	*
attribute_out2star	3.267091	1.20885	-0.03691	*

Table 40. Early Imperial 20km circuit model with Iron age origins attribute.

Effects	Estimates	Standard error	T-ratio	
reciprocity	11.41507	1.53122	0.04396	*
2-in-star	-7.392486	2.13442	-0.01468	*
2-out-star	-9.221778	1.85689	-0.00868	*
path2	4.821053	0.93025	-0.01785	*
030T	0.6563	0.17354	0.08788	*
isolates	-1.92352	0.9423	0.00995	*
AinS(2.00)	7.677137	2.32622	-0.00981	*
AoutS(2.00)	1.777674	1.35468	0.00768	
AinAout-star(2.00)	-0.044307	0.86487	0.02686	
attribute_interaction	-0.638658	1.19763	0.08372	
attribute_sender	0.038114	1.24334	0.08916	
attribute_receiver	1.370365	1.26671	0.08671	
attribute_interaction_reciprocity	-0.071999	1.39578	0.07913	
attribute_in2star	1.891361	1.44902	0.04275	
attribute_path2	-4.301578	1.08583	0.01711	*

attribute_out2star	8.984439	1.75245	0.00297	*

Table 41. Early Imperial 20km circuit model with three attributes: sites on elevation, sites with Iron Age origins, and sites with an urban status.

effects	estimates	stderr	t-ratio	
reciprocity	9.99	1.46	-0.05	*
2-in-star	-7.70	1.84	-0.04	*
2-out-star	-1.09	1.42	-0.05	
path2	2.05	1.14	-0.05	
030T	0.59	0.15	-0.12	
isolates	-1.25	0.99	-0.01	
AinS(2.00)	8.24	2.35	-0.02	*
AoutS(2.00)	0.55	1.56	-0.03	
AinAout-star(2.00)	-0.03	0.86	0.04	
Status_interaction	0.33	0.55	-0.01	
Elevation_interaction	-0.30	0.60	-0.06	
Occupation_interaction	0.29	0.52	0.03	
Status_sender	0.38	0.76	0.01	
Elevation_sender	0.96	0.73	-0.02	
Occupation_sender	-0.62	0.77	-0.01	
Status_receiver	-0.56	0.84	-0.01	
Elevation_receiver	-0.07	0.71	-0.06	
Occupation_receiver	0.05	0.88	-0.02	
Status_in2star	-0.13	1.72	-0.05	
Elevation_in2star	3.21	1.46	-0.02	*
Occupation_in2star	-2.12	1.57	-0.03	
Status_path2	0.24	1.35	-0.04	
Elevation_path2	-1.95	1.11	-0.03	
Occupation_path2	0.90	1.29	-0.04	

Status_out2star	0.12	1.43	-0.04
Elevation_out2star	0.99	1.01	-0.01
Occupation_out2star	-0.52	1.40	-0.04

Goodness of fit attribute models

Configurations included in the model are in bold. T-ratios express how good the model succeeds in reproducing the frequency of configurations in the observed network. T-ratios are calculated as (observed - mean) / standard deviation. In order to be a good fit the t-ratio for these configurations should be lower than 0.1, t-ratios for all other configurations should be between 0.1 and 2.

Table 42. Goodness of fit for circuit model of Early Imperial 20km network with transport network attribute. 100milion simulations, 1000 samples.

Effects	Observed	Mean	Standard deviation	T-ratio
reciprocity	**18**	**17.969**	**1.709**	**0.018**
2-in-star	**25**	**24.946**	**4.413**	**0.012**
2-out-star	**31**	**31.1**	**8.98**	**-0.011**
3-in-star	6	6.138	2.866	-0.048
3-out-star	16	17.253	13.364	-0.094
path2	**57**	**57.036**	**11.327**	**-0.003**
T1	2	1.686	1.388	0.226
T2	12	10.717	8.366	0.153
T3	12	11.379	8.482	0.073
T4	6	5.685	4.259	0.074
T5	6	5.685	4.25	0.074
T6	20	17.493	4.688	0.535
T7	44	41.868	8.678	0.246
T8	50	46.888	11.634	0.267
T9(030T)	12	12.086	8.736	-0.01
T10(030C)	4	4.041	2.903	-0.014
Sink	8	8.447	2.693	-0.166
Source	8	8.157	2.678	-0.059
Isolates	**110**	**110.032**	**4.189**	**-0.008**
AinS(2.00)	**22**	**21.972**	**3.301**	**0.008**
AoutS(2.00)	**24.125**	**24.136**	**5.003**	**-0.002**

AinS(2.00)	22	21.972	3.301	0.008
AoutS(2.00)	24.125	24.136	5.003	-0.002
Ain1out-star(2.00)	45	44.271	6.571	0.111
1inAout-star(2.00)	42.125	41.627	5.558	0.09
AinAout-star(2.00)	**31.875**	**31.858**	**3.458**	**0.005**
AT-T(2.00)	11	11.161	7.313	-0.022
AT-C(2.00)	11	11.194	7.281	-0.027
AT-D(2.00)	11	11.213	7.357	-0.029
AT-U(2.00)	11	11.089	7.261	-0.012
AT-TD(2.00)	11	11.187	7.334	-0.026
AT-TU(2.00)	11	11.125	7.286	-0.017
AT-DU(2.00)	11	11.151	7.306	-0.021
AT-TDU(2.00)	11	11.155	7.308	-0.021
A2P-T(2.00)	52	55.618	10.395	-0.348
A2P-D(2.00)	28.5	30.38	8.522	-0.221
A2P-U(2.00)	22.5	24.233	4.026	-0.431
A2P-TD(2.00)	40.25	42.999	9.236	-0.298
A2P-TU(2.00)	37.25	39.926	6.915	-0.387
A2P-DU(2.00)	25.5	27.306	5.604	-0.322
A2P-TDU(2.00)	34.333	36.744	7.159	-0.337
Transport_interaction	**2**	**1.941**	**1.985**	**0.03**
Transport_sender	**13**	**13.19**	**4.423**	**-0.043**
Transport_receiver	**9**	**9.038**	**3.283**	**-0.012**
Transport_interaction_reciprocity	1	0.754	0.91	0.27
Transport_activity_reciprocity	8	7.011	2.572	0.384
Transport_in2star	**5**	**5.06**	**3.152**	**-0.019**
Transport_path2	**18**	**18.385**	**11.16**	**-0.034**
Transport_out2star	**14**	**14.369**	**9.39**	**-0.039**
Std Dev in-degree dist	0.749	0.748	0.04	0.039

Skew in-degree dist	2.31	2.291	0.196	0.098
Std Dev out-degree dist	0.801	0.799	0.074	0.035
Skew out-degree dist	2.975	2.898	0.552	0.138
CorrCoef in-out-degree dists	0.84	0.84	0.042	0.014
Global Clustering Cto	0.194	0.187	0.121	0.051
Global Clustering Cti	0.24	0.232	0.151	0.054
Global Clustering Ctm	0.211	0.201	0.129	0.071
Global Clustering Ccm	0.211	0.202	0.129	0.065
Global Clustering AKC-T	0.212	0.194	0.119	0.146
Global Clustering AKC-D	0.193	0.181	0.112	0.106
Global Clustering AKC-U	0.244	0.224	0.14	0.149
Global Clustering AKC-C	0.212	0.195	0.118	0.141

Table 43. Goodness of fit for circuit model of Early Imperial 20km network with 'sites on elevation' attribute. 100milion simulations, 1000 samples.

Effects	Observed	Mean	Standard deviation	T-ratio
reciprocity	**18**	**17.987**	**1.781**	**0.007**
2-in-star	**25**	**25.46**	**4.405**	**-0.104**
2-out-star	**31**	**31.681**	**8.12**	**-0.084**
3-in-star	6	6.633	3.148	-0.201
3-out-star	16	17.831	11.791	-0.155
path2	**57**	**58.045**	**10.672**	**-0.098**
T1	2	1.777	1.299	0.172
T2	12	11.296	7.85	0.09
T3	12	11.977	8.003	0.003
T4	6	5.982	3.997	0.005
T5	6	6	4.005	0
T6	20	18.329	4.637	0.36
T7	44	43.07	8.763	0.106

T8	50	48.792	11.188	0.108
T9(030T)	**12**	**12.721**	**8.251**	**-0.087**
T10(030C)	4	4.237	2.754	-0.086
Sink	8	8.444	2.595	-0.171
Source	8	8.267	2.599	-0.103
Isolates	**110**	**110.104**	**3.924**	**-0.027**
AinS(2.00)	**22**	**22.281**	**3.223**	**-0.087**
AoutS(2.00)	**24.125**	**24.41**	**4.548**	**-0.063**
AinS(2.00)	22	22.281	3.223	-0.087
AoutS(2.00)	24.125	24.41	4.548	-0.063
Ain1out-star(2.00)	45	44.598	5.933	0.068
1inAout-star(2.00)	42.125	41.87	5.159	0.049
AinAout-star(2.00)	**31.875**	**31.827**	**3.332**	**0.014**
AT-T(2.00)	11	11.751	6.893	-0.109
AT-C(2.00)	11	11.735	6.898	-0.107
AT-D(2.00)	11	11.803	6.938	-0.116
AT-U(2.00)	11	11.676	6.836	-0.099
AT-TD(2.00)	11	11.777	6.915	-0.112
AT-TU(2.00)	11	11.713	6.863	-0.104
AT-DU(2.00)	11	11.739	6.883	-0.107
AT-TDU(2.00)	11	11.743	6.886	-0.108
A2P-T(2.00)	52	56.542	9.682	-0.469
A2P-D(2.00)	28.5	30.921	7.652	-0.316
A2P-U(2.00)	22.5	24.705	3.968	-0.556
A2P-TD(2.00)	40.25	43.731	8.476	-0.411
A2P-TU(2.00)	37.25	40.624	6.525	-0.517
A2P-DU(2.00)	25.5	27.813	5.17	-0.447
A2P-TDU(2.00)	34.333	37.389	6.638	-0.46
Elevation_interaction	**33**	**32.66**	**5.363**	**0.063**

Elevation_sender	**42**	**41.908**	**2.782**	**0.033**
Elevation_receiver	**41**	**40.785**	**3.323**	**0.065**
Elevation_interaction_reciprocity	12	11.874	2.622	0.048
Elevation_activity_reciprocity	18	17.884	1.808	0.064
Elevation_in2star	**20**	**20.257**	**4.366**	**-0.059**
Elevation_path2	**49**	**49.764**	**11.102**	**-0.069**
Elevation_out2star	**29**	**29.621**	**8.257**	**-0.075**
Std Dev in-degree dist	0.749	0.752	0.039	-0.079
Skew in-degree dist	2.31	2.326	0.206	-0.077
Std Dev out-degree dist	0.801	0.804	0.067	-0.044
Skew out-degree dist	2.975	2.954	0.5	0.042
CorrCoef in-out-degree dists	0.84	0.841	0.041	-0.01
Global Clustering Cto	0.194	0.197	0.118	-0.029
Global Clustering Cti	0.24	0.241	0.14	-0.005
Global Clustering Ctm	0.211	0.211	0.122	-0.006
Global Clustering Ccm	0.211	0.211	0.123	-0.005
Global Clustering AKC-T	0.212	0.203	0.113	0.072
Global Clustering AKC-D	0.193	0.19	0.109	0.026
Global Clustering AKC-U	0.244	0.232	0.129	0.099
Global Clustering AKC-C	0.212	0.203	0.113	0.073

Table 44. Goodness of fit for circuit model of Early Imperial 20km network with urban status attribute. 50milion simulations, 1000 samples.

Effects	Observed	Mean	Standard deviation	T-ratio
reciprocity	**18**	**18.132**	**1.834**	**-0.072**
2-in-star	**25**	**24.819**	**4.399**	**0.041**
2-out-star	**31**	**30.716**	**9.328**	**0.03**
3-in-star	6	6.056	2.697	-0.021
3-out-star	16	17.354	19.845	-0.068

path2	**57**	**56.654**	**10.533**	**0.033**
T1	2	1.603	1.364	0.291
T2	12	10.16	8.175	0.225
T3	12	10.775	8.24	0.149
T4	6	5.381	4.122	0.15
T5	6	5.372	4.114	0.153
T6	20	17.628	4.77	0.497
T7	44	41.858	8.72	0.246
T8	50	46.305	10.863	0.34
T9(030T)	**12**	**11.428**	**8.383**	**0.068**
T10(030C)	4	3.826	2.805	0.062
Sink	8	8.352	2.662	-0.132
Source	8	7.918	2.557	0.032
Isolates	**110**	**110.153**	**4.016**	**-0.038**
AinS(2.00)	**22**	**21.868**	**3.332**	**0.04**
AoutS(2.00)	**24.125**	**23.959**	**4.797**	**0.035**
AinS(2.00)	22	21.868	3.332	0.04
AoutS(2.00)	24.125	23.959	4.797	0.035
Ain1out-star(2.00)	45	44.12	6.242	0.141
1inAout-star(2.00)	42.125	41.805	5.482	0.058
AinAout-star(2.00)	**31.875**	**31.936**	**3.412**	**-0.018**
AT-T(2.00)	11	10.624	7.104	0.053
AT-C(2.00)	11	10.671	7.137	0.046
AT-D(2.00)	11	10.654	7.136	0.048
AT-U(2.00)	11	10.597	7.077	0.057
AT-TD(2.00)	11	10.639	7.119	0.051
AT-TU(2.00)	11	10.61	7.09	0.055
AT-DU(2.00)	11	10.625	7.104	0.053
AT-TDU(2.00)	11	10.625	7.104	0.053

A2P-T(2.00)	52	55.415	9.702	-0.352
A2P-D(2.00)	28.5	30.098	8.98	-0.178
A2P-U(2.00)	22.5	24.203	4.04	-0.422
A2P-TD(2.00)	40.25	42.756	8.946	-0.28
A2P-TU(2.00)	37.25	39.809	6.536	-0.392
A2P-DU(2.00)	25.5	27.151	5.58	-0.296
A2P-TDU(2.00)	34.333	36.572	6.867	-0.326
Status_interaction	**2**	**1.998**	**1.867**	**0.001**
Status_sender	**14**	**14.233**	**3.872**	**-0.06**
Status_receiver	**12**	**12.148**	**3.106**	**-0.048**
Status_interaction_reciprocity	1	0.731	0.854	0.315
Status_activity_reciprocity	8	9.083	2.508	-0.432
Status_in2star	**6**	**6.116**	**2.928**	**-0.04**
Status_path2	**18**	**18.341**	**9.772**	**-0.035**
Status_out2star	**14**	**14.109**	**9.712**	**-0.011**
Std Dev in-degree dist	0.749	0.747	0.039	0.068
Skew in-degree dist	2.31	2.288	0.178	0.124
Std Dev out-degree dist	0.801	0.795	0.076	0.079
Skew out-degree dist	2.975	2.87	0.684	0.153
CorrCoef in-out-degree dists	0.84	0.845	0.046	-0.115
Global Clustering Cto	0.194	0.181	0.121	0.101
Global Clustering Cti	0.24	0.22	0.147	0.135
Global Clustering Ctm	0.211	0.192	0.126	0.145
Global Clustering Ccm	0.211	0.193	0.127	0.138
Global Clustering AKC-T	0.212	0.186	0.116	0.222
Global Clustering AKC-D	0.193	0.176	0.112	0.155
Global Clustering AKC-U	0.244	0.213	0.136	0.228
Global Clustering AKC-C	0.212	0.187	0.117	0.214

Table 45. Goodness of fit for circuit model of Early Imperial 20km network with Iron Age origins attribute. 50milion simulations, 1000 samples.

Effects	Observed	Mean	Standard deviation	T-ratio
reciprocity	**18**	**17.988**	**1.82**	**0.007**
2-in-star	**25**	**25.048**	**4.582**	**-0.01**
2-out-star	**31**	**31.246**	**8.032**	**-0.031**
3-in-star	6	6.25	3.056	-0.082
3-out-star	16	16.54	10.37	-0.052
path2	**57**	**57.355**	**11.065**	**-0.032**
T1	2	1.694	1.268	0.241
T2	12	10.698	7.666	0.17
T3	12	11.259	7.803	0.095
T4	6	5.628	3.9	0.095
T5	6	5.688	3.937	0.079
T6	20	19.068	4.79	0.195
T7	44	42.881	9.195	0.122
T8	50	49.196	11.507	0.07
T9(030T)	**12**	**11.974**	**8.093**	**0.003**
T10(030C)	4	3.953	2.674	0.018
Sink	8	8.38	2.524	-0.151
Source	8	8.162	2.652	-0.061
Isolates	**110**	**110.071**	**4.061**	**-0.017**
AinS(2.00)	**22**	**22.032**	**3.366**	**-0.009**
AoutS(2.00)	**24.125**	**24.31**	**4.708**	**-0.039**
AinS(2.00)	22	22.032	3.366	-0.009
AoutS(2.00)	24.125	24.31	4.708	-0.039
Ain1out-star(2.00)	45	44.486	6.236	0.082
1inAout-star(2.00)	42.125	41.924	5.4	0.037
AinAout-star(2.00)	**31.875**	**31.968**	**3.18**	**-0.029**

AT-T(2.00)	11	11.195	6.915	-0.028
AT-C(2.00)	11	11.089	6.858	-0.013
AT-D(2.00)	11	11.251	6.961	-0.036
AT-U(2.00)	11	11.101	6.832	-0.015
AT-TD(2.00)	11	11.223	6.937	-0.032
AT-TU(2.00)	11	11.148	6.872	-0.022
AT-DU(2.00)	11	11.176	6.893	-0.026
AT-TDU(2.00)	11	11.182	6.9	-0.026
A2P-T(2.00)	52	56.062	10.107	-0.402
A2P-D(2.00)	28.5	30.585	7.567	-0.276
A2P-U(2.00)	22.5	24.393	4.173	-0.454
A2P-TD(2.00)	40.25	43.323	8.664	-0.355
A2P-TU(2.00)	37.25	40.227	6.897	-0.432
A2P-DU(2.00)	25.5	27.489	5.324	-0.374
A2P-TDU(2.00)	34.333	37.013	6.889	-0.389
Origins_interaction	**35**	**35.27**	**4.925**	**-0.055**
Origins_sender	**43**	**43.146**	**2.537**	**-0.058**
Origins_receiver	**42**	**42.181**	**3.189**	**-0.057**
Origins_interaction_reciprocity	14	14.159	2.488	-0.064
Origins_activity_reciprocity	18	17.928	1.831	0.039
Origins_in2star	**21**	**21.205**	**4.677**	**-0.044**
Origins_path2	**50**	**50.577**	**11.411**	**-0.051**
Origins_out2star	**30**	**30.308**	**8.064**	**-0.038**
Std Dev in-degree dist	0.749	0.749	0.041	0.017
Skew in-degree dist	2.31	2.296	0.201	0.071
Std Dev out-degree dist	0.801	0.801	0.066	0.01
Skew out-degree dist	2.975	2.888	0.434	0.199
CorrCoef in-out-degree dists	0.84	0.84	0.044	0
Global Clustering Cto	0.194	0.184	0.11	0.086

Global Clustering Cti	0.24	0.228	0.136	0.086
Global Clustering Ctm	0.211	0.198	0.117	0.105
Global Clustering Ccm	0.211	0.197	0.116	0.12
Global Clustering AKC-T	0.212	0.192	0.108	0.178
Global Clustering AKC-D	0.193	0.179	0.103	0.136
Global Clustering AKC-U	0.244	0.221	0.127	0.186
Global Clustering AKC-C	0.212	0.191	0.108	0.194

Table 46. Goodness of fit for circuit model of Early Imperial 20km network with three attributes: sites on elevation, sites with Iron Age origins, and sites with an urban status. 50milion simulations, 1000 samples.

Effects	Observed	Mean	Standard deviation	T-ratio
reciprocity	**18**	**18.098**	**1.757**	**-0.056**
2-in-star	**25**	**25.283**	**4.564**	**-0.062**
2-out-star	**31**	**31.095**	**8.585**	**-0.011**
3-in-star	6	6.492	3.301	-0.149
3-out-star	16	17.03	12.691	-0.081
path2	**57**	**57.379**	**11.44**	**-0.033**
T1	2	1.696	1.362	0.223
T2	12	10.665	8.214	0.163
T3	12	11.203	8.329	0.096
T4	6	5.591	4.165	0.098
T5	6	5.628	4.205	0.088
T6	20	18.386	4.952	0.326
T7	44	42.918	9.163	0.118
T8	50	48.024	12.045	0.164
T9(030T)	**12**	**11.833**	**8.594**	**0.019**
T10(030C)	4	3.936	2.846	0.022
Sink	8	8.361	2.505	-0.144
Source	8	8.197	2.614	-0.075

Isolates	**110**	**110.141**	**3.978**	**-0.035**
AinS(2.00)	**22**	**22.175**	**3.302**	**-0.053**
AoutS(2.00)	**24.125**	**24.165**	**4.757**	**-0.008**
Ain1out-star(2.00)	45	44.337	6.251	0.106
1inAout-star(2.00)	42.125	41.656	5.331	0.088
AinAout-star(2.00)	**31.875**	**31.916**	**3.251**	**-0.013**
AT-T(2.00)	11	10.939	6.985	0.009
AT-C(2.00)	11	10.911	6.951	0.013
AT-D(2.00)	11	10.993	7.029	0.001
AT-U(2.00)	11	10.861	6.904	0.02
AT-TD(2.00)	11	10.966	7.006	0.005
AT-TU(2.00)	11	10.9	6.943	0.014
AT-DU(2.00)	11	10.927	6.962	0.01
AT-TDU(2.00)	11	10.931	6.97	0.01
A2P-T(2.00)	52	55.992	10.131	-0.394
A2P-D(2.00)	28.5	30.394	7.922	-0.239
A2P-U(2.00)	22.5	24.592	4.053	-0.516
A2P-TD(2.00)	40.25	43.193	8.852	-0.332
A2P-TU(2.00)	37.25	40.292	6.817	-0.446
A2P-DU(2.00)	25.5	27.493	5.418	-0.368
A2P-TDU(2.00)	34.333	36.993	6.952	-0.383
Status_interaction	**4**	**3.54**	**2.993**	**0.154**
Elevation_interaction	**23**	**22.69**	**5.624**	**0.055**
Occupation_interaction	**26**	**25.163**	**6.041**	**0.139**
Status_sender	**14**	**13.254**	**4.985**	**0.15**
Elevation_sender	**35**	**34.872**	**4.045**	**0.032**
Occupation_sender	**36**	**35.283**	**4.75**	**0.151**
Status_receiver	**11**	**10.389**	**4.134**	**0.148**
Elevation_receiver	**35**	**34.913**	**3.533**	**0.025**

Occupation_receiver	**35**	**34.366**	**4.724**	**0.134**
Status_sender_missing	0	0	0	
Elevation_sender_missing	0	0	0	
Occupation_sender_missing	0	0	0	
Status_receiver_missing	0	0	0	
Elevation_receiver_missing	0	0	0	
Occupation_receiver_missing	0	0	0	
Status_interaction_reciprocity	2	1.4	1.407	0.426
Elevation_interaction_reciprocity	9	8.683	2.652	0.12
Occupation_interaction_reciprocity	8	8.027	2.711	-0.01
Status_activity_reciprocity	7	7.046	2.888	-0.016
Elevation_activity_reciprocity	17	17.089	2.049	-0.043
Occupation_activity_reciprocity	15	15.273	2.425	-0.113
Status_in2star	**7**	**6.616**	**4.156**	**0.092**
Elevation_in2star	**18**	**18.19**	**4.129**	**-0.046**
Occupation_in2star	**14**	**13.589**	**4.409**	**0.093**
Status_path2	**21**	**19.759**	**12.595**	**0.099**
Elevation_path2	**36**	**36.396**	**9.549**	**-0.041**
Occupation_path2	**36**	**34.755**	**12.13**	**0.103**
Status_out2star	**14**	**13.203**	**9.333**	**0.085**
Elevation_out2star	**18**	**18.186**	**6.299**	**-0.03**
Occupation_out2star	**21**	**20.176**	**8.872**	**0.093**
Std Dev in-degree dist	0.749	0.751	0.041	-0.035
Skew in-degree dist	2.31	2.313	0.215	-0.013
Std Dev out-degree dist	0.801	0.799	0.071	0.033
Skew out-degree dist	2.975	2.895	0.534	0.149
CorrCoef in-out-degree dists	0.84	0.842	0.044	-0.045
Global Clustering Cto	0.194	0.181	0.109	0.118
Global Clustering Cti	0.24	0.224	0.141	0.117

Global Clustering Ctm	0.211	0.195	0.119	0.131
Global Clustering Ccm	0.211	0.195	0.119	0.134
Global Clustering AKC-T	0.212	0.188	0.109	0.212
Global Clustering AKC-D	0.193	0.175	0.101	0.177
Global Clustering AKC-U	0.244	0.216	0.13	0.22
Global Clustering AKC-C	0.212	0.188	0.109	0.216

12. Appendix IV: case study 3

12.1. Lists of sites and forms used

Table 47. Sites in ICRATES database in alphabetical order.

Sites in ICRATES database in alphabetical order	
Aazaz	Ed-Dur
Abdera	Eisodei_tis_Theotokou_kai_Ayios_Petros_(26)
Adulis-Diodoros_Island	El_Aareime
Ain_Dara	Emporio
Aizanoi	Ephesos
Aleppo	Epiphaneia
Alexandreia	Eretria
Altinum	Gadara
Amathous	Gebel_Barkal
Amorion	Gerasa
Amphipolis	Gindaros
Amygdalea	Glyfada
Anemorion	Gortyn
Antikythera_shipwreck	Halikarnassos
Antiocheia_ad_Orontem	Hammath_Tiberias
Antiocheia_ad_Pisidiam	Haouar_enn_Nahr
Apamea	Hippos-Sussita
Apollo_Smintheion	Iasos
Argos	Isthmia
Arikamedu	Jalame
Arsameia_am_Nymphaios	Jebel_Khalid
Asagi-Dikenli	Jericho
Asea_Valley	Jerusalem
Ashkelon	Kallion
Assos	Kallirhoe
Athens	Kanatha
Athis_(Neocaesareia – Qasrin – Dibsi_Faraj)	Karamildan
Axum	Karanog
Ayios_Philon	Kastro_Tigani
Azotos_(Ashdod)	Kenchreai
B29iv	Kepia
Bab	Khirbet_ez_Souaine
BERB96-Findspot_420/Tract_443	Kition
BERB96-Findspots_500-522	Knossos
Berenice	Kommos
Berenike	Kopetra
Berytus	Kourion
Butrint	Kozluca
Byblos	Kucuk_Burnaz
Caesarea_Maritima	Kululu
Carthage	Kydonia
Corinth	Kythera
Cyrene	Labraunda
Damaskos	Lepcis_Magna
Danakaya	Leukos_Limen
Delos	Lidar_Hoyuk
Diaseli_Otzia_(55)	Limassol
Didyma	Malta
Diokaesareia	Mampsis
Doliche	Marina_el-Alamein
Dor	Maroni_Petrera
Dura_Europos	Meroe

Sites in ICRATES database in alphabetical order (continued)

Methymna	Pella	Tell_Ilbol
MS010	Pelusium	Tell_Kadrich
MS116	Pergamon	Tell_Kaffine
MS216	Perge	Tell_Kassiha
Mutatio_Heldua	Petra	Tell_Khibi
Myos_Hormos	Phalasarna	Tell_Noubbol
Nessana	Phaselis_shipwreck	Tell_Qaramel
Nikopolis	Philadelphia	Tell_Rahhal
Oboda	Porphyreon	Tell_Rifaat
Olympia	Porsuk	Tell_Sourane
ORO01-90/3_Vlastos/Hydragogeio	Priene	Tell_Soussiane
Oumm_el-Amed	Pylos	Tell_Zaitane
Palai-79D-10	Qara_Keupru	Tenos
Palai-79D-12	Qara_Mazraa	Thasos
Palai-79D-7	Qusair_as-Saila	Timna
Palai-79D-9	Resafa	Tourhleu
Palai-79X-1	Salamis/Constantia	Troia/Ilion
Palai-80E-22	Samaria-Sebaste	Umm_el-Tlel
Palai-80E-28	Samos_(Heraion)	Uruk
Palai-80E-30	Samothrace-Hieron	Veloukovo
Palai-80E-30_6	Sardis	Xanthos
Palai-80E-4	Scythopolis	Yel_Baba
Palai-80E-41	Seleukeia_ad_Tigrim	Zeugma
Palai-80E-48	Siphnos	
Palai-80X-8	SK7	
Palai-83D-23	SP1B	
Palai-83D-27	Sparta	
Palai-83D-58	Stobi	
Palai-83E-126	Sultantepe	
Palai-83E-128	Sumhuram	
Palai-83E-130	Sykea	
Palai-83E-18	Tall She Hamad/Magdala	
Palai-83E-20	Tanagra	
Palai-83E-26	Tanagra-TS21	
Palai-83E-53	Tanagra-TS4	
Palai-83E-63	Tarsos	
Palai-83E-67	Tel_Anafa	
Palai-83E-68	Tel_Mevorakh	
Palai-86D-16	Tell_Aajar	
Palai-86D-22	Tell_Aar	
Palai-86E-1	Tell_Aarane	
Palai-86E-126	Tell_Akhtareine	
Palai-86E-2	Tell_Atrib	
Palai-86E-4	Tell_Bahouerte	
Palai-86E-5	Tell_Banat	
Palai-86K-10	Tell_Bararhite	
Palai-86K-4	Tell_Berne_(West)	
Palaipaphos_Area	Tell_Botnan	
Palaityr/Tell_Arqa	Tell_el_Qoubli	
Panayia_Ematousa	Tell_Fafine	
Paphos	Tell_Hailane	
Patras	Tell_Haourane	

Table 48. List of all forms used grouped per fabric, with their standard typo-chronological lower and upper dates.

Standard Form	Fabric	Standard typo-chronological lower date	Standard typo-chronological upper date
EAA1	ESA	-150	-100
EAA10	ESA	-50	-25
EAA101	ESA	-100	-1
EAA102	ESA	-100	-1
EAA102/105	ESA	-100	-1
EAA104A	ESA	-50	50
EAA104A-B	ESA	-50	50
EAA104B	ESA	-50	50
EAA105	ESA	-100	-75
EAA106	ESA	1	25
EAA107	ESA	1	25
EAA108	ESA	1	100
EAA109	ESA	50	100
EAA11	ESA	-50	-1
EAA111	ESA	25	100
EAA114	ESA	25	75
EAA116A	ESA	70	120
EAA116A-B	ESA	70	120
EAA117	ESA	70	120
EAA12	ESA	-40	10
EAA12/32	ESA	-40	30
EAA13A	ESA	-50	25
EAA13B	ESA	-50	25
EAA14	ESA	1	25
EAA15A	ESA	-100	-50
EAA15A-B	ESA	-100	-50
EAA15B	ESA	-100	-50
EAA16	ESA	-175	-125
EAA17A	ESA	-150	-100
EAA17A-B	ESA	-150	-100
EAA17B	ESA	-150	-100
EAA18	ESA	-125	-75
EAA19A	ESA	-100	-50
EAA19A-B	ESA	-100	-50
EAA19B	ESA	-100	-50
EAA20	ESA	-150	-100
EAA21/22	ESA	-27	14
EAA22/43	ESA	-125	14
EAA22A	ESA	-125	10
EAA22A-B	ESA	-125	10
EAA22B	ESA	-125	10
EAA23	ESA	-100	-50
EAA2-3	ESA	-200	-1
EAA24	ESA	-100	100
EAA24-25	ESA	-100	100
EAA25	ESA	-100	100
EAA26-27	ESA	-10	30
EAA26A	ESA	-10	30
EAA26A-D	ESA	-10	30
EAA26B	ESA	-10	30
EAA26C	ESA	-10	30
EAA26D	ESA	-10	30
EAA27	ESA	-10	30
EAA28	ESA	-10	30
EAA28-30	ESA	-30	50
EAA29	ESA	-30	25

EAA2A	ESA	-150	-100
EAA2A-B	ESA	-200	-50
EAA2B	ESA	-200	-50
EAA3	ESA	-125	-1
EAA30	ESA	10	50
EAA30/33	ESA	1	50
EAA30/33-34	ESA	1	50
EAA31	ESA	1	25
EAA32	ESA	1	30
EAA33	ESA	1	50
EAA33/36	ESA	1	100
EAA34	ESA	25	50
EAA3-4	ESA	-125	20
EAA34/37	ESA	25	100
EAA35	ESA	40	70
EAA35/40	ESA	40	120
EAA35-37	ESA	40	100
EAA36	ESA	60	100
EAA36-37	ESA	60	100
EAA37A	ESA	60	100
EAA37A-B	ESA	60	100
EAA37B	ESA	60	100
EAA38	ESA	25	75
EAA38-39	ESA	25	100
EAA39	ESA	60	100
EAA40A	ESA	80	120
EAA40A-B	ESA	80	120
EAA40A-C	ESA	80	120
EAA40C	ESA	80	120
EAA41	ESA	60	100
EAA42	ESA	-10	30
EAA43	ESA	1	14
EAA44	ESA	1	50
EAA45	ESA	1	60
EAA46	ESA	1	25
EAA46-47	ESA	1	70
EAA47	ESA	10	70
EAA48	ESA	40	70
EAA49	ESA	40	70
EAA4A	ESA	-125	20
EAA4A-B	ESA	-125	20
EAA4B	ESA	-27	14
EAA50	ESA	60	100
EAA50-51	ESA	60	120
EAA51	ESA	70	120
EAA52	ESA	115	140
EAA53	ESA	75	125
EAA53-54	ESA	75	150
EAA54	ESA	75	150
EAA54/56	ESA	75	200
EAA55	ESA	100	175
EAA56	ESA	150	200
EAA57	ESA	100	150
EAA57/59	ESA	100	150
EAA57/60A-B	ESA	100	200
EAA57-59	ESA	100	150
EAA57-60B	ESA	100	200
EAA58	ESA	100	150
EAA58/60	ESA	100	200
EAA59	ESA	100	150
EAA5A	ESA	-125	25

EAA5A-B	ESA	-125	25
EAA5B	ESA	-125	25
EAA6	ESA	-125	-50
EAA60A	ESA	100	150
EAA60A-B	ESA	100	200
EAA60B	ESA	100	200
EAA61	ESA	100	125
EAA62	ESA	90	120
EAA65	ESA	80	120
EAA7	ESA	-50	-1
EAA8	ESA	-50	-1
EAA9	ESA	-50	-25
EAA9-10	ESA	-50	-25
EAA9-11	ESA	-50	-1
EAArara-a	ESA	-100	-75
EAAtarda-a	ESA	175	200
EAAtarda-b	ESA	175	200
EAAtarda-c	ESA	175	200
EAAtarda-d	ESA	175	200
EAAtarda-e	ESA	175	200
EAAtarda-e/f	ESA	175	200
EAAtarda-f	ESA	175	200
EAAtarda-g	ESA	175	200
EAAtarda-h	ESA	175	200
EAA1	ESB	-27	14
EAA1-3	ESB	-27	14
EAA13A	ESB	25	75
EAA13A-B	ESB	25	75
EAA13B	ESB	25	75
EAA16	ESB	-27	14
EAA17A	ESB	25	50
EAA17A-B	ESB	25	50
EAA17B	ESB	25	50
EAA18	ESB	25	75
EAA19	ESB	25	75
EAA2	ESB	-27	14
EAA21	ESB	-27	14
EAA22	ESB	-27	14
EAA3	ESB	-27	14
EAA30	ESB	1	25
EAA30-31	ESB	1	50
EAA30-32	ESB	25	75
EAA31	ESB	25	50
EAA32	ESB	25	75
EAA35	ESB	25	75
EAA36	ESB	25	75
EAA37	ESB	25	75
EAA38	ESB	1	25
EAA39	ESB	1	75
EAA4	ESB	1	25
EAA40	ESB	1	25
EAA5	ESB	25	50
EAA58	ESB	50	125
EAA58 early	ESB	50	75
EAA58 late	ESB	75	125
EAA6	ESB	25	50
EAA60	ESB	50	150
EAA60 early	ESB	50	90
EAA60 late	ESB	80	150
EAA62A	ESB	70	120
EAA62B	ESB	70	120

EAA63	ESB	70	120
EAA64/65	ESB	50	80
EAA65	ESB	50	80
EAA6-7	ESB	25	50
EAA7	ESB	25	50
EAA70	ESB	50	125
EAA70 early	ESB	50	75
EAA70 late	ESB	75	125
EAA71	ESB	70	120
EAA72	ESB	50	75
EAA74A	ESB	70	120
EAA74A-B	ESB	70	120
EAA74B	ESB	70	120
EAA76A	ESB	50	100
EAA76A-B	ESB	50	150
EAA76B	ESB	100	150
EAA77	ESB	100	150
EAA78	ESB	100	150
EAA79	ESB	75	125
EAA8	ESB	25	75
EAA80	ESB	80	150
EAA9	ESB	25	50
EAAL1	ESC	25	100
EAAL15	ESC	25	100
EAAL19	ESC	75	125
EAAL20	ESC	25	100
EAAL26A	ESC	25	100
EAAL26B	ESC	100	150
EAAL28	ESC	100	150
EAAL6	ESC	25	100
EAAL9	ESC	1	200
EAAL9A	ESC	1	100
LRP1	ESC	125	275
LRP1 early	ESC	125	200
LRP1 late	ESC	200	275
LRP1/2	ESC	100	275
LRP2	ESC	100	200
LRP3	ESC	100	300
LRP4	ESC	175	300
LRP5	ESC	200	250
M-SA1c	ESC	-100	-1
M-SB3	ESC	-125	-25
M-SB6	ESC	-200	-25
M-SB8	ESC	-100	25
M-SB9	ESC	1	200
M-SK1	ESC	-150	-50
M-SK2	ESC	-125	-100
M-SK3	ESC	-50	75
M-SK5	ESC	1	50
M-SKg1	ESC	-75	75
M-SKg2	ESC	-100	75
M-SKr10	ESC	-125	-1
M-SKr6	ESC	-25	75
M-SN1	ESC	-175	25
M-SN11a	ESC	-25	175
M-SN12	ESC	-100	75
M-SN12-13	ESC	-100	75
M-SN13	ESC	1	50
M-SN15c	ESC	-25	100
M-SN2	ESC	-100	75
M-SN20/N40	ESC	-50	75

M-SN21	ESC	-75	14
M-SN25	ESC	1	100
M-SN3	ESC	-75	100
M-SN33a	ESC	-75	100
M-SN33a-d	ESC	-75	150
M-SN33b	ESC	-25	25
M-SN33c	ESC	25	50
M-SN33d	ESC	25	150
M-SN34	ESC	-75	125
M-SN36	ESC	-75	75
M-SN37	ESC	-50	25
M-SN39a	ESC	-75	100
M-SN39a-d	ESC	-75	75
M-SN39b	ESC	-75	75
M-SN39c	ESC	25	50
M-SN39d	ESC	25	150
M-SN40	ESC	-25	75
M-SN43	ESC	-25	75
M-SN4b	ESC	-75	75
M-SN7	ESC	-75	25
M-SN8	ESC	-25	75
M-SS1	ESC	-175	25
M-SS2	ESC	-200	25
M-SS3	ESC	-75	25
M-SS5	ESC	-25	75
M-SS6	ESC	-150	-50
M-SS7	ESC	-175	-25
M-SS8	ESC	-125	-25
M-SSa10	ESC	-125	-25
M-SSa12	ESC	-50	75
M-SSa13a	ESC	-25	75
M-SSa14a	ESC	-100	200
M-SSa14b	ESC	1	125
M-SSa15	ESC	-25	50
M-SSa16	ESC	1	200
M-SSa17a	ESC	1	125
M-SSa2	ESC	-75	100
M-SSa20	ESC	1	200
M-SSa21	ESC	1	75
M-SSa22	ESC	50	175
M-SSa25	ESC	-25	75
M-SSa26	ESC	-25	100
M-SSa27a	ESC	50	175
M-SSa27a-c	ESC	50	200
M-SSa27b	ESC	50	175
M-SSa27c	ESC	50	200
M-SSa3	ESC	-50	75
M-SSa30	ESC	-50	-1
M-SSa31	ESC	-25	75
M-SSa4	ESC	-25	75
M-SSu10	ESC	-50	50
M-SSu13	ESC	1	100
M-SSu14	ESC	1	200
M-SSu15	ESC	1	200
M-SSu16	ESC	1	200
M-SSu18	ESC	25	200
M-SSu19a	ESC	-75	100
M-SSu19a-b	ESC	-75	75
M-SSu21	ESC	-50	175
M-SSu22a	ESC	75	300
M-SSu23a-b	ESC	-125	25

M-SSu25	ESC	-75	25
M-SSu5	ESC	-75	125
M-SSu9	ESC	-25	175
M-ST16	ESC	-75	75
M-ST17	ESC	-75	75
M-ST2	ESC	-125	25
M-ST20	ESC	-25	125
M-ST22	ESC	1	100
M-ST23	ESC	25	150
M-ST24b	ESC	40	200
M-ST26	ESC	100	200
M-ST27	ESC	-125	-25
M-ST3	ESC	-75	125
M-ST30	ESC	-75	75
M-ST31a	ESC	-75	75
M-ST31a-e	ESC	-75	75
M-ST31b	ESC	-75	75
M-ST31c	ESC	-75	75
M-ST31d	ESC	-75	75
M-ST34	ESC	-25	75
M-ST35	ESC	-25	75
M-ST36	ESC	-25	75
M-ST4	ESC	-75	100
M-ST6	ESC	-75	75
M-ST8a	ESC	-25	175
M-STs1b	ESC	-75	33
M-STs1c	ESC	-75	-25
M-STs3	ESC	-75	50
EAAP1	ESD	-100	-1
EAAP10	ESD	1	75
EAAP10/P11	ESD	1	150
EAAP11	ESD	50	150
EAAP11/P12	ESD	50	150
EAAP12	ESD	50	150
EAAP14	ESD	-100	-1
EAAP15	ESD	-100	-75
EAAP17	ESD	-100	-25
EAAP18A	ESD	-100	-1
EAAP18A-B	ESD	-100	-1
EAAP18B	ESD	-100	-1
EAAP19	ESD	1	100
EAAP2	ESD	-100	-1
EAAP20	ESD	-100	-75
EAAP21	ESD	-100	-25
EAAP22A	ESD	-25	25
EAAP22B	ESD	75	120
EAAP23B	ESD	1	25
EAAP24	ESD	1	25
EAAP25	ESD	1	25
EAAP26	ESD	25	100
EAAP28	ESD	1	100
EAAP29	ESD	100	150
EAAP3	ESD	-27	14
EAAP30A-B	ESD	100	150
EAAP31A	ESD	100	200
EAAP31B	ESD	100	200
EAAP33	ESD	-75	-25
EAAP33-34	ESD	-100	-1
EAAP34	ESD	-100	-1
EAAP36	ESD	-75	-25
EAAP37A	ESD	-100	25

EAAP37A-B	ESD	-100	25
EAAP37B	ESD	-100	25
EAAP40	ESD	90	150
EAAP41	ESD	100	200
EAAP42	ESD	175	300
EAAP44	ESD	-75	-25
EAAP44/X45	ESD	-75	14
EAAP47	ESD	-75	-25
EAAP48	ESD	-75	-25
EAAP49	ESD	-75	-1
EAAP4A	ESD	-27	14
EAAP4B	ESD	1	75
EAAP4B/P6	ESD	1	75
EAAP5	ESD	1	75
EAAP5/P6	ESD	1	75
EAAP50	ESD	1	25
EAAP54	ESD	70	100
EAAP6	ESD	1	75
EAAP6/P4B	ESD	1	75
EAAP7	ESD	1	25
EAAP7/P8	ESD	1	25
EAAP8	ESD	1	25
EAAP9	ESD	25	75

12.2. Exploratory network analysis case study 2

12.2.1. Introduction

This appendix presents an exploratory network analysis of the dataset used in case study 2. I decided to include this analysis as an appendix since it does not introduce innovative network measures compared to the previous two case studies (indeed the kind of analysis presented here is by far the most common way archaeologists have been applying network science techniques) and since its results are not directly comparable to the simulation output of the ABM. The results of this analysis and their implications for the aims of the case study are given in section 5.4.8.

In this case study I aim to better understand the significant differences in the distributions of tablewares in the Roman East. This will be done in part by analysing the distribution patterns of forms and wares. Although there are many different ways to explore such distributions (one of which is the exploratory data analysis performed in section 5.4), here I will restrict myself to exploring similarities and differences between the lists of sites that are part of these different distributions. Such an approach is by no means unique in exploratory network analysis, indeed it is by far the most common application of network techniques in our discipline (e.g. Golitko et al. 2012; Mills et al. 2013). Underlying these applications is an assumption that similarities and differences in artefact distribution patterns reveal something about the human behaviour that led to them. This is the assumption underlying this case study as well. A further assumption can be formulated that is necessary when aiming to compare the observed distribution patterns with the simulated ones produced by the agent-based model (section 5.5): artefacts with a similar distribution pattern have a higher probability of having been distributed through similar types of processes. Such an assumption will be tested as a hypothesis with the ABM but it cannot be tested through an exploratory network analysis. The latter merely aims to provide a representation of the observed tableware distribution patterns as network data and analysing the outcome, but making this assumption explicit is important when deciding on a suitable method for describing artefact similarity. In this section I will discuss a method for creating and analysing similarity networks. I will pay particular attention to the issue of the robustness of network analytical results: how sensitive are the results derived from network analysis techniques to our decisions of what constitutes the core of the network (in the sense of a subnetwork with particularly high similarity values)? Such an exploratory network analysis is a crucial step in seeing the pattern of interest in this case study through the lens of the assumption that similarities in tableware distribution patterns reveal similarities in their distribution mechanisms.

12.2.2. Similarity of forms' distributions: the Brainerd-Robinson coefficient

There are many ways of describing the similarity of artefact distributions (see Doran and Hodson 1975, 135-157; Shennan 1997, 222-234; for different approaches to similarity networks see Östborn and Gerding 2014). Many such measures can be derived from and manipulated as a matrix, as in figure 66a, which represents a list of tableware forms and their quantities attested at different sites. From this matrix we can derive two further matrices depending on the focus and aim of analysis: figure 66b shows sites and the number of forms they have in common with other sites; figure 66c shows tableware forms and the number of sites a pair of forms are both present at. In the exploratory network analysis I will focus on the latter representation since I aim to explore similarities and differences between forms, rather than sites. In figure 66c the similarity between forms is represented by the absence (values = 0) or presence (values > 0) of forms on sites. These values represent the number of sites a pair of forms is co-present on. Such presence-absence matrices are interesting for exploring the data without requiring further manipulation. A previous exploratory network analysis with the same ICRATES dataset was performed using exactly this approach (Brughmans and Poblome *in press*).

In this case study I will use another similarity measure commonly used in archaeology: the Brainerd-Robinson (BR) coefficient (Brainerd 1951; Robinson 1951; Cowgill 1990; Shennan 1997, 233-234). Rather than considering absolute numbers as in the presence/absence technique, the BR coefficient considers proportions. Although it is commonly used to compare the similarity of pairs of site assemblages, I will use it to compare the similarity between two forms' distributions. This measure as used in this case study therefore compares the proportion of all sherds of a pair of forms found in different sites. For every pair of forms, this measure sums up the absolute difference between proportions per site and subtracts this from 200 (the maximum possible difference for a pair of forms' distributions), providing a numerical similarity value between 0 and 200 where 0 indicates no similarity and 200 complete similarity, using the following equation:

$$\boldsymbol{S = 200 - \sum_{k=1}^{p} \left| P_{ik} - P_{jk} \right|}$$

Eq. 6. Brainerd-Robinson coefficient

where *P* is the percentage representation of site *k* in the distribution patterns of forms *i* and *j*. For example, the distributions given in figure 66a can be expressed as proportions (Fig. 67a) for which a BR similarity matrix representing the similarity of forms' distributions can be computed (Fig. 67b). In this case study, the BR coefficient is calculated using a script written in R by Matt Peeples (2011b).

Using the BR coefficient for forms' distributions rather than sites' distributions can only be justified if one understands what this equation does with the data and one interprets the results in light of this. Importantly, one must not forget that the BR coefficient used in this way does not use the proportion of a form in sites' assemblages at all. Most crucial is the difference between this approach and the presence/absence technique. Notice how in figure 66c forms A-B have a presence/absence value of 1 and B-C a value of 2, whilst in figure 67b forms A-B have a BR value of 106 and B-C a lower value of 94. Although one should be cautious not to over-interpret this small difference in the BR values, it is nevertheless clear that these two different approaches to comparing forms' distributions reveal very different things. Forms A-B are only co-present on one site but their distributions are marginally more similar than those of forms B-C. My motivation for using this measure in this rather unconventional way is threefold: firstly, because it does not merely suggest similarity based on co-presence but also takes into account possible similarities between the proportions of all sherds of pairs of forms co-present at a site; secondly, to compare differences in the results of exploratory network analyses of two different similarity measures applied to the same dataset (i.e. a comparison with the results published in Brughmans and Poblome *in press*); thirdly, as a methodological exercise to explore the potential of this unconventional approach for the study of similarity networks, since the BR coefficient is often used in the conventional way in exploratory network analysis by archaeologists (e.g. Golitko et al. 2012; Mills et al. 2013; Peeples 2011a).

One drawback of using the BR coefficient for creating similarity networks is that it does not calibrate the results based on the number of sherds attested for each form, rather it emphasizes the diversity of sites a form is attested on. This means that often forms for which few sherds are attested and which often have a very limited distribution are directly compared to forms for which a large number of sherds are included in the dataset, which leads to very high BR values for the former and possibly rather low values for the latter. However, the impact of this is immediately evident from the results and in this case study the number of sites a form is present at is considered more indicative of the wideness of its geographical distribution than the volume of that form attested at that site.

(a)

	Site A	Site B	Site C
Form A	2	0	0
Form B	10	3	6
Form C	0	8	4

(b)

	Site A	Site B	Site C
Site A	-	1	1
Site B	1	-	2
Site C	1	2	-

c)

	Form A	Form B	Form C
Form A	-	1	0
Form B	1	-	2
Form C	0	2	-

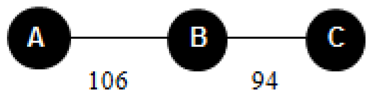
Fig. 66. Three matrices representing site assemblages: (a) tableware forms and their quantities at sites; (b) sites and the number of forms they have in common with other sites; (c) forms and the number of sites a pair of forms is co-present at.

(a)

	Site A	Site B	Site C
Form A	100%	0%	0%
Form B	53%	16%	31%
Form C	0%	67%	33%

(b)

	Form A	Form B	Form C
Form A	200	106	0
Form B	106	200	94
Form C	0	94	200



(c)

Fig. 67. (a) same matrix as figure 66a but showing percentages of forms' distributions rather than absolute numbers (e.g. 100% of all sherds of Form A are found on Site A); (b) Brainerd-Robinson coefficients of the same matrix, representing the similarity of forms' distributions; (c) network representation of (b).

12.2.3. Creating and exploring similarity networks

The matrix of BR values given in figure 67b can be represented as a network as in figure 67c. The diagonal of the matrix, representing perfect similarity of a forms' distribution to itself, is disregarded in the exploratory network analysis. In these one-mode similarity networks, forms are connected to each other if the similarity of their distributions as expressed by the BR coefficient is greater than zero. The meaning of the relationships of these networks should be derived from our discussion of the meaning of the BR values above: a relationship between a pair of tableware forms indicates they are co-present on at least one site and gives a measure for the similarity of the proportions of their respective total distributed volume at sites where they are co-present. It is important to note that these networks are undirected (since the BR similarity matrices are symmetric), which implies that questions concerning the directionality of flows of goods cannot be addressed with this approach to creating similarity measures. This is not considered problematic for this case study, since my interest here is merely in the similarity of distributions.

Tableware similarity networks are created per 25-year period. These networks will be explored as a whole, but also using different "thresholds" on the BR values (i.e. removing edges with a BR value lower than a certain threshold value; my motivation for using thresholds is discussed in more detail below). The networks will also be explored by grouping forms together according to their wares. This approach will provide insights into how similar or different the distribution patterns of wares are.

A range of network analysis measures will be applied to explore these networks, the technical definition of which has been introduced in previous chapters. It is important to note that in this case study I do not assume that indirect similarity (i.e. a path of two) allows for flows between two nodes. Therefore, no path-based measures (like closeness and betweenness centrality) will be used to explore the similarity networks: I will focus on nodes and their direct neighbours. A number of measures are used to identify the structural features of networks as a whole and compare them to each other: number of nodes, number of edges, connected components, network clustering coefficient, heterogeneity, average degree, and density. The number of nodes is the number of individual forms with a standard chronology that falls within a certain 25-year period that have been identified by archaeologists on sites and included in the ICRATES dataset. Connected components are groups of forms whose distributions are directly or indirectly similar to each other but completely dissimilar to forms in other groups. The network clustering coefficient is an indication of the tendency for forms to cluster together, which is suggestive of very similar distributions. The average degree is the average number of forms a given form has similarities in distribution patterns with. The density is a normalised version of average degree. Heterogeneity is an indication of the existence of forms with similarities to far more other forms in terms of distribution than other forms, i.e. hubs. Just two node-based measures will be used: node clustering coefficient and degree. The node clustering coefficient is here an indication of how similar all forms' distributions are that are directly similar to one form. The node degree is the number of forms that have a similar distribution to the node in question.

12.2.4. Network structure and sensitivity analysis

Introduction

This section explores the general structure of the created networks and performs a sensitivity analysis to see how this structure changes when one uses different thresholds of similarity values to create subnetworks. There are several reasons why it might be useful to cut off a chunk of the networks to analyse subnetworks: to identify and analyse nodes connected by the strongest relationships (i.e. the highest BR-values between tableware forms); to make a dense network more sparse for easier visual exploration; to test whether the analytical results one draws from a formal analysis of the whole network also hold true for subnetworks. Threshold values could either be selected with reference to a certain theory (as is the case for the thresholds used for the visibility analysis in case study 2 of this PhD project) or they could be determined arbitrarily using some quantitative approach. The latter approach has been used by many archaeological network analysts, either by selecting a minimum edge value (e.g. Golitko et al. 2012) or by comparing the distribution of similarity values of the observed network with those of simulated networks (e.g. Östborn and Gerding 2014; Peeples 2011a). I believe the strong variations in the number of forms per site (see Fig. 42) and the fragmentary nature of this dataset require one to first analyse the observed networks completely, and only analyse subnetworks derived with arbitrary threshold values to evaluate the robustness of the analytical results of the complete networks, i.e. to not attach any interpretative value on arbitrary thresholding but merely use it as an exploratory and sensitivity analysis tool. For this case study I do not believe it would be useful or even possible to argue for theoretical claims about what specific threshold value represents a strong similarity between forms' distributions and what not, other than the extreme BR-values of 0 and 200.

The 'complete networks' are networks per 25-year period of all forms and the similarity links between them of whatever strength. Below I discuss the distribution of the BR-values of all complete networks to argue for the selection of possibly useful (but nevertheless completely arbitrary) threshold values. Two different thresholds are then used to create subnetworks from these complete networks: a threshold on the average BR-value and on the average plus the standard deviation. I will also explore how the ranking of each node according to certain network measures (degree and clustering coefficient) changes with these changing thresholds. Such a sensitivity analysis is crucial since it will influence the interpretation of network patterns. These network patterns will subsequently be described per 25-year period for the periods between 25BC and 150AD.

Distribution of BR values

Histograms representing the distribution of BR values for each 25-year period summarise a wealth of information and can be used as an aid in deciding what threshold values to use (Figs. 68-69). Each histogram in these figures draws on all BR values for a single network (forms dated to a 25-year period), where the sum of the number of occurrences (y-axis) of a certain BR value (grouped per 10 on x-axis) is displayed as a bar. Figure 68 includes all of the values for each network and immediately shows the very heavy skew towards low BR values and the long tail. Indeed, for all networks a very high number of BR values are 0 or no more than 10. This is due to the very high number of sites that are included in this analysis (between 59 in 150-125BC and 162 in 1-25AD). As a result, the mean BR value is very low for all networks (solid vertical line in Figs. 68-69; table 24). This very high variation has a strong impact on the results of the standard deviation as well, which will be pushed quite high due to the existence of outliers (dashed vertical line in Figs. 68-69; table 25). The inter-quartile range could be considered more informative of typical observations in such cases (Shennan 1997, 44). However, due to these distributions' great spread the inter-quartile range itself is very low for many of the periods (table 24). Using the third quartile as a threshold value would (for most networks at least) not serve my purpose of bringing more visual clarity to the exploration of a dense network. A high threshold would be more suitable for that purpose, which is why I decided to use the following two threshold values: the mean and the mean + standard deviation (table 25). This decision implies that pairs of forms which are connected by BR values higher than the mean + standard deviation should be considered rather extreme outliers in the exploratory network analysis.

In figure 69 I decided to remove all BR values lower than 10, which reveals that for the networks up to 75AD the trend in the long tail is not very smooth either: there is more difference in the proportions of the variation for these networks and outliers are more frequent than for later period networks. The threshold values plotted on these distributions show a further distinction: the mean lies between 22 and 31 for networks up to 75AD and between 16 and 12 for later period networks. These differences might be indicative of the decrease in the diversity of ESC forms after 75AD. Before this date ESC has a high diversity of forms but many of these forms are co-present and have a limited distribution, which would result in high BR-values. It will be interesting to explore further down in this section whether this distinction is also reflected in some of the exploratory network measures.

These results suggest that the distributions of tableware forms are not very similar, given the low BR values. However, the results also suggest there is great variety among BR values larger than 10, and that a larger number of forms have a more similar distribution in the periods up to 75AD. This might be a result of the decrease in the distribution of ESA around this time. In the rest of this section I will explore the variance in BR values in more detail by using threshold values, and grouping forms according to wares.

Table 49: summary statistics of Brainerd-Robinson coefficients for the complete networks per period.

	150-125BC	**125-100BC**	**100-75BC**	**75-50BC**	**50-25BC**	**25-1BC**	**1-25AD**	**25-50AD**	**50-75AD**	**75-100AD**	**100-125AD**	**125-150AD**	**150-175AD**	**175-200AD**
Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0	0
First Quartile	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Median	0	6	0	0	0	0	0	20	0	0	0	0	0	0
Mean	22	30	22	31	27	26	23	28	23	16	15	14	14	12
Third Quartile	33	59	30	40	23	22	15	20	18	13	12	10	1	1
Maximum	167	200	200	200	200	200	200	200	200	200	200	200	200	200

Table 50: summary statistics of Brainerd-Robinson coefficients of complete networks per period. The mean and mean + standard deviation are suggested as thresholds for exploring the networks.

	150-125BC	**125-100BC**	**100-75BC**	**75-50BC**	**50-25BC**	**25-1BC**	**1-25AD**	**25-50AD**	**50-75AD**	**75-100AD**	**100-125AD**	**125-150AD**	**150-175AD**	**175-200AD**
Mean	22	30	22	31	27	26	23	28	23	16	15	14	14	12
St. Dev.	37	44	39	54	52	51	49	54	48	37	33	34	37	31
Mean + St. Dev.	59	74	61	85	78	77	73	82	71	54	48	48	51	43

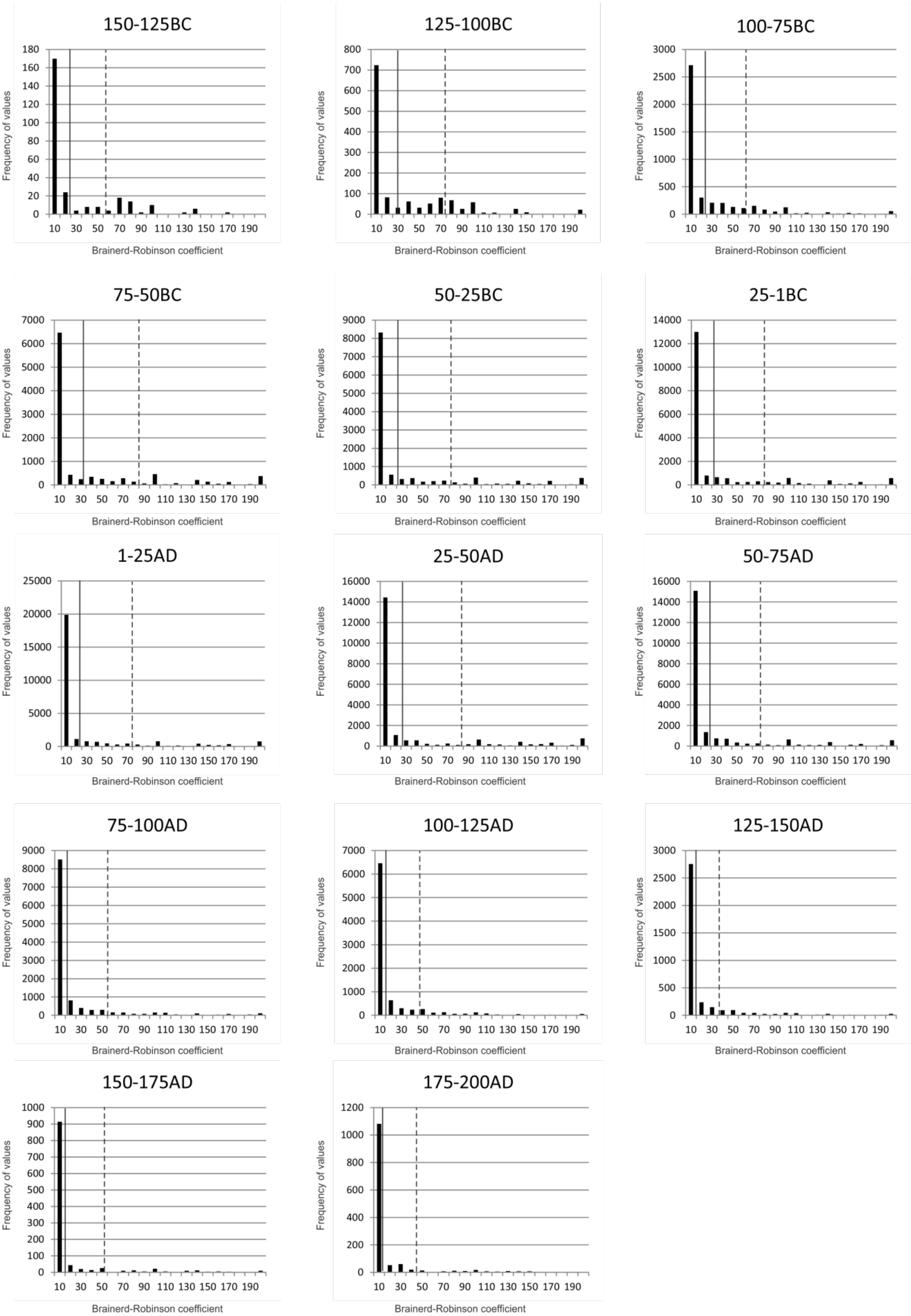


Fig. 68: distribution of Brainerd-Robinson values of form-form similarity matrices per period; full line: mean value; dashed line: mean + standard deviation.

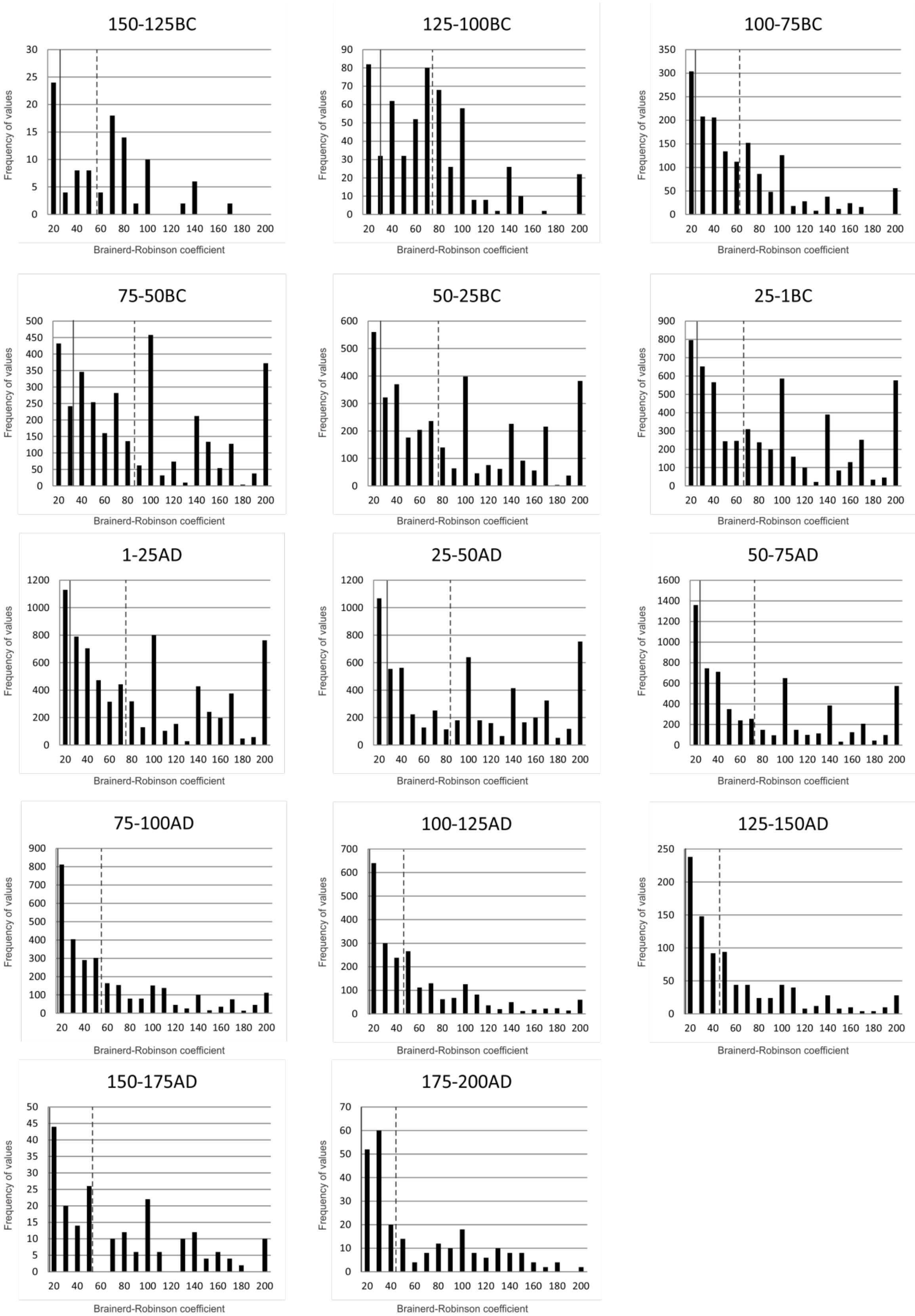


Fig. 69: distribution of Brainerd-Robinson values of form-form similarity matrices per period, excluding values between 0 and 10; full line: mean value; dashed line: mean + standard deviation.

Global network measures

Complete network: the network density remains rather high throughout time, although it is highest in the first few periods (Table 51; Fig. 70). This is possibly due to the high similarity in distributions of ESC forms and of ESA forms. Heterogeneity, on the other hand, increases through time, suggesting that forms with a similar distribution to many other forms and skewed degree distributions occur more frequently in the later periods. This suggests that the decreasing density creates a pattern of well-connected nodes and less-well-connected nodes. The groups of forms with a similar distribution become smaller over time. The clustering coefficient remains very high throughout all periods. This suggests that forms in general have a distribution that shows similarities with many other forms' distributions.

Table 51. Global network measures for the complete networks per 25-year period.

	150-125 BC	125-100 BC	100-75BC	75-50BC	50-25BC	25-1BC	1-25AD	25-50AD	50-75AD	75-100AD	100-125AD	125-150AD	150-175AD	175-200AD
Nodes	17	37	66	99	109	137	166	143	146	107	91	58	33	36
Nodes (incl. isolates)	17	37	66	100	111	137	167	146	149	111	97	64	36	39
Connected components	1	1	1	1	1	1	1	1	1	2	1	1	1	1
Average degree	8	20.4	31.3	42.8	45.2	53.9	58.7	53.3	54.7	39.2	34.7	21.5	9.8	10.1
Clustering coefficient	0.778	0.848	0.842	0.844	0.836	0.833	0.813	0.816	0.832	0.791	0.812	0.817	0.755	0.778
Density	0.5	0.568	0.482	0.437	0.418	0.396	0.356	0.376	0.377	0.369	0.386	0.377	0.305	0.287
Heterogeneity	0.427	0.343	0.443	0.454	0.522	0.541	0.553	0.537	0.537	0.566	0.575	0.561	0.64	0.589

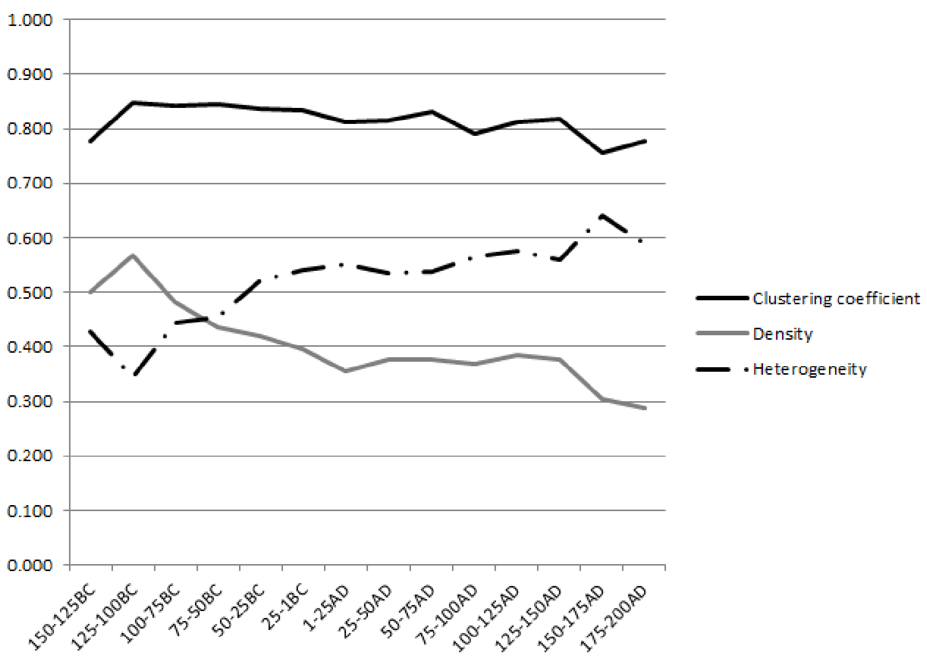


Fig. 70. Global network measures per 25-year period for the complete network.

Threshold mean: density is much lower when thresholding on the mean BR value (Table 52; Fig. 71). The density of networks decreases over time. The heterogeneity shows a similar increasing pattern as in the complete network. However, now the values are slightly higher, and there is a peak with many hubs in 25-50AD. This suggests that this threshold emphasizes differences in node degrees more. It seems that hubs are a strong pattern in these networks. The clustering coefficient is invariably very high. Nodes with this threshold are still very much clustered with other nodes with similar distributions.

Table 52. Global network measures for the networks per 25-year period with a threshold on the mean BR value.

	150-125BC	125-100BC	100-75BC	75-50BC	50-25BC	25-1BC	1-25AD	25-50AD	50-75AD	75-100AD	100-125AD	125-150AD	150-175AD	175-200AD
Nodes	17	37	64	94	105	132	163	141	143	103	90	57	31	36
Nodes (incl. isolates)	17	37	66	100	111	137	167	146	149	111	97	64	36	39
Connected components	2	1	1	1	2	1	2	3	2	2	1	1	1	1
Average degree	4.5	12.3	19.2	29.2	27.5	33.7	37.5	33.7	34	24	23	14.6	6.6	6.9
Clustering coefficient	0.659	0.856	0.792	0.832	0.788	0.827	0.791	0.788	0.784	0.751	0.771	0.771	0.777	0.726
Density	0.279	0.342	0.304	0.314	0.264	0.257	0.232	0.241	0.239	0.235	0.259	0.261	0.222	0.197
Heterogeneity	0.442	0.422	0.461	0.495	0.601	0.644	0.662	0.721	0.64	0.601	0.623	0.594	0.639	0.544

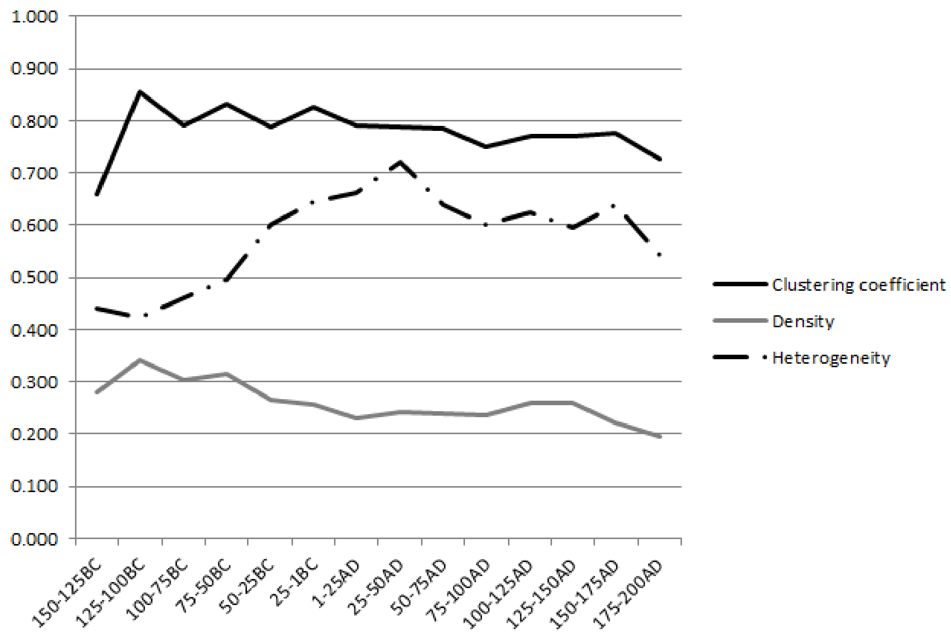


Fig. 71. Global network measures per 25-year period for the network with a threshold on the mean similarity value.

Threshold mean plus standard deviation: the density is again lower with this threshold, which is to be expected (Table 53; Fig. 72). However, the general trend of decreasing density over time remains, suggesting this is a robust pattern. The heterogeneity measure is very interesting. Although it still shows a similar trend as before, this is now much exaggerated: this threshold reveals a much stronger difference between nodes in terms of degree as time moves on. In particular between 50BC and 100AD when hubs are most common. This seems to be caused by one or two nodes that bridge dense cliques, mainly between ESC and a combination of ESA/ESB/ESD. The clustering coefficient shows a similar trend as with other thresholds, although it is overall slightly lower.

Table 53. Global network measures for the networks per 25-year period with a threshold on the mean + standard deviation BR value.

	150-125BC	**125-100BC**	**100-75BC**	**75-50BC**	**50-25BC**	**25-1BC**	**1-25AD**	**25-50AD**	**50-75AD**	**75-100AD**	**100-125AD**	**125-150AD**	**150-175AD**	**175-200AD**
Nodes	14	33	61	82	92	116	144	129	133	96	86	54	27	31
Nodes (incl. isolates)	17	37	66	100	111	137	167	146	149	111	97	64	36	39
Connected components	2	4	2	3	7	7	8	11	8	4	2	3	5	5
Average degree	3.9	6.4	9.9	18.9	18.3	23.1	25.1	24.4	20.4	12	11.3	6.9	3.9	3.8
Clustering coefficient	0.833	0.601	0.713	0.807	0.806	0.822	0.799	0.828	0.802	0.696	0.702	0.712	0.8	0.718
Density	0.297	0.201	0.164	0.233	0.201	0.201	0.175	0.191	0.154	0.126	0.134	0.129	0.148	0.127
Heterogeneity	0.414	0.572	0.515	0.771	0.817	0.877	0.962	0.901	0.946	0.758	0.597	0.59	0.688	0.44

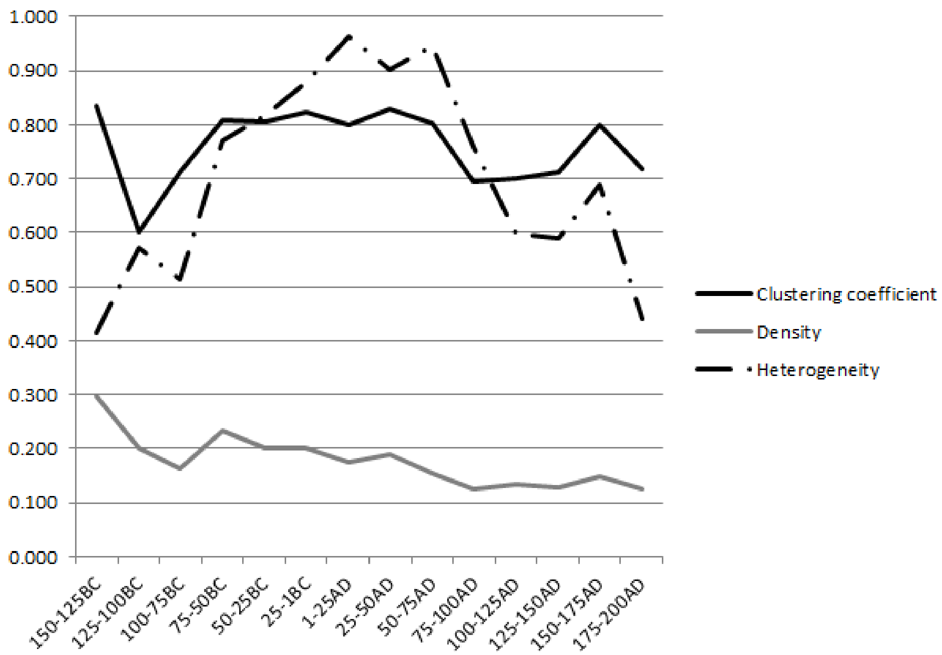


Fig. 72. Global network measures per 25-year period for networks with a threshold on the mean + standard deviation value.

Total number of nodes (forms): the networks of BR similarity values described above do not include isolated forms, i.e. forms whose distribution similarity with other forms' distributions as described by the BR coefficient is 0. However, only a marginal number of forms per period are not included in the networks (min 0, max 6; Table 51). The largest number of not included forms are in the second part of the period of study (25AD onwards) and the effect is stronger for the last four periods, which have fewer nodes: more forms in these periods have no similarity to other forms' distributions. The period between 50BC and 100AD has the highest number of forms (Figs. 73-75). However, the thresholds do not affect the number of forms that much. A threshold on the mean BR value only decreases the number of forms for a few periods, whilst the mean + standard deviation threshold decreases the number of forms for all periods only marginally.

The biggest difference is for the periods between 25BC and 75AD, and for ESA forms in particular. Up to 75AD and after 150AD the number of ESA forms is most affected by using a threshold on the mean + standard deviation value. This suggests that although most wares always have significantly similar distributions to some forms (mostly of the same ware), a large proportion of ESA forms have an overall extremely low similarity to all forms. This is not the case in the period 100-150AD when ESA forms show significant similarity with at least one other form, whilst a large proportion of ESD forms shows no significant similarity to any other forms.

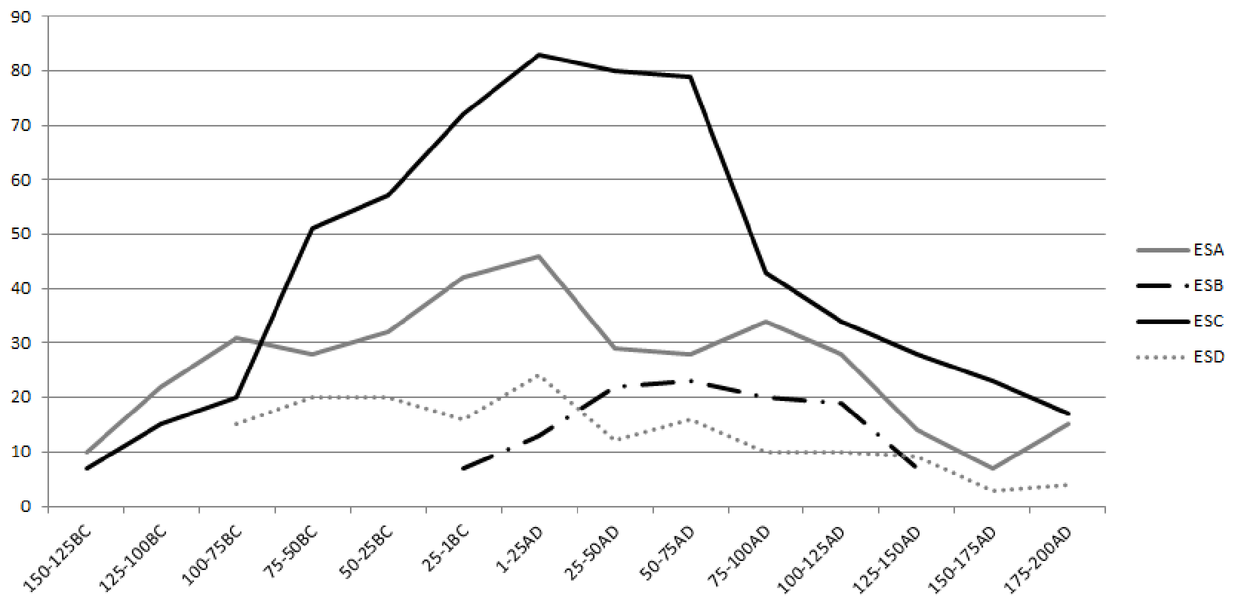


Fig. 73. Number of nodes (forms) per ware for the total network.

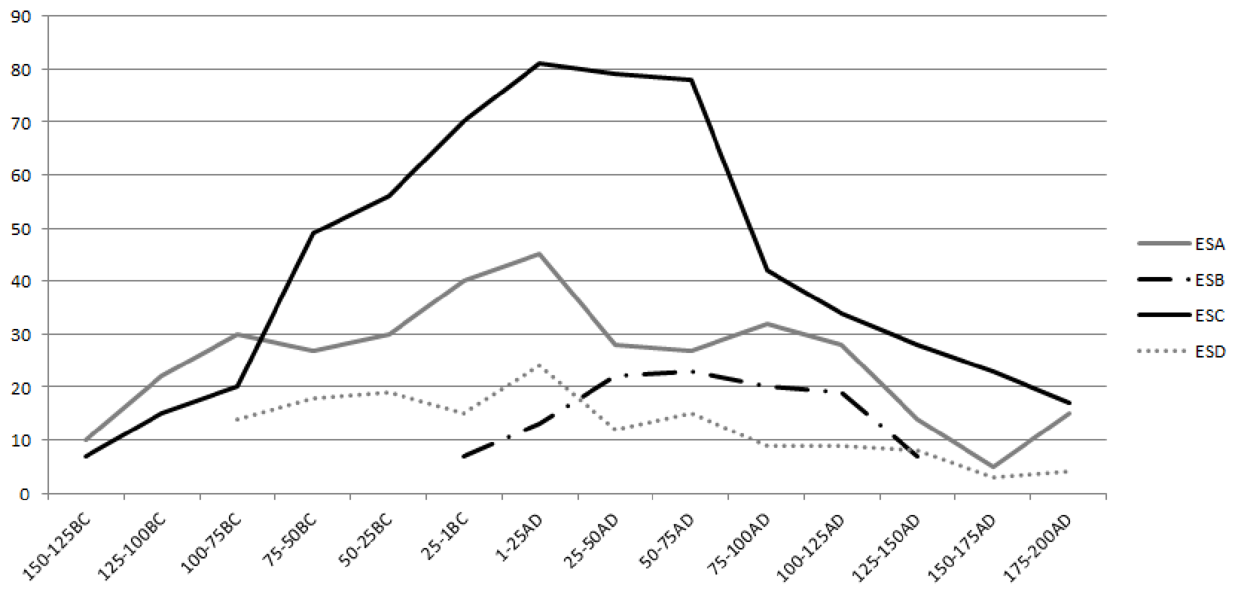


Fig. 74. Number of nodes (forms) per ware for the network with a threshold on the mean similarity value.

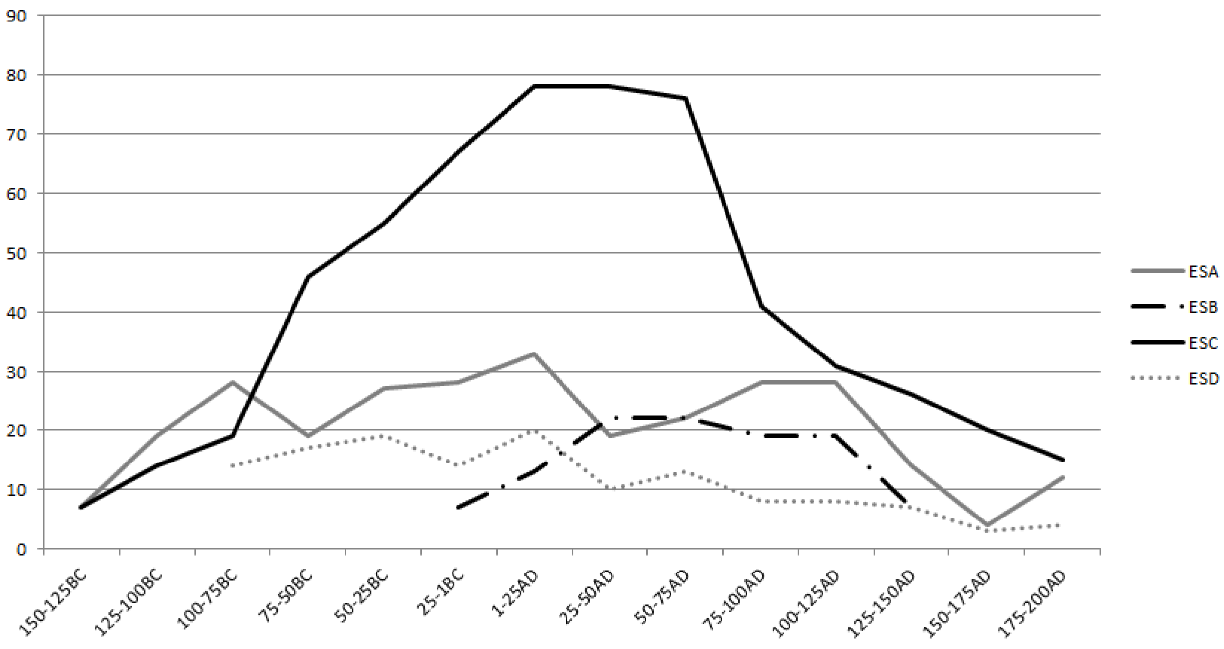


Fig. 75. Number of nodes (forms) per ware for the network with a threshold on the mean + standard deviation similarity value

Local network measures: ranking clustering coefficient and degree

The previous section discussed the sensitivity of global network measures to using thresholds, and this next section focuses on the impact it has on local network measures' results by tracing the changes in the ranking of node clustering coefficient and degree. The ranking shown in figures 76 and 77 was created as follows: node clustering coefficient and degree measures were calculated; the results were ordered from high to low and nodes were given a rank (equal values were given an equal rank); the number of places each node changed in the ranking was counted as it was traced from the complete network to that with a threshold on the mean, and further to that with a threshold on the mean + standard deviation (please note that this was only done for nodes that are present in the mean + standard deviation network); in order to be able to compare the proportional change of node rankings across different periods the results were normalized by dividing them by the maximum possible number of changes in ranking a node could undergo in going from the complete network to the network with a threshold on the mean + standard deviation (and multiplied by 100 to get a percentage); boxplots of these results for clustering coefficient and for degree were created.

Clustering coefficient (Fig. 76): the networks of all periods show strong similarities in their sensitivity to changing thresholds. The interquartile range is almost in all cases limited from

0% to 35/45%, whilst some few nodes see up to 80% change in ranking. In the periods 125-100BC and 1-25AD one node sees 100% change (ESC form M200ST27 and ESD form EAAP28 respectively). The period 25-50AD knows the highest range of change. The networks of period 150-125BC are very different from all others because of the limited number of nodes. On this network changing thresholds have extremely little effect. This sensitivity analysis therefore indicates that changes in the clustering coefficient ranking of nodes is common. However, nodes only rarely change their ranking dramatically by over 50%. The vast majority of nodes remain within 50% of their ranking position.

Degree (Fig. 77): the networks of all periods show strong similarities in their sensitivity to changing thresholds. The interquartile range is almost in all cases limited from 15% to 35%, whilst some nodes see change of up to 80% in the first period only. A few outliers are identified with only 50 or 55% change. Two periods show a different pattern. In the period 125-100BC change in degree is slightly higher, with an interquartile range of 15-45%. The most dramatic difference is seen in the networks of 150-125BC, with most nodes changing their rank between 0 and 65%. This is no doubt the result of the small number of nodes in this network. This sensitivity analysis therefore indicates that changes in the degree rankings are common. Almost every single node is affected by changing thresholds. However, the change in their ranking is overall very limited, rarely do nodes change more than 35% in the rankings. An exception to this is the period 150-125BC, which has significant change but also has a very low number of nodes.

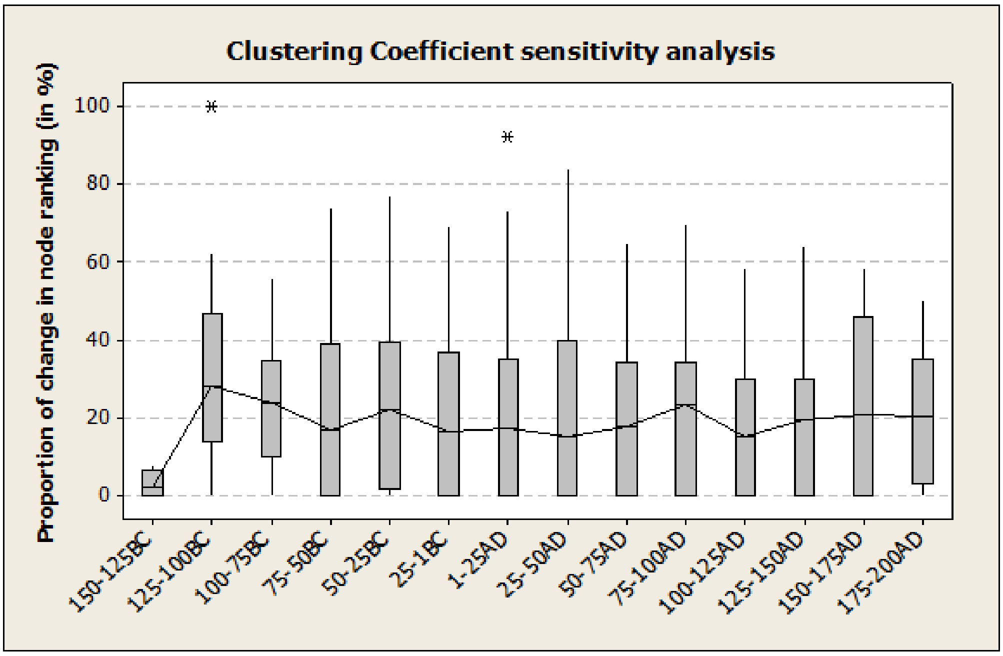


Fig. 76. Boxplot of the proportion of change in node ranking of the clustering coefficient.

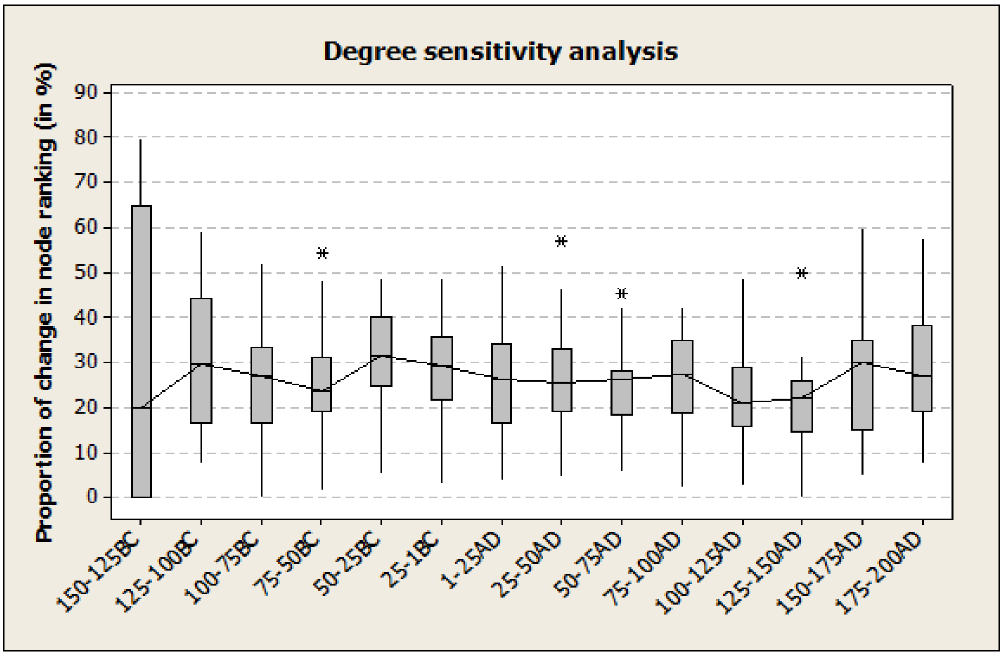


Fig. 77. Boxplot of the proportion of change in node ranking of the degree.

Description similarity networks per 25-year period

In this section the similarity networks are described in more detail for the 25-year period from 25BC to 150AD in which four tablewares circulated on a large scale in the Eastern Mediterranean. This description will be performed on the level of individual forms as well as for each ware, and will take the results of the sensitivity analysis of the impact of the suggested threshold values into account. In this description ‘very strong similarity’ refers to a

BR value higher than the mean + standard deviation; 'strong similarity' refers to a BR value higher than the mean; a 'homogeneous distribution' refers to forms of the same ware which are similarly distributed among sites.

25-1BC

ESA: has a homogeneous distribution but much less so than in previous periods. A number of forms' distributions are very similar to each other (EAA11, EAA12, EAA12/32, EAA13A, EAA13B, EAA21/22, EAA22/43, EAA22A, EAA22A-B, EAA22B, EAA2-3, EAA24, EAA24-25, EAA25, EAA26-27, EAA26A, EAA26A-D, EAA26B, EAA26C, EAA26D, EAA27, EAA28, EAA28-30, EAA29, EAA3, EAA42, EAA4A, EAA4A-B, EAA4B, EAA5A, EAA5A-B, EAA5B, EAA7, EAA8). All other forms' distributions are not very similar to each other. Some of the latter (EAA21/22, EAA22/43, EAA22A, EAA22A-B, EAA22B, EAA2-3) show a strong similarity with some ESB forms (EAA16, EAA21). Some other forms' distributions are more similar to ESD than to ESA or ESB (EAA102, EAA22B, EAA26-27, EAA26A-D). Of these only EAA22B is very similar to ESA forms. EAA26-27 is similar to many ESD forms, because this form is only present at Paphos. ESA forms' distributions are dissimilar to those of ESC forms.
ESB: not very internally homogeneous. One form (EAA21) shows strong similarity to ESA forms. Two forms (EAA2, EAA22) are very similar to ESC forms due to their presence on Assos and Ephesos.
ESC: overall very homogeneous. A few forms are less similar due to limited distribution, but not more similar to ESA or ESD. No forms show strong similarity with ESA. M-SSu25 is very similar to many ESD forms, because it is present on Paphos, the only sites with evidence of a number of ESD forms.
ESD: a group of ESD forms is pretty homogeneous, mostly because some are only present on Paphos. Three forms are dissimilar to all others (EAAP18A-B, EAAP33-34, EEP44-X45). Two forms are more similar to ESA (EAAP37A-B, EAAP37B).

1-25AD

ESA: not very homogeneous, less so than in the previous period, but still showing slightly more similarity of ESA forms' distributions than with forms of other wares. A number of forms are very similar to each other (EAA12, EAA42, EAA43, EAA44, EAA45, EAA46, EAA47, EAA4A, EAA4A-B, EAA4B, EAA5A, EAA5A-B, EAA5B; these are less than in the previous period, although there is some overlap). All other forms are not very similar to each other. Some of the latter (EAA12/32, EAA21/22, EAA30/33-34, EAA33/36) show a strong similarity with some ESB forms (EAA21, EAA30-31). Some other forms are more similar to ESD than to ESA or ESB (EAA26-27, EAA26A-D, EAA45). Of these only EAA45 is very similar to ESA forms. EAA26-27 is similar to many ESD forms, because this form is only present at Paphos. ESA forms are dissimilar to ESC forms. ESA forms show most similarity with ESB and ESD forms, not ESC.
ESB: rather homogeneous distributions of forms. One form (EAA30-31) shows strong similarity to ESA forms. Two forms (EAA2, EAA22) are very similar to ESC forms due to their presence on Assos and Ephesos. ESB is very dissimilar from ESD.

ESC: overall very homogeneous. A few forms are less similar due to limited distribution, but not more similar to ESA or ESD. No forms show strong similarity with ESA. M-SSu25 is very similar to many ESD forms, because it is present on Paphos, the only sites with evidence of a few ESD forms. Two ESB forms (EAA2, EAA22) are very similar to ESC forms due to their presence on Assos and Ephesos.
ESD: less homogeneous internally than in the previous period. A group of ESD forms is pretty homogeneous, mostly because some are only present on Paphos. Three forms are dissimilar to everything (EAAP44-X45, EAAP4B/P6, EAAP5/P6). ESD forms show much stronger similarity to ESA forms than in the previous period. One form is more similar to ESA (EAAP23B).

25-50AD

ESA: not very homogeneous internally, even less than previous period, but still slightly more homogeneous internally than with other wares. ESA forms show most similarity with ESB and ESD forms, not ESC. But that similarity is also less striking than in the previous period. There are no forms very similar with ESD. One ESB form (EAA30-31) shows strong similarity to ESA forms, because of its presence in Jerusalem. Two ESC forms (EAAL6, EAAL9A) show strong similarity to some ESA, because they are virtually exclusive to Jerusalem.
ESB: very internally homogeneous. One ESB form (EAA30-31) shows strong similarity to ESA forms, because of its presence in Jerusalem. One ESC form (M-ST22) is very similar to ESB, it is present on Assos and Ephesos. ESB is rather dissimilar from ESD.
ESC: very internally homogeneous. One ESC form (M-ST22) is very similar to ESB, it is present on Assos and Ephesos. Two ESC forms (EAAL6, EAAL9A) show strong similarity to some ESA, because they are virtually exclusive to Jerusalem. ESC and ESD are very dissimilar, with the exception of ESC form EAAL15 which is most similar to ESD and not at all to ESC, because it is only present at Lepcis Magna.
ESD: not very homogeneous, similar to previous period. ESD forms show much less similarity to ESA than in the previous period. ESC and ESD are very dissimilar, with the exception of ESC form EAAL15 which is most similar to ESD and not at all to ESC, because it is only present at Lepcis Magna.

50-75AD

ESA becomes slightly more internally similar. Similarities to ESD increase slightly. In particular forms EAA39, EAA41, EAA45 due to their presence on Pahos and Amathous. Similarities with ESB increase slightly, especially form EAA35/40 which is only recorded for Athens. ESA is still very dissimilar from ESC.
ESB: only slightly less internally homogeneous than in the previous period. Shows slightly more similarity to ESD and ESA than in previous period. In particular forms 58 early, 60, 60 early, 64/65, 65, 70 and 70 early are similar to ESA form EAA35/40 (only found in Athens) because they are all found together in Athens. Almost all ESB forms show strong similarity with one or more ESC forms. Just one form (30-32) shows strong similarity with ESD because it is only found on Berenice.
ESC: ESC becomes slightly less internally homogeneous. It is still very dissimilar from ESA and ESD. Some forms show similarities with ESB (EAAL20, M-SN11a, M-SN33a, M-SSa27a, M-ST22, M-ST24b), even though they have a wide distribution.

ESD becomes slightly more similar to ESB and ESA, and it is not very internally homogeneous. EAAP5/P6_ESD is similar to ESB because it is only found on Berenice. EAAP10 and EAAP28 are similar to ESA, both are co-present on Berenice, Amathous, Corinth, Panayia, Paphos, Tarsos. They have a wide distribution and are attested at sites with a high diversity.

75-100AD

ESA becomes slightly less internally homogeneous. Similarities to ESD increase slightly again, in particular for ESA forms 111, 36, 39, 40A-C, 40C, 41, 51, 53, 54, 62. Similarities to ESB increase again, in particular with ESA forms 35/40, 37A-B, 40A-C, 65. These are co-present at Athens and all but 35/40 have a wide distribution. Although overall ESA distribution is quite different from ESC, a few ESA forms (117, 33/36, 34/37, 38-39, 40A-C, 40C) show significant similarities with ESC.
ESB: almost all ESB forms are very similar to just a few ESA forms (35/40, 40A-C, 65) and to just a few ESC (M-SN33a, M-SN33d, M-SSa27a, M-SSa27a-c), few are very similar to ESD.
ESC becomes even less internally homogeneous. It remains dissimilar to ESD. It becomes slightly more similar to ESA (through EAAL6, EAAL9A, M-SB9, M-SSa-27a-c, M-ST8a) and ESB (through M-SN11a, M-SN33a, M-SN33d, M-SSa27a, M-SSa27a-c, M-ST24b).
ESD: almost all ESD forms (with the exception of p10/p11, p11/p12, p19 which are very dissimilar to ESD and show way more similarity to ESB) have strong similarities with ESA but less so with ESB (only p28 which has a wide distribution on sites with diverse assemblages like Jerusalem, Athens, Corinth) and extremely dissimilar to ESC.

100-125AD

ESA become slightly less internally homogeneous. It remains similar to ESB (through 35/40, 40A-C, 40C, 58/60, 65 (mostly similar to previous period)) and ESD (through 40A-C, 40C, 51, 52, 53, 54, 62 (mostly similar to previous period)), and shows slight similarity to some ESC forms, but not much more than in the previous period.
ESB: similarity of most ESB with just a few ESA (35/40, 40A-C, 58/60, 65 (similar to previous period)) and ESC forms. Some ESB (62b, 74a, 77) are particularly similar to ESA. Like in the previous period still quite dissimilar to ESD.
ESC (LRP2, LRP3, M-SN33d, M-SSa27a, M-SSa27a-c) is very similar to ESB, but less so to ESA (only ESA 40A-C, 40C are similar, present at Athens and Gortyn). It is very dissimilar to ESD.
ESD is very dissimilar to ESC and rather dissimilar to ESB. Most ESD (with the exception of p10/p11, p11/p12, p30a-b, found on sites with extremely limited assemblages) is similar to a few ESA forms (40C, 51, 52, 62).

125-150AD

ESA is less internally homogeneous. It becomes less similar to ESB and ESD and is completely dissimilar to ESC. Only one ESB form (77, widely distributed but highest proportion in Athens) is significantly similar to one ESA form (58/60, only found in Athens). Many ESD forms are similar to ESA 52 and 54.

ESB becomes internally very homogeneous. As in the previous period, many ESB forms are similar to ESC forms (LRP1, M-SN33d, M-SSa27a, M-SSa27a-c). As in the previous periods, ESB is not extremely similar to ESD.

ESC becomes less internally homogeneous. It is very similar to ESB, but not at all to ESD and ESA.

ESD is dissimilar to ESB and ESC but most forms are very similar to two ESA forms (52, 54). As in the previous period, with the exception of p10/p11, p11/p12, p30a-b, found on sites with extremely limited assemblages.

12.3. Agent-based model code

This appendix presents the Netlogo code of the ABM developed for case study 3. The code is highly annotated to explain how it works. Annotations are indicated by a semicolon, i.e. everything after the symbol ‘;’ should be considered an explanation of the code following it. The model itself is available as an electronic supplement. In order to run the ABM you will need to download and install Netlogo[36]. When opening the model in Netlogo the interface appears where the model can be initialised, run and where variables can be changed (Fig. 78). The ‘info’ tab holds concise information on how the model was constructed and how to use it. The ‘code’ tab contains the ABMs code written in the Netlogo language. This code is also included on the following pages in this section of the appendix.

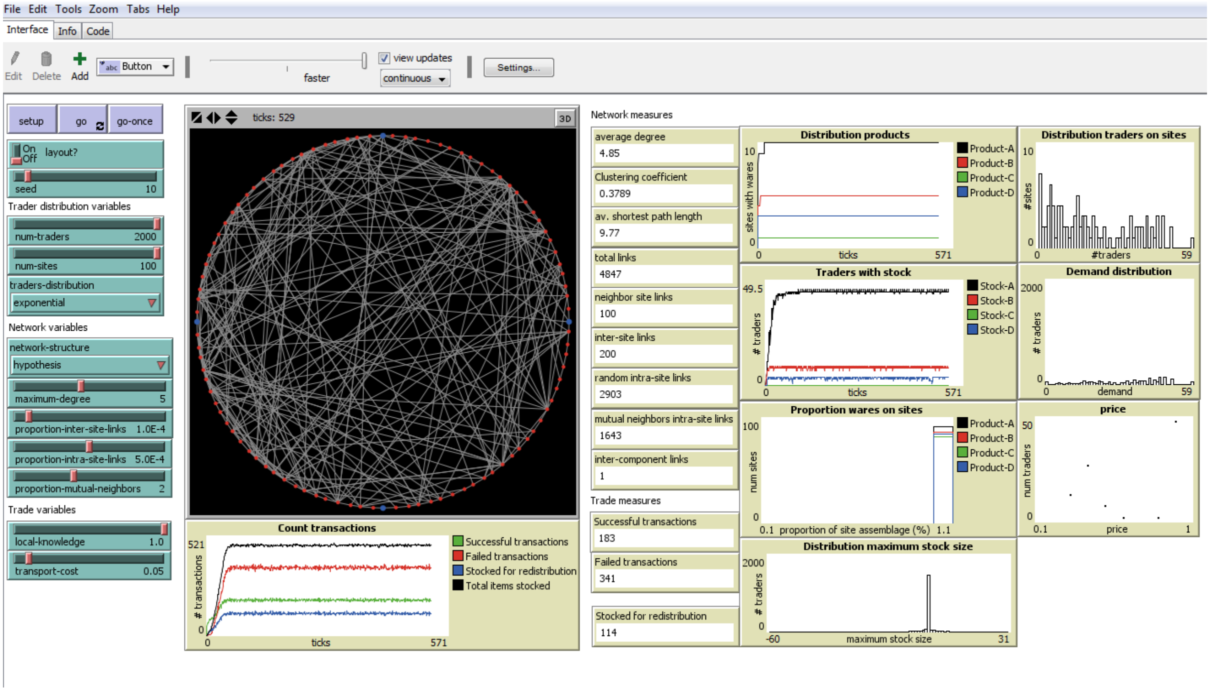


Fig. 78. Interface of the ABM created for case study 3. Sliders allow for changing the variable settings. The ‘setup’ button initialises the model, the ‘go’ button runs the model with the selected settings.

[36] Netlogo can be downloaded via this URL: https://ccl.northwestern.edu/netlogo/download.shtml (accessed 24-06-2014).

```
extensions [ network nw ]
; two network extensions are used in this model.
; The 'network' extension comes packaged with Netlogo 5.0.5
; the 'nw' extension needs to be downloaded from the Netlogo website and added to the
'extensions' folder: https://github.com/NetLogo/NetLogo/wiki/Extensions

breed [ traders trader ]
; the agents in this model who are positioned on sites, are connected by commercial links, and
trade products with each other over these links

breed [ sites site ]
; the sites, representing marketplaces on which traders are based, meet, and trade

globals
[
  av-degree
; reporter, parameter used to calculate and report the average degree of the network during
the setup procedure
  clustering-coefficient
; reporter, parameter used for calculating and reporting the clustering coefficient in the
setup procedure
  neighbor-site-links
; reporter, parameter used for reporting the number of links between neighboring sites in the
circular layout, step 1 of the connect-traders procedure
  random-inter-site-links
; reporter, parameter used for reporting the number of random inter site links in a setup of
the model, step 2 of the connect-traders procedure
  random-intra-site-links
; reporter, parameter used for reporting the number of random intra site links in a setup of
the model, step 3 of the connect-traders procedure
  mutual-neighbors-intra-site-links
; reporter, parameter used for reporting the number of mutual neighbors linked within sites in
a setup of the model, step 4 of the connect-traders procedure
  random-component-links
; reporter, parameter used for reporting the number of random links added to join components,
step 5 of the connect-traders procedure
  counter-trade
; reporter, parameter used in the trade procedures to keep track of how many transactions have
already been attempted
  counter-transaction
; reporter, parameter used in the trade procedures to count the number of successful
transactions
  counter-stock-redist
; reporter, parameter used in the trade procedures to count the number of items put in stock
with redistribution in mind per tick
  counter-no-transaction
; reporter, parameter used in the trade procedures to count the number of failed transactions
per tick
]

sites-own
[
```

```
  target-distribution
; counter parameter used in the setup procedure to identify the number of traders that need to
be present at a site under the selection distribution hypothesis
  producer-A ; parameter used to indicate the production centre of product-A
  producer-B ; parameter used to indicate the production centre of product-B
  producer-C ; parameter used to indicate the production centre of product-C
  producer-D ; parameter used to indicate the production centre of product-D
  volume-A ; the volume of product A deposited at a site
  volume-B ; the volume of product B deposited at a site
  volume-C ; the volume of product C deposited at a site
  volume-D ; the volume of product D deposited at a site
  prop-A ; reporter, parameter used to report the proportion of product A deposited on a site
  prop-B ; reporter, parameter used to report the proportion of product B deposited on a site
  prop-C ; reporter, parameter used to report the proportion of product C deposited on a site
  prop-D ; reporter, parameter used to report the proportion of product D deposited on a site
]

traders-own
[
  comp
; parameter used to identify which connected component a trader is part of. Used when creating
the initial trade network in the setup procedure
  trader-clustering-coefficient
; parameter used for calculating the clustering coefficient of nodes
  moved?
; parameter used to keep track of the traders who have already moved to a site in the setup
phase
  product-A ; the amount of product A an agent possesses
  product-B ; the amount of product B an agent possesses
  product-C ; the amount of product C an agent possesses
  product-D ; the amount of product D an agent possesses
  stock-A
; the amount of product A a trader saves for possible redistribution in the next tick
  stock-B
; the amount of product A a trader saves for possible redistribution in the next tick
  stock-C
; the amount of product A a trader saves for possible redistribution in the next tick
  stock-D
; the amount of product A a trader saves for possible redistribution in the next tick
  price
; the price a traders believes one item of the product is worth in his part of the network,
float between 0 and 1
  demand ; the demand a trader believes he can supply products for
  average-demand ; parameter used to calculate the average demand of a trader's link neighbors
  site-number
; parameter used to keep track of what site this trader is located at, used to reposition
traders after layout
  transport-fee ; the fee a seller has to pay when a buyer is located on another market
  known-traders
; parameter used to store the agentset a trader receives commercial information from in each
tick
```

```
  maximum-stock-size
; in each tick the number of items a trader is willing to obtain over his own demand for
redistribution in the next tick
]

;;;;;;;;;;;;;;;;;;;;;
;;;;;;; SETUP ;;;;;;;
;;;;;;;;;;;;;;;;;;;;;

to setup
  clear-all
  random-seed seed
  set-default-shape traders "person"
  set-default-shape sites "circle"
  create-sites num-sites
  create-traders num-traders
  layout-circle ( sort sites ) max-pxcor - 1 ; arrange sites according to a circular layout
  ask sites
  [
    set size 0.8
    set color red
  ]
  ask traders
  [
    set size 0.3
; size of traders should be smaller than size of sites, since 'traders-here' is used which is
very sensitive to patch and node size
    set color blue
    set moved? false
; ensure the model knows at the setup stage that none of the traders have moved to any of the
sites yet
  ]
  distribute-traders-on-sites
; move all traders to a site following a uniform, normal, or exponential distribution
  ask traders
  [
    set site-number [who] of one-of sites-here
; make a note of which site a trader is located at (needed to reposition traders after layout)
    set demand 0 ; initialise demand per trader by setting it to 0
   ]
  nw:set-context traders links
; set the context the NW extension procedures will be working in
  connect-traders
; connect traders in five steps: 1) connect a pair of traders on neighboring sites
; 2) connect a proportion of randomly selected node pairs;
; 3) connect randomly selected node pairs on the same site
; 4) connect a proportion of randomly selected pairs with a mutual neighbor (loop 3 and 4
until the average degree is reached)
; 5) add links within sites until the network consists of one connected component
; OR create a random network with the same number of links
  select-production-sites
; select which sites are production centres of each product. When possible, these production
```

```
sites will be evenly spaced among the circular layout of sites
  reset-ticks
end

;;;;;;;;;;;;;;;;;;;
;;;;;;;;;; GO ;;;;;
;;;;;;;;;;;;;;;;;;;

to go
  ifelse layout?
  [
; if layout is turned on, traders will be positioned following a spring-embedded layout
algorithm.
; No other procedures take place and ticks don't increase, to allow for the experiment to
continue from the point when layout was turned on
    layout
  ]
  [
    reposition-traders
; if layout was turned on before, then move the traders back to their sites.
    determine-demand
; each trader increases its demand by one (as long as it is lower than the number of traders
at the site).
    discard-part-of-stock
; a fixed proportion of a trader's stock from last tick is deposited on its site, the rest can
be traded again this tick
    produce
; each trader at a production site obtains X newly produced items of the locally produced
product if its total number of items of all products is less or equal than its demand.
    trade-and-consume
; a single transaction will occur for each item of each product: after a successful
transaction the item is deposited on the buyers' site, when the seller is not connected to any
potential buyers it puts the item in its stock, when the buyer notices the average demand in
its personal network is higher than its own demand it puts the item in its stock.
    tick
  ]
end

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;; SETUP PROCEDURES ;;;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

to select-production-sites
; select tableware production sites and space them evenly accross the circular layout
  let spacing round ( num-sites / 4 )
; identify the spacing between production sites
; make each of these sites a producer of a certain product
  ask site 0 [set producer-A true set color blue set size 1.2]
  ask site spacing [set producer-B true set color blue set size 1.2]
  ask site (spacing + spacing) [set producer-C true set color blue set size 1.2]
  ask site (spacing + spacing + spacing) [set producer-D true set color blue set size 1.2]
  if remainder num-sites 4 != 0
```

```
; make the user aware of the bias in the spacing of sites in cases where the number of sites
is not divisibe by the number of production centres
  [user-message ( word "The number of sites is not divisible by the number of wares -->
producer-sites are not evenly spaced. Click OK to proceed" )]
end

to distribute-traders-on-sites
; the trader distribution (number of traders per site) is initialized depending on the
hypothesis
  if traders-distribution = "uniform"
; uniform distribution of traders on sites
  [setup-uniform-distribution]
  if traders-distribution = "normal"
; normal distribution of traders on sites
  [setup-normal-distribution]
  if traders-distribution = "exponential"
; exponential distribution of traders on sites
  [setup-exponential-distribution]
end

to connect-traders
  if network-structure = "random" [connect-random-network]
; create a random network connecting traders. This option allows for comparing the results of
the hypothesises being tested with those of the same processes working on a random network
  if network-structure = "hypothesis"
; The commercial network of traders is set up in five steps reflecting the hypotheses being
tested
  [
    connect-traders-adjacent-sites
; 1) we ensure that at least one pair of traders is connected between every pair of adjacent
sites in the circular layout
    connect-random-traders
; 2) a proportion of randomly selected pairs on different sites is connected
; 3) on each site a number of randomly selected pairs of traders are connected
; 4) on each site a number of pairs of traders with mutual contacts on the same site are
connected
    connect-to-av-degree
; (steps 3 and 4 are looped until the average degree is reached)
    single-connected-component
; 5) we ensure that the network consists of one connected component, i.e. that there are no
isolated traders and multiple components
  ]
end

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;; GO PROCEDURES ;;;;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

to reposition-traders
  if layout? = false [ask traders [move-to site site-number]]
; ensure the traders are located at the correct site
end
```

```
to determine-demand
; each site has a maximum demand (=num-traders*num-traders), demand increases at a constant
rate up to this maximum (representing new demand due to e.g. pots breaking)
  ask traders
; if a trader's demand is lower than the number of traders at its site, increase its demand by
1
  [if demand < (count traders-here)
    [set demand demand + 1]]
end

to discard-part-of-stock
; traders discard a constant proportion of their stock as a cost/risk/punishment of
redistributing items
; all of their remaining stock is added to their volume of the product and the stock is set to
0
; NOTE that increasing the amount of product at the end of the tick ensures it does not
decrease the demand of the agent
; Next tick the agent will have the same demand as always and will strive to satisfy this
demand as well as being able to sell the items he obtained to other traders who have a
positive demand.
ask traders
  [
    let discard-A round(stock-A * 0.14)
    let discard-B round(stock-B * 0.14)
    let discard-C round(stock-C * 0.14)
    let discard-D round(stock-D * 0.14)
    set product-A product-A + (stock-A - discard-A)
    set product-B product-B + (stock-B - discard-B)
    set product-C product-C + (stock-C - discard-C)
    set product-D product-D + (stock-D - discard-D)
    set stock-A 0
    set stock-B 0
    set stock-C 0
    set stock-D 0
    ask sites-here
    [
      set volume-A volume-A + discard-A
      set volume-B volume-B + discard-B
      set volume-C volume-C + discard-C
      set volume-D volume-D + discard-D
    ]
  ]
end

to produce
  ask sites with [producer-A = true]
  [ask traders-here ; all traders on the production site obtain X newly produced items...
    [if (product-A + product-B + product-C + product-D) < demand
; ... if its total possession of all products is less than its demand
      [set product-A (product-A + round(demand - (product-A + product-B + product-C + product-
D)))]]]
; increase product by however much it differs from demand
```

```
  ask sites with [producer-B = true]
  [ask traders-here ; all traders on the production site obtain newly produced items...
    [if (product-A + product-B + product-C + product-D) < demand ; ... if its total possession
of all products is less than its demand
      [set product-B product-B + round(demand - (product-A + product-B + product-C + product-
D))]]] ; increase product by however much it differs from demand
  ask sites with [producer-C = true]
  [ask traders-here ; all traders on the production site obtain newly produced items...
    [if (product-A + product-B + product-C + product-D) < demand
; ... if its total possession of all products is less than its demand
      [set product-C product-C + round(demand - (product-A + product-B + product-C + product-
D))]]] ; increase product by however much it differs from demand
  ask sites with [producer-D = true]
  [ask traders-here ; all traders on the production site obtain newly produced items...
    [if (product-A + product-B + product-C + product-D) < demand
; ... if its total possession of all products is less than its demand
      [set product-D product-D + round(demand - (product-A + product-B + product-C + product-
D))]]] ; increase product by however much it differs from demand
end

to trade-and-consume
  ask traders
  [
; determine the neighboring traders each trader obtains commercial information from in this
tick
    identify-known-traders
    price-setting ; estimate the price for one item
  ]
  ask traders [set-maximum-stock-size]
; determine how many items a trader is willing to obtain over his own demand for
redistribution in the next tick
; each item in all traders' possessions will be considered in turn in a random order. Each
item is either deposited as a result of a successful transaction, or it is added to a trader's
stock
  let product-on-market sum [ product-A + product-B + product-C + product-D] of traders
; count the total number of items of all products
  set counter-trade 0
; counter used to keep track of the total  number of items considered in each tick
  set counter-transaction 0
; counter to report the number of successful transactions per tick
  set counter-stock-redist 0
; counter to report the number of items put in stock with redistribution in mind per tick
  set counter-no-transaction 0
; counter to report the number of failed transactions per tick
  while [counter-trade < product-on-market]
; repeat until all items on the market have been considered for trade
  [
; for each potential transaction randomly select whether an item of product A, B, C or D will
be traded
    let p random 4
    if p = 0
```

```
    [trade-product-A]
    if p = 1
    [trade-product-B]
    if p = 2
    [trade-product-C]
    if p = 3
    [trade-product-D]
  ]
end

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;;;;;; TRADE PROCEDURES ;;;;;;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

; the trade-product procedures are exactly the same for each product with the exception that
only agent parameters relevant to one specific product are updated in each procedure
to trade-product-A
  if count traders with [product-A > 0] > 0 ; if there are traders with this product ...
  [ask one-of traders with [product-A > 0] ; select a seller at random
    [let potential-buyers link-neighbors with [(demand > 0) or (maximum-stock-size > 0)]
; trade can only happen with link neighbors that have a positive demand or are willing to
stock items for redistribution
      ifelse count potential-buyers = 0
      [
; if there are no buyers for the item then put all of your items in stock
        set stock-A stock-A + product-A
        set maximum-stock-size maximum-stock-size - product-A
        set counter-trade counter-trade + product-A
        set counter-no-transaction counter-no-transaction + product-A
        set product-A 0
      ]
      [ ; if there are potential buyers for the item then consider selling it
        let seller-price price
        let buyer-price 0
        ask potential-buyers [add-transport-cost]
; determine whether a transport fee applies in cases where the buyer is not located on the
same market as the seller
        let likely-buyer one-of potential-buyers with-max [(price - transport-fee)]
; select the buyer that offers the highest profit (the most likely buyer of the item)
        ask likely-buyer [set buyer-price (price - transport-fee)]
        ifelse (buyer-price - seller-price) >= 0
; if the seller can make a profit then complete the transaction
        [
          set counter-transaction counter-transaction + 1
          set product-A product-A - 1 ; process the transaction for the seller
          ask likely-buyer ; process the transaction for the buyer, two possible actions...
          [
            ifelse demand < 1
; ... 1) if the buyer's demand is 0 and the average demand is higher than its own demand
(captured by a positive maximum-stock-size) then store the item for redistribution in the next
tick (with the prospect of trading it for a higher profit)
            [
```

```
              set stock-A stock-A + 1
              set maximum-stock-size maximum-stock-size – 1
              set counter-stock-redist counter-stock-redist + 1
            ]
            [
; ... 2) if the buyer's demand is not 0 then sell it to a consumer who deposits the item at a
site
              set demand demand – 1
              ask one-of sites-here [set volume-A volume-A + 1]
            ]
          ]
        ]
        [ ; if the seller cannot make a profit then move this item in stock
          set product-A product-A – 1
          set stock-A stock-A + 1
          set maximum-stock-size maximum-stock-size – 1
          set counter-no-transaction counter-no-transaction + 1
        ]
        set counter-trade counter-trade + 1
      ]
    ]
  ]
end

to trade-product-B
  if count traders with [product-B > 0] > 0 ; if there are traders with this product ...
  [ask one-of traders with [product-B > 0] ; select a seller at random
    [let potential-buyers link-neighbors with [(demand > 0) or (maximum-stock-size > 0)]
; trade can only happen with link neighbors that have a positive demand
      ifelse count potential-buyers = 0
      [ ; if there are no buyers for the item then put all items in stock
        set stock-B stock-B + product-B
        set maximum-stock-size maximum-stock-size - product-B
        set counter-trade counter-trade + product-B
        set counter-no-transaction counter-no-transaction + product-B
        set product-B 0
      ]
      [ ; if there are potential buyers for the item consider selling it
        let seller-price price
        let buyer-price 0
        ask potential-buyers [add-transport-cost]
; determine whether a transport fee applies in cases where the buyer is not located on the
same market as the seller
        let likely-buyer one-of potential-buyers with-max [(price - transport-fee)]
; select the buyer that offers the highest profit (the most likely buyer of the item)
        ask likely-buyer [set buyer-price (price - transport-fee)]
        ifelse ((buyer-price - transport-fee) - seller-price) >= 0
; if the seller can make a profit then complete the transaction
        [
          set counter-transaction counter-transaction + 1
          set product-B product-B - 1 ; process the transaction for the seller
```

```
          ask likely-buyer ; process the transaction for the buyer, two possible actions...
          [
            ifelse demand < 1
; ... 1) if the buyer's demand is 0 and the average demand is higher than its own demand
(captured by a positive maximum-stock-size) then store the item for redistribution in the next
tick (with the prospect of trading it for a higher profit)
            [
              set stock-B stock-B + 1
              set maximum-stock-size maximum-stock-size – 1
              set counter-stock-redist counter-stock-redist + 1
            ]
            [
; ... 2) if the average demand is lower than my own demand then sell it to a consumer who
deposits the item at a site
              set demand demand – 1
              ask one-of sites-here [set volume-B volume-B + 1]
            ]
          ]
        ]
        [ ; if the seller cannot make a profit then move this item in stock
          set product-B product-B – 1
          set stock-B stock-B + 1
          set maximum-stock-size maximum-stock-size – 1
          set counter-no-transaction counter-no-transaction + 1
        ]
        set counter-trade counter-trade + 1
      ]
    ]
  ]
end

to trade-product-C
  if count traders with [product-C > 0] > 0 ; if there are traders with this product ...
  [ask one-of traders with [product-C > 0] ; select a seller at random
    [let potential-buyers link-neighbors with [(demand > 0) or (maximum-stock-size > 0)]
; trade can only happen with link neighbors that have a positive demand
      ifelse count potential-buyers = 0
      [ ; if there are no buyers for the item then put all items in stock
        set stock-C stock-C + product-C
        set maximum-stock-size maximum-stock-size - product-C
        set counter-trade counter-trade + product-C
        set counter-no-transaction counter-no-transaction + product-C
        set product-C 0
      ]
      [ ; if there are potential buyers for the item consider selling it
        let seller-price price
        let buyer-price 0
        ask potential-buyers [add-transport-cost]
; determine whether a transport fee applies in cases where the buyer is not located on the
same market as the seller
        let likely-buyer one-of potential-buyers with-max [(price - transport-fee)]
```

```
; select the buyer that offers the highest profit (the most likely buyer of the item)
        ask likely-buyer [set buyer-price (price - transport-fee)]
        ifelse ((buyer-price - transport-fee) - seller-price) >= 0
; if the seller can make a profit then complete the transaction
        [
          set counter-transaction counter-transaction + 1
          set product-C product-C - 1 ; process the transaction for the seller
          ask likely-buyer ; process the transaction for the buyer, two possible actions...
          [
            ifelse demand < 1
; ... 1) if the buyer's demand is 0 and the average demand is higher than its own demand
(captured by a positive maximum-stock-size) then store the item for redistribution in the next
tick (with the prospect of trading it for a higher profit)
            [
              set stock-C stock-C + 1
              set maximum-stock-size maximum-stock-size – 1
              set counter-stock-redist counter-stock-redist + 1
            ]
            [
; ... 2) if the average demand is lower than my own demand then sell it to a consumer who
deposits the item at a site
              set demand demand – 1
              ask one-of sites-here [set volume-C volume-C + 1]
            ]
          ]
        ]
        [ ; if the seller cannot make a profit then move this item in stock
          set product-C product-C – 1
          set stock-C stock-C + 1
          set maximum-stock-size maximum-stock-size – 1
          set counter-no-transaction counter-no-transaction + 1
        ]
        set counter-trade counter-trade + 1
      ]
    ]
  ]
end

to trade-product-D
  if count traders with [product-D > 0] > 0 ; if there are traders with this product ...
  [ask one-of traders with [product-D > 0] ; select a seller at random
    [let potential-buyers link-neighbors with [(demand > 0) or (maximum-stock-size > 0)]
; trade can only happen with link neighbors that have a positive demand
      ifelse count potential-buyers = 0
      [ ; if there are no buyers for the item then put it in stock
        set stock-D stock-D + product-D
        set maximum-stock-size maximum-stock-size - product-D
        set counter-trade counter-trade + product-D
        set counter-no-transaction counter-no-transaction + product-D
        set product-D 0
      ]
```

```
        [ ; if there are potential buyers for the item consider selling it
          let seller-price price
          let buyer-price 0
          ask potential-buyers [add-transport-cost]
; determine whether a transport fee applies in cases where the buyer is not located on the
same market as the seller
          let likely-buyer one-of potential-buyers with-max [(price - transport-fee)]
; select the buyer that offers the highest profit (the most likely buyer of the item)
          ask likely-buyer [set buyer-price (price - transport-fee)]
          ifelse ((buyer-price - transport-fee) - seller-price) >= 0 ; if the seller can make a
profit then complete the transaction
          [
            set counter-transaction counter-transaction + 1
            set product-D product-D - 1 ; process the transaction for the seller
            ask likely-buyer ; process the transaction for the buyer, two possible actions...
            [
              ifelse demand < 1
; ... 1) if the buyer's demand is 0 and the average demand is higher than its own demand
(captured by a positive maximum-stock-size) then store the item for redistribution in the next
tick (with the prospect of trading it for a higher profit)
              [
                set stock-D stock-D + 1
                set maximum-stock-size maximum-stock-size – 1
                set counter-stock-redist counter-stock-redist + 1
              ]
              [
; ... 2) if the average demand is lower than my own demand then sell it to a consumer who
deposits the item at a site
                set demand demand – 1
                ask one-of sites-here [set volume-D volume-D + 1]
              ]
            ]
          ]
          [ ; if the seller cannot make a profit then move this item in stock
            set product-D product-D – 1
            set stock-D stock-D + 1
            set maximum-stock-size maximum-stock-size – 1
            set counter-no-transaction counter-no-transaction + 1
          ]
          set counter-trade counter-trade + 1
        ]
      ]
    ]
end

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;; DISTRIBUTE TRADERS ;;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

to setup-uniform-distribution
  ask sites [set target-distribution ceiling (num-traders / num-sites)]
; determine the number of traders that should be present at each site. Round up, this ensures
```

```
we have as many sites with the target number of traders as possible.
  ask traders with [moved? = false]
; ask each trader that has not been moved to a site yet in turn
  [ifelse count sites with [count traders-here < target-distribution] = 0
    [
; move to a randomly selected site if all sites already reached their desired number of
traders
      move-to one-of sites
      set moved? True
    ]
    [
; if some sites did not yet reach their desired number of traders then move to one of these
sites
      move-to one-of sites with [count traders-here < target-distribution]
      set moved? True
    ]
  ]
end

to setup-normal-distribution
  ask sites
  [set target-distribution ceiling (random-normal (num-traders / num-sites) floor((num-traders
/ num-sites) / 3 ))]
; round up, this ensures we have as many sites with the target number of traders as possible.
; a normal distribution is created with a mean equal to the mean number of traders per site,
and the standard deviation of the mean divided by three and rounded down
; this ensures all sites will have a positive number of traders, since 99.9% of all values lie
within three standard deviations from the mean
  ask traders with [moved? = false]
; ask each trader that has not been moved to a site yet in turn
  [ifelse count sites with [count traders-here < target-distribution] = 0
    [
; move to a randomly selected site if all sites already reached their desired number of
traders
      move-to one-of sites
      set moved? True
    ]
    [
; if some sites did not yet reach their desired number of traders then move to one of these
sites
      move-to one-of sites with [count traders-here < target-distribution]
      set moved? True
    ]
  ]
end

to setup-exponential-distribution
  ask sites
  [set target-distribution ceiling (random-exponential (num-traders / num-sites))]
; round up, this ensures we have as many sites with the target number of traders as possible.
; an exponential distribution is created with a mean equal to the mean number of traders per
site
```

```
  ask traders with [moved? = false]
; ask each trader that has not been moved to a site yet in turn
  [ifelse count sites with [count traders-here < target-distribution] = 0
    [
; move to a randomly selected site if all sites already reached their desired number of
traders
      move-to one-of sites
      set moved? True
    ]
    [
; if some sites did not yet reach their desired number of traders then move to one of these
sites
      move-to one-of sites with [count traders-here < target-distribution]
      set moved? True
    ]
  ]
end

;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;; CONNECT TRADERS ;;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;

; if the experiments representing the hypotheses are compared with a random network with the
same density then do the following procedure
to connect-random-network
  connect-traders-adjacent-sites connect-random-traders connect-to-av-degree single-connected-
component
  let num-links count links
  ask links [die]
  repeat num-links [ask one-of traders [create-link-with one-of traders with [self != myself
and (link-with myself = nobody)]]]
end

; if the experiments represent the hypotheses then do the following five steps:
to connect-traders-adjacent-sites
;; FIRST, connect one trader on each site to a trader on the next site in the circuar layout
  set neighbor-site-links num-sites
  let site-counter 0
  while [site-counter < ( num-sites - 1 ) ]
  [
    ask one-of traders with [site-number = site-counter] [create-link-with (one-of traders
with [site-number = (site-counter + 1)])]
    set site-counter site-counter + 1
  ]
;; connect a trader on the first site to one on the last site (since this is not done in the
previous step)
  ask one-of traders with [site-number = 0]
  [create-link-with (one-of traders with [site-number = (num-sites - 1)])]
end

to connect-random-traders
;; SECOND, create a variable number of inter-site links by randomly selecting a pair of
```

```
traders on different sites and connecting them
;; in order for the proportion of inter-site links to be similar irrespective of the number of
traders I adopt the approach by Jin et al 2001
;;"at each time-step, we choose np * r0 pairs of vertices uniformly at random from the network
to meet.
;; if a pair meet who do not have a pre-existing connection, and if neither of them already
has the maximum number of connections
;; then a new connection is established between them."
;; np = 1/2*N *(N-1) (in this procedure called node-pairs1)
;; s0 is a variable (called r0 by Jin et al.; in this procedure called proportion-inter-site-
links)
  let node-pairs1 ( ( num-traders / 2 ) * (num-traders - 1) )
; this is called np in jin etal 2001, i.e. the number of node pairs
  set random-inter-site-links ceiling(node-pairs1 * proportion-inter-site-links)
; show the number of random links in this setup of the model
  let time 0
  while [time < random-inter-site-links] ; np * s0 pairs of vertices are selected at random
  [ask one-of traders with [count link-neighbors < maximum-degree]; select node1
    [
      let node2 one-of traders with [ self != myself and (link-with myself = nobody) and
[site-number] of self != [site-number] of myself and count link-neighbors < maximum-degree]
; identify a node that is not myself, is not a link-neighbor, and is located on another site
      create-link-with node2
      set time time + 1]]
end

to connect-to-av-degree
  set random-intra-site-links 0
  set mutual-neighbors-intra-site-links 0
  while [av-degree <= (maximum-degree - ((maximum-degree / 100) * 10))]
; the following two processes of link creation are looped until the average degree is reached
  [
; THIRD, create links between pairs of randomly selected traders on the same site
; step 1 of the Jin etal 2001 simplified model starts here, but modified to only select
traders on the same site
; "at each time-step, we choose np * r0 pairs of vertices uniformly at random from the network
to meet.
; if a pair meet who do not have a pre-existing connection, and if neither of them already has
the maximum number of connections
; then a new connection is established between them."
; np = 1/2*N *(N-1) (called node-pairs2 in this procedure)
; r0 is a variable (called proportion-intra-site-links in this procedure)
    let node-pairs2 ( ( (count traders) / 2 ) * ((count traders) - 1) )
; this is called np in jin etal 2001
    repeat (node-pairs2 * proportion-intra-site-links)
; np * r0 pairs of vertices are selected at random
    [if any? traders with [count link-neighbors < maximum-degree]
      [ask one-of traders with [count link-neighbors < maximum-degree]
        [if any? (traders-here with [self != myself and (link-with myself = nobody) and count
link-neighbors < maximum-degree])
          [
```

```
; identify a node that is not myself, is not a link-neighbor, does not have the maximum
degree, and is located at the same site as node1
            create-link-with one-of (traders-here with [self != myself and (link-with myself =
nobody) and count link-neighbors < maximum-degree])
            set random-intra-site-links random-intra-site-links + 1]]]]
; FOURTH, nodes with a mutual contact on the same site are invited to become connected
; step 2 of the Jin etal 2001 simplified model starts here
; at each time-step, we choose NmR1 vertices at random, with probabilities proportional to Zi
( Zi - 1 ).
; for each vertex chosen we randomly choose one pair of its neighbors to meet, and establish a
new connection between them
; if they do not have a pre-existing connection and if neither of them already has the maximum
number of connections.
; mutual-neighbors = 1/2 * SUMi ( Zi * ( Zi - 1) ). Parameter Nm in Jin etal 2001 and
represents the number of mutual neighbors.
; proportion-mutual-neighbors (r1 in Jin et al 2001)
    let mutual-neighbors ((sum([( count link-neighbors ) * ( ( count link-neighbors ) - 1 )]
of traders)) / 2) ; calculate nm
    repeat (mutual-neighbors * proportion-mutual-neighbors) ; repeat nm * r1 times
    [ask one-of traders ; each time randomly select a trader
      [
        let possible-mutual-neighbors link-neighbors with [[site-number] of self = [site-
number] of myself and count link-neighbors < maximum-degree]
        if any? possible-mutual-neighbors
        [
          let node1 one-of possible-mutual-neighbors
          ask node1
            [
              let possible-node2 possible-mutual-neighbors with [self != node1 and (link-with
node1 = nobody)]
              if any? possible-node2
              [
              create-link-with one-of possible-node2
              set mutual-neighbors-intra-site-links (mutual-neighbors-intra-site-links +
1)]]]]]]
end

to single-connected-component
; FIFTH, we ensure that the network consists of one connected component, i.e. that there are
no isolated traders and multiple components
; this to reflect the theoretical possibility in the static network that objects produced
anywhere can end up in any site
; to enforce this the network nodes need to be able to be connected by a path
  set random-component-links 0
  let length-c length nw:weak-component-clusters
; calculate the number of components and save this as length-c
  let components nw:weak-component-clusters
; identify the components as a list of different agentsets
  let number-components n-values length-c [?]
; create a list of length-c counting from 0 to (length-c - 1). This list will be used to give
each component a different number
```

```
; loop through the two lists just created: the components and the numbers
  (foreach components number-components
    [
      let cluster ?1
      let cluster-no ?2
      ask cluster ; for each trader in the component
      [set comp cluster-no] ; give them a number according to the component they are in
    ]
  )
  let count-components length-c
  while [count-components > 1]
; repeat the following process of link creation between pairs of traders located on the same
site but in different network components as long as there are multiple connected components
  [ask one-of sites ; randomly select a site
    [
      let f min [comp] of traders-here
      let g max [comp] of traders-here
      if f != g ; if there is a difference in the components its traders belong to ...
      [ask one-of traders-here with [comp = f] ; ... select two traders in different
components and create a link between them
        [ask one-of other traders-here with [comp = g]
          [
            create-link-with myself
            set random-component-links random-component-links + 1]
        ]
      ]
    ] ; re-calculate the number of components
    set length-c length nw:weak-component-clusters
; calculate the number of components and save this as length-c
    set count-components length-c
    set components nw:weak-component-clusters
; identify the components as a list of different agentsets
    set number-components n-values length-c [?]
; create a list of length-c counting from 0 to (length-c - 1). This list will be used to give
each component a different number
; loop through the two lists just created: the components and the numbers
    (
      foreach components number-components
      [
        let cluster ?1
        let cluster-no ?2
        ask cluster ; for each trader in the component
        [set comp cluster-no] ; give them a number according to the component they are in
      ]
    )
  ]
end

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;; PRICE SETTING ;;;;;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
```

```
to identify-known-traders
; every tick each trader will only have commercial information available from a fraction of
its link neighbors. These traders are identified here to ensure they remain the same
throughout the tick.
  set known-traders n-of ceiling((count link-neighbors) * local-knowledge) link-neighbors
end

to price-setting
; the trader estimates what the current price for one item of a product is, based on the
demand and supply of a proportion of his link-neighbors.
  find-average-demand ; determine the average demand of this fraction of neighbors
  let sum-prod sum[product-A + product-B + product-C + product-D] of known-traders
; determine the total supply of this fraction of neighbors
  let average-supply ((sum-prod + product-A + product-B + product-C + product-D) / ((count
known-traders) + 1))
; determine the average supply of this fraction of neighbors + myself
  set price (average-demand / (average-supply + average-demand))
; determine the trader's price estimate
; the price of the product as perceived by the actor is the average demand divided by the
average supply plus the average demand (to normalise the result)
end

to find-average-demand
  let sum-demand sum[demand] of known-traders
; determine the total demand of this fraction of neighbors
  set average-demand ((sum-demand + demand) / ((count known-traders) + 1))
; determine the average demand of this fraction of neighbors + myself
end

to set-maximum-stock-size
  let sum-prod (product-A + product-B + product-C + product-D)
; determine total number of items possessed
  ifelse sum-prod > demand ; if this is higher than trader's demand ...
  [set maximum-stock-size round(average-demand - sum-prod)]
; ... set its maximum stock size to the difference with the average-demand
  [set maximum-stock-size round(average-demand - demand)]
; if not, then set its maximum stock size to the difference with the trader's demand
end

to add-transport-cost
; potential buyers observe that the seller will need to pay a transport fee if the pair is not
on the same market. This fee is stored in the transport-fee parameter of the potential buyer
  set transport-fee 0
; ensure possible transport fee values of previous transactions are not considered by setting
the parameter to 0
  if [site-number] of myself != site-number [set transport-fee transport-cost]
; per experiment the transport-fee is set by the variable transport-cost
end

;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;; LAYOUT ;;;;;;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;
```

```
to layout
; the number 3 here is arbitrary; more repetitions slows down the model, but too few gives
poor layouts
  repeat 3 [
; the more nodes we have to fit into the same amount of space,
; the smaller the inputs to layout-spring we'll need to use
    let factor sqrt count traders
    ; numbers here are arbitrarily chosen for pleasing appearance
    layout-spring traders links (1 / factor) (7 / factor) (1 / factor)
    display  ; for smooth animation
  ]
  ; don't bump the edges of the world
  let x-offset max [xcor] of traders + min [xcor] of traders
  let y-offset max [ycor] of traders + min [ycor] of traders
  ; big jumps look funny, so only adjust a little each time
  set x-offset limit-magnitude x-offset 0.1
  set y-offset limit-magnitude y-offset 0.1
  ask traders [ setxy (xcor - x-offset / 2) (ycor - y-offset / 2) ]
end

to-report limit-magnitude [number limit]
  if number > limit [ report limit ]
  if number < (- limit) [ report (- limit) ]
  report number
end

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;; Network measures ;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

to-report report-neighbor-site-links
  report neighbor-site-links
end

to-report report-random-inter-site-links
  report random-inter-site-links
end

to-report report-random-intra-site-links
  report random-intra-site-links
end

to-report report-mutual-neighbors-intra-site-links
  report mutual-neighbors-intra-site-links
end

to-report report-random-component-links
  report random-component-links
end

to-report report-av-tdegree
  set av-degree sum([count link-neighbors] of traders) / num-traders
```

```
  report av-degree
end

to-report number-tlinks
  report count links
end

to-report average-shortest-path-length
  report network:mean-link-path-length traders links
end

; clustering coefficient (cc) measure, based on small-world model in Netlogo library
; for an alternative calculation see:
; http://www.ladamic.com/netlearn/NetLogo4/SmallWorldWS.nlogo
to-report in-neighborhood? [ hood ]
  report ( member? end1 hood and member? end2 hood )
end

to-report find-clustering-coefficient
  let traders-with-links traders with [ count link-neighbors != 0 ]
  ifelse all? traders-with-links [count link-neighbors <= 1]
  [
    ; it is undefined
    ; what should this be?
    set clustering-coefficient 0
  ]
  [
    let total 0
    ask traders-with-links with [ count link-neighbors <= 1]
    [
        set trader-clustering-coefficient "undefined"
    ]
    ask traders-with-links with [ count link-neighbors > 1]
; only when more than one neighbor exists will cc be calculated
    [
      let hood link-neighbors
      set trader-clustering-coefficient (2 * count links with [ in-neighborhood? hood ] /
                                   ((count hood) * (count hood - 1)) )
      ; find the sum for the value at turtles
      set total total + trader-clustering-coefficient
    ]
    ; take the average
    set clustering-coefficient total / count traders-with-links with [count link-neighbors >
1]
    report clustering-coefficient
  ]
end
```

13. Glossary

This glossary contains definitions of terms used in this thesis. Some of these are explained using figure 79 as an example.

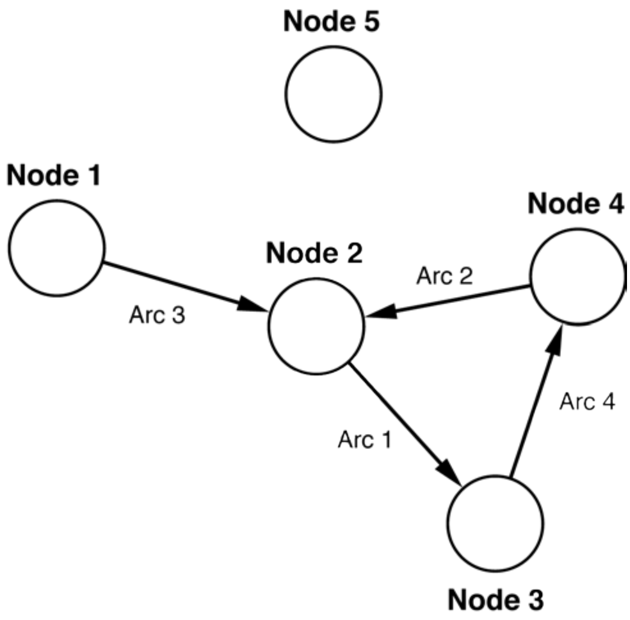


Fig. 79. Example network used in some of the definitions below to clarify network concepts.

Arc: in network jargon an arc is a directed line joining two points in a network.

Average degree: the degree of a node is the number of edges connected to it and the average degree is therefore the average of all degree scores in a single network (Newman 2010, 133-136; Nooy et al. 2005, 63-64).

Average shortest path length: a shortest path or geodesic path in network terms is the shortest route over the network that runs from one vertex to another along the edges of the network and the average shortest path length is the average of all shortest path scores between all possible pairs of vertices in the network (Newman 2010 136-140; de Nooy et al. 2005, 127). For example, the shortest path from Node 1 to Node 4 in figure 79 is 3, and the average shortest path length of the network is 1.667.

Betweenness centrality: the proportion of all shortest paths between pairs of other vertices that include this vertex (de Nooy et al. 2005, 131).

Centrality: [something that refers to a number of measures, including the ones defined elsewhere, e.g., betweeness, closeness, degree, eigenvector]

Closeness centrality: “the number of other vertices divided by the sum of all distances between the vertex and all others” (Nooy et al. 2005, 127; Sabidussi 1966).

Clustering coefficient: the Watts-Strogatz *clustering coefficient* measures the average probability that two neighbours of a vertex are themselves neighbours as a ratio of the number of edges between the neighbours of a given node and the maximum number of edges that could possibly exist between these neighbours (Newman 2010, 262-266; Watts and Strogatz 1998, 441).

Cohesion: network cohesion (usually called density) is the fraction of the maximum possible number of edges in the network that is actually present (Newman 2010, 134; Wasserman and Faust 1994, 101-103).

Complete network: “a network with maximum density” (Nooy et al. 2005, 63).

Components: connected parts of a network (Nooy et al. 2005, 66-70). See also weak component and strong component.

Degree: the degree of a node is the number of edges connected to it and the average degree is therefore the average of all degree scores in a single network (Newman 2010, 133-136; Nooy et al. 2005, 63-64). For example, Node 2 in figure 79 has a degree of 3. See also indegree and outdegree.

Degree distribution: the degree distribution represents the fraction of nodes in a network with a certain degree (Albert and Barabási 2002, 49; Newman 2010, 243-247).

Density: network density (sometimes called cohesion) is the fraction of the maximum possible number of edges in the network that is actually present (Newman 2010, 134; Wasserman and Faust 1994, 101-103).

Diameter: The network diameter is the length of the longest geodesic path between any pair of nodes in the network (Newman 2010, 140; Wasserman and Faust 1994, 111-112).

Edge: in network jargon an edge is an undirected line joining two points in a network.

e.g.: example given.

Ego-network: a node, the nodes it is directly connected to, and the connections between them.

ERGM: exponential random graph models, introduced in detail in chapter 4.

Geodesic path: a shortest path or geodesic path in network terms is the shortest route over the network that runs from one vertex to another along the edges of the network and the average shortest path length is the average of all shortest path scores between all possible pairs of vertices in the network (Newman 2010, 136-140; Nooy et al. 2005, 127). For example, the shortest path from Node 1 to Node 4 in figure 79 is 3, and the average shortest path length of the network is 1.667.

Graph: “a graph is a set of vertices and a set of lines between pairs of vertices” (Nooy et al. 2005, 6).

Hubs and authorities: in citation networks Hubs are publications that cite many other publications and many good authorities in particular. Authorities are publications that are cited by many other publications and by good hubs in particular (Kleinberg 1999).

i.e.: *id est*.

Indegree: “the indegree of a vertex is the number of arcs it receives” (Nooy et al. 2005, 64). For example, Node 2 in figure 79 has an indegree of 2. See also degree and outdegree.

Input domain: represents the number of all other vertices that are connected to a given vertex by a path (Nooy et al. 2005, 193).

Isolates: nodes that are unconnected to other nodes within a sample or population.

M-slices: M-slices include vertices linked by lines with a value equal to or greater than *m* (Nooy et al. 2005, 109).

Network: “a network consists of a graph and additional information on the vertices or the lines of the graph” (Nooy et al. 2005, 7). “A network is, in its simplest form, a collection of points joined together in pairs by lines” (Newman 2010, 1).

Node: in network jargon a node or vertex is a point in a network.

One-mode network: “In a one-mode network, each vertex can be related to each other vertex.” (Nooy et al. 2005, 103).

Outdegree: the outdegree is the number of arcs a vertex sends (Nooy et al. 2005, 64). For example, Node 2 in figure 79 has an outdegree of 1. See also degree and indegree.

Path: “a path is a walk in which no vertex in between the first and last vertex of the walk occurs more than once” (Nooy et al. 2005, 67). For example, a path from Node 1 to Node 4 in figure 79 passes over arcs 3, 1 and 4. The difference with a walk is that in a path none of these steps can be performed multiple times.

RMSE: root-mean-square-error.

Semipath: “a semipath is a semiwalk in which no vertex in between the first and last vertex of the semiwalk occurs more than once” (Nooy et al. 2005, 67). For example, a semipath from Node 1 to Node 4 in figure 79 passes over arcs 3, 1 and 4. The difference with a semiwalk is that in a semipath none of these steps can be performed multiple times.

Semiwalk: “a semiwalk from vertex u to vertex v is a sequence of lines such that the end vertex of one line is the starting vertex of the next line and the sequence starts at vertex u and ends at vertex v” (Nooy et al. 2005, 67). For example, a semiwalk from Node 1 to Node 4 in figure 79 passes over arcs 3, 1 and 4. The difference with a semipath is that in a walk some of these steps can be performed multiple times.

Shortest path: a shortest path or geodesic path in network terms is the shortest route over the network that runs from one vertex to another along the edges of the network and the average shortest path length is the average of all shortest path scores between all possible pairs of vertices in the network (Newman 2010, 136-140; Nooy et al. 2005, 127). For example, the shortest path from Node 1 to Node 4 in figure 79 is 3, and the average shortest path length of the network is 1.667.

Strong component: “a strong component is a maximal strongly connected subnetwork” (Nooy et al. 2005, 68). For example, figure 79 contains three strong components: nodes 2, 3, and 4 (because they are connected by a path); secondly the isolated node 5; and thirdly node 1 (which cannot be connected to by a path).

Strongly connected network: “a network is strongly connected if each pair of vertices is connected by a path” (Nooy et al. 2005, 67). For example, figure 79 is not strongly connected because Node 5 is isolated.

Two-mode network: “In a two-mode network, vertices are divided into two sets and vertices can only be related to vertices in the other set.” (Nooy et al. 2005, 103).

Vertex (vertices pl.): in network jargon a vertex or node is a point in a network. In this special issue the term node is preferred over vertex.

Walk: “a walk is a semiwalk with the additional condition that none of its lines are an arc of which the end vertex is the arc’s tail” (Nooy et al. 2005, 67). For example, a walk from Node 1 to Node 4 in figure 79 passes over arcs 3, 1 and 4. The difference with a path is that in a walk some of these steps can be performed multiple times.

Weak component: “a (weak) component is a maximal (weakly) connected subnetwork” (Nooy et al. 2005, 68). For example, figure 79 contains the (weak) component consisting of nodes 1, 2, 3 and 4.

Weakly connected network: “a network is (weakly) connected if each pair of vertices is connected by a semipath” (Nooy et al. 2005, 68). For example, figure 79 is not (weakly) connected because Node 5 is isolated.