University of Southampton
Faculty of Environmental and Life Sciences
Geography and Environmental Science

Network Analysis of Simulated and Real Indigenous Irrigation Systems

Alex Stokes

Thesis for Master of Philosophy

May 2020

University of Southampton

Abstract

Faculty of Environmental and Life Sciences

Geography and Environmental Science

Thesis for Master of Philosophy

Network Analysis of Simulated and Real Indigenous Irrigation Systems

Alex Stokes

Small-scale Indigenous Irrigation Systems (IIS) are water-sharing societies which have been observed to persist for long periods of time finding a dynamic equilibrium with the environment. This persistence is thought to be mostly due to the institutions and system structures which evolve to maintain stability despite internal and external changes. They have been described as the most ancient and ubiquitous example of public infrastructure system, however, the way in which they grow, evolve and maintain stability is a contested and also controversial topic. The study of IIS is interdisciplinary and generally classified under the umbrella term Social-Ecological Systems.

Advances in computing performance and software have enabled simulation modelling to be quicker, cheaper and more accessible. Alongside this, recent scientific understanding of complex systems and network theory has led to new interdisciplinary theories on universal scaling and preferential attachment based growth in systems of many interacting components. This research aims to harness this potential by building an abstract simulated IIS in a generative agent-based model environment.

The model assumes that the IIS network grows through preferential attachment to optimise space, which is the mechanism of growth found in previous models and a common assumption in biological systems as it increases efficiency. Different growth strategies relating to local or global information and stochastic processes are tested and find a range of space optimal configurations.

The results find that all models form rooted planar tree networks. Stochastic processes are important for the model to search for different model configurations and finding almost optimal configuration. The optimal solution is found from collecting global information of the network and selecting growth glob-

ally, however this is computationally expensive, and inefficient so unlikely to be found in real systems. The networks formed through random asynchronous processes follow scale-free laws which coincide with similar scaling exponent as other sub-critical networks such as rivers.

Two vastly different real-world IIS networks are then analysed and compared with the simulated models. The real-world networks show differences comparable to the differences found in the simulated models. The reasons for these differences are speculated to be due to a number of factors including geomorphology and managerial arrangements, however given the limited data collected no firm conclusion can be made. It is recommended that further data is collected and analysed in order to confirm this.

Contents

List of Tables

List of Figures

Research Thesis: Declaration of Authorship

Print name: Alex Stokes

Title of thesis: Network Analysis of Simulated and Real Indigenous Irrigation Systems

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

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. None of this work has been published before submission.

Signature: A. Stokes

Date: 05/05/2020

Chapter 1

Introduction

Agriculture is often regarded as a key human innovation enabling emergence of civilisation (Lev-Yadun et al., 2000). The maintenance of this food supply is a primary driver of stability and persistence for any society through time. Given this importance, a large area of research has been devoted to the study of societal persistence.

Social-Ecological-Systems (SES) is a term to describe the coupling and distinction between humans and the environment (Levin et al., 2012). Persistence of this coupled system depends on the stability of both the social and environment systems. The coupling of SES is studied over a range of time-lengths and spatial-scales from small-scale traditional systems to large-scale modern systems. SES have traditionally been studied qualitatively recording the institutions and cultural practises which have evolved in a particular system enabling stability (Ostrom, 1990).

Indigenous Irrigation Systems (IIS) a subset of SES, focus on the water sharing aspect of self-sustaining traditional SES for agricultural purposes. IIS across the planet have found ways to persist for prolonged periods through the construction of robust water transport networks and SES institutions. Some of the most prominent examples of IIS can be found as part of the rice terraces in South East Asia. Specific well-studied examples that are still in operation today include the Subak System in Bali, Indonesia and the Banaue Rice Terraces in Ifugao, Philippines. IIS are also present across the world, in South Asia, the Middle East, Africa and South America.

The Subak System in Bali is perhaps one of the most studied IIS and one of the first to be quantified into a computer model (Lansing and Kremer, 1993). However, these pioneering computer models were restricted by both computational power and a lack of modelling tools and frameworks to conduct research. This led to the first models being predominately static and not accounting for growth and evolution of the system. There is continual work on the Subak System up to the present day to further understand whether it was organised by a central authority or through local decentralised community initiatives (Lansing et al., 2009) and how it maintains an optimal state through feedback between the social and ecological systems (Lansing et al., 2017).

Advances in complex systems theory and network theory have discovered universal laws across disciplines. Universal scaling laws have been uncovered within systems in nature and society in terms of their space-filling organisation and size distribution (West et al., 1997; Batty, 2009), but importantly the scaling laws are different for biological and sociological systems (Bettencourt et al., 2007). Related to this is the ubiquity of power law distributions in scale-free networks across nature and society (Barabási and Albert, 1999). A system, when modelled as a network and having a power law (or skewness) in terms of the connections of each node is now often seen as a signature of a self-organising unregulated natural network (Wang et al., 2019). For a network to have such a signature it has been found that both growth and preferential attachment are required for this signature to emerge (Barabási and Albert, 1999).

Alongside (and also in part enabling) these scientific discoveries are the advancement of computer performance and associated simulation tools. High performance computing power has become much easier and cheaper to access and computer simulation tools have been built and optimised for the purpose of modelling complex systems and networks. This has allowed for scientific simulation experiments of many complex interacting components that is common across all scientific disciplines.

This research aims to apply Generative Agent Based Modelling (Epstein and Axtell, 1996) to grow an artificial society of an IIS. An agent can be a physical or virtual entity that can act, perceive its environment (in a partial way) and communicate with others (Ferber, 1999). The 'Budding Model' is a hypothesised model for growth in the Subak System whereby local communities expand based on local conditions which in turn leads to the system as a whole reaching an optimal state (Lansing et al., 2009). This can be thought of as a form of space-optimisation to grow and increase available space for the IIS to distribute water. However many other factors influence IIS growth such as geology, geomorphology, ecology, climate and culture. Space-optimisation is commonly observed in biological systems and one of the factors giving rise to universal scaling (West et al., 1997). Different growth algorithms are tested on the model to see which is optimal for space-optimising and computational efficiency. An algorithm which is computationally inefficient in this analysis is comparable to an ineffective social institution.

Simulating the growth of an IIS forms an irrigation network which when compared to universal network scaling laws and real-world IIS networks can show whether the simulation follows the same behaviour as found in the real world and other space-optimising systems in nature. If the simulation is shown to be scale-free then the model can be extended and applied to larger social systems.

This thesis includes the following chapters:

- Chapter 2 - Literature review;
  Summary:
  - A definition of an indigenous irrigation system (IIS) is provided.
  - IIS are present in many locations throughout the world.
  - Commonalities and differences between these systems are outlined and the factors leading to these are discussed.
  - Detailed information on the Subak System in Bali and the Qanat system in the Middle-East is provided.
  - The idea that the network structure could be an emergent property of many contributing factors in each system is discussed.
  - The managerial arrangements are highlighted as one factor of interest. Local versus global optimal growth, individual versus community and planned verses self-organised are identified as interesting trade-offs in different systems.
  - Different methods to study IIS include interviews, field mapping, remote sensing and simulation modelling.
  - IIS are changing in the modern world, which makes it difficult to study them.
  - Research gaps highlighted.
  - Definition of complex systems and networks and an introduction to theories regarding their growth and stability.
  - Examples of previous complex system modelling techniques and their application such as equation based models, cellular automata, agent based models and network models.
  - Examples of other networks across disciplines that have commonalities with IIS.
- Chapter 3 - Research questions, aims and objectives;
  Summary:
  - Research Questions deriving from the research gaps found in the literature review.
  - Aims and objectives added to achieve answering these questions.
- Chapter 4 - Space Optimising Growth in Simulated Indigenous Irrigation Systems;
  Summary:
  - This chapter simulates a space optimising network in order to understand how managerial rules lead to different network structures.

- Network growth based on local deterministic or local stochastic rules produce networks with low space optimising capabilities.
- Network growth based on global selection or global stochastic rules produce networks with high and optimal space optimising capabilities.
- Comparison with other networks such as rivers find similar scaling exponents.

- Chapter 5 - Empirical Data and Analysis on Indigenous Irrigation Systems;
  Summary:
  - Data is collected from two real world IIS networks - The Subak system in Bali and the Qanat system in Iran.
  - Data was collected through remote sensing using high resolution images.
  - The irrigation networks were then analysed using the same scaling exponent as the simulated networks.
  - The network structures formed by the two networks were noticeably different and comparable to the simulated networks in Chapter 4. This gives some supporting evidence to the managerial rules in operation for each of the systems but it is by no means conclusive.
  - Other factors may also play an important role which are not factored into the simulation such as geomorphology.

Chapter 2

Literature Review

The literature review has been broken down into the following Sections:

- Section 2.1 introduces the field of study of Indigenous Irrigation Systems including definitions, case studies, general models of IIS and research gaps.
- Section 2.2 reviews scientific schools of thought relevant to studying IIS focusing on interdisciplinary tools such as complex systems and network theory.
- Section 2.3 focuses on the many different methods for modelling complex systems, some successful examples of their application and the methods which are most useful for studying IIS.
- Finally, Section 2.4 looks at other systems with similar processes and properties to IIS.

2.1 The Study of Indigenous Irrigation Systems (IIS)

2.1.1 Introduction

The focus of this research is on the growth, structure and stability of Indigenous Irrigation Systems. This first section gives an overview of what an Indigenous Irrigation System is, examples of IIS, gaps in understanding IIS and approaches for studying IIS.

2.1.2 Definition of an Indigenous Irrigation System

An Indigenous Irrigation System can be defined as (Groenfeldt, 2004):

1. The physical structure of water capturing devices (diversion weirs, dams, or wells), conveyance devices (canals, aqueducts, tunnels, flumes), and control structures (gates, outlets, dividers) by which water is delivered to agricultural fields.

2. The management arrangements for designing, constructing and maintaining the physical works, allocating and distributing water among the users, resolving disputes, and addressing emergencies or other unforeseen circumstances.

Indigenous can be defined as 'Born or produced naturally in a land or region; native or belonging naturally to (the soil, region, etc.)' (Oxford University Press, 2015). Indigenous Irrigation Systems represent an applied aspect of Indigenous Knowledge (IK); a systematic body of knowledge acquired by local people through accumulation of experiences, informal experiments and intimate understanding of the environment in a given culture (Rajasekaran et al., 1993).

IIS have been described as the most ancient and ubiquitous example of public infrastructure system (Yu et al., 2015). They represent not only a human technical achievement increasing yields (Hunt et al., 1976), but have major social consequences (Kelly, 1983) and possibly enabling the urban revolution (Hunt et al., 1976).

2.1.3 Example of Indigenous Irrigation Systems

Over the last few thousands of years, Indigenous Irrigation Systems have developed in different forms with the common purpose of distributing water for growing the food of people living a predominately sedentary lifestyle. IIS can therefore be viewed as a key component in the emergence of many civilisations.

IIS exist or have existed across the planet, including the Hohokam system in Arizona, USA (Murphy, 2012), irrigation in the Miju catchment, Yunnan, China (Crook and Elvin, 2013; Elvin, 2002), Suranga Irrigation in South Karnataka and Northern Kerala, India (Crook et al., 2013), Wadi Faynan, Jordan (Crook, 2009), the Pumpa Irrigation System in Nepal (Cifdaloz et al., 2010), the Indigenous Water Management System in Bhaktapur City, Nepal (Gautam et al., 2018), the Subak System in Bali, Indonesia (Lansing, 1987), Pokot, Northwest Kenya (Davies, 2008), the hill-furrow irrigation system of the Marakwet Escarpment, Kenya (Watson et al., 1998; Davies, 2009), the Qanat System, Iran (Bonine, 1996), The East Mitidja scheme, Algeria (Laoubi and Yamao, 2009), Karez Irrigation, Maywand District, Kandahar Province, Afghanistan (Egitto, 2013), Acequia irrigation communities, Taos Valley, New Mexico, USA (Cox and Ross, 2011), Mount Kilimanjaro, Tanzania (Gillingham, 1999; de Bont et al., 2019), Sonjo Irrigation, Tanzania (Adams et al., 1994), Wadi Laba Spate Irrigation, Eritrea (Mehari et al., 2005), Kuhl irrigation, Kangra Valley, Western Himalayas (Baker, 2005), Andean Irrigation Systems, Peru (Trawick, 2001), the Tank Cascade Systems, Sri Lanka (Geekiyanage and Pushpakumara, 2013) and the Ifugao Irrigation System, Northern Philippines (Araral, 2013).

These systems operate at many different scales, the Ifugao system, northern Philippines occupies 4,000 km^2 of terraced slopes (Araral, 2013) whereas the Qanat irrigation system may support a single small village (1km^2 (see figure 2.1)) (English, 1998).

2.1.4 Case Studies of Indigenous Irrigation Systems

The following paragraphs give further detail of different IIS located around the world.

Qanat System, Middle East

Qanats are one of the most significant hydraulic technologies of the pre-modern arid desert environments in the Middle East (English, 1998). Given the harsh desert conditions, there is little surface water. To counter this the an underground pipe called a Qanat is dug to transfer water from highlands to villages for agricultural use. A well shaft usually 50m deep, is sunk into the groundwater recharge zone usually in sediments of alluvial valleys at the base of highlands. A tunnel larger enough for a person to walk through transports the water down a shallow gradient under the force of gravity to the village (Bonine, 1996). Shafts are dug from the surface down to the Qanat tunnel every 50 to 100m to allow for Qanat workers to breathe, excavate material and provide access for repairs (English, 1998). Most Qanats flow as a direct response to precipitation, reflecting the permeability of the rock and soil into which the groundwater recharge occurs. A schematic sketch of a single qanat flowing into a village and distributing water for agricultural use is shown in Figure 2.1, (Bonine, 1996).

The origins of the Qanat system are thought to date back 2,500 years in the mountains of Kurdistan (English, 1998), although specific examples of these systems may not persist for such long periods of time, or abandoned and re-

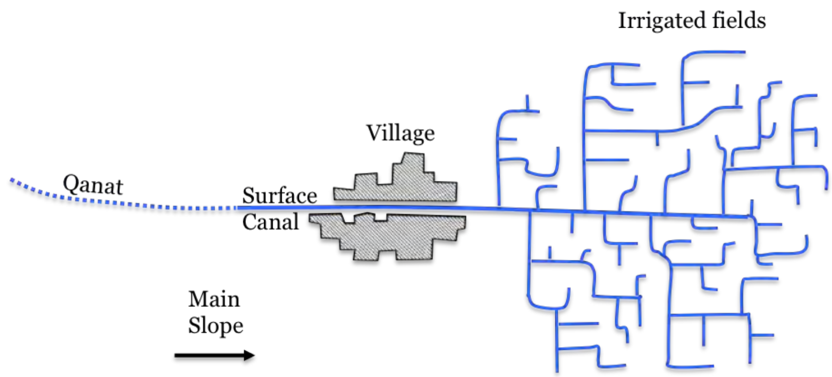


Figure 2.1: A schematic example of the Qanat system from (Bonine, 1996).

established at a later date. Variations of the Qanat system are present in Afghanistan (Karez) (Egitto, 2013), Morocco (Khettara) (Faiz and Ruf, 2010), Iran (Qanat) (English, 1998), North Africa (Fughara) (English, 1998), Syria (qanat Romani) (Lightfoot, 1996) and Spain (Galerias) (English, 1998). Qanat technology is also present in Central and South America in Mexico, Peru and Chile (English, 1998), Central Asia and western China (English, 1968).

Figure 2.1 shows a simplified example of a Qanat system, in reality many of the systems are much more complicated with multiple qanats flowing into groups of villages, Figure 2.2 (Faiz and Ruf, 2010).

The construction of a qanat depends on two decisions by the Qanat Specialist (*Muqanni*); the location of the mother well (*Madari*) which is the furthest point from the settlement and the slope between the *Mardari* and the settlement (English, 1998). The *Muqanni* poses indigenous knowledge of the area allowing them to decide based on local slope conditions, topography, vegetation, groundwater and the proposed destination.

Qanats, like all IIS require a nexus of ecological and social conditions including high levels of social cohesion in terms of water allocation (English, 1998). The managerial arrangements of the Qanat societies have been described as individualistic (Geertz, 1972).

The Subak Irrigation System, Bali

The Balinese Subak irrigation system is thought to have been in operation for over 1000 years (Lansing, 1996). The main staple food in Bali, rice, is grown in paddy fields fed by rainfall guided irrigation channels (Lansing and Kremer, 1993). Direct evidence for rice cultivation in Bali is from rice phytoliths (rigid, microscopic structures made of silica, found in some plant tissues and persisting

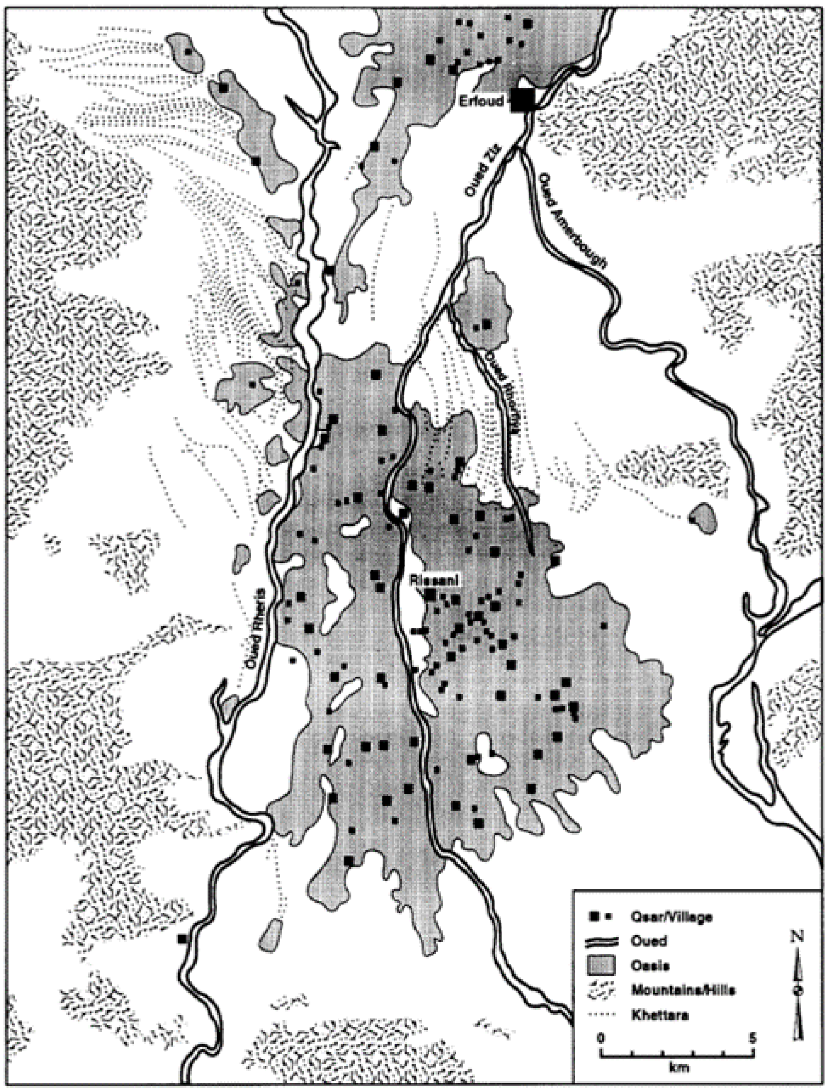


Figure 2.2: Extract from Faiz and Ruf (2010) showing multiple Qanat channels flowing into a large collection of villages.

after the decay of the plant) found in the sediments dated to 1 AD (Lansing et al., 2009). There are two main technologies used for irrigation expansion in pre-colonial Bali. Small weirs are constructed upstream, and larger storage dams are created downstream (Lansing et al., 2009). It has been extensively studied (Janssen, 2007; Lansing and Miller, 2005; Wijermans and Schlüter, 2014) and is thought to provide an example of a sustainable, resilient community managed agricultural system. However this is disputed, Wittfogel (1957) suggests central bureaucratic organisations are necessary to coordinate a network, whilst Lansing (1991) suggests they can self-organise through local interactions.

Geertz (1980) provides some of the first detailed and systematic reviews of the Subak System. The area which the Subak system covers is approximately 3,450 km^2 on steep slopes formed of volcanic sediments. At the crest of the slope are three volcanic cones 1500-3000m high. The slopes are defined by deep-cut river gorges running from the crest down to the sea. Geertz (1980) focuses on one example village of the Subak - Tihingen. The society consists of settlement units termed *Bandjar* and agricultural units termed *Subak*. The Subaks adjacent to a particular Bandjar tend to collectively owned by the villagers. The geomorphology tends to have a strong control over the shape and form of both the Bandjar and Subak. The river gorges form reasonably linear channels down the slope, the Bandjar and Subak are located on the spurlines in-between, presumably it is the driest area and least susceptible to flooding. The size of the system means that downstream Subaks suffer losses from percolation and evaporation, however downstream Subaks routinely take advantage of excess flows from neighbours and local springs (Lansing et al., 2009).

The work of Stephen Lansing in Bali provides some of the first and best examples of the application and usefulness of simulation for gaining insight into social-ecological systems in particular indigenous agricultural systems. His work mainly focuses on the social aspects of how cooperation can emerge in a system without any centralised control mechanism (Lansing and Miller, 2005). The simulations model 'Subak' scale relationships, where a Subak is a local-level farmer's association. These are the level at which decisions over water allocation are made (Lansing, 1987). For instance the district of Badung, an area 115 km by 40 km consists of 151 Subaks. The strong dependence on rainfall for rice cultivation leads to difficulties given its variability both spatially and temporally on Bali (probably partly due to the steep topography present on the island). Rice pests can also have a strong effect on rice yields. Stability of these two factors seems to be the main reason for the emergence of the system (Lansing and Miller, 2005). If adjacent Subaks synchronise their fallow periods they reduce the amount of pests thus reducing the damage they could cause. Due to constraints on the water available, only a limited number of Subaks can synchronise their fallow periods leading to a decentralised system of localised coordination of fallow periods. Lansing and Kremer (1993) and Lansing and Miller (2005) construct a numerical model of this synchronisation finding that cooperation over water increased with increased pressure of pests.

It is uncertain whether similar cooperation emerges in other IIS (Janssen, 2007).

To understand how such a system might evolve, Lansing et al. (2009) propose a budding model whereby canal irrigation systems expand downstream as a result of local initiatives, describing it as being an example of a Complex Adaptive System (CAS) - expansion, stability and persistence depend on reactions to environmental and social processes. This is an alternative to a centrally planned model (Wittfogel, 1957), which is governed by centralised bureaucrats. In order to test the budding model genetic data was collected from farmers from different Subaks. The results of the genetic data found that Subaks located further upstream in their irrigation system contained greater genetic diversity than those downstream suggesting they were created first, giving support to the budding model.

In contrast to the Qanat system, the managerial arrangements of the Subak system are said to be community based (Geertz, 1972).

2.1.5 Factors influencing the evolution of Indigenous Irrigation Systems

The difference in geology, geomorphology, climate, ecology and culture create endless novelty in how IIS evolve. These interwoven factors influence the emergent properties of IIS.

Geertz (1972) highlights this in his seminal work *The Wet and the Dry*, a comparative study of traditional agriculture in Bali and Morocco. Bali has a tropical climate, plentiful water supply, with a highly collective approach to organisation of irrigated fields, whereas Morocco, an arid country has a much more individual, property based approach to water regulation.

The physical irrigation network structure of the IIS, could be particularly important in containing details of the emergent properties from factors influencing the system. As IIS grow and persist for longer periods of time, the network is likely to become more efficient and the dominant factors more apparent. For Qanats, digging channels underground allows access to water as their is little present at the surface, for more efficient water transfer and reducing evaporation. Whereas in Bali, there are large amounts of rainfall and surface water, so there is not a need to dig underground tunnels.

As water can only usually flow downhill without sophisticated technology which is not available for IIS, the majority if not all IIS are located on or at the base of a slope, (English, 1998; Araral, 2013; Lansing et al., 2009). Geomorphology will therefore play a key role in where IIS are located. The geology directs the groundwater flow and is also a controlling factor on the geomorphology. Climate will have a large affect as a continual supply of water is required to sustain a population. A seasonal rainy season would mean that reservoirs would be more effective for storing water in times of droughts such as the tank system in Sri Lanka (Geekiyanage and Pushpakumara, 2013). Finally the ecology will affect the types of crop which are available and the pests which cause poor harvests (Lansing and Kremer, 1993).

The Pokot irrigation system in Kenya combines a nomadic lifestyle with an irrigation system (Davies, 2008). Mapping of the irrigation networks shows that communities move between locations, perhaps to escape soil degradation or climatic changes.

The Ifugao and Subak Systems experience large amounts of rainfall and can evolve on steep terrain through the use of terraces to maximise the land where water supply is plentiful (Araral, 2013). The cultural factors such as collective or individual organisation are also likely to have an effect on physical network, however given IIS are often community based, strong social cohesion is enduring (English, 1998; Ostrom, 1990).

2.1.6 The Influence of Managerial Arrangements

One prominent uncertainty regarding Indigenous Irrigation Systems is validating the managerial arrangements for how they have been set up or organised (Lansing et al., 2009). A more hierarchical system, will have greater central control (Wittfogel, 1957) over the growth and organisation, whereas a decentralised system is more likely to have preferential locally driven growth by different parts of the system in a so called 'Budding Model' (Lansing et al., 2009). This is a large area of research, in particular studying how IIS can evolve to a stable state through decentralised management as this contradicts a lot of modern economic theory, which suggests privatisation is the only way of sustaining common pool resources (Hardin, 1968). A number of rules have been proposed which appear to be common across most systems persisting for prolonged periods (Ostrom, 1990). In contrast, the IIS systems in East Africa the social institutions employ corporate power, where large segments of the population act against monopolisation, however this is not an egalitarian society as each group has a varied amount of power (Davies, 2009). This seems like a form of tribalism.

However, the collection of data to validate the managerial practice under which IIS grow can be very difficult and open to multiple interpretations. Genetic markers of populations located within IIS have been collected in the Subuk System to offer support to the 'Budding Model' (Lansing et al., 2009). The findings suggest that the 'Budding Model' is correct as populations with similar genetic markers can be found downstream of one another, giving an argument that the system has grown based on local initiatives.

The managerial arrangements may affect the network properties of the IIS. A planned network would be expected to take a more ordered shape, whereas a system which grows through a decentralised (self-organising) process is likely to be more random (or natural). However this has not been explored in the IIS literature and represents a research gap. A comparison can be drawn with modern urban systems, which when planned and built in a short space of time, tend to form regular patterns such as the majority of cities in the USA, whereas cities which have developed over much longer periods in a less planned manner tend to have a more random pattern, such as UK cities.

2.1.7 Growth and Persistence in Indigenous Irrigation Systems

Many IIS have been operating for considerably long periods of time. The Ifugao System and Subak System are thought to both be approximately 2,000 years old (Araral, 2013; Lansing et al., 2009). The origins of the Qanat system are thought to date 2,500 years in the mountains of Kurdistan (English, 1998), although specific examples of these systems may not persist for such long periods of time, or might be abandoned and re-established at a later date.

Given IIS are not modern systems, their growth and persistence might be similar to biological systems, which grow following a sigmoidal curve to an eventual population limit. Archaeological evidence of the Subak System has found that older weirs are located further upstream (Lansing et al., 2009), but there is little information available about the rate of growth. The managerial arrangements of each IIS may also have an effect on this growth rate. Unplanned systems may follow feedback based growth, reacting to the changing factors influencing the efficiency of the system (Lansing et al., 2009) which in turn would be reflected in the age of parts of the system. A planned system may be set up in one go based on indigenous knowledge of the planner so the growth would unlikely follow a sigmoidal curve.

The large area covered by the Subak System shows that if water is not a limiting factor then it may be limited by geographical space, whereas a small qanat system in a very hot arid area is likely to be limited by water availability. However both systems may still grow following a sigmoidal curve.

2.1.8 Inequality in Indigenous Irrigation Systems

Indigenous Irrigation Systems form part of the earliest civilisations; there has been an ever-present water allocation dilemma leading to differences in inequality. Inequality over long periods leads to societies becoming stratified (Davies, 2009), which is one of the defining characteristics of civilisations in general. Conflict between different parts of society is one postulated reason for societal collapse (Araral, 2013; Motesharrei et al., 2014).

The managerial arrangements implemented in different IIS may lead to different levels of inequality. For example, the Subak System in Bali has a community based managerial model, the crops harvested from the subaks surrounding a village is shared amongst the village, whereas the irrigation systems in Morocco have a much more individual based approach (Geertz, 1972). Logically it may be the case that inequality can develop more readily in an individual based system than a community based one, but the Subak system comprises of many communities; which might show different levels of inequality.

Inequality in terms of intergenerational wealth transfer has previously been studied in traditional agricultural systems (Borgerhoff Mulder et al., 2009; Bowles et al., 2010), finding lower intergenerational wealth transfer in hunter-gatherer and horticultural populations, whereas for pastoral and agricultural societies this is higher. This is thought to be due to the higher amount of material wealth in pastoral and agricultural societies which is past through generations allowing wealth to accumulate in certain families or communities through time. This is also related to ownership in society (Geertz, 1972). IIS can be both horticultural and agricultural so the inequality may vary between systems under study.

Yu et al. (2015) explores a model of how irrigation system design can lead to different levels of inequality and maintenance between two villages connected by a shared water supply. Two design variations are explored, upstream-downstream access and equal common pool access. This takes a game theoretic approach of farmers being either opportunists and conformists in terms of using water and maintaining the infrastructure. Exploration of the model finds that the equal common pool access model is only stable if the majority of farmers are conformists. Both systems collapse if all farmers have an opportunistic strategy (they do not contribute to maintaining the infrastructure). Multiple basins of attraction exist when there are different ratios of opportunists and conformists, and also if the maintenance costs are high. In general, the model shows similar behaviour to other models of cooperator-defector behaviour; increased conformist behaviour leads to lower inequality and more efficient infrastructure, whereas increased opportunistic behaviour leads to higher inequality and lower efficiency infrastructure.

Other models have looked at emergence of inequality in a generalised society through agent based and equation based models. Motesharrei et al. (2014) introduce a model of a society using an extension of the predator-prey equations for groups of humans with different characteristics (elites, commoners, workers and non-workers) and the environment. The model shows outcomes where high levels of inequality can lead to collapse of the society. *Sugarscape* is an agent based model which simulates the origin of inequality in artificial societies by assigning a distribution of behaviours to agents and then simulating their interaction through time (Epstein and Axtell, 1996). This model has found similar distributions of inequality which are present in modern societies. Agent based models of this kind have not yet been applied specifically to IIS.

2.1.9 A General Model for Indigenous Irrigation Systems

Despite stark differences in IIS, the definition of Indigenous Irrigation Systems (Section 2.1.2) shows that they have two main components in common - a (spatial) physical network of channels distributing water and (social) management arrangements. This is a universal solution to the problem of supplying water to a sedentary human population.

Physical IIS networks appear to be space-optimising; in general they grow to use all available space for agriculture (given social and environmental constraints). Given they are distribution systems, they form a branching hierarchical network, with reduced channel size down the system. These two properties have been found previously to be common properties of biological systems (West et al., 1997), but not in IIS.

Biological distribution networks tend to be very efficient, for example the parent branch cross-sectional area is the sum of the daughter branches (West et al., 1997). It is already known this is not the case for IIS. The size of branches has also reportedly been largely dependent on the technology available to create them (Lansing et al., 2009). In particular for the Qanat System the canals are large enough to fit a person through, to dig them but also for maintenance (English, 1998).

The structure of the physical network formed by IIS tends to form tree networks, Figure 2.1. This is not surprising given they are distribution systems. The changing properties of these tree networks between IIS and the reasons for this has not yet been analysed. It may contain a large amount of information on the evolution of an IIS. A general simplified model of an IIS network will offer a base which can then be compared to the many different types seen in reality.

The idealised example of an IIS which has organised dominantly by managerial constraints is a system which has a single input and organised without outside intervention to reach a homoeostatic state with minimal environmental factors. For example, the climate would be reasonably constant, leading to a constant inflow of water, the geomorphology would be simple, a shallow planar slope for example and the ecology would have little effect, such as no pests. In reality there are no systems which organise in such as minimal fashion, but by understanding this ideal example, strong emergent features can be found which may not be as obvious in real-world systems. The network formed by this idealised example, are likely be planar for maximum efficiency (the edges between nodes do not overlap), and rooted (it grows from one point) which would define it as a rooted tree network.

There are two end points in the managerial organisation, one which is controlled by a central power (Wittfogel, 1957) and one which organises through local initiatives (Lansing et al., 2009). Two models giving properties of a planned ordered irrigation network and a self-organised local network can then be compared against real-world systems. But there are also differences in managerial arrangements depending on if the system is divided into individual agents or community based.

2.1.10 Indigenous Irrigation Systems in the modern world

Given the far reaching effect of globalisation and increased connectivity of the modern world, there are few IIS left untouched to study. Most IIS today use a mixture of indigenous and modern practices, and are linked by varying degrees to national, regional and global markets (Mabry and Cleveland, 1996). Due to the inherent complexity of studying real-world systems, particularly in the modern globalised highly interconnected world, simulation can play a more important role in developing simplified models of how systems of many agents organise. The power of this method over others is that the simulation can be thought of as being alive and evolving. Hypotheses can be tested on the simulation which would not be possible real systems.

The difference between Figure 2.1 and Figure 5.5 highlights this issue. Figure 2.1 shows an idealised drawing of the Qanat system which is isolated with no connectivity to the outside world, whereas Figure 5.5 shows a real world system found in the modern day with modern tarmacked roads connecting the IIS to nearby cities.

Modern technology also has an influence. English (1998) outlines the change that is happening in the region with respect to the Qanat system. Population growth and agricultural expansion have heightened demand for water. This has led governments to abandoning the indigenous irrigation practices for more productive modern pumping technologies and dams. In particular the Qanats are largely disappearing and being replaced by deep wells.

Increases in population related to globalisation and migration have led to many IIS changing over the past few hundred years. The Pokot System in Kenya has been moved to wet highland areas away from arid area in the past 200 years, the Pumpa system in Nepal was initiated in 1968 by migrant communities (Cifdaloz et al., 2010) and the IIS in Mount Kilimanjaro (Gillingham, 1999) which has extended downstream since the 1950s under population pressure.

The changes which occur to IIS from modern influence are profound and long lasting. Many studies are dedicated to gathering knowledge, conserving and enabling smooth transitions avoiding collapse of SES and ecosystems into new regimes in the modern world. Regime shift analysis is a tool predominately used to analyse ecosystems, lakes in particular (Carpenter and Bennett, 2011; Dearing et al., 2012). These have been applied to IIS, most notably in Bali (Lansing et al., 2014), which found evidence of alternative social states between Subaks, but no evidence of regime shifts suggesting transitions in traditional societies are likely to be rare. However, a much larger regime shift in IIS is taking place due to the effects of globalisation. In relation to water resources, Cole and Browne (2015) qualitatively assess the impact of tourism on water allocation in the Subak System, Bali finding that there is obvious indications that it is water resources are overstretched.

Quantitatively, different regime change behaviour has been found for SES (Dearing et al., 2014), when the boundary of a particular regime is reached. In an effort to quantify the changes occurring in IIS in the modern world, the framework should be applied to them. As it is very difficult to measure regime change behaviour in real-world systems (Lansing et al., 2014), models of IIS can be used at first to find the parameters under which change certain types of change occur.

Recent research has also looked at the effect of climate change on IIS, including adaptation strategies to mitigate against impacts. Kihila (2018) presents a review of adaptations globally and a focus specifically on communities in Tanzania. Adaptations and coping strategies include - changing farming practises and migration to less water stressed areas.

2.1.11 Approaches to studying Indigenous Irrigation Systems

IIS consist of many interacting components, which grow and evolve over time in response to both internal interaction and environmental influence. The previous section stated that the physical network IIS may hold information on the managerial arrangements organising the system. In order to study this phenomena, different scientific approaches are available.

Qualitative approaches can be useful for collecting information which is difficult to quantify such as the managerial arrangements through carrying out interviews or questionnaires for example. The physical irrigation network can be mapped through quantitative methods such as aerial photograph interpretation (API) or ground truthing in person through field mapping. There are many initial sketches of irrigation systems which give an indication of the general properties of certain networks although these are not accurate and unlikely to stand up to scrutiny (For example Figure 2.1).

Once reliable data has been collected, statistical approaches can provide a quantification of the different IIS networks. This can aid with classification of the network types, although will not be useful for linking managerial arrangements with the specific network without further information on the managerial arrangements. However unless the data has been collected through time, which is unlikely, it will not provide information on the growth of the system, which is key to understanding how the managerial arrangements lead to the network emerging.

Complex Systems, a quantitative inter-disciplinary field of study which seeks to understand how systems of many interacting components lead to emergent behaviour can provide a relevant framework and modelling techniques for studying IIS. Generative simulations of complex systems are useful for showing how simple interactions can lead to emergent large-scale behaviour. This technique would be very applicable for looking at how different managerial arrangements of the IIS can affect the network properties of the system grown, as other influential environmental factors are minimised. Also given the effects of the modern world on IIS, there are many difficulties with studying them. Computer simulations, a product of the modern world, provide an isolated environment in which IIS can be built, allowed to evolve and analysed. The output can be compared against snapshot data collected from real systems and provide arguments for and against whether the arrangements implemented in the model are reflected in reality.

2.1.12 Summary

This section has given an introduction into the study of Indigenous Irrigation Systems. Firstly, the definition of IIS consists of both the managerial arrangements and the physical distribution network. Despite this common definition, IIS are very varied in their arrangement across the planet. The main factors influencing their development are geology, geomorphology, climate, ecology and culture (or managerial arrangements). Two systems described in detail from contrasting parts of the world, the Subak System in Bali, and the Qanat System in Iran highlight the differences in IIS.

The following research gaps have been found in the literature:

- IIS seem to consist of space-optimising branching tree networks forming a hierarchical structure. Quantitative analysis of different IIS networks has not yet been carried out to confirm this, which in turn can be used to find commonalities and differences between IIS.
- No studies have looked at the relationship between managerial arrangements and network structure of IIS.
- There is no recognised model for how IIS tend to grow, do they follow a biological type sigmoidal curve, sociological expansion? Is this related to the managerial arrangements?
- Previous studies have looked at the development of inequality in simplified game-theoretic models of IIS, but not for larger network based simulations.
- No quantitative models have been built to look at the effects of globalisation on IIS.

2.2 Review of Relevant Schools of Thought

This section gives an overview of schools of thought relevant to studying Indigenous Irrigation Systems. These are focused on approaches for studying complex systems and networks as this is the chosen modelling method given the outcome of the literature review. These are high level frameworks which offer general models of how systems of many interacting components may behave. and are relevant to a systems based approach which can be applied to Indigenous Irrigation Systems.

2.2.1 Complex Systems

The field of complex systems tends to be an ill-defined concept, most likely due to the generality of the subject. A complex system may be defined as a system with many degrees of freedom or many interacting components, which interact in a non-linear fashion. The three body problem in the physical sciences is a starting point of when a system becomes complex as there is no closed form solution as there is for two bodies (Šuvakov and Dmitrašinović, 2013). Instead methods to solve the three body problem under particular conditions tend to involve numerical simulations with certain parameters (Li et al., 2017). The three body problem in itself is a large topic of study which this thesis will not delve into. However, it does serve as an example of how complex, chaotic systems can arise from the simplest of systems. In reality, most complex systems have many more interacting bodies than three, but a lot of the same theoretic and analytical frameworks are still applicable.

Bar-Yam (1998) states the field of Complex Systems seeks to increase our ability to understand the universality that arises when systems are highly complex and defines a complex system as one which consists of interconnected or interwoven parts. To put this into context, a simple or complicated system is one in which the whole is the sum of the parts; they act in a linear fashion. Therefore, to describe the system behaviour, each part can be summed together. In a complex system, not only do the parts need to be understood, but their interaction with one another too due to the inherent non-linearity present. A complex system therefore requires a holistic systems-based approach.

Most if not all natural systems are complex, but many assumptions are made such as linearising the interactions of the components in order to simplify them. This is usually valid for explaining the behaviour of systems within given constraints but not applicable when the system is close to a regime shift such as a pendulum swinging out of control or a social-ecological system nearing collapse (Scheffer et al., 2001).

When defining a complex system, the following components should be considered: (Bar-Yam, 1998):

Space The structure and spatial extent of the system. The structure may relate to its formation, its boundaries and how it interacts with other systems.

Time The length of time that dynamical processes take in the system. The ways in which the complex system changes in response to changes in their environment.

Organisation is present across all disciplines and a key concept/observation in any scientific endeavour. Two theoretical extreme fixed points can be thought of when thinking in terms of organisation. A self-organised system is a system which organises without any intervention. A planned system is one that is organised completely through external or centralised control. Neither of these pure forms exist in nature but any system lies on a spectrum between the two. Another spectrum involves top-down and bottom-up organisation, the degree to which bottom-up and top-down processes control the organisation of the system.

Complexity To what degree is the system complex. There are many definitions of complexity; one common definition is the amount of information needed in order to describe the system (Bar-Yam, 1998). Therefore more random systems will have a higher complexity, s these require more information to describe whereas a more ordered systems which can be described more easily will have a lower complexity. The fractal dimension can be used to measure complexity as it gives a precise metric for complexity and irregularity of a given object (Corbit and Garbary, 1995).

These general characteristics are important properties to take into account when studying complex systems, but they are very general, and need to be built upon when studying a particular complex system, for instance when looking at a particular social-ecological system, the different variables in the systems need to be accounted for and the time-scales over which they operate.

Relevance to IIS Indigenous Irrigation Systems consist of many interacting farmers which construct and manage a physical network through *time* and *space*. How IIS systems *organise* is difficult to measure and the source of great debate in the research community (Lansing et al., 2009; Wittfogel, 1957). IIS are thought to have different organisation, including centrally planned (Wittfogel, 1957), community managed (Lansing et al., 2009) or based on individual decision making (Geertz, 1972). The complexity of IIS can be difficult to measure as managerial arrangements are difficult to quantify. The physical IIS network may be easier to measure and the fractal dimension being one such way of measuring this.

2.2.2 Scaling in Complex Systems

Many complex systems are scalable, they exhibit common (fractal) properties at a range of scales (Mandelbrot, 1983). A system which exhibits scaling, can be divided up into levels, with each level being a different scale. Levels are often thought of in terms of hierarchy for instance in a company or even academic organisation. However, they may also be thought of in terms of a "container view" such as minutes are a finer level to hours, but form part of hours and also an "emergent view" whereby processes arise at a higher level through interactions taking place at a lower level (Wilensky and Resnick, 1999).

Scaling has been found to change for different systems (Bettencourt et al., 2007), but also dependent on the way it is measured (Bettencourt et al., 2007; Barabási and Albert, 1999).

Living biological systems have been found to follow universal scaling laws, referred to as allometric scaling (West et al., 1997). An explanation for this scaling is proposed to be due to the transport of materials through space-filling fractal networks of branching tubes (West et al., 1997). In the model, these tubes are area-preserving, the cross-sectional area of a parent branch is equal to the sum of the daughter branches, which gives rise to scaling exponents (β) for different biological variables as shown in in the following equation:

$$y = kX^{\beta} \tag{2.1}$$

where Y is a biological variable dependent on body mass X and k is a constant, and β is the scaling exponent.

Based on the model, predictions of many biological variables including metabolism, capillary radius and capillary density are almost identical to empirical data (West et al., 1997). The scaling exponents (β) have been found to scale at quarter powers, which is due to the area preserving nature of the model and are all β<1 (West et al., 1997).

The equation has been extended to study resource scaling in a city of a given size through time (Bettencourt et al., 2007):

$$Y(t) = Y_0 N(t)^{\beta} \tag{2.2}$$

where,

Y can be a material resource such as energy or infrastructure or a measure of social activity such as wealth or pollution at time (t), Y_0 is a normalisation constant, $N(t)$ is the population at time t and scaling exponent $^{\beta}$ reflects the general dynamic rules across the system. The addition of time into the equation is of importance as many systems will change as they grow. This is true for both cities and biological systems.

Empirical data of cities found that infrastructure (such as petrol stations) scales with an exponent $\beta<1$, but sociological data (such as wealth or patents) scales with an exponent $\beta>1$ (Bettencourt et al., 2007). When plotted through time, variables with $\beta<1$ will have sub-linear growth reaching a carrying capacity (as is the case for biological systems) and variables with $\beta>1$ will have super-linear growth leading to accelerating growth. This seems to be due to spatial constraints acting on biological and infrastructure systems.

Figure 2.3 shows equation 2.1 plotted with k=1 and $0.1<\beta<1$, showing sub-linear to linear growth (at $\beta=1$). Figure 2.4 shows equation 2.1 plotted with k=1 and $1.0<\beta<2.0$, showing increasing super-linear growth.

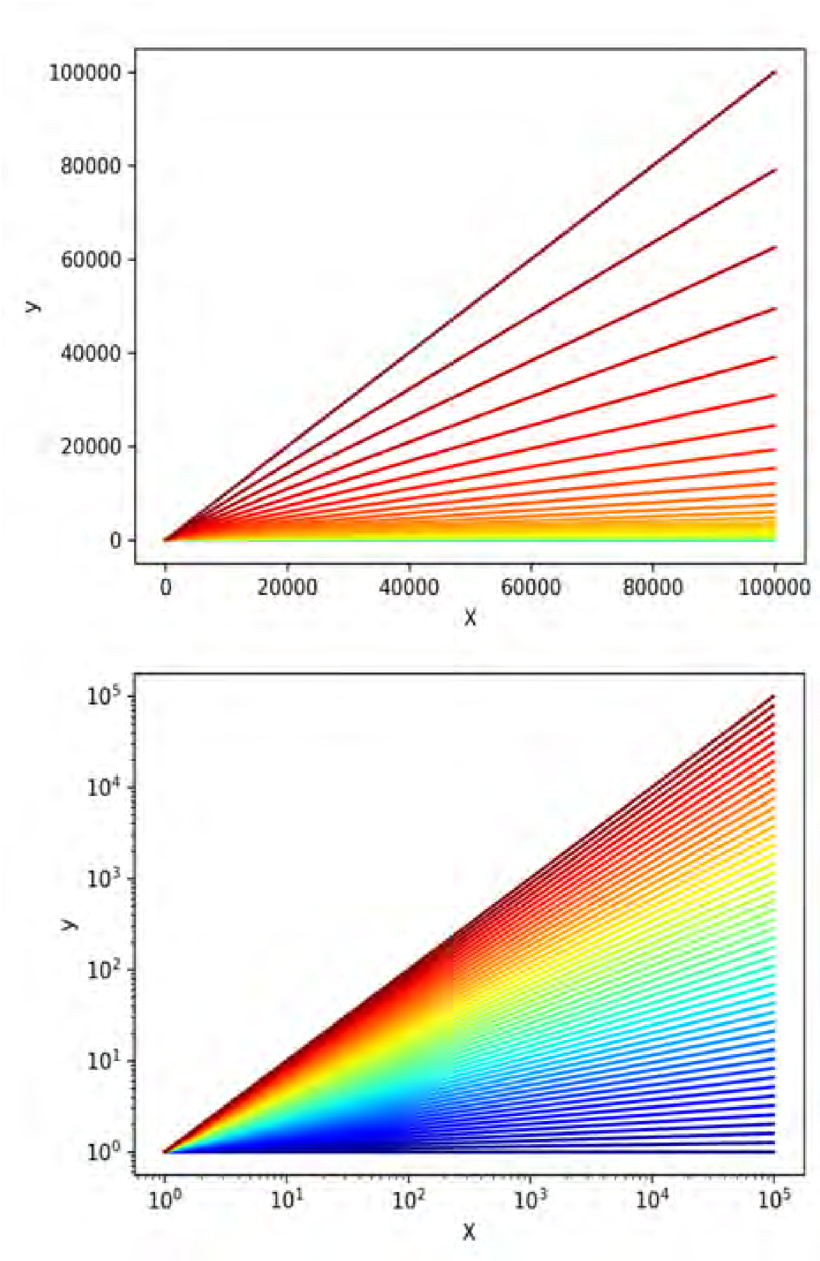


Figure 2.3: Plot of equation 2.1 with k=1, and $0.1<\beta<1.0$ (blue to red). *Top*:Plot on regular graph. *Bottom*: Plot on log-log graph.

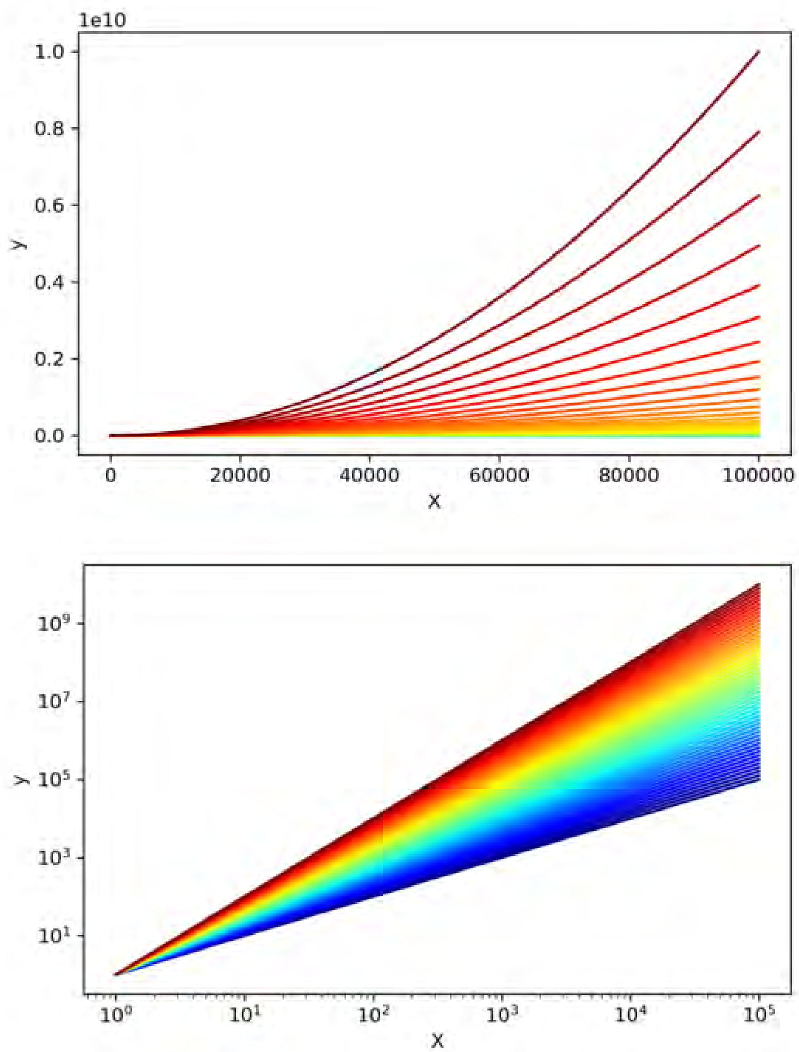


Figure 2.4: Plot of equation 2.1 with k=1, and $1.0<\beta<2.0$ (blue to red). *Top*:Plot on regular graph. *Bottom*: Plot on log-log graph.

Similar research has been conducted on abstract tree branching processes, more detail is given in Section 2.4.3. Tree networks which are spatially constrained (river systems or plants) are limited to 'critical' behaviour, whereas non-spatial *mean-field* tree networks (such as the internet) can behave super-critically (see Section 2.4.3 for definition). A different way of measuring the exponent is used in networks using the probability distribution of nodes downstream (Caldarelli et al., 2000; Jun and Hübler, 2005; Barabási and Albert, 1999).

IIS form a network of channels which distribute water. The research cited in this section studied biological systems and cities, finding scaling laws apply to both but with different exponents. There are many commonalities with the transport systems which exist in both biological and urban system and Indigenous Irrigation Systems. IIS in some respects can be viewed as a gap between purely biological systems and social-economic systems, and the scaling of these systems might be reflect this.

2.2.3 Complex Adaptive Systems (CAS)

CAS are a subset of Complex Systems, more generally synonymous with the life sciences (Gell-Mann, 1995; Holland, 1992; Lansing, 2003). In CAS the patterns at a higher level emerge from the interaction and selection at a lower level (Levin, 1998), a 'Bottom Up' process. They are adaptive, implying that the system can survive changing conditions. They are said to share three characteristics: evolution, aggregate behaviour and anticipation (Holland, 1992) suggesting a CAS grows memory or wisdom over time. The term has been applied to Indigenous Irrigation Systems in the past (Lansing, 2003) as their dynamics seem to resemble other more fundamental CAS such as the Daisyworld model (Lovelock and Margulis, 1974; Lansing et al., 1998). Given there are many factors which influence the evolution IIS, which may fluctuate over time, for the the IIS to persist, it needs to adapt to these perturbations. Adaptive behaviour should therefore be considered when studying, constructing models and analysing IIS.

2.2.4 Resilience and Stability

Defined by Holling (1973) as 'determining the persistence of relationships within a system and is a measure of the ability of these systems to absorb changes of state variables, driving variables, and parameters, and still persist'. The resilience view of the world contrasts that of a stable one. A stable view emphasises equilibrium and the maintenance of the predictable whereas a resilience perspective is concerned with domains of attraction and the need for persistence. Definitions of resilience and stability tend to be similar (McCann, 2000), it is often the trend of a particular discipline which leads to one becoming more favourable. Resilience is often used more in the context of ecology (Dakos et al., 2015), whereas stability is more frequently used in mathematical analysis (Strogatz, 2015).

Other terms which are generally synonymous with resilience and stability are robustness-fragility and vulnerability. Robustness is commonly defined as 'the maintenance of some desired system characteristics despite fluctuations in behaviour of its component parts or its environment' (Carlson and Doyle, 2002). Robustness-fragility is often concerned with the trade-offs between increasing the robustness of a system for a set of problems whilst this will increase the fragility to others. In engineering, a robust system will typically perform as efficiently with respect to a chosen set of criteria than its non-robust counterpart, however its performance will not drop off as rapidly when confronted with perturbations (Anderies and Janssen, 2011).

Vulnerability is often described as the degree to which a system is susceptible to, or unable to cope with, adverse effects such as of climate change, including climate variability and extremes to moderate damages, to take advantage of opportunities, or to cope with the consequences (IPCC, 2001).

Differences between robustness and resilience are the extent to which (non-structural) changes in dynamics may be introduced into a system under the impact of perturbations (Young et al., 2006) although other authors do not make a definite difference between them (Fleischman et al., 2010). Vulnerability refers to situations in which neither robustness nor resilience enables a system to survive without structural changes, in such cases either the system does adapt structurally or it is driven to extinction (Young et al., 2006).

Resilience and stability are important topics for studying IIS as they often persist for long periods of time and must withstand changing conditions in order to persist. Complex adaptive systems must have resilience. The resilience can be thought of as the pathways which the CAS can move between in order to persist.

Measures of Stability

Stability (and resilience) is often measured both quantitatively and qualitatively in the form of fixed point analysis (Strogatz, 2015; Scheffer et al., 2012). This is particularly useful when applied to non-linear problems with many interacting components, like most real world systems. Fixed points can either be unstable or stable and represent maxima or minima on a landscape which describes a system. This landscape could be one-dimensional or multi-dimensional depending on the variables in the model. A stable minima fixed point represents an attractor or sink which the system may move towards, like a boulder rolling down a valley side to the base. When the system is at this stable fixed point it is resilient to forces acting on it, as would a boulder at the base of a valley. The system could be perturbed, or the environment could change; the boulder might be pushed by a human, or erosion overtime might create favourable conditions for the boulder to move again, changing the fixed point into an unstable one. Unstable fixed points represent maxima, or peaks. This might be at the crest of a hill where rock outcrop is protruding with a large tension crack at its rear and just about to fail and move to a more stable position. Of course, rock outcrop at the crest of a hill might also be stable, so a rock at the crest of a hill will not necessarily be at an unstable fixed point.

2.2.5 Adaptive Cycles

Related to Complex Adaptive Systems and Resilience is the theoretical framework of Panarchy and Adaptive Cycles (Holling, 2001). Panarchy describes a concept that explains the evolving nature of complex adaptive systems. It involves hierarchies of human systems and natural systems interwoven together. They interact causing continuous adaptive cycles of growth or exploitation (r), conservation (κ), collapse or release (Ω) and reorganisation (α), Figure 2.5. These cycles exist within each scale over time and space. For example, in societies this could be cycles between the following institutions - small groups or individual decisions, policies and contracts, law, constitution and culture.

Figure 2.5 shows the adaptive cycle. The potential can be thought of as the wealth of the system, it determines the range of future options possible; increased potential gives the system greater ability to change. Connectivity can also be thought of as controllability, increased connectivity decreases the degree to which the system can control its own destiny. Resilience is related to the amount of adaptive capacity the system has. This plot is conceptual and theoretical although many of the features of it have been found in studies of real world systems. The growth phrase follows a sigmoidal curve from exploitation to conservation and is common across all biological systems (Bettencourt et al., 2007). Gardner and Ashby (1970) found that large complex systems will become unstable at a certain level of connectedness which is shown in the collapse stage between conservation and release. This is also touched upon in work on how systems of different size share the commonality of becoming rigid or locked into certain patterns over time making it more difficult for them to respond to changing scenarios and lead to their collapse (Scheffer and Westley, 2007). Multiple studies of lake systems show how they reorganisation (regime shift or critical transitions), this is the stage between release and reorganisation of the adaptive cycle (Wang et al., 2012; Scheffer et al., 2001, 2012). This reorganisation often follows increased human interaction with the system (Scheffer et al., 2001), which can be viewed as increased connectedness.

The Adaptive Cycle is also moving through time, so although it appears like a continuous infinity loop, it is not. In Figure 2.5 this is shown by a disconnect between the reorganisation and exploitation parts of the cycle.

As an IIS evolves over time, and is exposed to internal and external perturbations, it will go through adaptive cycles to gain resilience against changing conditions. It would be very difficult to measure adaptive cycles, and it is a theoretical framework so may not exist entirely in reality. A scenario might be that a IIS grows to the carrying capacity of the water supply in its environment,

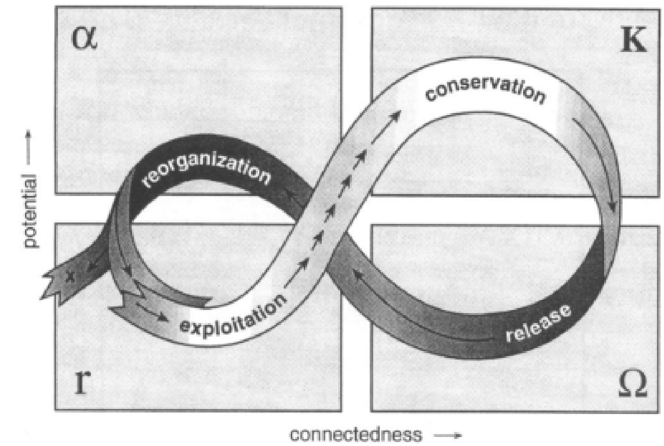


Figure 2.5: Conceptual drawing of the Adaptive Cycle in Panarchy taken from (Holling, 2001). α, κ, r and Ω represent the four ecosystem functions.

and then is hit by a large drought leading to the system collapsing and reorganising in an adaptive cycle. The managerial arrangements may alter based on this collapse in a form of adaptation.

2.2.6 Social-Ecological Systems

An Indigenous Irrigation System (IIS) is defined as a subset of the broader topic of Social-Ecological Systems (SES). SES are two distinct but coupled systems, humans and the environment (Levin et al., 2012). The study of SES is important for the persistence of complex ecosystems and societies on earth (Rockström et al., 2009). This is currently a pressing issue given scientific consensus that we are now in a new geological epoch - The Anthropocene; defined as the period in which human activities rival global geophysical processes (Steffen et al., 2011). In response to this, the 'Planetary Boundaries' framework defines a safe operating space for humanity to regulate the stability of social-ecological systems (Steffen et al., 2015; Dearing et al., 2014).

The behaviour of social and ecological systems are fundamentally different. Scaling relationships of biological systems have been found to follow quarter-power scaling with a scaling exponent less than 1 leading to sub-linear growth (West et al., 1997). Social (urban) system properties such as wealth and knowledge are found to increase at much quicker rates with scaling exponents above 1 (Bettencourt et al., 2007), leading to super-linear growth. This super-linear growth often leads to boom-collapse relationships (Bettencourt et al., 2007). The scaling exponent emphasises the differences between social and ecological systems. Enforcement of the planetary boundaries contains the ever increasing wealth and knowledge to the social system, limiting negative effects on the ecological systems.

IIS are social-ecological systems but are not urban. It is therefore expected that their scaling exponent will be less than 1, meaning they are an example of a systems that reaches a long term population limit, which might be why they are able to persist for thousands of years.

2.2.7 Complexity and Stability

Whether increasing diversity leads to increased stability is an open question in ecology and complexity science (McCann, 2000). Complexity tends to be synonymous with diversity, for instance a more diverse ecosystem requires more information to describe it which is the definition of complexity in Section 2.2.1.

Early empirical work found that more diverse ecosystems were more stable (MacArthur, 1955). However theoretical mathematical work found that larger increasingly interconnected systems were more likely to be unstable (May, 1972). But this result is based on several interacting with one another, with each species having a probability of being unstable. As connectedness and interaction increases the system is more likely to become unstable. This is generally saying that as a system of connected species interacts, if this interaction has high strength and high connectedness, then if only a small proportion of the system is unstable, this will lead to the whole system becoming unstable. Although this simple model is elegant and clear with its conclusions, it does not represent what has been observed in the empirical data. It is possible that this model does not represent all the interactions and adaptabilities which are inherent in living systems. For instance, if one species is unstable, instead of this instability cascading through the entire ecosystem, perhaps it leads to a loss of connections as other species *adapt* in order to persist. This is similar to more recent work showing that CAS often re-organise in order to persist (Scheffer et al., 2001). So, the model lacks the presence of evolutionary principles which if added may show that the system will fluctuate between stability and instability as it evolves and adapts. This school of thought has also been applied to real world human systems such as the financial industry to explain that the reasons for financial crisis are due to highly interconnected high strength interactions (May et al., 2008).

Following on from this, May (1972) also states that the organisation of the system is of importance for increasing stability. That is, if a system is arranged into blocks (is modulated) then the probability of stability is increased.

Generally, recent research suggests that increases in diversity, on average, gives rise to ecosystem stability (McCann, 2000). Diversity for the sake of diversity does not seem to increase stability (May, 1972). Instead it seems that a system which can maintain novel redundancies which are in turn capable of differential response dependent on changing conditions increases stability. This is of course related to Resilience. A system with increased Resilience will also have increased stability. This also relates to the Adaptive Cycle, which shows changing stability with connectedness and potential.

Therefore, would more complex IIS have increased resilience and stability?

2.2.8 Self-Organisation and Criticality

Self-Organisation is a key concept in complex systems science. It was originally developed in physics and chemistry stating that systems consisting of many interacting components at a lower level will interact in such a way that they organise to produce emergent behaviour at a higher macroscopic level (Bonabeau et al., 1999). It has a much wider application though across the life sciences and humanities. Models of Self-Organisation often show that the interaction of simple agents can explain complex collective behaviour (Bonabeau et al., 1999). Well studied examples of self-organisation in biology are bees, termites and ants (Bonabeau et al., 1999). The study of Self-Organisation has extended to artificial systems, for both the study of natural systems, but also in the development of engineered robotic systems (Bonabeau et al., 1999). This self-organising behaviour is capable of solving collective problems which might underlie why it evolved in nature (Nowak et al., 2010). Self-Organisation relies on four basic system properties: Positive Feedbacks, Negative Feedbacks, Amplification of Fluctuations and multiple interactions (Bonabeau et al., 1999).

The related theory of Self-Organised Criticality (SOC) states that large systems of many components organise to a poised critical state which is out of equilibrium (Bak et al., 1987). When in this state, minor disturbances may lead to events of all sizes termed avalanches. According to the theory these events will follow an inverse power law of few large events and many smaller events. The Sandpile model is the first and one of the simplest examples of self-organised criticality (Bak et al., 1987). It consists of sand being dropped continually to the same point. Over time a pile of sand is formed and once the sides reach the friction angle of sand it is said to be at a critical state. This sounds like a very simple example, but self-organised criticality has been observed in many systems (Bak and Paczuski, 1995) such as Earthquakes, Computer models such as the Game of Life computer model (Bak et al., 1989), Economies, Cities and river systems (Van De Wiel and Coulthard, 2010).

Self-Organisation and Criticality can be related to the exploitation and conservation stages of Adaptive Cycles (Section 2.2.5). Figure 2.5 shows a system growing to this criticality state (The Conservation Stage). Many fields of science are devoted to predicting or understanding when a large event might occur when in the Conservation Stage, as it could have large consequences for the system and for those connected to it. For example the study of lakes has found early warning signs at the point when the lake is about to change regime and collapse (Scheffer et al., 2001).

Self-organisation has been studied in IIS previously (Lansing et al., 2017), finding power law relationships in spatial patterns in the Subak System result from feedback between the local ecology and farmer decisions in a self-organising process. The power law is a signature of self-organised criticality which results from local agent adaptations driving the system to a global optima. This is unlike the sandpile model which is exogenously driven (Lansing et al., 2017). If IIS evolve in reasonably isolated conditions, then they are likely to exhibit many characteristics of a self-organising system, but in reality many do not, the Qanat system found across the middle east is a technology which has been introduced to arid areas as a solution to the water supply problem. However each Qanat system tends to be isolated allowing for self-organising processes to occur. The long standing debate of self-organisation and planned processes in IIS (Lansing et al., 2009; Wittfogel, 1957) can be studied by observing signatures of self-organisation in IIS such as power law scaling.

These self-organising systems all require energy so can be thought of as open systems which are driven out of equilibrium forming dissipative structures (Prigogine and Nicolis, 1985). The energy in the system allows them to produce ordered structures or patterns having the effect of negative entropy.

2.2.9 Common Pool Resources and Public Goods

Indigenous Irrigation Systems involve the distribution of a resource amongst many farmers or agents. This can be viewed as a common pool resource (CPR) problem.

In general CPR describes the allocation of pooled resources within a population and is most often studied in the humanities such as in social science and economics. It is not often referred to in the life sciences, but it can also be readily applied to natural systems. For example, a common resource has to be allocated to different parts of a plant to enable it to function properly in a stable manner.

A public good is one which all enjoy in common in the sense that each individual's consumption of such a good leads to no subtractions from any other individual's consumption of that good (Samuelson, 1954). The key issue of Common Pool Resources and Public Goods is whether individuals can suppress their self interest for an end result in which all individuals benefit to a greater degree than if acting independently (Levin, 2014). Evolutionary thinking can aid in helping societies address some of its greatest issues such as system sustainability as many common pool problems have been solved previously in biological and small-social systems (Levin, 2014).

Following on from this, the famous essay 'The Tragedy of the Commons' Hardin

(1968) discusses the problem of overpopulation of humans and the effect of it on our CPR, the planet. The essay states that in order for humans to overcome the problem of overpopulation they need to abandon their freedom to breed and this should be enacted with mutual coercion, mutually agreed upon by the people affected. Many small-scale societies overcome the Tragedy of the Commons through the establishment of norms. Ostrom (1990) famously defined a number of common rules between societies that avoid the Tragedy of the Commons.

'The Tragedy of the anti-Commons' has been proposed as an opposite effect of the Tragedy of the Commons (Heller et al., 1998). In this scenario, over protection by individual players leads to the resource not being used at all.

Brede and Boschetti (2009) defines a spectrum between tragedy of the commons and tragedy of the anti-commons. Brede and Boschetti (2009) then explore this in a game-theoretic framework and an evolutionary simulation. The simulation found two stable fixed points, one being the tragedy of the commons, the other being the anti-tragedy of the commons. In-between the two are locally stable fixed points where a few individuals initiate obstructive policies which stop overuse of the resource. These fixed points are however fragile to perturbations such as a population increase.

Previous research has already found that IISs solve the Common Pool Resource problem in different ways. The Moroccan irrigation systems rely on individual private property rights whereas the Subak system relies on community ownership (Geertz, 1980). The reason for this difference might be due to the amount of water supplied into the system.

2.2.10 Social Norms

IISs solve CPR problems through different cultural practises (individual property rights/community based farming practises) (Geertz, 1980). The cultural practise is a high level process, at a lower level this interaction between individuals takes the form of social norms, such as with adaptive cycles.

Social norms can be viewed as the 'memes' which are associated with a certain cultural practise. They are representative or typical patterns and rules of behaviour in a human group (Ehrlich and Levin, 2005). Norms are of great interest in social science, environmental science and policy research due to the effect that human behaviour has on both social systems and ecological systems. Government policy can aid in changing human behaviour to avoid less desirable social and ecological states (Kinzig et al., 2013). However such analysis and subsequent policies may have detrimental affect on individual freedom.

In order to model such behaviour, simple cellular automata can be used to abstractly model the development of social norms (Schelling, 1971; Hartnett et al., 2016; Ehrlich and Levin, 2005). Figure 2.6 shows the basic idea of modelling

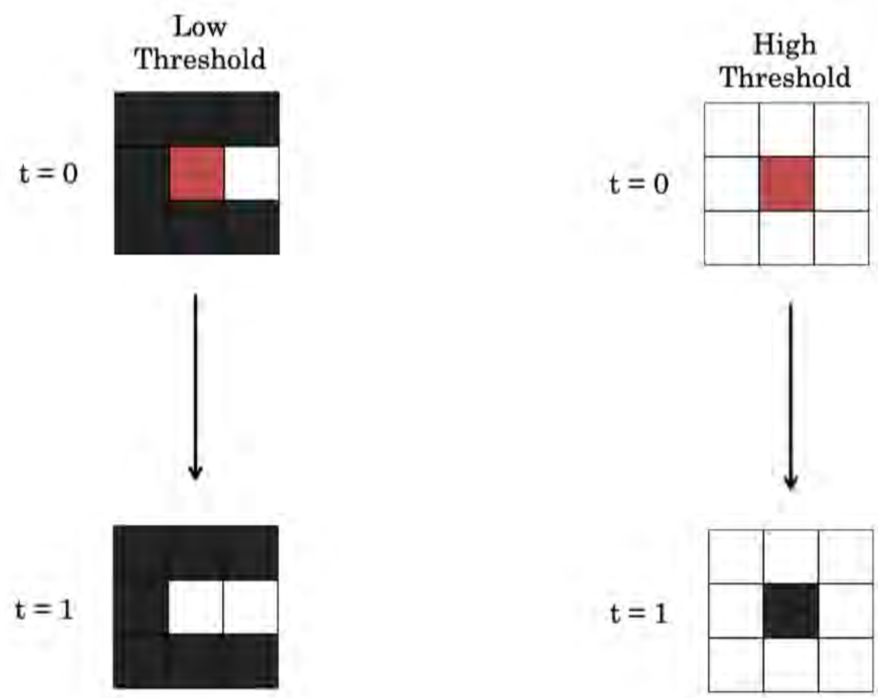


Figure 2.6: Basis of modelling thresholds using cellular automata. The agent in the red square will change state based on its local (Moore neighbourhood).

norm changes in social systems. An agent with a low threshold is one which will change their state with even a low number of neighbours with a different state. A high threshold is an agent which will not change their state even if all their neighbours have an alternative state. This example looks at the changing of norms based on the opinions of neighbouring cells. But a similar method can be used to look at models of segregation in humans (Schelling, 1971). This is a related method to that used in game theory simulations (Axelrod and Hamilton, 1981; Nowak and May, 1992a).

As culture is one of the interwoven factors which leads to the emerging properties of an IIS, the cultural practises should be studied. In addition, if the system is to be modelled through the interaction of agent farmers, the social norms have to be known and compared to the model for validation purposes.

2.3 Methods for modelling Complex Systems

The previous section provided a review of relevant high level theories which are useful frameworks for studying dynamic systems of many interacting components (complex systems) such as IIS. The following sub-sections firstly give an overview of different approaches to modelling complex systems and then more detail on models and applications which are more relevant to the research questions of this thesis.

2.3.1 Modelling and its limitations

The paper 'Why Model?' focuses on some of the largest misconceptions about modelling (Epstein, 2008). Prediction and validation are often thought to be key attributes of a robust model. Whilst prediction is often thought of as the foremost reason to model, Epstein (2008) gives many more reasons to model other than to predict. These include:

- Explaining;
- Guide Data Collection;
- Illuminate Core Dynamics;
- Suggest Dynamical Analogies;
- Discover New Questions;
- Illuminate Core Uncertainties;
- Challenge the Robustness of Prevailing Theory Through Perturbations;
- Offer Crisis Options In Near-Real Time;
- Train Practitioners;
- Education;
- Reveal the Simple to be Complex and Visa-Versa.

Models are used in both qualitative and quantitative social science. In qualitative social science, modelling is implicit - it is undertaken via thought experiments. In quantitative social science, it is explicit, assumptions are laid out in detail so we can study what they entail (Epstein, 2008).

Modelling can be very useful for studying IIS. Many IIS of interest have persisted for long periods of time, and often without considerable outside intervention. Studying them in the modern globalised, interconnected world is very difficult given many more factors influence their behaviour. It is also very difficult to deduce the factors leading to their current form. Computer models offer a tool in which the factors of influence can be controlled.

2.3.2 Mathematical Origin

The origin of the modern field of modelling complex systems can be traced back to the work of Henri Poincare (Chenciner, 2012). The three body problem stands as one of his most important fundamental contributions to the field which finds that for a system consisting of three bodies or more, there is no general closed form analytical solution to find the positions and velocities at any point in time without using numerical or simulation based methods. This forms the basis of chaos theory, in which chaos is defined as deterministic, sensitive to initial conditions and unpredictable. The need to model or simulate complex systems can be seen to come from this problem of not being able to find closed form solutions.

The three body problem is particularly important for this research as many of the qualitative ways of analysing dynamic systems originate from studies on three body systems (Strogatz, 2015). As scholars realised that there was no closed form solution for the interaction of three or more bodies, a shift in analysing such systems occurred from trying to find the exact positions of each of the bodies in time, to asking more general questions about the system such as "Will the system be stable forever?" (Strogatz, 2015). This expanded into dynamic systems theory which can be in turn applied to many more much complex systems. Seminal applications of stability analysis of many body systems include the work on dissipative hydrodynamic flow (Lorenz, 1963) and more recently of ecological systems (Scheffer et al., 2012).

There are many ways to model complex systems, and in particular social-ecological systems (Verburg et al., 2015). This section outlines general tools used to model complex systems. All the modelling methods discussed below can be thought of as numerical models. Their differences lie in the different applications of mathematics and whether they rely on data or not. Models may be theoretical (which tend to be more general) or applied, when attempting to model a specific system.

Modelling techniques include:

- Equation based models;
- Deterministic process-based biophysical models;
- Cellular Automata Models;
- Simple Toy social-ecological models;
- Network Models;
- System Dynamic Models.

2.3.3 Equation based Models

Equation based models in life science do not have as strong a foundation as in physical science and engineering (May, 2004). May (2004) and May (1976) provide overviews of some efforts to build simplified equation based models which can capture observed patterns and processes seen in reality without being lost in the detail. Many of the most powerful and also most fundamental models in biology are the simplest and also the most general in that they can be applied universally. These simple models offer building blocks and clarity on which to build theory and applications for the real world.

Examples of equation based models used in the life sciences include:

2.3.4 Lotka-Volterra Equations

The work of Lotka and Volterra was one of the first examples of equation based models applied to the life sciences. This explored, in the form of two coupled non-linear, first order differential equations how two species interact in a predator-prey model. A variation of the equations can be summarised as:

$$\frac{dx}{dt} = ax - \beta xy \tag{2.3}$$

$$\frac{dy}{dt} = \delta xy - \gamma y \tag{2.4}$$

where x and y are the two populations and a, β, δ and γ are constants. Sayama (2015) provides a clear summary on how such equations can be derived using simple mathematical functions.

These equations have since been extended to look at more complex ecosystem dynamics with many more agents and interactions (Gilpin and Ayala, 1973). The coupling has also been applied to social-ecological systems using the HANDY model (Motesharrei et al., 2014) and even the coupling between two societies (Roman et al., 2017).

2.3.5 Interaction Matrices

For modelling ecosystems consisting of a large number of components interaction matrices are often used (May, 1972). They can therefore be thought of as an agent based modelling platform. These have been useful in studying the stability of complex systems in relation to size and connectivity (May, 1972; Allesina and Tang, 2011), an area of interesting study in both ecology and economics (May et al., 2008). May (1972) and initially Gardner and Ashby (1970) find that large (n >>1) connected systems are inherently unstable. This has similarities to later work on how a virus spreads across a network in that the probability of becoming unstable can be thought of as similar to the probability of an infection spreading across the network (May, 2004). The notion of increased instability with increased connectivity was later explored and theorised qualitatively in Panarchy theory and adaptive cycles (Holling, 2001). More recent studies within the theoretical field have looked at combining models of ecosystems with a large number of components with models of homoeostasis to understand how regulation might emerge in such large systems (Dyke and Weaver, 2013). They can then be analysed with linear stability analysis to provide information about when a particular system is stable or unstable (May, 1972).

As IIS contain many interacting agent farmers, interaction matrices could be useful for modelling. There is a lot of flexibility when using interaction matrices as multiple matrices can be used for many variables in the system and many coding languages are optimised for computing in matrices. There are no known examples of interaction matrices being applied to modelling IIS.

2.3.6 Spatial Models

Most systems in the real world tend to interact spatially. Systems which are spatially constrained (river systems, ecological systems, transport networks) can be categorised separately to systems where non-spatial processes are more dominant (the world wide web, neural networks, numerical modelling). A non-spatially constrained system being one where information is transferred at such a high rate that the spatial configuration is not necessarily important. Models of the internet contain structural information about the network but this is not often spatially constrained as it is not an important factor (Caldarelli et al., 2000). Whereas, in a spatially constrained system, the information transfer is much slower (relative to the system size) so the spatial structure is more important - local interactions are important in such a system. Simulation models such as cellular automata and agent based models also allow for spatial interaction of the system to be explicitly added, which tends to not be as easy in equation based models. Examples of spatial models include earth surface dynamics such as sand dune processes (Nield and Baas, 2008), river flood models (Coulthard et al., 2013) and coupled ocean-atmosphere simulation models which have been developed mainly for simulating climate change (Jungclaus et al., 2006).

All factors influencing the evolution of an IIS are interwoven and spatially constrained; geology, geomorphology, climate, ecology and culture. The physical network of the IIS will also be constrained spatially, as it will most likely be built on a downhill slope and a space-optimising manner to make best use of space and resources.

2.3.7 Cellular Automata

Cellular Automata (CA) are one of the simplest frameworks in which issues related to complex systems, dynamics and computation can be studied (Mitchell et al., 1993). At its most basic level, a CA consists of a spatial lattice of cells, each of which, at time t, can be in one of k states. The CA has a fixed rule which is used to update each cell based on its nearest neighbours (Mitchell et al., 1993). They can be used to explore abstract phenomena such as 1-dimensional CA (Wolfram, 1984) which can be used to look at simplistic but fundamental mathematical and computational spatial problems. Wolfram (1984) updated the CA row by row through time from a rule set which relates to the nearest cells in the previous time step. By mapping this 1-dimensional CA through time and producing a 2-dimensional pattern, many interesting complex phenomena emerge. These have been systematically numbered as different rule sets from 0 to 255 and classified into four main classes - convergence towards a uniform state, convergence towards a repetitive or stable state, appear to remain in a random state and CAs which form areas of repetitive or stable states but also form structures that interact each other. Classes three and four tend to be most interesting where chaotic, fractal and aperiodic behaviour can be observed. Not only are these simulations interesting from a abstract mathematical point of view, they are also visible in nature. For example the shell of the Mollusc *Conus Textile* appears to resemble the pattern produced by Rule Set 30 (Wolfram, 1984).

CAs can also be studied in 2-dimensions. Conway's Game of Life is the classic example of this (Adamatzky, 2010). On the surface, the Game of life appears simple, but even though it was first implements in the 1960s, it is still being actively researched and explored to this day (Beer, 2018). A recent applied example of 2D CA models has been used to shed light on the patterns found on skin scales in lizards (Manukyan et al., 2017). The study goes further by linking a continuous reaction-diffusion equation to the emergence of discrete values in a cellular automata leading to the pattern formed on the skin scales. This also poses an interesting overlap between a fundamental simplistic computer model of a cellular automata which has finite states and no movement processes to most natural systems which commonly involve movement through the system leading to a change in the finite state.

The application of these 2D Cellular Automata (which may involve movement processes) is almost infinite with examples including physical geographic processes such as river basin development and flood dynamics (Coulthard et al., 2002), fundamental evolutionary biological processes in spatial games (Nowak and May, 1992a), human geographic processes such as urban development (Clarke and Gaydos, 1998; Benenson and Itzhak, 2012) and plant development processes such as the canalization-based vein formation in a growing leaf (Lee et al., 2014).

2.3.8 Revisited examples of Cellular Automata

Simple simulations can be undertaken to explore the evolution of different behaviours from initial conditions based on the above method. From an initial random configuration figure 2.7 different thresholds can be tested. If all agents are treated as homogeneous, a low threshold for all agents (from figure 2.6)

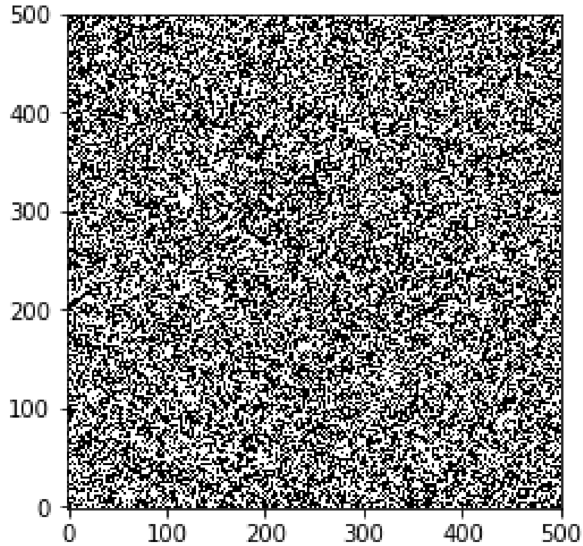


Figure 2.7: Initial configuration for a social norms simulation with two states - black or white.

after many iterations leads to a constantly dynamic random simulation, ie the same as the initial configuration figure 2.7. However, if all agents have a high threshold, the simulation remains static, and the same as the initial configuration, figure 2.7. Interesting dynamics occur at a moderate threshold, figure 2.8. At this point an agent will change their state with a probability of a half if half of their neighbours have a different state to them. A video of this simulation can be found at `https://youtu.be/HokoPnIPvas`. Figure 2.9 shows the dynamics of the system after 500 time steps. This has a lot of similarities to self-organising behaviour in other chemical and biological systems, which is usually termed morphogenesis (Turing, 1952). Another well known example is stripes on a zebra. Following the same method, but instead using a Von Neumann neighbourhood for the cellular automata, stripes form rapidly `https://youtu.be/jrBRsrAAgSk`.

Recent examples of more complex pattern formation processes using a cellular automata framework include implementations of the Grey-Scott equations (which is a simple reaction-diffusion model) (Gray and Scott, 1985; Bartlett and Bullock, 2015; Adamatzky, 2018) and the Navier-Stokes equations (the flow of incompressible fluids) (Wolf-Gladrow, 2000).

The evolution of an IIS forms a pattern which can be modelled using a cellular automata. Different rules of formation can be tested and the pattern formed by

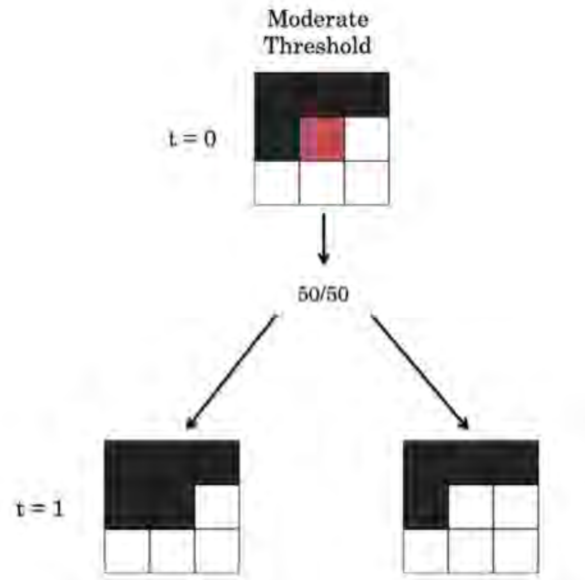


Figure 2.8: Moderate threshold scenario - Agents will switch behaviour to that of the majority of their neighbours. At the point where they have an equal number neighbours in opposition camps, their choice will depend on a coin toss.

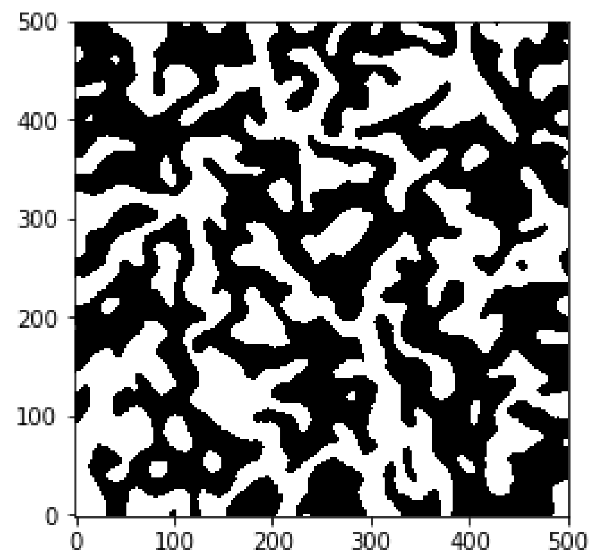


Figure 2.9: Final configuration after 100 time steps for a social norms simulation with two states - black or white.

real-life systems can be compared back to the model.

2.3.9 Agent Based Models

Agent Based Models (ABMs) can be viewed as an extension or type of Cellular Automata in that each agent is an entity which is updated based on its own rule set and other agents in their neighbourhood. The agents in ABMs are usually allowed to move in the model and the update of each cell may not be dependent on their immediate neighbours (Batty, 2007). Due to this rule based often adaptive nature, ABMs are usually applied more readily to models in biological and social systems. Examples of which include models of Jaguar habitats (Watkins et al., 2014) and social norms on payment for ecosystem services (Chen et al., 2012).

The use of computer models to generate *Artificial Societies* has been researched for many years. The origins of can be traced back to abstract models of local interaction in cellular automata for the purpose of researching self-reproducing machines by John Von Neumann in 1950 (Langton, 1984), one of the first attempts at simulating artificial life. This was not a model of a society though, the first artificial society model used a cellular automata framework to understand how segregation can emerge in societies (Schelling, 1969) and is also one of the first examples of an agent based model. The model showed that using just a few simple rules on local interacting agents leads to the emergence of a pattern, a segregated society. The Boids model of flocking behaviour in birds shows how local interactions can lead to large scale behaviour such as with flocking birds (Reynolds, 1987). The next advancement aided partly by an increase in computer power was the creation of 'The Sugarscape' which is explained in detail in the book 'Growing Artificial Societies' (Epstein and Axtell, 1996). Agent Based Models have since become part of the wider discipline of generative modelling (Epstein, 2006; Adamatzky, 2018). The philosophy of this type of modelling is discussed further in Section 2.3.17.

ABMs can be modelled in object-orientated computer coding languages such as Python with the agents defined as an *Instance* of a *Class*. Each instance stores a number of unique values about that particular agent, and after each time step the agent's state is updated based on the rules of the system and the unique values of that particular agent, and the agents in its neighbourhood. ABMs offer great power in exploring systems of many interacting (heterogeneous) agents, which would not be able to be explored in qualitative models, thought experiments or even using equation based models. Agent based models differ from system dynamics models in that they deal with individual agents and not aggregate behaviour (Gilbert, 2008). In a sense higher level behaviour emerges from the interaction between agents at a lower level - they are *Bottom-up models*.

IIS have not previously been modelled using an agent based model. In addition to the pattern forming nature of IIS which can be modelled using cellular automata, agent based modelling can add more factors of influencing in the resulting simulated IIS formed, such as social norm based decisions.

ABMs and cellular automata offer a more visual approach to modelling as the simulations explicitly occur in a spatial environment which can be animated to show how a particular outcome unfolds in the simulation, and not simply a graph output. This makes the approach more accessible and easier to explain to non-modellers. However agent based models tend to require a lot code whereas equation based models can be explained using a few equations usually allowing them to be clearer and more concise which in turn allows for easier repeatability.

2.3.10 Simulated Toy Models

These are generally conceptual models with no direct empirical input and can be described as theoretical and exploratory. They are useful for developing theories and underlying rules of complex systems that can then be applied and tested on empirical data. Examples of such models are evolutionary selfish-cooperative prisoners' dilemma games (Axelrod and Dion, 1988; Nowak and May, 1992a), Models of Segregation (Schelling, 1969), the sugarscape model of artificial societies (Epstein and Axtell, 1996), the Daisy World model of planetary homoeostasis (Watson and Lovelock, 1983), small-world networks (Watts and Strogatz, 1998), voter decision making models (Ehrlich and Levin, 2005; Hartnett et al., 2016), landscape evolution models (Coulthard, 2001), aeolian processes (Nield and Baas, 2008) and flocking behaviour (Reynolds, 1987).

If a model is built to simulate general behaviour of IIS, then a toy model could initially be a useful tool for theoretical and exploratory work. Toy models focusing on a pattern based interaction could be used to simulate the evolution of different IIS networks without having to add many other factors which would make modelling very difficult and drawing conclusions uncertain due to the large amount of factors.

2.3.11 Network Theory

Network theory is a framework which seeks to model and analyse through the use of networks. A network fundamentally consists of nodes and connecting edges. It is a popular tool to use given its simplicity. When simulating a Complex System a framework incorporating networks is often most appropriate. In fact the use of networks has greatly aided recent discoveries in complex systems, for example, 'small-world' networks appear to be ubiquitous in nature and society (Watts and Strogatz, 1998). Many systems in nature and society, when analysed as networks follow scale-free power law distributions, for example the internet (Barabási and Albert, 1999).

The growth of networks theory in many different fields of study is often attributed to the general process of preferential attachment (Barabási and Albert, 1999). This has been found to be effective at producing power-law network degree distributions which are common across nature (Barabási and Albert, 1999). Both growth in a network and preferential attachment are key to the development of power-law scaling in a network. Preferential attachment can be defined differently depending on the system that is being modelled. In the original network model (Barabási and Albert, 1999); preferential attachment is applied by the nodes having more edges connected to them having an increased probability that additional nodes added to the network will connect to them. It is a *'rich gets richer'* type model.

The notion of social and ecological networks has been present for a long time in qualitative social science and in an abstract form of mathematics in graph theory, however only recently has mathematics been applied to ecological and social networks (Watts and Strogatz, 1998). The simple nature of graph theory makes it very easily applicable, especially since it relies on reasonably simple mathematics. For instance a network can be modelled as an interaction matrix in which a number of properties are defined about the interaction of different agents in the system. Initially networks were not studied spatially, although this is becoming more common in the present day.

In terms of Complex Adaptive Systems, networks have been used in conceptual ecological systems such as food networks, and to an extent in system dynamics modelling. Attempts at modelling ecological systems with networks tend to be fraught with difficulty due to the large amount of interacting factors present. However tools such as system dynamics allow for a system's network to be built and visualised with little difficulty although calibrating and including the relevant variables may prove difficult.

Relating back to the section on adaptive cycles (Section 2.2.5), stability and network connections is explored in the ecological/sociological theory of Panarchy (Holling, 2001). This is a qualitative framework for exploring Complex Adaptive Systems. It can be generally thought of as a network based model, which relates to how systems expand and collapse with changes in 'potential' and 'connectedness' and to the scaling of the system, and how these cycles at different levels in the system interact with one another. When modelling resource distribution systems with altering levels of connectivity, it might be possible to see these cycles occurring within the system under study.

Power laws are commonly observed in networks such as the internet, in terms of links between web pages (Albert et al., 1999). That is, there are a few web pages with strong connectivity to most of the internet and many web pages with weak connectivity. A network model which attempted to explain this found that it required growth through preferential attachment to websites with more links already to give similar results (Albert et al., 1999).

There are also efforts to uncover universality in networks - the fundamental building blocks of different types of networks (Milo et al., 2002). The Watts-Strogatz Small-Worlds network is one such example, a network somewhere between a regular network and random network which appears to be prevalent across different systems in nature and society (Watts and Strogatz, 1998).

Networks are very applicable for modelling IIS. Both the physical network and the managerial arrangements can be studied as networks. There are maps and conceptual sketches of IIS, but none have been analysed using the tools of network theory.

2.3.12 Spatial Networks

Many agent based models can be viewed as being similar to spatial networks. Of course not all agent based models have to be spatial, and not all spatial networks have to be agents. But by focusing on a network perspective of such models the connectivity between the agents is one of the more important aspects of the model.

Firstly, a spatial network is defined as one with nodes and edges which are constrained by some geometry and are usually embedded in a two or three dimensional Euclidean space, which in turn has important effects on their topological properties and consequently on the processes which take place on them (Barthélemy, 2011).

Just like Cellular Automata, Spatial Networks can be implemented across a wide range of topics. Examples include transport, rivers (Kyungrock and Kumar, 2008), internet, power grids, social and neural networks (Barthélemy, 2011). These systems are spatially constrained but do not require local interaction such as with cellular automata.

Transport networks in cities can be analysed as networks with each trip between locations viewed as a link (Batty et al., 1999). Spatial interaction matrices can then be built to analyse the flow of people between areas, and probabilistic predictions of movement through the system made. These are sometimes termed as gravitational models with each area having a certain 'mass' which might relate to the average salary in that area and so will lead to a greater flow of trips to that area.

Section 2.4 gives further examples of spatial networks, particularly tree and planar networks which are more similar to IIS networks.

2.3.13 Spatial Network Models and Analysis

Analysis of Spatial networks (and networks in general) helps in quantifying and characterising the structure and behaviour of the system. This might be for identifying keystone species in an ecosystem (Solé and Montoya, 2001), fragility in financial systems (May et al., 2008) and analysing dynamics of networks such as in gene regulatory networks (Karlebach and Shamir, 2008).

Since the subject has increased in popularity, there have been efforts to formalise the description of different types of network and the ways to analyse them. An overview of types and analysis is given below.

Firstly networks consist of nodes (or vertices) and edges. The nodes are points and edges are lines connecting the points. They can be classified in terms of their direction, weighting, self loops and planarity (Newman, 2010):

- Direction applies to the edges. They can be undirected, meaning no information about flow direction, or directed meaning direction information is contained within them. If networks are directed, then information will flow in the direction stated, this can be uni or bi-directional. A directed network which contains flows which loop back to the same node are called cyclic. If these are not present then the network is described as acyclic.

- Weighting can apply to both the nodes and edges. This may act as a multiplier increasing the size of information being passed through certain nodes or edges. This is particularly important for neural networks - when they undergo training the weights are increased or decreased for the problem in hand.

- A self loop or self edge is when a edge is connected back into the same node. This might apply to something like regulation to a gene in gene network.

- Planarity refers to whether a network is planar or not. A Planar network is one which can be drawn without any edges crossing one another.

Networks can be simply modelled using an edge list. Each entry contains two values relating to two nodes which are connected by an edge. However for modelling and analysis they are not particularly useful. Instead Adjacency matrices are more appropriate (Newman, 2010). The adjacency matrix $\mathbf{A}$ maps the interactions between elements A_{ij} so there is a 1 in the matrix if there is an edge between them else 0. If multiple edges (multi-edge network) are present between two nodes, then the A_{ij} value in the adjacency matrix will increase. If a node has a self loop then the value in the adjacency matrix is equal to two. This will always be on the diagonal of the matrix.

The Adjacency matrix $\mathbf{A}$ of a graph is written with elements A_{ij} as follows:

$$A_{ij} = \begin{cases} 1, & \text{if there is an edge between } i \text{ and } j. \\ 0, & \text{otherwise.} \end{cases} \tag{2.5}$$

If the network is weighted or directed then this can also be represented in a matrix. The conventions of which are widely adopted are discussed in (Newman, 2010). An undirected network will have a symmetrical matrix, whereas a directed network will have an asymmetrical matrix.

Types of analysis of networks include path length, degree, clustering coefficient, community detection and largest connected component (Gosak et al., 2018). More detail is provided below:

- A path is a continuous route across a network. The length is usually the number of edges constituting the path. For spatial network the physical distance is obviously of importance too.

- The degree (k) of a node is the number of edges connected to it. There are many statistical analysis techniques related to the degree. The *Mean Degree* is the average number of edges each node has in a network. The *Connectance (ρ)* of an network is the fraction of the total possible edges that are actually present in a network. This is most easily measured for a simple network (one without multi-edges or self-edges) as there is less ambiguity over the total number of possible edges. This will always be in the range $0 \leq \rho \leq 1$. If ρ tends to 0 with increased n the the network is characterised as *sparse*, however if ρ tends to a constant with increased n, then it is characterised as *dense*. The *Degree Distribution* gives a distribution of the number of nodes in a network which have a certain degree. Each value in the distribution is represented as j_k/n where j is the number of nodes for a certain degree (k) and n is the total nodes in the network. It is therefore a probability that a randomly chosen node has degree k. Two networks with the same number of nodes and edges can have the same degree distributions even if the structure is different. Most real world networks have a right-skewed degree distribution.

- The Clustering Coefficient measures the average probability that two neighbours of a node are themselves neighbours which is essentially looking for triangles in a network (Newman, 2010). Some networks tend to have higher clustering coefficients than random networks of similar size and comparable degree such as social networks where as others have lower such as the world wide web. Clustering Coefficients can also look for other sorts of network structure than triangles, although this is the most common and so allows for easier comparison between networks.

- Community Detection is a more sophisticated tool which is used to search for the naturally occurring groups in a network regardless of their number or size, which is used primarily as a tool for discovering and understanding the large-scale structure of networks (Newman, 2010). In some respects it can be viewed as a more sophisticated form and scale-free version of the clustering coefficient. There is no universal accepted way of detecting communities. One such way is to detect modularity in the system, which has an increased value when edges are more connected to one type of node. An automated algorithm would find out these different types on its own. Also it might be useful to defined the maximum size for a community detection algorithm for whatever the problem being tackled is.

- The largest connected component is the maximum set of nodes such that each pair are connected by a path. It reflects the fraction of nodes that are interconnected either directly or indirectly and how many of them are isolated. This is informative when looking at the susceptibility to perturbations (Gosak et al., 2018).
- Network Resilience, like the broad definition of resilience given previous is a system's ability to adjust its activity to retain its basic functionality, when errors, failures or environmental change occur (Gao et al., 2016). Common ways of measuring the resilience of a network analyse the effect of node removal, edge removal or global fluctuations (Gao et al., 2016). When nodes are removed, the resilience (or fragility) can be measured as the proportion of nodes in the largest connected node cluster in the network (Solé and Montoya, 2001).

2.3.14 Complexity and Entropy in Spatial Networks

A more applied example of spatial network analysis is to look at how a certain activity N is distributed across a number (n) of adjacent areas where N_i is the activity in each area i. This might be a city for example. This activity could be one of many things, for example economic activity, number of a petrol stations, parks and so on. A measure of the complexity (W) of the system can be found using the following equation:

$$W = \frac{N!}{\prod_i N_i!} \tag{2.6}$$

If all the activity took place in the first area, then W would equal 1, else if the activity was equally distributed amongst all areas, then the value would depend on the size of N and n. The most often cited definition of complexity is that of *Kolmogorov Complexity* in Computer Science. The Kolmogorov Complexity of a sequence is the length (in bits) of the shortest computer program that prints the sequence and then halts. So increased complexity and a higher value of W means more information is needed to describe the system.

The Shannon Entropy H of the system can also be calculated. Firstly a probability is found for each of the areas, $P_i = \frac{N_\text{i}}{N}$ which is then used in the following equation,

$$H = -\sum_i p_i \; ln \; p_i \tag{2.7}$$

The Entropy is at a maximum when the probability of activity occurring across all areas is the same (p_i=$1/n$) and a minimum when it is all in one area. The principle of maximum entropy states that the probability distribution with the largest entropy, gives the best representation of our state of knowledge of the system (Purvis et al., 2019). Studies which agree with this principle will follow it when carrying out additional analysis or modelling. The method for finding the maximum entropy can be found in (Batty, 2009).

Both entropy and complexity increase with n. There is therefore a clear link between Complexity and Entropy which is discussed further in (Grunwald and Vitanyi, 2004). Complexity focuses on the the object and the amount of information required to reproduce it, where as Entropy focuses on the probability of producing that object from a random signal.

These two measures are also related to others in the Economics such as inequality. The Gini Coefficient G is one of the most popular and as defined as:

$$G = \frac{1}{n}(n + 1 - 2(\frac{\sum_{i=1}^{n}(n + 1 - i)y_i}{\sum_{i=1}^{n} y_i})) \tag{2.8}$$

where n is the number of areas, and y_i is the activity value in area i. The Gini Coefficient will be 1 if all the activity is in one area, 0 if it is completely equally distributed and somewhere in between for other distribution patterns.

These methods can be useful for analysing the development of inequality in IIS. The common use of the Gini Coefficient means that any measurement of can be easily compared, this is particularly useful for simulation models.

2.3.15 Scaling in Spatial Systems

Empirical data shows that many systems in nature and society follow scaling laws in terms of the size distribution, cities being one. For example in any nation there are many more small cities than large ones. The universal scaling equation which is often widely quoted for systems following a power law relationship (Batty, 2009) is as follows:

$$p_i = KP_i^{-\phi} \tag{2.9}$$

where,

p_i is the probability of finding a system of size P_i occurring given this distribution. K is a normalisation constant and ϕ is the scaling parameter (which is negative as for this distribution as the power law is negative, for example city size). Batty (2009) provides further information on this equation.

Within these systems, there is a power law distribution for the activities taking place in small cities and larger ones. Both biological and human systems have been found to follow these scaling relationships (West et al., 1997; Bettencourt et al., 2007). This is generalised in the the equations in the Section on Scaling (Section 2.2.2).

This same relationship is likely to be found for IIS. There are likely to be few large IIS and many smaller ones.

2.3.16 Distribution Systems as Planar Rooted Tree Networks

As mentioned in the previous section, one of the main reasons which is thought to underlie scaling (particularly in biological systems) is the transport of materials through linear networks that branch to all parts of the organism (West et al., 1997), is a type of distribution network. Given the vast amount of processes and systems which can be identified as Distribution Networks there are many types of network to describe them. Distribution systems will be directed (the analysis of them does not have to be), but they can be uni- or bi-directional. A hierarchical distribution system will most likely take the form of a tree. A tree is a connected, planar, network which has no closed loops (Newman, 2010). If this network originates from one node, it is termed *Rooted*, with the origin being the *Root Node*. The assumption that it is planar (planar meaning that the network can be drawn without edges crossing) can be relaxed in some cases. Connected means that every vertex in the network is reachable from every other. Non-hierarchical or less hierarchical distribution systems may not be like a tree at all, and instead a small-worlds network with hubs.

By studying distribution systems as networks commonalities between different systems can be found and greater understanding of the reasons why certain distribution systems for particular functions are beneficial. Which when applied can lead to improved network design and management of existing networks.

All IIS are distribution networks. Based on the information collected on them so far they appear to be rooted, planar networks. Further empirical network can be collected to confirm this.

2.3.17 Discussion of approaches and Philosophical underpinnings

Throughout this section a number of different approaches to undertaking scientific modelling have been introduced. These approaches can usually be categorised into three different types of philosophical reasoning - Inductive, Deductive and Abductive. Inductive reasoning (at least in the social sciences) relies on assembling macroeconomic data and estimating aggregate relations (Epstein, 2006). Deductive Reasoning relies on reductively applying general rules across a domain until a true conclusion is met. Abductive reasoning is similar to deductive reasoning but seeks to find the most likely explanation, so has an element of uncertainty involved.

Examples of induction based models are those which are built from data *Data Driven Modelling*. Equation 2.9, the scaling relationship probability of finding a city of a certain size is inductive as it has been fitted to data. However Equation 2.1, a model of scaling relationships within biological systems takes similar form to Equation 2.9, but has been derived using deductionist reasoning from a simplistic model of a branching system. Machine learning models which often produce probabilistic outcomes without finding a true complete understanding in the model can be thought of as being abductive.

Agent Based Modelling has been described as having differences to both inductive and deductive approaches. This is described as a generative approach (Epstein, 2006) which is inspired by the definition of syntactic theory which seeks minimal rule systems that are sufficient to generate the structures of interest (Chomsky, 1965). The main questions which agent based and generative models are effective at answering are on the topic of how certain regularities can arise from decentralised local interaction of heterogeneous autonomous agents (Epstein, 2006). This is not similar to inductive reasoning which as stated previously relies on data to produce models of aggregate relations. Epstein (2006) argues that generative models are a form of deduction as each model can be deduced to understand all the interactions taking place and a true answer found, but deduction does not imply generative, as deductive theories can rely on equilibrium models which are not generative.

The modelling approach used in this research is generative agent based modelling.

2.3.18 Model Validation

All models should undergo a process of validation. The way in which this is done depends largely on the type of model. For example a machine learning model which is built to classify images of a particular object can easily be validated by testing it on correctly identifying that particular object.

However, abstract (often general) complex systems models may not be so easy to validate. The approach taken in this research of generative modelling relies on a set of initial parameters and possibly stochastic processes to grow an artificial society. The parameters will be derived from previous models of the system under study, but as the systems under study consist of many interacting components, not all processes or factors will be accounted for. Validation can take the form of common measures from both the model and the real world system. For examples power law or Pareto Distributions is one such commonality found between the Sugar-scape simulation of human systems and real world systems (Epstein and Axtell, 1996). The gradient of the power law line is a common comparative measure (Jun and Hübler, 2005). The limits of validation should be stated, including assumptions and components not accounted for in the model.

This is especially important when considering the application of the model to the real world system.

A data driven approach such as machine learning may prove much easier to validate and have greater predictive power, but the model itself is often poorly understood a 'black box', often due to it being very complex. On the other hand a generative model which is simulated in isolation away from the data, may prove more difficult to validate and offer less predictive power, but the model is better understood meaning that the core processes of the system are better understood, which in turn can be compared back to the real world system.

2.4 Examples of Systems with similar properties to IIS

As this research is taking an interdisciplinary Complex Systems based approach, systems with common properties to IIS should also be considered. By finding commonalities between IIS and other systems in nature it provides evidence that the processes and structures formed in IIS are universal. The following systems analysed are planar rooted tree networks which are generally thought to be similar to IIS. In addition different types of planar rooted tree network are discussed, those which distribute and resource and those which collect it.

2.4.1 Examples of Spatial Distribution and Collection Networks

A distinction should be made between planar rooted tree networks which distribute, collect or do both. All types might direct growing through preferential attachment, but if the process is a collection system then the evolution and underlying processes which create it might be quite different to a distribution system. Examples of planar rooted collection systems are leaf venation (Lee et al., 2014) and river system formation (Kyungrock and Kumar, 2008). Although dealing with systems on much different scales and with completely different building blocks, the processes are surprisingly common. Both models involve the addition of a resource which creates a non-equilibrium environment. In the case of the leaf it is the production of auxin and for a river it is water. The changing concentration of this resource as it flows down a gradient through the system leads to the creation of canals which allow for more efficient transport through the system, which in turn creates a planar rooted tree network. In leaves, the canals are created by cells changing state to become specialised for transport and in river systems erosion of the surface material leads to the creation of channels. Although within the literature on leaf venation there is debate about the role of 'pre-pattern' in the formation of these networks (Dimitrov and Zucker, 2006). It can also be argued that a 'pre-pattern' exists for river systems too, although this is not a design in the same way as for a leaf but a history of previous processes that have occurred on a landscape which can affect future river system evolution. Kyungrock and Kumar (2008) and Lee et al. (2014) both use randomness in their models, Lee et al. (2014) uses it to ensure cell division is not synchronised. Kyungrock and Kumar (2008) uses randomness in the sediment erosion in the generation of river networks and argues that the inherent randomness is sufficient to generate patterns under evolutionary dynamics. A key distinction between models which are built as networks and adds nodes using randomness leading to preferential attachment, and resource collective models such as the river system which rely on physical equations with randomness leading to the emergence of preferential attachment scale-free networks.

A distribution network which is planar and rooted on the other hand will build out from a central root node. There might be collection aspects to such a network, as the building of the network will require information about the environment in order to build an efficient structure. The growth of a tree or plant gives an example of a distribution system which has both distribution and collection aspects to it. An approach which has been very successful in modelling plants and trees and other branching self-similar structures is L-Systems (Lindenmayer, 1968; Prusinkiewicz, 2004). This can be thought of as a 'pre-pattern' technique. L-Systems is a language which is similar in some ways to the Chomsky hierarchy for formal language theory (Jäger and Rogers, 2012). It starts with a simple set of rules which through recursion will produce a series of code each time step which is basically a set of instructions for drawing the growth of the system at that time step. L-Systems might be similar to genetic code, which also must contain a set of instructions, a pre-pattern or blueprint. They have been extended to by adding stochastic processes into their formation leading to slight differences every time the model is run, but do not include any computation regarding optimality of growth, or growth direction based on local conditions. A spatial growing environment would be required for this. An iterative L-System which draws the system after each time step and feeds back information to the rules based on the growth could allow for this. L-Systems also focus on fractal self-similarity and scaling which is not the focus of this section.

Jun and Hübler (2005) conduct an experiment on the self-organisation of an electromechanical system. A power supply is connected to a source electrode at the centre of the system and a boundary electrode, which flows round the edge of the system, in this case it is circular. A number of conducting particles are placed in the system which in this case are stainless steel ball bearings. A constant supply is passed through the system until the particles organise into a stable state. Different initial configurations of the particles are tested. Multiple runs found the particles went through three stages of development. Firstly they formed strands, which allowed them to complete the circuit connecting the electrode at the centre of the system to the boundary, and then geometric expansion by the particles space-filling while maintaining the network topology.

The arrangement of the network is very dependent on initial conditions, slight differences in the the arrangement of the particles can lead to radically different final network topologies. The paper does not give a clear reason for the geometric expansion, but it is most likely to form the most efficient state to dissipate the current through, given the initial conditions. The resulting patterns formed by the particles produce statistically robust network features, in terms of the number of termini and branch points and also finds the network usually produces trees. The system has commonalities to biological systems such as the power supply requiring to be constant after the steady state is reached and the particles organising into a geometric space-filling network.

The planar network models described in this section so far have looked at both distribution and collection systems which do not necessarily follow the standard definition of preferential attachment which is increased probability of connecting to a node with greater connections (Barabási and Albert, 1999). This is most likely due to the models being spatially constrained, which means that a node added to the system will not directly attach to the most popular root node, but the most efficient path to it. So it is more likely to connect to the shortest path to the root node. This is assuming there are not irregularities in the environment which may change this path.

Finally for a previous example on data collection and analysis of distribution networks, Papadopoulos et al. (2018) provides a comparison of two biological distribution systems, mycelial fungi and vasculature from the surface of rodent brains. The two systems vary in terms of growth, transport mechanisms and environment in which they exist but are both planar, in two-dimensional space and transport fluid and nutrients. In order to gather data on these networks, the rodent vasculature network was traced by hand from images and the mycelial network was extracted digitally. The networks were then stored in adjacency matrices which allows for efficient analysis. This paper highlights the commonalities between two distribution systems but also the variability in how they have evolved which could reflect differences in function, environmental condition or development. The vasculature network is organised for low cost, high efficiency distribution, whereas the mycelia forms more expensive but in turn more robust networks. It would be interesting to see if any form of comparison can be made with IIS as well.

2.4.2 Space-Filling in Networks

'Space-Filling' (or space-optimisation) is the structural features of a distribution network which are able to supply all parts of a particular system, for example the capillaries in a cardiovascular system (Hunt and Savage, 2016) and biological systems in general (West et al., 1997). Hunt and Savage (2016) look at different strategies to build networks which fill space in cardiovascular networks. Some strategies do not lead to biologically adaptive structures as they are too inefficient - they deliver blood too slowly, require too much construction material

or use too much power to move blood around the system. They find that a trade-off between minimum use of materials and minimum path length enforce thresholds for balance in the network configuration which match the empirical data better. Hunt and Savage (2016) find that within the vascular system, there is a sharp descent from optimal conditions in vascular networks. This implies that vascular networks are under strong selection for space-filling and efficiency. Given the length of time which such systems have evolved over it is possibly unsurprising that the optimal state has been reached.

From the literature review it seems like IIS also have space-filling properties. But whether they have a similar level of efficiency to biological systems is uncertain. Given the networks in biological systems have evolved over billions of years they are likely to be more efficient than an IIS that has developed only over thousands of years.

2.4.3 Branching in Tree Networks

The general mathematical study of the *Branching Processes* has been undertaken for many years for studying processes from particle physics, river systems, vein and lymphatic channels in living systems and population biology (Caldarelli et al., 2000; Harris, 1964). Branching processes in nature often follow fractal structures (Caldarelli et al., 2000), which is characterised by having similar properties at all length scales (Mandelbrot, 1983).

The probability distribution $P(k)$ of tree size k for a network is a widely used form of analysis for comparing different tree networks (De Los Rios, 2001). The exponent τ which is defined as $P(k) \sim k^{\tau}$ is a universal character used to compare tree networks. It is equal to the gradient of the probability distribution P(k) on a log-log plot. The value of τ relates to the growth rate, the hierarchy and centrality of the network (De Los Rios, 2001). A network which has a branching ratio m,

$$m = \sum_n n p_n \tag{2.10}$$

where, p_n is the probability of n new events occurring. When m = 1, each generation of branching is on average identical to the last (De Los Rios, 2001), this is described as being *critical* behaviour and leads to $P(k) \sim k^{-3/2}$ (Harris, 1964). If m>1, then the system is said to be *supercritical* leading to $P(k) \sim k^{-2}$ (De Los Rios, 2001). Increasing the value of m, will lead to the system having a much quicker growth rate, a stronger hierarchical structure and increased centrality.

The value of τ has been measured for many real-world networks. For river systems it has been measured as τ=1.43 +/-0.02 (Cieplak et al., 1998). The lower range of values is thought to be due to river systems approaching a state of minimal energy dissipation (Cieplak et al., 1998). For the internet, τ=1.9 +/-0.1 (Caldarelli et al., 2000).

Therefore, the river system is classified as a Critical (or slightly sub-critical system) and the internet as a supercritical system. This is interesting particularly relating back to the difference between human and natural systems and also the difference between spatial and non-spatial systems.

Given IIS are spatial systems, similar to river systems, they are likely to have sub-critical to critical properties.

2.4.4 Summary

The review of the literature has found many commonalities across different types of planar network. All the systems under study produced networks which increased efficiency for the particular system and were modelled as open dissipative systems with an energy gradient (leaf venation, river systems and electromechanical particles). L-Systems being the only exception to this which is a pattern forming system.

Computational modelling of such systems leads to different insights as oppose to experiments. Both the leaf venation and river system models have to add randomness to the model, this induces asynchronous behaviour, which seems to be required for the system to search for efficient networks. In real world experiments this is not required as perfect order is never present so asynchronous behaviour is inherent in real-world systems. To some extent this can be a limitation of computer models, but by engineering asynchronous behaviour, the results tend to match real world systems.

As stated in the section on leaf venation, there is discussion on whether a leaf has a pre-pattern which influences the way in which leaf veins form. The same question can be asked about indigenous irrigation systems. Whilst an irrigation system will be planned to some extent, it is unlikely that given the amount of effort taken to build one it is unlikely to be rebuilt again and again. But if a number of systems are built local to one another then knowledge gained in building one is likely to be passed on to building the next. This means that there might be a memory or blueprint or event gene for a particular type of indigenous irrigation system in a particular area so an evolutionary approach to modelling can also be taken to looking at such systems.

Chapter 3

Research Questions

This section outlines the main research questions established from the gaps in the literature.

3.1 Research Gaps

As a reminder the research gaps are as follows:

- IIS seem to consist of space-optimising branching tree networks forming a hierarchical structure. Quantitative analysis of different IIS networks has not yet been carried out to confirm this, which in turn can be used to find commonalities and differences between IIS.

- No studies have looked at the relationship between managerial arrangements and network structure of IIS.

- There is no recognised model for how IIS tend to grow, do they follow a biological type sigmoidal curve, sociological expansion? Is this related to the managerial arrangements?

- Previous studies have looked at the development of inequality in simplified game-theoretic models of IIS, but not for larger network based simulations.

- The long term stability of IIS has not been explored from a network perspective. How do the managerial arrangements and network structure interact to maintain stability for long periods?

- No quantitative models have been built to look at the effects of globalisation on IIS.

3.2 Research Questions

3.2.1 Chapter 4 - Space Optimising Growth in Indigenous Irrigation Systems (IIS)

1. Previous models of Indigenous Irrigation Systems have hypothesised growth based on a 'Budding Model' where local parts of the network will expand downstream based on local conditions (Lansing et al., 2009). This contrasts to a model in which a central power controls the growth. A generative agent-based model is built in this Chapter to ask whether growth based on local conditions leads to an optimal system, or is a central power required?

2. By producing an optimal system, the network will optimise the amount of space which the system can distribute to. This is a common property of self-organising dissipative systems in nature. Will the simulated IIS have common network properties to other such networks such as power law scaling?

3.2.2 Chapter 5 - Empirical Data Collected on Indigenous Irrigation Systems

1. Do the networks formed by real life IIS share common characteristics to the simulated IIS, therefore validating the model?

Chapter 4

Space Optimising Growth in Simulated Indigenous Irrigation Systems (IIS)

4.1 Introduction

The literature review introduced a number of general models and theories for studying Complex Systems and Networks. Numerical methods for modelling these systems are then introduced which focus on cellular, agent-based and network models. Finally case studies of the systems to studied for this research are reviewed with a focus on Indigenous Irrigation Systems, but also other systems which share common features.

This Chapter aims to add new knowledge by building a simulation model of an indigenous irrigation system which can be used to test existing theories on their organisation and compare to other networks with similar properties.

The STRESS Guidelines for simulation modelling have been followed to enable transparent reporting and repeatability (Monks et al., 2019).

4.2 Research Questions

The model specifically aims to answer the following Research Questions:

1. Previous models of Indigenous Irrigation Systems have hypothesised growth based on a 'Budding Model' where local parts of the network will expand downstream based on local conditions (Lansing et al., 2009). This contrasts to a model in which a central power controls the growth. A generative agent based model is built in this section to ask whether growth

based on optimal local conditions leads to an optimal global system, or is a central power required?

2. By producing an optimal system, the network will optimise the amount of space which the system can distribute to. This is a common property of self-organising dissipative systems in nature. Will the simulated IIS have common network properties to other such networks found in nature such as power law scaling?

4.3 The Model

The model created in this section is a generative pattern formation cellular automata network. It relies on local rules which when iterated over time leads to a pattern emerging. This allows for simplicity when coding with minimal assumptions and leads to precise answers for a given pattern forming. As stated in the introduction and research questions, the goal is to look for efficient states of a space filling network by testing different theories of organisation of an Indigenous Irrigation System.

Pattern formation models based on local rules have previously been used for studying segregation in artificial societies (Schelling, 1969; Collard et al., 2013), evolutionary game theory (Nowak and May, 1992b), flocking behaviour (Reynolds, 1987), and the simplest models of artificial life such as Game of Life (Adamatzky, 2010). Cellular Automata is not usually associated with networks as networks tend to model non-spatial interaction and cellular automata tend to model local spatial interaction. Yang and Yang (2007) is the only article found which attempts to model networks in cellular automata but also includes non-spatial interaction. As mentioned in the literature review, networks (specifically tree networks) are often an emergent feature of local interaction in cellular automata models such as in rivers (Kyungrock and Kumar, 2008), leaf vein formation (Lee et al., 2014), heat conduction (Yu and Li, 2006) and electro-mechanical systems (Jun and Hübler, 2005). The common underlying principle of these systems is the dissipation of energy down gradients. This is largely due to the network being for distribution rather than collection. A large-scale computational model of an IIS has previously been built to understand how local synchronisation can lead to optimal yields of the system as a whole, but it is static and does not simulate growth (Lansing and Kremer, 1993). Small-scale infrastructure models have been built to explore how individual decisions of farmers (conformists or opportunists) affect the efficiency in terms of maintenance and income inequality of the system (Yu et al., 2015). The model in this section differs, by forming a tree network using a cellular automata, but does not use the process of energy dissipation down gradients, but instead grows based on space filling principles for local conditions.

It is set up to have one initial distribution root node (or cell) which is placed

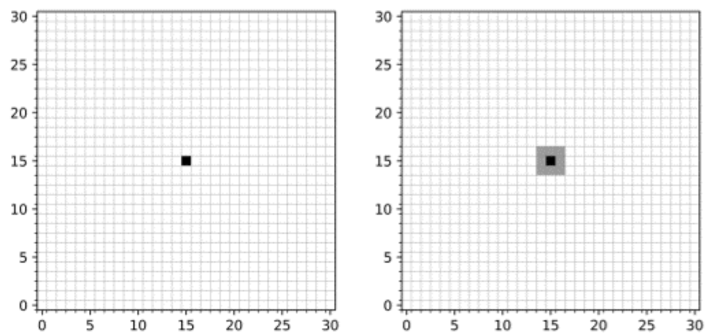


Figure 4.1: *Left:* Initial distribution cell of space optimising algorithm on a 30 by 30 grid with single starting point in the centre (15, 15) at time step 0. *Right:* Initial configuration of single Distribution Cell (black) and 8 Receiver Cells (grey) using a Moore Neighbourhood with r=1.

in the centre of a large enough grid so the growing network is not effected by boundary conditions. To search for local rules on a space filling network the system has to be allowed to grow, such as the preferential attachment model (Barabási and Albert, 1999). This is done by attaching further distribution nodes to those already present in the network preferentially. Preferential attachment in this network is based on space-filling, but instead of filling space another type of cell is added to the network, receiver cells. The network grows by adding distribution cells which lead to the highest amount of receiver cells. This is explained further below.

Consider a two dimensional grid with a single starting cell at the centre (x_0, y_0) at time step t=0, Figure 4.1 (*Left*). The size of the grid is irrelevant as the simulation is voided if the system grows beyond the boundary.

Each distribution cell at (x_0, y_0) will add all the receiver cells in empty cells in its Moore Neighbourhood ($N^{M}_{(x0,\ y0)}$), which is defined as:

$$N^{M}_{(x0,y0)} = (x, y) : |x - x_0| \leq r, |y - y_0| \leq r \tag{4.1}$$

The size of the Moore Neighbourhood is given by: $(2r+1)^2$. For this model $r=1$. The Receiver Cell at (x_0, y_0) is not included in the Moore Neighbourhood as this is the location of the distribution cell. So the maximum number of Receiver Cells for each Distribution Cell is 8, giving a ratio of 1:8. This is shown in Figure 4.1 (*Right*). Systems with larger neighbourhoods can be explored with r>1, but this is beyond the scope of this section.

The space which the system grows over is homogeneous - the cell sizes are all the same and all cells have the same value. The growth of other parts of the system are the main effect on whether adding a Distribution Cell to a particular one will increase the number of Receiver Cells. This is not the case in reality, the environment is heterogeneous. There are ecological and geomorphological factors which play a role in the optimal network growth.

4.3.1 Growing the System

Growing the network will allow for efficient states to be found. This is done by adding Distribution Cells to those already in the network. For this model, it is assumed that a distribution cell at (x_0, y_0) can only expand directly adjacent horizontally and vertically (not diagonally), also known as the Von Neumann neighbourhood ($N^{V}{}_{(x0,\ y0)}$) which is defined as:

$$N^{V}_{(x0,y0)} = (x, y) : |x - x_0| \leq r + |y - y_0| \leq r \tag{4.2}$$

The size of the Von Neumann Neighbourhood is given by: *2r(r+1)+1*. For this model *r=1*. The Cell at (x_0, y_0) is not included as this is the location of the current Distribution Cell. So the maximum number of Distribution Cells which can be connected to one is 4, giving a ratio of 1:4. This is shown in Figure 4.2.

Model Rules

So far the rules have been outlined for where Distribution and Receiver Cells can be added.

As this model is attempting to find the most efficient states for a growing planar network to test either the 'Budding model' or central authority model, Distribution Cells will only be added if they increase the amount of Receiver Cells in the network.

Each time step the Distribution Cell selected will check its Von Neumann Neighbourhood and test by adding a Distribution Cell in each available position. A position is available if there is not already a Distribution Cell present. It is also available if a Receiver Cell is present in that position; the Receiver Cell will be removed in that position and a Distribution Cell added. A Distribution Cell is added if it increases the number of Receiver Cells in the system. As this model is looking at increased efficiency Distribution Cells which increase the number of Receiver Cells by the most are added. Once a Distribution Cell is added, its associated Receiver Cells are also added.

Figure 4.3 shows the initial configuration of the model (*Left*) and the first options in which the model can grow (*Right*). As can be seen, all the options lead to the same increase in Receiver Cells. Three will be added, but one taken away where the new Distribution Cell is added.

This same basic algorithm is iterated to add more Distribution Cells and Receiver Cells. The only difference in the models constructed in the remainder of

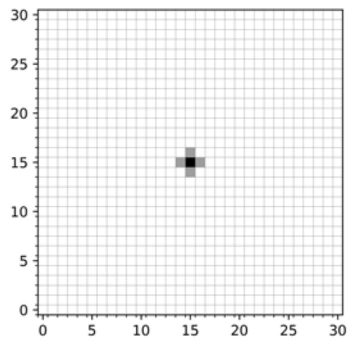


Figure 4.2: Initial configuration of single Distribution Cell (black) at coordinates (x_0, y_0) The Von Neumann neighbourhood is highlighted in grey, which is the locations where additional Distribution Cells can be added.

this section is subtle changes in the preferential attachment algorithm, that is changes in terms of which Distribution Cell is selected to be added.

Four alternative models of growing the network are explored and analysed to see which is most efficient. The changes reflect how much information is shared across the system in deciding the placement of distribution cells. This is to test to what extent the 'Budding model' or central authority model is more efficient. These are as follows:

- Local Deterministic and age dependent;
- Local Stochastic and age dependent;
- Global and Local stochastic;
- Global Selection with local stochastic growth.

Firstly *Local* and *Global* are defined. *Local* means that preferential attachment at each time step is decided locally for each Distribution Cell in the network (Figure 4.3, *Right*). As such, Distribution Cells are effectively added in parallel in a synchronous manner. *Global* means that preferential attachment is decided globally; the control of growth is taken at the level of the whole network. Examples in each of the models will provide more information on how this is applied.

A description of each model is given below.

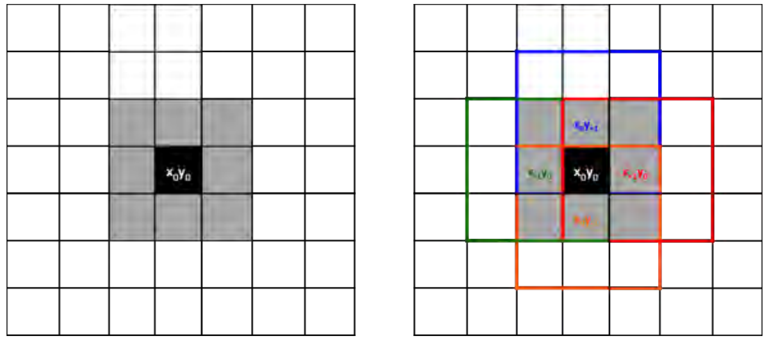


Figure 4.3: *Left:* Initial Distribution Cell (black) with Receiver Cells (Grey). *Right:* Locations which new Distribution Cells can be added with the corresponding Receiver Cells are highlighted in different colours.

4.3.2 Local Deterministic and Age Dependent (LDA)

A deterministic algorithm is used (without any stochastic processes) and by being age dependent the system preferentially allows Distribution Cells which have been present in the system for longest to grow. This is the simplest growing scenario explored as it produces the same result each time. A preferential growth direction is added for the Von Neumann Neighbourhood in the following order (Left, Up, Right, Down) to allow for scenarios where the maximum amount of Receiver Cells can be added in multiple directions. This creates slight asymmetry in the network as can be seen in Figure 4.4.

This can be thought of a strict example of the 'Budding Model' where each distribution cell grows according to their own local environment without any communication with other cells.

4.3.3 Local Stochastic and Age Dependent (LSA)

This model is very similar to the Deterministic and Age Dependent model (Section 4.3.2). The only difference being that in scenarios where there are multiple growth directions for a given Distribution Cell, the next Distribution Cell is chosen from a uniform distribution with probability 1/n where n is the number of growth directions. It is therefore another example of a strict 'Budding Model'.

4.3.4 Global and Local Stochastic (GLS)

The global stochastic growth model disregards any preferential growth to older cells in the system. It will select Distribution Cells anywhere in the network which increase the number of Receiver Cells at using a probability of 1/n where n is the number of growth options. The model also has local stochastic processes

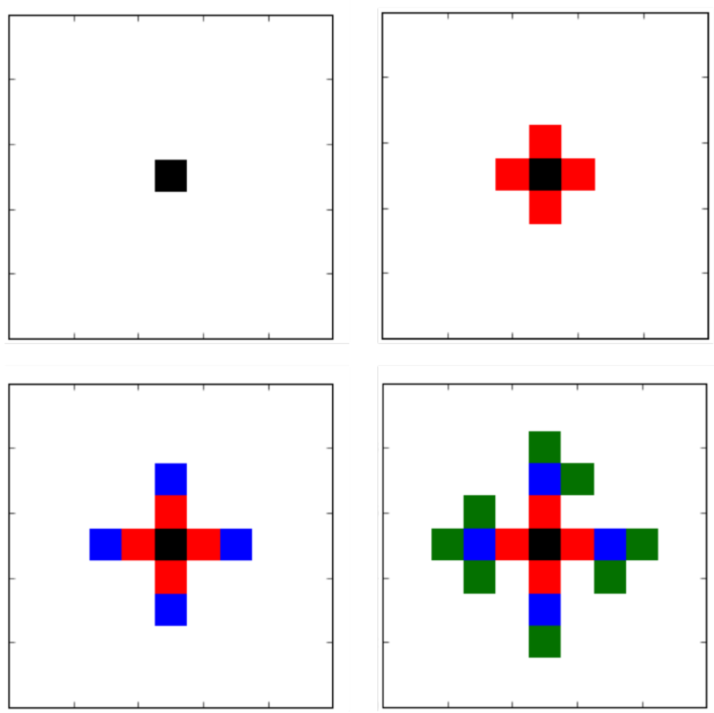


Figure 4.4: Deterministic and age dependent Model, showing the growth over four time steps with colour coding for each time step.

similar to the LSA model. The position where additional Distribution Cells is added is selected from a uniform random distribution of positions which will increase the number of Receiver Cells in the network locally. However there is no selection of Distribution Cells which increase the number of Receiver Cells by the most.

This model is different to the strict 'Budding Models' seen in the first two models. The growth of the system is not synchronised, so it can described as asynchronous and contains inherent randomness.

4.3.5 Global Selection and Local Stochastic (GS)

The Global Selection Model gathers information on all current Distribution Cells at each time step and selects the Distribution Cell which will increase the Receiver Cells in the system by the maximum amount. If there are multiple Distribution Cells which give the same maximum Receiver Cells, then one is selected at random, with a probability of 1/n where n is the number of growth options. It is locally stochastic, so if for one particularly Distribution Cell there are multiple directions to grow which will give the same maximum amount of Receiver Cells, then one is chosen at random.

This model more closely resembles a central authority controlling the growth of the network. All the information of the network is collected and the growth is decided based on which distribution cells will increase the space the most. The model can be adapted to be deterministic by preferentially selecting one direction to grow in.

Further information on the model can be found in Appendix A.

4.4 Results

The results of the simulations need to be presented in a way that allows the Research Questions to be answered. Firstly an output from each simulation for 1000 Distribution Cells is shown in Figure 4.5 which at first glance gives an indication of how each model performs. The Local models (Top Left and Top Right) form well-ordered patterns, where as both global models (Bottom Left and Bottom Right) form seemingly more disordered patterns.

The following sections present results from each simulation in order to answer the two Research Questions in the introduction.

4.4.1 Research Question 1 Results

The first Research Question asks whether the 'Budding model' of network growth based on local conditions and decisions or the 'central authority model'

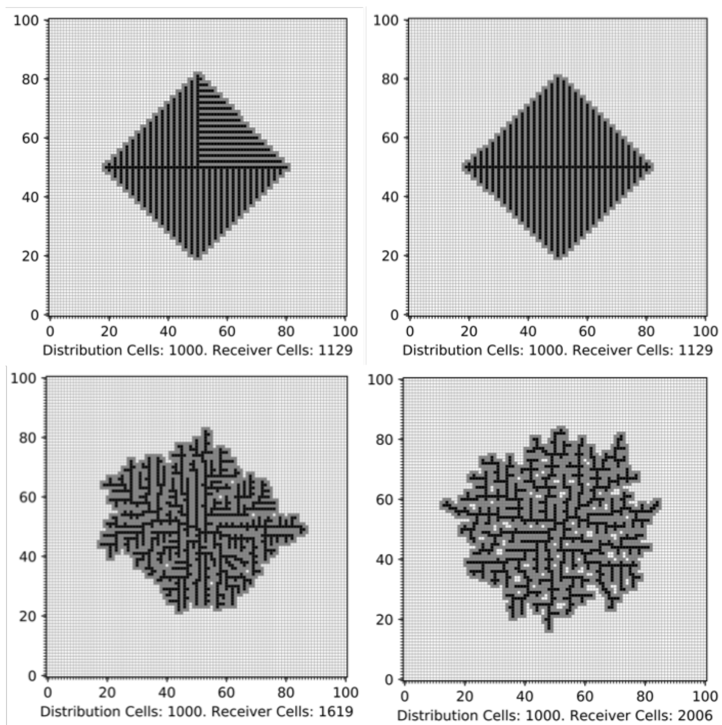


Figure 4.5: Each of the Models run for 1000 Distribution Cells. The number of Receiver Cells for each model is given. *Top Left:* Local Deterministic and Age Dependent, *Top Right:* Local Stochastic and Age Dependent, *Bottom Left:* Global and Local Stochastic and *Bottom Right:* Global Selection with Local Stochastic Growth.

of growth based on a central decision maker will give an optimal solution.

There are different ways of measuring optimality. The most obvious is to look at space-optimisation, which is the ratio of Receiver Cells to Distribution Cells; the higher the ratio the more optimal the model is. Another way would be to look at the time involved for the model to grow the network. A model which takes longer is less efficient at completing the task and analogous can be drawn to the additional work required in the real system.

Figure 4.5 shows that for 1000 Distribution Cells, the GS model gives the optimal ratio, followed by the GLS model and the LDA and LSA models. Many runs of the different models have subsequently been simulated to see if this is the same property at different scales, Figure 4.6. This shows all the models (generally) follow linear growth, with the GS model always producing the optimal arrangement. The GS model follows a simple linear equation of:

$$N_{RC} = 2N_{DC} + 6 \tag{4.3}$$

p-value = 3.6e-149, R^2 = 1.000 (using scipy.stats.linregress).

where, N_{RC} is the number of Receiver Cells and N_{DC} is the number of Distribution Cells. If the growth of the network continues to follow that seen in Figure 4.3 - each Distribution Cell added increases the total Receiver Cells by two, then this equation is followed.

The GLS model fits a linear equation of (Using Python module *numpy.polyfit.*):

$$N_{RC} = 1.6N_{DC} + 19 \tag{4.4}$$

p-value = 4.07e-21, R^2 = 0.998 (using scipy.stats.linregress).

for lower values but loses efficiency with larger values (creating a logistic curve), likely due to the randomness and lack of selection in the model.

The LDA and LSA models follows the linear equation (Using Python module *numpy.polyfit.*):

$$N_{RC} = 1.09N_{DC} + 47 \tag{4.5}$$

p-value = 2.56e-33, R^2 = 0.999 (using scipy.stats.linregress).

for lower values, but again loses efficiency at larger values like the GLS model.

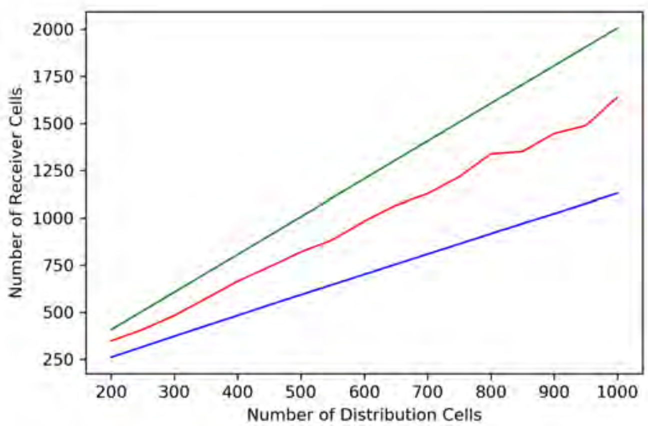


Figure 4.6: Plot of Distribution Cells and Receiver Cells for each of the models. *Blue* - LSA/LDA Model, *Red* - GLS Model and *Green* - GS Model.

To test the second type of optimality, the growth of the model was timed for each model for different sizes. The results are shown in Figure 4.7. The results have been plotted on a log-scale due to the GS model taking approximately $10^{2.25}$ longer to complete showing the optimality of the GS model comes at a cost of time to (see Figure 4.8).

4.4.2 Research Question 1 Discussion

The results show that finding a model which gives the best performance in space-optimising comes at a cost in terms of computation time. This section discusses the models in the context of the research question.

The LDA and LSA models are described as strict 'Budding models' in which the decision for each Distribution Cell to expand downstream is kept strictly local and growth is preferentially dependent on age in the system. The time taken to produce patterns is fast but the ratio of Receiver Cells to Distribution Cells is low. The way that the model has been coded means that each cell is completely synchronised, each cell which can grow each time step is allowed to grow. Although in reality, the growth of different parts of the system may occur at a similar rate given they have similar technology, but it will not be exactly synchronised. The pattern formed by the model is very ordered and similar to what a planned system might look like. These models are therefore deemed to not represent a growth and pattern which would occur outside of a computer model.

The GS simulation finds the optimal space-optimising system but at a much greater computation cost. It shows that collecting all the information and mak-

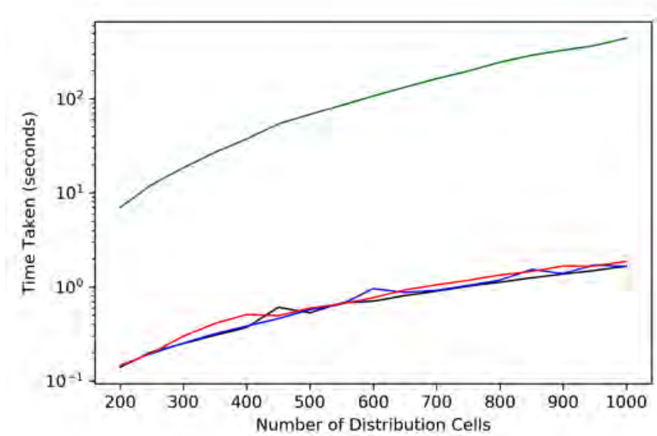


Figure 4.7: Time Taken for a simulation of each model for different system sizes. *Black* - LDA Model, *Blue* - LSA Model, *Red* - GLS Model and *Green* - GS Model.

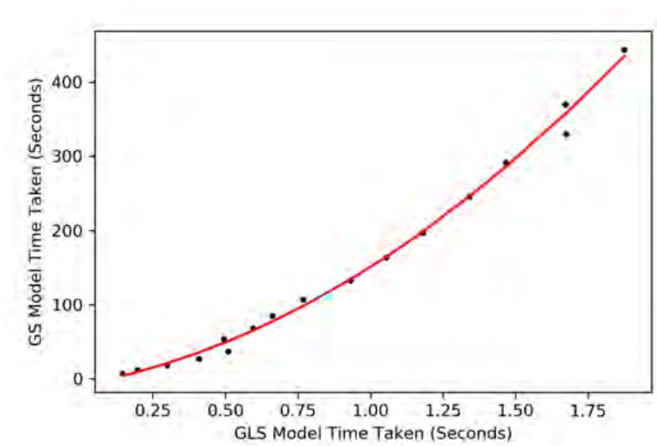


Figure 4.8: Prediction of the time taken for the GS model (y) using the GLS model (x), the best fit line was found using *numpy.polyfit*, $y = 88.3x^2{+}70.7x{-}8$. p-value = 9.55e-13, $R^2 = 0.969$ (using scipy.stats.linregress).

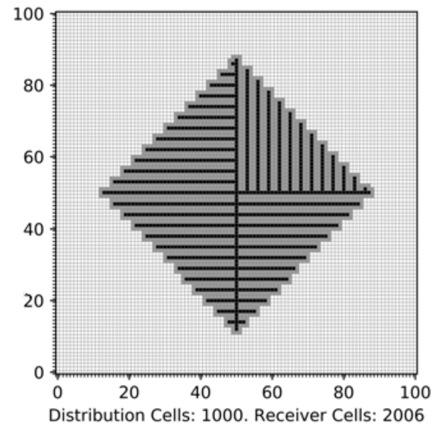


Figure 4.9: Deterministic Ordered version of the GS Model.

ing a decision such as with a central authority can be suboptimal when considering the time it takes.

The GLS model relaxes the age dependent assumption and instead relies on randomness. The effect of this is that different parts of the system grow at different rates, however older cells are still closer to the initial cell and younger cells at the edge of the system. The effect of the randomness increases the space optimisation of the system without increasing the time taken. The randomness essentially causes the model to behave in an asynchronous manner allowing it to search for sub-optimal solutions. The use of randomness in Complex Systems modelling has previously been highlighted for its importance (Kyungrock and Kumar, 2008; Lee et al., 2014). Nature appears to be asynchronous and simulations which model it should be too. The GLS model seems to be the best representation of the Budding Model, it relies on local information, but with added randomness to simulate natural behaviour.

In a system which relies on a central authority, it is likely that all the local information will not be collected such as in the GS model. Much of it would be assumed. This would mean that the model would perform quicker, and would likely have a much ordered shape such as the LDA and LSA models. The GS model can be modified to be deterministic producing an ordered state, as shown in Figure 4.9. Also in the current form of the GS model, the information is collected each time step; this is inefficient and can be improved by only collecting and updating information from parts of the network which have changed.

It is hypothesised that a budding model will perform better in a heterogeneous environment with vastly changing local conditions (such as ecological, geomorphological, or sociological) whereas a central authority system will perform better in homogeneous environment where greater assumptions can be made. The central authority will also be required to have knowledge or a design of a system which can be repeated, similar to a 'pre-pattern' in leaf growth.

As mentioned in the model description (Section 4.3), the value of r in Equation 4.1 is constant = 1 for all iterations of the model. Increasing the value of r produces space-filling networks more similar to IIS, as the distribution cells cover proportionally a much smaller area than the receiver cells (see Appendix A.3 for an initial exploration in increased values of r). There is likely to be similar simplistic linear equations for the GS model for higher r values.

To find a general solution, Equation 4.3 has to be written in terms of r. One version of this would be:

$$N_{RC} = 2rN_{DC} + (2r+1)^2 - (2r+1) \tag{4.6}$$

The working for this equation can be found in Appendix A.3. This equation holds as long as the network grows at maximum efficiency. The number of Receiver Cells ($\mathrm{N_{RC}}$) will scale proportionally to the number of Distribution Cells ($\mathrm{N_{DC}}$), given a constant value of r. As can be expected from the equation, the ratio of distribution cells to receiver cells always scales linearly. This equation also only gives the number of Distribution and Receiver Cells, it does not give information about the network structure which is particularly important for comparison with real-world networks.

As this pattern forming model is not time dependent, it is merely producing patterns for a given number of distribution cells, the growth rate is not measured or constrained and so cannot be related to scaling processes through time such as in (Bettencourt et al., 2007).

4.4.3 Research Question 2 Results

This section presents results to answer the second research question which asked whether the simulated IIS will have common properties with other complex networks such as power law scaling. To answer this question, network analysis techniques from Jun and Hübler (2005) are used. This includes, cells downstream, number of branches, and number of termini, these are shown in Figure 4.10. A characteristic universal exponent (τ) is also measured for each model (Caldarelli et al., 2000; Cieplak et al., 1998; Jun and Hübler, 2005), which is the absolute value of the slope when plotting the probability distribution against the cells downstream. Section 2.4.1 provides an overview of the experiment undertaken on an electromechanical system by Jun and Hübler (2005). Although the experiment is on a completely different system, it has many commonalities with The Model. The system consists of many components and organises into

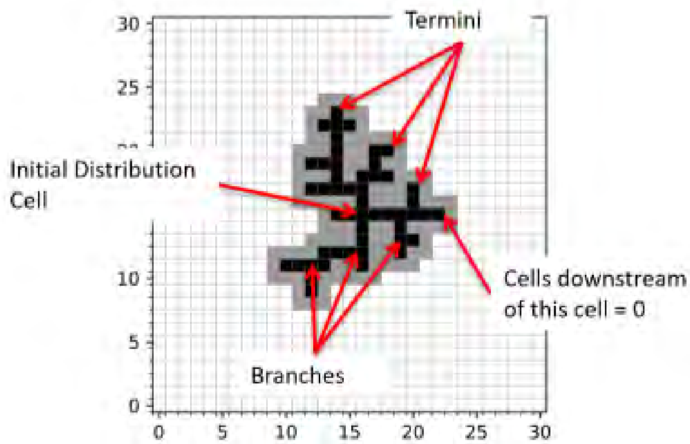


Figure 4.10: Definitions of *Termini*, *Branches* and *Cells Upstream* for Planar Tree Networks. Downstream is taken as being towards the termini. Taken from Jun and Hübler (2005).

a space-optimising state and starts from a single source which organises the system. Figure 4.11.

For each model, simulations were undertaken with a Distribution Cell range between 100 and 2000 (inclusive) with an interval of 100 giving a total of 20 simulations per model. Lines of best fit are drawn for the scatter plots using the linear regression function of the sklearn module in python. The power law line of best fit is found using the python module *Powerlaw* (Alstott et al., 2014), which uses a combination of maximum-likelihood fitting methods with goodness-of-fit tests (Clauset et al., 2009). The x_{min} and x_{max} values which are the starting and end points of the power-law behaviour in the data is estimated visually from the probability distribution plot, allowing for a more accurate estimate of the power law curve fit. The power law curve is compared to the data for significance using the KS statistic which is a goodness-of-fit measurement (Clauset et al., 2009), generating a *p-value*. A significance threshold of p>0.1 is used, similar to Clauset et al. (2009).

The results of the LDA model are shown in Figures 4.12 and 4.13. Similar to the Distribution Cells verses the Receiver Cells plot in Figure 4.6, the branches, termini and average path length follow logistic growth, Figure 4.12. However, a straight line linear model has still been applied to allow for easier comparison. Figure 4.13 shows the probability distribution plot. Both the histogram and log-log plot show no evidence that the process follows power law scaling for this model.

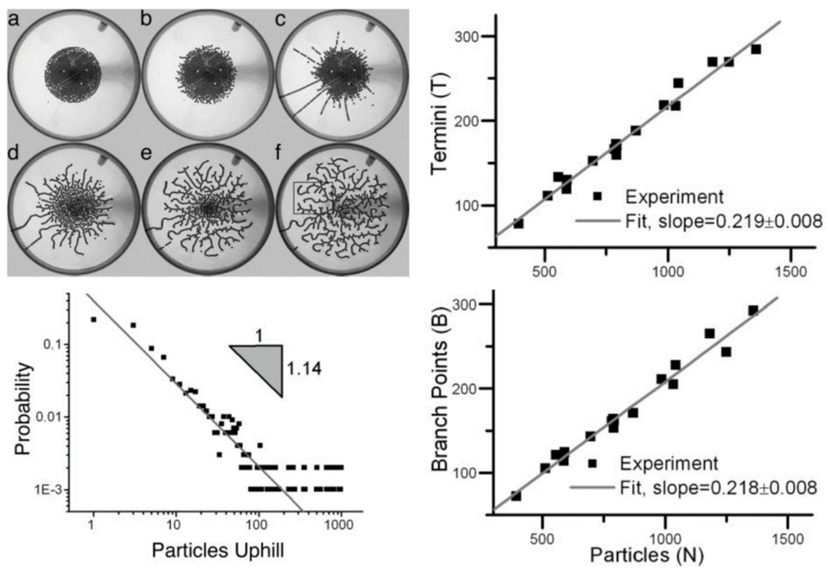


Figure 4.11: Extract taken from Jun and Hübler (2005). *Top Left:* Time sequence of single experimental run. *Top Right:* Number of Termini plotted against number of particles in the network. *Bottom Left:* Density Distribution P(n) of the number of particles uphill for a steady state network (Uphill is the same as Cells Downstream). *Bottom Right:* Number of Branch Points plotted against number of particles in the network.

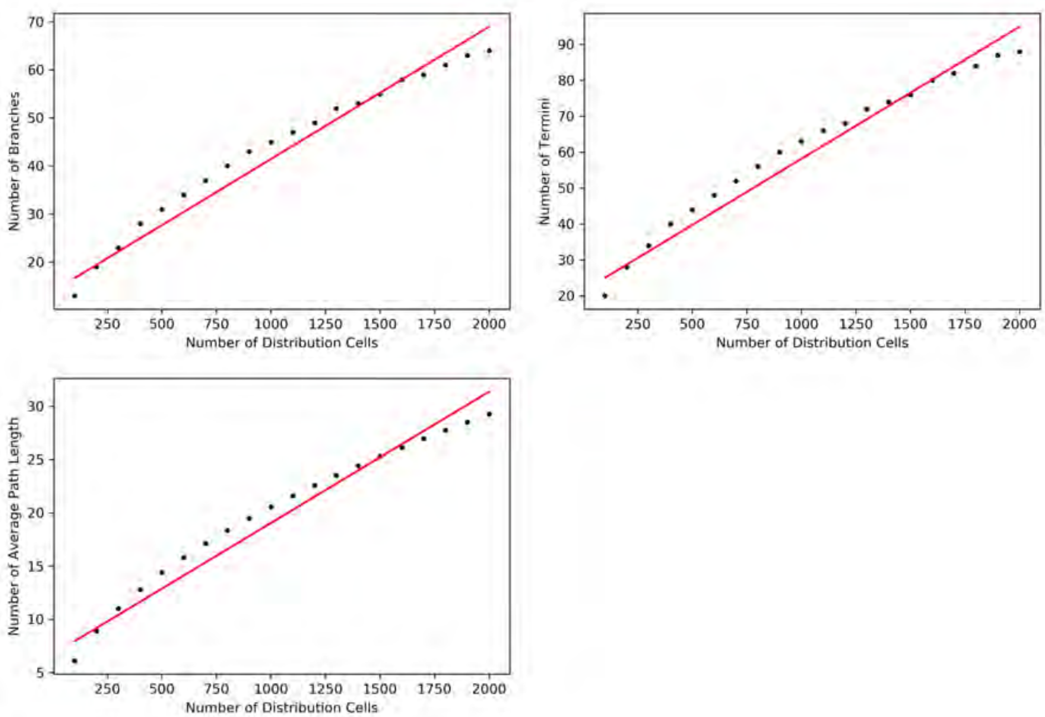


Figure 4.12: Results from LDA Model. *Top Left:* Number of Branches plotted against the number of Distribution Cells. The slope gradient is 0.026. *Top Right:* Number of Termini plotted against the number of Distribution Cells. The slope gradient is 0.034. *Bottom Left:* Average Path Length plotted against the number of Distribution Cells. The slope gradient is 0.012.

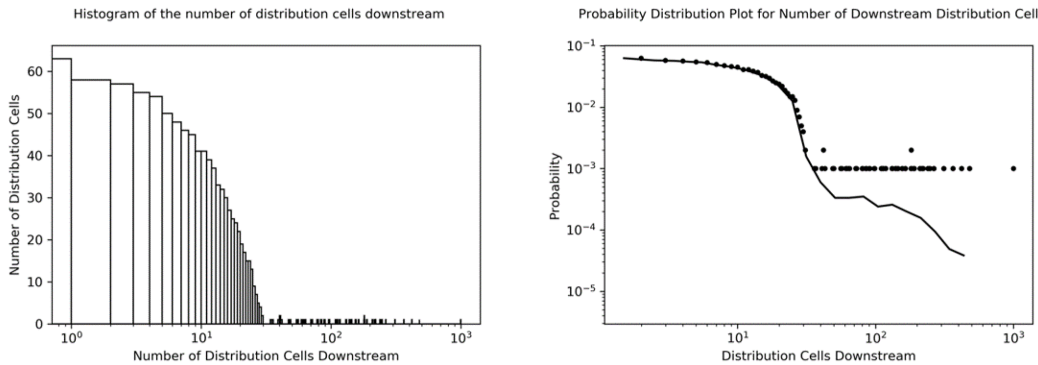


Figure 4.13: Results from LDA Model with 1000 Distribution Cells. *Left:* Histogram showing the Number of Distribution Cells Downstream for each Distribution Cell. *Right:* Probability Distribution for Number of Distribution Cells Downstream. No power law observed. The black dots are a probability distribution with linearly spaced bins and the black line is the probability distribution using logarithmically spaced bins from the powerlaw module.

The results of the LSA model are shown in Figures 4.14 and 4.15. The termini and average path length in Figure 4.14 are almost identical as the LDA model, but the number of Branches bifurcates into two curves. From visual inspection, power law relationship does not seem to apply to this model.

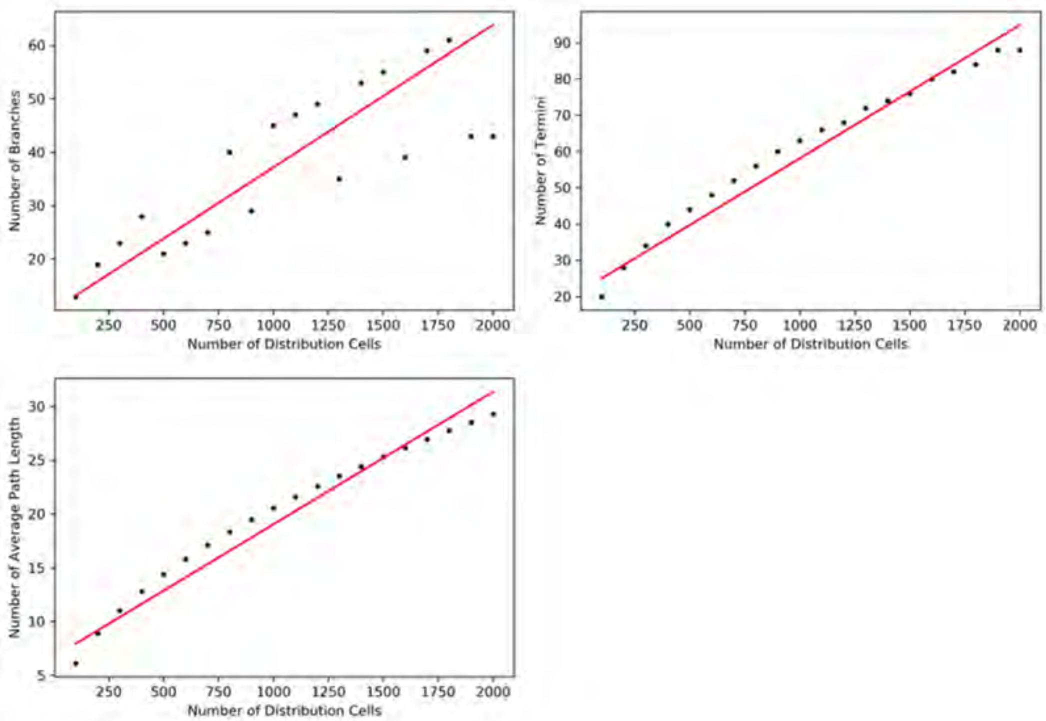


Figure 4.14: Results from LSA Model. *Top Left:* Number of Branches plotted against the number of Distribution Cells. The slope gradient is 0.024. *Top Right:* Number of Termini plotted against the number of Distribution Cells. The slope gradient is 0.035. *Bottom Left:* Average Path Length plotted against the number of Distribution Cells. The slope gradient is 0.012.

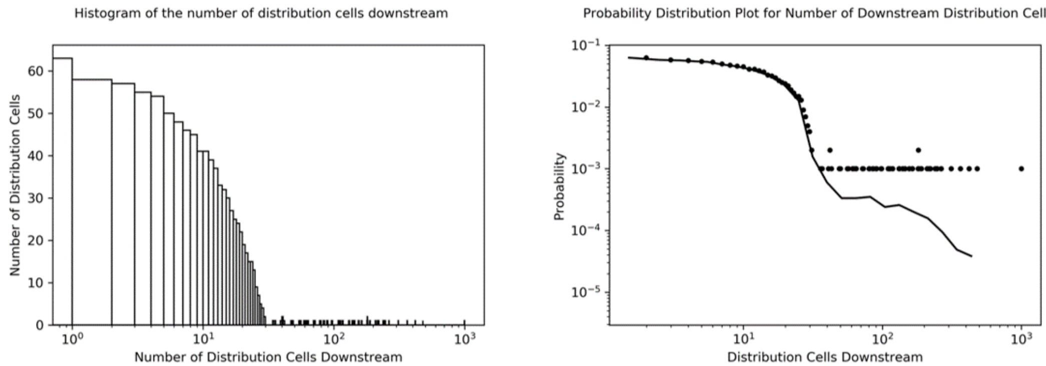


Figure 4.15: Results from LSA Model with 1000 Distribution Cells. *Left:* Histogram showing the Number of Distribution Cells Downstream for each Distribution Cell. *Right:* Probability Distribution for Number of Distribution Cells Downstream. No power law observed. The black dots are a probability distribution with linearly spaced bins and the black line is the probability distribution using logarithmically spaced bins from the powerlaw module.

The results of the GLS model are shown in Figures 4.16 and 4.17. Unlike the LDA and LSA models the branches, termini and average path length show a linear relationship with number of distribution cells, Figure 4.16. The data also fits the power law curve in Figure 4.17 with τ=1.49. Figure 4.18 shows the results of 1000 simulations for 50 Distribution Cells using the GLS algorithm and produces a normal distribution.

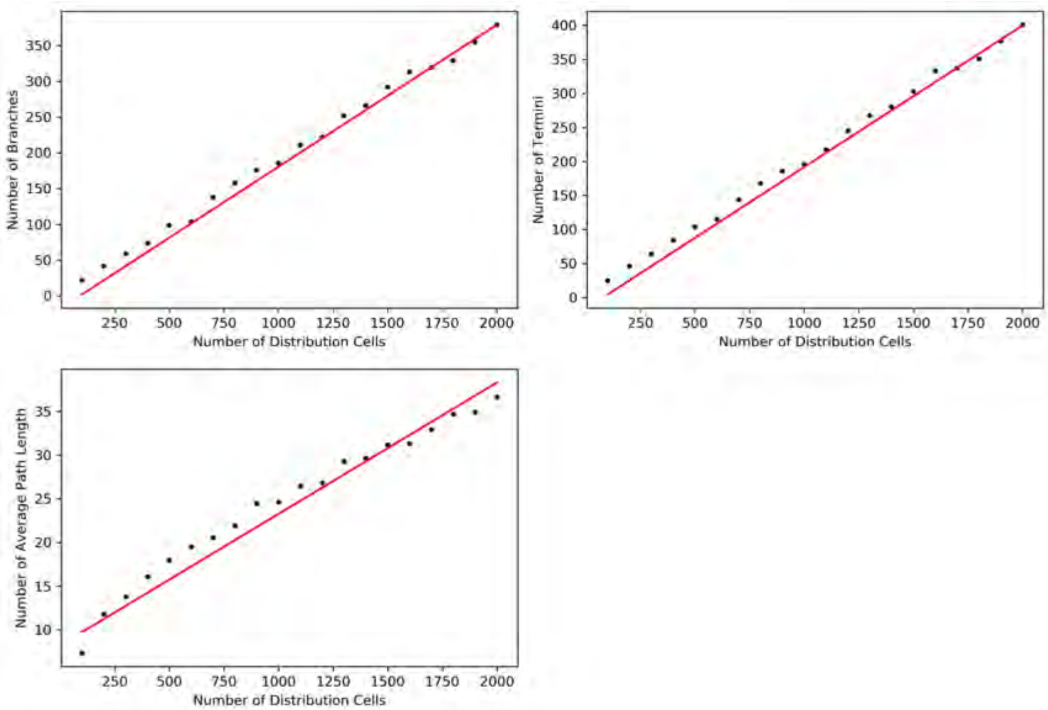


Figure 4.16: Results from GLS Model. *Top Left:* Number of Branches plotted against the number of Distribution Cells. The slope gradient is 0.19. *Top Right:* Number of Termini plotted against the number of Distribution Cells. The slope gradient is 0.20. *Bottom Left:* Average Path Length plotted against the number of Distribution Cells. The slope gradient is 0.014.

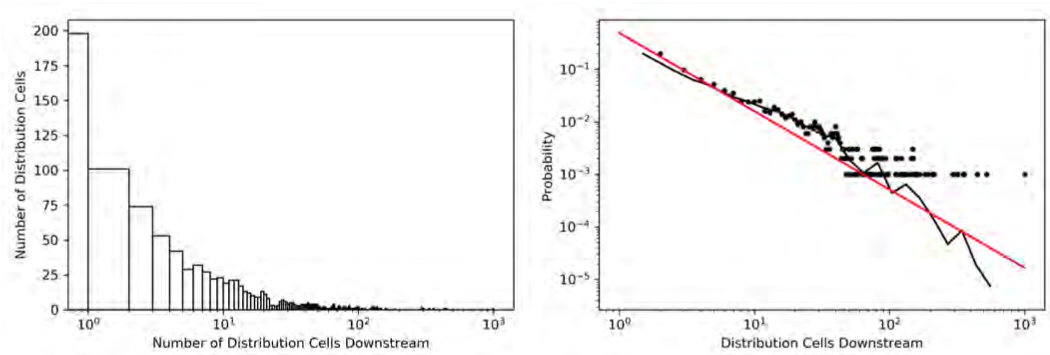


Figure 4.17: Results from GLS Model with 1000 Distribution Cells. *Left:* Histogram showing the Number of Distribution Cells Downstream for each Distribution Cell. *Right:* Probability Distribution for Number of Distribution Cells Downstream. The slope gradient of the power law curve is -1.49. (τ=1.49). Using a x_{min} value of 1, *p-value* = 0.17. The black dots are a probability distribution with linearly spaced bins and the black line is the probability distribution using logarithmically spaced bins from the powerlaw module.

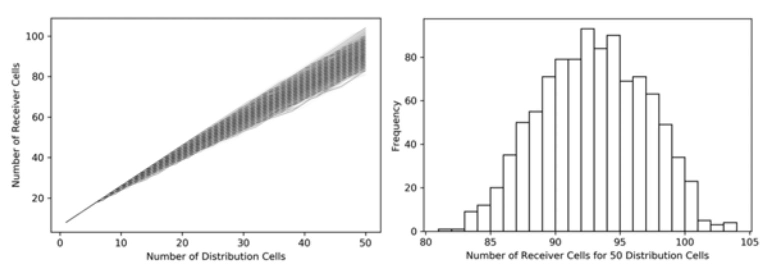


Figure 4.18: 1000 simulations of fifty Distribution Cells using the GLS model. Left - Plot showing the number of Receiver Cells to Distribution Cells. Right - Histogram showing the variation in number of Receiver Cells.

The results of the GS model are shown in Figures 4.19 and 4.20. The relationships are similar to the GLS model. The τ value is similar with a value of 1.47, Figure 4.20.

This section has so far given a comprehensive set of results for each of the models. There is a large amount of variation between the models, and to further understand these differences they have to be put into context of other experiments and real world systems. The value of τ, which is defined as $P(k) \sim k^{-\tau}$ is plotted for each model for different sizes, Figure 4.21.

As stated in Section 2.4.3, $\tau<1.5$ for sub-critical networks, $\tau=1.5$ for critical networks and $\tau>1.5$ for super-critical networks (De Los Rios, 2001). τ for the GLS and GS models produce networks which on average are critical. For both models, τ generally decreases for increased system size. Table 4.1 shows a comparison of τ between the GLS and GS models and other systems of interest. Table 4.2 shows a comparison of branches and termini with the electro-mechanic system in Jun and Hübler (2005).

Model/System	τ (Range)	Reference
GLS Model	1.54 (1.50-1.63)	-
GS Model	1.50 (1.46-1.56)	-
River System	1.43 (1.41-1.45)	(Cieplak et al., 1998)
Electro-mechanical System	1.14 (1.10-1.33)	(Jun and Hübler, 2005)
Internet	1.9 (1.8-2.0)	(Caldarelli et al., 2000)

Table 4.1: Value of τ for different models and systems. τ for the GLS and GS models are taken as the mean of values from Figure 4.21

.

Model/System	Branches gradient	Termini gradient	Reference
LDA Model	0.026	0.034	-
LSA Model	0.024	0.035	-
GLS Model	0.19	0.20	-
GS Model	0.17	0.18	-
Electro-mechanical System	0.22	0.22	(Jun and Hübler, 2005)

Table 4.2: Branches and Termini values taken from Figures 4.17 and 4.20

.

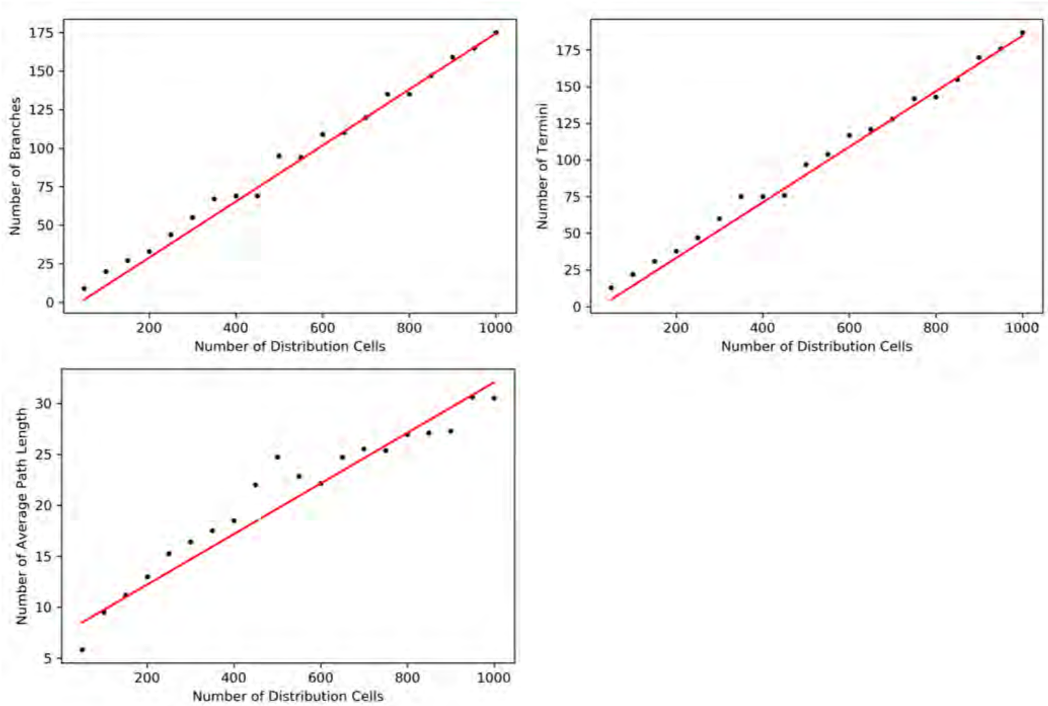


Figure 4.19: Results from GS Model. *Top Left:* Number of Branches plotted against the number of Distribution Cells. The slope gradient is 0.17. *Top Right:* Number of Termini plotted against the number of Distribution Cells. The slope gradient is 0.18. *Bottom Left:* Average Path Length plotted against the number of Distribution Cells. The slope gradient is 0.02.

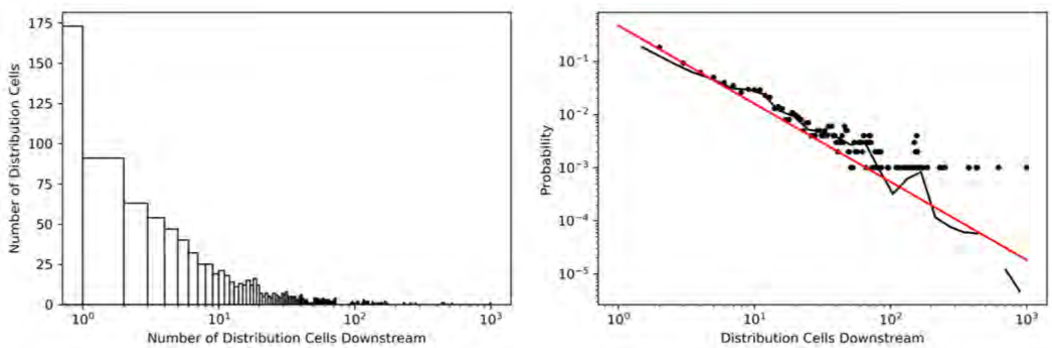


Figure 4.20: Results from GS Model with 1000 Distribution Cells. *Left:* Histogram showing the Number of Distribution Cells Downstream for each Distribution Cell. *Right:* Probability Distribution for Number of Distribution Cells Downstream. The slope gradient of the power law curve is -1.47 (τ = 1.47). Using a x_{min} value of 1, *p-value* = 0.15.

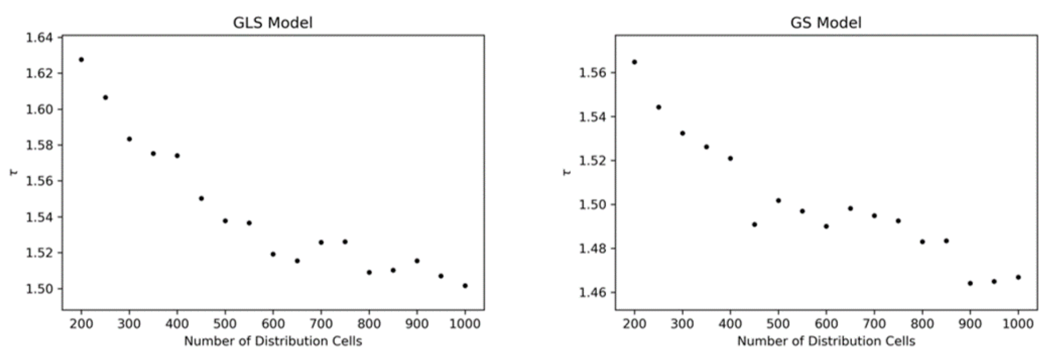


Figure 4.21: Value of τ for each of the different models found for a range of system size.

4.4.4 Research Question 2 Discussion

Research Question 2 asks whether the models created will have similar properties to other networks found in nature such as power law scaling.

To answer this question, additional analysis was carried out on each of the models and then compared with other experimental and empirical data. Two types of analysis were carried out - probability distribution of cells downstream and branches and termini analysis.

The LDA and LSA models did not exhibit any reasonably correlated power law distributions, Figures 4.13 and 4.15. This adds further evidence that the models do not reflect real-world systems and might only occur in a computer model. The analysis of the branches and termini found neither model follow a linear relationship as found in the literature and measure of the gradient was much different to the GLS and GS models and also the literature, table 4.2.

The GLS and GS models exhibit power law distributions, Figures 4.17 and 4.20. These are generally defined as sub-critical in terms of the τ exponent (De Los Rios, 2001). The results are comparable to those found in other sub-critical systems, see Table 4.2. The mean τ value for the GLS and GS models of 1.54 and 1.5 respectively are most similar to river systems (1.43). This reflects the optimality of the systems to a state of minimal dissipation. However, both the GLS and GS models exhibit a greater range (1.50-1.63) and (1.46-1.56) than river systems (1.41-1.45). This shows that the structure of the models can vary even if the optimality is reasonably constant. This might in part be due to the fact that GLS and GS models have a uniform cell size, meaning that all the network channels formed are the same size. River Systems on the other hand form channels of changing sizes dependent on the flow, which can be viewed as much more efficient, and may lead to τ being relatively constant. The Electro-mechanical system also exhibits a larger range but with lower values of τ (1.10-1.33) (Jun and Hübler, 2005). The reason for these lower values are thought to be due to finite scaling and dynamical effects such as friction stopping the system from reaching an optimal state (Jun and Hübler, 2005).

There are a few values of τ which are greater than 1.5 for the GLS and GS models. This is found for smaller systems, where less data is present. For larger systems the value seems to remain constant.

The analysis of Branches and Termini of the GLS and GS models found comparable results to the Electro-mechanical system albeit marginally lower values, Table 4.2. This again indicates commonalities between these models and other complex systems. Data on branches and termini is not available for other networks.

The analysis of the models show that the GLS and GS network properties have many commonalities with other trees which form space optimising networks.

4.5 Conclusion

The previous sections discussed the models in relation to answering the research questions. This section highlights the limits of the models and whether they can be validated.

The abstract simulation was built in order to test two alternative models of growth - local-orientated and global-orientated, which has previously been applied in Indigenous Irrigation Systems. The global models performed best (GLS and GS). The GLS model relied on randomness to produce asynchronous behaviour whereas the GS model used total information, but at a cost of additional work. These also had comparable properties to other dissipative tree systems such as river systems and experiments on electro-mechanical systems.

Real data on the network of Indigenous Irrigation Systems has not been collected. It is uncertain whether physical network data would find comparable results to the models due to the models only being applied in an abstract space-optimising framework. Real systems operate in a physical, social, ecological and geomorphological framework which may give wildly different network properties. The abstract model also only operates at one scale, the cell size is uniform across the model. Real-world systems operate at multiple scales across the framework. Despite all these differences, the model does highlight the usefulness of randomness in simulations of natural systems and a network of a spatially constrained tree which are not commonly built.

The model has not been tested against allometric scaling laws of biology (West et al., 1997). This is due to these models being reliant on area-preserving properties as the system branches, which is not present in this model. The model could however be compared with τ values of biological systems. However, it is unlikely that the behaviour of an Indigenous Irrigation System will be entirely like a biological system. Biological systems tend to maximise efficiency, given they have evolved over billions of years, and the models of them reflect this (West et al., 1997). An indigenous irrigation system on the other hand may have only evolved over at most, thousands of years. The channels of indigenous irrigation systems are not designed to have the most efficient cross-sectional area for the water to be carried, nor are they area preserving. But, there may still be aspects of IIS which are common with biological systems, such as space-filling, and downstream channels are likely to be narrower than upstream channels.

Chapter 5

Empirical Data and Analysis on Indigenous Irrigation Systems

The literature review found open questions regarding the evolution of the physical IIS network based on certain managerial arrangements. A generative agent based model was then built to look at the effect of using local or global information on the formation of an irrigation network, which seeks to optimise space.

This Chapter presents empirical data which has been collected on IIS in an effort in answering the following research question:

Do the networks formed by real life IIS share common characteristics to the simulated IIS, therefore validating the model?

Data has been collected using Remote Sensing, predominantly Google Earth.

5.1 A Remote Sensing Case Study of the Subak System in Bali, Indonesia

The Subak System, Bali, Indonesia was referenced a lot in the Literature Review, and the research questions regarding the influence of managerial arrangements on irrigation network structure were based on models of the Subak System (Lansing et al., 2009). It therefore seems logical to attempt to map at least part of the the Subak System in order to analyse the network structure.

A Satellite image of Bali is shown in Figure 5.1. The three volcanic cones are present in the north central part of the island with the river gorges flowing down to the sea from them. Figure 5.2 shows the relationship of the Bandjar

Figure 5.1: Satellite image of Bali. The area enclosed by the white line is where Subak Culture is predominately located. Taken from Google Earth (November 2019).

(settlements), Subak (agriculture) and river gorges. The gorges are dark green colour, likely reflecting dense natural vegetation, the Bandjar are brown and the Subak are light green. This is similar to the explanation provided by Geertz (1980), with the geomorphology being a controlling factor in the development of the system.

The area enclosed in *Box a*, Figure 5.2 is shown in Figure 5.3. This highlights the space-filling nature of the Subak System, with the Subak agriculture using the majority of the land in this area. The bottom image (Figure 5.3) is an interpretation of the irrigation network, using the image alone. Each of the Subak terraces is delineated by a line, as each terrace is more or less horizontal, this approximately maps the contours of the Subak. The river channels (blue) flow in incised channels on the valley floor. The main irrigation channels (yellow) flow down the spurlines, with minor irrigation channels (orange) branching off down slope to the river channels.

Figure 5.2: Increased scale image of Subak System. Relationship observed between the Bandjar (settlement), subak (agriculture) and river gorges in the North West part of the Subak System. *Box a* is the location of Figure 5.3. Taken from Google Earth (November 2019).

Figure 5.3: *Top* - Example of the Space-Filling Nature of Subak Agriculture. Located at *Box a* on Figure 5.2. *Bottom* - Interpretation of the Irrigation Network. Key - *Blue*: Natural River Channel, *Yellow*: Main Irrigation Channel, *Orange*: Minor Irrigation Channel and *Red*: Bandjar Settlement. Upstream is towards the Northwest and downstream is towards the Southeast.

5.2 A Remote Sensing Case Study of the Qanat System in Joupar, South-east Iran

The literature review also focused on an example of the Qanat system in the Middle East including a schematic drawing of a Qanat irrigation network, Figure 2.1 (Bonine, 1996). This irrigation system has many environmental differences to the Subak System, making it an interesting system to compare with.

Figure 5.4 shows the location of the village Joupar, in South-East Iran. This has been chosen due to the availability of high-resolution which distinctly show the Qanat tunnels flowing to the village. It is located 2 km north of a mountain range with a peak (Kuh-e Jupar) of 4150m and 100km west of the Dasht-e Lut Desert, which has recorded one of the highest recorded land surface temperature on Earth reaching 70.7°c (Mildrexler et al., 2011). Figure 5.5 shows the village, with multiple Qanat tunnels flowing from the mountain range (the source) into a single joining to form a single tunnel at the village. The village buildings are predominately located upstream of the irrigated fields.

An interpretation of the irrigation network mapped from the image is shown in Figure 5.6. There is uncertainty when making this interpretation, some of the irrigation channels can be mapped with high certainty, whereas other are obscured by vegetation. Walls and shadows created by walls can also be misinterpreted as irrigation channels. Images over multiple years were observed to increase the certainty of the interpretation.

5.3 Results

The networks shown in Figures 5.3 and 5.6 have been annotated by hand. In order to analyse them, they have to spatially quantified. This is done by first pinpointing each node spatially in terms of longitude and latitude (Figure 5.7) and then adding information on the edges. These were then imported into Python and plotted using the Networkx module (Figure 5.8).

To compare the networks to the synthetic networks created in Chapter 4 the same analysis is undertaken. Figure 5.9 shows the power law and histogram for the Qanat Network. This seems to show a power law relationship, although the data set is small, and the power law seems to only fit over one order of magnitude. The τ value is 1.66 which is greater than 1.5 so the network is categorised as super critical.

Figure 5.10 shows the analysis for the Subak Network. The power law scaling does not seem to fit at all. It is therefore fair to disregard this as showing powerlaw scaling behaviour.

Figure 5.4: Location of Joupar in Iran (Box A), see Figure 5.5. Image taken from Google Earth, November 2019.

Figure 5.5: An example of a Qanat system, Joupar, South-east Iran. A: The hilly area to the South is the source area of the system. The village and irrigated fields are to the North. B: A number of spoil heaps outlining the path of the Qanat. C: The irrigated fields. Channel paths can be inferred by the linear features present. Images taken from Google Earth, May 2015.

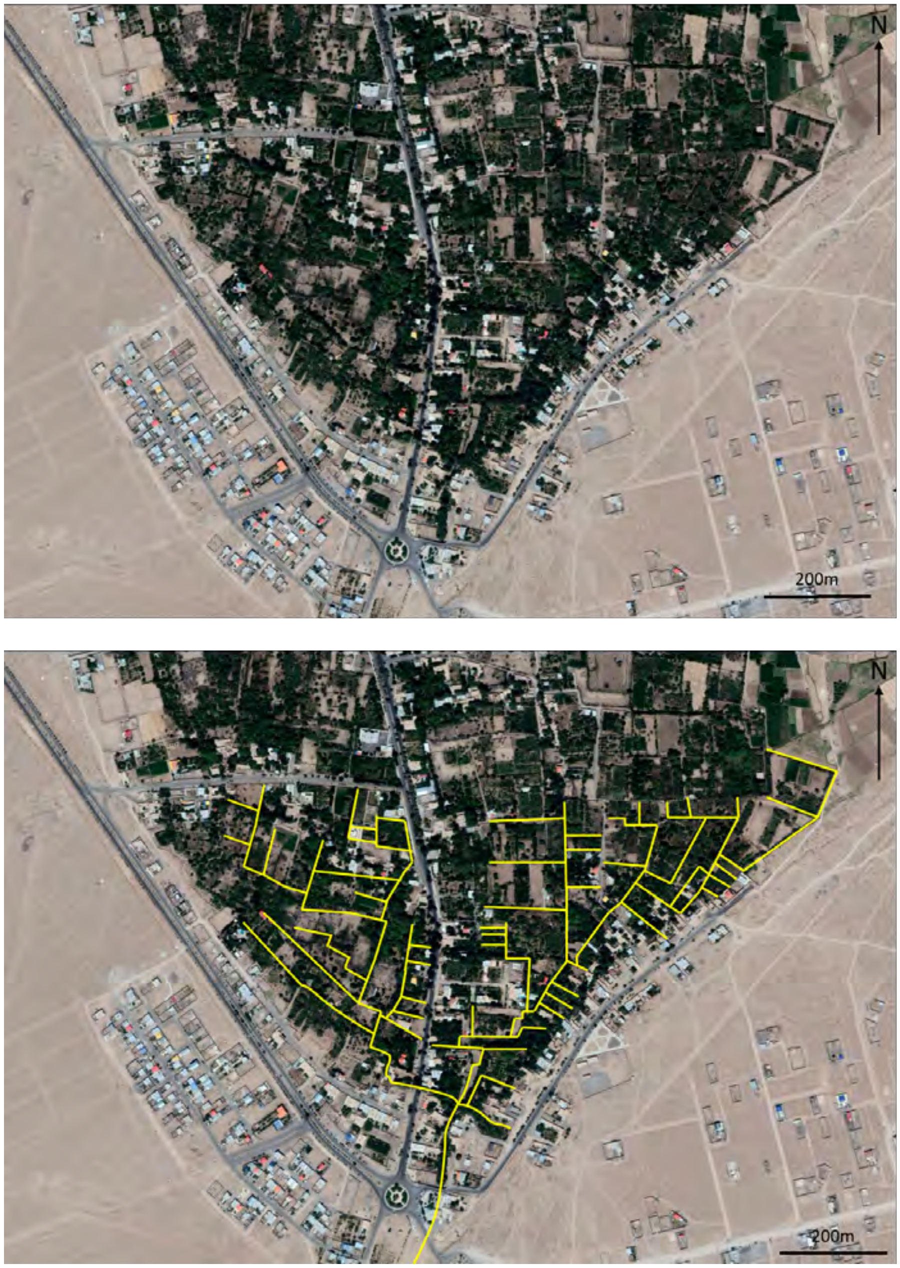


Figure 5.6: *Top* - Increased scale image of the Qanat system, Joupar. *Bottom* - *Yellow*: Interpretation of the Irrigation Network from Remote Sensing. Image taken from Google Earth, November 2019.

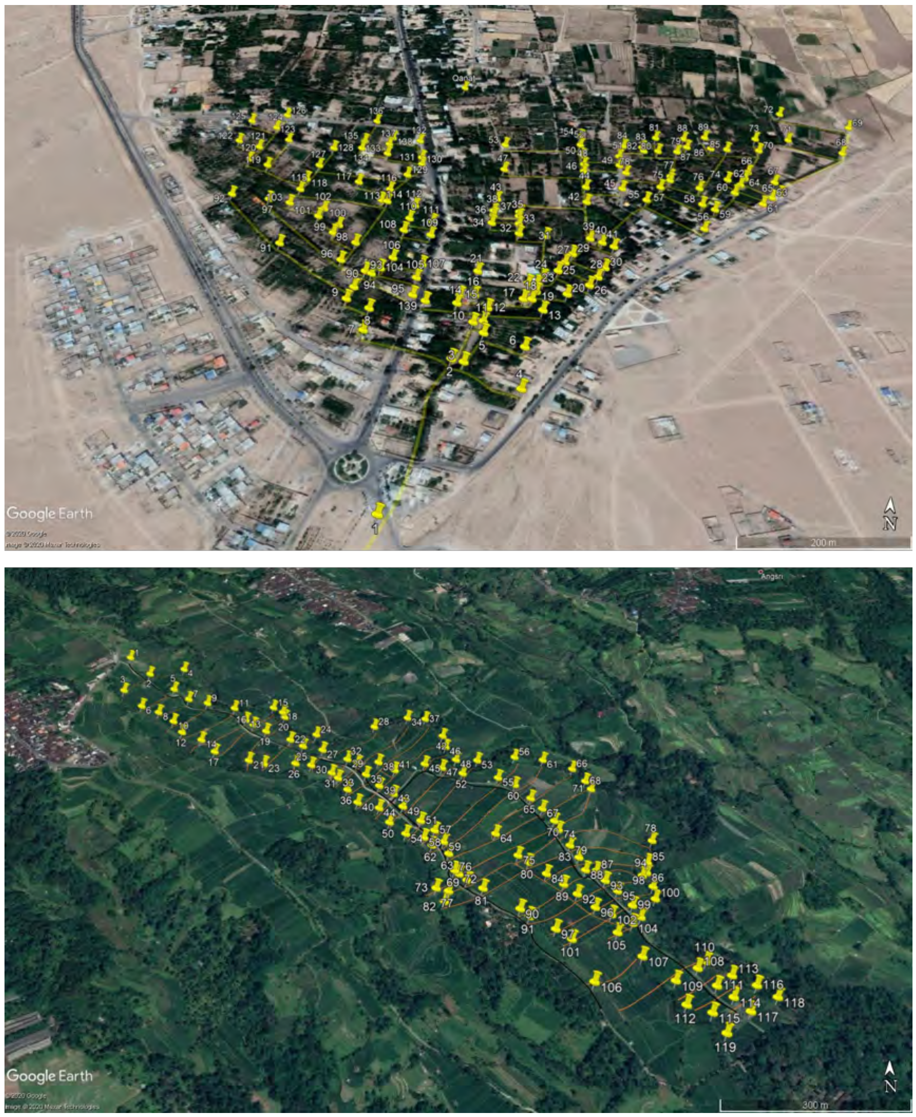
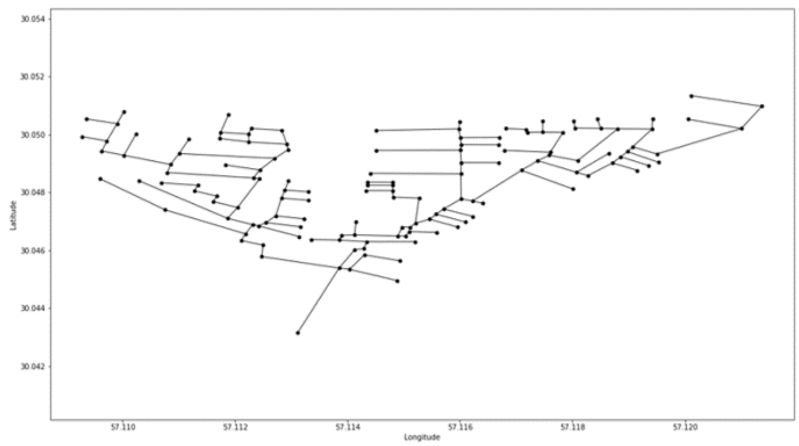


Figure 5.7: *Top* - Node Locations for the Qanat Irrigation Network. *Bottom* - Node Locations for the Subak Irrigation Network. Image taken from Google Earth, November 2019.



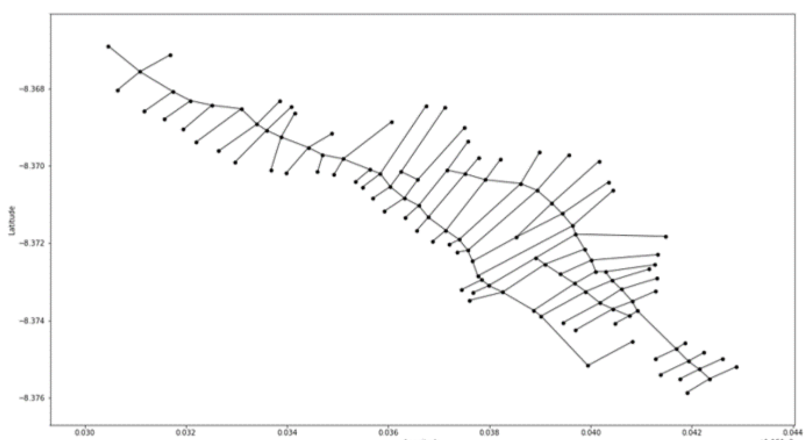


Figure 5.8: Quantified Networks using the Networkx module. *Top* - The Qanat Irrigation Network. *Bottom* - The Subak Irrigation Network.

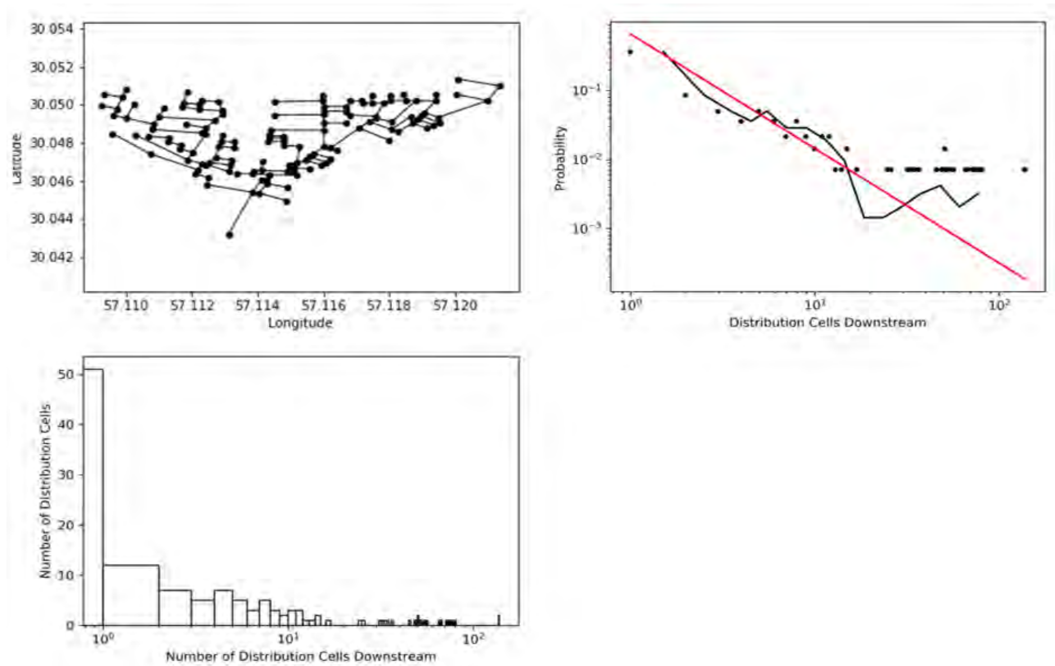


Figure 5.9: Scaling Analysis of the Qanat Irrigation Network. *Top Left:* Spatial Network. *Top Right:* Powerlaw Distribution for nodes downstream of each node. The slope gradient of the power law curve is -1.66 (τ = 1.66) with a *p-value* = 0.11.

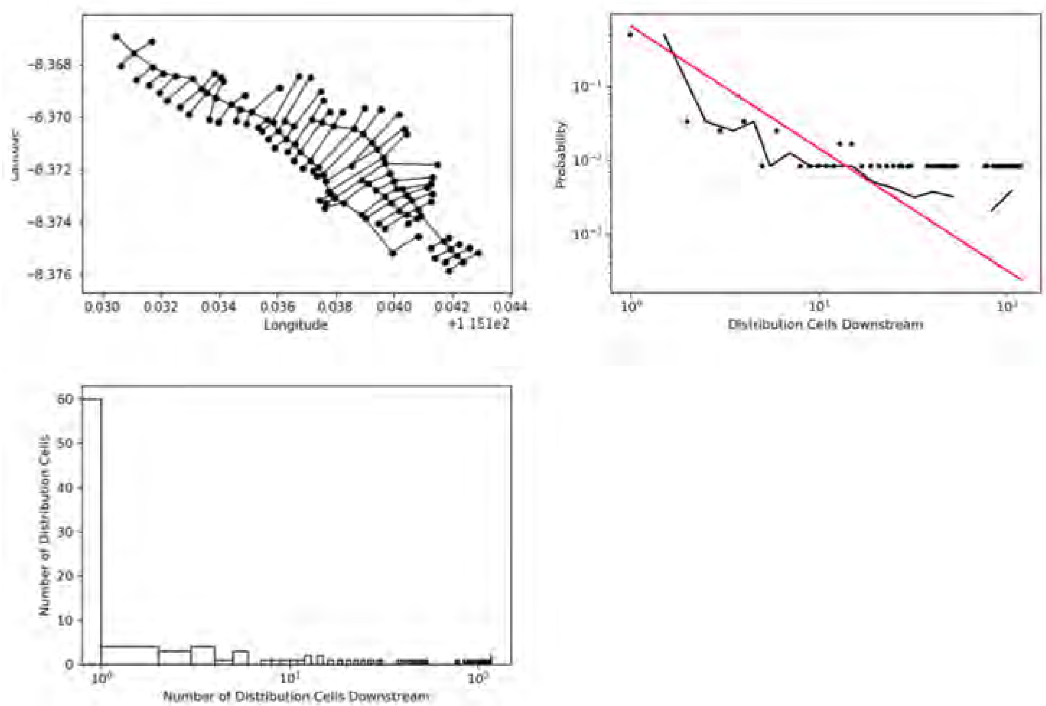


Figure 5.10: Scaling Analysis of the Subak Irrigation Network. *Top Left:* Spatial Network. *Top Right:* Powerlaw Distribution for nodes downstream of each node.

5.4 Discussion

Firstly, two small datasets have been analysed in this section. Drawing any sort of conclusion from these would be unscientific, however I still think it is possible to discuss these initial findings and offer insights, comparisons and possible interpretations for differences between the datasets.

An observation based comparison between the Qanat and Subak networks shows common features but also many differences. Firstly both networks start from an initial node, they have a single origin. The Qanat network branches out into many nodes, whereas the Subak network consists of main 'backbone' nodes which feed minor nodes. The Subak network seems to follow a form more similar to the ordered synthetic networks as shown in Figure 4.9. The Subak network also does not follow a power law scaling relationship which is to be expected from such as a network as shown in Figures 4.13 and 4.15, which have main 'backbone' nodes. On the other hand the Qanat network shows a power law scaling relationship, although it is weak and perhaps only present over one order of magnitude. This is more similar to the global stochastic and global section models as shown in Figures 4.17 and 4.20.

As discussed previously there are many factors which lead to differences in the networks formed by Indigenous Irrigation Systems including management, geomorphology, geology and climate. The two systems could not be located in more diverse climatic regimes with the Subak being located in the water rich sub-tropics and the Qanat in an arid desert. The geomorphology is also vastly different with the Qanat being located on a shallow planar slope and the Subak on a steep mountain slope with linear gorges formed from intense flooding activity. These factors will definitely have an influence on the structure of the irrigation network formed. The Qanat for example is dug underground from the source to the irrigation network in order to minimise evaporation. The linear gorges eroded into the mountain slope of the Subak are likely to lead in turn to linear irrigation networks which follow the topography. This could be one strong underlying reason for the 'backbone' nodes formed in the Subak.

The space optimising properties were explored in greater detail for the cellular simulation networks in Chapter 4. As the network was simulated on an array this allowed for the detailed analysis of space optimising properties and an analytical solutions to be obtained. The network information obtained in this chapter is in the form of node and edge data. This has not been converted to cellular data making it more difficult to measure properties of space optimisation. This could be done in the future to allow for further information on optimality of the real world networks and validate the simulated models.

Finally, the influence of managerial arrangements can be discussed. The synthetic simulations generated in Chapter 4, rely solely on these. Parallels between them and the empirical data have already been drawn. However some of these parallels might be due to coincidence and actually be due to other influences such as geomorphology. From a sociological perspective, previous research has found that the Qanat system is related to a more individualistic culture, and the Subak to a more community based culture (Geertz, 1972). This may also have an influence on the network formed. To argue this, it is worth questioning why the Qanat which is on a relatively planar slope, does not form a network with strong 'backbone' nodes, but instead one that seems to grow based on local rules. An explanation is that due to the limited water supply, the network has grown relatively slowly over time, dependant on increases in supply such as adding further feeder channels to the network. The supply of the water, being a necessity to those in the system leads to individuals or families controlling supply to parts of the network allowing it to grow over time based on agreements. The network is therefore not pre-planned, and extended based on local conditions in a similar fashion to the random or selection networks in Chapter 4. The Subak has strong geomorphological controls on its formation, but it can also be argued that due to the plentiful supply of water and community based culture, a irrigation network can be pre-planned forming one which is more ordered. This would explain why the Qanat Network follows power law scaling to a greater degree and the Subak does not follow it whatsoever.

5.5 Conclusion

The discussion of the reasons for the differences between the Subak and Qanat irrigation systems are preliminary, and the findings are by no means conclusive. However, I think the differences that have been found so far offer appetite to gather further data and analysis to see if interpretation provided in the discussion holds for further networks. Ground-truthing may also be required to validate the network structure. This will offer insight into the accuracy of remote sensing methods.

Chapter 6

Summary and Further Work

6.1 Summary

This thesis has aimed to explore the network properties of Indigenous Irrigation Systems to see if the managerial arrangements of the system lead to different network characteristics. To undertake this aim, simulation models of space optimising networks were constructed using different managerial rules. Real world IIS network data was then collected remotely from two different systems and compared to one another and the properties of the simulated networks.

The simulated networks focused on the effect of local, global, deterministic and stochastic rules on the emergent properties of the network structure. These particular rules were chosen given discussions in the literature over whether such systems rely on local initiatives or centralised planners (Lansing et al., 2009). The space optimality properties and comparisons to other similar networks in nature were used to suggest the likelihood of each network type occurring in reality. Networks which relied on local deterministic rules tended to result in low space optimising properties. Planned networks and globally selective or globally stochastic networks gave much greater space optimising properties. But these networks also differed in terms of scaling. Planned networks tend to not follow power law scaling, whereas globally selective and stochastic networks do, and also in a comparable manner to systems in nature such as rivers.

The data collected from the real world IIS networks was analysed using the same scaling tools as the simulated networks. This found that the Qanat system followed a power law (more closely), and the Subak system did not. This could be an indication that the Qanat system follows more globally selective/stochastic managerial rules whereas the Subak system follows more planned managerial rules. Although given the large ranged of factors influencing network formation this result is not conclusive. There is however appetite to further explore these factors.

6.2 Further Work

The literature review highlighted other research gaps which have not been answered in this thesis. These are as follows:

- There is no recognised model for how IIS tend to grow, do they follow a biological type sigmoidal curve or sociological expansion? Is this related to the managerial arrangements?
- Previous studies have looked at the development of inequality in simplified game-theoretic models of IIS, but not for larger network based simulations.
- No quantitative models have been built to look at the effects of globalisation on IIS.

A lot of the data that is required to address these research gaps is difficult to attain. For example growth of IIS would require age dating. Whilst there has been some age dating collected from the Subak system (Lansing et al., 2009), this is not conclusive and is unlikely to give detailed information about the type of growth experienced by IIS. Information on the development of inequality is also difficult to attain as it partly requires field data collection and the effects of globalisation and outside forces on these systems are likely to have influences on their current form.

Bibliography

Adamatzky, A. (2010). *Game of Life Cellular Automata*. Springer-Verlag, London.

Adamatzky, A. (2018). Generative complexity of Gray–Scott model. *Communications in Nonlinear Science and Numerical Simulation*, 56:457–466.

Adams, W. M., Potkanski, T., and Sutton, J. E. G. (1994). Indigenous irrigation Tanzania in Sonjo ,. *The Geographical Journal*, 160(1):17–32.

Albert, R., Jeong, H., and Barabási, A.-L. (1999). Diameter of the World-Wide Web. *Nature*, 401(6749):130–131.

Allesina, S. and Tang, S. (2011). Stability Criteria for Complex Ecosystems. *Nature*, 483(7388):12.

Alstott, J., Bullmore, E., and Plenz, D. (2014). Powerlaw: A python package for analysis of heavy-tailed distributions. *PLoS ONE*, 9(1).

Anderies, J. M. and Janssen, M. a. (2011). The fragility of robust social-ecological systems. *Global Environmental Change*, 21:1153–1156.

Araral, E. (2013). What makes socio-ecological systems robust? An institutional analysis of the 2000 year-old Ifugao society. *Human Ecology*, 41(6):859–870.

Axelrod, R. and Dion, D. (1988). The Further Evolution of Cooperation. 242(4884):1385–1390.

Axelrod, R. and Hamilton, W. (1981). The evolution of cooperation. *Nature*, 292(5821):291–292.

Bak, P., Chen, K., and Creutz, M. (1989). Self-organized criticality in the 'Game of Life'. *Nature*, 342:780–782.

Bak, P. and Paczuski, M. (1995). Complexity, contingency, and criticality. *Proceedings of the National Academy of Sciences of the United States of America*, 92(15):6689–6696.

Bak, P., Tang, C., and Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/f noise. *Physical Review Letters*, 59(4):381–384.

Baker, M. (2005). The Kuhls of Kangra: Community-Managed Irrigation in the Western Himalaya. 2005:151–154.

Bar-Yam, Y. (1998). *Dynamics of Complex Systems (Studies in Nonlinearity) Variational Principles and the Numerical Solution of Scattering Problems*, volume 12.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(October):509–512.

Barthélemy, M. (2011). Spatial networks. *Physics Reports*, 499(1-3):1–101.

Bartlett, S. and Bullock, S. (2015). Emergence of Competition between Different Dissipative Structures for the Same Free Energy Source. *Ecal 2015*, (1999).

Batty, M. (2007). *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals*. The MIT Press.

Batty, M. (2009). *Cities as Complex Systems: Scaling, Interaction, Networks, Dynamics and Urban Morphologies*. Springer, Berlin.

Batty, M., Xie, Y., and Sun, Z. (1999). Modeling urban dynamics through GIS-based cellular automata. *Computers, Environment and Urban Systems*, 23(3):205–233.

Beer, R. D. (2018). On the Origin of Gliders. *The 2019 Conference on Artificial Life*, (30):67–74.

Benenson, E. H. and Itzhak (2012). The Schelling Model of Ethnic Residential Dynamics : Beyond the Integrated - Segregated Dichotomy of Patterns. *Journal of Artificial Societies and Social Simulation*, 15(2012):1–23.

Bettencourt, L. M. a., Lobo, J., Helbing, D., Kühnert, C., and West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences of the United States of America*, 104(17):7301–7306.

Bonabeau, E., Dorigo, M., and Theraulaz, G. (1999). *Swarm Intelligence: From Nature to Artificial Systems*. Oxford University Press, New York.

Bonine, M. E. (1996). Qanats and rural societies: Sustainable agriculture and irrigation cultures in contemporary Iran. In *Canals and Communities: Small-Scale Irrigation Systems*, pages 183–209. University of Arizona Press.

Borgerhoff Mulder, M., Bowles, S., Hertz, T., Bell, A., Beise, J., Clark, G., Fazzio, I., Gurven, M., Hill, K., Hooper, P. L., Irons, W., Kaplan, H., Leonetti, D., Low, B., Marlowe, F., McElreath, R., Naidu, S., Nolin, D., Piraino, P., Quinlan, R., Schniter, E., Sear, R., Shenk, M., Smith, E. A., von Rueden, C., and Wiessner, P. (2009). Intergenerational wealth transmission and the dynamics of inequality in small-scale societies. *Science (New York, N.Y.)*, 326(5953):682–8.

Bowles, S., Smith, E. A., and Mulder, M. B. (2010). The emergence and persistence of inequality in premodern societies: Introduction to the special section. *Current Anthropology*, 51(1):7–17.

Brede, M. and Boschetti, F. (2009). Commons and anticommons in a simple renewable resource harvest model. *Ecological Complexity*, 6(1):56–63.

Caldarelli, G., Marchetti, R., and Pietronero, L. (2000). The fractal properties of Internet. *Europhysics Letters*, 52(4):386–391.

Carlson, J. M. and Doyle, J. (2002). Complexity and robustness. *Proceedings of the National Academy of Sciences of the United States of America*, 99 Suppl 1:2538–2545.

Carpenter, S. R. and Bennett, E. M. (2011). Reconsideration of the planetary boundary for phosphorus. *Environmental Research Letters*, 6:014009.

Chen, X., Lupi, F., An, L., Sheely, R., Viña, A., and Liu, J. (2012). Agent-based modeling of the effects of social norms on enrollment in payments for ecosystem services. *Ecological Modelling*, 229:16–24.

Chenciner, A. (2012). Poincare and the Three-Body Problem. *Poincare, 1912-2012, Seminaire Poincare XVI (2012)*, pages 45–133.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.

Cieplak, M., Giacometti, A., Maritan, A., Rinaldo, A., Rodriguez-Iturbe, I., and Banavar, J. R. (1998). Models of fractal river basins. *Journal of Statistical Physics*, 91(1-2):1–15.

Cifdaloz, O., Regmi, A., Anderies, J. M., and Rodriguez, A. a. (2010). Robustness, vulnerability, and adaptive capacity in small-scale socialecological systems: The pumpa irrigation system in Nepal. *Ecology and Society*, 15(3).

Clarke, K. C. and Gaydos, L. J. (1998). Loose-coupling a cellular automaton model and GIS: Long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science*, 12(7):699–714.

Clauset, A., Shalizi, C. R., and Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703.

Cole, S. and Browne, M. (2015). Tourism and Water Inequity in Bali: A Social-Ecological Systems Analysis. *Human Ecology*, 43(3):439–450.

Collard, P., Mesmoudi, S., Ghetiu, T., and Polack, F. (2013). Emergence of frontiers in networked schelling segregationist models. *Complex Systems*, 22(1):35–59.

Corbit, J. D. and Garbary, D. J. (1995). Fractal Dimension as a Quantitative Measure of Complexity in Plant Development. *Proceddings: Biological Sciences*, 262(1363):1–6.

Coulthard, T. J. (2001). Landscape evolution models: a software review. *Hydrological Processes*, 173(15):165–173.

Coulthard, T. J., Macklin, M. G., and Kirkby, M. J. (2002). A cellular model of Holocene upland river basin and alluvial fan evolution. *Earth Surface Processes and Landforms*, 27(3):269–288.

Coulthard, T. J., Neal, J. C., Bates, P. D., Ramirez, J., de Almeida, G. A., and Hancock, G. R. (2013). Integrating the LISFLOOD-FP 2D hydrodynamic model with the CAESAR model: Implications for modelling landscape evolution. *Earth Surface Processes and Landforms*, 38(15):1897–1906.

Cox, M. and Ross, J. M. (2011). Robustness and vulnerability of community irrigation systems: The case of the Taos valley acequias. *Journal of Environmental Economics and Management*, 61(3):254–266.

Crook, D. (2009). Hydrology of the combination irrigation system in the Wadi Faynan, Jordan. *Journal of Archaeological Science*, 36(10):2427–2436.

Crook, D. and Elvin, M. (2013). Bureaucratic control of irrigation and labour in late-imperial China: the uses of administrative cartography in the Miju catchment, Yunnan. *Water History*, 5(3):287–305.

Crook, D., Tripathi, S., and Jones, R. T. (2013). The sustainability of Suranga irrigation in South Karnataka and northern Kerala, India. *1st World Irrigation Forum, Conference Proceedings*, (May):10.

Dakos, V., Carpenter, S. R., Nes, E. H. V., Scheffer, M., and Dakos, V. (2015). Resilience indicators : prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370:20130263.

Davies, M. (2008). The irrigation system of the Pokot, northwest Kenya. *Azania: Archaeological Research in Africa*, 43(1):50–76.

Davies, M. (2009). Wittfogel's dilemma: Heterarchy and ethnographic approaches to irrigation management in Eastern Africa and Mesopotamia. *World Archaeology*, 41(1):16–35.

de Bont, C., Komakech, H. C., and Veldwisch, G. J. (2019). Neither modern nor traditional: Farmer-led irrigation development in Kilimanjaro Region, Tanzania. *World Development*, 116:15–27.

De Los Rios, P. (2001). Power law size distribution of supercritical random trees. *Europhysics Letters*, 56(6):898–903.

Dearing, J. a., Bullock, S., Costanza, R., Dawson, T. P., Edwards, M. E., Poppy, G. M., and Smith, G. M. (2012). Navigating the perfect storm: research strategies for socialecological systems in a rapidly evolving world. *Environmental management*, 49(4):767–75.

Dearing, J. A., Wang, R., Zhang, K., Dyke, J. G., Haberl, H., Hossain, M. S., Langdon, P. G., Lenton, T. M., Raworth, K., Brown, S., Carstensen, J., Cole, M. J., Cornell, S. E., Dawson, T. P., Doncaster, C. P., Eigenbrod, F., Flörke, M., Jeffers, E., Mackay, A. W., Nykvist, B., and Poppy, G. M. (2014). Safe and just operating spaces for regional social-ecological systems. *Global Environmental Change*, 28:227–238.

Dimitrov, P. and Zucker, S. W. (2006). A constant production hypothesis guides leaf venation patterning. *Proceedings of the National Academy of Sciences*, 103(24):9363–9368.

Dyke, J. G. and Weaver, I. S. (2013). The emergence of environmental homeostasis in complex ecosystems. *PLoS computational biology*, 9(5):e1003050.

Egitto, A. C. (2013). *REMOTE SENSING ASSESSMENT OF KAREZ IRRIGATION SYSTEMS AND ARCHAEOLOGICAL RESOURCES IN MAYWAND DISTRICT, KANDAHAR PROVINCE, AFGHANISTAN*. PhD thesis, University of Kansas.

Ehrlich, P. . R. and Levin, S. A. (2005). The evolution of norms. *PLoS biology*, 3(6):e194.

Elvin, M. (2002). *The impact of clearance and irrigation on the environment in the Lake Erhai catchment from the ninth to the nineteenth century*. Number 23.

English, P. (1968). The Origin and Spread of Qanats in the Old World. *Proceedings of the American Philosophical Society*, 112(3):170–181.

English, P. (1998). Qanats and lifeworlds in Iranian plateau villages. *Yale F&ES Bulletin*, pages 187–205.

Epstein, J. (2008). Why model? *Journal of Artificial Societies and Social Simulation*, 11(4):6.

Epstein, J. M. (2006). *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton University Press, Princeton, New Jersey, USA.

Epstein, J. M. and Axtell, R. (1996). *Growing Artificial Societies*. The Brookings Instituiton, Washington, D.C.

Faiz, M. E. and Ruf, T. (2010). An Introduction to the Khettara in Morocco: Two Contrasting Cases Mohammed. In Schneier-Madanes, G. and Courel, M. F., editors, *Water and Sustainability in Arid Regions*, number April. Springer.

Ferber, J. (1999). *Multi-Agent System: An Introduction to Distributed Artificial Intelligence*. Addison Wesley Longman, Harlow.

Fleischman, F. D., Boenning, K., Garcia-Lopez, G. A., Mincey, S., Schmitt-Harsh, M., Daedlow, K., Lopez, M. C., Basurto, X., Fischer, B., and Ostrom, E. (2010). Disturbance, response, and persistence in self-organized forested communities: Analysis of robustness and resilience in five communities in Southern Indiana. *Ecology and Society*, 15(4).

Gao, J., Barzel, B., and Barabási, A.-L. (2016). Universal resilience patterns in complex networks. *Nature*, 530(7590):307–312.

Gardner, M. R. and Ashby, W. R. (1970). Connectance of large dynamic (cybernetic) systems: critical values for stability. *Nature*, 228:784.

Gautam, D., Thapa, B. R., and Prajapati, R. N. (2018). Indigenous water management system in Nepal: cultural dimensions of water distribution, cascaded reuse and harvesting in Bhaktapur City. *Environment, Development and Sustainability*, 20(4):1889–1900.

Geekiyanage, N. and Pushpakumara, D. (2013). Ecology of ancient tank cascade systems in island Sri Lanka. *Journal of Marine and Island Cultures*, 2(2):93–101.

Geertz, C. (1972). The wet and the dry: Traditional irrigation in Bali and Morrocco. *Human Ecology*, 1(1):23–39.

Geertz, C. (1980). Organisation of the Balinese subak. In *Irrigation and Agricultural Development in Asia: perspectives from the social sciences*, pages 70–90.

Gell-Mann, M. (1995). What is complexity. *Complexity*, 1(1):16–19.

Gilbert, N. (2008). Agent-Based Models. *SAGE Publications*, 153(153):98.

Gillingham, M. E. (1999). Gaining Access to Water : Formal and Working Rules of Indigenous Irrigation Management on Mount Kilimanjaro , Tanzania. *Natural resources journal*, 39(1995):419–441.

Gilpin, M. E. and Ayala, F. J. (1973). Global Models of Growth and Competition. *Proceedings of the National Academy of Sciences*, 70(12):3590–3593.

Gosak, M., Markovič, R., Dolenšek, J., Slak Rupnik, M., Marhl, M., Stožer, A., and Perc, M. (2018). Network science of biological systems at different scales: A review. *Physics of Life Reviews*, 24:118–135.

Gray, P. and Scott, S. K. (1985). Sustained oscillations and other exotic patterns of behavior in isothermal reactions. *Journal of Physical Chemistry*, 89(1):22–32.

Groenfeldt, D. J. (2004). Building on Tradition : Indigenous Irrigation and Sustainable Development in Asia. *Water, Cultural Diversity and International Solidarity, Symposium Proceedings edited by Corinne Wacker. Zurich Working Papers in Social Anthropology*, 14:1–26.

Grunwald, P. and Vitanyi, P. (2004). Shannon Information and Kolmogorov Complexity. *arXiv preprint cs/0410002*, pages 1–54.

Hardin, G. (1968). The Tragedy of the Commons. *Science*, 162(June):1243–1248.

Harris, T. E. (1964). *The Theory of Branching Processes*. Rand Corporation.

Hartnett, A. T., Schertzer, E., Levin, S. A., and Couzin, I. D. (2016). Heterogeneous Preference and Local Nonlinearity in Consensus Decision Making. *Physical Review Letters*, 116(3):1–5.

Heller, M. A., Eisenberg, R. S., Hardin, G., Demsetz, H., Merges, R., Nelson, R., Lehman, D., Lempert, R., and Nisbett, R. (1998). Can patents deter innovation? The anticommons in biomedical research. *Science (New York, N.Y.)*, 280(5364):698–701.

Holland, J. (1992). Complex Adaptive Systems. *Daedalus*, 121:17–30.

Holling, C. S. (1973). Resilience and Stability of Ecological Systems. *Annual Review of Ecology and Systematics*, 4(1973):1–23.

Holling, C. S. (2001). Understanding the Complexity of Economic, Ecological, and Social Systems. *Ecosystems*, 4:390–405.

Hunt, D. and Savage, V. M. (2016). Asymmetries arising from the space-filling nature of vascular networks. *Physical Review E*, 93(6):1–16.

Hunt, R. C., Hunt, E., Ahmed, G. M., Bennett, J. W., Cleek, R. K., Coy, P. E. B., Glick, T. F., Lewis, R. E., MacLachlan, B. B., Mitchell, W. P., Partridge, W. L., Price, B. J., Roder, W., Steensberg, A., Wade, R., and Wellmann, I. (1976). Canal Irrigation and Local Social Organization [and Comments and Reply]. *Current Anthropology*, 17(3):389–411.

IPCC (2001). Climate change 2001: impacts, adaptation, and vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Published for the Intergovernmental Panel on Climate Change. page 1042.

Jäger, G. and Rogers, J. (2012). Formal language theory: Refining the Chomsky hierarchy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1598):1956–1970.

Janssen, M. a. (2007). Coordination in irrigation systems: An analysis of the Lansing-Kremer model of Bali. *Agricultural Systems*, 93:170–190.

Jun, J. K. and Hübler, A. H. (2005). Formation and structure of ramified charge transportation networks in an electromechanical system. *Proceedings of the National Academy of Sciences of the United States of America*, 102(3):536–540.

Jungclaus, J. H., Keenlyside, N., Botzet, M., Haak, H., Luo, J. J., Latif, M., Marotzke, J., Mikolajewicz, U., and Roeckner, E. (2006). Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM. *Journal of Climate*, 19(16):3952–3972.

Karlebach, G. and Shamir, R. (2008). Modelling and analysis of gene regulatory networks. *Nature Reviews Molecular Cell Biology*, 9:770–780.

Kelly, W. W. (1983). Concepts in the Anthropological Study of Irrigation. *American Anthropologist*, 85(4):880–886.

Kihila, J. M. (2018). Indigenous coping and adaptation strategies to climate change of local communities in Tanzania: a review. *Climate and Development*, 10(5):406–416.

Kinzig, A. P., Ehrlich, P. R., Alston, L. J., Arrow, K., Barrett, S., Buchman, T. G., Daily, G. C., Levin, B., Levin, S., Oppenheimer, M., Ostrom, E., and Saari, D. (2013). Social Norms and Global Environmental Challenges: The Complex Interaction of Behaviors, Values, and Policy. *Bioscience*, 63(3):164–175.

Kyungrock, P. and Kumar, P. (2008). Emergence of self-similar tree network organization. *Complexity*, 13(4):30–37.

Langton, C. G. (1984). Self-Reproduction in Cellular Automata. *Physica D: Nonlinear Phenomena*, 10(1-2):135–144.

Lansing, J., Kremer, J., and Smuts, B. (1998). System-dependent Selection, Ecological Feedback and the Emergence of Functional Structure in Ecosystems. *Journal of theoretical biology*, 192(3):377–391.

Lansing, J. S. (1987). Balinese "Water Temples" and the Management of Irrigation. *American Anthropologist*, 89(2):326–341.

Lansing, J. S. (1991). *Priests and programmers: technologies of power in the engineered landscape of Bali.* Princeton University Press, Princeton, New Jersey, USA.

Lansing, J. S. (1996). Simulation Modeling of Balinese Irrigation. In *Canals and Communities: Small-Scale Irrigation Systems*, pages 139–156. University of Arizona Press.

Lansing, J. S. (2003). Complex Adaptive Systems. *Annual Review of Anthropology*, 32(1):183–204.

Lansing, J. S., Cheong, S. A., Chew, L. Y., Cox, M. P., Ringo Ho, M.-H., and Arthawiguna, W. A. (2014). Regime Shifts in Balinese Subaks. *Current Anthropology*, 55(2):232–239.

Lansing, J. S., Cox, M. P., Downey, S. S., Janssen, M. A., and Schoenfelder, J. (2009). A robust budding model of Balinese water temple networks. *World Archaeology*, 41(1):112–133.

Lansing, J. S. and Kremer, J. N. (1993). Emergent Properties of Balinese Water Temple Networks : Coadaptation on a Rugged Fitness Landscape. *American Anthropologist*, 95(1):97–114.

Lansing, J. S. and Miller, J. H. (2005). Cooperation, Games, and Ecological Feedback: Some Insights from Bali. *Current Anthropology*, 46(2):328–334.

Lansing, J. S., Thurner, S., Chung, N. N., Coudurier-Curveur, A., Karakaş, Ç., Fesenmyer, K. A., and Chew, L. Y. (2017). Adaptive self-organization of Bali's ancient rice terraces. *Proceedings of the National Academy of Sciences*, 114(25):6504–6509.

Laoubi, K. and Yamao, M. (2009). A typology of irrigated farms as a tool for sustainable agricultural development in irrigation schemes: The case of the East Mitidja scheme, Algeria. *International Journal of Social Economics*, 36(8):813–831.

Lee, S. W., Feugier, F. G., and Morishita, Y. (2014). Canalization-based vein formation in a growing leaf. *Journal of Theoretical Biology*, 353:104–120.

Lev-Yadun, S., Gopher, A., and Abbo, S. (2000). The cradle of agriculture. *Science*, 288(5471):1602–1603.

Levin, S., Xepapadeas, T., Crépin, A.-S., Norberg, J., de Zeeuw, A., Folke, C., Hughes, T., Arrow, K., Barrett, S., Daily, G., Ehrlich, P., Kautsky, N., Mäler, K.-G., Polasky, S., Troell, M., Vincent, J. R., and Walker, B. (2012). Social-ecological systems as complex adaptive systems: modeling and policy implications. *Environment and Development Economics*, 18(02):111–132.

Levin, S. A. (1998). Ecosystems and the Biosphere as Complex Adaptive Systems. *Ecosystems*, 1(5):431–436.

Levin, S. A. (2014). Public goods in relation to competition, cooperation, and spite. *Proceedings of the National Academy of Sciences*, 111.

Li, X., Jing, Y., and Liao, S. (2017). *The 1223 new periodic orbits of planar three-body problem with unequal mass and zero angular momentum*, volume 1.

Lightfoot, D. R. (1996). Syrian qanat Romani: history, ecology, abandonment. *Journal of Arid Environments*, 33(3):321–336.

Lindenmayer, A. (1968). Mathematical models for cellular interactions in development. II. Simple and branching filaments with two-sided inputs. *Journal of theoretical biology*, 18(3):300–315.

Lorenz, E. N. (1963). Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, 20:130–141.

Lovelock, J. E. and Margulis, L. (1974). Atmospheric homeostasis by and for the biosphere: the gaia hypothesis. *Tellus*, 26(1-2):2–10.

Mabry, J. B. and Cleveland, D. A. (1996). The relevance of indigenous irrigation: A comparative analysis of sustainability. In *Canals and Communities: Small-Scale Irrigation Systems*, pages 227–260. University of Arizona Press, Tucson AZ.

MacArthur, R. (1955). Fluctuations of Animal Populations and a Measure of Community Stability. *Ecology*, 36(3):533.

Mandelbrot, B. B. (1983). *Fractals and the Geometry of Nature*, volume 173. WH Freeman, New York.

Manukyan, L., Montandon, S. A., Fofonjka, A., Smirnov, S., and Milinkovitch, M. C. (2017). A living mesoscopic cellular automaton made of skin scales. *Nature*, 544:173.

May, R. M. (1972). Will a large complex system be stable? *Nature*, 238:413–414.

May, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature*, 261(5560):459–467.

May, R. M. (2004). Uses and Abuses of Mathematics in Biology. *Science*, 303(February):790–793.

May, R. M., Levin, S. a., and Sugihara, G. (2008). Complex systems: ecology for bankers. *Nature*, 451(February):893–895.

McCann, K. S. (2000). The diversity-stability debate. *Nature*, 405(6783):228–33.

Mehari, A., Schultz, B., and Depeweg, H. (2005). Where indigenous water management practices overcome failures of structures: The Wadi Laba spate irrigation system in Eritrea. *Irrigation and Drainage*, 54(1):1–14.

Mildrexler, D. J., Zhao, M., and Running, S. W. (2011). Satellite finds highest land skin temperatures on Earth. *Bulletin of the American Meteorological Society*, 92(7):855–860.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science*, 298(5594):824–827.

Mitchell, M., Crutchfield, J. P., and Hraber, P. (1993). Dynamics, Computation, and the "Edge of Chaos": A Re-Examination.

Monks, T., Currie, C. S., Onggo, B. S., Robinson, S., Kunc, M., and Taylor, S. J. (2019). Strengthening the reporting of empirical simulation studies: Introducing the STRESS guidelines. *Journal of Simulation*, 13(1):55–67.

Motesharrei, S., Rivas, J., and Kalnay, E. (2014). Human and nature dynamics (HANDY): Modeling inequality and use of resources in the collapse or sustainability of societies. *Ecological Economics*, 101:90–102.

Murphy, J. T. (2012). Exploring complexity with the Hohokam Water Management Simulation: A middle way for archaeological modeling. *Ecological Modelling*, 241:15–29.

Newman, M. (2010). *Networks: An Introduction*. Oxford University Press.

Nield, J. and Baas, A. (2008). Investigating parabolic and nebkha dune formation using a cellular automaton modelling approach. *Earth Surface Processes and Landforms*, 33(5):724–740.

Nowak, M. A. and May, R. M. (1992a). Evolutionary Games and Spatial Chaos. *Nature*, 359:826–829.

Nowak, M. A. and May, R. M. (1992b). Evolutionary games and spatial chaos.

Nowak, M. a., Tarnita, C. E., and Wilson, E. O. (2010). The evolution of eusociality. *Nature*, 466(7310):1057–1062.

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.

Oxford University Press (2015). Oxford English Dictionary.

Papadopoulos, L., Blinder, P., Ronellenfitsch, H., Klimm, F., Katifori, E., Kleinfeld, D., and Bassett, D. S. (2018). *Comparing two classes of biological distribution systems using network analysis*, volume 14.

Prigogine, I. and Nicolis, G. (1985). Self-Organisation in Nonequilibrium Systems: Towards A Dynamics of Complexity. In Hazewinkel, M., Jurkovich, R., and Paelinck, J., editors, *Bifurcation Analysis*. Springer, Dordrecht.

Prusinkiewicz, P. (2004). Modeling plant growth and development. *Current Opinion in Plant Biology*, 7(1):79–83.

Purvis, B., Mao, Y., and Robinsion, D. (2019). Entropy and its application to urban systems. *Entropy*, 21(1).

Rajasekaran, B., Martin, R. A., and Warren, D. M. (1993). A Framework for incorporating indigenous knowledge systems into agricultural extension organisation for sustainable agricultural development in India. *Proceedings of the Annual Conference of the Association for International Agricultural and Extension Education, 1993.*

Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. *Computer Graphics*, 21(4):25–34.

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S., Lambin, E., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R. W., Fabry, V. J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., and Foley, J. (2009). Planetary boundaries: Exploring the safe operating space for humanity. *Ecology and Society*, 14(2).

Roman, S., Bullock, S., and Brede, M. (2017). Coupled Societies are More Robust Against Collapse: A Hypothetical Look at Easter Island. *Ecological Economics*, 132:264–278.

Samuelson, P. A. (1954). The Pure Theory of Public Expenditure. *The Review of Economics and Statistics*, 36(4):387–389.

Sayama, H. (2015). *Introduction to the Modelling and Analysis of Complex Systems*. Open SUNY Textbooks, New York.

Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., and Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, 413(6856):591–596.

Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., van de Koppel, J., van de Leemput, I. a., Levin, S. a., van Nes, E. H., Pascual, M., and Vandermeer, J. (2012). Anticipating Critical Transitions. *Science*, 338(October):344–348.

Scheffer, M. and Westley, F. (2007). The evolutionary basis of rigidity: locks in cells, minds, and society. *Ecology and Society*, 12(2):36.

Schelling, T. C. (1969). Models of segregation. *The Amercian Economic Review*, 59(2):488–493.

Schelling, T. C. (1971). Dynamic Models of Segregation. *Journal of Mathematical Sociology*, 1:143–186.

Solé, R. V. and Montoya, J. M. (2001). Complexity and fragility in ecological networks. *Proceedings of the Royal Society B: Biological Sciences*, 268(1480):2039–2045.

Steffen, W., Persson, Å., Deutsch, L., Zalasiewicz, J., Williams, M., Richardson, K., Crumley, C., Crutzen, P., Folke, C., Gordon, L., Molina, M., Ramanathan, V., Rockström, J., Scheffer, M., Schellnhuber, H. J., and Svedin, U. (2011). The anthropocene: From global change to planetary stewardship. *Ambio*, 40:739–761.

Steffen, W., Richardson, K., Rockström, J., Cornell, S., Fetzer, I., Bennett, E., Biggs, R., Carpenter, S. R., de Wit, C. a., Folke, C., Mace, G., Persson, L. M., Veerabhadran, R., Reyers, B., and Sörlin, S. (2015). Planetary Boundaries: Guiding human development on a changing planet. *Science*, 347(6223).

Strogatz, S. (2015). *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. CRC Press, Boca Raton, 2nd edition.

Šuvakov, M. and Dmitrašinović, V. (2013). Three classes of newtonian three-body planar periodic orbits. *Physical Review Letters*, 110(11).

Trawick, P. B. (2001). Successfully governing the commons: Principles of social organization in an Andean irrigation system. *Human Ecology*, 29(1):1–25.

Turing, a. M. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 237(641):37–72.

Van De Wiel, M. J. and Coulthard, T. J. (2010). Self-organized criticality in river basins: Challenging sedimentary records of environmental change. *Geology*, 38(1):87–90.

Verburg, P. H., Dearing, J. a., Dyke, J. G., Leeuw, S. V. D., Seitzinger, S., Steffen, W., and Syvitski, J. (2015). Methods and approaches to modelling the Anthropocene. *Global Environmental Change*.

Wang, R., Dearing, J. A., Doncaster, C. P., Yang, X., Zhang, E., Langdon, P. G., Yang, H., Dong, X., Hu, Z., Xu, M., Zhao, Y., and Shen, J. (2019). Network parameters quantify loss of assemblage structure in human-impacted lake ecosystems. *Global Change Biology*, 25(11):3871–3882.

Wang, R., Dearing, J. a., Langdon, P. G., Zhang, E., Yang, X., Dakos, V., and Scheffer, M. (2012). Flickering gives early warning signals of a critical transition to a eutrophic lake state. *Nature*, 492(7429):419–422.

Watkins, a., Noble, J., Foster, R. J., Harmsen, B. J., and Doncaster, C. P. (2014). A spatially explicit agent-based model of the interactions between jaguar populations and their habitats. *Ecological Modelling*.

Watson, A. J. and Lovelock, J. E. (1983). Biological homeostasis of the global environment: the parable of Daisyworld. *Tellus B*, 35B:284–289.

Watson, E. E., Adams, W. M., Mutiso, S. K., The, S., Journal, G., Mar, N., Watson, E. E., Adams, W. M., and Mutiso, S. K. (1998). Indigenous Irrigation, Agriculture and Development, Marakwet, Kenya. *The Geographical Journal*, 164(1):67–84.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–2.

West, G. B., Brown, J. H., and Enquist, B. J. (1997). A General Model for the Origin of Allometric Scaling Laws in Biology. *Science*, 276(5309):122–126.

Wijermans, N. and Schlüter, M. (2014). Modelling cooperation in Bali irrigation. In *Advances in Computational Social Science and Social Simulation*.

Wilensky, U. and Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, 8(1):3–19.

Wittfogel, K. A. (1957). *Oriental despotism: a comparative study of total power*. Yale University Press, New Haven, Connecticut, USA.

Wolf-Gladrow, D. A. (2000). Lattice-Gas Cellular Automata and Lattice Boltzmann Models - An Introduction. *PoLAR*, page 308.

Wolfram, S. (1984). Computation theory of cellular automata. *Communications in Mathematical Physics*, 96(1):15–57.

Yang, X.-S. and Yang, Y. Z. L. (2007). Cellular Automata Networks. In Adamatzky, A., Bull, L., Costello, B. L., Stepney, S., and Teuscher, C., editors, *Proceedings of Unconventional Computing 2007*, pages 280–302. Luniver Press.

Young, O. R., Berkhout, F., Gallopin, G. C., Janssen, M. a., Ostrom, E., and van der Leeuw, S. (2006). The globalization of socio-ecological systems: An agenda for scientific research. *Global Environmental Change*, 16(3):304–316.

Yu, B. and Li, B. (2006). Fractal-like tree networks reducing the thermal conductivity. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 73(6):1–8.

Yu, D. J., Qubbaj, M. R., Muneepeerakul, R., Anderies, J. M., and Aggarwal, R. M. (2015). Effect of infrastructure design on commons dilemmas in social-ecological system dynamics. *Proceedings of the National Academy of Sciences*, 112(43):201410688.

Appendices

Appendix A

Supplementary Information for Chapter 4

This Appendix provides Supplementary Information on the methods, discussion and analysis for Chapter 4 to allow for greater clarity and repeatability.

A.1 Supplementary Methods

The code for the simulation model was written in Python 3.7.3 using Spyder 3.3.3 Python Development Environment (`https://www.spyder-ide.org/`). It was written from scratch without extending previous code, but required many revisions to refine it to the current version. A number of Python modules are used to increase the speed of the simulation (*numpy*), plot data (*matplotlib*), store model outputs (*pandas*) and analyse the data (*sklearn, scipy, powerlaw*).

All versions of the code use a 2 dimensional array to map the spatial properties of the network. Earlier versions mapped directly onto an array, whilst later versions store information in class instances as the model became more complicated, which made the simulation more like an Agent Based Model. This is also more efficient than earlier versions which iterated over the whole array each time-step which can be very computationally expensive particularly if the array is large, however instead the simulation iterates over the list of agents. So the computation time relates to the number of agents in the list and not the size of the area used to map spatial properties. This also means the array can be much larger than the number of agents so the boundary conditions do not have to be considered.

The parameters for the model are given in Table A.1. Definitions and explanations for r and the model type are given in the main text.

Parameter	Definition	Range in Model
A_s	Size of spatial array	100-200
D_0	Location of first Distribution Cell	Usually $A_s/2$
$N^M{}_{ij}$	Size of Moore Neighbourhood of each Distribution Cell	8
$N^V{}_{ij}$	Size of Von Neumann Neighbourhood of each Distribution Cell	4
D_{max}	Maximum number of Distribution Cells	100-2000
M	Model Type	LDA, LSA, GLS, GS

Table A.1: Parameters for the Model in Chapter 4

.

To aid in communicating the processes in the model, pseudo-code is used. This is written in a series of steps which are implemented in order when the simulation is executed. The Pseudo-Code for the model is as follows:

1. Parameters initialised: A_s, D_0, $N^M{}_{ij}$, $N^V{}_{ij}$, D_{max}, M. The lists to store distribution cells and receiver cells initialised).

2. Classes to hold information are initialised. Two types: Distribution Cells and Receiver Cells. Each Class Instance of Distribution Cell has the following class objects: *(position, path length, upstream distribution cell, downstream distribution cell(s), total distribution cells downstream)*. Each Class Instance of Receiver Cell has the following class objects: *(position, associated distribution cell, id)*.

3. DateFrame initialised to store model output.

4. (If the code is running multiple instances of the model with different parameters, for example the number of agents, then an additional loop is added for multiple simulations.)

5. The initial Distribution Cell (Location based on Parameters). A class instance is added for this Distribution Cell.

6. Function implemented to add Receiver Cells for this Distribution Cell. Checks the cells in the Distribution Cell neighbourhood, adds Receiver Cells in empty cells. Class instances for each Receiver Cell are added.

7. Loop implemented to add Distribution Cells until the maximum number is reached. A new *AgentsExpand* list is initialised. Distribution Cells are added to the list as they are added to the network. As the following functions are implemented some Cells are removed from it.

8. A variation of the function is used depending on the model (LDA, LSA, GLS, GS). All the different models check to see if adding distribution cells

to the Von Neumann neighbourhood of the current Distribution Cells in the network lead to an increase in the number of receiver cells for the network, and will only add a distribution cell if it leads to an increase. The variation between the models depends on the Distribution Cell selected to be added.

9. For the LDA and LSA models, the Distribution Cell with the lowest *Path Length* is selected. If more than one Distribution Cell in the Von Neumann neighbourhood increase the amount of receiver cells, then the LDA model selects based on predefined order priority, whereas the LSA model selects at random.

10. The GLS model does not rely on the lowest *path length*, it will select any Distribution Cell and any direction for that Distribution Dell at random which increases the number of receiver cells.

11. The GS model iterates through all distribution cells in the *AgentsExpand* list adding the Distribution Cell and direction which increases the number of Receiver Cells by the most. If more than one Distribution Cell increase the number by the same maximum amount, then one is selected at random.

12. As each of these models iterates through the Distribution Cells, when one is added it is also added to the *AgentsExpand* list, but if when tested a Distribution Cell cannot increase the number of Receiver Cells at all, then it is removed from the *AgentsExpand* list. This is particularly important for the GS model as it computationally more expensive. When a new Distribution Cell is added, a new class instance is added, and new class instances for the associated receiver cells too.

13. When the number of Distribution Cells reaches the the maximum number as set out in the parameters, the loop is exited, data collected into the DataFrame, graphs plotted and the simulation is ended.

The full code can be found at `https://sites.google.com/site/alexjohnstokes/home/appendix-a`. The code contains many outputs, such as graphs, databases and animations.

A.2 Model Testing

All the model functions have been thoroughly tested, so the number of bugs if any are assumed to be minimal. Tests involve running parts of the code separately with a small number of Distribution Cells in which the output is known. Once the output is as expected it can be scaled up to larger numbers of Distribution Cells.

A.3 Derivation of Generalisation in Equation 4.6

Equation 4.6 is a generalisation for the number of Receiver Cells (N_{RC}) in a maximum space-optimising network, written in terms of the number of Distribution Cells (N_{DC}) and the Distribution Cell neighbourhood size (r).

Given all space optimising networks ($r \geq 1$) will grow linearly if the network is grown with maximum efficiency, the linear equation for each network can be found by the first two steps of each network.

For *r=1*,
$\mathrm{N_{DC}}$ = 1, 2, 3, 4, 5
$\mathrm{N_{RC}}$ = 8, 10, 12, 14, 16

For *r=2*,
$\mathrm{N_{DC}}$ = 1, 2, 3, 4, 5
$\mathrm{N_{RC}}$ = 24, 28, 32, 36, 40

For *r=3*,
$\mathrm{N_{DC}}$ = 1, 2, 3, 4, 5
$\mathrm{N_{RC}}$ = 48, 54, 60, 66, 72

Working leading to General Solution

Equation 4.3 gives an analytical solution for *r=1*
$\mathrm{N_{RC}}$=2$\mathrm{N_{DC}}$+6
Equation 4.1 gives the size of the Moore neighbourhood in terms of r as $(2r+1)^2$.
Substituting this in leads to:
$\mathrm{N_{RC}}$=2$\mathrm{N_{DC}}$+$(2r+1)^2$
For *r=1*,
$\mathrm{N_{DC}}$ = 1, 2, 3, 4, 5
$\mathrm{N_{RC}}$ = 11, 13, 15, 17, 19

This result is greater than that found in the model. But it still finds linear growth. This is due to the fact that in this model as DC are added, RC cells are removed and that each DC added will not add the full Moore neighbourhood.

By adding an extra term, the correct solution can be found:
$\mathrm{N_{RC}}=2\mathrm{N_{DC}}+(2r+1)^2-(2r+1)$

However to find a general solution (for any value of r) another extra r has to be added:
$\mathrm{N_{RC}}=2r\mathrm{N_{DC}}+(2r+1)^2-(2r+1)$.

To test the general solution equation. The model has been modified to allow for experiments with $r \geq 1$. By testing only the maximum spacing optimising

GS model, this can be compared to the analytical general solution. Table A.2 shows that the analytical and simulated solutions produce exactly the same results. For reference the final output of each of the simulated solutions is shown in Figure A.1.

r	Number of Receiver Cells (Analytical Solution)	Number of Receiver Cells (Simulated Solution)
1	206	206
2	420	420
3	642	642
4	872	872
5	1110	1110
6	1356	1356
7	1610	1610
8	1872	1872
9	2142	2142
10	2420	2420
11	2706	2706
12	3000	3000

Table A.2: Comparison between analytical and simulated solutions for the GS model. The value of r is changed for each simulation but the number of distribution cells is kept constant at 100.

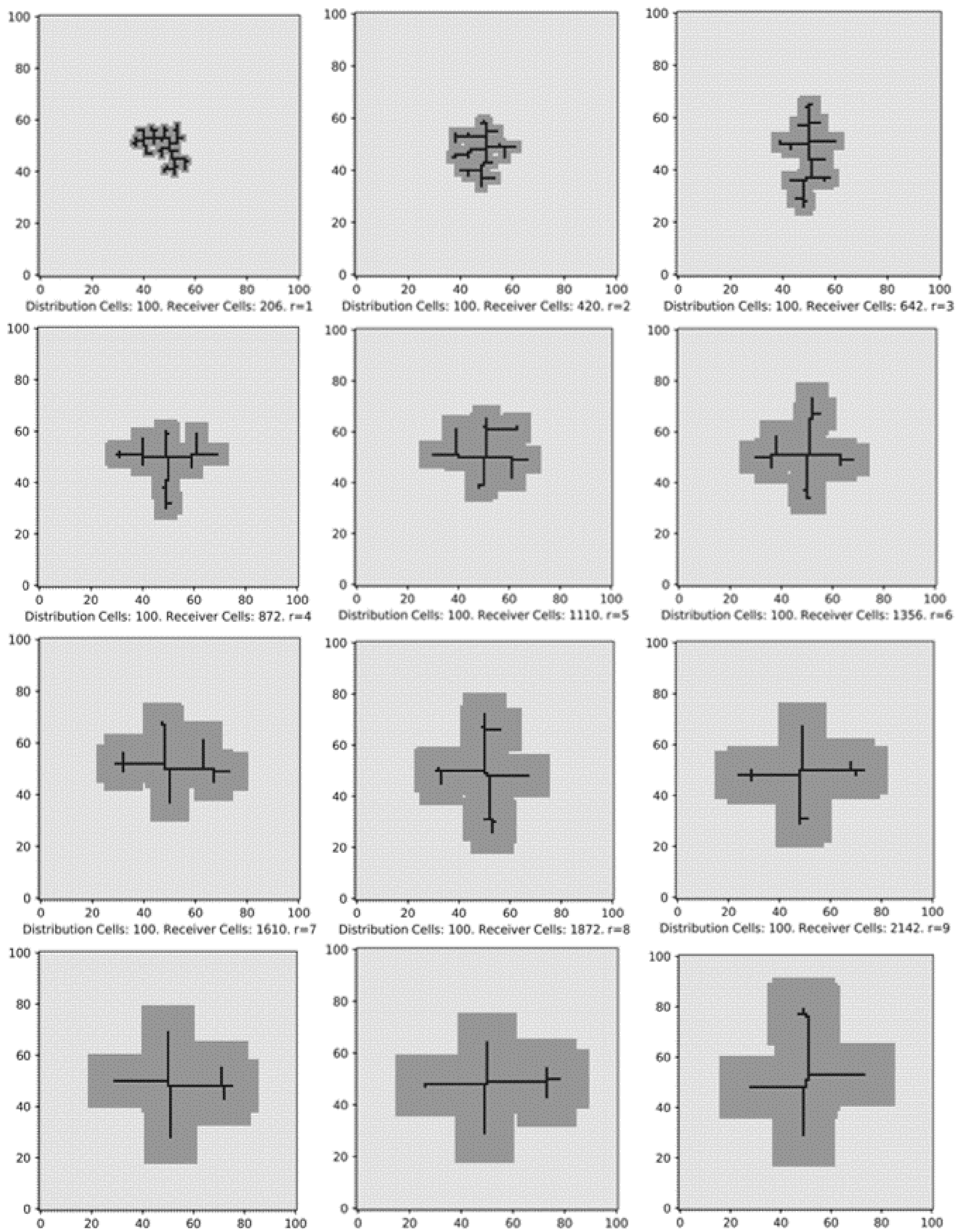


Figure A.1: Output of each of the simulated solutions from Table A.2

.