University of Southampton

Faculty of Environmental and Life Sciences

Geography and Environmental Science

Modelling Spatiotemporal Dynamics of Population, Flooding and Road Travel for Enhanced Risk Assessment: A case study of York, UK

Volume 1 of 1

by

Kate Emily New

ORCID ID 0000-0002-2065-3282

Thesis for the degree of Doctor of Philosophy

August 2020

University of Southampton

<u>Abstract</u>

Faculty of Environmental and Life Sciences

Geography and Environmental Science

Thesis for the degree of <u>Doctor of Philosophy</u>

Modelling Spatiotemporal Dynamics of Population, Flooding and Road Travel for Enhanced Risk Assessment: A case study of York, UK

by

Kate Emily New

The number of people travelling around an urban area varies greatly throughout the course of a day, week or year. As a result, the daily timing of a flood event, relative to these mobility patterns, is a critical factor in how a flood affects the local population. However, population mobility is seldom considered in flood risk assessments, meaning an important aspect of risk is missed. This thesis aims to investigate how daily variation in the travelling population, pluvial flood onset timing, and pluvial flood magnitude, interact to cause spatial and temporal variation in disruption to journey times and destination disruption across the urban area of York, UK. The population has been considered in subgroups, aligned to age and economic activity characteristics as these have been shown to have distinctive temporal characteristics. In each chapter, commuters and/or primary school children are the population groups selected as exemplars for analysis.

This thesis comprises three analysis chapters. The first analysis chapter's goal was to develop a framework for combining spatiotemporal population flow data with GIS network analysis, using journeys to primary schools as an example. The second analysis chapter examined how estimated flood-related commuter travel disruption was affected by different approaches to spatiotemporal population modelling. The third analysis chapter's goal was to assess if the time of flood onset is more important than flood magnitude for disrupting commuter and school travel in York. Overall, this thesis has provided evidence that for sudden onset flooding scenarios, timing of flood onset is a greater determinant of hazard-related travel disruption than flood magnitude. The semi-dynamic framework could be applied to other urban areas to model the effects of pluvial flood events, as it works at the scale of a local authority, and to other types of hazard which disrupt travel like landslides, earthquakes and fallen trees.

Table of Contents

Table of Tables

Table of Tables

Table of Figures

Table of Equations

Table of Equations

Research Thesis: Declaration of Authorship

Print name:	Kate Emily New

Title of thesis:	Modelling Spatiotemporal Dynamics of Population, Flooding and Road Travel for Enhanced Risk Assessment: A case study of York, UK

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

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. Parts of this work have been published as:

Chapter 6 has been presented as a short paper at the GISRUK 2019 conference under the title 'Integrating spatiotemporal dynamics for modelling disruption to road travel in flood events'
Chapter 7 has been presented as a paper at the GISRUK 2020 conference under the title 'Spatiotemporal modelling of flood-related impacts on daily population movement'
These papers are located in Appendix B

Signature:		Date:	17.08.2020

Acknowledgements

I would like to firstly thank my three supervisors, Jim Wright, Alan Smith and Sally Brown for all their support whilst I have researched and completed my thesis. From the initial research ideas through to finishing the final document, you have all brought so much enthusiasm and encouragement which without I would not have been able to complete my PhD.

I would also like to thank the Economic and Social Research Council and the South Coast DTP for funding my research (Grant number ES/J500161/1) and providing training and networking opportunities over the last few years, particularly the knowledge transfer partnership I undertook.

I have found my research to be a technical challenge which has broadened the scope of my quantitative skills, and I would like to thank those who have assisted in this journey. Firstly, to Jeremiah Nieves who wrote the R code for the spatially weighted Monte Carlo analysis used in Chapters 5, ,6 and 7. You were able to create a code which implemented my idea for analysis but did not have the R coding skills to enact, and I am very grateful for the time you spent completing this. I also want to thank Jeremiah for the encouragement and technical chats we had as I learnt Python over the past year or so. I would also like to thank the university staff involved with the Population 247 projects at Southampton, Sam Cockings, for allowing me to use the SurfaceBuilder247 model, David Martin, Andy Harfoot, Andy Campbell-Sutton and Alan Smith, for teaching me how to implement it and assisting the revision of my background layers under the new methodology. Spatiotemporal population modelling has been a challenge but with your help I have learnt so much and developed an interest in an aspect of Geography I had not explored before. Finally, I would like to thank the staff at Ambiental Risk Analytics, particularly Paul Drury and Evan Shallcross, for hosting me for a knowledge transfer internship and allowing me use of their flood model and data. I thoroughly enjoyed my three months stay in Brighton and loved working on projects with you all.

There are other staff and postgrads at the university I would also like to thank. Dianna Smith for her encouragement and kindness through the PhD process and in non-research projects like demonstrating and the PhD bootcamp. To all the staff who have had me as a demonstrator on modules and to Andrew Phillips for helping me to successfully apply for Associate Fellowship of the Higher Education Academy, thank you for allowing me the opportunities to develop my teaching skills and for the fun that has been had along the way. One of the best things about doing a PhD at Southampton has been working in the PGR hub with all my wonderful fellow PhD students. I have been very lucky to have worked alongside such a fabulous group of people who have provided a social network throughout my almost four years in the department. In particular I would like to thank Laura Crossley, Angharad Butler-Rees, Rebecca Collins, Josie Foster and Roseanna Mayfield for the tea breaks and chats which have kept me going through the challenging moments of research. I would also like to thank Jo Wilkin who inspired me to get coding and introduced me to the world of Python. Finally, thank you to Sarah Westbury for cheering me on over the last few years, your support has been invaluable.

My final thanks are to my family. To my Mum and Dad who have supported me throughout my whole life and provided me the opportunities to complete my university degrees, I couldn't have got here without you. And finally, to my fiancé Jon who encouraged me to take the PhD offer and has cheered me on throughout the whole process. Finishing my PhD during a national lockdown due to a global pandemic was never how I imagined finishing my thesis. It has been a hard few months and your positivity and ability to cheer me up has meant I have at last finished my thesis. I hope that submitting my PhD marks the start of things improving and we are able to have our rescheduled wedding soon.

“Every Village, Town or City is not merely a Place in Space, but a Drama in Time”

- Patrick Geddes

Definitions and Abbreviations

AADF - Annual Average Daily Flow

AAL – Average Annual Loss

ABMs – Agent Based Models

API – Application Programming Interfaces

DEFRA – Department for Environment, Food and Rural Affairs

DEM – Digital Elevation Model

DTM – Digital Terrain Model

EA – Environment Agency

EAD – Expected Annual Damage

ESRI – Environmental Systems Research Institute

FEH – Flood Estimation Handbook

FT1 – Flood time 1

FT2 – Flood time 2

FT3 – Flood time 3

GIS – Geographic Information Systems

HERE – Here Technologies

LA – Local Authority

LLFA – Lead Local Flood Authority

LSOA – Lower Super Output Area

MAUP – Modifiable Areal Unit Problem

MSOA – Middle Super Output Area

NUTS – Nomenclature of Territorial Units for Statistics

Definitions and Abbreviations

OA – Output Area

OSM - Open street map

ReFH – Revitalised Flood Hydrograph

UN – United Nations

Chapter 1 Introduction

1.1 Project Background

Flooding is a major global natural hazard with three billion people affected between 1990 and 2010 (Smith, 2013). The attractiveness of floodplains for urban development has resulted in a spatial concentration of populations in flood risk zones (Smith, 2013), with an estimated one billion people living in these environments (Di Baldassarre et al., 2013), amplifying flood risk exposure. Consequently, research into the environmental, social and economic aspects of flood risk is very active. Risk theory is used as a conceptual framework for assessing the consequences of hazards, such as floods, and is often represented through a risk equation (Equation 1.1).

$$\text{Risk} = \text{Hazard} \times \text{Exposure} \times \text{Vulnerability} \qquad \text{Equation 1.1}$$

Whilst spatial aspects of risk are always considered, as natural hazards are inherently spatial phenomenon, the temporal properties of risk are often neglected in research (Aubrecht et al., 2013a; Freire & Aubrecht, 2012). The temporal aspects of the risk equation are critical though as risk is not a static or ubiquitous concept and continually evolves. In particular, population movement is a dynamic factor at a local level (Dawson et al., 2011), as the interaction between where people are and the time of hazard onset can substantially change the exposure and vulnerability elements of risk (Freire et al., 2013; Aubrecht et al., 2014; Terti et al., 2015).

In the past it had been difficult to capture temporal population dynamics due to the difficulties of knowing the location of individuals at a given point in time and the effect of daily, weekly and seasonal fluctuations on population counts (Stepanov and Smith, 2009). The decennial census provides invaluable demographic data, however it represents a static night time population distribution (Smith et al., 2014). Spatiotemporal population modelling has been developed to address this issue, including the Population24/7 project in the UK which developed a framework using the SurfaceBuilder247 model (Martin et al., 2015) and Dynapop in the USA (Aubrecht et al., 2014). These can model daily, weekly and seasonal movements of the target population by combining census data with administrative data sets. The availability of temporally dynamic population data can be applied to enhance risk assessments.

The UK government describes flooding from all types as the greatest natural hazard risk to the UK (Cabinet Office, 2011). An estimated 5.4 million properties, 1 in 6 of those in England, are at risk of flooding with an average damage of £1.1 billion a year (Priestley & Rutherford, 2016). The endangerment of life, physical and mental health issues, economic damage to property and agricultural crops, loss of access to services and damaged transport networks are some of the consequences of flooding in the UK (Pitt, 2008; Cabinet Office, 2011). The government has committed to spending £2.5 billion between 2015-2020 on flood defences (Priestley & Rutherford, 2016). However, this capital investment is only expected to reduce flood risk by 5% overall (Priestley & Rutherford, 2016), so flood defences are not the only strategy which is required. Improving societal and infrastructure resilience and adapting to flood events is vital.

Pluvial flooding is the type of flooding caused by heavy rainfall and is prevalent in urban areas where natural drainage is reduced by impervious urban surfaces (Li et al., 2018). These events often occur over a short period of time (Melo et al., 2015), therefore reducing the ability for flood forecasts and warnings (Li et al., 2018; Boeing et al., 2019). Pluvial flood events in urban areas disrupt road networks as they can cause road links to be impassable for anywhere between a few hours to a couple of days, if there is insufficient drainage for the rainfall. These are small, localised events when compared to fluvial flooding and do not get as high profile news coverage. However, they can be significant for the affected communities. Within the context of climate change, it seems apparent that increasingly intense rainfall is very likely (Boeing et al., 2019), therefore increasing the likelihood of disruption to road users. Continued population increases and urban expansion also increase exposure of people and capital in at risk areas (Mechler & Bouwer, 2014). Therefore, methods for estimating the impact of pluvial flooding on road travellers is useful.

For UK infrastructure, natural hazards, especially pluvial flooding, are one of the top risks to maintaining normal function (Cabinet Office, 2011; Carter, 2015). Road networks are one aspect of infrastructure which are affected by flooding. They are 'lifelines' as roads provide access to health services, economic opportunities and leisure activities (Chen et al*.,* 2007; Platt 1995; Nyberg & Johansson 2013), so any closure can result in negative impacts for individuals, businesses and government. Figure 1.1 illustrates how a number of major roads (Motorways and A roads) coincide with the historical flood data layer (the combined maximum extents of all flood events since 1946). Whilst this does not necessarily mean that all these flood events resulted in road inundation, Figure 1.1 does show that large parts of the country's major road networks are at risk of flooding. During emergency situations, road networks are often the primary means of evacuating an area and providing vital resources (Rogelis, 2015), therefore maintaining function is essential. When damaged it is a priority to repair road links swiftly, however, this comes at a great cost. For example the emergency repairs to roads following Storms Desmond and Eva (2015) cost £40 million (Hankin et al., 2016).

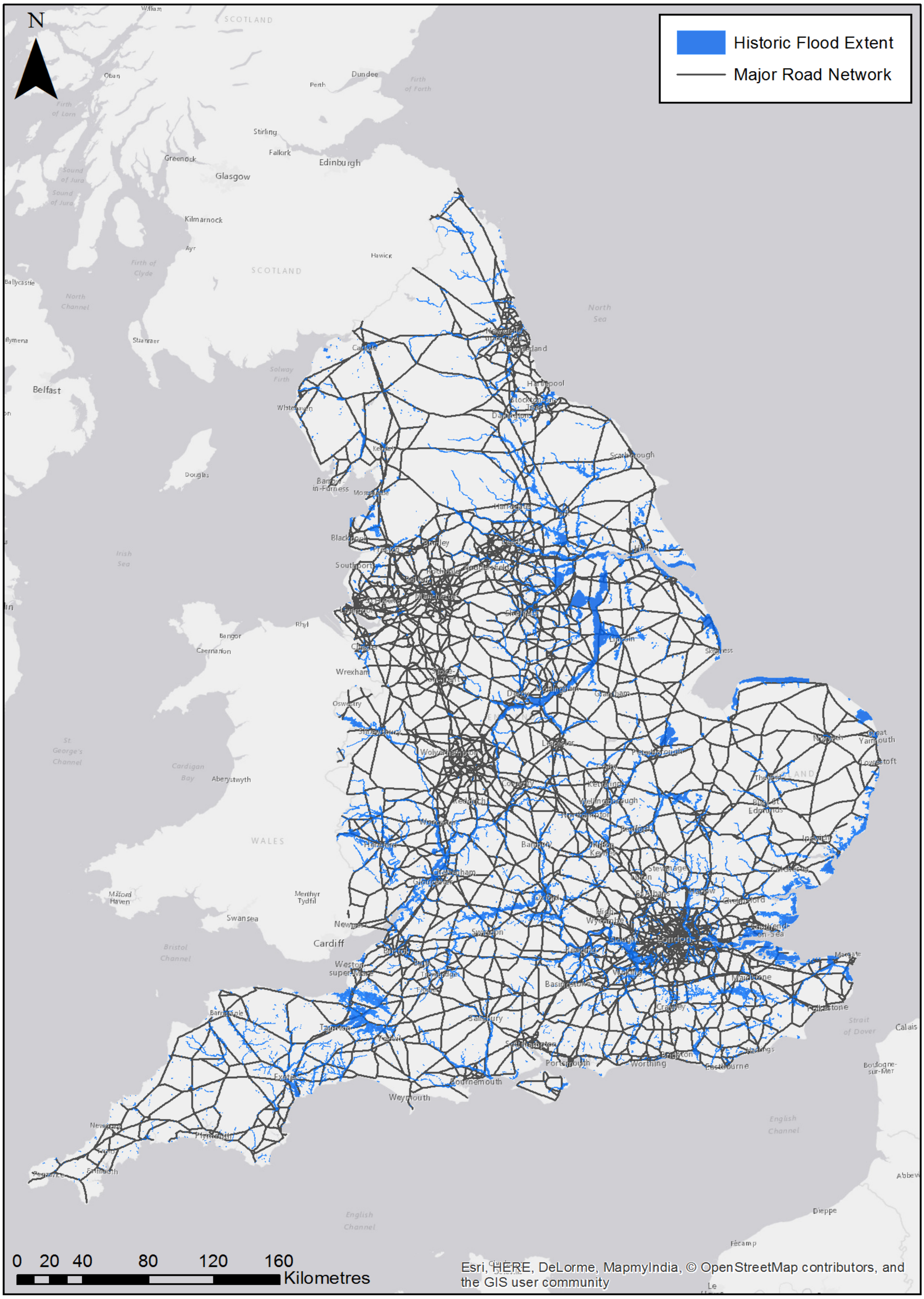


Figure 1.1 Map showing the Major Road Network in England (A roads and Motorways) compared to the historic flood extent map. Contains public sector information licensed under the Open Government Licence v3.0.

Flooding on the roads can affect a greater geographic area and population than the footprint of the flood itself due to the restrictions on travel and congestion. The 2007 flooding in Gloucestershire trapped up to 10,000 people overnight in their cars as the M5 and other surrounding roads flooded and there were no alternative routes. In terms of economic consequences, it is estimated the summer 2007 floods caused £674 million in damage to critical infrastructure, with a knock-on effect of £4bn in damage to the economy overall (Pitt, 2008; Cabinet Office, 2011). There are numerous examples of recent events which have negatively affected national infrastructure including the 2007 summer floods, 2009 Cumbrian Floods, 2013/14 winter floods in the South West and the 2015 winter floods in the North of England (Cabinet Office, 2011).

1.2 Statement of the Research Problem

As described in Section 1.1, the risk of flooding affecting road travel is widespread and the number of people travelling around urban areas varies greatly throughout the course of a day, week or year. As a result, the daily timing of a flood event, relative to these mobility patterns, is a critical factor in how a flood affects the local population. The dynamic nature of this problem increases the complexity of modelling the effects, and population mobility is seldom considered in flood risk assessments. Another source of complexity is that the conceptualisation of risk itself (Equation 1.1) should consider the spatial and temporal dynamics of each of its constituent parts. This thesis will attempt to encapsulate these complex spatial and temporal dynamics under one framework to further the development of spatiotemporal risk assessments for disruption to road networks.

The case study of pluvial flooding in York, UK has been chosen for the development of the spatiotemporal risk analysis framework, described in more detail in Section 4.1.1. Pluvial flooding has been chosen as the temporal scale of these events is short and they often occur with little pre-warning, meaning they are suited towards analysing the interaction with daily population movements. The population has been considered in subgroups, aligned to age and economic activity characteristics as these have been shown to have distinctive temporal characteristics. In this thesis the travel of commuters to work and primary school children to school is modelled.

1.2.1 Policy context

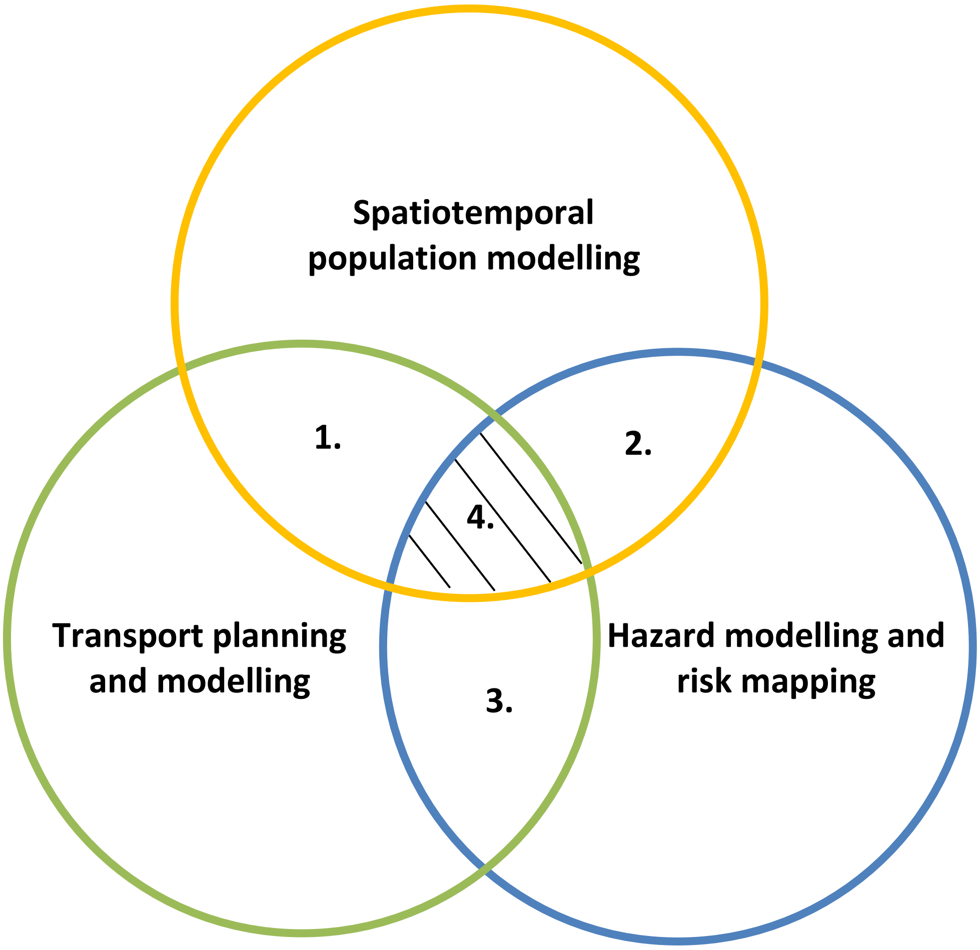
As the case study site is York, England, relevant government policy documents which apply to England were considered. The Pitt Review (Pitt, 2008) was a milestone in UK flood risk policy and was commissioned following the summer 2007 floods. The key recommendations from this report have translated into UK flood legislation, in compliance with European directives, and are referenced in subsequent government reports (for example Keeping the Country Running: Natural Hazards and Infrastructure, Cabinet Office, 2011). Table 1.1 outlines specific policy documents which are applicable to this PhD at both the national and local level of government.

Table 1.1 List of policy documents utilised in this thesis

Document	**Use in PhD**
The Pitt Review (Pitt, 2008)	Used for core understanding of UK government policy and the specifics of transport issues in flood events.
Keeping the Country Running: Natural Hazards and Infrastructure (Cabinet Office, 2011)	Guidance on past issues, current government strategies, definitions of critical infrastructure
National Flood Resilience Review (Cabinet Office, 2016)	Most recent review of flooding, includes sections on transport infrastructure resilience.
Surface Water Management Action Plan (DEFRA, 2018)	Most recent review which contains policy on the management of pluvial flooding.
Surface Water Management Plan (City of York Council, 2012)	Understanding specific issues and policy to pluvial flooding in York
City of York Strategic Flood Risk Assessment (City of York Council, 2013)	Background information on flood risk assessments in York
York 13 August 2018 Section 19 Surface Water Flood Investigation (City of York Council, 2019)	The most recent pluvial flood event which led to an investigation, used for comparing flood areas and understanding the disruption caused by a historical event.

1.2.2 Conceptual interdisciplinary diagram

The research problem of floods causing disruption to travellers and those living in the surrounding area is a complex one requiring the inclusion of several disciplines' theory and methods; population modelling, flood risk theory and modelling and transport studies (Figure 1.2). To achieve the aims of this research project, all three were combined producing innovation in the merging of each.



1. **Innovation: Linking gridded spatiotemporal population data to network analysis**
2. **Innovation: Improvement in the representation of people in travel for risk analysis**
3. **Innovation: Incorporating time of day of travel disruption into hazard risk assessments**
4. **Innovation: Bringing these disciplines together in one framework to provide insights on flooding disruption to populations travelling around an urban area**

Figure 1.2 The interdisciplinary elements and innovations which are incorporated in this thesis

1.3 Research Aims and Objectives

The overall goal of this research is:

To determine how daily variations in the urban population in travel interact with the time of pluvial flood onset and magnitude to influence spatial and temporal disruptions to journey times and destinations.

To address this goal, the following aims and objectives have been set for each results chapter:

Chapter 5

Aim: To develop a framework for combining spatiotemporal gridded population data with GIS network analysis, using journeys to primary school as an example.

1. Create a baseline spatiotemporal representation of primary school children's location across the study site, including the in-travel population.
2. Evaluate the use of spatiotemporal gridded data for representing population in travel.
3. To develop a method for converting the output from the SurfaceBuilder247 model into origin-destination flows for use in GIS network analysis.

Chapter 6

Aim: To assess the differences between using census origin-destination data and spatiotemporal population derived origin-destination data to measure the impact of a pluvial flood scenario on travel in York.

1. Generate pluvial flood scenarios for road travel disruption testing.
2. Analyse the differences in origin locations between census-derived origin-destination flows and flows simulated from workplace catchments using Chapter 5's methodology.
3. Examine the impact of the two different approaches to modelling spatiotemporal population flows on the simulated road travel disruption from pluvial flood hazards.
4. Test the inclusion of floods as obstructions in the network analysis tool.
5. Test methods for the aggregation and presentation of the results.

Chapter 7

Aim: To analyse whether the time of pluvial flood onset is more important than flood magnitude for disrupting commuter and primary school travel in York.

1. Analyse the extent to which pluvial flood magnitude and flood onset time affect journey times due to road network disruption.
2. Assess whether commuter's or school children's travel to their destination is affected the most.

3. Identify schools and workplaces which are most susceptible to disruption from flooding and the effect pluvial flood magnitude and pluvial flood onset time has on identifying them.

1.4 Novelty of this Thesis

As highlighted in Section 1.2.2, this thesis is interdisciplinary, and these innovations lead to the following novel aspects of the research:

- An original empirical assessment of hazard magnitude versus timing of flood onset in relation to disruption to the population in travel.
- It generates temporally specific origins and destinations from the output of the SurfaceBuilder247 spatiotemporal model for use in network analysis for the first time.
- It highlights how 'critical roads' can be viewed from the perspective of which population group is using a road network for a specific purpose, rather than from a topological perspective.
- Introduces a semi-dynamic workflow through syncing the temporal properties of population, traffic and flood data.
- A transferable and scalable methodology has been developed for assessing destination susceptibility to travel disruption

1.5 Thesis Structure

This chapter has introduced the research problem addressed in this thesis and stated the aims and objectives of the research. The novelty of the project has also been highlighted.

Chapter 2 provides the conceptual background for this research, including risk theory and the spatial and temporal components of risk. How risk is defined in the wider literature is discussed before the definition of risk which is applied in this work is stated. The application of risk theory within the context of the research problem is the third section.

Chapter 3 reviews the literature relevant to the three disciplines required for understanding the effect of pluvial flooding on road travel. The first section covers hazard modelling and theory for pluvial flooding, summarising pluvial flood policy and history in England, then the modelling methods used to simulate flood events and assess their risk. The second section reviews spatiotemporal population modelling, rooted in population geography, its origins, implementation and alternative new data sources for modelling people travelling across an area. This section also reviews the methods for spatiotemporal population modelling. The third section focusses on transport planning and modelling, summarising key concepts which are applied in this thesis. Data and methods for analysing people in travel and network analysis are also presented. The fourth section draws the three disciplines together to the specific issue of modelling disruption to road travellers from pluvial flooding by applying temporally specific population data. The final section presents the evidence gap which has been highlighted through the review of the literature.

Chapter 4 discusses the specific data, methods and models applied in this thesis. The study design, including choice of case study, population groups and flood scenarios are covered in the first section. The second section describes the models used for pluvial flood modelling, spatiotemporal population modelling and network analysis in turn. The third section presents an overview of the thesis and the methodology of each of the three analysis chapters which follow.

Chapter 5 is the first results chapter and focusses on the development of a framework for converting spatiotemporal population data, from SurfaceBuilder247, into a point format for use in GIS network analysis. The York study area is used, and the focus is on primary school children. A spatially weighted Monte Carlo analysis was used to convert the gridded data into a point format, meaning the temporally specific population data could be used in network analysis.

Chapter 6 is the second results chapter. The method developed in Chapter 5 is applied to a different population group, commuters, and compared to census origin-destination flow data. One pluvial flood scenario is introduced to gain an initial understanding of the impact of flooding in the network analysis model and resultant travel times.

Chapter 7 is the final results chapter and addresses the central question of whether the time of day a pluvial flood event occurs at has as much of an effect on travel times as increasing the magnitude of a flood event. A semi-dynamic methodology is applied where three flood onset times (6am, 7am and 8am) are tested for two flood magnitudes (1 in 30 year and 1 in 100 year). The routes to schools and commuters were modelled using network analysis and the travel times for each destination across the scenarios were compared to a non-flood baseline.

Chapter 8 analyses the results presented in the previous three chapters to identify key contributions which have emerged from this research. The transferability and scalability to other sites, natural hazards, population groups and transport systems is then discussed before the uncertainties and limitations of this work. The final two sections of the chapter cover the future recommendations for research and recommendations for policy which can be made.

Chapter 9 summarises the work conducted in this thesis and states how the aims and objectives of the thesis have been met. The contribution of this research and future research directions are also summarised.

Three appendices follow the conclusion. Appendix A describes the population data applied in the spatiotemporal modelling part of this, Appendix B the conference papers submitted during this research project and Appendix C additional figures for Chapter 6.

Chapter 2 Conceptual Framework

This chapter sets out the research problem addressed in this thesis and the risk theory which underpins it. Risk is an integral part of society and a person's evaluation of risk is an important determinant of human behaviour (Slovic et al., 2004). Natural hazards like floods, hurricanes, earthquakes, landslides and volcanic eruptions, are some of the most dangerous events to human health and wellbeing. Hence judging the risks associated with them is vital. This chapter is divided into three sections; Section 2.1 defines risk and its constituent parts, Section 2.2 outlines the spatial and temporal aspects of risk and Section 2.3 covers the application of risk theory to the research problem set out in Chapter 1.

2.1 Defining Risk

There are several ways of defining risk and the components which contribute to a risk. This section compares the different definitions of risk and states the definition used in this thesis. The hazard, exposure and vulnerability elements of risk are then covered.

2.1.1 Risk

Risk is complex and there are various perspectives on its framing; from the 'risk equation' of the natural sciences (Romieu et al., 2010; Fuchs et al., 2013) to Beck's 'Risk society' concept (Beck, 2002) in the social sciences. Therefore, the definition of risk is often specific to the context of the analysis being conducted. There are two perspectives which natural hazards can be viewed from; 'hazard risk', which focuses on the physical hazard and the damage the physical processes cause, and 'disaster risk', which focuses on the socioeconomic factors that combine with the physical hazard to cause a disaster. The natural sciences have previously focussed on the 'hazard risk' perspective, which involves linking physical hazards to their expected damage, through calculating the vulnerability of physical objects (Romieu et al., 2010; Fuchs et al., 2013). This is likely due to the relative ease of calculating physical hazard impacts and economic exposure compared to quantifying social vulnerability (Fekete, 2009; Koks et al., 2015). Following this perspective, risk can be defined in the natural sciences as the expected damages or potential loss to an exposed system from a hazard with a given probability (Aubrecht et al., 2013a; Yin et al., 2016). This probabilistic view of risk (Carrão et al., 2016) can be expressed in the form of Equation 2.1 and is applied by many researchers in the hazards literature (for example Eiser et al., 2012; Klijn et al., 2015; Kron, 2005; Yin et al., 2016).

Risk = probability(of hazard) x consequences(of hazard) Equation 2.1

This is useful in studies of physical objects, like buildings, or economic exposure. However, this definition does not account for the social factors of risk. The United Nations International Decade for Disaster Reduction in the 1990s marked the shift from a hazard-centric approach to using risk as a central component of disaster management (Cutter et al., 2008), providing a more encompassing view of a situation and the uncertainty associated with risk analysis (Aubrecht et al., 2013a). Since the International Decade for Disaster Reduction, it has been argued that socioeconomic variables are a bigger influence on a 'disaster' than the 'natural' hazard (Aubrecht et al., 2013a; Wisner et al., 2004). Therefore, by framing risk in a way which captures social vulnerability, it becomes an assessment of 'disaster risk' rather than just 'hazard risk' (Romieu et al., 2010). This acknowledges that disasters are not just natural hazards (Aubrecht et al., 2013a), but a combination of natural processes and human actors. The most recent UN framework, 'The Sendai Framework for Disaster Risk Reduction 2015-2030' (Nations Office for Disaster Risk Reduction United Nations, 2015), cements this focus on disaster risk management into the next two decades.

To encapsulate disaster risk, a common definition considers risk as a product of three components; hazard, exposure and vulnerability (Equation 2.2) (Mechler & Bouwer, 2014; Koks et al., 2015; Freire et al., 2015). This is based upon Crichton's 'Risk Triangle' (Crichton, 2008) and is favoured by spatial planners and social scientists (Albano et al., 2015). In some applications this equation differs by only including hazard and vulnerability (Equation 2.3), as vulnerability can be considered to incorporate exposure (Romieu et al., 2010). However, the United Nations Office for Disaster Risk Reduction considers exposure as independent (Aubrecht et al., 2013b) and Klijn et al., (2015) argue the understanding of risk is richer for this division.

Risk = Hazard x Exposure x Vulnerability Equation 2.2

Risk = Hazard x Vulnerability Equation 2.3

2.1.2 Hazard

A hazard can be described as a threatening event (Kron, 2005) or phenomenon which can result in harm to people, economic assets or the environment. In line with the concept of 'hazard risk' described above, hazard is often addressed in a quantitative way, through a measure of the probability of occurrence or the magnitude of the hazard event. In reality, it is rare for there to be a single hazard in isolation and there are often compound hazards triggered by the main hazard. For example, in a flood situation there are hazards associated with the floodwater itself, affected by its depth and velocity, but also secondary hazards, like waterborne diseases and debris. All possible hazards require consideration to fully understand the risk of a situation and ensure measures to mitigate one hazard do not lead to increased danger from another related hazard. However detailed examination of all hazards is difficult and often not possible within the scope of a study, leaving the main hazard as the focus. Finally, a hazard does not exist in isolation as it is a component of risk and the magnitude, spatial and temporal properties of it affects the level of exposure and vulnerability of a society.

2.1.3 Exposure

Exposure is considered the amount of population, infrastructure and other assets within a potential hazard area (Hirabayashi et al., 2013). Population data are used to calculate the number of people within a hazard zone (Smith et al., 2014). Whilst this does bring humans into the consideration of risk, it is more of a 'head count' of the number of people in an area and does not consider the social factors of why risk varies between population sub-groups in depth. Exposure often includes an economic measure of the damage which could be inflicted by a hazard of a given size (Jongman et al., 2014). Asset exposure is a key part of calculating Average Annual Loss (AAL), the value used in the insurance industry to assess the risk of insurance losses (Hallegatte et al., 2013). So, whilst hazard tends to be framed within the physical sciences, exposure is often framed within economics and the quantifiable number of people in the area affected by a hazard. As discussed in Section 2.1.1, 'disaster risk' involves understanding social as well as physical components in risk. Therefore exposure is closely linked with vulnerability, as a population may be equally exposed to a hazard but the unique vulnerability of certain individuals may mean the risk to them is greater (Wisner et al., 2004).

2.1.4 Vulnerability

If exposure is a numerical count of what is in a hazard zone and the economic value of it, vulnerability captures the nuances of society and the underlying reasons of why risk is unequal. It is not simply a function of proximity to the hazard, which is why it is different from exposure and should be considered separately, as stated in Equation 2.2. There is no universally agreed term of vulnerability (Papathoma-Kohle et al., 2011), partly due to the wide range of disciplines the term is applied in and the fact it is difficult to quantify as it involves a multitude of complex qualitative factors (Fekete, 2009). Vulnerability is defined here as the degree a person is able to help themselves during a hazard event, influenced by the presence, or lack of, resources to cope with a hazard (McFadden et al., 2006; Civil Contingencies Secretariat, 2008).

The presence or absence of resources is important as it can be argued that vulnerability originates from underlying social conditions which limit access to resources, rather than the hazard event (Cutter et al., 2008). Wisner et al., (2004) conceived of the 'Pressure and Release' model of disasters which is a tool for understanding vulnerability. They argue a hazard is a 'triggering event', and a disaster only arises due to entrenched social structures, like macro-economic policies, unsafe living conditions and lack of access to resources. When a hazard strikes, it acts as a tipping point in already precarious lives caused by social vulnerability. Whilst Wisner et al.'s (2004) model was designed for developing countries, it still has relevance for the UK as those who are poorer or less physically fit are often those who are identified as being most vulnerable to events like floods.

Vulnerability feeds back into exposure, as those who are most vulnerable can end up living in cheaper housing, which has been shown to be a factor in hazard exposure (Wisner et al., 2004). Vulnerability is also shaped by risk perception, as how a person views their vulnerability can influence their behaviour and potentially lead them to experience higher exposure if they underestimate their own vulnerability. The 'levee effect' describes how people do not think they are vulnerable to flooding once a defence is in place (Wisner et al., 2004). But defences may be breached and are only designed for a certain level of flood, so the actual vulnerability is higher than those who live there are likely to perceive (Pitt, 2008). Therefore, this makes them vulnerable as they may not take sufficient action during an event.

2.2 Spatial and Temporal Dimensions of Risk

As described in Section 2.1.1, the appropriate definition of risk which can be applied to this research problem is that risk is a function of hazard, exposure and vulnerability (Equation 2.4). However, risk is not a static concept and fluctuates both spatially and temporally when comparing the same hazard event (Mechler & Bouwer, 2014). Mechler and Bouwer (2014) define risk as a *dynamic* function of hazard, vulnerability and exposure, and the word 'dynamic' is crucial as it acknowledges that risk is an evolving concept through space and time. The dynamic nature of risk needs to be considered for effective risk management; however, it is a challenging task as the scales of spatial and temporal change are complex.

$$\text{Risk} = f(\text{Hazard},\text{Exposure},\text{Vulnerability}) \qquad \text{Equation 2.4}$$

In this thesis the spatial and temporal dynamics of risk are under consideration, so Equation 2.4 needs extending. The majority of academic literature does not acknowledge this in their definition of risk (Freire and Aubrecht, 2012). Therefore for this thesis, it is necessary to adapt the 'risk equation', as it does not explicitly include space or time (Hu et al., 2017). Variations over space and time are present in all three components of the risk equation (Lugeri et al., 2010). Therefore, in Equation 2.5 this is represented, where S represents spatial aspects and T temporal aspects, and this definition will be applied in the remainder of this thesis. Whilst both space and time are important, time is of particular interest in this work as examining risk over time adds a dynamic quality to the research findings.

$$\text{Risk} = f(\text{Hazard}_{ST}, \text{Exposure}_{ST}, \text{Vulnerability}_{ST}) \quad \text{Equation 2.5}$$

Variations over space and time are present in all three components of the risk equation (Lugeri et al., 2010), which means risk itself is not fixed in time nor equal over geographic space. Figure 2.1 is a conceptual model which details how hazard, exposure and vulnerability combine to form risk and examples of the spatial and temporal components.

The spatial impact of a hazard primarily lies with the location of the hazard relative to the population, infrastructure and assets. Often, secondary hazards are triggered by the primary hazard, for example water-borne disease from standing water, creating hazards chains. Temporal factors of hazards include the time of day and year it occurs at, the return period of a given magnitude and, for hydro-meteorological hazards, the impact of climate change. Exposure is the sum of assets, infrastructure and people in the hazard zone and the spatial distribution of these elements, particularly of people, is highly time-dependent (Aubrecht et al., 2012). The temporal dimension of exposure is at several scales; the shorter term of daily and seasonal fluctuations, often due to population change (Freire et al., 2013; Smith et al., 2014; Smith et al., 2015), or the longer term of years to decades, due to urban expansion and demographic changes. Thus fluctuating exposure can lead to changes in risk, regardless of changes in the hazard severity (Aubrecht et al., 2012; Freire et al., 2013; Liu et al., 2014; Mechler and Bouwer, 2014).

The spatial context of vulnerability is the location of vulnerable groups, the services they require and how easily they are accessed (Rodríguez-Núñez et al., 2014; Dawson et al., 2011). These are considerations present outside of a hazard scenario, but during a hazard event may become more acute (Hemingway & Priestley, 2006). The temporal context is whether vulnerable groups are moving, or whether they are static. This would influence the level of exposure they could face at different times of the day or year. An example of a spatially mobile group would be tourists and school children (Smith et al., 2015), whilst those admitted to hospital as in-patients or in residential care homes are less likely to be moving away from their location. Demographic changes, particularly the increase in the number of older people (Donner & Rodríguez, 2008), is another factor which will cause vulnerability of a population to evolve.

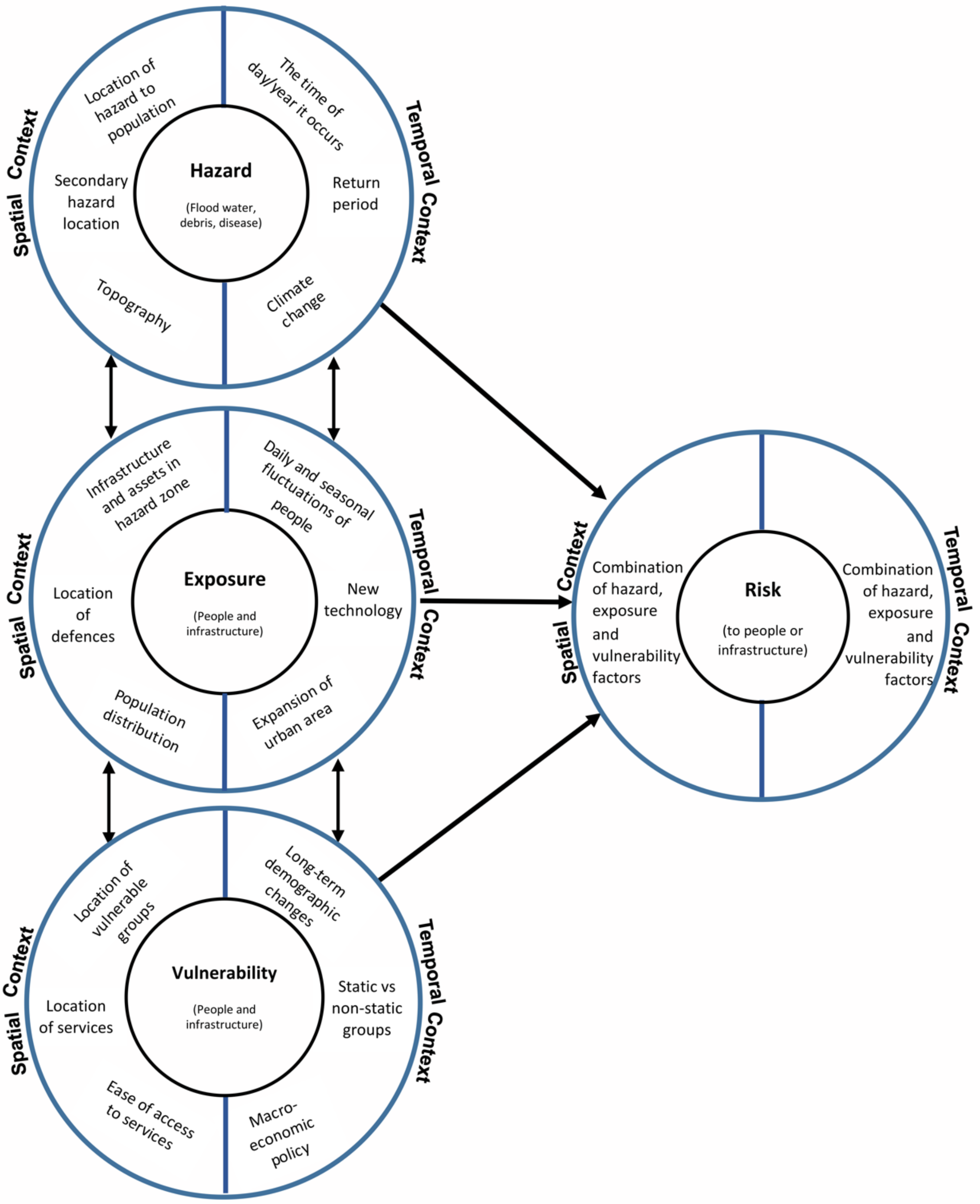


Figure 2.1 Conceptual diagram showing how hazard, exposure and vulnerability combine to generate risk with their spatial and temporal dimensions, from the perspective of flood risk. Each element of the risk equation is detailed in the centre, with examples of the temporal and spatial context given around the outside. Arrows show the connections between elements.

In addition to existing as its own entity, the three components of risk interact with each other. The spatial and temporal context of a hazard influences exposure, as shifts in the location of the hazard lead to different levels of exposure. Exposure influences vulnerability in turn (Kienberger et al., 2013), as at some level all who are in an area exposed to a hazard are vulnerable compared to those outside of it. However, this connection is mutual, with the temporal context of vulnerability moderating exposure. For example macro-economic policy causing an unequal distribution of resources (Wisner et al., 2004) means those who are more socially vulnerable often live in areas particularly exposed to hazards where property is cheaper (Kaźmierczak & Cavan, 2011). This interaction of vulnerability, exposure and hazard has been highlighted previously in other conceptual models. For example Wisner et al.'s (2004) pressure and release model and Cutter's (1996) 'Hazards of place model' both demonstrate the complex interactions of natural hazard risk. However, these conceptual models do not identify the temporal component of risk as explicitly as the conceptual model in Figure 2.1.

Flooding is considered a 'known hazard' as it occurs in the same locations (Wisner et al., 2004) and the probability of a given magnitude can be calculated statistically. Thus, conventional flood risk assessments look at the probability of flooding and its consequences (Dawson et al., 2011). However focussing solely on physical parameters does not capture the complexity of human and natural processes which interact to form flood risk, as it is not a simple function of probability of exposure (Donaldson et al., 2013). Over the last decade it has become apparent increases in flood risk are primarily due to social factors, rather than changes in flood magnitude (Mechler & Bouwer, 2014). Hence, including socio-economic factors to describe the vulnerability and exposure of affected populations has become more commonplace (Romieu et al., 2010).

2.3 Application of Risk Theory to the Research Problem

This section takes the risk theory covered previously in this chapter and applies it specifically to the research problem outlined in Section 1.2.

2.3.1 Mapping the research problem onto the risk equation

To restate the research problem, this thesis aims to analyse how the time of pluvial flood onset and flood magnitude affects the disruption to travel across road networks for two groups of the population. Figure 2.2 highlights how Equation 2.5 can be applied to this research problem and questions the results in this thesis should answer. Developing the risk equation in this way, particularly with the consideration of time, for use as a conceptual structure to hazard-related transport research project is novel. This framework can be applied to other hazards other than flooding, for example earthquakes and landslides.

Risk	f(**Hazard$_{ST}$**,	**Exposure$_{ST}$**,	**Vulnerability$_{ST}$**)
Risk of restricted movement increasing travel time and reducing time spent at work and school. Are there places in space and time which can be identified as having greater risk?	**Spatial:** The extent and depth of floodwater. Does a larger magnitude rainstorm lead to a greater disruption? **Temporal:** The time of flood onset and the duration of the rainfall. How do changes in the time of flood onset affect disruption	**Spatial:** The population head count by location on the road network; the location of flooded roads; the location of workplaces and schools. Are some locations identified as being exposure hotspots? **Temporal:** The varying counts of commuters and children using the road depending on the time of day. Does exposure vary in the time period	**Spatial:** The geographic concentration of a vulnerable group (primary school children). How is this group affected? Which destinations are most susceptible to disruption across the city? **Temporal:** The changes to vulnerability over time. Do changes in time of flood onset alter the disruption seen to primary school children or change which destinations are

Figure 2.2 The application of the risk equation to the research problem of this thesis, and questions to answer with the results presented in later chapters.

2.3.2 Spatial and temporal scales in the research project

As Figure 2.1 shows, there are spatial and temporal components to hazard, exposure and vulnerability. These factors operate at different spatial and temporal scales, for example macro-economic policy is a larger spatial and temporal scale than the location of individuals at a given time of day. Figure 2.3 summarises the spatial and temporal scales of hazard, exposure and vulnerability applied in this thesis. The spatial scale is defined by probable limits to the size of flood models, the likely average distances travelled by daily commuters (15 kilometres (Office for National Statistics, 2014)) and emergency service catchments. Macro-economic policy and climate change operate at a spatial scale which is larger than the research problem considered in this thesis. Pluvial flood events occur over reasonably short timescales (hours to days) therefore this is a logical threshold for the temporal scale. In addition, the data which are being used is census data from 2011-2016 and rainfall data from the last few decades. Hence, these data are only reasonably applicable to near-future scenarios (up to 10 years after the data reference point) where substantial changes to the demographic makeup of the population and hydrological cycle has not occurred. Applying this temporal threshold to exposure and vulnerability, short-term changes like daily movement and the current structure of the road network are included in the scope of this project. However longer processes like urban expansion, infrastructure updates and demographic changes are not considered as they exceed the temporal threshold.

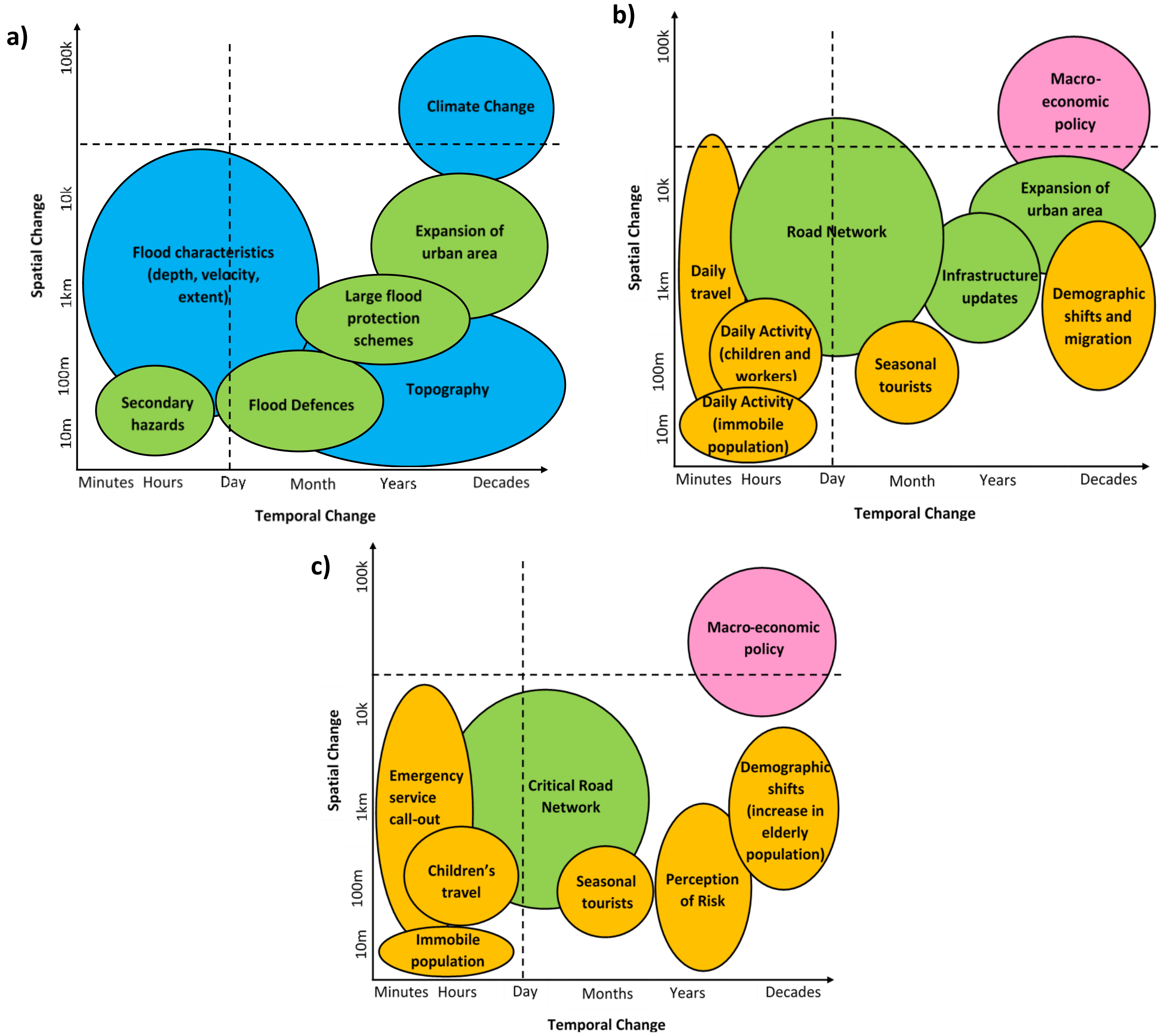


Figure 2.3 Plots showing the spatial and temporal characteristics of a) Hazard, b) Exposure and C) Vulnerability in this research project. The dotted lines indicate the boundaries of the temporal and spatial scales considered in this project. Connected circles indicate a link between factors. Blue: Environmental Factors and Processes. Green: Man Made Structures. Yellow/orange: People. Pink: Policy

2.3.3 Hypothesised impact of varying timing of flood onset

This section describes the likely effects of changing the time of flood onset and the impact on risk and road travel. Table 2.1 demonstrates some scenarios that might be observed when keeping the hazard characteristics constant, except the timing of onset. Scenario 1 happens during the day and Scenario 2 happens at night, which changes the number of people who are affected; therefore, exposure and vulnerability change and risk varies. Seasonal impacts are also important due to the potential of increased populations in summer due to tourism, or decreased population during university holidays, resulting in different total population numbers and the number of vulnerable people (Scenario 3 and 4). A flood which occurs early in the morning will impact morning commuting traffic (Scenario 1), however if it has cleared in the evening then the number of journeys taken will rise, although not to normal levels as some were unable to reach their normal daytime destinations. In comparison, a flood that strikes in the afternoon (Scenario 3) will not impact the morning commute but will limit evening travel and may cause above normal levels of journeys into the evening as the flood waters recede and roads are reopened. These examples demonstrate the changes to risk from shifting the time a flood occurs at. This level of temporal resolution, changing the time of flood onset during a single day, has not often been explored in wider literature and policy documents.

Table 2.1 Hypothetical examples to illustrate how keeping the hazard element as a comparator, whilst varying the exposure and vulnerability elements, changes the overall population at risk.

	Hazard	**Exposure**	**Vulnerability**	**Risk**
Scenario 1	1 in 100-	Daytime population Road network	Daytime population e.g children, elderly, students Road access	750 people 100 Vulnerable
Scenario 2	1 in 100-	Night-time population Road network	Night-time population e.g children, elderly Road access	500 people 70 Vulnerable
Scenario 3	1 in 100-	Summer Daytime population Road network	Summer population e.g. tourists Road access	1500 people 300
Scenario 4	1 in 100-	Winter Daytime population Road network	Winter population Road access	600 people 100

2.4 Conclusion

The interdisciplinary nature of this project requires more than just assessing the probability and consequences of a flood event in order to fully understand flood risk. A disaster risk perspective with vulnerability and exposure is required, explicitly distinguishing the hazard, exposure and vulnerability components. It has also been shown how spatial and temporal aspects of risk need to be considered for a dynamic perspective. By using the risk equation with specific spatial and temporal components, this focusses the research on these aspects.

Chapter 3 Literature Review

This chapter is divided into five sections. The first three summarise the relevant literature around the three key disciplines dawn upon in this thesis; flood risk, spatiotemporal population modelling and transport modelling. The fourth section examines literature on the specific research problem where these disciplines overlap. The final section draws these disciplines together and summarises existing evidence concerning the research gap that this thesis is addressing.

3.1 Review of Flood Risk Concepts

Following on from Chapter 2 and the conceptualisation of risk, this section focusses on the specific issues associated with pluvial flood risk management in England.

3.1.1 Mechanisms of flooding

A flood can be defined as “the temporary covering of land by water as a result of surface waters escaping from their normal confines, or as a result of heavy precipitation” (Kron, 2005, p.58). There are three main types of flooding in the UK; coastal, fluvial and pluvial. Coastal flooding occurs when storms cause sea water to surge onto the land, and the combination of storm surge and point in the tidal cycle determine how high it will be. Fluvial flood events are from riverine sources and occur when water exceeds the rivers bankfull capacity, due to heavy rainfall. This can either be rapid through rainfall directly into the channel, which is more common in smaller streams in steeper catchments, or rainfall in the wider catchment running into the river over a longer period of time. The ability to forecast heavy rain events and the lag time between rainfall and flooding is often sufficient in the UK’s rivers to be able to provide warnings up to a few days before the high river levels, meaning the surrounding population can prepare for disruption. Finally, pluvial flooding is the result of heavy rainfall from convective rainstorms onto an impermeable surface causing flooding in the place it falls (Boeing et al., 2019). These floods are common in urban areas due to extensive changes to the natural land cover decreasing permeability (Li et al., 2018) and often occur over a short period of time (Melo et al., 2015), reducing the ability for flood forecasts and warnings (Li et al., 2018; Boeing et al., 2019). Due to their shorter time frames than fluvial floods, pluvial floods are less likely to be predicated and more likely to cause unexpected disruption to travellers on the road network.

Pluvial flood risk emerges from the interaction between biophysical (duration and intensity of rainfall) and human factors (impermeable surfaces in urban areas, vulnerability of exposed population, installation of complex infrastructure) (Jenkins et al., 2017). Therefore, the risk cannot be determined on the characteristics of the flood hazard alone. The number of pluvial floods will likely to increase in severity as a result of climate change, due to greater potential extremes in rainfall (Houston et al., 2011). The characteristics of pluvial flood hazards mean they are a suitable hazard to select for this thesis and are a relevant issue in urban areas.

3.1.2 Recent English flood history

For fluvial floods, it has been described that the UK has experienced 'flood rich' and 'flood poor' periods over the last century. For example, Wilby and Quin (2013) found that there were three flood rich episodes in river catchments; 1908-1934, 1977-1988 and 1998 onwards. National trends for pluvial flood events, particularly sudden onset localised urban floods, are not apparent as there is a high year-to-year climate-led variation in rainfall and thus pluvial flood frequency (Stevens et al., 2016). There has a been an increase in insurance loss for all flood events in England over the last century, which is largely explained by a growth in exposure rather than a change in hydrological regime (Stevens et al., 2016). Over the 20th century, the population of the UK grew from 38.2 million to 59.1 million, with the number of properties rising from 7.7 million to 24.8 million (Stevens et al., 2016). This has resulted in large expansions to urban areas, including into the floodplains, and the impermeable material they contain. Rapid urban development of the past few decades has been reliant on an ageing drainage infrastructure (Boeing et al., 2019) in combination with a reduction in land surface permeability. This has led to an increase in property exposure to pluvial flooding and an increase in exposure to people travelling along the road networks. Presently, the EA estimates 3.2 million properties in England are currently at risk of pluvial flooding (DEFRA, 2018), a higher number than at risk from fluvial and coastal (Boeing et al., 2019). Despite the increased risk, the recent UK government surface water management plan concluded pluvial flooding has been poorly understood and the risks not communicated to the public effectively (DEFRA, 2018).

The effect of predicted changes on weather and climate extremes is an important issue when planning for natural hazards, particularly for pluvial floods as they have been identified as the type of flooding most likely to increase in severity with climate change (Houston et al., 2011). Continued population increases and urban expansion are also likely to increase exposure of people and capital in at risk areas, a dominant cause of predicted increases in financial losses over the next century (Mechler & Bouwer, 2014). These biophysical and socioeconomic changes mean assessing the risks of pluvial flooding will become increasingly important.

3.1.3 English flood policy

Over the last century, there have been three paradigms used as the basis of flood policy: 'land drainage' (drain land in order to build on it), 'flood defence' (build 'hard' defences to prevent water reaching developments) and 'flood risk management' (measures to reduce the amount of flood water and reduce the building on at risk land) (Penning-Rowsell et al., 2014). More recently the approach of 'learning to live with water' has become more common in policy agendas (McEwen et al., 2017) as it recognises not all flooding can be prevented (Wisner et al., 2004). Flood risk management requires negotiation of socially acceptable levels of risk and exposure (McEwen et al., 2017), and it is not a simple function of level of exposure as there is a mixture of human and biophysical processes which determine the risk (Donaldson et al., 2013). The following summary of recent flood risk legislation or regulation, with a focus on pluvial flooding, demonstrates this need to balance human and physical factors of risk, and to develop a greater understanding of flooding in order to minimise disruption to society.

The Pitt Review (Pitt, 2008) was a milestone in UK flood risk policy. Commissioned following the 2007 summer floods, its aim was to review the decisions and strategies implemented in 2007 and to provide recommendations to improve the countries flood risk response. Whilst this was in response to a fluvial flood event, the recommendations have influenced all areas of flood risk policy. Following the Pitt Review, the Flood and Water Management Act (2010) was passed which sought to provide clarity of the roles and responsibilities of the EA, local government, water and sewerage companies and a framework for assessing flood risk. This legislation established that the EA has strategic responsibility for flooding on main water courses and coastal erosion whilst Lead Local Flood Authorities (LLFAs) were established to take responsibility for flood risk from surface water, groundwater and ordinary watercourses in local areas (Bevan, 2018).

Several reports were commissioned by the government in the following decade to meet recommendations from the Pitt Review, and those relevant to this research are highlighted here. Firstly, the report 'Keeping the Country Running: Natural Hazards and Infrastructure' was published by the Cabinet Office in 2011. A main issue in 2007 was infrastructure being impacted by flood water. From roads becoming impassable, with the notable example of motorists stranded on the M5, to sewerage and water treatment centres becoming inoperable due to flood water, leading to water shortages. Secondly, the 'National Flood Resilience Review' (2016) reviewed the steps that had been taken to increase flood resilience since 2007 and what future action should be taken. In this report, the government committed to consider issues related to pluvial flooding in a future report specifically on the topic. In 2017, pluvial flooding in an urban area was added as its own risk to the National Risk Register (Cabinet Office, 2017).

The most recent report for England on pluvial flooding is the ‘Surface Water Management Action Plan’ published in 2018. This review stated that pluvial flooding causes significant disruption to people’s lives and livelihoods and will become a growing challenge with more frequent heavy storms, due to climate change, over the next few decades, compounded by increased urban development (DEFRA, 2018). To manage this risk water needs to be drained effectively by maintaining and improving where necessary current drainage systems and including sufficient drainage in any new urban developments (DEFRA, 2018). However, the spatial pattern of pluvial flood risk is less obvious than fluvial or coastal flood areas meaning awareness is lower, leading to less mitigation action currently in place. The report made several recommendations to improving the capability of dealing with pluvial flood risk. These include first, risk communication needs to be improved to stakeholders through pluvial flood risk mapping and developing better systems for short term forecasts. Current methodologies are unable to determine where heavy rainfall will be with the necessary levels of precision for a useful forecast lead time (DEFRA, 2018). Second, infrastructure resilience needs to be prioritised and improved through reviewing the resilience of current infrastructure against extreme pluvial events and enacting change to increase resilience. Third, the roles and responsibilities of all organisations need to be clarified and LLFA need to be guided on local flood risk management strategies and their role in overseeing work. Finally, local government capacity will be built through advice and support to develop the skills of staff.

The following governmental structure currently exists in England for managing flood risk. The Department for Environment and Rural Affairs (DEFRA) leads the policy for flood and coastal erosion risk management. The Cabinet Office is involved with emergency response planning (under the UK Civil Contingencies Act 2004) and the Department for Communities and Local Government are involved with land-use and planning policy (under the Local Government Act 2000). As outlined in the Flood and Water Management Act (2010), the Risk Assessment and Management Authorities are the EA, LLFAs, District and Borough Councils, Coast protection authorities, water and sewerage companies, internal drainage boards and highways authorities. These authorities are required to co-operate with each other and act within the framework of national and local flood risk strategies whilst sharing information with each other (Local Government Association, 2020; Pitt, 2008). The EA provides the strategic overview of flooding for the country and risk management activities for main rivers and the coast, and works with the Met Office to provide flood forecasts and warnings (Local Government Association, 2020). LLFAs are county councils and unitary authorities with the lead in managing local flood risks from surface water, ground water and ordinary watercourses. LLFAs are required to; prepare and maintain a local flood risk management strategy, carry out works to manage local flood risks, maintain a register of assets, investigate significant local flooding and publish the results and play a lead role in emergency planning and recovery after a flood event (Local Government Association, 2020). Highways authorities have the lead responsibility for providing and managing highway drainage and roadside ditches under the highways act 1980 (Local Government Association, 2020).

The availability of flood insurance is also an important component of English flood risk policy. Rising insurance losses from flood events over the last decade led to rising insurance losses, with the 2015 floods in Northern England alone causing over £1.3 billion in insured damages (Marsh et al., 2016). This has affected the willingness of insurance companies to insure homeowners on floodplains (Lamond et al., 2009). The perception of increasing flood risk and growing economic loss led to the creation of FloodRe (FloodRE, 2020), a reinsurance scheme created to ensure homeowners could purchase home insurance (which includes flood insurance) for a reasonable price. This formalised a previously informal agreement between the government and insurers, as there was a danger of insurers ceasing affordable cover for flood risk properties.

3.1.4 Flood transport policy

General flood policy was summarised in Section 3.1.3 and this section summarises flood policy as it specifically relates to travel disruption, the focus of this research. In the UK, the Department for Transport's 2014 Transport Resilience Review (Brown, 2014) is the most recent government report on natural hazard impacts on transport. In the wake of the Pitt Review (Pitt, 2008), the Transport Resilience review recommended that local authorities define a 'resilient network' which will be given maintenance priority in order to reduce the economic impact of any disruption from events like floods. The National Security Strategy (Cabinet Office, 2011) states the criteria for determining whether local infrastructure is critical, and therefore should be targeted to be part of the resilient network, is whether the loss of it would cause, or be likely to cause, a local emergency (Cabinet Office, 2011). This guidance could be incorporated into flood risk reviews as it shows that critical roads are location specific and could be effectively identified through examining their importance to the community, rather than through traditional network techniques to identify critical links. Further discussion is given in Section 3.3.3.

One of the key themes of the Surface Water Management plan (DEFRA, 2018) is to make infrastructure, including road networks, resilient to pluvial floods with future action to review the resilience of infrastructure networks against extreme pluvial floods stated as an outcome of the report. This demonstrates the intersection of flood and transport policy and the importance of understanding the effect of pluvial flooding on road networks. However, a standard method has not been created to assess this resilience. The government has identified this is an area which requires review but has not published the results of any reviews.

The depth of flood water which is considered disruptive is important when creating simulations to determine disruption. It is widely cited as between 25-30cm depth in the UK and other countries like the US (Li et al., 2018; Arrighi et al., 2019; Pregnolato et al., 2017b). This is because it is the height of the air inlet on the average car (Li et al., 2018) and water entering a car engine can cause engine failure, trapping motorists in flood water. Consequently, government, insurers and breakdown companies have created warning emails and social media images to instruct drivers not to traverse flood water which is deeper than 30cm. This guidance has resulted in this threshold applied to academic papers on travel disruption from road closure. Most studies use a closed/open switch at 30cm depth, but Pregnolato et al., (2017b) developed a disruption-depth function which modelled the safe travel speed through floodwater up to 30cm depth, to provide a more nuanced being and realistic approach to flooding disruption.

3.1.5 Methods

This section describes the methods used to model floods and flood risk assessments.

3.1.5.1 Hydrological and hydraulic models

Traditionally, flood inundation maps are created by driving a hydraulic model with available streamflow data or the output from a hydrological model. However the entire stream network does not have gauges meaning there is not quality data available for all sites which need modelling in the UK (Rajib et al., 2020). To address this lack of information, hydrological models are used alongside hydraulic models. Hydrological models simulate streamflow for the upstream boundary conditions and the hydraulic models simulate the flood event extent to produce, depth and velocity results to create flood inundation maps (Rajib et al., 2020). Both these types of models are common tools used for a wide variety of purposes as they allow multiple scenarios to be run and estimate the likely effect of different size floods on the area of interest (Balica et al., 2013). Improvements in computing power and availability of high resolution data (Dottori et al., 2013) means the application of flood models has been a mainstay of flood risk work over the last decade. In this section the theory behind hydrological and hydraulic models will be reviewed to provide context to the modelling work described in Chapter 4.

Hydrological models are used to estimate how rainfall will result in runoff over a catchment, through sets of equations and parameters which describe the watershed and how water moves (Devia et al., 2015). At a minimum, rainfall data and drainage area are required, with watershed characteristics like soil properties, land use, topography and soil moisture sometimes required (Devia et al., 2015). There are some empirical models which use measured data for each parameter, however due to the difficulties of having data for all watersheds of interest, most models are conceptual and instead rely on equations to replicate physical processes. Hydrological models can be classified as lumped, semi-lumped or distributed (Devia et al., 2015), referring to how the watershed is considered in the model. Lumped models take the catchment as homogenous whilst distributed models have cells in which the rainfall in and runoff out is calculated for a heterogenous representation of an area. Hydrological models can provide the input to a hydraulic model by generating a hydrograph or river flow stage curve which can be used as an input boundary condition to the hydraulic flood model (Kjeldsen et al., 2013).

The Revitalised Flood Hydrograph model (ReFH) is a lumped conceptual rainfall-runoff model which is the de facto standard for flood hydrograph estimation in the UK (Kjeldsen et al., 2013). The model consists of three components; a loss model which creates an excess rainfall hyetograph, a unit-hydrograph based model for routing excess rainfall in the catchment and a baseflow model representing the flow in the river before the onset of a storm (Kjeldsen et al., 2013). The ReFH has been designed with the fewest parameters as possible so it can be applied to ungauged catchments across the whole UK (Kjeldsen et al., 2013). The Flood Estimation Handbook (FEH) web service provides access to a national database of catchment descriptors for the whole UK. These descriptors include rainfall-depth-frequency data for any catchment larger than $0.5km^2$ to enable hydrological modelling (Kjeldsen et al., 2006; UK Centre for Ecology & Hydrology, 2020). The output of the ReFH is flow hydrographs at the catchment outlet, which can then form the input into a more detailed hydraulic model representing the catchment with local topography (Kjeldsen et al., 2013). This is the hydrological model used in this thesis to provide boundary data for the hydraulic model (See Chapter 4 for details of its application).

Hydraulic models are designed to produce representations of the depth, velocity and area covered by flood water as a result of the input hydrology. There are two main divisions in how hydraulic models represent the channels and floodplains of the study area, 1D and 2D models. 1D models calculate flow in the downstream direction only through one dimensional equations, for example the St. Venant equations (Gugat & Leugering, 2003). This assumption of downstream flow works well for river channels and the channel is modelled through cross sections where channel depth and width are known (Balica et al., 2013). The 1D flow equations calculate water flow through the cross-sections based on topography and river flow data. Having calculations at cross sections rather than at every part of the channel increases computational efficiency, so it is the preferred method for modelling fluvial flooding over large areas, particularly in homogenous catchments (Castellarin et al., 2009; Leandro et al., 2009). 2D models calculate flow both in the downstream and perpendicular direction, often on a grid matrix, and use two dimensional hydraulic equations. 2D models provide better data for flood flow, depth and extent in places with more complicated flow paths like dense urban areas (Gallegos et al., 2009). The use of 2D equations in each grid cell does increase computational time and power which impedes its use over large model domains. It has therefore become common to combine the two model types to have a coupled 1D/2D model, where a 1D model with cross sections is used for the river channel and a 2D model for the floodplain (Leandro et al., 2011). Overland flow is not the only consideration in urban areas, and some flood models have coupled sewer models to the 2D flood model to capture the dynamics between surface water and sewer systems (Leandro et al., 2011). For modelling pluvial flooding, a 2D model would be required to capture the dynamics of water around a complex urban environment.

Hydraulic flood models can also be categorised by how they represent flow. Steady state models represent changes in flow spatially but keep flow temporally constant and are useful when just the maximum extent of floodwater is required for hazard assessments. Unsteady state models vary flow both spatially and temporally and are more accurate when modelling the effect of flood water on flood defences to test the likelihood of failure (Büchele et al., 2006). For pluvial flooding, there are not often local flood defences designed specifically for the prevention of pluvial flooding, rather drainage mechanism are installed, so a steady state model is sufficient to understand pluvial risk over a large urban area during a rainstorm (Leandro et al., 2011).

Many flood models exist and the type of model which is suitable depends on the size and characteristics of the study area, the purpose of the modelling work and the available software and computational power. A selection of hydraulic flood models, including the one used in this thesis, are compared in Table 3.1. For large areas, “low-complexity” inundation mapping tools which rely on topography, roughness parameters and simple hydraulic equations have been used frequently due to lower computational requirements (Rajib et al., 2020) However, they are less accurate than coupled 1D/2D models, so it is a trade-off of computational time compared to accuracy (Rajib et al., 2020; Sosa et al., 2020). The complexity of urban environments means they have always been a challenge to model (Ghimire et al., 2013) as surface flow is strongly influenced by topography. Therefore, the availability of accurate Digital Elevation Models (DEMs) or Digital Terrain Models (DTMs) at appropriate resolutions are central to accurately simulating flooding (Noh et al., 2018; Ghimire et al., 2013).

As with all models, hydraulic and hydrological models are a simplified representation of reality and there is uncertainty associated with their outputs (Montanari and Di Baldassarre, 2013). The main source of uncertainty is flow data (Dottori et al., 2013) as it is the extreme events, which are hardest to measure, that are of most interest. The flooding in urban areas from rainfall is often caused by intense, short rainfall events causing field observations to often be non-existent or difficult to collect, meaning model calibration is a challenge (Leandro et al., 2011). There is also typically only a small set of floods which have been recorded, which means simulations are required to fill in these data gap (Emanuelsson et al., 2014).

Table 3.1 Comparison table of some well-known flood models, including the model applied in this thesis (Flowroute – i™). Adapted from a Table in the authors MSc dissertation (Rawlings, 2016)

Model	Strengths	Weaknesses
HEC-RAS Developer: United States Army Corp of Engineers 1D hydrologic modelling for channels and flood banks	- Free - Easy to set up and use - Output compatible with GIS software	- Does not model 2D domains well - No technical support provided by the manufacturer - Does not include urban drainage systems like sewers explicitly
LISFLOOD FP Developer: Bates and De Roo, Bristol University 2D model for hydrodynamic modelling of floodplain inundation (Bates et al., 2010)	- Free software for academics - Coupled 1D/2D model - Outputs include depth, elevation and velocity compatible with GIS software - Can be applied to large areas (Rajib et al., 2020)	- No support except via online tutorial exercises - Does not include urban drainage systems like sewers explicitly
Flood Modeller Developer: Jacobs 1D and 2D solvers for 1D/2D flood models of channels and floodplains	- Free version allows for small modelling domains to be used. - Can run 1D, 2D and linked models - Provides depth, extent, velocity and flow outputs which can be exported to GIS software - Support available via online tutorials and forums	- Unclear error message - Larger modelling domains require 'pro' license - Support offered by a software support email with a pro license, not available for free users
Flowroute-i™ Developer: Ambiental Risk Analytics A steady state 2D flood model	- Guidance in using the software available - Suitable spatial resolution for pluvial flooding (5m) over large urban areas - A 2D model which can model fluvial, pluvial and tidal flooding - Provides depth and velocity outputs - Has been validated by the EA - Accounts for drainage, flood defences, ground roughness and infiltration - Results are compatible with GIS platforms	- Limited application to academic settings - Does not include urban drainage systems like sewers explicitly

3.1.5.2 Flood risk assessments

A flood risk assessment (FRA) is an analysis which reviews a development, urban area or region against the probability of a flood event occurring and the consequences for the site itself and the surrounding areas (Emanuelsson et al., 2014; Ambiental Environmental Assessment, 2020). All types of flooding can be considered in an FRA. Often academic papers will focus on one type (for example pluvial) whilst the FRA required for planning development sites in the UK require consideration of groundwater, fluvial, pluvial, tidal and sewer sources (Ambiental Environmental Assessment, 2020). FRAs can be used to enable decisions on whether developments should be built (Ambiental Environmental Assessment, 2020), how to design flood defences to protect existing infrastructure, test which strategies are the best performing at keeping people safe and alleviating flood risk (Woodward et al., 2014). Insurance companies also conduct risk assessment models to understand the value and risk of their portfolio for investment decisions (Hammond et al., 2015). Risk management often involves weighing up the benefits and costs of action to mitigate significant consequences (Emanuelsson et al., 2014), as some action may not be required if the cost of action is outweighed by the consequence of a flood. The consequences of a flood are often quantified in monetary terms, but the less easily quantifiable social and environmental impacts need consideration as well (Emanuelsson et al., 2014).

Two categories of flood risk assessments are discussed here. Firstly, academic papers where the flood risk across a town, city or region is assessed, often in economic terms. Secondly, there are very localised flood risk assessments which are required by local authorities as part of the planning process in the UK for any development which falls within modelled flood zones.

When assessing flood risk across an area, risk-based decision making relies upon estimates of risk now and in the future, in the context of a range of flood risk management options (Blanc et al., 2012). Within the risk assessments, the risks and consequences from flooding can be categorised in many different ways. There are short term versus long term risks, for example shorter term risk include the risk to life in the immediate flood event, property and infrastructure damage whilst economic impacts tend to be longer term (Hammond et al., 2015). The risk can also be classed as tangible, those which are quantified, for example economic loss from damaged properties, or intangible, for example long-term mental health problems (Hammond et al., 2015). Finally, there are direct and indirect risks, direct risks arising from contact with the floodwater and its immediate vicinity and indirect risks to the wider area which is not itself flooded (Hammond et al., 2015). How risk is defined, and which risks are of interest, varies between studies which leads to a variety of methods of assessing flood risk.

There are some broad commonalities in the approaches to assess flood risk, as outlined in Blanc et al., (2012), Hammond et al., (2015) and Emanuelsson et al., (2014) . Firstly, the flood scenarios of interest are determined through identifying the flood risk level of the study area and then the input rainfall and hydraulic simulations are completed. Secondly, the vulnerability of the study area is measured and suitability criteria for assessing the consequences of the flood are decided. Finally, the flood event is combined with the vulnerability and consequence assessment, which could be through numerical scoring of factors to produce an overall risk score like Emanuelsson et al., (2014). Risk assessment methodologies can also incorporate qualitative measures as well as quantitative. Sperotto et al., (2015) developed a regional risk assessment methodology for Venice through collecting the knowledge of local stakeholders, with workshops and surveys to identify places at risk and the measures which the community felt would be best placed to help.

For quantifying direct economic costs, depth-damage curves are applied to estimate the costs related to building damage and business interruption (Blanc et al., 2012) particularly in insurance models. Depth-damage curves estimate the amount of loss a building type would expect for a given depth of flood water. In the UK, the National Property Dataset allows for the classification of individual properties by age, building type and social class of residents who use them (Hammond et al., 2015), for which depth-damage functions have been generated (Penning-Rowsell et al., 2005). These depth-damage curves are frequently used to ascertain economic risk, however there is uncertainty associated with their results. Depth is not the only factor in the level of damage to a building. Velocity of the water into and around the building is a significant factor in the damage (Kreibich et al., 2009; Hammond et al., 2015) but is not included in the economic risk calculation. Depth-damage curves also do not consider the indirect costs from flooding like interruption to supply chains or the social, health and environmental consequences (Blanc et al., 2012).

Flood risk can be measured as Expected Annual Damage (EAD) in order to balance the expected costs of flooding in a given year against the cost of mitigation measures for stakeholders (Woodward et al., 2014; Olsen et al., 2015). One example is Woodward et al's (2014) paper which designed a risk assessment for the Thames Estuary, UK. EAD was calculated through a model which considered the hydraulic load of defences, the condition of flood defences in combination with realistic flood scenarios. The result was a map which showed the estimated EAD value for each zone, identifying in a quantitative way which areas were at greatest risk to losses due to flooding. Olsen et al., (2015) tested several methods for calculating EAD and found the method itself had a minor impact on the risk analysis results. They applied a GIS-based risk model which incorporated asset classes (for example residential, commercial, roads) with each having a flood depth threshold for when damage starts to occur, and the unit cost of the damage calculated from previous insurance pay-outs for that class. The shift in damage costs as a function of return period was more important, which in their study this threshold was around the 1 in 10 year return period, which they posit could be related to the capacity of the sewerage system (Olsen et al., 2015).

In flood risk literature, there has been great focus on tangible flood damage which can be economically calculated (Hammond et al., 2015) whilst understanding of the intangible impacts has been neglected. The indirect effects on infrastructure networks beyond the direct flood area is one of these. Infrastructure such as telecommunications network, transport networks, power lines, water networks and food production should all be considered in flood risk assessments (Hammond et al., 2015; Conrad et al., 2006), but rarely are due to the difficulties in capturing these indirect effects. Physical and mental health are another intangible impact where the risk to a population is difficult to capture and predict. To combine the intangible health impacts with tangible economic impacts, Hammond et al., (2015) argue that a quantification of health impacts would be needed, but this is difficult to calculate and controversial to attach an economic value to human life. There are examples though of multicriteria frameworks which explore combining all impacts, for example Kubal et al., (2009) Zhou et al., (2012).

Under the UK National Planning Policy Framework and EA guidance, a FRA is required for any developments which are proposed in a flood risk zone (Ambiental Environmental Assessment, 2020). Flood zone maps are generated through simulating flood events of a given magnitude and estimating the area which would be affected and can provide an overview of the flood risk for an area. In the UK, the EA's flood maps have divided the country into zones as an easily accessible way of determining risk for flooding. For fluvial flooding, Zone 1 is land with less than a 1 in 1000 annual probability of a flood (<0.1%), Flood Zone 2 is land with between a 1 in 100 and 1 in 1000 annual probability of flooding (1%-0.1%) in any year and Flood Zone 3 is land with a 1 in 100 or greater annual probability of river flooding (>1%) (Ambiental Environmental Assessment, 2020). For pluvial flooding, a similar planning map is available with very low, low, medium and high-risk categories used as zones (Gov.UK, n.d.).

The advice for FRAs from the UK government outlines the following steps for completing a FRA (DEFRA & EA., 2019):

- Firstly, research the development site through contacting the EA and local planning authorities and check if it is likely to be affected by flooding

- Secondly, complete the risk assessment with details of the site, the flood risk from all sources and the estimated level of defence which is required.
- Finally, submit the flood risk assessment and planning application to the local planning authority.

For pluvial flood management at vulnerable sites, the requirements of the local authorities surface water management plan also need to be considered in the planning application (DEFRA & EA., 2019). Often pluvial flooding is neglected when FRAs are considered, despite it being a large issue in urban areas due to the potential of new developments obstructing overland flow pathways (Ambiental Environmental Assessment, 2020). However, pluvial models with specific depth calculations to ensure flow paths are not obstructed are being requested more frequently (Ambiental Environmental Assessment, 2020). A key part of a FRA for development sites is the consideration of how people are able to enter and leave the site if there is flood water (DEFRA & EA., 2019; Ambiental Environmental Assessment, 2020). Buildings themselves may be secure and not be flooded, but if routes of entry and exit are blocked then the site is not safe for use. This requirement is to only consider access and evacuation of the development site and not any issues of isolation at the local area around the site due to surrounding roads being impassable.

As with all models, there are many uncertainties associated with the methods and outputs of flood risk assessments. A key uncertainty is that in most locations there is not an extensive flood data set of past flood events, therefore hydraulic simulations are required to generate data (Emanuelsson et al., 2014). Whilst these models are calibrated to real flood events, predicting large events where there are few examples or projecting into future floods under climate change, there will always be a degree of uncertainty in the depth, velocity and coverage of a flood event. The scale and precision of flood inundation data relies heavily on very precise local topography data, particularly in dense heterogenous urban areas (Emanuelsson et al., 2014), so access to such data is vital for a reliable flood risk assessment (Ambiental Environmental Assessment, 2020; Emanuelsson et al., 2014). As discussed, intangible impacts or asset dependencies across networks are generally not considered in FRAs (Emanuelsson et al., 2014) meaning they do not capture the complete picture of risk for the area of study.

3.2 Review of Spatiotemporal Population Modelling

As highlighted in Chapter 2, the spatial location of a population in relation to a hazard is often the main factor in the scale of impact (Freire et al., 2013). Therefore, it is imperative to have information on where the target population is and where they travel to at the finest temporal resolution possible. For the vast majority of natural hazard studies, residential population data are drawn from decennial national census data, which do not adequately capture the temporally variant nature of population in their calculation of exposure. Although many national censuses do also record place of work for the economically active population, and in some countries place of higher education for those studying (Shabou et al., 2017), these are seldom used in natural hazard planning. The field of spatiotemporal population modelling has rapidly expanded over the last decade as its importance in various fields including public health, urban planning, hazard risk has been recognised. New advances in technology provide opportunities for inclusion of such data models into natural hazard risk assessments and this thesis provides an attempt at linking these disciplines together. This section reviews 'traditional' population data from the census, spatiotemporal modelling principles and methods as well as the alternative 'big data' sets and microsimulation models.

3.2.1 Census data

Census data have been the standard for population research (Aubrecht et al., 2014), providing a cross-sectional snap shot of population location at set temporal intervals (Smith et al., 2014). In the UK, a census has traditionally been conducted decennially, with the last census taken in 2011. The census collects social, economic and health data which are applied in a wide range of purposes. However, processing vast amounts of data takes a couple of years before release for public use, meaning the data are several years old when made available for research. There are five nested levels for census geography in England and Wales, relevant to flood risk management in England (Figure 3.1). Full unit postcode level is the finest resolution of census data, however for data protection purposes there are only very limited head counts by gender data available at this level. Output Area (OA) is the finest resolution available for data. To ensure confidentiality, sometimes data are suppressed at the OA level.

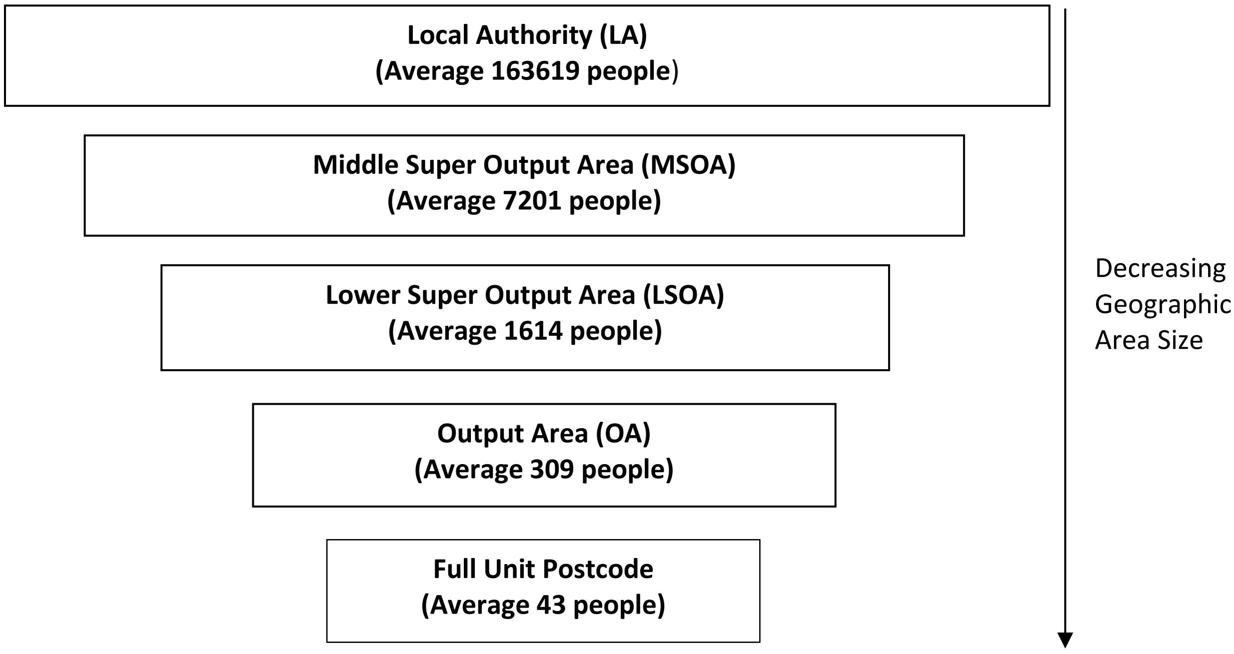


Figure 3.1 Diagram of the census geography hierarchy applied in 2011. (Office for National Statistics, 2016)

Whilst censuses have provided an invaluable resource to many researchers, there are both spatial and temporal issues with using census units as measures of population density. Spatially, the administrative areas used to aggregate populations are not of equal geographic area and boundaries can change between censuses. This causes complications as the choice of boundaries can have a greater effect on the output analysis than the actual variation in the study phenomenon (Smith, 2015). This is referred to as the 'Modifiable Areal Unit Problem (MAUP)' (Openshaw & Taylor, 1979; Openshaw, 1984) and it is why it is important to use appropriate area sizes in analysis (Harland & Stillwell, 2010). Traditional choropleth mapping is subject to the MAUP (Holt et al., 2004).

Temporal issues also arise because, as noted previously, the spatial distribution of population fluctuates at a temporal resolution of minutes to days (Aubrecht et al., 2014; Bian and Wilmot, 2015) and census data are collected decennially. Census data are often thought of as the 'night-time' residential population of an area (Aubrecht et al., 2013a; Smith, 2015), as the survey is sent to everybody's home address, which is not where the majority of the population spend their day. Whilst the census does provide information on workplace zones and if there are students living away from home, it does not easily provide a record of where every person spends their time. Consequently the data can only provide limited information on spatial and temporal changes in the population at the resolution which matters to risk managers and spatial planning (Renner et al., 2018). In the UK, Mid-Year Estimates (MYEs) for population age are produced in order to update the census data annually; however, the finest resolution available is Lower Super Output

Area (LSOA) meaning spatial resolution is lost. This only provides an estimate of the night-time population once a year, a temporal resolution which is still too coarse for detailed risk analysis. Furthermore, MYEs only provide a population estimate by gender and age and do not include any other population characteristics. In light of this, there has been increasing demand for population data which have a finer temporal scale and are independent of administrative areas (Aubrecht et al., 2013b).

3.2.2 Spatiotemporal population data

Traditionally, population data are often presented in the visual form of a choropleth map based on census zones. Holt et al., (2004) state that population mapping has two purposes; to cartographically portray population density within a study area and as a basis for deriving quantitative population estimates. On the surface, choropleth maps do fulfil these requirements. However, choropleth maps create the illusion of sudden changes across boundaries and uniform population densities, which is untrue, especially in rural areas where large parts do not have any population present (Tapp, 2010). Hence they do not provide a 'realistic' depiction of population location (Wright, 1936). Therefore, alternative ways to display population data have been created which rely on interpolation techniques as they improve the representation compared to census zones. Spatial interpolation is the process of estimating the value of properties at unsampled sites by using existing data points from the area (Heywood et al., 2011). There are many methods of interpolation and they can be broadly split into two groups; point-based interpolation and areal interpolation.

Point-based interpolation takes point data and produces a gridded data set (Hengl et al., 2007) without the restriction of census boundaries. There are many point-based interpolation methods and one of the most popular is kernel density estimation (Danese et al., 2008). This interpolation method involves a fine grid being placed over the point data (Borruso, 2005) and a kernel is assigned to each data point. An algorithm calculates the density of the kernels through a bandwidth (or radius) for the observation point (Fotheringham et al., 2000; Anderson, 2009). A surface representing the variation of point density (typically census zone centroids with population counts as weights) across the area is created (Borruso, 2005), which can be used to show 'hotspots' in the spatial pattern of variables (Anderson, 2009; Gerber, 2014). In order to account for edge effects, the analysis needs to be extended one kernel beyond the study area (Bracken & Martin, 1989). Applications of kernel density analysis include population modelling (Martin et al., 2015), crime analysis (Gerber, 2014), spatial epidemiology (Gatrell et al., 1996), population distribution, social segregation (O'Sullivan & Wong, 2007), animal ecology (Hemson et al., 2005) and urban modelling (Danese et al., 2008). An advantage of kernel density estimation methods is that they are ‘volume preserving’, which means the number of people in the source zones translates exactly to the target zone. However, as these are locally specific models the results are not transferable to other geographic areas (Gerber, 2014) which does restrict the wider applicability of studies. Also, zone centroid location is critical to the analysis as irregular centroid position can lead to bias (Wu et al., 2005; Tapp, 2010). Other examples of point-based interpolation techniques include kriging, spline functions and inverse distance weighting (Harman et al., 2016).

Interpolation methods have primarily been focussed on spatial interpolation, rather than temporal interpolation, although time is a key concept in creating accurate population surfaces. Investigation into spatiotemporal dimensions of interpolation are not just restricted to population modelling, it is also a relevant research issue in environmental sciences, climate prediction, and hydrology. ‘Time-geography’ is the key underlying theory for temporal interpolation methods. ‘Time-Geography’ is a branch of human geography, first defined by Torsten Hagerstrand, and was prominent in the 1970s. The theory states that time is a resource for social life and a key measure of productivity. It is complementary to space, as we move through space over time, so both should be considered in analyses as the challenges and characteristics of each need to be addressed simultaneously (Bogaert, 1996; Pred, 1977). As part of the theory, constraints to the spatial-temporal movement of people are defined (Pred, 1977):

1. Capability constraints such as how much free time an individual has for an activity; and the limitation of distance they can cover in a given time on transport available
2. Authority constraints: not all space can be occupied by anyone e.g. private property
3. Individuals cannot do multiple activities at once so must not be duplicated in models
4. Time is taken moving between activities, so needs considering in any analysis of how individuals spend their days

Whilst the importance of including time in analyses is clear, actually including it in population models has proved challenging as much geographical software was designed primarily for spatial analysis. Although there has been development of specific spatiotemporal GIS packages, for example ‘spacetime clusters’ in ArcGIS Pro (ESRI, 2020f), R packages for handling and analysing spatiotemporal data (Pebesma et al., 2020) and STEMgis, a temporal GIS system (Discovery Software, 2020). Additionally, traditional raster and vector data structures limit the way temporal data can be handled, resulting in ‘snap shots’ of change between time periods. To address this problem, geostatistics can be extended to provide a joint tool for space-time analysis (Kyriakidis & Journel, 1999). Space-time kriging is one such tool (Bogaert, 1996). It yields optimal predictions for any point in a space-time domain and the result of the interpolation can be displayed as an animation or series of maps (Heuvelink et al., 2016). Other techniques include conditional autoregressive regression modelling, which takes into account spatial and temporal autocorrelation in data. So, these methods can take an interpolated surface at specific time intervals and compare changes between them (Kyriakidis & Journel, 1999), which is an improvement on interpolation from a single census time period. But such interpolated outputs depict gradual long-term change, not the fine scale of population movement in real-time.

The recognition of the limitations of census data and the availability of other administrative and remotely sensed data sources has led to an increased demand in spatiotemporal population modelling (Aubrecht et al., 2014). The principle behind spatiotemporal population data modelling is to mimic the flows of people as they conduct their daily routines to produce data of population counts for different times in a day or year in an area. These models assume population groups with similar characteristics share similar movements, determined by the location of economic, social and leisure facilities (Batista e Silva et al., 2018), which can be used to assign daily temporal routines (Bian & Wilmot, 2015). Bhaduri et al., (2007, p.106) provide useful definitions for identifying the groups within night-time and daytime populations, as shown in the Equation 3.1 and Equation 3.2.

Night time = Night time Residential Population + Night time workers + Tourists + Business Travelers (+ Immobile Population) Equation 3.1

Daytime Population = Workers + School Children +Tourists + Business Travelers + Residual Night time Population (+ Immobile Population) Equation 3.2

Within recent literature, the groups identified in Equation 3.1 and Equation 3.2 are commonly used, with night-time residents, workers, students and daytime residents as staples *(*Aubrecht et al., 2014; Bian and Wilmot, 2015; Freire et al., 2013) and tourists and shoppers recognised as important sub-groups with high mobility which are more challenging to model (Bian and Wilmot, 2015; Smith et al., 2015). In relation to hazard preparation, tourists may not have knowledge about local hazards to be able to react quickly to a warning (Cahyanto et al., 2014) so are an important group to account for in a risk assessment.

Such data are often gridded as this provides a more meaningful representation of settlements than choropleth maps of census zones (Martin et al., 2009). It also offers a consistent sampling grid over time, minimising the effect of the MAUP (Martin et al., 2009). Examples of spatiotemporal population distribution projects include Gridded Population of the World (Deichmann et al., 2001) , Global Rural Urban Mapping Project (Balk et al., 2006), LandScan USA project (Bhaduri et al., 2007), DynaPop (Freire et al., 2015) WorldPop (Tatem, 2017) and Population 24/7 (Martin et al., 2009, Martin et al., 2015). More details are given in Section 1.2.4.1.

3.2.3 New forms of data

Future advances in spatiotemporal modelling of population will be from incorporating new non-conventional data sources into spatiotemporal models (Batista e Silva et al., 2018). Types of non-conventional data for population in travel include mobile phone data, from GPS and smartphone applications which track movement (Malleson et al., 2018), social media data, from platforms like Twitter, Flickr, foursquare (Malleson & Andresen, 2015) and automated 'big data' like footfall count, automatic number plate recognition and travel smart cards (Heppenstall et al., 2016; Crols & Malleson, 2019). This section will briefly describe the main new forms of data that are emerging and examples of their application to spatiotemporal population modelling.

Mobile phones have become an increasingly popular source of population data as they allow movement to be followed in near-real time. Globally, 96% of people have a mobile phone (Deville et al., 2014), a figure which is only increasing, providing a very large data set. These data are especially valuable in countries where reliable census data at good spatial resolution is absent (Deville et al., 2014). Mobile phones provide the location of an individual via one of two ways: either through use of the cellular network which identifies the cellular zone the handset is in, or through the handset's GPS chip. For cellular zones, phones are recognised when they receive/place a call or text message or connect to the internet (Calabrese et al., 2011), so when not in use the phone's location is not known, meaning some movement will inevitably not be captured. Also, these data are anonymised to the number of people within the cellular zone so cannot be used for individual modelling (Deville et al., 2014). Most modern phones are GPS enabled, which can provide a much greater spatial and temporal resolution on an individual's location. For example, in Herrera et al.'s (2010) study the phone's GPS updated its location every 3 seconds. GPS does not require the phone to be in use, just turned on, so a greater amount of the population's movement can be tracked. Herrera et al., (2010) did a large scale experiment in the San Francisco Bay area and found reliable data on traffic conditions could be gained when as little as 5% of the cars on a road had GPS enabled phones. A decade later, this number of cars travelling with smart phones is a lot higher, with widespread use of navigation apps for real-time directions. Therefore, mobile phone data can be a powerful tool for applications including disaster monitoring, disease mapping (Tatem et al., 2014), and road travel behaviours (Deville et al., 2014).

Smartphones and the applications which are installed on them are another source of spatiotemporal data on the population. Malleson et al., (2018) gathered data generated from a smartphone application in order to analyse the daily routes of pedestrians in the Greater Boston area, USA. The app from which the movement data were generated was downloaded by the user for the purpose of allowing the user to track their daily routine, so there was consent for the data to be collected. Malleson et al., (2018) mapped the routes taken using an algorithm to map them onto roads, with 94% of routes successfully matched to a road path. The data generated from the smartphone app was of sufficient quality and spatiotemporal resolution to examine pedestrian asymmetry. Travellers' paths were about 20% longer than the direct shortest route between places and people changed their routes between origin and destination 15% of the time. This study is one example of how the widespread use of smartphones is producing new data sets which can provide greater insight to daily population movement and the routes they take.

Mobile phone GPS data are collated by companies like Google to generate estimates of the number of people at a facility for a given time of day. The 'popular times' data appears underneath the details of many businesses and attractions on Google and show the relative busyness of a place at different times of day over the course of a week. This could be extremely useful data for retail and tourist spatiotemporal models; however, it is not currently available through the API due to its commercial value.

The main advantage of mobile phone data over census data are the increase in temporal resolution (Deville et al., 2014). The spatial resolution is comparable to the spatial resolution which can be achieved through downscaling census data sets, but mobile phone data opens the potential for real-time monitoring of traffic for better management or locating population flows in emergencies (Deville et al., 2014). Studies have shown that mobile phone data can capture the differences between weekday/weekend movements which again is an improvement on census data (Calabrese et al., 2011; Herrera et al., 2010). Secondly, a much larger sample of the population is achieved than through travel surveys, a full order of magnitude in Herrera et al.'s (2010) survey, as people are increasingly leaving larger 'digital footprints' (Calabrese et al., 2011). The data from smartphones via location tracking apps can also be much more accurate than travel diaries (Malleson et al., 2018) as they do not require the participant to remember to complete the diary.

However, there are some substantial issues with mobile phone derived data which still need resolving. Firstly, the 'digital divide' means it cannot be considered a fully representative data set of population demographics, with potentially serious biases towards the younger generation (Tatem et al., 2014; Malleson et al., 2018). Mobile phone uptake has shown to be low amongst older people and those who are the most marginalised in society (Deville et al., 2014), meaning their representation is lower than from census data sets. Also, the youngest children are unlikely to have a phone of their own, so these data may not be suitable for tracking travel to schools without assumption on how many children are in travel per adult smartphone reading. Therefore, there is a divide between people who have and do not have a phone and who is represented in any data. Additionally, GPS location data are only recorded whilst a phone is turned on, meaning users could choose to switch their phones off and would not appear in any data.

Secondly, there are data privacy and ethical concerns with the use of individual mobile phone data. Currently, the use of these data involves a rigorous anonymisation process before it can be used for academic research, so individual movements cannot be identified (Deville et al., 2014; Calabrese et al., 2011). For example, Herrera et al., (2010) had to 'degrade' their data as the temporal and spatial resolution was too fine and identified individuals, when their study was to provide an estimate of overall traffic flow on a road. Thirdly, the market share of the mobile phone data provider can limit the sample size (Calabrese et al., 2011). Mobile phone plans may limit how often someone uses their phone if they do not have a large phone call or data allowance, meaning not all movement may be logged (Calabrese et al., 2011). Using data from smartphone apps which work across a variety or networks and phones is one way to resolve this issue and is more feasible now than in 2011 when Calabrese et al., (2011) did their work. Finally, mobile phone data does not contain the demographic data that the census does, unless it is connected to other data sets (Herrera et al., 2010). So, whilst it is useful for looking at travel movement as a whole, as yet it cannot indicate how different sub-groups of the population move or why someone has undertaken a journey in the same way a travel survey does (Longley et al., 2015). There are some examples where survey data has been combined with mobile phone data to generate travel patterns for groups of the population (for example Arai et al., 2014). Future data may facilitate more of this type of analysis.

Following on from mobile phones, social media and the data the population publish on these platforms has drawn great interest in recent years as a potential proxy for population movement. There are a number of social media sites where users publish location data which can be viewed publicly or accessed via an Application Programming Interfaces (APIs), including Twitter, foursquare and Flickr (Malleson & Andresen, 2015). In recent years, Twitter data in particular has been a focus of spatiotemporal population studies. Geotagged tweets have been applied as proxies for movement to overcome data restrictions of mobile phone networks, as the data are available via an API. Twitter has an option where users can geotag the location they send the tweet from, based on GPS, and the accuracy of this function is quoted as 10-20m in various sources (Jurdak et al., 2015).

Jurdak et al., (2015) used Twitter data to study international migration, finding they could divide their twitter users into two groups, those who tweeted from the same place predictably and then those who were random. So, they argued that these data can be useful for tracking those who have migrated or move around a lot, which would not been seen at the same temporal resolution in other data sets. They attempted to introduce social characteristics to the data by using facial recognition software to classify people based on age and gender in order to map migration patterns. This is not a substitute for the wealth of data a census can provide but does show how more than just location can be estimated from twitter data. Longley et al., (2015) used the forenames and surnames of twitter users to define age, gender and ethnicity by cross-referencing twitter handles against detailed name data bases. As with facial recognition, Longley et al., (2015) recognised difficulties with this approach as the names or photos used may not be the account holders' own, or someone may have a name which the profiling database does not consider 'typical' for their age, gender or ethnicity.

These studies are interesting in their exploration of new data for modelling people's travel, however there are still many restrictions on the amount of data available and assumptions in these data which mean they cannot replace 'traditional' census and administrative data sets. Twitter data has also been used to assess the crime rate of a city, including the ambient non-residential population, rather than basing crime rate solely on the residential census population. Malleson & Andresen (2015) used Twitter data of messages posted in Leeds with GPS coordinates and crime data for 'street crimes' to build a model which searched for spatiotemporal clusters of crime in Leeds. They found clusters for specific space and time points, for example the area around the University between 21:00 and 02:00 on a Saturday/Sunday, which registered significant levels of crime which would not likely have been picked up in traditional crime rate data based on census-derived residential data. This study demonstrates the value of incorporating 'new' social media data with 'traditional' data sources to enhance the spatiotemporal understanding of an issue.

The advantage of geotagged tweets is that the data are more accessible than mobile phone data as it can be downloaded via Twitters API with relative ease. However, it is estimated that 1% of tweets are geotagged, and the twitter API allows someone to download up to 1% of all tweets which match their search criteria (Singleton et al., 2011). Thus, any data used in research is a small subsection of the population on Twitter. Twitter is also predominantly used by younger 'tech savvy' people, which Longley (2015) found was not representative when compared to the census data for London. However, Jurdak et al., (2015) argue that it is younger adults who are most likely to migrate so twitter could still provide useful data for this application. Currently, there are no reliable means of extrapolating the data up to the scale of the population at large, so these data are only applicable for the area and people surveyed. Furthermore, these data rely on users tweeting whilst conducting activities, rather than from their homes. So, there is the chance that the activity being tweeted about is not where the geotagged location indicates. A main issue with social media data are that there is little data to validate it against in order to know how it represents the day-time population (Malleson & Andresen, 2015). Where available, social media data could be compared to footfall data to see how much movement is captured as Crols & Malleson (2019), however there are difficulties in this as well as it is not known what proportion of the footfall data represents people with smartphones or Twitter accounts.

3.2.4 Methods

This section gives a brief overview of the main methods used in spatiotemporal mapping and modelling within the literature.

3.2.4.1 Available spatiotemporal models

Two approaches are emerging in spatiotemporal population methodologies, those that extend mapping technologies within a time-enabled framework, and those that attempt to track population movement through ‘big data’ (Martin et al., 2015). This section will focus on the former approach as this is the one applied in this thesis. Over the last two decades, gridded population models have been developed with increasing sophistication. The first such model output was the Global Population of the World (GWP) project and the Global Rural Urban Mapping Project (GRUMP) in 1990, and now there are several different models from projects across the world (see Table 3.2). These models have similar aims, however they utilise different interpolation methods which means they are appropriate in different situations (Galway et al., 2012). In this thesis, SurfaceBuilder247 is used as it is the only spatiotemporal population model available for England at a fine spatial and temporal resolution (Martin et al., 2015). The National Population Database (Health and Safety Laboratory, 2020) is a source of spatiotemporal population data for England, however it is not freely available to the public and is only available in night/day time slices, rather than specific hours and day types.

Table 3.2 Available gridded population models and projects

Project name	Year	Source	Geographical Scope and Resolution
Global Population of the World (GWP)	1990, 1995, 2000, 2005, 2010, 2015	Socioeconomic Data and Applications Center (SEDAC)	Global, 5km grid resolution
Global Rural Urban Mapping Project (GRUMP)	1990, 1995,2000	Socioeconomic Data and Applications Center (SEDAC)	Global, 1km grid resolution
Landscan USA	2000-present	Oak Ridge National Laboratory	Global, 1km grid resolution
DynaPop	2014-present	(Aubrecht et al., 2014)	Up to 100m grid spatial resolution, up to hourly temporal resolution.
WorldPop	2013-present	WorldPop, University of Southampton	South America, Africa, Asia. Up to 100m grid resolution.
Population 24/7	2009-present	(Martin et al., 2009)	England and Wales population model, can be run at up to 100m grid cell resolution (depending on data sources)
National Population Database	2014-present	Health and Safety Laboratory	GIS tool for Great Britain, 100m point resolution
ENhancing ACTivity and population mapping (ENACT)	2016-17	EU Science Hub	Europe. 100m grid cell resolution.
Population Density Maps	2016-present	Facebook, in partnership with the Centre for International Earth Science Information Network at Colombia University	30m grid cell resolution

Applied as part of an ESRC funded project (UK Research and Innovation, n.d.), SurfaceBuilder247 is a gridded spatiotemporal population model available to academics (Martin, 2011). This modelling framework can be applied to a range of applications which require spatiotemporal population data, including emergency planning for natural hazards (Martin et al., 2015), hence its suitability to this PhD research project. The model uses census data, administrative data sources and government travel data. Briefly, SurfaceBuilder247 determines where the origins and destinations of a synthetic population are, and then for the specific time of day chosen calculates the number of each population sub-group in every grid cell, either residing there at an origin, destination or in travel on their way to an origin/destination. The interpolation method used within the SurfaceBuilder247 model is based on an adaptive kernel density approach developed by Martin (1996). It takes population-weighted centroids of small areas and redistributes them onto a regular grid (Martin et al., 2009). The width of redistribution kernel is adjusted according to typical areal extent zones in that locality and a distance decay function is used in the redistribution (Martin et al., 2009). The Surfacebuilder247 framework presents a volume-preserving model, so the number of people present in the census origin input file is the same as in the output grid (Martin et al., 2015). It does not show the movement of individuals between places, as confidentiality restrictions on census data mean the initial inputs are aggregated beyond the individual level. Rather, the changing totals in grid cells of population groups are seen when each time slice is compared. More detail on how this model works is provided in Section 4.2.2.

3.2.4.2 Microsimulation models

An alternative method to aggregate spatiotemporal population models is to apply a microsimulation model to investigate spatiotemporal population movement. Aggregate models, like spatial interaction models, have the disadvantage of lacking behavioural traits for individuals who compose the large complex populations of cities and respond to urban environments (Heppenstall et al., 2016). Agent based models (ABMs) represent individuals at multiple spatial and temporal levels who respond to their surroundings, often through behavioural triggers within a sensing radius. This generates complex emergent behaviours as agents have the ability to evolve and make decisions over time and space (Heppenstall et al., 2016). The advantages of ABMs are that they can represent the micro-dynamics of individuals, which cannot be captured as easily in statistical models (Heppenstall et al., 2016). ABMs can represent multiple spatial relationships and have the flexibility of incorporating equations and statistical techniques of other models (Heppenstall et al., 2016). An example of using an ABM to model spatiotemporal population movement is Crols and Malleson (2019) who applied the Simulating Urban Flows ABM to model population groups moving into and around a town centre. Census data for commuters were combined with a time use survey to parameterise the movement rules of agents, with other population groups iteratively added. The accuracy of their model was assessed by comparing it to footfall data in the town centre which tracks how many phones attempt to connect to the WiFi network. Disparities between the datasets were presented to local stakeholders who provided reasons for this by identifying groups which needed to be included in the model to more accurately reflect the footfall data. This paper demonstrates how the combination of traditional and new data sets can be effective in providing higher resolution spatiotemporal models of population movement.

Like all models, there are issues with ABMs. The main issue is that of calibration and validation of ABMs, as one method cannot be applied to all as the processes in an ABM occur at different spatial and temporal scales (Heppenstall et al., 2016). A large amount of data would also be required to validate potentially thousands of agents (Heppenstall et al., 2016), and often there is no suitable validation data set (Crols & Malleson, 2019). The ontology of the model can be validated, for example Crols & Malleson (2019) did this through presenting the results to local stakeholders, however there will always be a deviation from the real world in the model. Finally, ABM models can be very computationally intensive which could limit their application to large geographic areas.

3.3 Review of Transport Planning and Network Analysis

In order to research how flood events affect travel, an understanding of theory in transportation engineering is required. This section focusses on how people's journeys are conceptualised, how road networks are modelled and how the risk of disruption is considered. As this research is specifically focussed on the functioning of road networks, the wider context of network science needs reviewing first as it forms the theoretical basis of the methods within transport modelling.

3.3.1 Introduction to network science

Network science is a broad, interdisciplinary field of scientific research which emerged during the 'quantitative revolution' of the 1960s (Ducruet & Beauguitte, 2014) and gained popularity at the end of the 20th century (Fang et al., 2007; Börner et al., 2007). Network science can be defined as the understanding of the connections between elements, in the physical and social world, and is important in numerous research fields, including biology, social sciences, ecology, physics, computer science and mathematics (Börner et al., 2007; Baggio et al., 2010). Networks can be physical (e.g. road networks) virtual (e.g. the internet) or imagined (e.g. social networks) (Baggio et al., 2010). In all types of network, the ontology to describe the elements are the same. There are 'nodes' which connections are made from, which could be origins and destinations in physical networks, or people in a social network. 'Links', sometimes referred to as 'edges', are what connect nodes, and can be physical links like roads, virtual links on the internet or the relationships between people (Kong et al., 2010; Gross & Yellen, 2014; Demšar et al., 2008).

Graph theory forms the foundation of network science with the concepts and terminology adapted from it (Barnes & Harary, 1983). Graph theory is a branch of mathematics in which 'graphs' are a representation of a network (Demšar et al., 2008; Barnes & Harary, 1983). Graphs can be directed, which means that the direction of travel along the edge is important, or weighted, where a numerical value is assigned to the edges and vertices to denote a variable (Demšar et al., 2008). The interaction of vertices and edges can be assessed in several ways. Shortest path is a common function and these algorithms determine the shortest route between two vertices (Kong et al., 2010). Centrality measures are values which describe how important a vertex is to the structure of the network, a higher centrality value means it has a high impact on other nodes if removed (Demšar et al., 2008), which is useful in calculating the impact of disruption. Finally, betweenness is the extent to which a node is needed to link the rest of the network, nodes with high betweenness values have a large impact on flow across the graph (Demšar et al., 2008). Graphs in mathematical models observe the topological properties of a network; however, in this form do not contain the spatial properties of many real-world networks. Graph theory has been used to create GIS network tools in order to examine networks as spatial phenomena (Demšar et al., 2008).

Networks can be simple and consist of a few nodes and links, however real-life networks are often complex systems where the impacts of disruption are non-linear (Dawson et al., 2011). Whatever the type of network or complexity level, networks have broad properties which remain the same (Baggio et al., 2010). One of these is network topology, the geometric connections of elements in a network which are independent of a spatial coordinate system (Heywood et al., 2011). Connectivity is an ‘essential’ component of networks (Reggiani et al., 2015), as how networks connect and the status of these connections determines the functions which can be performed. Due to the importance of connection in networks, the exploration of the effects of network disruption is key, particularly for high consequence, low probability events (Berdica, 2002).

Transport networks are complex system as they are the interface between physical, technological and social spheres. The layout of road networks is shaped by the social and economic connections between places, and the roads represent the physical manifestation of a societal desire to travel to destinations. Like all networks, when road networks are disrupted this reduces its overall performance and impacts the running of society. Consequently, the study of road network disruption and how to adapt to it is a highly active research area. The rest of this section addresses theories in modelling transport networks specifically and how the issue of network disruption is considered.

3.3.2 Transport network theories

Transport networks include roads, train lines and public transport routes which are along fixed paths, but also travel by boats and aeroplanes, which follow routes not physically marked. Road networks have the same underlying properties as all transport networks, with their own unique features. Transport models are networks which contain properties specific to the transport mode, including speed restrictions, travel times, one-way roads and turn restrictions.

Three distinct classifications of transport model scale can be identified (Alaeddine et al., 2015; Lämmel et al., 2010; Borrmann et al., 2012). Firstly, microscopic models cover the individual behaviours of people and vehicles to simulate interactions and emergent behaviour, making ABMs the tool of choice. This can result in high computational cost which restricts the spatial area it can be applied to. Mesoscopic models allow the tracking of small aggregate groups but they do not account for behaviour and interactions of the agents (Alaeddine et al., 2015; Borrmann et al., 2012). Finally, macroscopic models treat the population as larger aggregate groups (for example travel between cities in a region) and again do not include behavioural interactions. The size and type of model applied depends on the purpose of the research, the scale of network and the available computing power.

Traffic and congestion are a feature of road networks, something not present in all types of networks (Sohn, 2006). The impact of individual behaviour also has a large impact on the system, as individual decisions interact to form system wide features e.g. the morning travel peak formed by the structure of the typical working day. This decentralised decision making (Nagurney, 2011) adds challenge to constructing transport models as users of the system do not always make logical decisions and it is impossible to account for all individual decisions. From these principles, a few key conceptual models are used within transport planning and form the basis of computing workflows. Two of these models are described in the next subsection.

3.3.2.1 Conceptual transport models

First conceived in the 1960s, the ‘four stage transport model’ approaches transport planning from a ‘person-perspective’ (McNally, 2007; Ortuzar & Willumsen, 2011; Jones, 2012). The four stages represent the broad stages of an individual’s decision when planning travel (Figure 3.2). Trip generation is determining what journey needs to be undertaken and by whom. In transport planning this has largely been done through travel survey questionnaires to gain an understanding of how often different groups of people travel daily and weekly (Jones, 2012). The next stage is trip distribution, which assigns a destination to these potential journeys. If survey data exists then this can be used, if not, spatial interaction models have been used to determine the likely flows between origins and destinations. The third step is modal split, representing the decision concerning mode of transport (e.g. car, bus, bike or foot) to reach the desired destination. Finally, trip assignment is where a route needs to be selected based on the destination and mode of transport used. Often this will be the one which is calculated to be the quickest or most direct.

This model has largely remained unchanged since it was created, its simplicity providing flexibility in application to a wide range of scenarios. Like all models, the four-stage model has flaws and situations where it is not suitable. First, it is not easily applicable to journeys with multiple destinations as it was based on the concept of a singular origin and destination pair (Jones, 2012). Secondly, it does not include human behaviour which can have a large impact on the destinations, modes and routes taken (Jones, 2012). People do not always choose to take the shortest logical route. However, the four stage model remains a good base for transport models and is still widely used (Ortuzar & Willumsen, 2011; Jones, 2012).

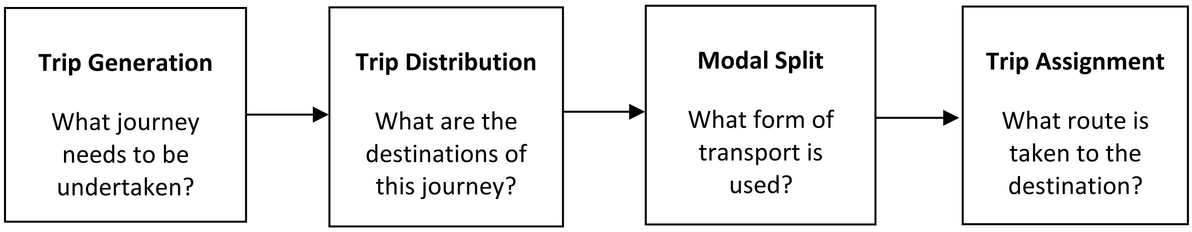


Figure 3.2 The Four-Stage Transport Model

Emerging in the 1980s, activity-based models use the decisions made from activity choice as the basis for transport modelling (Wang & Cheng, 2001), considering measures of attractiveness and impedance (Bocarejo & Oviedo, 2012). Activity based models can include the demand for services, which Reggiani et al., (2015), among others, attest is the basis of travel, as people rarely undertake journeys for the sake of travelling.

3.3.3 Key Concepts for transport network disruption

There are key transport network concepts in relation to transport network disruption which need defining, especially as they often overlap and are measured in similar ways. This section covers the network concepts as well as other key terms found within transport literature.

Society works when its networks are functional. Road networks are the lifelines of society (Chen et al., 2007; Platt 1995; Nyberg & Johansson 2013), hence, even minor disruptions can have negative effects. Disruptions can be caused by natural events, accidents, planned engineering works or terrorist attacks, and each type of disruption has its own warning period and adaptive measures. There has been a rise in the body of literature assessing the impacts of weather on travel (Liu et al., 2014a) as the increased frequency of extreme weather events has a high likelihood of causing significant future problems (Mitsakis et al., 2014). Society cannot afford absolute protection, so identifying what nodes and links are critical to the normal functioning of the network and prioritising this infrastructure is essential (Freiria et al., 2015).

Critical network links can be defined as those which when unavailable have the greatest impact on network function. There is not one method which is used to define critical links (Cats & Jenelius, 2015) as it largely depends on the researcher definition of 'critical', and the linked concepts of vulnerability and accessibility discussed later in section. Identifying critical links is important in all network systems in order to identify where resources for maintenance and protection should be prioritised to ensure the continuation of social and economic functions (Sohn, 2006; Jaroszweski et al., 2015). When considering transport networks, they are not aspatial and links which may not appear critical through the network scan methods may in fact be critical to local populations. A social science perspective of who uses the network is also important. This is particularly true in disruption events.

The origins of network disruption methods are from graph theory. There are a set of methods which can be grouped as topological analyses (Khademi et al., 2015), as they are primarily aspatial methods based on graph theory and do not consider the geography of a disruption (Pregnolato et al., 2016b). The basic method is to remove one link at a time to systematically examine the impact on the whole networks function from their removal (Sohn, 2006). Eliminating singular links and testing the impact is a long process (Chandra & Quadrifoglio, 2013) and it does not provide a holistic view of the tolerance of the network as a whole (Duan & Lu, 2014). Hence, other measures have been developed which are quicker or are more targeted in their network attacks (Chandra & Quadrifoglio, 2013). A common technique is the 'full network scan' approach, which removes each link at a time but in an automated way (Freiria et al., 2015). However, this is still computationally expensive so cannot be used for large data sets (Freiria et al., 2015). These methods can work well for network problems which are not for physical networks e.g. internet communications, however spatial methods are better for physical networks with physical disruptions.

Jenelius (2010a) proposes that research needs to go beyond those which are important in the normal functioning network, as when disruption occurs to critical links, re-routing is required (Jenelius, 2010a). So, these roads can be considered secondary critical links and their importance may only become apparent in disruption scenarios. Extrapolating the critical links theory further, research should consider a particular population group, as different routes and services are used by different groups of people, therefore the importance of a road link would vary depending on the perspective of the group of people.

Vulnerability and accessibility are terms which are linked to defining critical links in networks. The term vulnerability is mostly used in relation to the risks posed by severe transport disruptions (Cats & Jenelius, 2014) and was first formalised by Berdica (2002). Chen et al., (2015), Reggiani et al., (2015), Rupi et al., (2015) and Muriel-Villegas et al., (2016) state there lacks consensus on a definition for network vulnerability due to the numerous methods which try to capture the consequences of network disruption. In general, the definitions agree vulnerability is a measure of the susceptibility and impact of disruption on the network, independent of probability. Early conceptualisations of vulnerability for networks were related to the concept of reduced levels of accessibility as a consequence of network failure or degradation, however it has evolved to consider the wider consequences of failure (Muriel-Villegas et al., 2016).

Jenelius (2010a) takes the concept of network vulnerability further by considering network equity; defining importance through how the network design can lead to inequitable distribution of the consequences from disruption. Their example is that if a link closure results in one area being affected substantially more than others, the link could be considered more important i.e. critical to maintain. Jenelius (2010b) identifies horizontal and vertical equity, where horizontal equity ensures the costs and benefits to be equally distributed across all individuals whilst vertical equity provides special consideration to vulnerable groups. This theory could be applied to hazard risk assessments, as those who are most vulnerable are often prioritised in plans, therefore road links which connect to these groups could be considered critical even if they are not topologically critical.

Accessibility is the key for any transport network, since people need to be able to access the services they require. Accessibility is a useful measure for a number of reasons. Firstly, it can be used to calculate the opportunities available to an individual or group, like access to workplaces, shops, education and other amenities. Secondly it is important in emergency situations, where the level of access to affected groups needs to be assessed quickly to provide an appropriate response (Schnebele et al., 2014). Khademi et al., (2015) state there are similarities in the definitions of vulnerability and accessibility, and the link between accessibility and vulnerability was first made by D'Este and Taylor (2003). Common themes in accessibility definitions are the focus on positive connections networks provide between places and opportunities, rather than a more negative view of weak links and nodes given in vulnerability (Bocarejo & Oviedo, 2012; Ford et al., 2015a; Coles et al., 2017). Where they do correlate is that they can both be measured as a product of disruption, with the accessibility before and after an event calculated. The concept of potential and realised accessibility is also a consideration as analyses may indicate there is a route to provide access to a service, however other barriers like affordability of travel, availability of service and consumer choice can prevent potential accessibility becoming realised accessibility (Nieves, 2015). These factors can often be included through generalised cost indices.

There are many papers which address network accessibility for the general population, including Jenelius, 2010; Jenelius 2009; Jenelius & Mattsson, 2012; Rodríguez-Núñez et al., 2014; Sohn, 2006; Taylor & D'Este, 2003; Taylor et al., 2006; Chen et al., 2007, Chen et al., 2015; Balijepalli & Oppong, 2014. However, there are few studies which look at accessibility for certain groups. Nyberg and Johansson (2013) modelled how accessibility for elderly residents is affected by storm-felled trees in Southern Sweden, whilst Coles et al., (2017) focussed on elderly residents' access to care but this time in the context of flooding in York. Looking at key sub-groups of the population and which roads they use to access services can therefore be used to identify critical links those groups need to conduct their daily routines.

3.3.4 Transportation-related geospatial data systems

There are many geographic data sets which can be used in UK transport modelling and they can be classified into three categories; government data, open source data and commercial data.

3.3.4.1 UK Government sources

There are a number of data sources for data on traffic flow, speed, road location and accessibility available for use in UK transport studies. For traffic flow, the Highways Agency Traffic Information System collects information about traffic speeds and flows for UK motorways at a 15 minute temporal resolution (Jaroszweski et al., 2015) which can be accessed for transport planning and research. Another available data source is Annual Average Daily Flow (AADF) data for traffic points around the country, and they provide an average count of the number of vehicles using that section of the road by vehicle type. The UK national average value for generalised cost was designed to quantify travel time in transport planning (Postance et al., 2017) and can be used for major roads. The UK Department for Transport also has an accession GIS tool which calculates service accessibility to services via public transport for different groups (Ford et al., 2015b), useful data for assessing equality of service provision.

In terms of GIS data for building network models, Ordnance Survey have developed several road network data products. OS MasterMap Highways Network is a comprehensive GIS data set of all roads in Great Britain, available via license to businesses and under the Public Sector Geospatial Agreement to public sector organisations. These data include information on any road restrictions, turnings and road ownership. Ordnance Survey do provide an open source roads data set, OS Open Roads which provides basic information on road location and topology, but without the additional information like average road speed data for different times of day. Finally, the National Public Transport Access Nodes (NaPTAN) is a data set for Great Britain of public transport access points e.g. bus stops, rail stations, underground entrances. These data are for more specialised purposes but open the potential for modelling public transport networks.

3.3.4.2 Open source data and software

There are a number of open source data and software package and data which can be used in transport analysis. The main advantage is that they are free to use so can be used by more people than government or commercial software/data. However, as open source software and data are developed in a crowd-sourced manner, data set coverage and updating may be variable across a country, unlike Ordnance Survey whose mandate requires that all parts of the country be mapped to a consistent basis. For example, for Open Street Map (OSM) many contributors would not have training in surveying meaning their insertions and updates to the data base may be inaccurate or incomplete (Basiri et al., 2016). This can be mitigated through control processes to identify user errors (Basiri et al., 2016). This section describes some of the open source data and software packages relevant to transport network modelling.

Open street map (OSM) is a crowd sourced GIS project and includes transport data which can be freely downloaded and used in projects. For example, Gil (2015) created a multimodal urban transport network model as there are data for private and public transport options. The OSM data consist of nodes (points) and ways (polylines), which have tags associated with them to denote if they are part of a transport structure (Gil, 2015; Dingil et al., 2018). Therefore, these data can be used for network modelling for roads and public transport. The OSMnx Python software package is one example of a way to extract and convert the OSM transport data into a transport graph for network analysis (Dingil et al., 2018). A benefit of OSM transport data are that it is available globally, although there may be differences in detail and the frequency of updates. It also provides data for ‘soft modes’ of transport (cycling, walking) as well as road network data (Gil, 2015). In terms of positional accuracy, Gil (2015) found that for the Netherlands it was not substantially different from official data. However, a challenge with OSM data are quality issues with the attribute data, which are required for identification of transport nodes and ways (Gil, 2015).

OpenTransportMap (Jedlicka et al., 2016) is a platform built on OSM data and allows for transport data for Europe to be viewed, downloaded, incorporated into online maps and to be used for network routing (Jedlicka et al., 2016). It also includes dynamic traffic volumes calculated from Annual Average Daily Traffic flows, daily traffic volumes and hourly traffic volumes processed in the OmniTrans software. The data can be downloaded for Nomenclature of Territorial Units for Statistics (NUTS) level 3 for European countries and contains information on road names, travel mode, capacity and in some cases traffic volume. However, it does not contain average speed data or speed data for specific times of day, restricting its use for temporally specific analysis. Openrouteservice (The Heidelberg Institute for Geoinformation Technology, 2019) is another example of an open routing platform based on OSM data. Accessed via an API, several services are provided including directions, service area catchments and time-distance matrixes. The service allows for up to 2500 origin-destination pairs can be run per day for free. As with OpenTransportMap, the coverage is global as the underlying OSM data allows this. However, again it does not contain temporally specific traffic data to allow for temporally specific routing by itself.

QGIS is a well-known, open source GIS software which has several network routing plugins. These include Online Routing Mapper, which allows for routes to be generated with online services like Google Directions, Here, MapBox and OSRM; Networks, a processing algorithm for linear networks and multimodal routing; QNetwork, a network analysis tool using the Geographic Network Model and Road Graph Plugin, which solves the shortest path problem (QGIS, n.d.). With these plug-ins routes can be simulated either through python coding or the graphical user interface to generate shortest path routes. QGIS can also access OSM data via the Overpass Turbo API for use within the software. R is another widely used open source software package, although the majority of its applications are for statistical purposes. However, a transport package called stplanr has been developed to provide functions for transport routing problems, but also to manipulate and clean transport data sets (Lovelace & Ellison, 2018). R can make use of open street map data and access data via API's, so it has the potential to build powerful transport models for routing.

3.3.4.3 Commercial data and software

There is a range of commercial software and data which can be used for transport applications, covering regional-scale geographic analysis to planning for an individual road junction. ArcGIS is perhaps the most well-known GIS software as it is used by academics and industry. A license can be bought for the network analyst toolbox which provides a range of origin-destination and service area tools (ESRI, 2020b). Historical traffic data can be incorporated into network analysis through accessing the network routing server, available in ArcGIS online and ArcGIS Pro desktop (ESRI, 2018). The use of this service does require credits that the user pays for. This service does allow a user to easily specify the time of day they wish to run the route for, adding valuable temporal dimensions to the routing. Another advantage is open source GIS data and government data are often in a format compatible with ArcGIS software. ArcGIS therefore is a good software for running network analyses in, however the network analyst license and credits for historical traffic data can be prohibitively expensive for some users (see Chapter 4 for more discussion on this).

Google collects a vast amount of data on population movement, traffic and business opening hours. This is due to tracking mobile phones via GPS, and the live traffic can be seen when using Google maps online. The Google Routes function (Google, 2020) allows businesses to access google map data for routing to enable efficient route planning. It is updated 25 million times a day to allow for real-time location analysis, either through directions for routes or distance matrix's, to provide travel times and distances for one or more locations (Google, 2020). The distance matrix API facilitates the calculation of travel time and distance for a matrix of origins and destinations. Use of the API is on a pay-as-you go model, with a range of 0.005 – 0.004 USD per route calculated. TomTom also provides a routing and traffic service for developers through its API, in addition to a partnership with Apple, and again there is a cost price associated with accessing the data. HERE are another company which provides traffic data and routing services and they have a partnership with ESRI to provide the traffic data for ESRI's routing service (ESRI, 2020d).

In transport engineering, software tends to be designed for modelling small parts of a road network rather than a whole city. For example, PTV VISSIM is a traffic model which is good for testing the traffic flow at specific junctions, as it models all individual road users. This works for the purpose of many engineering projects which are adapting the existing road network but could not be used to look at traffic flow across a whole city. Two other examples of specific transport packages include Assessment of Roundabout Capacity and Delay (ARCADY) which is for roundabout modelling and Priority Intersection Capacity and Delay (PICADY) designed for testing the capacity, queues and accident risk at new unsignalled junctions.

3.3.5 Methods

This section gives a brief overview of the methods used for measuring proximity, flow, accessibility and critical roads in networks.

3.3.5.1 Proximity metrics

Proximity is a measure of how close two points are and there are three ways this can be calculated; distance (by Euclidean, Manhattan or network metrics), travel time (derived from a network or cost surface) and economic cost (measured through direct costs and value of time). Euclidean distance is the simplest and can be calculated either through a straight line between an origin-destination (Figure 3.3) pair, or using a buffer if considering service areas (Apparicio et al., 2008). In some cases, Euclidean distance and network distance can have a high correlation, so can be used as a proxy (Boscoe et al., 2012). However, Euclidean distance can never be completely accurate unless a road is a singular straight line. Manhattan distance is similar, although it calculates a distance between two points through the horizontal and vertical distances (Figure 3.3). Manhattan distance approximates travel in grid-based cities (Charreire et al., 2010) however, in other types of city Euclidean distance is preferable (Apparicio et al., 2008).

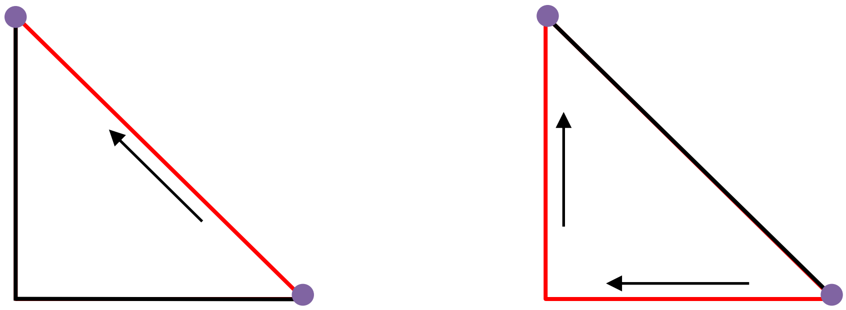


Figure 3.3 Euclidean distance (left) and Manhattan distance (right). The red lines show the distance measured in each.

The more advanced way of measuring distance travelled is using the road network itself based on the principles of graph theory (see Section 3.3.1). The shortest distance between two points on a network can be calculated using network topology principles (Postance et al., 2017). This proximity metric is more realistic than Euclidean or Manhattan distance as it is accounts for the spatial road network configuration, rather than being a straight-line approximation. Therefore, it is more common in transport and travel analyses to use software packages such as ArcGIS, QGIS and R which have network analysis functions. There are numerous shortest path algorithms; however, Dijkstra's algorithm (Dijkstra, 1959) is perhaps the most well-known (Wang, 2012). As it is the basis for the Network Analyst toolbox from ArcGIS used in this thesis, the algorithm workings will be described here.

Dijkstra's algorithm can be represented by G= (V,E) where G is the graph, V is the set of all vertices or nodes, E is the set of all edges (Eneh & Arinze, 2017). The algorithm is designed for weighted graphs, i.e. networks where there is a cost associated with each link, for example seconds of travel time (Jiang et al., 2014). To calculate a shortest-distance path, the origin node and the destination node are chosen. The algorithm works by first identifying nodes adjacent to the origin node and calculating weighted distances to these adjacent nodes, along the network edges. The path with the least weighted distance cost to an adjacent node is identified and this intermediary node becomes the focus for the next step. In the next step, adjacent nodes around the new focus node are identified and the cumulative weighted distances from the origin node are calculated. The node which the shortest cumulative weighted distance along the network edges from the two steps is identified, and this is the next node in the journey. These steps are repeated until the destination node has been reached and the minimum weighted distance pathway between the origin and destination is calculated. A path from the origin node to a destination node is said to be the shortest path if its total cost is the minimum among all the possible paths (Chen, 2003). Dijkstra's algorithm is the fastest known shortest path algorithm for directed graphs with non-negative weights (Eneh & Arinze, 2017). From Dijkstra's (1959) original publication of the algorithm there have been many revisions and improvements to the algorithm, including Jiang et al., (2014), Chen (2003), Wang (2012) and Eneh & Arinze (2017). One example in commercial software is the ESRI network analysis toolbox. This has several routing algorithms which are based on Dijkstra's, with additions to allow multiple-origin multiple-destination calculations (OD Cost Matrix and Closet Facility), and reductions to the performance time by including Hierarchical Routing to reduce the number of vertex pairs for which the path has to be calculated (ESRI, 2020a).

It is not always the distance which is covered which is the important factor in an analysis. In instances like calculating service areas for emergency vehicles, it is the travel time to get to a destination which is the key metric. The time taken to travel between two points across the network can be calculated from the distance covered and the speed of travel, taken from average road speeds if data are not available. The fact network distance measures do not tend to include individual behaviour is a limitation (Sohn, 2006) as it presents an optimum functioning of the network rather than the complexity of individual decisions combining e.g. travelling to/from work at the same time which leads to congestion. Therefore travel times are often an underestimation of reality (Borrmann et al., 2012). Calculating individual routes also fails to consider the interdependencies of links and how drivers affect each other whilst travelling (Jenelius & Mattsson, 2015).

Cost surfaces are another method of measuring proximity without using a network data set. These are raster files where the value in each grid cell is the measure of impedance across the surface. For example, Wood and Schmidtlein (2013) combined elevation, land cover and slope to create a speed impedance value across the raster surface, and when an average walking speed was applied, the time to evacuate from the area could be calculated. The cost surface therefore allows the travel cost, often in minutes, between an origin and destination to be calculated (Burns & Inglis, 2007) and the shortest suitable path found (Wood & Schmidtlein, 2013). One advantage of a cost surface is it allows physical and socioeconomic factors to be combined quantitatively (Jenelius, 2010a) and it has strong links to the generalised cost metric.

Finally, proximity between two points can be measured in terms of economic cost. In transport studies, this is often measured through generalised cost functions which include a set of cost components of a journey (Ford et al., 2015a; Davidson & Davidson, n.d.). There are two main components; the direct monetary costs and the cost of the time taken for the journey. Monetary costs are the direct costs of trips, which include the cost of petrol and the price of a ticket. This can either be from data for public transport ticket costs or can be based on distance with the price of petrol per mile. The distance can be calculated either by network analysis or cost surfaces, as described above. Secondly, the valuation of the costs accrued from the time an individual spends on a journey are included. The Value of Time is the value the average person places on an hour of their time, which could include their hourly wage for commuter journeys, although this varies for different socio-economic groups and the purpose of the journey (Ford et al., 2015a). This measure is used in the calculation of journeys and to understand the relative savings when improvements to transport networks are modelled, for example assessing schemes like the HS2 rail project (Department for Transport, n.d.).

3.3.5.2 Accessibility and network disruption measures

As described in Section 3.3.3, accessibility is a key transport concept and links to disruption metrics. Therefore, there are ways to measure accessibility based on proximity metrics described in Section 3.3.5.1. Service areas represent the area which is served by a specified service (e.g. hospitals, pharmacy, supermarket) and it is a key function in business models. Accessibility can be considered in terms of how many opportunities a transport system provides (Bocarejo and Oviedo, 2012) and service areas are one way of conducting research to ascertain which services are available. GIS software allows changes to the position of services or the network to be made to see how access changes. The generalised cost metric discussed in the previous section can also be used to calculate the cost of a delay to journeys through the extra time taken (Jenelius et al., 2011; Postance et al., 2017) the quality of public transport service (Currie, 2010). An example of a study which includes Value of Time calculations for measuring the disruption to travellers across a city are Kilgarrif et al., (2019), who calculate the costs to commuters in Galway, Ireland due to a flood event. They used Subjective Value of Travel Time values for three subsamples of the Irish population based on the 2011 census of population for Ireland. They could then calculate how much additional cost the flood event caused to commuters’ journeys. Ford et al., (2015a) created a GIS-based accessibility tool using a generalised cost measure in order to calculate accessibility in London to destinations like employment locations.

Introduced in 1946, Spatial Interaction Models, also known as gravity models, are a ‘prevailing framework’ for predicting flows between places (Simini et al., 2012). Based on Newton’s gravitational theory, the underlying concept is that the number of trips from an origin to a destination can be calculated through the 'attractiveness' of a destination, balanced against the relative attractiveness of alternatives. Equation 3.3 is the standard equation for Spatial Interaction Models (Ortuzar & Willumsen, 2011) where T_{ij} is the number of trips between origin *i* and destination *j*, O*i* is the number of trips available from the origin, D*j* the total number of trips attracted to that destination, *f*(C*ij*) a generalised cost function (often a distance decay function) and A_i and B_j balancing factors. Spatial Interaction Models are used to measure trip distribution in relation to current or past demand (Tsekeris & Stathopoulos, 2006) therefore require underlying data to provide validation of the calculated flows. These models can be singly constrained, where either the origin or destination totals are known, or doubly constrained, where both origin and destination totals are known (Simini et al., 2012). Another data requirement is for the generalised cost function, which can be as simple as distance, or include retail turnover, income, population numbers or employment data (Vickerman, 1974). A major limitation of this method is the amount of data which are required to fit the parameters of the model. It cannot be run if no data on flow between origin-destination pairs is available (Simini et al., 2012) and often relies on historical demand data for predictions (Tsekeris & Stathopoulos, 2006). Attempts have been made to create dynamic Spatial Interaction Models with real-time travel data (Tsekeris & Stathopoulos, 2006), however mostly historical data are used for planning purposes

$$T_{ij} = A_i \, O_i \, B_j \, D_j \, f(c_{ij}) \qquad \text{Equation 3.3}$$

Within transport geography, GIS based network analysis is a common method for measuring access and the impact of disruption (Ford et al., 2015a). Based on the network distance metrics described in section 3.3.5.1, GIS software can calculate routes between origins and destinations. Access can be measured by how much of an area is within an 'accessibility threshold' (Salonen et al., 2012) which could be by time or distance. Network analysis is often used in scenarios to see how access is affected when part of the network is closed (Nyberg & Johansson, 2013). The number of people who can/cannot now be reached along the roads, or get to their destinations, before and after applying the blockage, can therefore be calculated (Nyberg & Johansson, 2013). Often it is assumed the shortest route is taken which again, does not reflect every persons decision as people will take alternative routes to the 'logical' one (Sohn, 2006). Also, GIS network models do not consider traveller behaviour, so it is considered as the lower bound of travel times (Borrmann et al., 2012). But the strengths of GIS network analysis are that it can spatially model the real world, putting graph theory into context, but can also enable socioeconomic data to be linked to network data sets (Liu & Zhu, 2004).

As covered in Section 3.2.4.2, agent-based models can be used to model individual vehicle or person travel. This technique is common in evacuation scenarios (Wood & Schmidtlein, 2013), where knowledge of the individual moving through the space is important. The main difference between this and GIS network models is that travel behaviour is incorporated, which should provide a more accurate measure of travel times and routes. Queuing theory can also be included which simulates traffic (Lämmel et al., 2010), something which non-behaviour methods lack. The main advantage is this inclusion of driver behaviour as individual decisions are an important factor in the capacity and performance of the network (Brown, 2014). However, agent-based models are computationally intensive and require a lot of data to parameterise agent behaviours.

3.3.5.3 Overall travel model limitations

As with the other simulation models discussed in this chapter, there are limitations and uncertainties associated with travel models. An almost ubiquitous assumption in network routing models is that drivers have optimal knowledge of the network, and disruptions which are on the network, when applying a shortest-path function (Arrighi et al., 2019). This assumption means that the results are limited to being a minimum estimation of disruption (Arrighi et al., 2019), as in reality there would likely be more congestion and disruption as people do not have perfect real-time knowledge and some will ignore diversion signs. This links to the assumption often found in models that whilst an area is disrupted, the remaining network functions as normal (Cats & Jenelius, 2014). A disruption does indirectly impact the network around it through congestion propagating over the whole system (Cairns et al., 2002; Pregnolato et al., 2017a; Nagurney, 2011), again meaning the results of travel models are often an underestimation of actual disruption (Postance et al., 2017). Whilst ideally these assumptions would not have to be made, in reality they are used in models due to the difficulty of creating a model which incorporates traffic conditions and individual driver behaviours over a whole urban area. However, continuous improvements in computing power and ABMs may mean these issues are resolved.

The validation of route changes and travel times from disruption scenarios presents difficulties. Often there is not an alternative data set which demonstrates the traffic for a specific disruption event, for example ESRI and Google provide average travel times for day types but not specific events. In some cases, like Pregnolato et al., (2017b), there is sensor data for traffic available for periods where there has been disruption, in this case from pluvial flooding. However, not all roads are installed with traffic sensors, especially local roads, so this is not suitable for all locations. In addition, this would not be possible for risk planning exercises for events which have not recently happened in an area, but stakeholders wish to prepare (for example a terrorist incident). Expert judgement and stakeholder engagement is used in many fields so should be successful in this application as well as Crols & Malleson (2019) demonstrate. However, not all research studies will have the time and budget to invest in this level of validation.

3.4 The Impact of Flooding on Population Mobility and Transport Networks

The previous three sections have summarised the relevant literature in each of the three academic areas applied to this thesis. This section will now draw together the key strands where these disciplines overlap to inform the knowledge base upon which this thesis is structured.

To summarise the research issue, roads underpin mobility within urban areas, both during day-to-day life and emergencies, and fast connections to critical infrastructure and services is important (Arrighi et al., 2019). Pluvial flash floods are a common hazard in urban environments due to extensive impervious surfaces (Li et al., 2018) and identifying areas more likely to be cut off, or with reduced access, is fundamental for flood risk management and preparedness (Arrighi et al., 2019). Over half of flood victims in post-industrial countries are motorists trapped by road flooding (Terti et al., 2017; Debionne et al., 2016), consequently, daily mobility is one of the primary causes of population exposure and vulnerability to flash floods. There are numerous studies of population exposure in homes and offices. However, mobility aspects are not systematically included in studies assessing human exposure and vulnerability in flood risk measurement. Population density data are often used assuming a static distribution, which do not match the rapid dynamics of a flash flood model (Shabou et al., 2017). However, the availability of spatiotemporal population models and new 'big data' are leading to dynamic population models. The combination of all these dynamic elements remains a challenge for researchers and planners.

3.4.1 Applying risk theory to transport disruption

Risk theory, as outlined in Chapter 2, can be applied to transport network assessments for travel disruptions. In general, risk theory is implicitly, but not explicitly, utilised in transport network disruption studies. Examples of where risk theory has explicitly been applied are Berdica (2002), Rupi et al., (2015) and Pregnolato et al., (2017a), as they consider risk as a function of disruption probability and consequences, and use it to structure their analyses. As detailed in Chapter 2, this physical sciences definition of risk works well for physical structures and economic impact, however it is not as good at capturing the social elements of risk, or perhaps the social dynamics behind transport decisions. Using the risk = f(Hazard,Exposure,Vulnerability) conceptual equation, as detailed section 2.4.1, should provide a better foundation to combine transport modelling methods with spatiotemporal population data, considering specific population groups and their needs, which is one objective of this research. Indeed Arrighi et al., (2019) also use hazard, exposure and vulnerability to assess risk as it more functional and incorporates the natural and built environments (Arrighi et al., 2019), with this research including the social component of risk in the form of demographic data.

Hammond et al., (2015) state that there is clear scope for traffic models to be combined with flood models to improve estimates of the disruption to travel, and therefore calculate the cost of delay. In addition, demographic data or a population model should also be included to fully understand the risk to those travelling on the road network as it is the effect on people, rather than the physical infrastructure, which is important to many stakeholders. Knowing the number of motorists exposed on flood prone road sections (Debionne et al., 2016), and the characteristics of the population in travel, would provide a more realistic evaluation of the degree of exposure.

The role of individual behaviour in creating the dynamics of a transport system have been demonstrated earlier in this chapter; therefore, risk behaviour should also be considered as part of disruption planning. Liu et al., (2014a) found in their study that people chose closer destinations for shopping and leisure activities in adverse weather. Furthermore, Terti et al., (2015) argue that perception of warnings and environmental factors strongly depends on contingent conditions. For example, if a hazard occurs to disrupt a homeward journey, a person is more likely to persevere to get to their destination compared to if the same event increased the difficulty of a journey into work.

At present, mobility aspects are not often taken into account when assessing human exposure and vulnerability to natural hazards, with residential population data used most of the time (Debionne et al., 2016). There are some examples of specific disaster risk management research which examine finer temporal resolution population data. Bian and Wilmot (2015) compared the population exposed to a chemical spill during a normal weekday and a weekend festival day. Their data set was created by using land use data to assign a percentage of the population to each sub-area, and time-periods used to generate movement of population groups. However, these data are specific to New Orleans, and only for the two scenarios they tested, so it is not applicable to other areas. Kunwar et al., (2014) created their own temporal population data, using open source data; however, this is also study specific and only related to their work on evacuation of ten cities in the UK. Smith et al., (2015) investigated spatiotemporal changes to exposure from flooding in Cornwall by including the seasonal fluctuations in population from tourism. However, those conducting journeys were not explicitly considered and how the flooding would affect them.

3.4.2 Modelling disruption to travel over space and time

Most studies on travel disruption consider transport networks from an engineering point of view, focussing on network connectivity, rather than measuring the impact of road closures on the population which use them (Pfurtscheller & Genovese, 2019). The inclusion of multidisciplinary methodologies comprised of hydrology, transportation and have proved effective in many case studies (Li et al., 2018) therefore, the inclusion of spatiotemporal population data should enhance these measures further.

Roads are built to run at near maximum capacity to minimise costs, which means there is little redundancy in the system to cope with disruptions (Mossoux et al., 2019). When a disruption scenario occurs, the location of this is very important. If this is at a 'critical link' then the impacts will be much greater than a non-critical link. The spatial distribution of the network disruption is one factor (Postance et al., 2017), but time dimensions are 'excessively important' in transport systems, as fluctuations in demand and supply evolve on the daily to decadal scale (Reggiani et al., 2015). For example, the morning and evening 'rush hours' when most people are travelling to/from home and work are the busiest times on England's road networks, so a disruption has a greater impact than one occurring on a Sunday morning. The costs associated with a network link are time-dependent, as there are peaks and troughs in supply and demand, leading to the spatial distribution of critical and vulnerable links to evolve over time (Cats & Jenelius, 2014). Time is also important for travel behaviour, as Liu et al's., (2014a) paper shows that previous studies indicate commuters' travel activity is less variable in adverse weather conditions than non-commuters' travel. Space and time are interlinked concepts in transport disruption studies, and when capturing these dynamics in models often space is continuous whilst time is measured as snapshots (Wang & Cheng, 2001). The element of demand on the road network could be generated from data supplied by spatiotemporal population models as they produce time-specific information on the location of many demographic groups.

The majority of network disruption studies look at which links are critical or vulnerable, but without the specific context of time of day/year and the relative demand on the network. When time is considered, it is often travel time to a destination and how that changes, rather than the actual time the disruption occurs and how this interacts with demand to produce varying impacts. There are some studies which have taken this view. Freire et al., (2013) looked at how day and night time population distributions affected hazard exposure and time taken to evacuate after a Tsunami warning in Lisbon, using shortest route calculations. Their results showed differences in total evacuation time with a night-time evacuation being quicker, probably due to a smaller number of people present in the city. However, they only considered daily fluctuations and did not investigate seasonal fluctuations or the differences for population sub-groups. Kunwar et al.,

(2014) applied temporal population data to a model of total evacuation of ten UK cities to test how total evacuation time of agents varies with and without intervention. Again, they did not explore how routes change or compare the impacts from a specific day or time of year. Pregnolato et al., (2017a) examined the impact of disruption on an evening commute, due to the presence of validation data, however they do not examine the impact of flood events at other times of day. Shabou et al., (2017) used an activity-based ABM (MobRISK model) to combine travel behaviours with weather disruption to understand disruption. Their method required census data and travel activity survey data to assign daily activity programs to the population, then locate different activity areas, and mobility is generated when individual agents attempt to implement their activity programs. Exposure was measured by the likelihood of crossing flood roads along each individuals route (Shabou et al., 2017). These studies show that combining hydrology, transport and population is possible and a useful endeavour.

For spatial models, scenario testing is the most common way of testing disruption to discover which links are most critical (Pregnolato et al., 2016b). Examples include Pregnolato et al.'s (2017a) paper on accessibility during flooding of the road network in Newcastle-upon-Tyne, UK and Yin et al., (2016) combining a hydrological model with a transport risk assessment to see how disruption affects traffic flow. In order for comparison, a baseline scenario must be run initially before any disruption is included (Pregnolato et al., 2016b), a common method of measuring disruption to travel from flooding (Kilgarriff et al., 2019; Wei et al., 2018). But, there could be a very large number of scenarios, so a few are picked to test the impact of a disruption event (Smith et al., 2014a). If the change in population is the key variable, then the hazard event should be kept the same (Bian & Wilmot, 2015). There is no universal methodology for understanding the impact of disruption, therefore the results of the critical road analysis cannot be transferred to other scenarios or locations (Mitsakis et al., 2014).

As with all the models examined in this chapter, there are common limitations to models used for modelling spatiotemporal disruption to travel. Firstly, incorporating a dynamic flood, traffic and population model is challenging. Often, a dynamic flood model with a static travel model (Versini et al., 2010) or vice versa is common, reducing the realism of the simulation. Not including a dynamic flood model could lead to an overestimation of the impact of flooding on travel, due to localised changes in flood water depth meaning roads may be passable sooner than predicted (Chang et al., 2010). Secondly, it is hard to estimate the indirect costs of infrastructure damage on commuters. One London study estimated the flood related disruption to major roads at peak times in London was at least 146,000 Euros per hour (Arkell and Darch 2006).

3.5 The Evidence Gaps

Several gaps in the understanding of spatiotemporal aspects of flood hazards and population in travel are apparent from this literature review:

- Focussing specifically on the representation of sub-groups of the population in travel in spatiotemporal population modelling, which has received limited attention.
- Combining gridded spatiotemporal population data with network analysis, which not been done before, as far as the author is aware.
- Quantifying how the timing of flood onset interacts with flood magnitude to change the impacts on road travel disruption, through examining changes to time of flood onset for a short time window, like the morning commute.
- Linking the 'critical roads' concept in literature, based on topology, to the requirement for local authorities to identify which links are critical for resilience, based on need of the population (Cabinet Office, 2011)

The chapters that follow seek to generate evidence or methods to address these gaps.

Chapter 4 Methods and Data

This chapter provides an overview of the study design and methodology used in the thesis. The first section covers the study design and choice of case study location. The second section describes the three main existing modelling tools utilised in this thesis and their application to this work. Finally, the overall structure of the results chapters and the methodologies of each (Chapters 5, 6 and 7) is given. More detailed and specific methodologies are included in each results chapter.

4.1 Study Design

4.1.1 Case study site

York has been selected as the study site due to its population size (just under 200,000 in 2011 (City of York Council, 2018), which is a sufficient size to demonstrate a new methodology of using spatiotemporal population data with network analysis and flood modelling, but not too large the data processing is unmanageable. York is also a reasonably self-contained city in terms of commuting patterns, with 71% of those who work in York also usually resident in York (2011 census, WU01EW - Location of usual residence and place of work by sex (MSOA level)), unlike other more complex conurbations like Manchester or London. York has a long history of fluvial flooding and regularly experiences small floods, due to being situated at the confluence of the Rivers Ouse and Foss. The most recent large fluvial flood event was December 2015, when there was substantial flooding and the Foss Barrier failed, leading to 600 properties being flooded (Environment Agency, 2016). York, like most urban areas, is also at risk from pluvial flooding and has experienced pluvial flood events in the past, as described in Section 4.1.2. However, there is less knowledge of the effects of pluvial flooding than the fluvial flood risk (City of York Council, 2012). There is only one previous academic study, to the author’s knowledge, which assesses the impact of flooding in York on access to services. Coles et al., (2017) demonstrated the relatively small number of roads which are flooded in York can cause significant changes to the accessibility of care homes to emergency services. Hence, it has been shown that there are impacts on travel in York, and this project will extend Coles et al.’s (2017) work by investigating disruption to different population subgroups and consider how the time of flood onset is important.

Figure 4.1 shows the geographic area which data were gathered for the spatiotemporal population modelling of York. The red square captures the desired study area which is the York Local Authority. In addition, a buffer zone is required for the SurfaceBuilder247 population model to allow for travel from outside of the study area into the study area (further explanation in Section 4.2.2). A 20 km buffer was chosen as 82% of the people working in York travelled this distance or less (2011 census table for York Local Authority WP702EW - Distance travelled to work (Workplace population)), but it also would include the furthest likely distance school children would travel into York. All the output areas whose population weighted centroid fell within the study area and buffer zone were included in the input for the SurfaceBuilder247 model, although only the study area is displayed in the model output.

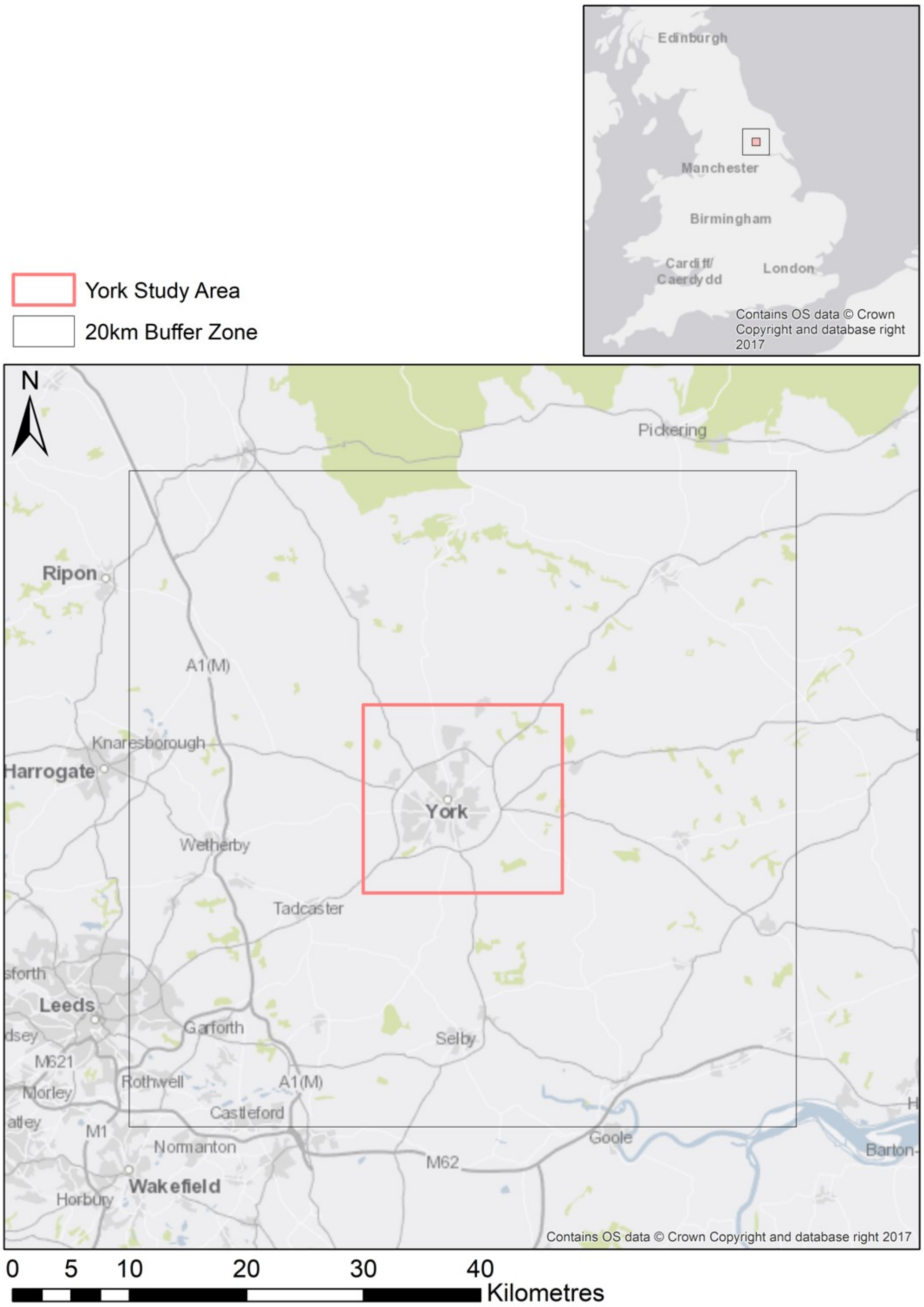


Figure 4.1 Study area and buffer zone around York, UK. Contains public sector information licensed under the Open Government Licence v3.0.

4.1.2 Hazard selection and York pluvial flood history

York is known for fluvial flooding with the low-lying historic city centre at the confluence of the River's Ouse and Foss regularly flooding (Environment Agency, 2020). However, there are several reasons for choosing pluvial flooding rather than fluvial flooding. Firstly, pluvial flooding has the potential to affect a much greater area of York than fluvial flooding. The floodplain in York from the Rivers Ouse and Foss does not include many of the primary roads, whilst any part of the urban area has the potential for pluvial flooding. There have also been previous instances of York's outer A roads being affected by pluvial flooding causing disruption for several hours (City of York Council, 2012). Additionally, a study of pluvial flooding has a greater transferability as it is an important issue for many urban areas in the UK (Kendon et al., 2018). Secondly, pluvial flood events are less predictable than fluvial flood events so the chance for preparation by drivers is limited (Li et al., 2018; Boeing et al., 2019). Pluvial storm events also tend to be shorter in duration than fluvial flood events, usually a couple of hours, which fits with the goal of this research to study the impact of an event at different times of the day. Thirdly, pluvial events are predicted to increase over the next century with climate change. Kendon et al.'s (2014) and Chan et al.'s (2018) climate models predicted increases in rainfall intensities in winter, and a greater number of short duration rainfall events in summer in the UK, with the EA's 'accumulation threshold' of 30mm/hr expected to be exceeded more frequently in the future (Kendon et al., 2014).

City of York Council (2019) state that since 2000, there have been notable pluvial flood events in the summers of 2007, 2012, 2013, 2014 and 2018, although other smaller incidences of pluvial flooding would have occurred much more frequently. There are flood warning systems available from the Environment Agency providing notice if a fluvial flood event is likely. However, pluvial flooding is harder to predict and can affect any part of the city, indeed the previous pluvial flood events have occurred in different locations on each occasion (City of York Council, 2019). In York there is not considered to be a link between places at risk from fluvial flooding and pluvial flooding, as pluvial flooding occurs mostly due to ineffective or blocked drains and/or heavy rainfall (City of York Council, 2019).

Pluvial flood policy for York is based on the national guidelines set out in Section 3.1.3, and the City of York Council is the LLFA for the study area of this thesis. The most recent intense summer rainfall event which led to a Surface Water Flood Investigation, as specified in the Flood and Water Management Act (2010), was on the 13th August 2018. This intense summer rainstorm resulted in >40 mm of rain falling on a large area of the city in just over 15 minutes, at a rate in excess of 100 mm/hr in some locations, the consequences of which were 123 locations reporting flooding including many roads, gardens and a small number of properties and businesses (City of York Council, 2019). Heavy rainfall had been predicted at >32 mm/hr; however, the duration of rainfall was greater than expected, overwhelming drainage systems and resulting pluvial floods (City of York Council, 2019). Whilst this flood event is less severe than a fluvial flood event in York, it does outline the ongoing difficulties of prediction and mitigation of pluvial flood events. Research applying climate change models specifically for Yorkshire suggest that increasingly intense rainfall is a very likely outcome (Boeing et al., 2019), which is expected to lead to pluvial flooding in urban areas.

The recommendations of the Surface Water Flood Investigation report include that the city of York and flood risk partners, should develop a strategy to increase understanding of pluvial flood resilience within the local population and internal highways design teams, as the effect on roads was one of the most significant. It was also recommended to develop and fund a network of digital rain gauges in the city to provide data for a better warning system (City of York Council, 2019). The local paper reported that funding to install electronic flood signs around the River Ouse, to warn people about the risk of fluvial and pluvial flooding, had been approved, making use of technology to track, respond to and investigate heavy rainfall events (Laversuch, 2020). Academically, Yorkshire's Integrated Catchment Solutions Programme (iCASP) has been set up as a NERC funded programme to target research and expertise to understand how the latest advances in probabilistic rainfall forecasting and high resolution hydrodynamic modelling can be combined into probabilistic, hyper-local forecasts with short lead times for making decisions (Boeing et al., 2019).

4.1.3 Population group selection

Whilst all of the population in an urban area can be affected by pluvial flooding and are at risk of experiencing negative consequences, some groups experience a greater risk due their age, socioeconomic status or health characteristics making them more vulnerable (as discussed in Section 2.1.4). Within the field of natural hazard risk, there are countless papers which describe which groups of the population are most vulnerable to hazard events (Cutter et al., 2008; Wisner et al., 2004; McFadden et al., 2006; Civil Contingencies Secretariat, 2008). Broadly, there is consensus over which groups are most vulnerable, although this does vary depending on the nature of the hazard (Wisner et al., 2004). In this thesis, the UK Government's guidelines for defining vulnerable groups, as outlined in the report 'Identifying People who are Vulnerable in a Crisis' (Civil Contingencies Secretariat, 2008), are applied as it is the guidance policy makers use in England. This will aid the transference of results to the non-academic sector. The groups identified in this report broadly concur with the academic literature (for example Cutter et al., 2003).

The groups identified by the Civil Contingencies Secretariat (2008) (Table 4.1) were considered for implementation within the SurfaceBuilder247 model. The model framework requires the population to be divided into mutually exclusive groups that share temporal characteristics, where the location of their origins and destinations is known, in addition to the typical amount of time they spend at each location. These conditions meant several vulnerable groups identified by the Government (e.g. the mobility impaired, sensory impaired, or homeless) were not suitable for inclusion as the required data were not available, or they could not be defined as a mutually exclusive group. Children of primary school age were chosen as the vulnerable group for modelling as they are a mutually exclusive group and there are sufficient data to model their daily movements.

The second population subgroup studied in this thesis' analysis are commuters going to workplaces in York aged 18-64 years. This will be the majority of the residential population in York. This group was chosen in order to capture the effects on workplaces due to flooding disruption. The commuting patterns and locations of where people travel to work is contained withing the 2011 census and other data sets like the quarterly labour force survey.

Table 4.1 Vulnerable groups identified by Civil Contingencies Secretariat (2008)

Group	Reason considered vulnerable
Children	Children, particularly young children, lack knowledge on how to deal with emergency situations. Local authorities have a duty of care whilst children are at school.
Older People	Certain sections of the elderly community, especially those in ill health, require regular medical attention or support.
Mobility impaired	Cannot react to emergency situations as quickly as those who are not mobility impaired
Sensory impaired	For example, blind or reduced sight, deaf and speech impaired. May need extra communication measures to receive warnings or help evacuating.
Homeless	Unlikely to have standard infrastructure used for warnings (phones, television) and do not have financial reserves to cope with an emergency
Tourists	Are unfamiliar with the area and procedures, would not be signed up to local warning systems.

4.2 Models, Tools and Data Sources

This section provides an overview of the models, tools and software applied in this thesis. Specific details on how each is applied to the analysis in each chapter is discussed in Chapter 5, 6, and 7.

4.2.1 Pluvial flood modelling

The Flowroute – i^{TM} flood model, created by Ambiental Risk Analytics, was used to generate pluvial flood scenarios for York, UK. Flowroute-i^{TM} is a steady state 2D flood model designed to produce accurate flood event data for local, regional and national projects. Previously it has been used for creating national flood maps for countries including the UK, Ireland and Australia and local flood analysis including pluvial flood modelling for Hull, UK. It was included in the EA's benchmarking exercise on the capabilities of 2D flood modelling software available in the UK, and the Flowroute-i^{TM} software package was found to be appropriate for supporting flood risk management activities (Néelz & Pender, 2013).

For this thesis, Flowroute – i^{TM} was used to create flood data for two pluvial storm events in York; a 1 in 30 year storm event and a 1 in 100 year event. These two return periods were chosen as both storms would lead to disruption to the road network, with representation of a smaller but more frequent event (1 in 30 year) and a larger but less frequent event (1 in 100 year). This is similar to the EA's pluvial flood maps which show a 1 in 30 and 1 in 200 year event (City of York Council, 2019). As this flood modelling is based on design hydrographs, it does not simulate a specific historical flood event but instead represents where flooding is likely to occur under different storm scenarios. The geographic area used in the flood modelling was based on the study area defined for the SurfaceBuilder247 population model, with the same 20km buffer area for capturing catchment drainage into the city. The input data required for Flowroute-i^{TM} is outlined in Table 4.2 and the specific data for the York study area were applied.

Table 4.2 Data required for input to Flowroute- i™ model and their source

Data	**Source**	**Description**
Digital Terrain Model (DTM)	Environment Agency LIDAR data, 5m resolution, BlueSky Lidar	A raster which represents the height above sea level of the terrain for the study area. The resolution was resampled to 5m due to the size of the area required for flood modelling.
Cellmap file	Ambiental Risk Analytics	A raster which represents the land classifications of the model area. The cells are given a value of either 0 (water), -1 (rural) or -2 (urban) to denote land use class for different infiltration values to be applied.
Manning's roughness file	Ambiental Risk Analytics	A raster which represents the manning's roughness coefficient for each cell in the model area. This classification is based on the OS Vectormap data product which has been reclassified to the appropriate manning's roughness coefficient for each land use type.
Storm Hydrograph	ReFH computer programme (UK Centre for Ecology & Hydrology, 2020) used with permission from Ambiental Risk Analytics	A text file which represents the rainfall (mm/hr) at each time step for input to the flood model. This is generated from storm hydrograph profiles for the catchment of the River Ouse upstream of York.

To model pluvial flood scenarios for York, a set of simulated storm events were required to provide rainfall inputs to the model. Catchment data from the Centre for Ecology and Hydrology's Flood Estimation Handbook (FEH) web service (UK Centre for Ecology & Hydrology, 2020) was downloaded for use in the Revitalised Flood Hydrograph (ReFH) computer programme in order to generate storm hydrographs for the catchment. The FEH was the primary source of the methodological approach used in designing the pluvial flood scenarios. The storm events chosen were summer storms for a 1 in 30 and 1 in 100 year magnitude. Using the ReFH computer programme flood hydrographs, two storm profiles were created by taking the peak rainfall rate and applying it for one hour. The rainfall then stops and the model continues to run for a further two hours simulation time to allow for overland flow and drainage (Figure 4.2). These data were input into the model as a text file. A rainfall event of an hour fits the aim of researching the effect of sudden onset pluvial events and is compatible with focussing on a short time frame for analysis in Chapter 7 (the morning commute). Flowroute – i™ models the depth and velocity of water in each cell during the simulation, and a raster for both depth and velocity across the study area is produced at each 15 minutes of simulation time. Additional final maximum depth and velocity rasters, which represent the maximum value each for each cell at some point during the simulation, were also generated for each scenario.

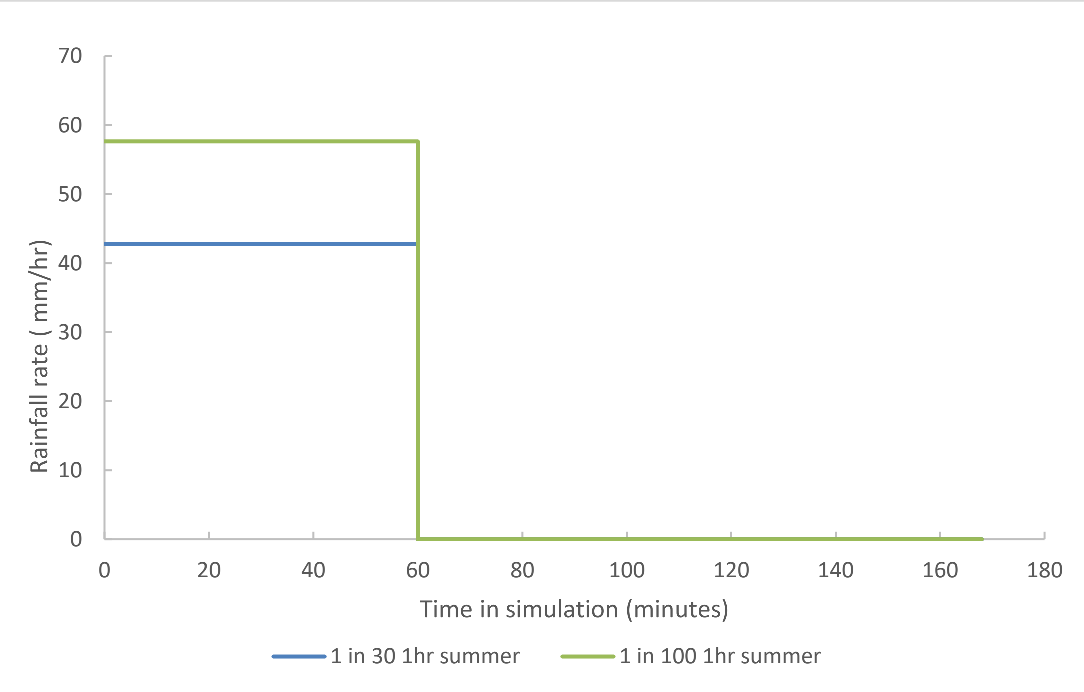


Figure 4.2 Storm hydrograph profile created from the ReFH catchment data. These data are used as a boundary input into the flood model. The graph shows a storm of peak rainfall intensity for an hour before the rain stops. Contains data which are copyright of UKCEH and Ambiental Risk Analytics, used with permission.

The output raster data for the two scenarios were processed before they were used within the network analysis. In the network model, the flood data are represented as barriers to the network where travel through them is prevented. Therefore, the exact modelled flood depths across the city were not required. A vector file for each time step of the flood scenario, which contained the areas where the floodwater met the conditions to prevent travel of vehicles along the road was sufficient. Two criteria were applied to the raster data. Firstly, the depth of water is a key factor in whether a road is impassable and a depth of 30cm was applied as the threshold, a value used in other studies (Li et al., 2018; Arrighi et al., 2019; Pregnolato et al., 2017b) and in UK government advice as the maximum depth for safe driving. Secondly, small patches of water can be easily circumnavigated, or very small patches could be model artefacts. Therefore, any areas of water which were smaller than 125m^2 were removed. Finally the data were further cleaned to remove flooding across the city's bridges, as Coles et al., (2017) did for their study in York. This was necessary because there is no record of the bridges flooding in York and bridges are often poorly represented in the DTMs used in flood models (Coles et al., 2017). After this processing the areas of flood water which would pose an issue to road travel for each 15-minute time step for both pluvial flood scenarios were left.

4.2.2 Population model

The approach for creating spatiotemporal population data developed by Martin et al., (2015) using the SurfaceBuilder247 software tool (Martin, 2011) is applied in this thesis. A full description of the background, data and framework can be found in Martin et al., (2015) and this section will summarise the model as it is relevant to this thesis.

The Population24/7 project (University of Southampton, 2011) developed a framework for spatiotemporal analysis based on previous work of an adaptive kernel density approach for building gridded population models by Martin (1989) and Martin (1996) (Martin et al., 2009). Figure 4.3 diagrammatically represents the spatiotemporal framework. The framework is split into two parts, the data system and the analysis system. The division of the data system and the data required for each domain is given in Table 4.3.

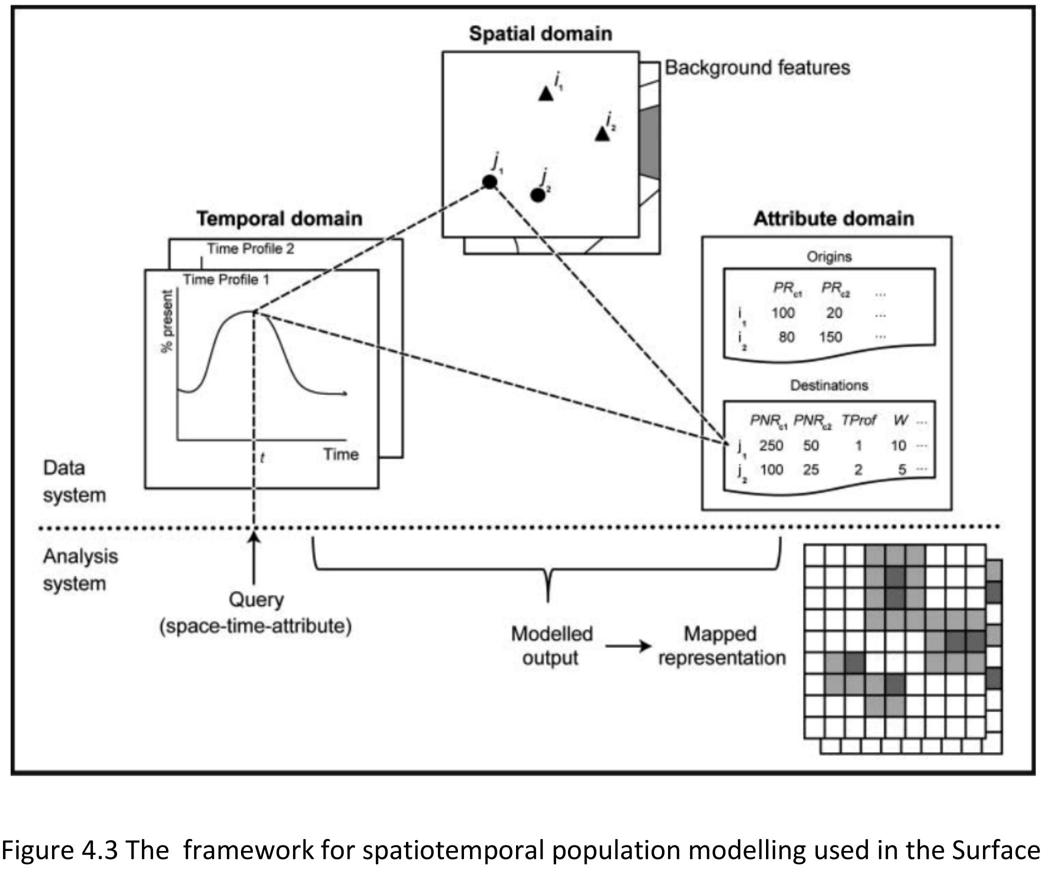


Figure 4.3 The framework for spatiotemporal population modelling used in the SurfaceBuilder247 software (Figure from Martin et al., 2015). The two parts (the data system and analysis system) are divided by the dotted line. Permission to reproduce this figure has been granted by Taylor & Francis Group.

Table 4.3 Description of data classifications required for each domain shown in Figure 4.3 and examples of data sources which could be used to capture them.

Domain	Data Classification	Description	Example data source
Spatial domain	Origins	The usually resident population located at census population weighted centroids. Includes mobile and immobile groups (e.g. those in residential facilities)	Resident population data from the Census for divisions like output areas, source the Office for National Statistics. Communal establishment resident data.
	Destinations	Locations where non-resident populations travel to for an activity e.g. work, education, health care, retail, leisure	Census data like workplace data, Edubase data, hospital episode data
	Population in travel	Those who are in transit between two locations at the time of day being modelled. A background layer containing features which people travel on (roads) and where people cannot be (land use classes like water) is required.	Time profiles for activities; AADF data for weighting background layer from the Department for Transport (NTM; Department for Transport, 2009). OS Strategi data for land use classes.
Temporal domain	Time profiles	Data which informs when the population is present at origins or destinations. Each destination type will have a daily time profile with the percentage of the destination population present at each time of day.	Local education authority and survey data (schools) Quarterly labour force survey (workforce)
Attribute domain	Origins	Associated information with the origins and destinations including the population subgroups which attend a destination (e.g. age group attending a school).	Local education authority data
	Destinations	Destinations each have a catchment or wide are dispersal area where the population travels to the destination from.	Travel to work census data

The spatial domain represents the 'containers' of human activity which can be divided into origins and destinations (Table 4.3). There is also a third group of the population who are in transit between locations and this can be estimated by examining the numbers of people arriving and departing from destinations in adjacent time periods (Martin et al., 2015). To provide a spatial structure for the distribution of the population a background layer is created (Figure 4.3). The background layer has two functions. Firstly, it masks out features where no population can be located e.g. water bodies (Martin et al., 2015). Secondly, it contains information on the major road network in order to represent population in transit (Martin et al., 2015), taken from Highways England (Highways England, n.d.). Two types of traffic data are combined in the background layer. Firstly, Average Annual Daily Flow (AADF) data from the Department of Transport which represent the average number of each type of vehicle passing along a section of road (Department for Transport, 2018b). Secondly Department for Transport traffic flow time periods covering the 7-day week with 17 time periods within a single day which have similar traffic levels (Martin et al., 2015). Together, these two types of data are able to simulate weights for the strategic road network for any time specified in the model input, so busier roads have more people travelling on them. This results in a raster layer which contains gridded traffic-derived weights that are used to estimate counts of population in travel in grid cells. The adjustments to the background layer from the original Martin et al., (2015) project are outlined in Section 5.2.1. The attribute domain is linked to the spatial domain as it provides the information for the origins and destinations (Martin et al., 2015). This includes the population sub-groups who attend a destination or the catchment area from which population to supply to destination can be drawn from Table 4.3.

The temporal domain is the collection of time profiles used to inform the population present at each origin and destination for a given time of day (Table 4.3). A data library of time profiles is built to cover each type of destination and the types of day which are of most interest (Martin et al., 2015). With a large time profile database, this approach allows for a more flexible choice in time of day to be used in the analysis query rather than restricting the analysis to predefined reference periods (e.g. day time or weekday morning) (Martin et al., 2015). As Figure 4.4 shows, the time profiles used in this thesis have data at 15-minute intervals and represent the percentage of the population at a destination at the specific time. The data is derived from the Quarterly Labour Force Survey (UK Data Service, 2020).

The entire modelled population must be represented in the spatial domain at any time point for volume preservation i.e. that the number of people does not change but rather everyone is redistributed across the area (Martin et al., 2015). This means a geographic area larger than the study area, and the accompanying data, needs to be input into the model to allow for a buffer zone of people travelling into and out of the study area during the day (Martin et al., 2015). This is recommended to be at least 20 km (Martin, 2017,pers. comm.).

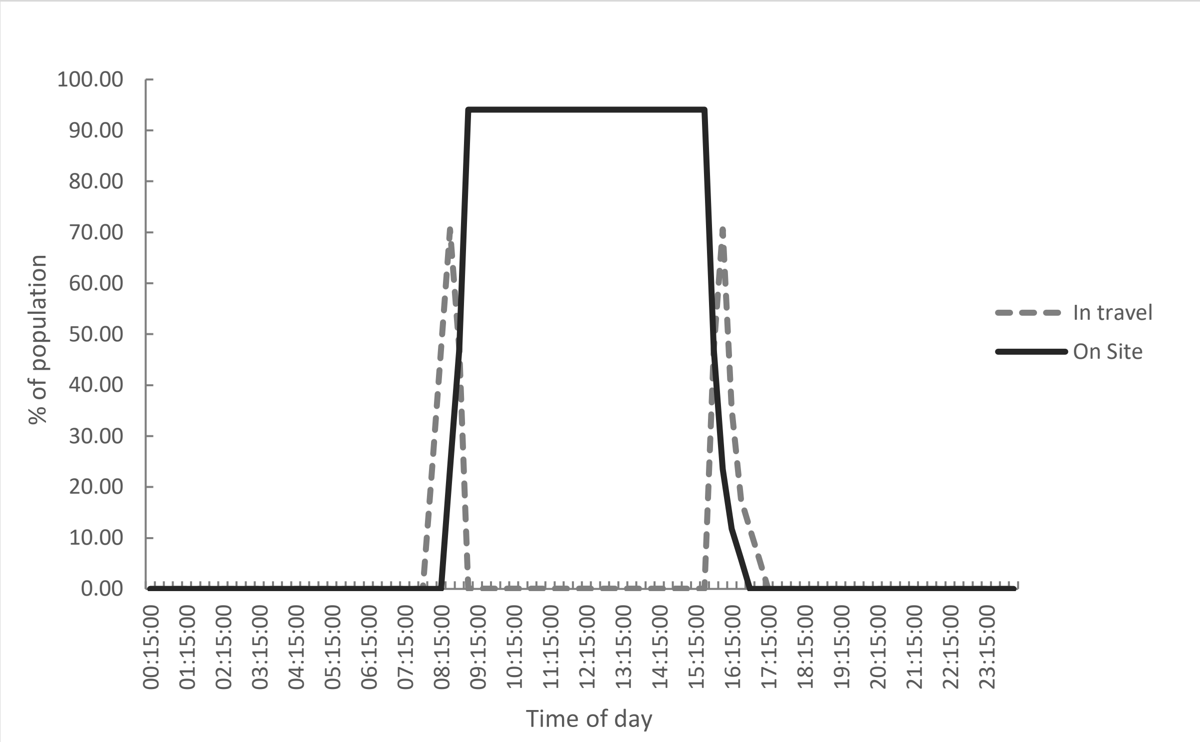


Figure 4.4 Time series profile for primary school children. The dashed line is the percentage of the population subgroup who are in travel at that given time of day and the solid line is the percentage at their destination site. This profile shows that children travel to school between 8am and 9am and 94% of the population group are present a school for the day.

The analytical system (Figure 4.3) provides the means to interrogate these data to answer specific spatiotemporal queries (Martin et al., 2015). Origin population is reallocated either to the nearest destination or to meet known proportions of population travelling in successive distance bands (for example those created from travel to work data) (Martin et al., 2015). This approach allows for overlapping catchment areas; however, it does not account for more complex travel patterns, like a traveller not going to the nearest applicable destination. The product of the model is a series of rasters which represent each population sub-groups distribution at the time of day specified. The spatial units of this raster should reflect the spatial resolution of the source data, and for origin populations drawn from census output areas this is a 200 m grid cell (Martin et al., 2015). The grid cells contain the estimated number of people located there at that time of day. A gridded output has the advantage of stability over time and is easier to integrate with other data sets which are commonly gridded formats, for example environmental data (Martin et al., 2015).

Applying the SurfaceBuilder247 software to this research, Table 4.4 describes the main data required for this thesis. The chosen year of analysis was 2016 so the data closest to this year was used. The origin files for the residential population counts per output area are derived from the 2011 England and Wales census and projected to 2016 using mid-year estimates at the LSOA level. Seven population sub-groups were used to cover the entire population, dividing the population by age, based on educational and workplace stage, which have similar temporal and spatial patterns. The groups were split into the following ages; 0-4, 5-10,11-17, 18-64 students, 18-64 non-students, 65-79 and over 80. For destinations, each was given a wide area dispersion radius which specifies the area that people should be drawn from, for workplaces this was from travel to work data and for schools this was from Edubase catchments. For a detailed description of all data sources for the SurfaceBuilder247 model see Appendix A.

Table 4.4 Sources for the main data used in the SurfaceBuilder247 tool when implemented in this thesis. The codes in brackets refer to 2011 census tables the data was acquired from.

Data	Source	Year	Description
Population totals and age structure	Office for National Statistics (QS103EW)	2011	Residential population at origins
Mid-Year Estimates of York's population	Office for National Statistics	2016	Data applied to residential population to scale to 2016
Communal establishment residents	Office for National Statistics (QS420EW and DC4210EWla)	2011	Data for the 'immobile population' who are in facilities which mean they do not leave during a typical day e.g. care home, prison
Workplace zone and travel to work	Office for National Statistics (via Nomis) (WP605EW and WP702EW)	2011	Location of workplace zones and the distance travelled by commuters
Hospital admissions	NHS digital	2016	The number of people in hospital for the day type
Schools	Edubase York City Council Independent School's council	2017	The location of schools and the de facto catchments
Traffic data (AADF and traffic flow for time of day)	Department for Transport (via data.gov)	2016	Data for the background layer weightings to enable distribution of the population in travel
Time profile	Quarterly Labour Force Survey	2006	Proportion of population group in travel or at work

4.2.3 Network analysis tool

The network analysis tool used in this research is the closest facility tool, part of ESRI's ArcGIS network analysis toolbox (ESRI, 2020a). ArcGIS routing tools have been utilised in other similar research projects. For example, Kilgarriff et al., (2019) used ArcGIS routing tools for their study of the impact of flooding on commuters in Galway. Pregnolato et al., (2015) used ArcGIS to calculate least cost path routes between origin and destination locations according to the shortest time. By overlaying water depth estimates onto the network, they could assess the impacts of the flood through comparing the perturbed travel times from the flood event to the non-flood event. ArcGIS network analysis tools have been used in relation to other natural hazards as well as floods. Toma-Danila (2018) used ArcGIS network analysis to study the implications of urban road network failure due to earthquakes. Toma-Danila (2018) states that ArcGIS one of the most advanced and used GIS platforms for road network modelling, due to its flexibility to be used in many different situations whilst being a specialised tool for routing algorithms. ArcGIS can handle large networks and process routes across an urban area. As Ahmed et al., (2017) attest, the software allows for time to be used as the criteria for shortest route, more useful than distance in real-world travel studies, and a specific travel start time and day type can be specified. This allows for temporal differences in travel times across the study area to be explored. It is for these reasons why it was selected for this research.

The closest facility tool locates the nearest 'facility' (destination) for an 'incident' (origin) and returns a shapefile of the route and the journey time and distance travelled. The analysis enables users to find the routes for multiple origins to multiple destinations with parameters including; specifying the number of destinations to search for each origin, mode of transport, impedance value which the shortest path is calculated on (time or distance) and time of day for the journey. This tool is based on the well-known Dijkstra's algorithm (see Section 3.3.5.1), which is used to find the shortest path between two locations (Debionne et al., 2016; Li et al., 2018). ESRI's closest facility analysis uses a modified version of Dijkstra's algorithm to reflect real-world transportation rules like one-way streets, turn restrictions and barriers on the network (ESRI, 2020a). It is also applied in conjunction with a hierarchical path solver which improves calculation time and is designed to reflect travel behaviour. Hierarchy is based on the fact that not all road segments are used equally, since some roads are major transport routes with more people travelling on them and some are small minor roads with few cars (ESRI, 2020a). The hierarchical path solver searches from the origin and destination location for entry points to higher-level roads, for example A roads in a UK city, and then tracks along these segments until the two meet (ESRI, 2020a). This means a smaller number of roads are searched which results in quicker solve. The hierarchical approach is also a better reflection of driver behaviour as main roads are more often used when traversing a city than local roads.

The server version of this tool was used over the desktop version as it allowed for the use of ESRI's historical traffic information. As the focus of this research was on how the time of day of a flood event affects the travel across the city, including traffic data for specific times of day was important. ESRI's traffic data allows for any time of day to be specified for a typical day of the week when running the closest facility query. The traffic data are collected by HERE who have a partnership with ESRI (ESRI, 2020d). HERE collect billions of GPS and cell phone probe records per month which is compiled to compute accurate speeds, utilising sensor data to augment the probe data collection where available (ESRI, 2019). Historical traffic information is generated from the average of observed traffic speeds over the past three years (ESRI, 2019). These traffic data are representative of 'normal' conditions and do not reflect the traffic conditions and congestion caused by a flood event in the city. However, in the absence of city-wide traffic data for every road during a flood event it was applied to the network analysis conducted for the flood scenarios to keep that variable the same. The network analysis using traffic data also requires subscription to ArcGIS online and uses credits per closest facility analysis (ESRI, 2020e) which does restrict the use of these data. It is worth noting though that other routing tools with traffic data, for example Google routing, also use a credit-based system.

Barriers are a term for any object, in this case water, which impedes travel on the road network and can be represented as a point, line or polygon feature (ESRI, 2020b). There are two options for barriers: either they slow down travel along the link by a set scale factor, or travel is not permitted (ESRI, 2020b). Barriers were used to represent the flooding across the study area to capture the effect on the road system. As described in Section 4.2.1, the raster output from the Flowroute-i™ model was converted into polygon files and processed to represent the areas where flooding fit the criteria for causing disruption to the road network. These polygon files were included as impassable barriers in the closest facility analysis in all flood scenarios.

For Chapter 5 (modelling spatiotemporal flows), the analysis was conducted manually via ArcGIS online whilst initial analysis for primary school children using the closest facility tool with traffic data were conducted. For Chapters 6 (Evaluation of spatiotemporal flow data) and 7 (the effect of flooding on travel), a Python script was written to enable batch processing of a folder of data for one scenario, a much quicker computational process than individually running data using ArcGIS online.

4.3 Overall Methodology

4.3.1 Thesis overview

The following three chapters contain the analysis conducted during the course of this research project and the results inform the discussion in Chapter 8. Figure 4.5 provides an overview of how the research chapters are structured. Chapter 5 presents the first piece of analysis in this thesis and examines how to create spatiotemporal flow data from the spatiotemporal population data. Chapter 6 takes the approach developed in Chapter 5 and applies it to commuters before introducing a flood scenario. The final piece of analysis is presented in Chapter 7 where the main aim of the thesis is addressed. In this chapter, the spatiotemporal flow data created in chapters 5 and 6 is applied to different flood scenarios, to see the effect of flood onset time and magnitude.

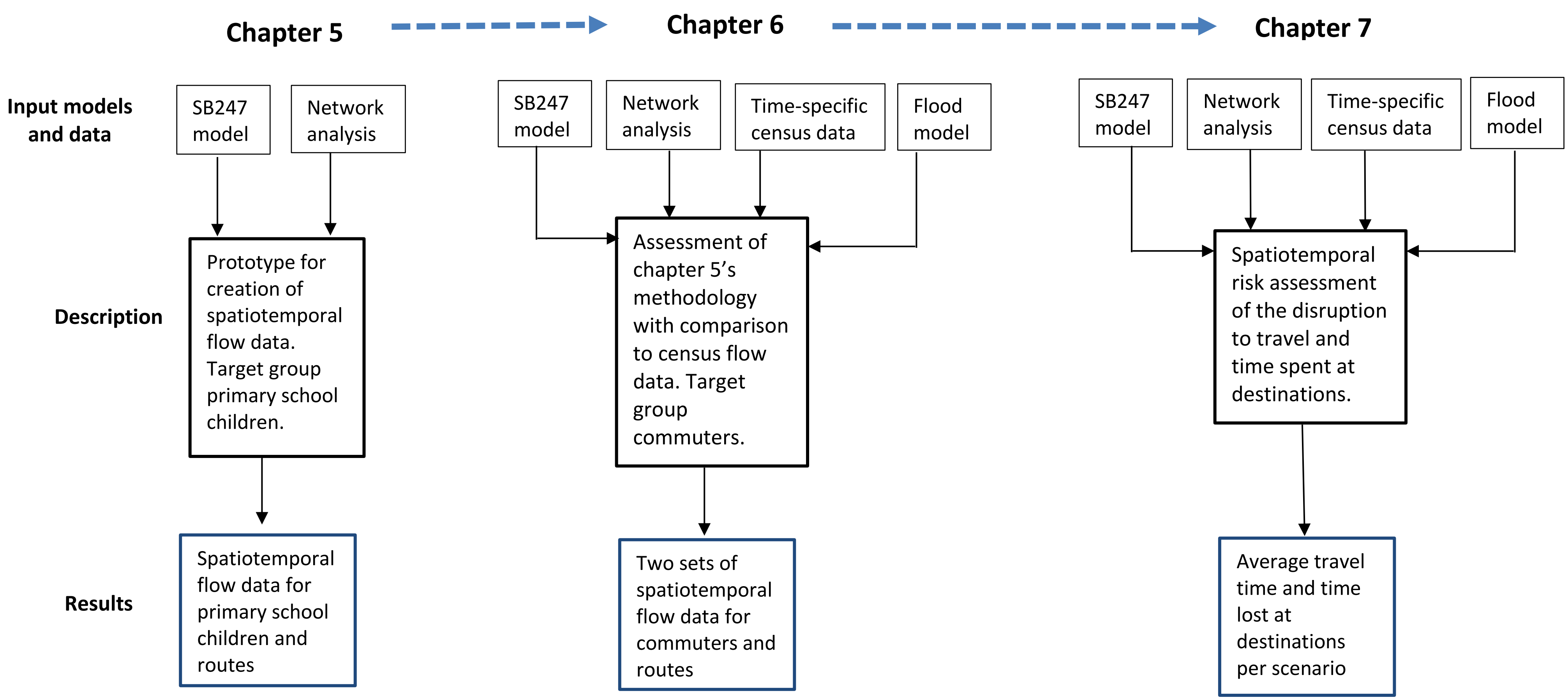


Figure 4.5 Overview of the analysis structure of the thesis and the interaction of the three results chapters. The top row of boxes are the input models, the middle row the description of the analysis chapter and the bottom row the results. The work of one chapter build into the analysis of the next chapter.

4.3.2 Modelling implementation by chapter

Chapter 5 is the initial prototype of a spatiotemporal travel methodology (Figure 4.6). Chapter 6 will be applying the prototype to commuters (Figure 4.7). Chapter 7 introducing flood risk scenarios (Figure 4.8). Due to the substantial number of possible scenarios from combining models, only a few scenarios tend to be chosen to illustrate research objectives (Smith et al., 2014a). Hence, in this research a couple of flood scenarios are selected for testing as it would be beyond the scope of this study to investigate all likely scenarios. Instead, the merit here it to evaluate and demonstrate the transferability of the approach developed, using a set of reasonable exemplar scenarios.

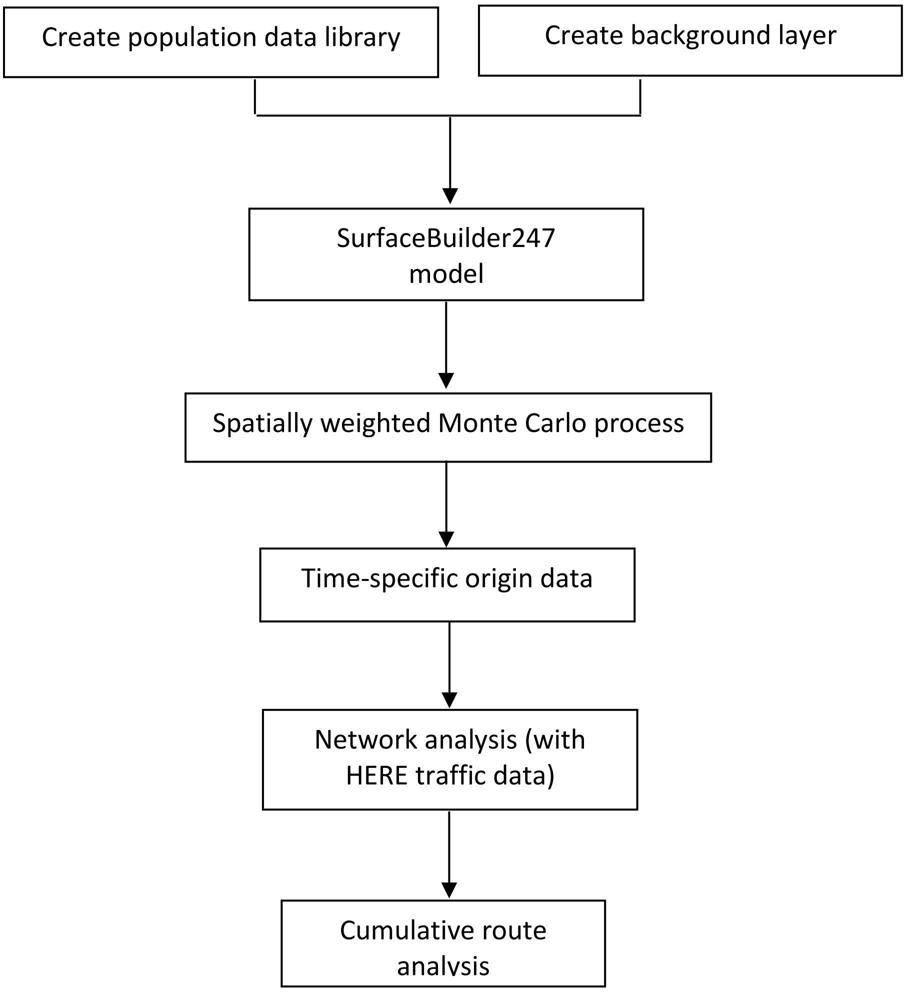


Figure 4.6 Chapter 5 overall methods structure

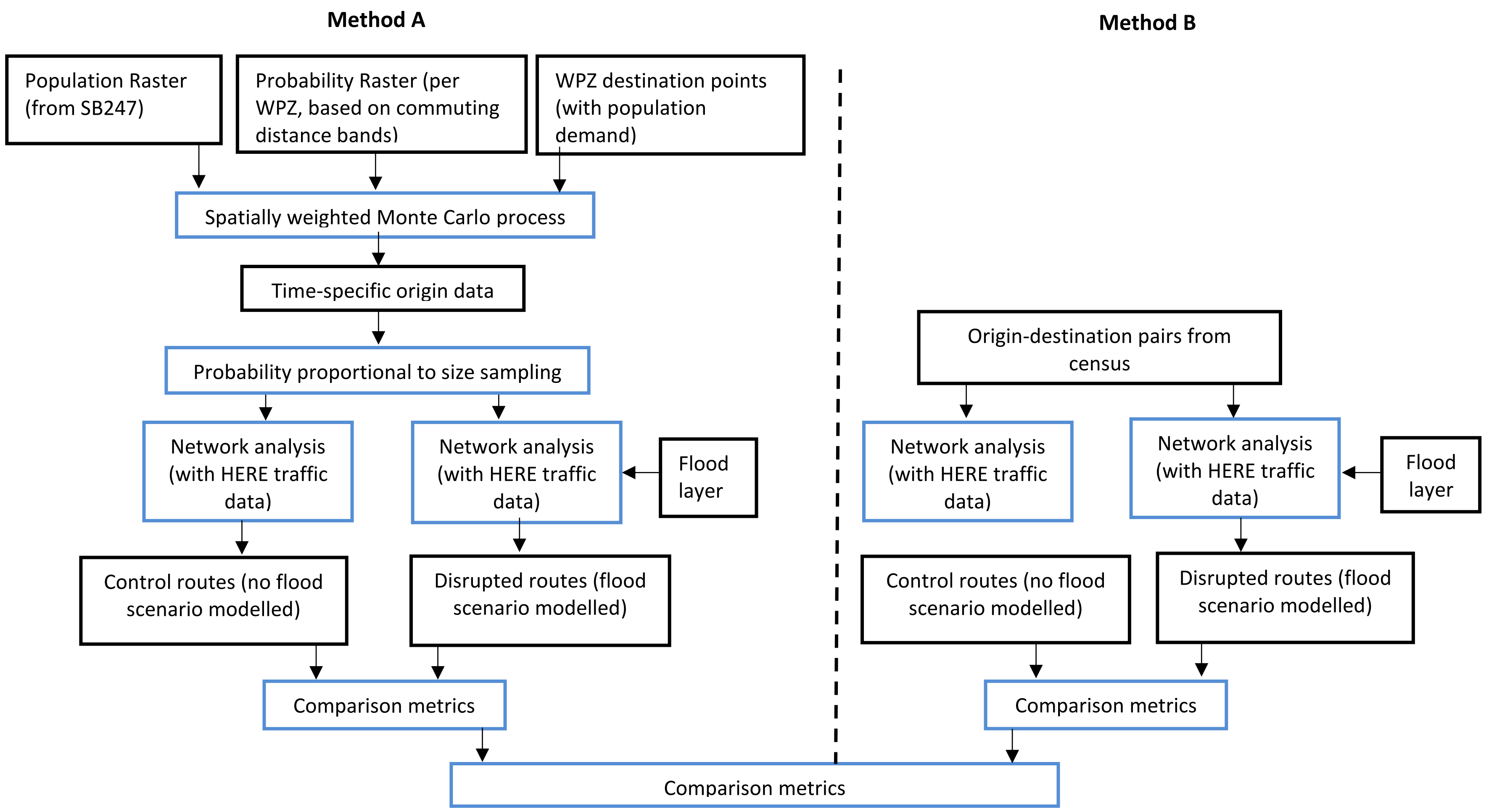


Figure 4.7 Methodological approach of Chapter 6.Blue boxes are analysis processes, black boxes are data

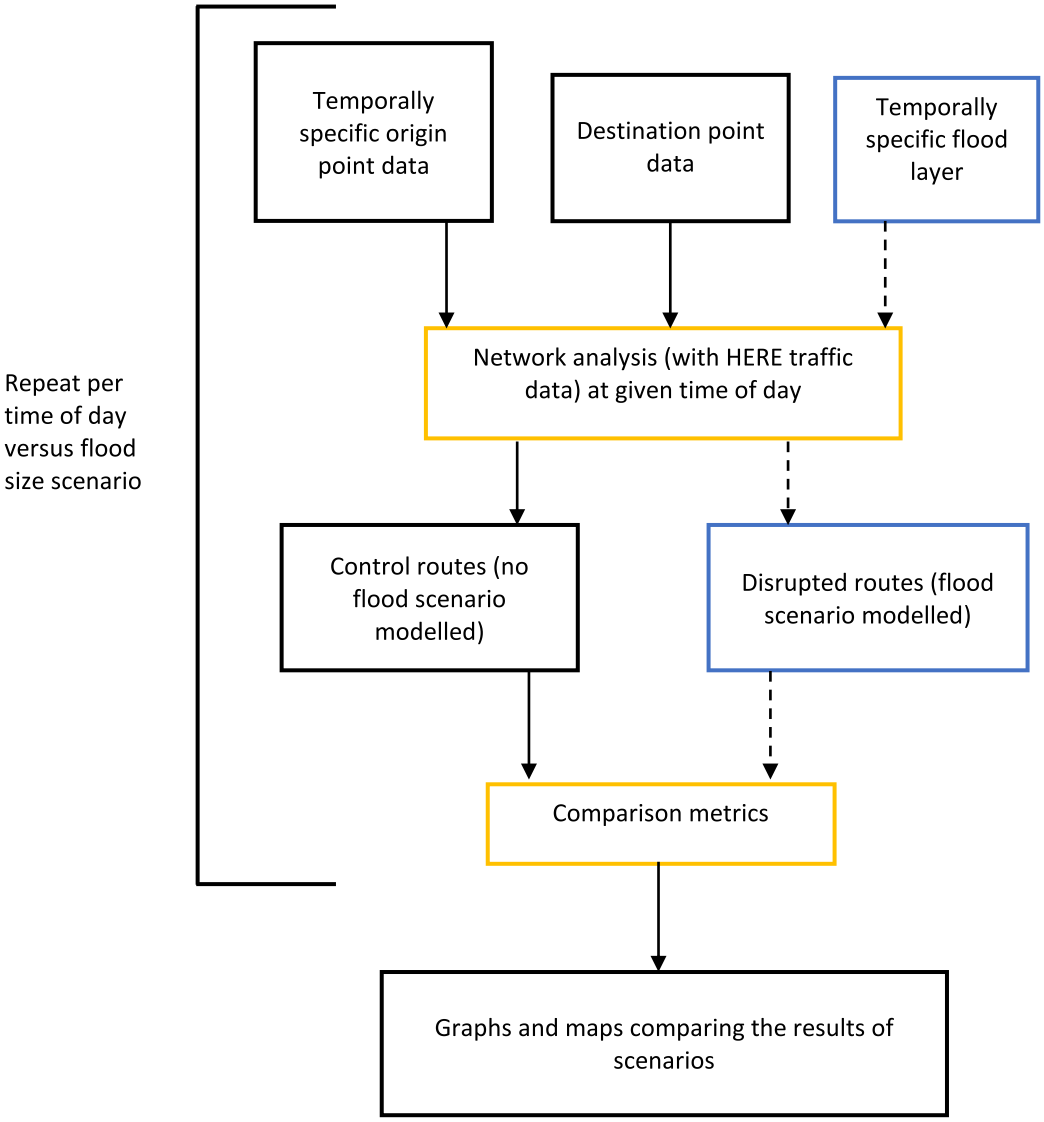


Figure 4.8 Overall methods structure for Chapter 7

Chapter 5 Modelling Spatiotemporal Population Flows

5.1 Introduction and Background

Accurate location information on population groups underpins numerous research fields including health geographies, retail planning, political science and emergency planning. Traditionally, census data for usual place of residence and place of work for the study population have been used in research (Aubrecht et al., 2013a; Smith, 2015; Debionne et al., 2016). However, this does not represent the true dynamic nature of population mobility (Aubrecht et al., 2014; Bian & Wilmot, 2015; Renner et al., 2018). Most people make journeys every day, be that from homes to places of work, schools, medical facilities or leisure and retail activities. The average number of trips undertaken per person in 2016 in England was 774 (Department for Transport, 2017), or just over two per day, with 62% of these by car (Department for Transport, 2017). There has also been an increase in commuting distances as people are living further away from their place of work (Bates, 2015), and this 'hypermobility' is not captured well within traditional population data sets. Over the last decade, administrative data sets and data from emerging technologies have become more important as they can provide detailed information between census years (Einav & Levin, 2014). For example, school enrolment data can capture the numbers of children travelling to a school and the increased use of smartphones amongst the population provide the opportunity to monitor footfall through shopping areas (Malleson et al., 2018). Data of this nature, when used in conjunction with census data, enables a richer understanding of population travel.

Journeys are individual decisions, but the data provided for them are often aggregated in time and space. Whilst administrative data can rectify some of the temporal issues with census data, there are spatial issues surrounding the aggregation of data for both census and administrative data sets. For example, in England the census is collected at an individual level but aggregated for spatial analysis into output areas in order to protect privacy. As these areas are designed to include 40 households as the minimum (Office for National Statistics, n.d.), they are not spatially regular as densely populated areas will have output areas of a smaller physical size compared to rural output areas. Furthermore, between censuses there can be boundary changes meaning a household may be aggregated into a different zone in the 1991 or 2001 census compared to the 2011 census, reducing the temporal stability of data from census zones. Spatiotemporal gridded population data provides a way of combining census boundaries with other administrative geographies, as well as addressing the spatiotemporal stability of aggregation zones. As described in Table 3.2, there are a number of spatiotemporal gridded population projects and data sets, some of which are for specific purposes and others for global use. These provide an alternative form of population data and are easier to link to other data which typically use a raster format, for example environmental data. Some of these gridded population models produce data to represent population at different times of day (LandScan (Bhaduri et al., 2007), Population 24/7 (Martin et al., 2015), Dynapop (Aubrecht et al., 2014)). These data have led to valuable research in infectious disease spread (Strano et al., 2018), natural hazards (Renner et al., 2018; Smith et al., 2015; Aubrecht et al., 2014), population estimates in developing countries (Wardrop et al., 2018; Weber et al., 2018), maternal health inequalities (Ruktanonchai et al., 2016) and migration (Tatem & Smith, 2010).

One instance where the location and travel patterns of a population is key knowledge is natural hazard planning. As discussed in Chapter 2, it has been shown that the time of day when a natural hazard event occurs is a large determinant of the scale of impacts (Aubrecht et al., 2012). Within natural hazard literature, certain population sub-groups have been identified as particularly vulnerable, including children, the elderly, disabled persons and tourists (Cutter et al., 2003; Civil Contingencies Secretariat, 2008). Knowing how these vulnerable sub-groups move throughout the day is therefore of particular interest, as it enables risk mitigation measures to be specifically tailored towards them. The availability of spatiotemporal gridded population data which identifies vulnerable population sub-groups opens the potential for their movement and travel to be studied.

However, spatiotemporal gridded population data do not show the routes taken by individuals to travel across an area, rather they show how the number of people in each grid cell changes between time periods and where those 'in travel' are likely to be. Vector-based analyses, like network analysis in GIS, have been used to generate journeys from census output area centroids for example, Comber et al., (2008); Langford & Higgs, (2010); Higgs et al., (2012). This provides a reasonably fine spatial resolution and has demographic data attached, but it is only valid for the time the census was taken. Alternative data sources like Smart Card travel data have been used to model individual journeys across urban areas (Han & Sohn, 2016; Wang et al., 2017) which does capture individual spatiotemporal dynamics, however this does not provide detailed demographic information on the travellers. Linking spatiotemporal gridded population data to travel models enables the demographic data from the census to be preserved, but also allows a temporal resolution more similar to that used with transport software, combining social science research with transport planning. Therefore, the aim of this chapter is to assess how spatiotemporal gridded population data can be linked to existing GIS network analysis tools in order to provide more detailed insights on population travel. As far as the author is aware, there is no literature expressly discussing how to improve the representation of travel from spatiotemporal gridded population data and linking it to vector-based network analysis, making this research novel. The first three objectives of the thesis are fulfilled in this chapter:

Aim: To develop a framework for combining spatiotemporal gridded population data with GIS network analysis, using journeys to primary school as an example.

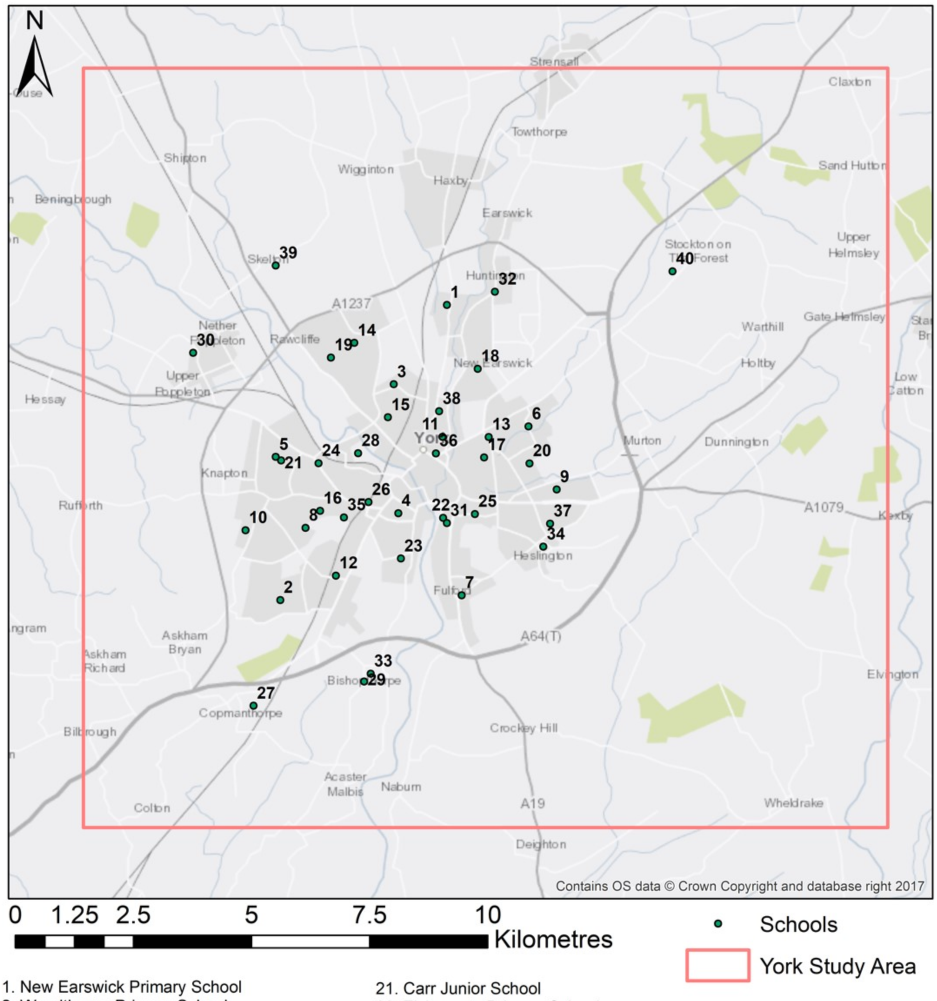
Objectives:

1. Create a baseline spatiotemporal representation of primary school children's location across the study site, including the in-travel population.
2. Evaluate the use of spatiotemporal gridded data for representing population in travel
3. To develop a method for converting the output from the SurfaceBuilder247 model into origin-destination flows for use in GIS network analysis

In this chapter, primary school children (aged 4-11 years) are the group which have been selected for study in order to develop the methodological framework. This group has been chosen because they are a 'data rich' population group (Harland & Stillwell, 2010) with information on daily travel patterns during the working week available. Education is compulsory for all children in this age group, therefore the vast majority of children age 4-11 will travel to a school and back again on a typical term-time day. Whilst the relative numeric changes in the primary school population can be modelled over space and time through the existing spatiotemporal framework in the SurfaceBuilder247 model, the specific routes individuals take across space is much more complicated to model. In England, school admissions were traditionally arranged through school catchments, designated geographic areas for each school based primarily on distance to school, sibling attendance and religious affiliation, where the children who met these criteria attended the school (Burgess et al., 2011). However, the introduction of the 1988 Education Reform Act, later strengthened by the 2006 Education and Inspection Act, has enabled parental choice in school selection and made it simpler for children to attend schools outside of their catchment area (Burgess et al., 2011). The increase in parental choice, along with the creation of academies (Academies Act 2010), has led to progressive marketization of school choice, making published school catchment areas unreliable as representations of student home location. Parsons et al., (2000) found 35% of children in their study area did not go to their catchment school and this is likely to be a larger proportion now through the introduction of parental choice measures. Therefore, an alternative method is needed to calculate where children travel to school from.

5.1.1 Study Site

As described in Section 4.1.1, York is the location used for this research. Within the York study area, there are 40 state primary schools which educate pupils age 4 to 11 years (Figure 5.1). Independent schools are not included as there is no publicly available information on their pupils' home addresses are. However, as pupils educated at independent schools account for 7% of school children (Independent Schools Council, 2017), they represent a minority of children.



1. New Earswick Primary School
2. Woodthorpe Primary School
3. Burton Green Primary School
4. Scarcroft Primary School
5. Carr Infant School
6. Hempland Primary Academy
7. St. Oswald's C of E Primary School
8. Hob Moor Community Primary School
9. Osbaldwick Primary Schools
10. Westfield Primary Community School
11. Park Grove Primary Academy
12. Dringhouses Primary School
13. Heworth C of E Primary Academy
14. Lakeside Primary School
15. Clifton Green Primary School
16. Acomb Primary School
17. Tang Hall Primary Academy
18. Yearsley Grove Primary School
19. Clifton with Rawcliffe Primary School
20. St. Aelred's Roman Catholic Primary School
21. Carr Junior School
22. Fishergate Primary School
23. Knavesmire Primary School
24. Poppleton Road Primary School
25. St. Lawrence's C of E Primary Academy
26. St. Paul's C of E Primary School
27. Copmanthorpe Primary School
28. St. Barnabas C of E Primary School
29. Archbishop of York's C of E Junior School
30. Poppleton Ousebank Primary School
31. St. George's Roman Catholic Primary School
32. Huntington Primary Academy
33. Bishopthorpe Infant School
34. Lord Deramore's Primary School
35. Our Lady of Queen of Martyrs Roman Catholic Primary School
36. St. Wilfrid's Roman Catholic Primary School
37. Badger Hill Primary Academy
38. Haxby Road Primary Academy
39. Skelton Primary School
40. Stockton-on-the-Forest Primary School

Figure 5.1 Map of the location of the primary schools in York. Contains public sector information licensed under the Open Government Licence v3.0

5.2 Methods

In this chapter, the SurfaceBuilder247 model outputs for population in travel are compared to an extended methodology framework, based on the four-stage conceptual transport model (Ortuzar & Willumsen, 2011; Jones, 2012) (see Section 3.3.2.1 for a full description). The SurfaceBuilder247 model is described in Section 4.2.2 and a full explanation of the theory is given in Martin et al., (2015). The model's implementation in the context of the York case study will be described here. To allow the spatiotemporal gridded population data from SurfaceBuilder247 to be used in network analysis, a spatially weighted Monte Carlo procedure was designed to produce statistically likely origins for each primary school destination. This is described in Section 5.2.2.

5.2.1 Application of SurfaceBuilder247 software

Briefly, SurfaceBuilder247 model produces gridded data layers which represent the number of people present in each cell at a given point in time, whether they are at an origin, destination or in travel. The model was used to create a series of spatiotemporal population grids for all population sub-groups at the times of day which were of interest in this study. The data produced for the population sub-groups who attend primary school (0-4, 5-10 and 11-17 groups) are used in this chapter to develop spatiotemporal population flow data.

The origin data for the residential population was generated as described in Section 4.2.2. Destinations in this Chapter are state run primary schools (Figure 5.1) and a point data set was generated of each school's location with the number of pupils in attendance taken from local authority data (City of York Council, 2020). Determining a catchment area from which population can be selected by the SurfaceBuilder247 model and the spatially weighted Monte Carlo process for each school is challenging, due to the complications of school admissions (Section 5.1). A 'de facto' catchment can be defined as the radius from a school where a given percentage of the pupil intake was drawn from in the previous year (Burgess et al., 2011). This provides a reasonable estimation of where pupils are travelling from and can be used as the wide area dispersion. Edubase data for the Euclidean radii encompassing 95% of the children from the last intake was utilised, extracted from Locrating (2017) data used was for the 2017 school intake, the closest available data to the target year. Schools collect data on their pupils home location and mode of travel to school which is stored in the National Pupil Database (Department for Education, 2018). Using these data would provide more realistic catchments than the catchment radii used in this chapter. However, there are several reasons why the generalised catchment radii are appropriate in this instance. Firstly, the radii fit well with the existing wide area dispersion function in the SurfaceBuilder247 model. Secondly, there are strict data protection measures when using data from the National Pupil Database, so uploading such data to ArcGIS Online for the routing calculations would be in breach of these terms. Finally, the intention is to use the methodology developed for groups other than school children where individual data on origins are not available, therefore this method is preferable as it is more easily transferable.

The Population24/7 NRT project (UK Research and Innovation, n.d.) updated the methodology for creating the background layer, a key data input to the SurfaceBuilder247 model which determines the location population in travel are distributed to. Therefore, the background layers used in this thesis were updated in line with this project. To summarise, the main updates were as follows. Background layers for a specific day type e.g. normal working weekday, bank holiday Monday, Saturday's in December, were able to be created due to Traffic England's National Transport Model time period data becoming available to support this (Department for Transport, 2018a). Minor roads were also introduced to give a more detailed spatial representation of those in travel around York.

To create the updated background layers, Annual Average Daily Flow (AADF) data for the target year (2016) was downloaded, which provides the average number of each vehicle type which passes a count point each day for that year. The count point data were then combined with the National Transport Model data which denotes the day type and time period (Department for Transport, 2018a). The count point data were then classified by road type and whether it was urban or rural. Finally, for each time period the count point data are interpolated using Inverse Distance Weighting to produce a gridded surface of the number of people per grid cell for each type of road, i.e. motorways and A roads, based on the OS open roads data set. The separate raster files were overlaid to generate one background layer at a 200m resolution. As Figure 5.2 shows, the SurfaceBuilder247 output for the population in travel is noticeably different with a greater spatial dispersion of the 'in travel' population reflecting the inclusion of minor roads.

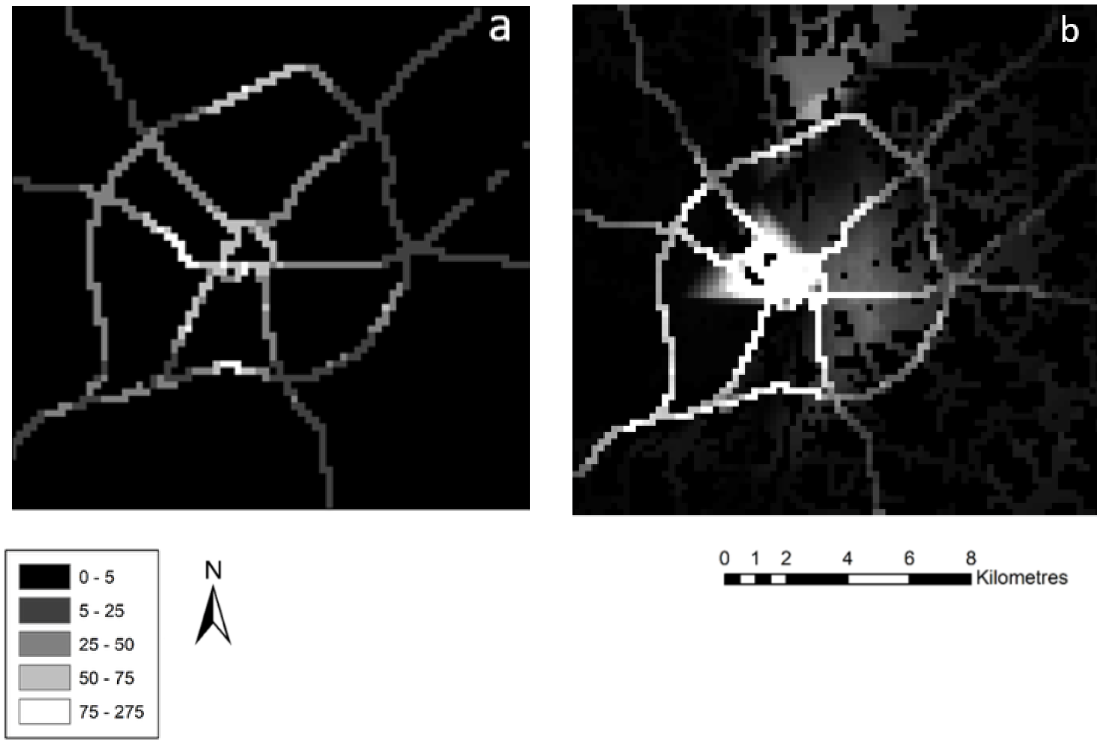


Figure 5.2 Distribution of the total population 'in travel' with a) the original background layer and b) the updated background layer. Contains public sector information licensed under the Open Government Licence v3.0

As described in Section 4.2.2, a time profile for each type of destination and day type is required. For schools, nobody is present during the early morning and between 0800-0900 hours the destination fills up with school children. They remain on site until 1500 hours when they start to leave and return to their home destinations. Schools should be at near full capacity as children registered at the school must attend, although there will be some absences due to illness and other events. SurfaceBuilder247 therefore seeks to fill 94% of the school capacity with the correct population age groups and there should not be unfilled places at the destination.

5.2.2 Adapted four-stage model

The 'four-stage model' is a classic conceptual model in transportation engineering (see Section 3.3.2.1). This conceptual model was chosen as the foundation for this chapter's methodology for its simplicity and adaptability (Ortuzar & Willumsen, 2011; Jones, 2012), and because it is well known within transportation research, forming the basis of the UK government's National Transport Model (Thiessen, n.d.). McNally (2007) describes this model as approaching transport planning from a 'person-perspective', which makes it suitable for combining with spatiotemporal population data. Figure 5.3 illustrates the adapted four-stage model proposed in this chapter, henceforth known as the four-step approach.

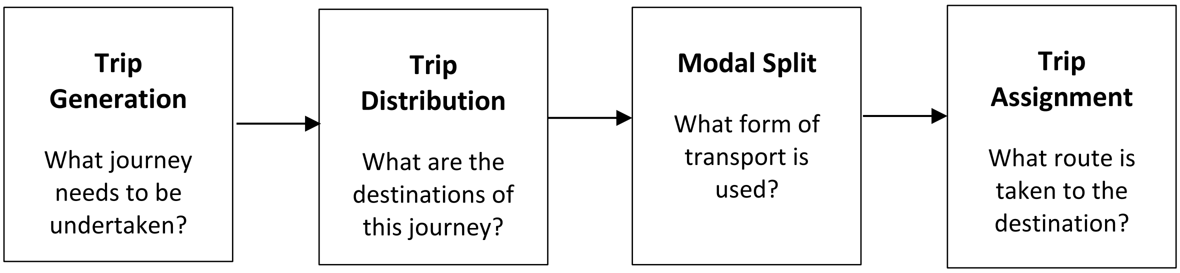


Figure 5.3 The four-stage transport model

The **trip generation** stage, defining the number of people who need to travel, is fulfilled with the spatiotemporal population grids for the target group generated by SurfaceBuilder247. Here a spatiotemporal population grid for a time period before travel occurs e.g. 6 am is used, indicating the number of people available in each grid cell able to travel.

Trip distribution identifies the destinations to fulfil the travel requirements from the trip generation stage. As SurfaceBuilder247 does not indicate where individuals have travelled to and from, an alternative model is required. One solution would be to create a spatial interaction model, as these are designed to calculate the flows of people between origin and destinations and are widely used, including for school travel (Harland & Stillwell, 2010). However as discussed in Section 5.1, 'de facto' catchments are being used rather than individual data. An alternative solution was required to fulfil the trip distribution stage. An original spatially weighted Monte Carlo simulation was created in R (Nieves, 2018, pers. comm.), using computer cluster techniques from Stevens et al., (2015) and the WorldPop research group at Southampton University (World Pop, 2020). The Monte Carlo simulation code selects individuals to fulfil demand at school destinations, based on the probability of a grid cell providing children to that destination, until all the schools are full (Figure 5.4). Each school has a probability raster, similar to Sofianopoulou et al.,(2012) and Nieves (2015), based on the primary school catchment areas used for wide area dispersions in the SurfaceBuilder247 model, maintaining consistency through the framework. It assigns greatest probability to those grid cells that fall within the radii of where 95% of the intake are from and all grid cells outside of this are assigned a proportion of the remaining 5% of the probability. The other input data into the spatially weighted Monte Carlo simulation are the gridded population data for each age group, and a shapefile of the school destinations with the number of children in each age group it requires. The number of children was 94% of the total school places, introducing the assumption that there are some absences, as defined in the SurfaceBuilder247 model.

As described in Figure 5.4, the simulation takes one school destination at a time and assigns individuals to it using the probability raster for that school. Once an individual has been selected from a grid cell they are no longer available to any other school destination to ensure that the population volume is preserved. In urban areas like York, there is a large overlap of school catchment areas (Figure 5.5). This has been found in other studies, for example Burgess et al., (2011) found there was an average of 19 schools within a 3km radius of a pupil in an urban area. Consequently, the order in which the schools are run through the simulation is of great importance as those at the beginning have the greatest available population to select from, and those at the end the least. This could mean that schools with smaller catchments have not enough people available if other school destinations have already selected the population. The schools were therefore ranked by the area of their school catchment and the simulations run starting with the school with the smallest catchment and descending to the one with the largest. This meant the schools with the greatest geographic constraint had an available population.

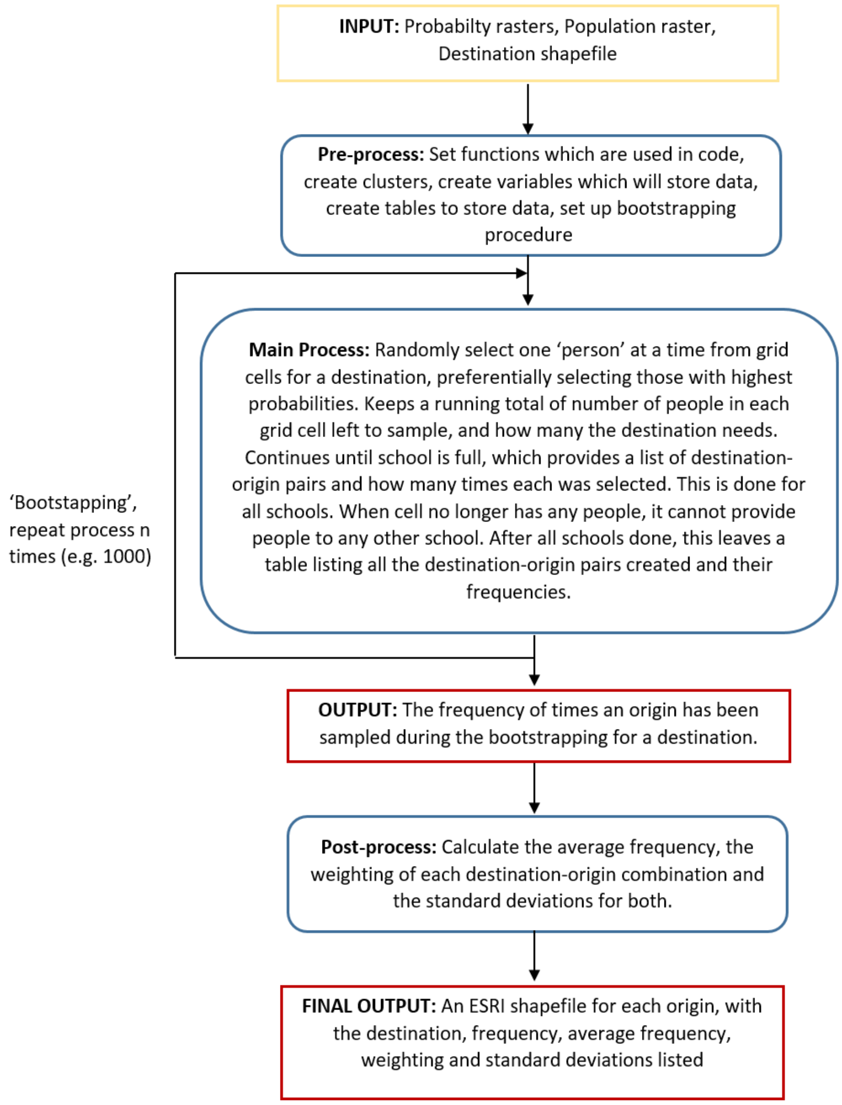
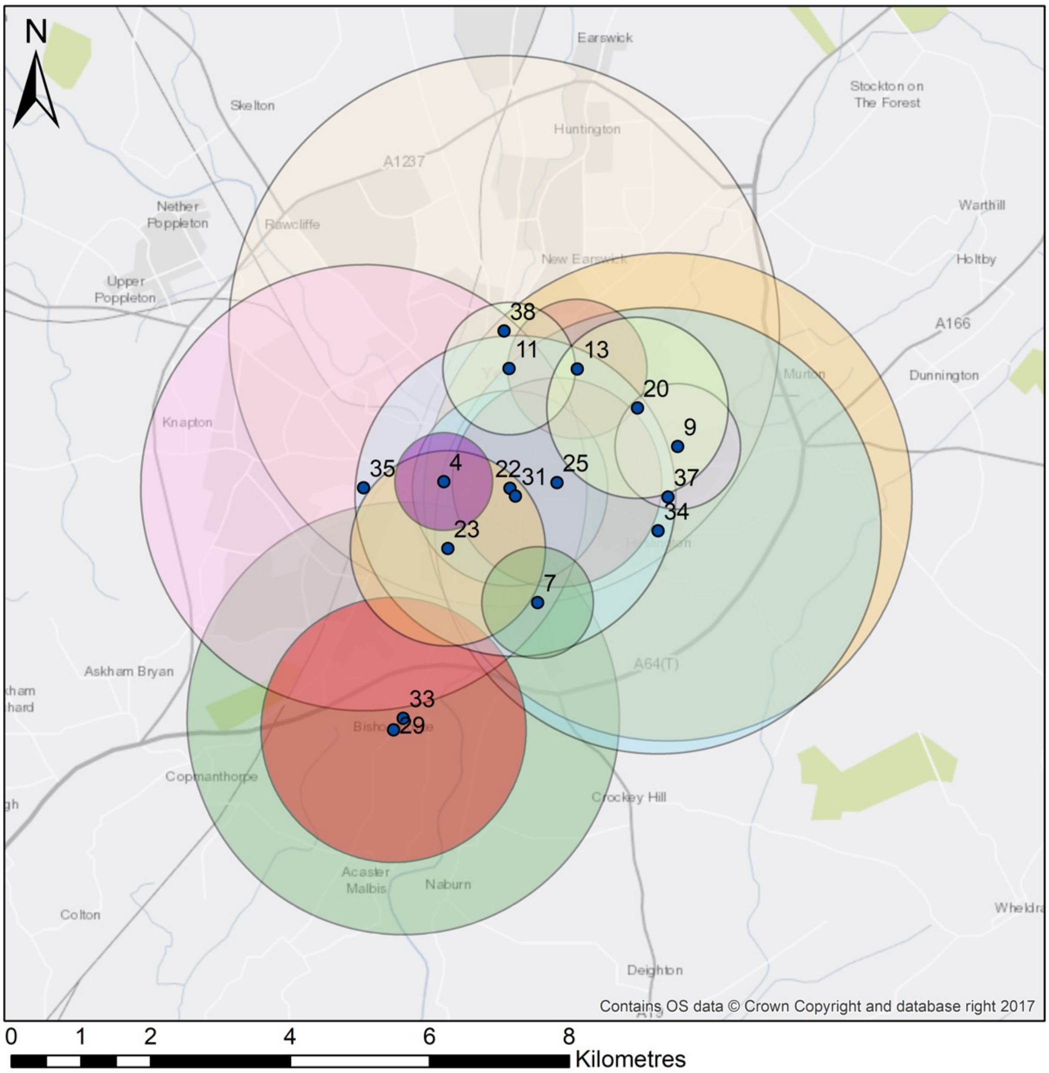


Figure 5.4 Flow chart of the procedure of the spatially weighed Monte Carlo process applying a bootstrapping technique.



4 Scarcroft Primary School
7 St Oswald's C of E Primary School
9 Osbaldwick Primary School
11 Park Grove Primary Academy
13 Heworth C of E Priamry Academy
20 St Aelred's Roman Catholic Primary School
22 Fishergate Primary School
23 Knavesmire Primary School
25 St Lawrence's C of E Primary Academy
29 Archbishop of York's C of E Primary School
31 St George's Roman Catholic Primary School
33 Bishopthorpe Infant School
34 Lord Deramore's Primary School
35 Our Lady Queen of Martyrs Roman Catholic Primary School
37 Badger Hill Primary School
38 Haxby Road Primary Academy

Figure 5.5 School catchment overlap for selected primary schools in York. Contains public sector information licensed under the Open Government Licence v3.0

A bootstrapping process was applied to run the simulation 1000 times for each school, balancing the requirements for statistical accuracy and constraints of computing power. A bootstrap process is a random sampling method with replacement. It works by randomly selecting data points to fulfil a sampling requirement and noting the data value. Over many repetitions, an average of the small samples of the data allows estimation of the characteristics of a population (Brownlee, 2018). In the application in this thesis, the data are the origin population grid which is sampled to generate an average of where people travel to the destination from.

The output of the spatially weighted Monto Carlo process was an origin shapefile for each destination, which specifies the frequency of selection and a statistical weighting reflecting this frequency for the origin destination pair. The average frequency also allows the average number of children assigned to that school in a simulation from the grid cell to be calculated. The weighting of the origin-destination pairs was used as the first step in selecting only origins which were more frequently selected. A plot of the origin weights from largest to smallest showed a long tail end in the distribution, and these origins which all had low weightings were removed from further analysis. Origin-destination pairs averaging less than 1 child over all simulations were excluded. Upon review of the data, for the 0-4 and 11-17 age groups it was not possible to restrict origins to just those providing at least 1 child. As the population required at each destination was small, in large catchment areas the origins were only fractions of children e.g. 0.34. In these cases, 0.1 was used as threshold and the sum of the origins was checked against the destinations, again finding a close match to the total required.

Modal Split defines the method of transport used for journeys. Since equivalent data were unavailable locally for York, a cross-tabulation of school travel mode by distance band for England (Table 5.2) was used to estimate modal split (Department for Transport, 2014), based on the Euclidean distance between each origin-destination pair. The origin-destination pairs were split into two categories, motorised and non-motorised travel. Motorised travel included those who travelled by car and bus, non-motorised was walking and cycling. This is due to travel by public transport and bike being a very small percentage of primary school children, so they can be subsumed into categories with cars and walking respectively. This classification was also used by Easton & Ferrari (2015).

Table 5.1 Travel mode by distance band for Primary schools in England

Classification	Distance band	Motorised	Non-motorised
1	Under 1 mile (up to 1.6km)	20%	80%
2	1 to under 2 miles (1.6-3.2km)	72%	28%
3	2 to under 5 miles (3.2 up to 8km)	99%	1%
4	5 miles and over (over 8 km)	100%	0%

Trip assignment was fulfilled by ArcGIS Online's network analysis function. The online version of the software package was used as it allows for mode of transport to be easily specified and the historic traffic for a time of day to be included in driving times. The historical traffic data are based on the average travel speeds of cars over the last three years, recorded in 15-minute intervals (ESRI, 2018), and allows for time of day to be considered throughout the framework. For each school destination, the travel times and distances for the walking origins and driving origins were calculated for the three population groups; 0-4 years, 5-10 years, 11-17 years. The result were shapefiles for each mode, origin destination pair and age group which represented the travel time or the shortest route at the time specified.

5.2.3 Comparison of population in travel modelled via two approaches

To address the second objective, the results of the four-step approach were transformed into a raster data set showing the number of children in travel at 8:15am comparable to the output from SurfaceBuilder247. To consolidate the route information, maps were produced to show the number of children who use each road. This identifies spatial patterns for which roads are used most in allowing children access to school.

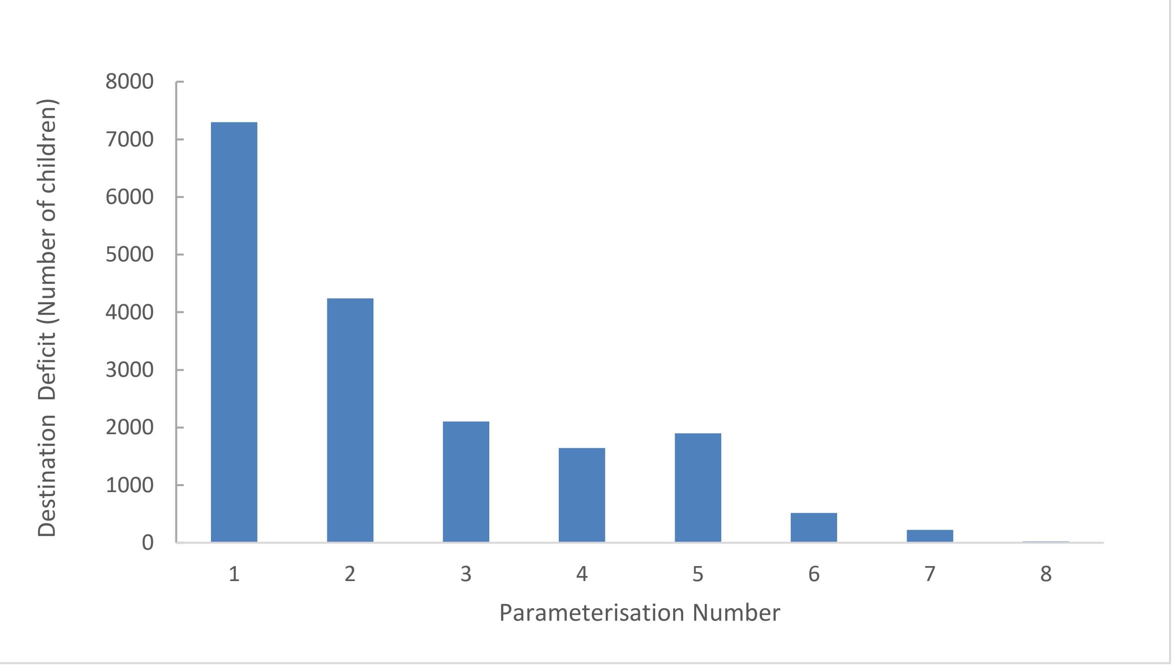
5.3 Results

In this section, the results from the development of the four-step approach for linking gridded spatiotemporal population data to vector-based network analysis and the network analysis route results will be presented.

5.3.1 Outputs from preliminary population models

The initial results from the SurfaceBuilder247 model did not succeed in capturing the dynamics of school travel as there were 4240 unfilled school places in York's primary schools (Figure 5.6 and

Figure 5.7) which could not be met from origin populations, an 11% deficit. With adjustment of the model parameters and the input data, the deficit reduced through successive reparameterisations of the models as shown in Figure 5.6. Figure 5.7 shows that there was a spatial nature to those schools with unfilled places, as in V1 the majority of the schools were on the South East side of the city, and V3 and V5 the two schools were within 1km of each other. The change between V1 and V3 is the greatest and there is a change in which schools have unfilled places.



Parameterisation	Notes
1	First model run (with MYE)
2	Adjusted MYE values in origins
3	Have school population split as 5% 11-17 and 95% 5-10
4	Use the York Local Authority school population data for target York schools
5	Expanded study are and buffer zone so they include the catchment radiuses for all schools
6	Have school population split as 7% 0-4, 86% 5-10 and 7% 11-17
7	Have school population split as 10% 0-4, 83% 5-19 and 7% 11-17
8	Have school population split as 13% 0-4, 80% 5-10 and 7%11-17

Figure 5.6 Graph showing the model version and the unmet demand at primary schools in York. The table describes the changes made in each successive parameterisation of the model version.

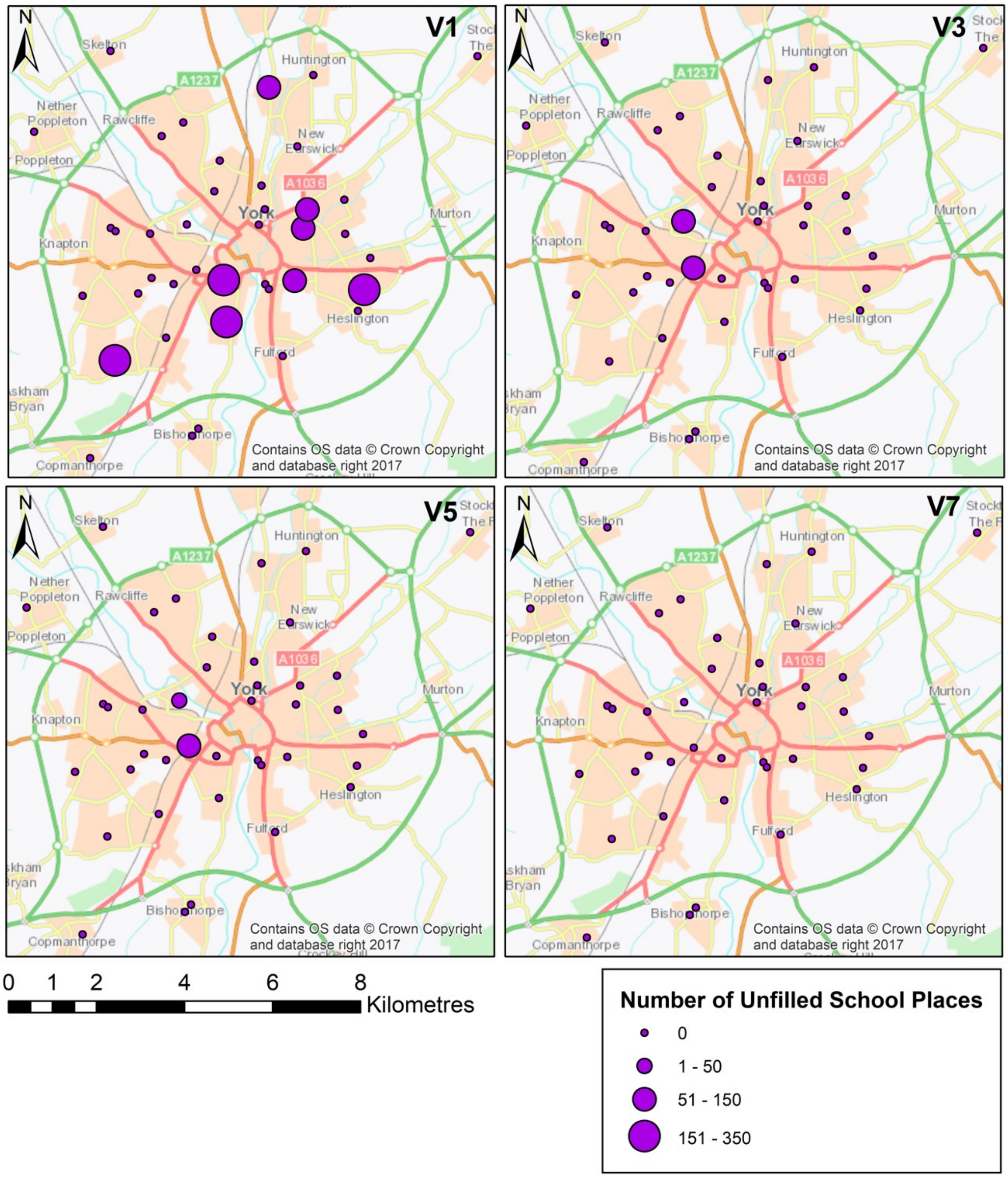


Figure 5.7 Maps showing the location of primary schools in York and the number of unfilled demand at the school. The version number is denoted by the Vx in the top right-hand corner of each map. Contains public sector information licensed under the Open Government Licence v3.0

5.3.2 Comparing SurfaceBuilder247 output to the four-step approach output

Objective 2 was to compare the outputs of SurfaceBuilder247 and the four-step approach to see how they represent population in travel. Figure 5.8 shows that the SurfaceBuilder247 output has a high weighting of school children in travel on the main A roads around York and the arterial roads into the city centre. The results of the four-step approach show a greater use of minor roads, with a large proportion of children not using the main ring road for travel. There is also a greater homogeneity in the number of people per grid cell with a more even spread across the city, particularly for walkers (C).

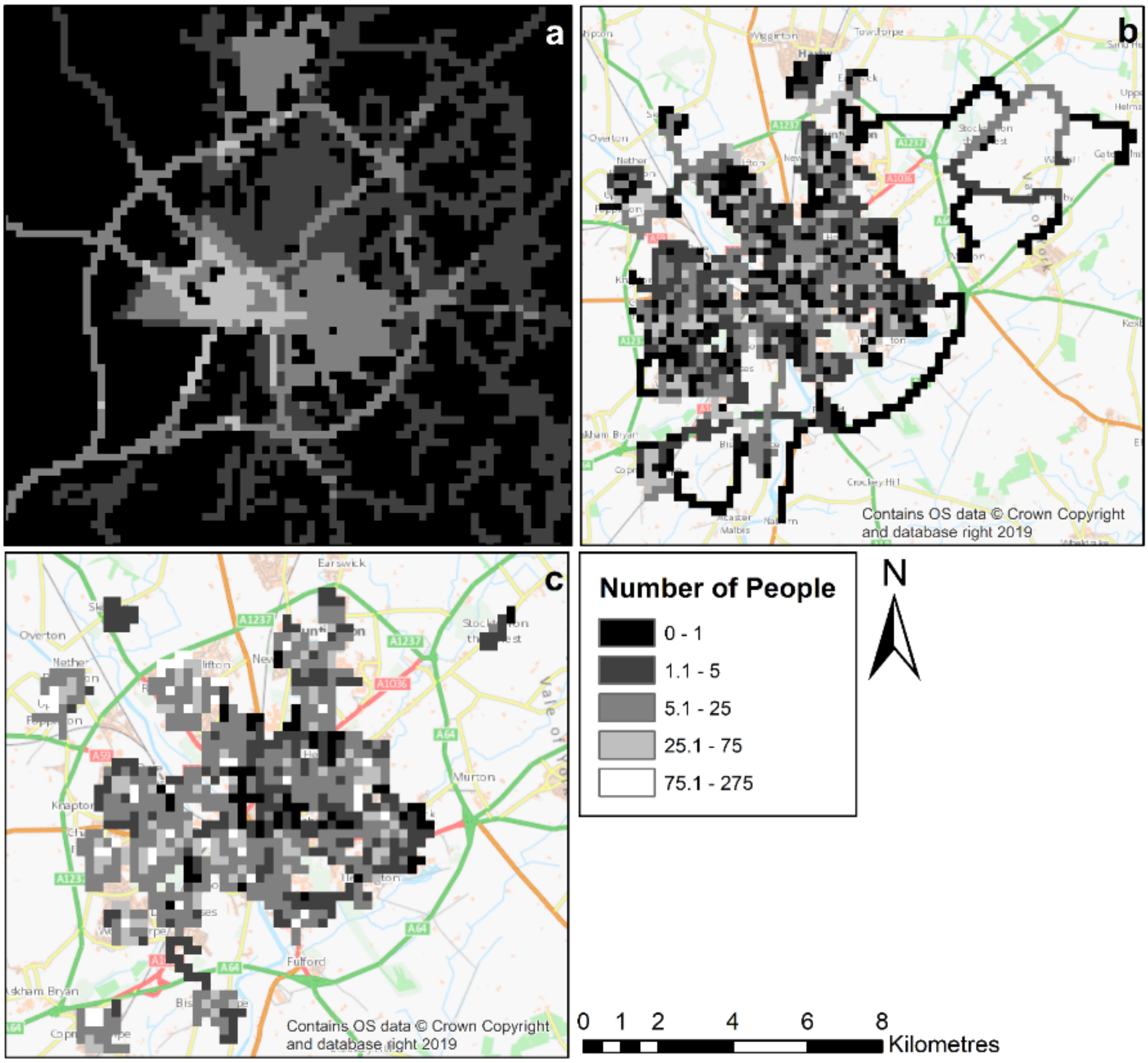


Figure 5.8 Comparison raster maps of the primary school population in travel at the same time point across York. a) is the output from SurfaceBuilder24/7, b) the output of the network analysis for those travelling by car and c) the output of the network analysis for those walking. Contains public sector information licensed under the Open Government Licence v3.0

5.3.3 Network Analysis Results

Figure 5.9 and Figure 5.10 present the results of the network analysis for school travel on a typical Monday morning, dividing road sections into quintiles based on the modelled number of children travelling along a road section. The roads used by those who walk are more clustered around destinations than those used by car travellers, particularly for schools in the villages on the edge of York (Figure 5.9). The concentration of children travelling on roads is also higher for walking routes, with 465 the greatest number on a section compared to 109 for car travel. There is also a difference between walking and motorised transport, with car journeys tending to be longer and utilising a greater proportion of the road network. In both forms of travel, it is local roads which are used more than major A roads. It can be inferred that the greater the number of children using a road, the more important the road is to school travel, so those in the highest quintile can be considered the most critical.

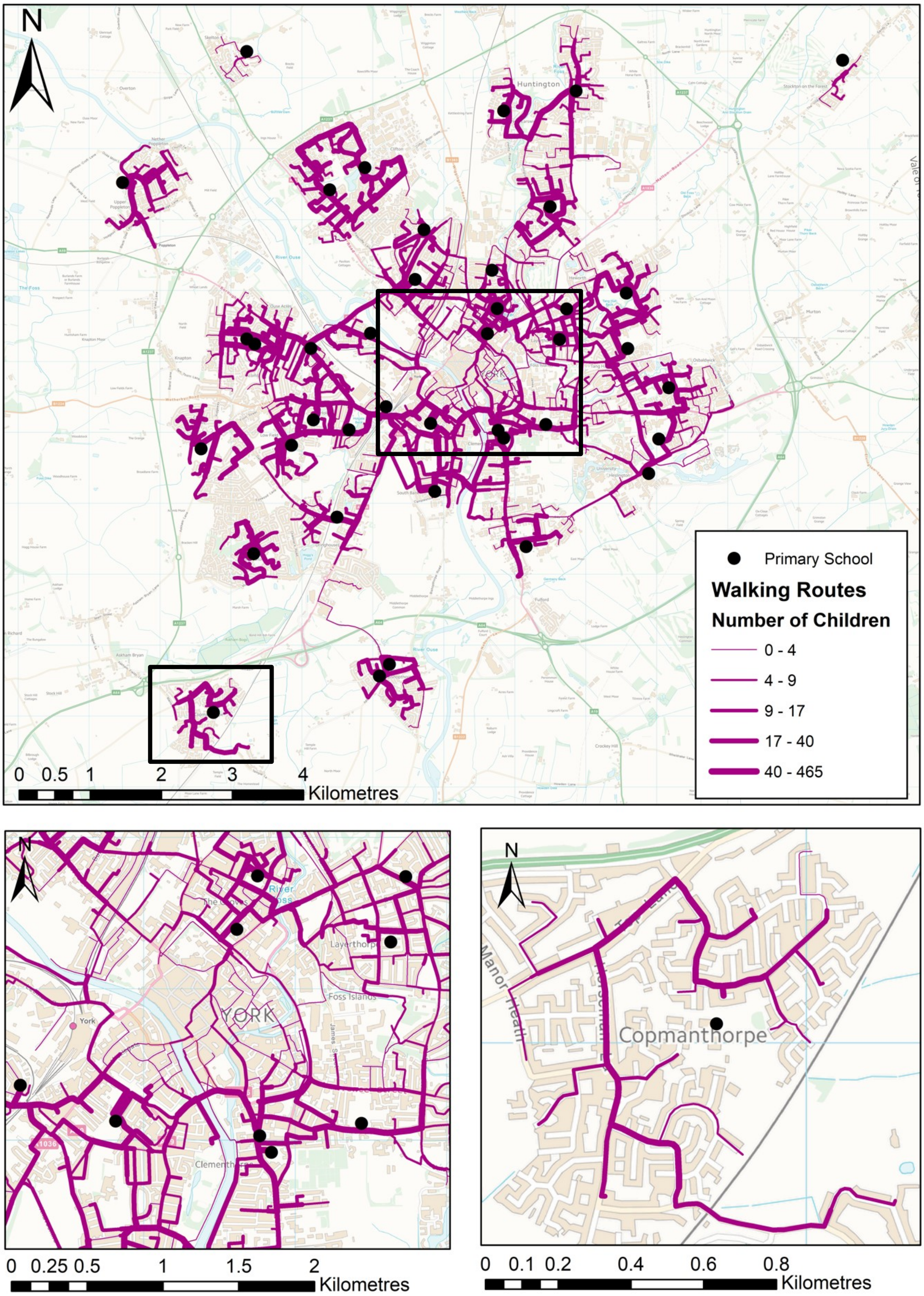


Figure 5.9 The cumulative number of children using each road segment for those walking to school. Inset image extents shown on main map. Contains public sector information licensed under the Open Government Licence v3.0

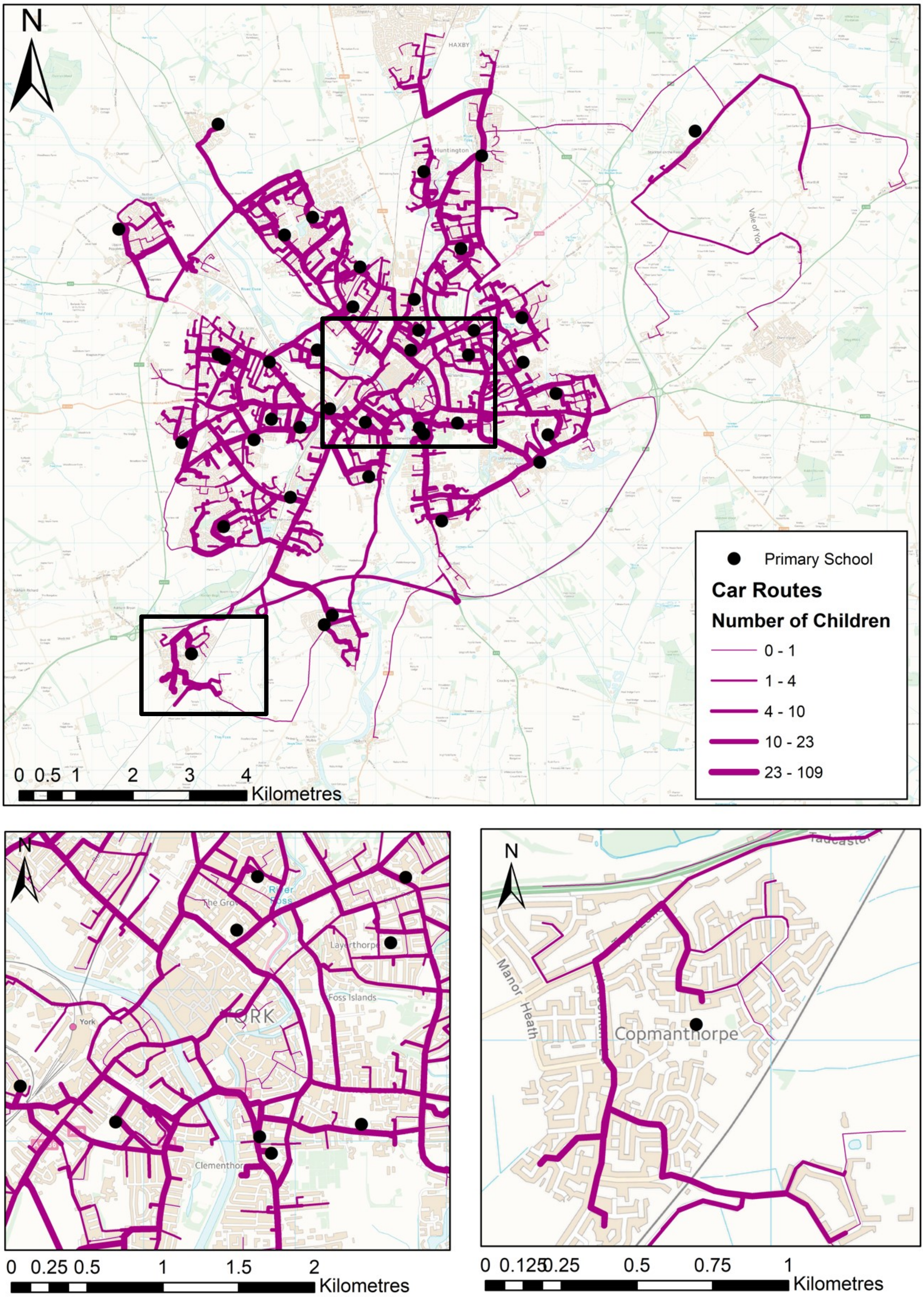


Figure 5.10 The cumulative number of children using each road segment for those being taken by car to school. Inset image extents shown on main map. Contains public sector information licensed under the Open Government Licence v3.0

5.4 Discussion

5.4.1 Configuration of SurfaceBuilder247 model

Primary school children are a unique group in the wider population, as the vast majority of people in this age group follow a well-defined routine during the weekdays (going to and from school). The SurfaceBuilder247 parameters mean that all school destinations should be filled on any working weekday. If there is unmet demand at a destination this is likely due to discrepancies between school administrative (destinations) and census data sets (origins). The initial parameterisation of the SurfaceBuilder247 model had 11% of the school destination places unfilled in the output grid, demonstrating the mismatch between the census data and the school data. There was also a spatial pattern to schools with unmet demand (Figure 5.7) as urban areas have a limited number of children with multiple destinations competing for them. By adjusting the model parameters and the input data, a parameterisation where all the demand in schools was met was achieved.

There are several factors which contributed to this discrepancy. Firstly, the proportional split in the numbers of children from each age group attending primary schools was revised. Between V1 and V2, adjusting the school population so that 5% were from the 11-17 category, leaving 95% as aged 5-10 rather than 100%, halved the number of unfilled places. This establishes that the proportional split of population groups at school destinations is an important parameter and needs to be as accurate as possible for the model to work. To say that a primary school is filled by 5-10 year olds only is not true, and will partly explain why there is a deficit of 5-10 year olds for these destination places as some of the school capacity is filled by 4 and 11 year olds. The challenge is that during the school year, the proportion of 4 year olds and 11 year olds changes, adding an extra temporal dimension to this modelling. In this example, data were found which stated that 80% of children enrolled at primary schools are aged 5-10 (Department for Education, 2016). Therefore, the model was adjusted to fulfil this by V7.

Secondly, potential issues with the school destination data (from Edubase) were investigated. It was assumed that the census data would be correct for the number of children in York at the time it was taken, and adjustments were made to the administrative data for the destination numbers. The issue could therefore be that more children were being required at some schools than the population available to them from the origin census data. It is possible that the school data are affected by a lag in de-registration of outgoing pupils once they transfer to another establishment. This is similar to the health sector, where there is a lag in deregistration from practice registers as patients move between practices. This means that population headcounts from general practice registers often exceed headcounts from census data, because the lag inflates register sizes (Smallwood & Lynch, 2010). Whilst school registers may be affected by similar lags as pupils change schools, head counts of children from school registers nationally are in fact lower than head counts of children from census data, but there can be local variations in such discrepancies (Office for National Statistics, 2013). Some children do go to school outside of their local authority, so would not be captured by the population data for the study area. In York, cross-border flows account for a 0.8% increase in children registered at schools (Department for Education, 2016) which is not enough to account for the discrepancy seen in these initial model results. Thirdly, the census data and school capacity data were from different years, 2011 for the census data and 2017 for school data. Whilst the census data were scaled using the 2016 Mid-Year Estimate, this will not be the exact number of people present in York as it is still an estimate. However, it is unlikely the Mid-Year Estimate is so inaccurate it accounts for the 11% discrepancy in itself. Consequently, a data set from York Local Authority on school enrolment for the target year (2016) was used instead of national Edubase data as the number of pupils was lower, therefore a better match to the origin census data.

Finally, it should be noted that the spatial concentration of schools had an impact. As Figure 5.5 shows, the school catchments in York greatly overlap, meaning there is high competition between destinations for students. The total number of children over the study area and buffer zone was always enough to fulfil school demand; however, there was not enough within the catchments of some of the schools in York. Figure 5.7 demonstrates the schools with unfilled places were clustered together, supporting the theory that urban areas with lots of schools and a finite number of children can cause model problems. What can be taken from this is that in an urban area it is important to get different data sets as closely matched as possible.

5.4.2 Comparing the modelled results of population in travel

Objective 2 was to compare the output of the four-step approach to the output of SurfaceBuilder247 for the representation of people in travel. The results shown in Figure 5.8 have demonstrated that the four-step approach can provide a more detailed representation of the spatial distribution of population in travel than SurfaceBuilder247. This is largely due to the nature of the background layer in SurfaceBuilder247 placing the people in travel across the whole study area, based on the relative weightings of roads, which is not specific to population sub-groups. The four-step approach therefore is an improvement on SurfaceBuilder247 in its current form as it represents travel on the local roads which would be used for school travel.

5.4.3 Defining critical roads

Jenelius (2009, 2010a) define a critical link as one which many people typically use whilst travelling around an area. Therefore, the four-step approach could be used to identify critical roads, as it identifies which are used by the most people within a population group. The results show it is mostly smaller roads rather than strategic network which are important for schools. There are more clusters of the road network for walking routes than car routes, which is unsurprising as walking was the mode used for shorter distances. As discussed previously, the studies by Nyberg and Johansson (2013) and Coles et al., (2017) have looked at defining critical links for a sub-group. Both of these studies look at access for the elderly. Therefore, this research has examined a new group. The four-step approach could be used for any population sub-group with the right data, so after running multiple groups a detailed picture of road use would emerge. A strength of these results is that they are for the whole city which is a useful scale for local authority planning. Also, displaying the results as the number of children is an accessible way of displaying the data to stakeholders.

5.4.4 Preliminary assessment of the four-step approach to modelling population in travel

The four-step approach is a novel framework which can be used to bring spatiotemporal population dynamics to GIS network analysis and has achieved Objective 1. The greatest challenge for this chapter was converting the gridded spatiotemporal population data into a format suitable for network analysis which, as far as the author is aware, has not been addressed within academic literature before. The spatially weighted Monte Carlo simulation has the same underlying principles of Spatial Interaction Models and microsimulation techniques, so the logic for the method is sound. The difficulty lies in validation as the data used as the input for SurfaceBuilder24/7 is the best data available and there is not an alternative data set for comparison. Nevertheless, the development of a code which assigns destinations to origins is a useful advancement to the field of population modelling as it opens the potential to link these data to other techniques.

Like all models, there are limitations to the four-step approach. It is based on modelled data so the results can never be used as an exact representation of population in travel or for real-time emergency management. A key assumption is that people of a certain group follow the same movement patterns (Bian & Wilmot, 2015), when in reality transport networks are a collection of individual decisions (Nagurney, 2011). However, this is an issue with all transport models of this scale (Borrmann et al., 2012; Sohn, 2006; Jones, 2012) and the approach is still a good tool for city-scale transport planning.

The four-step approach is an aggregate modelling approach which means it cannot account for individual behaviour, a key determinant of travel (Jones, 2012), like an agent based model would. Whilst detail would be gained with a micro-scale approach, individual-level micro models would be more computationally expensive than the network analysis for areas as large as York. Also, micro-scale models need to contain behavioural choice-making data (Harland & Stillwell, 2010) which is difficult to come by and challenging to validate. When considering the size of the study area (a city) this aggregate 'meso-scale' framework is the best option as it provides detailed enough information for decisions to be made on a city-wide basis (Alaeddine et al., 2015), the scale local government is ultimately working at.

Journeys with multiple stops cannot be included as part of the four-step approach, as the original concept is based on a principle of journeys being from one origin to a destination (Jones, 2012) which does limit its applications. However, for school travel, it is more likely a child is just travelling to the school rather than going via other amenities, so a simplified travel model is appropriate. A linked limitation is that the results are just for primary school children and do not represent the adults which will be travelling with them. Assumptions could be made about how many adults are travelling these routes for every child, but further investigation is needed to find empirical data to base these on. Validation of the results is a challenge for SurfaceBuilder247 and the four-step approach, as the best available data are used in the models and there is not an independent data set to verify the results against. Whilst individual level data like the National Pupil Database and mobile phone data do exist (Singleton et al., 2011; Alexander et al., 2015; Deville et al., 2014), this is at a different scale to the census data the models use so would not be suitable for validation.

The four-step approach has been designed to work for any population group which can be modelled in SurfaceBuilder247 or other spatiotemporal gridded population models, as it is intended for a wider use than just modelling school travel. The fact it only requires origin data, destination data and destination catchments mean it is a transferable framework as these are standard data required for transport analysis. Therefore, this four-step approach could be used in transport planning as it introduces demographic data to transport analysis allowing more detailed scenarios to be tested. The identification of critical roads could be used within risk management frameworks as society cannot afford absolute protection; therefore, prioritising resources on those nodes and links which are essential is necessary (Freiria et al., 2015). Introducing disruption scenarios to the network analysis part of the framework would allow policymakers to assess how different population group's travel is affected differently, enabling policy to be designed with specific groups in mind.

There are several aspects to this work which can be considered original. Firstly, incorporating the four-stage conceptual transport model with gridded spatiotemporal population modelling has not been attempted before. This chapter has shown it is an appropriate and successful conceptual model for population transport analysis and could be used more widely within Geography. Secondly, how the SurfaceBuilder247 model represents population in travel has not been evaluated before or compared to an alternative method, which this chapter has provided. This work has also looked at the representation of primary school children in the model and how the destination parameters and input data can have a large effect on model outcomes. Thirdly, it has applied code components for population modelling developed by Stevens et al., (2015) to a new situation, demonstrating the transferability of these techniques. Finally, many papers address network functions for the population as a whole including Jenelius (2010b), Rodríguez-Núñez et al.,(2014), Sohn (2006), Taylor et al., (2006); Chen et al., (2015) and Balijepalli & Oppong, (2014). However, this is one of few studies which develops a method to examine the impact on specific population groups, and the only one to use the detailed demographic data provided by gridded populations.

5.5 Conclusion

This chapter has provided a framework for linking gridded population data to vector-based network analyses and shown how it can be used to map routes taken by specific population groups. This is a transferable framework which can be used for other population groups and with different types of gridded data other than SurfaceBuilder247. In terms of the case study area, it has provided more detail on which roads in York are used for school travel at a finer scale than SurfaceBuilder247 alone can currently. The objectives set have therefore been met and this original research provides a sound basis for development in the remainder of the thesis.

Chapter 6 Evaluation of Approaches to Modelling Spatiotemporal Population Flows and Flood-Related Travel Disruption

This chapter has been presented as a conference paper at the GISRUK 2019 conference in Newcastle. See Appendix B for paper.

6.1 Introduction and Background

The majority of fatalities caused by pluvial flash-floods in post-industrial countries are due to people travelling in vehicles (Debionne et al., 2016; Arrighi et al., 2019; Terti et al., 2017; Shabou et al., 2017) rather than those exposed whilst in a residential building. However, daily mobility aspects are not frequently considered in flood exposure and risk assessments (Debionne et al., 2016). As flood risk assessments can now involve dynamic flood models, it follows that dynamic population data should also be included to understand exposure and risk fully (Shabou et al., 2017). Batista e Silva et al., (2018) state that multi-temporal population grids from the ENACT project, similar to the ones created by the SurfaceBuilder247 model, are useful for assessing exposure to natural hazards and planning and modelling transport. This chapter therefore tests this assertion by seeing whether spatiotemporal population grids can be used in a common network analyst tool for the purposes of assessing the effect of a flood event on travellers.

In previous studies of pluvial flash floods on road networks, using census data for origin-destination pairs is a common way of determining travel flow, for example Debionne et al., (2016). However, this does not capture daily flow dynamics in the way the spatiotemporal data used in Chapter 5 does. Alternative methods to include a dynamic nature to the census data include Shabou et al., (2017), who added travel activity data to their census data to include a measure of daily travel activity. In other studies, traffic data has been used to provide the dynamic quality to the road network, either average data for that day (Li et al., 2018) or for a specific flood event (Pregnolato et al., 2017a). However, using traffic flow or count data do not reveal anything about the demographic characteristics of road users, information which would be useful to those involved in risk planning like councils and local resilience forums. In this chapter I will attempt to use a third type of data, spatiotemporal population data created through SurfaceBuilder247, to assess the effect on intra-city journeys from a flash flood event in York.

Chapter 5 introduced a method for adapting spatiotemporal gridded population data to create time-specific origins for use in ArcGIS network analysis tools. However, assessment of these data was not possible due to a lack of reference data for school children. In this chapter, the focus is on commuters and there is origin-destination flow data for workplace travel recorded in the England and Wales census. This can be used to as a comparison to the origin-destination flow data produced from the spatiotemporal gridded population data. In this chapter, the origin data will be compared to identify differences in spatial locations. The origin data will then be used in a cloud-based proprietary transport network tool (ArcGIS closest facility tool), without a flood event, to measure how commuting routes and journey times compare between the two. This process will then be repeated with a flood event acting as a barrier to the roads to assess the effect of a flood on commuter travel times. Finally, the sensitivity of the parameters of the network analyst tool are tested. This chapter has the following aim and objectives.

Aim: To assess the differences between using census origin-destination data and spatiotemporal population derived origin-destination data, to measure the impact of a pluvial flood scenario on travel in York.

Objectives:

1. Generate pluvial flood scenarios for road travel disruption testing
2. Analyse the differences in origin locations between census-derived origin-destination flows and flows simulated from workplace catchments using Chapter 5's methodology
3. Examine the impact of the two different approaches to modelling spatiotemporal population flows on the simulated road travel disruption from pluvial flood hazards
4. Test the inclusion of floods as obstructions in the network analysis tool
5. Test methods for the aggregation and presentation of the results

6.2 Methods

6.2.1 Overview

An interdisciplinary methodology of hydrology, transport and GIS has been applied effectively by previous researchers to this problem area (for example Li et al., 2018). A similar approach is taken in this chapter. The study area is the city of York, as described in Section 4.1.1, focussing on intra-city travel as this is the area covered by the SurfaceBuilder247 population data. There are also restrictions to the number of origin-destination pairs which can be analysed with the ArcGIS network analyst server, which restricts the size of study area to a city rather than a region. Whilst this does not account for all travellers who come into York for work, 71% of those who work in York also live in York (ONS census table WU01EW for MSOA level). Therefore, the majority of work trips are included in this area. I will analyse pluvial flash floods as they are a common hazard in urban environments, due to the prevalence of impervious surfaces (Li et al., 2018). Two approaches to creating population flow data for conducting network analysis were applied; Method A, which utilises raster data from the SurfaceBuilder247 model and Method B, which uses 2011 workplace origin-destination census data. Figure 6.1 provides a diagrammatic overview of both procedures.

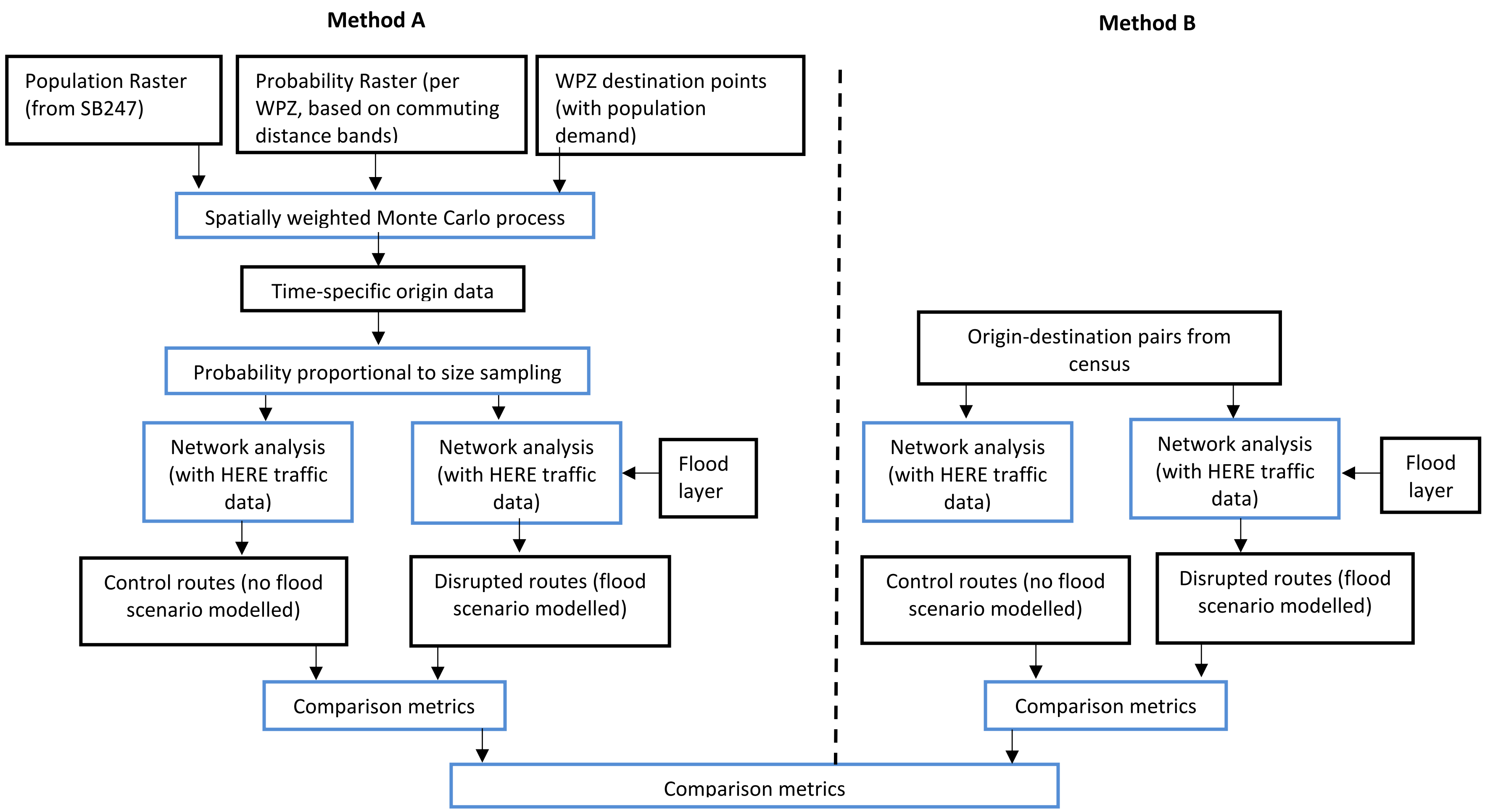


Figure 6.1 Methodological approach of Chapter 6.Blue boxes are analysis processes, black boxes are data

6.2.2 Creating flood data

The flood data used in this chapter was modelled using the Flowroute-i^{TM} software from Ambiental Risk Analytics. For a full description of the model and the pluvial flood modelling strategy, see Section 4.2.1. In this chapter, the one-hour summer storm event for a 1 in 30 year return period was applied as the disruption scenario, taking the maximum depth raster. This was post-processed to remove any areas of flood water which were too shallow to cause impedance to vehicles or had an extent smaller than 125m^2. There is a range for the threshold depth of water which makes a road impassable to a car, with the estimation of 20-30cm of water cited in the UK and US as the air inlet height on a typical car (Li et al., 2018). Pregnolato et al., (2017b) created a depth-disruption curve which indicates 30cm is sufficient water to stop an average car from crossing a road section. In this chapter I have used 25cm depth, a median between all estimates. As the main focus of this chapter is the testing of population flow data and the network analysis tools one flood scenario is used with the intention for the effect of different flood scenarios to be explored in Chapter 7.

6.2.3 Adapting the spatially weighted Monte Carlo approach (Method A)

An R script for a spatially weighted Monte Carlo analysis (Nieves, 2018, pers. comm.) was developed and presented in Chapter 5, to simulate the population origins which are likely to provide people to a destination from spatiotemporal gridded population data and destination catchments (see Section 5.2.2). As the spatiotemporal gridded population data are the same format for commuters as school children, this script was adapted to generate origins for commuters. Three sets of data are required for the script; (i) a population grid, (ii) a shapefile of destinations with population demand and (iii) probability surfaces denoting areas where those working at a destination are likely to live. The population grid was the output of SurfaceBuilder247 for the working population aged 18-64. The destinations were the centroids of workplace zones within the study area using the most recent 2011 census geography. As the study area does not cover the area where all people who work at the selected workplace zones live, the population demand for each destination was reduced to reflect the estimated proportion of workers commuting from within the study area. The distance travelled to work (workplace zone) is provided as a 2011 census output table (WP702EW – Distance travelled to work (workplace population)). These data were used to calculate the proportion of each selected workplace zone that commuted from outside of the chosen study area. This cumulative percentage was then the proportion of the number of people working at that workplace zone who should be sought for each destination, ensuring the number of people required did not exceed those present in the population grid.

The surfaces depicting probability of an adult working in a given workplace zone were based on the same method used to create the probability surfaces for each school in Section 5.2.2. The travel to work distance bands were again used as the basis for creating the probability surface. Each WPZ destination had 3-4 distance bands represented within the study area, depending on its location within the study area. To calculate the weighting of each band, the population percentages were divided by the number of grid cells in each distance band. As no other information is available to indicate if a grid cell has a higher likelihood than any other to provide to the destination, it is assumed all grid cells have an equal chance of supplying the destination.

The script took 10 days to run, longer than for school children due to more destinations to process and greater numbers of people to be selected. The results were 1166 origins for each destination, equating to the number of grid cells with population present in the study area. The weighting of these origins followed the pattern of the probability surface however, with the points most picked in the 2km ring around the destination. This meant a ‘tail end’ could not be discerned to remove from analysis like in Chapter 5. However, the number of origins needed to be reduced, as 1166 would be unrealistic number to model for the number of people in the destination and computationally expensive when subsequently calculating routes using network analysis. Probability Proportional to Size sampling was used to pick a selection of origins for each destination. The number of origins to find was 50% of the number of people at the WPZ, assuming two people travelled from each origin. This produced a sampled selection of origins for each destination, reflecting likely commuter travel between that origin and destination.

6.2.4 Applying census data (Method B)

Method B uses 2011 census data of origin-destination flows for commuters, taken from the WF02EW data table, for the output area and workplace zone levels. These data are directly reported by 2011 census respondents, rather than the simulated data of Method A, although with the caveat that output areas with very small numbers of commuters are swapped for data protection. As the majority of origin-destination pairs in the census data consist of 1-3 travellers, it is likely these data protection measure has influenced the data used. However, the general spatial pattern should be maintained. The data were processed using python scripts to remove any origin-destination pairs where either the workplace zone or output area were not in the study area, and then any pairs where the commuter flow was zero. A shapefile of origins was then created for each destination, with between 80-150 origins for each workplace zone. In this chapter, only one time of day is considered in the model, 7am on a typical working Monday, so additional temporal information is not added to the census data.

6.2.5 Network Analysis

The origin-destination data generated by Method A and B were implemented in the ArcGIS network analysis tool 'closest facility', accessed via the online server using a python script, as described in Section 4.2.3. There are many optional parameters but the ones of interest here are the inclusion of barriers and hierarchy function. To investigate how flood events change the closest facility routing, the 1 in 30 year flood was incorporated as a barrier in the closest facility analysis for both origin data sets. The flood layer represented all flood water deemed impassable (as described in Section 6.2.2.) so was classified as an impassable barrier in the tool parameters. Ideally, flood water which would slow down traffic, but still allows travel, would have also been included, based on the depth-disruption function created by Pregnolato et al., (2017b). However in practice, network analysis via the ArcGIS server is limited to a maximum of 2000 road segments being intersected by barriers (ESRI, 2020b), meaning only the impassable flood barriers were able to be included.

The hierarchy function is included to attempt to mimic the preference for driving on main roads over local roads. This algorithm works in conjunction with Dijkstra's algorithm (Section 4.2.3) to direct the calculated route onto main roads as soon as possible, and complete as much of the journey on higher order roads (ESRI, 2020a). By default, this is turned on in the network analysis tools. However, after noticing some unexpected results between the non-flood and flood routes for some destinations, the effect of hierarchy was tested by turning it off and re-running the analysis to see how the results were affected. Validation of this work is difficult, but sensitivity tests can be used to examine the influence of parameters on the model (Arrighi et al., 2019).

Overall, seven scenarios were analysed using the closest facility server tool in order to compare flood to non-flood travel times, between origin-destination data creation methods and to test the sensitivities of the hierarchy function. The scenarios are described in Table 6.1 below. The time of day used in the network analysis was the departure time from the destination, based on the SurfaceBuilder247 spatiotemporal population data. Journey times in the network analysis were calculated from this point in time for completed origin-destination journeys.

Table 6.1 Description of scenarios run for data collection in this chapter

Scenario	Description
1	Method B origin-destination pairs, hierarchy on, typical Monday morning at 7am, no flooding
2	Method B origin-destination pairs, hierarchy off, typical Monday morning at 7am, no flooding
3	Method B origin-destination pairs, hierarchy on, typical Monday morning at 7am, flood layer including bridge coverage
4	Method B origin-destination pairs, hierarchy on, typical Monday morning at 7am, flood layer not including bridge coverage
5	Method B origin-destination pairs, hierarchy off, typical Monday morning at 7am, flood layer not including bridge coverage
6	Method A origin-destination pairs, hierarchy on, typical Monday morning at 7am, flood layer not including bridge coverage
7	Method A origin-destination pairs, hierarchy on, typical Monday morning at 7am, no flooding.

6.2.6 Comparing the data

Several techniques were used to process the raw origin-destination route data and compare between scenarios. To compare the origin spatial pattern for each destination, the directional distributional tool from ArcGIS was used. This comparison is designed to show if the catchments of the origins from Method A were similar to those of Method B, as it is assumed the census-derived data is a correct measure of catchment. This technique fits an elliptical catchment that encloses 95% of population-weighted origins, when 2 standard deviations are applied, estimating the x and y axis lengths for each destination's ellipse (ESRI, 2020c). The x and y axis lengths were used to calculate a ratio reflecting the compactness of each destination catchment. A value of close to 1 indicates a more circular catchment and a value further away from 1 a more elliptical catchment. To statistically test for differences in modelled catchments, the axis ratios were compared using a paired sample t-test for Method A versus Method B.

To test the effect that differing origin data sets had on the desired routing outcomes, the closest facility tool was run for both origin data sets to the same destinations. These data were used to calculate an average travel time to each destination, aggregating the data into 215 values (one for each destination) per origin data set. These two data sets were then compared to see if the destinations with the highest or lowest average travel times were the same between the two data sets. This aggregation also allowed the comparison between travel times for the non-flood baseline and the flood scenario, to see the effect flooding had.

The hierarchy property of the closest facility tool was tested by running the analysis with the function turned on and the function turned off, for both the non-flood routing scenario and the flood routing scenario, using Method B's origin data. The average travel time and distance for each destination were then calculated for the non-flood and flood scenarios, before the difference between the two for each destination was calculated.

Finally, the spatial route data were consolidated to ascertain which roads were most frequently used in the routing simulations and compare the difference in outcome between the two origin data sets. Each route was given a count of 1 and then combined to give each road segment a value for the number of times it was used.

6.3 Results

6.3.1 Comparison of method origin data

The directional distribution tool showed that the catchments were generally elliptical (a sample is shown in Figure 6.2) and as expected, the catchments for the 215 destinations covered the study area of York. Visually, Method A has a much smaller catchment sizes than Method B, indicating that the origin data sets are substantially different. In the sample shown, the direction of the long axis also differs in the catchments between the two datasets. The mean ratio of the ellipticals for Method A was 1.042 and for Method B 0.875. This demonstrates that the catchments for Method A were generally closer to circular in shape than for Method B. This could be due to the underlying weighting layer used in Method A containing circular catchments. A paired sample t-test was conducted on the ratio values which showed a significant difference between the ratio values of the destinations generated by each method (p value of <0.001). This shows that Method A has not created an origin data set which is similar in spatial size of catchment to Method B (the census data). This is likely due to the underlying probability surfaces used in Method A having higher weightings close to the destination point, meaning more origins were selected which are nearer to the destination than is reflected in the census data of Method B.

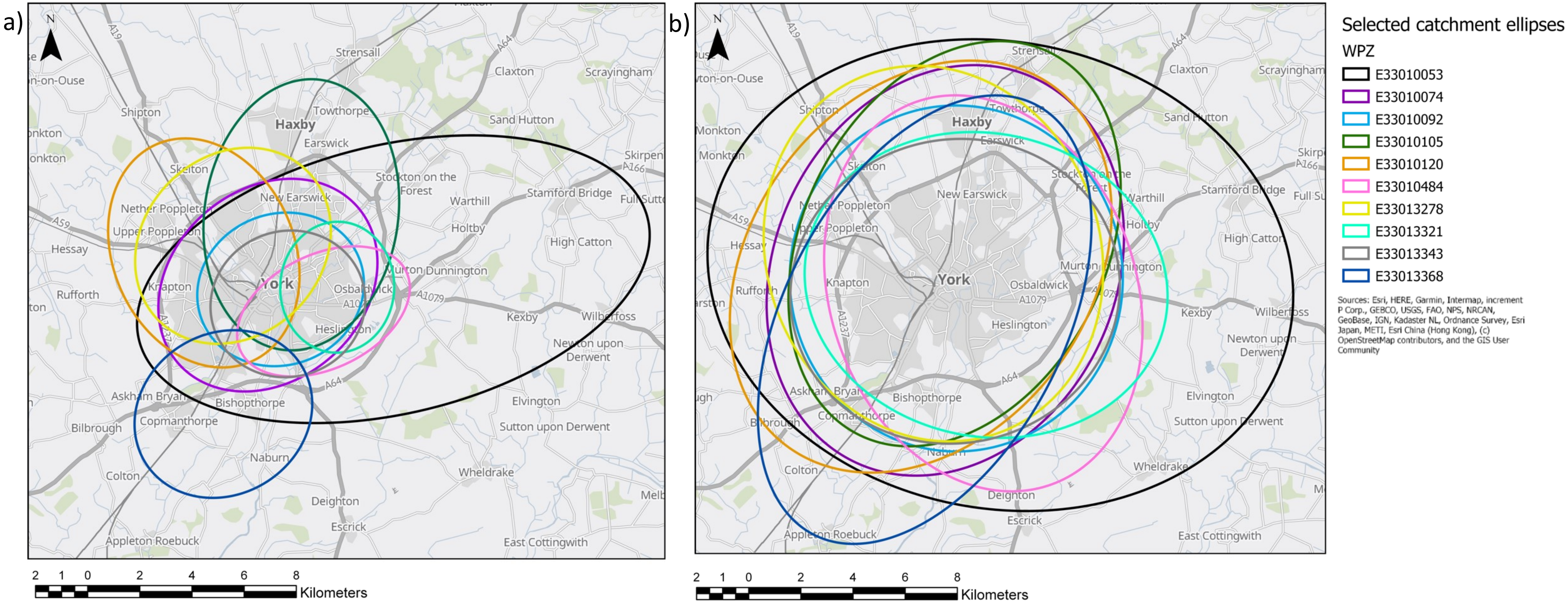


Figure 6.2 Directional distribution elliptical catchments for 10 Workplace Zones, each elliptical cover 95% of the origin data for that destination (2 standard deviations).

a) is Method A, b) is Method B. Contains public sector information licensed under the Open Government Licence v3.0

Figure 6.3 compares the average travel time to each destination in the non-flood baseline scenario for Method A and B (scenario 6 and 1 Table 6.1). The overall trend is that the destinations with the longest average travel times are the same for each method. There is a cluster of the majority of the 215 WPZs with travel times between 4 and 10 minutes, with 10 WPZ with an average travel time above 10 minutes. This data does show that average travel times in Method B are longer than for Method A for the same destinations by around 1 or 2 minutes. This aligns with the results that the catchment areas in Method B are larger, therefore travel time to the destination would be longer.

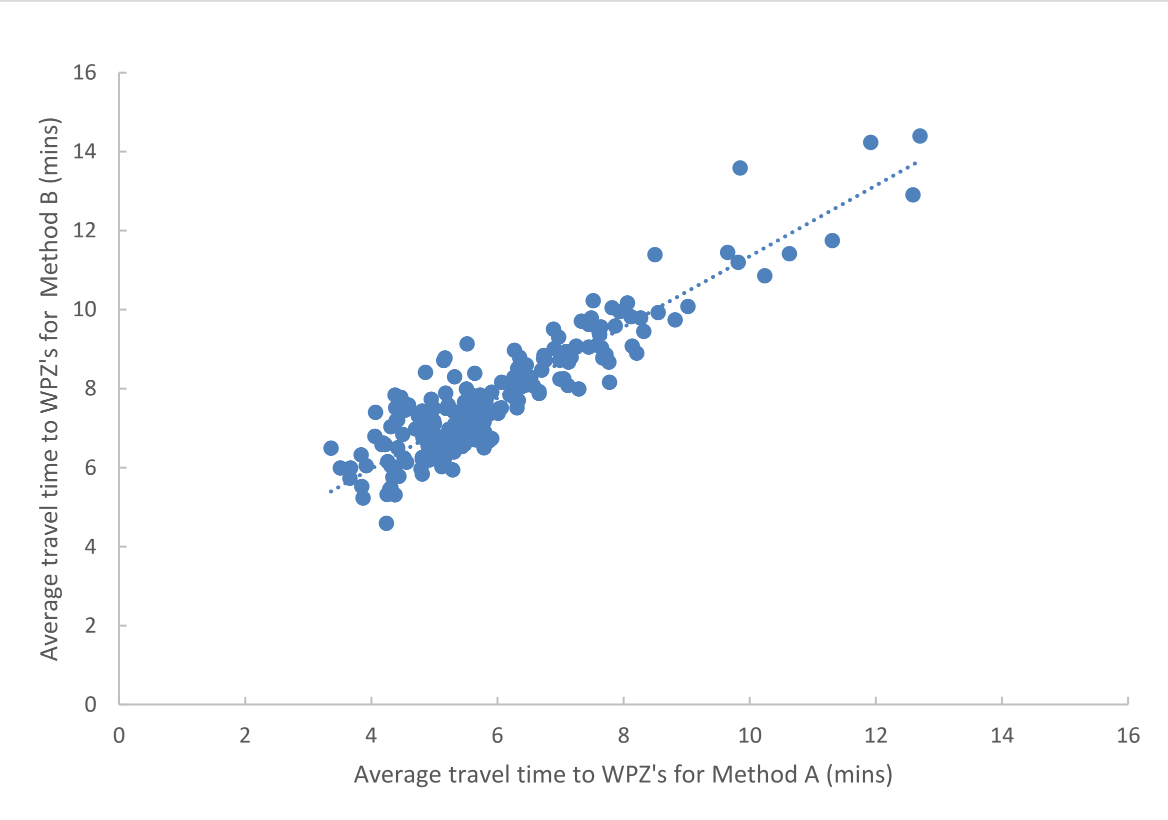


Figure 6.3 A comparison of the average travel times for each WPZ in Method A and Method B.

A comparison of the usage of road segments is presented next. Method B had a higher number of routes using segments. However, this could be due to Method B producing more origin-destination pairs therefore more journeys were modelled. In both data sets, main roads like the A1237 around York and the A19 into York were used more than minor roads. This is unsurprising given the hierarchy properties of the closest facility analysis and the longer journeys compared to school children (Figure 5.10). Overall, a similar spatial usage exists, indicating which roads are used by commuters the most, regardless of the input origin data set.

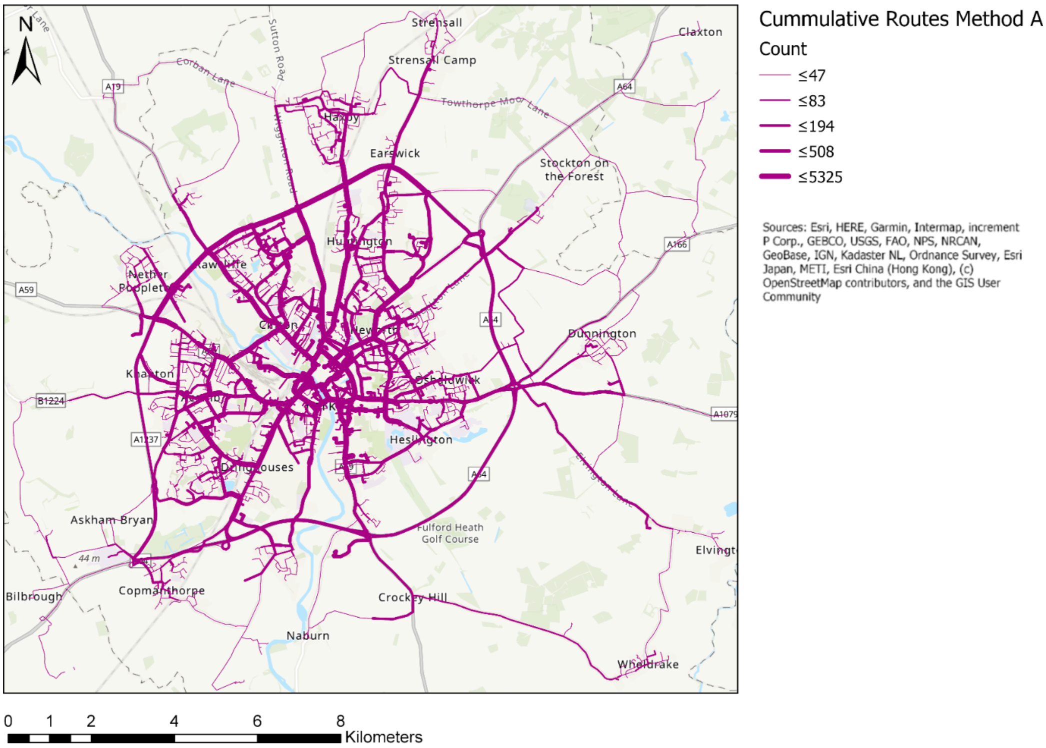


Figure 6.4 Cumulative route map for origins to workplace destinations for Method A.

Contains public sector information licensed under the Open Government Licence v3.0

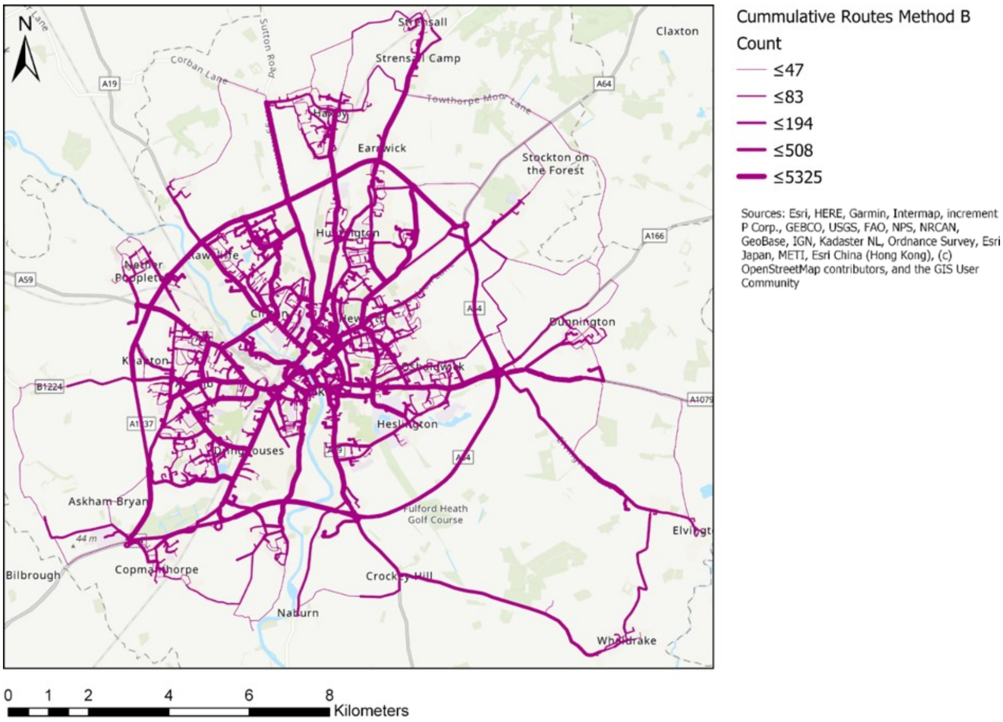


Figure 6.5 Cumulative route map for origins to workplace destinations for Method B. Contains public sector information licensed under the Open Government Licence v3.0

6.3.2 Modelled flood event

The chosen flood event (one hour summer storm for a 1 in 30 year return period) resulted in small patches of flooding across the city and areas where roads were deemed to be impassable (Figure 6.6). Inset A shows one area of flooding, which while small, is at the start of a one-way road system which meant travel into this part of the city centre was not possible. This is one example of pinch points in the city network. Whilst the flood model did place flood water on the bridges over the city's rivers, these were removed due to the low likelihood that they would in fact be flooded in a real situation. The DEM data which underlies the flood model does not always capture bridges well, so this may account for this discrepancy.

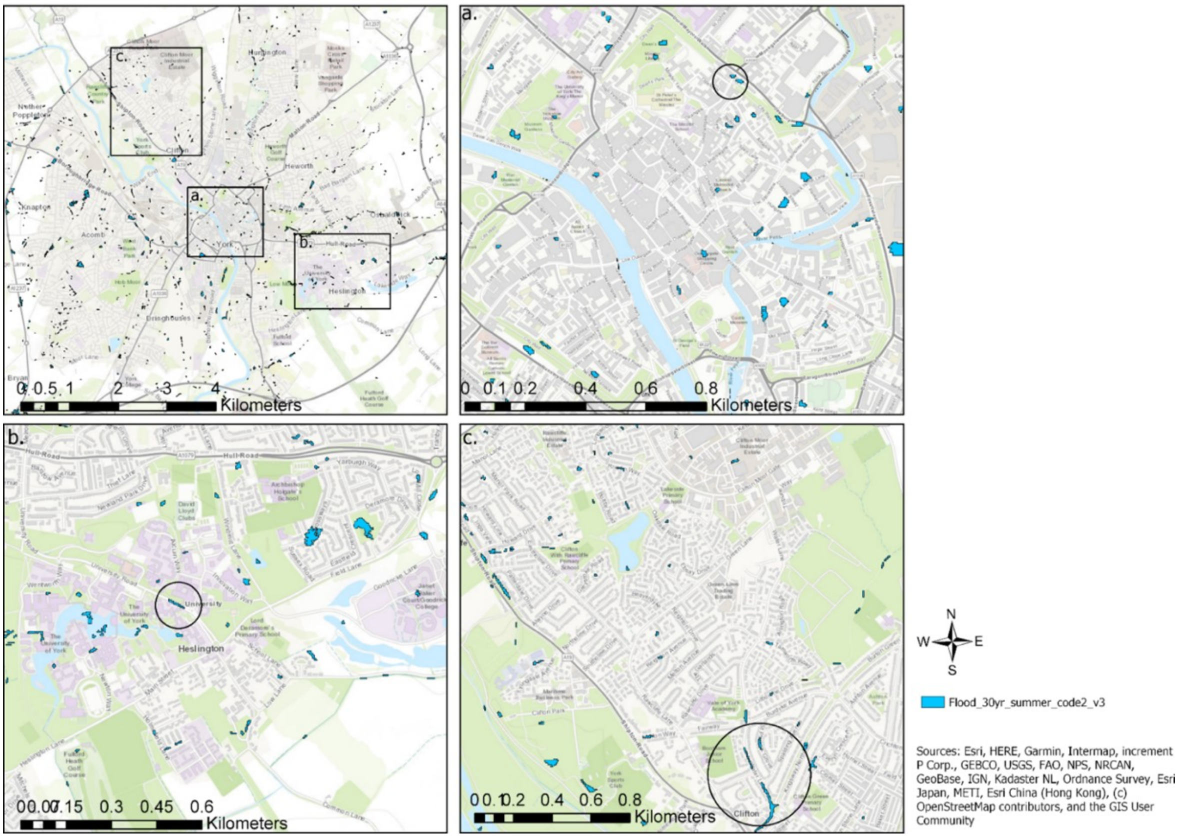


Figure 6.6 Modelled flood area for York for a 1 in 30 year flood event for a 1hr duration, summer hydrology. Insets show part of the city where flooding has caused roads to be blocked (circles). Contains public sector information licensed under the Open Government Licence v3.0
Contains data which are copyright of Ambiental Risk Analytics, used with permission.

6.3.3 Sensitivity of network analysis hierarchy setting

Figure 6.7 are graphs of the average distance difference plotted against average time difference, of the flood scenarios compared to the baseline, for the workplace zone destinations, using origins produced by Method B. Figure 6.7 shows that time and distance increased when flooding was introduced, which would be expected as the flood layer should cause disruption. There are some cases where distance decreases, but time increases. This could possibly be due to the flood route involving local roads which are a shorter distance to the destination but are slower due to lower speed restrictions. The hierarchy function also recalculates values for roads each time the analysis is run, so when the flood layer is included the roads will be given a different hierarchy value to the non-flood simulation (Applebaum, 2019, pers. comm.), this is therefore another likely cause. The removal of the hierarchy function means fewer destinations appearing in the upper left quadrant of the graph where time increased while distance decreased. In this case, all roads are treated as equals so the effect of adding the flood layer (and changing the hierarchy values) is not seen. The general spread of results does not change substantially from when the hierarchy is turned on.

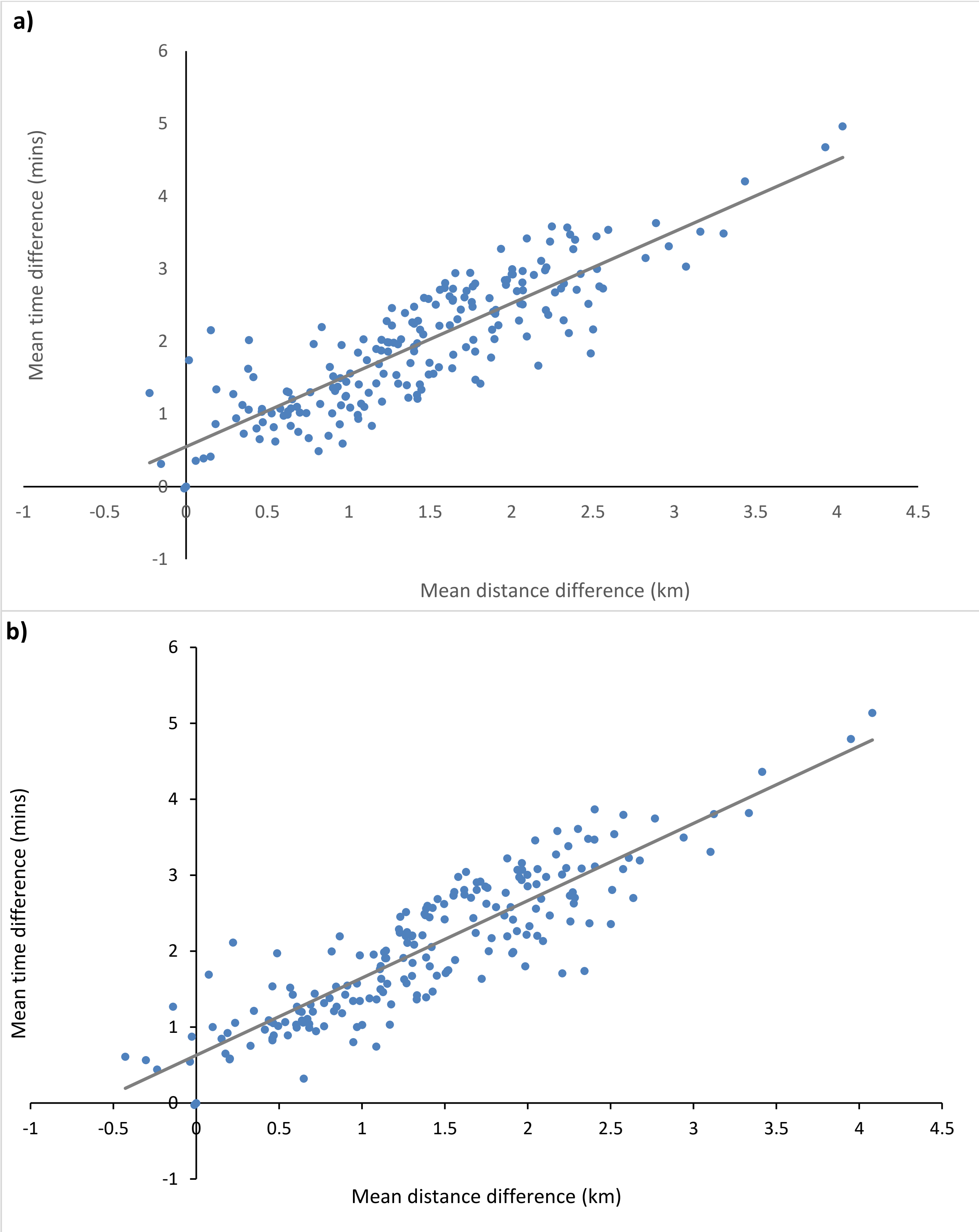
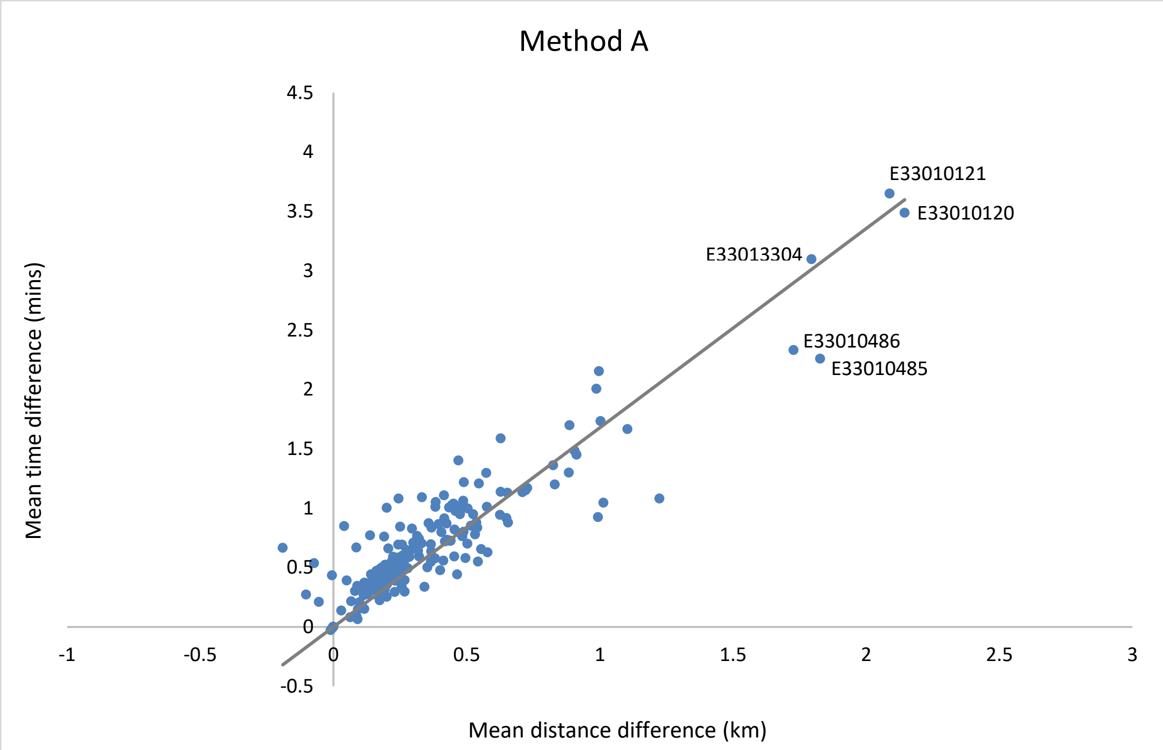


Figure 6.7 Modelled journey times with hierarchy function turned on (a) and turned off (b), for a flood scenario where floods which cover bridges were included. (dots represent workplace zones)

6.3.4 Effect of flooding on network analysis

Figure 6.8 demonstrates there was a general trend of distance and time increasing after the flood event was introduced to the closest facility analysis for both Method A and B. There are some destinations where distance decreases as time increases, partly explained by the hierarchy function as demonstrated in Section 6.3.3. Figure 6.8 identifies the five destinations which have increases in distance and time larger than the general cluster of the data. These five are the same in both Method A and Method B, demonstrating that the WPZ with longest average travel times will be identified despite changes in the origins. Appendix C contains maps showing the spatial distribution of workplace zones and the average travel time to reach them.



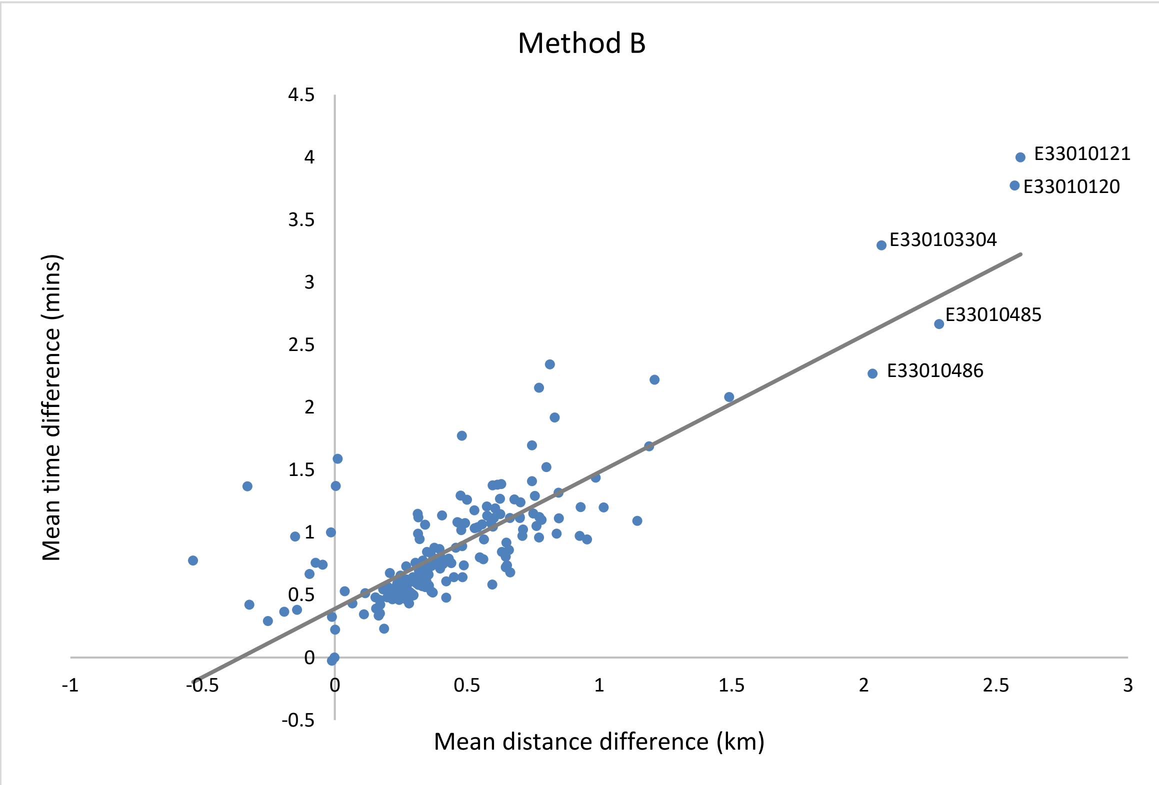


Figure 6.8 Comparative graphs of mean distance and mean time to each destination for Method A and Method B.

The spatial routing data were aggregated for Method A (Figure 6.9) and Method B (Figure 6.10) for travel during the flood scenario. In both data sets, there are a notable absence of any routes travelling on the Northern section of the A1237 near Earswick. In the non-flood scenarios (Figure 6.4 and Figure 6.5) this was one of the road segments with the highest usage. Leeman Road (alongside York station) also does not have any road users during the flood event having previously been a well-used road.

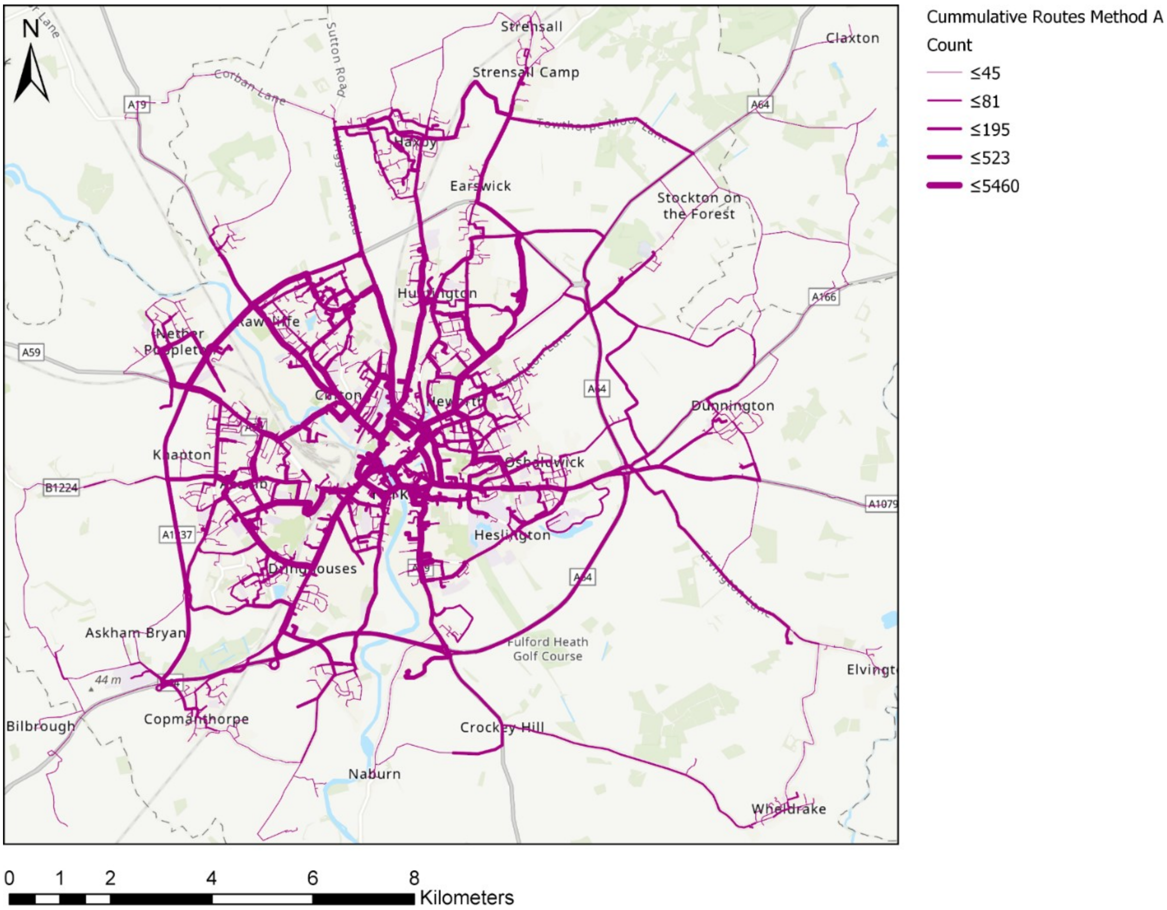


Figure 6.9 Cumulative routes in flood event for Method A. Contains public sector information licensed under the Open Government Licence v3.

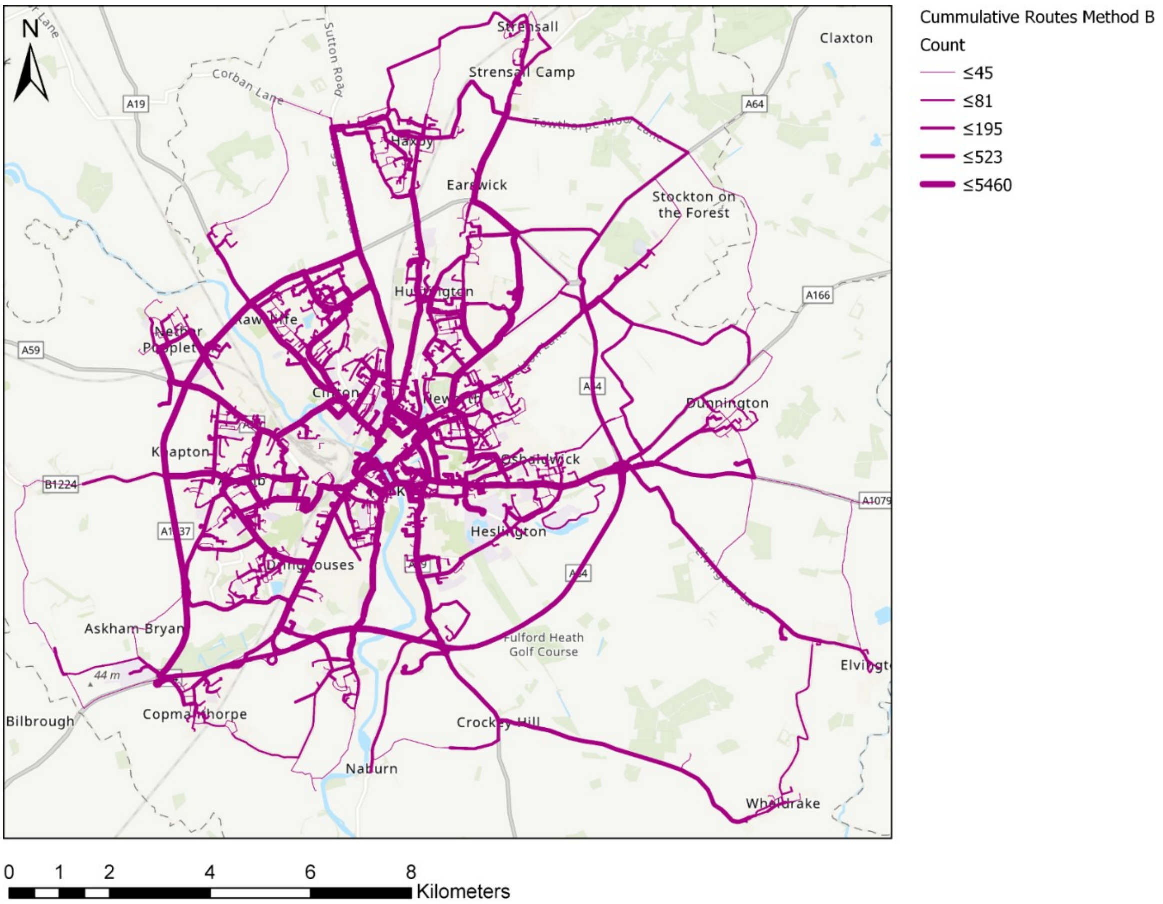


Figure 6.10 Cumulative routes in flood event for Method B. Contains public sector information licensed under the Open Government Licence v3.0

6.4 Discussion

6.4.1 Comparison of the origin data

The second objective of this chapter was to compare how the origins from Method A, designed in Chapter 5, compared to the origins from the census data for commuter flows (Method B). The origin selection process in Method A did not work as well as expected and a clear set of origins which were statistically more likely were not as easily distinguishable as in Chapter 5. This led to using the proportional population sampling approach to reduce the number of origins modelled in the closest facility routing tool as the 1166 origins initially selected was much greater than the number of people required.

The results of the paired t-test and directional distribution tool have shown that the two data sets are significantly different (Section 6.3.1), therefore Method A was unable to produce origin data similar to that in Method B. The smaller catchments of Method A's data are likely due to more origins being selected closer to the destination than in the small area census data of Method B. This lies in the properties of the underlying probability grids created for Method A. These were based on travel band data for the percentage of the working population living within set distances from their workplace zone. An assumption of creating the raster probability surface was that any grid cell within one of these distance bands has as equal a chance of providing to the destination as any other grid cell within the same band. Therefore, the proportion of the working population required in the workplace zone from that band was divided equally amongst all grid cells within it. The first distance band (<2km) was the smallest and had the lowest number of grid cells, with the number of grid cells increasing through the distance bands. This meant that the area with the highest weighting was the distance band closest to the destination and the lowest the furthest distance band. This skews the data, so more origins were selected closer to the destination than appears in Method B's data.

The resultant routes generated from these origin data were also compared. Figure 6.5 and Figure 6.6 show broadly the same spatial pattern of which roads are used most frequently in the routes, demonstrating that on an aggregate level the differences in the origins do not have a marked effect on which roads are shown to be used most by commuters. The difference is more notable in Figure 6.4 where there are differences of 1-2 minutes in average travel times between the two data sets. This is perhaps because those destinations which take longest to travel to, perhaps because they lie on the outside of the main city centre, will always be those with the longer travel time.

6.4.2 Sensitivity of the network analysis tool

The hierarchy function does influence the closest facility routing results with more destinations showing an average change of shorter distance, but greater time, once a flood event is introduced. This sensitivity test has shown that the hierarchy does have an effect; however, it is a useful parameter to keep as it is designed to mimic human behaviour of preferring to travel on larger roads.

6.4.3 Scoping of network tool behaviour with flood scenarios

The general pattern of Figure 6.8 is not surprising, it is expected that introducing a flood event would cause disruption by increasing travel time and that the destinations with the largest increases in travel time are generally those with large increases in distance too. The actual time disruption is not very large, but this could partly be a product of the journeys being intra-city, meaning they are shorter. The network analysis model also does not account for congestion caused by drivers having to take alternative routes, so the travel times presented here are likely to be underestimations of the actual disruption (Borrmann et al., 2012).

There are spatial clusters of destinations which were affected more than others (Appendix C), indicating that these places would have more disruption to workplaces due to difficulties in reaching the area. Whilst there are differences between the two methods, they show similar results when identifying places where the largest impact of flooding occurs. This is useful as it means the extremes of the data distribution are consistent and there is more confidence in the results, identifying places most disrupted if this method was used for flood risk assessments.

6.4.4 Limitations and uncertainties

There are limitations in the analysis presented in this chapter which are shared with other studies. Firstly, not including dynamic congestion within the traffic model (Versini et al., 2010) will lead to an underestimation of travel times, and could potentially affect the routes some travellers would be assigned. Secondly, the closest facility tool uses algorithms which assume travellers take the least-cost path which is known to be an underestimation of travel time and disruption (Li et al., 2018). Thirdly, validation of the outcomes of the flood disruption to travel time is difficult as often there is not ‘real-world’ traffic data to compare the results to (Arrighi et al., 2019).

6.5 Conclusion

Comparing the outcomes of Method A and B, it has been shown there are significant differences. The origins generated from the Method A were affected by the probability surfaces used to dictate the likelihood of a grid cell providing to an origin, leading to an underestimation of distance from origin to destinations. So, whilst Method A worked well for primary school children in Chapter 5, it has not worked well for generating reliable origin data for commuters. Therefore, in Chapter 7 an alternative method of combining census data with time profiles will be used to capture the spatial accuracy of the census data and incorporate a dynamic element. One objective of this Chapter was to assess the differences between using spatiotemporal population data (method A) and census origin-destination data to measure the impact of a pluvial flood scenario on travel in York. This has been achieved as it has been shown that whilst there are differences in the origins, similar results appeared for the travel times and routes to destinations when the flood scenarios were introduced.

Chapter 7 Spatiotemporal modelling of flood-related impacts on daily population movement

This chapter has been presented as a conference paper at the GISRUK 2020 online conference. See Appendix B for paper.

7.1 Introduction and Background

The previous two chapters have examined how spatiotemporal flow data can be created from spatiotemporal population data and compared the results to census origin-destination data. This chapter is applying these spatiotemporal flow data to the issue of pluvial flood events disrupting road travel, incorporating the principles highlighted in Chapter 2, in order to create a proof of concept for spatiotemporal flood and transport risk analysis.

As discussed in Chapter 6, the majority of fatalities in post-industrial countries caused by pluvial flash-floods are due to people travelling into floodwaters in vehicles (Debionne et al., 2016; Arrighi et al., 2019; Terti et al., 2017; Shabou et al., 2017). Traditional flood risk assessments focus on the characteristics of the hazard (e.g. magnitude, velocity, spatial location) and do not consider the effect of population characteristics. Risk literature (for example Aubrecht et al., 2012; Wisner et al., 2004; Freire et al., 2013; Liu et al., 2014a) asserts that time of day is a critical factor in the scale of impact, due to the movement of people throughout the day leading to different demands on road use. Therefore, the hypothesis for this chapter is that the time of day of flood onset is as important as the size of the flood.

In this thesis, risk is defined as a function of hazard, exposure and vulnerability over space and time (Section 2.2). This research is thus positioned in the 'disaster risk' view of natural hazards by integrating time-specific estimates of population in travel to provide what Aubrecht et al., (2013) describe as a more 'encompassing' view of risk. In Chapter 2, a risk equation (Equation 7.1) was developed as a foundation for considering the spatial and temporal nature of risk and these principles are applied in this chapter.

$$\text{Risk} = f(\text{Hazard}_{ST},\text{Exposure}_{ST},\text{Vulnerability}_{ST}) \quad \text{Equation 7.1}$$

As set out in Section 1.3, the third aim of the overall thesis will be fulfilled in this chapter with several objectives to structure the analysis.

Aim: To analyse whether the time of pluvial flood onset is more important than flood magnitude for disrupting commuter and primary school travel in York.

1. Analyse the extent to which pluvial flood magnitude and flood onset time affect journey times due to road network disruption.
2. Assess whether commuter's or school children's travel to workplace or school is affected the most.
3. Identify schools and workplaces which are most susceptible to disruption from flooding and the effect pluvial flood magnitude and pluvial flood onset time has on identifying them.

This chapter presents a proof of concept for combining spatiotemporal flow data with flood data and network analysis to quantify the effect of time of flood onset and size of flood hazard. It focuses on the 18-64 aged commuter population and primary school children aged 4-11, as described in Sections 4.1.3 and 4.2.2.

7.2 Study Design, Data and Methods

7.2.1 Choice of case study scenarios

The population groups modelled in this analysis are the same as those in Chapters 5 and 6, primary school children (aged 4-11) and commuters aged 18-64. These two groups are chosen because their travel patterns are known through travel surveys and one is considered a vulnerable group (primary school children). To decide on a suitable time period of study, the number of people in travel at a given time of day were calculated from time profiles applied in the population modelling (Figure 7.1) and showed the interaction of primary school children and commuters in travel during the morning. The time period chosen for study was 07:00 – 09:00 hours on a typical Tuesday weekday morning (an 'average' day of the week), correlating to the morning commute. This was because this time period involves substantial variation in both school children and commuters travelling in a confined time frame, enabling examination of the effects of changes in time of flood onset relative to the population movement. Similarly, Pregnolato et al., (2017a) simulated travel in the morning and evening commutes for their study of flood disruption to travel, as the greatest economic disruption occurs from interruption to commuter journeys.

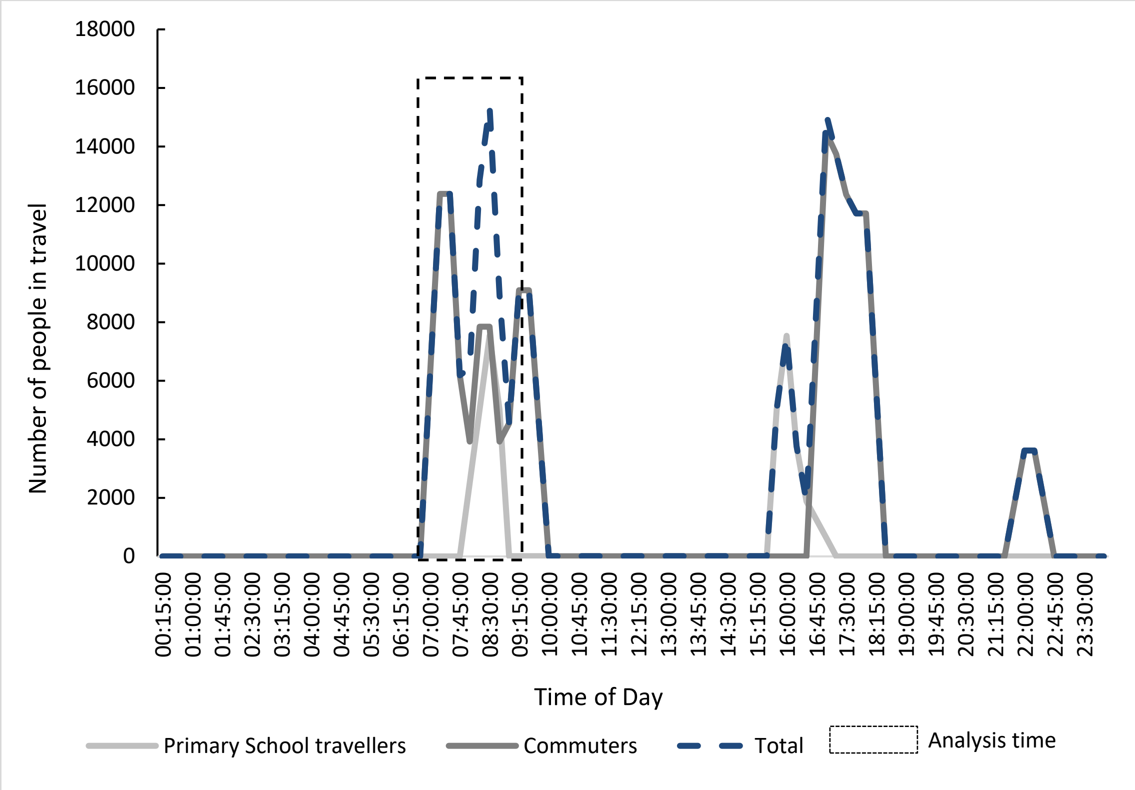


Figure 7.1 The number of people in travel on a normal working weekday. The box highlights the time period selected in this chapter for analysis

Both the time of the flood event onset and the magnitude of the flood were variables when designing the hazard component of the scenarios for simulation. The flood events modelled were 1 hour duration summer storms with 2 hours additional simulation time to capture the flooding after the rainfall stops (see Section 4.2.1 for full pluvial flooding methodology). As the research interest is the effect of flood events on road travel, it was critical to examine the temporal coincidence of the peak travel period with peak flooding. As the target time period was the morning travel period (07:00-09:00 hours) flood onset times were selected based on this.

A scenario based approach was taken to test the hypothesis as outlined in Figure 7.2 and Table 7.1. First, a 'baseline' set of data were created to be a control comparison to the flood scenarios. This represents 'business as usual' without any flood disruption. Two factors, time of flood onset and magnitude of flood event, were then varied in turn. Three times to introduce the start of the flood scenario were chosen to cover the morning commute, 6am (flood time 1 (FT1)), 7am (flood time 2 (FT2)) and 8am (flood time 3(FT3)) (Figure 7.2). As flooding does not appear immediately as rainfall begins, the rainfall was introduced an hour before the time period of interest, so there would be flooding in the city at 7am in the first scenario. On Figure 7.2, the duration of the rainfall in the scenario is highlighted by the blue lines. These were chosen as they systematically shifted peak flooding to coincide with, or avoid, the morning travel patterns and should reveal hourly temporal changes in the travel data. Two flood magnitudes were modelled, 1 in 30 and 1 in 100 year events, both for summer hydrology, and introduced at each of the three flood onset times. This resulted in seven scenarios being modelled in total, including the non-flood baseline (Table 7.1). The data used in the network model was the flood area exceeding 30 cm depth and 125m^2 at each time point. As discussed in section 4.2.1, a snapshot of the flood area was available for every 15 minutes in the simulation period. This meant a different set of flood data could be selected to align with the time elapsed since the rain began. For example, FT1 (6am), the network analysis run at 7am would use the flood data for 1 hour into the simulation. Chang et al., (2010) state that studies on travel disruption can overestimate the impact of flooding due to not modelling the receding water at a fast enough rate, particularly if one static flood outline is used. Therefore, this method of using a series of flood layers for each time point during the flood scenario means it is more likely changes in which roads are open or closed during the course of the flood event are captured. As Figure 7.2 shows, there were five 'sampling points' the network analysis was run at for each of the 7 scenarios. This meant there were consistent time points results could be compared at between scenarios.

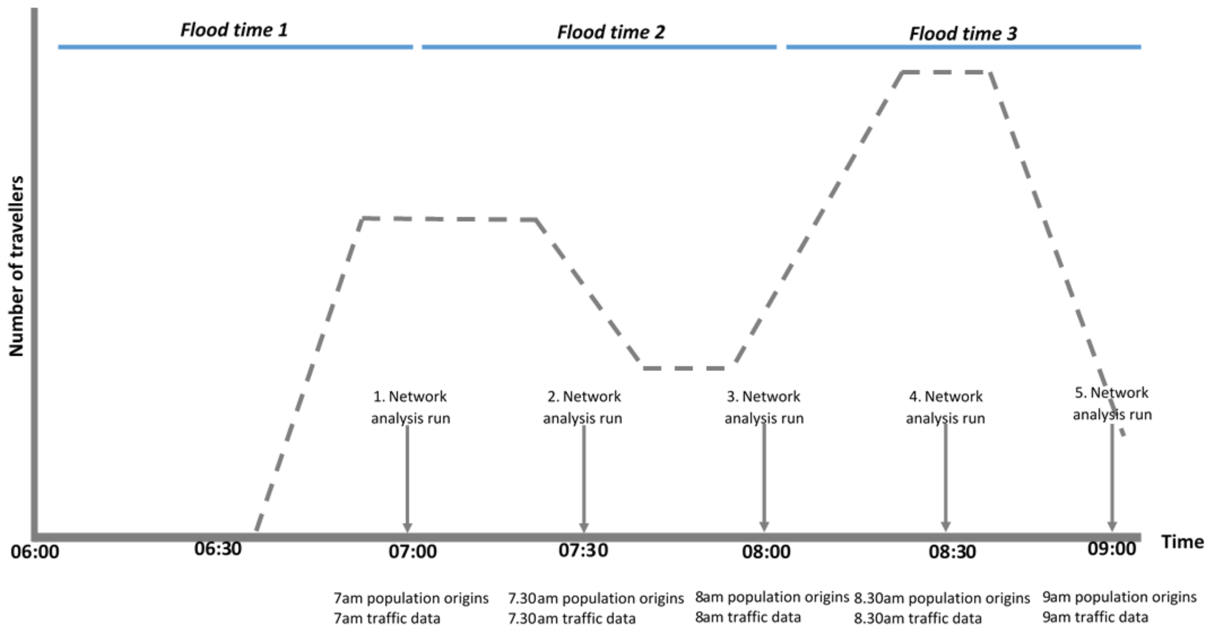


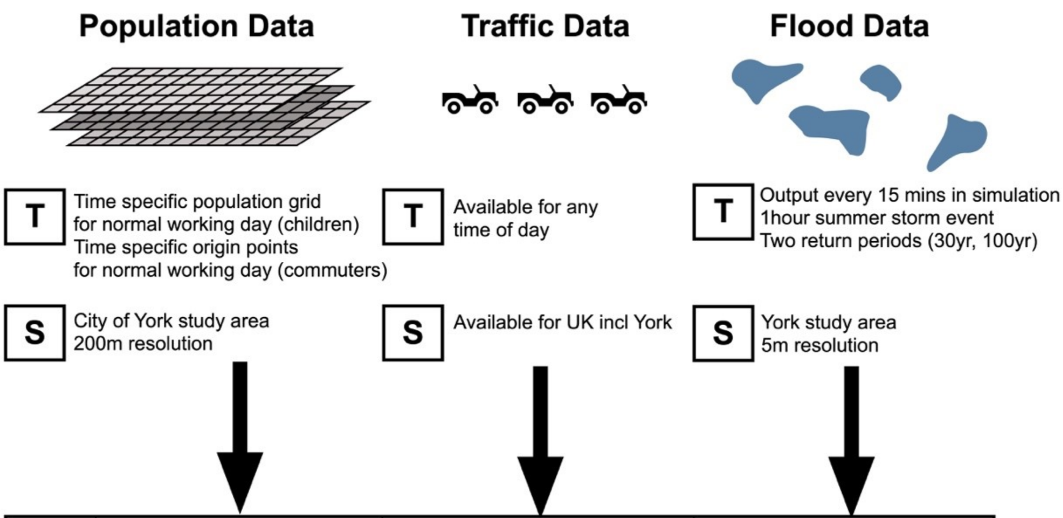
Figure 7.2 Conceptual graph of the population in travel during a typical weekday morning and the flood hazard onset for each flood scenario, the time where rainfall occurs is highlighted in blue.

Table 7.1 Scenario structure used in this chapter

Scenario Name	Description
Non-flood Baseline	Network analysis with no flood layer included, a baseline to compare the flood results to for changes.
Flood time 1, 30yr flood	Network analysis with flood layer for a 1 in 30 year event, flood onset at 6am.
Flood time 2, 30yr flood	Network analysis with flood layer for a 1 in 30 year event, flood onset at 7am.
Flood time 3, 30yr flood	Network analysis with flood layer for a 1 in 30 year event, flood onset at 8am.
Flood time 1, 100yr flood	Network analysis with flood layer for a 1 in 100 year event, flood onset at 6am.
Flood time 2, 100yr flood	Network analysis with flood layer for a 1 in 100 year event, flood onset at 7am.
Flood time 3, 100yr flood	Network analysis with flood layer for a 1 in 100 year event, flood onset at 8am.

7.2.2 Overview of study design

A semi-dynamic workflow has been created to utilise the temporal dynamics of the datasets available. Three main types of data are utilised in this chapter; population, traffic and flood. Each of these has a temporal and spatial element which are synchronised, as shown in Figure 7.3. For each modelling scenario (Table 7.1) the datasets are input into a routing algorithm to calculate the origin-destination routes for each workplace or school. Within the scenario, routes are calculated at five time points: 07:00hrs, 07:30hrs, 08:00hrs, 08:30hrs and 09:00hrs (Figure 7.2). The output is a set of routes and associated time data for those origin-destination pairs, for the school children and commuters.



Scenario Time	Population	Traffic	Flood
7am	7am population origins	Run network analysis for 7am traffic time	Flood area for 60 minute mark of 3 hr flood simulation
7:30am	7:30am population origins	Run network analysis for 7:30am traffic time	Flood area for 90 minute mark of 3 hr flood simulation
8am	8am population origins	Run network analysis for 8am traffic time	Flood area for 120 minute mark of 3 hr flood simulation
8:30am	8:30am population origins	Run network analysis for 8:30am traffic time	Flood area for 150 minute mark of 3 hr flood simulation
9am	9am population origins	Run network analysis for 9:00am traffic time	Flood area for 180 minute mark of 3 hr flood simulation

Outputs * Route shapefile of each origin - destination journey
* Time and distance values for each route

Figure 7.3 Overview diagram of how the flood disruption scenarios are structured in this chapter to achieve a semi-dynamic work flow, using the example of flood onset time 1.

7.2.3 Flood modelling

As described in Section 7.2.1, two flood magnitudes for a 1 hour summer storm were modelled, a 1 in 30 year event and a 1 in 100 year event. These two return periods were chosen as both storms would lead to disruption to the road network, with representation of a smaller but more frequent event (1 in 30 year) and a larger but less frequent event (1 in 100 year). This is similar to the EA's pluvial flood maps which show a 1 in 30 and 1 in 200 year event (City of York Council, 2019). A full description of the Flowroute-i^{TM} model and its functionality were given in Section 4.2.1. In summary, Flowroute – i^{TM} models the depth and velocity of water in each cell during the simulation and a raster for both depth and velocity across the study area is produced for each 15 minutes of simulation time. In the network model, the flood data are represented as barriers to the network where travel through them is prevented. Therefore, the exact modelled flood depths across the city were not required, only a vector file for each time step of the flood scenario which contained the areas where the floodwater met the conditions to prevent travel of vehicles along the road. Two conditions for 'disruptive flooding' were set, a depth of 30 cm or greater and an area of 125 m^2 or more. The depth threshold has been discussed previously (see Sections 3.1.4, 4.2.1 and 6.2.2) and 30 cm was used in this chapter as it is the depth considered no longer safe for most cars (Pregnolato et al., 2017b). After this processing, the areas of flood water which would be considered disruptive to car travel remained, one shapefile per 15-minute time step in the three hour simulation, for both flood magnitudes.

The geographic area used in the flood modelling was based on the study area defined for the SurfaceBuilder247 population model, with the same 20km buffer area for capturing catchment drainage into the city. As there is the restriction on the number of barriers which can be included in the network analysis tool (ESRI, 2020b), the area of the flood output used was clipped to the area of the study area (the city of York as shown in Figure 4.1 and Figure 5.1). Including flooding for the entirety of this square study area worked for the 1 in 30 year storm event; however, the area included too many flood barriers for the 1 in 100 year event, due to more flooding from greater rainfall. This meant the area of flooding included had to be adjusted in order for the network analysis to run, excluding any flooding outside of the main A1237/A64 York Road surrounding the city. This led to some differences in routing for origins outside of the main city between the two flood scenarios. However, it did not affect the main part of the study area.

7.2.4 Time-specific origin data

A data library of time-specific origin data was created for commuters and school children which could be applied to all flood scenarios. By creating time-specific origins for each of the time points (07:00, 07:30, 08:00, 08:30 and 09:00), the temporal nature of population movement could be captured. Chapter 6 resolved that the spatially weighted Monte Carlo technique did not work well for commuters, therefore the alternative method of combining census origin-destination data with the time profiles, based on the Labour Force Survey data, of travel patterns was used. To create the time-specific origins, some of the principles of the spatially weighted Monte Carlo simulation were replicated. The 2011 census provided the population weighted centroids of the output areas from which commuters travelled to each workplace zone (WPZ) centroid, with the associated number of people (Nomis (2020) data table WF02EW). This was the total number of commuters over the entire day, so using the time profiles for the percentage of workers in travel at a given time the proportion of people travelling at each time point was calculated for each destination. This gave the total number travelling to the destination at a given time, however the number per origin was required. A spatially weighted random selection process was created using Python to select the number of people required from the output area origins to fulfil the travel demand to each destination. Similar to the spatially weighted Monte Carlo process in Chapter 5, the origins and their number of people were listed for each destination, and one person selected at a time and the origin OA code stored. This continued until the number of people for the destination was selected and the relevant origins extracted as a shapefile. An origin could be selected more than once, with those origins with the highest number of people given a higher chance of selection. This spatially weighted random selection script was run for all destinations at each of the five time points (07:00, 07:30, 08:00, 08:30 and 09:00) to generate five sets of origin data points per destination which now had time-specific populations attached.

The time-specific origins for school children were created using the spatially weighted Monte Carlo simulation, as this worked effectively in Chapter 5 and there are no alternative freely available origin-destination data on travel to schools. The time profile data from the SurfaceBuilder247 model has travel to school occurring between 08:00-09:00 hours, so only two time points (08:00 and 08:30 hours) were modelled for school children. For primary schools, three population groups from the SurfaceBuilder247 model are required; 0-4 years, 5-10 years and 11-17 years. For each of the three children population groups, origins for each time point (08:00, 08:30 hours) were created.

7.2.5 Traffic data and network analysis

As outlined in Section 4.2.3, the ArcGIS network analysis 'closest facility' tool was the method used for route analysis in this thesis. The script developed in Chapter 6 for accessing the server version of the tool with historic traffic information (source HERE) was utilised again. Incorporating traffic data addresses limitations affecting baseline scenario modelling in studies like Coles et al., (2017), who were unable to include traffic data in their study of flooding in York. In this thesis, the traffic data are applied to the flood scenarios, however it will be an underestimation of travel time as it does not account for the increased travel time due to congestion on surrounding roads to the flood disruption.

7.2.6 Assessment of flood-related travel disruption

The methodological structure of the analysis for this chapter is set out in Figure 7.4. The origin, destination and flood layer (if applicable) were used in the network analysis, producing shapefiles of the routes taken between each origin-destination pair and the distance and travel time of the route. This procedure was repeated for each destination across each time point (07:00-09:00 hours) in the scenario for commuters and school children. This process was repeated for each of the 7 scenarios.

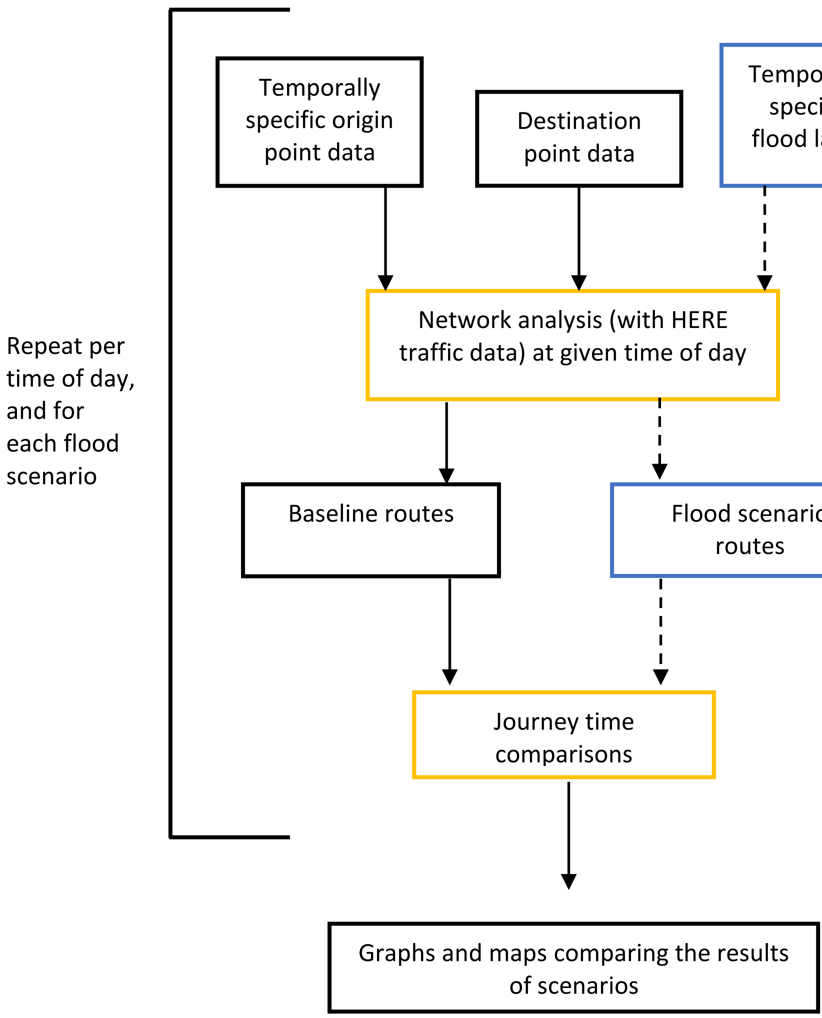


Figure 7.4 The analysis workflow of Chapter 7. Black boxes indicate inputs and outputs, the blue box optional flood data, the yellow boxes processes and the dashed arrows steps included for flood scenario analysis.

The result of this analysis cycle were thousands of route shapefiles ready for comparison between each scenario and the baseline, with the challenge to create readable graphics from the complexity. To answer the main research question, whether flood onset has more of an effect on travel disruption than flood magnitude, the total time lost at destinations for commuters and school children across the city for each scenario was calculated. This measure is used as a value of the cost caused by the flood event. Those origins where travel to the destination was not possible were given a time value of 480 minutes per person, the value of missing a day's work. The travel time for each origin-destination journey was the chosen metric for comparison as time is a common measure of journeys (Wei et al., 2018; Kilgarriff et al., 2019). Python scripts were written to sum total time lost at work or school overall per destination and scenario. A Spearman's Rank correlation test was conducted to see if those destinations with the highest time lost at school or work were consistent across the scenarios. These data were then visualised to highlight spatial patterns in the destinations affected by the largest loss of time. Maps of the roads used at each time point in the scenarios were created by merging the relevant route shapefiles together and adding up the number of people traversing a section of road. This value was then plotted as proportional line symbols on the map.

7.3 Results

7.3.1 Modelled flood events

Figure 7.5 (1 in 30 year storm) and Figure 7.6 (1 in 100 year storm) show the places where the flood model predicted there would be floods greater than 30 cm deep and 125 m^2 in size. At 30 minutes into the scenario, there are only small areas where the flooding meets the disruption criteria. From 60 minutes disruptive flooding is clearly visible on the map in both scenarios. Overall, the 100 year flood has a greater flood area than the 30 year flood, but both show disruptive flooding in the same locations across the city, providing confidence that these are the places which are at risk of pluvial flooding.

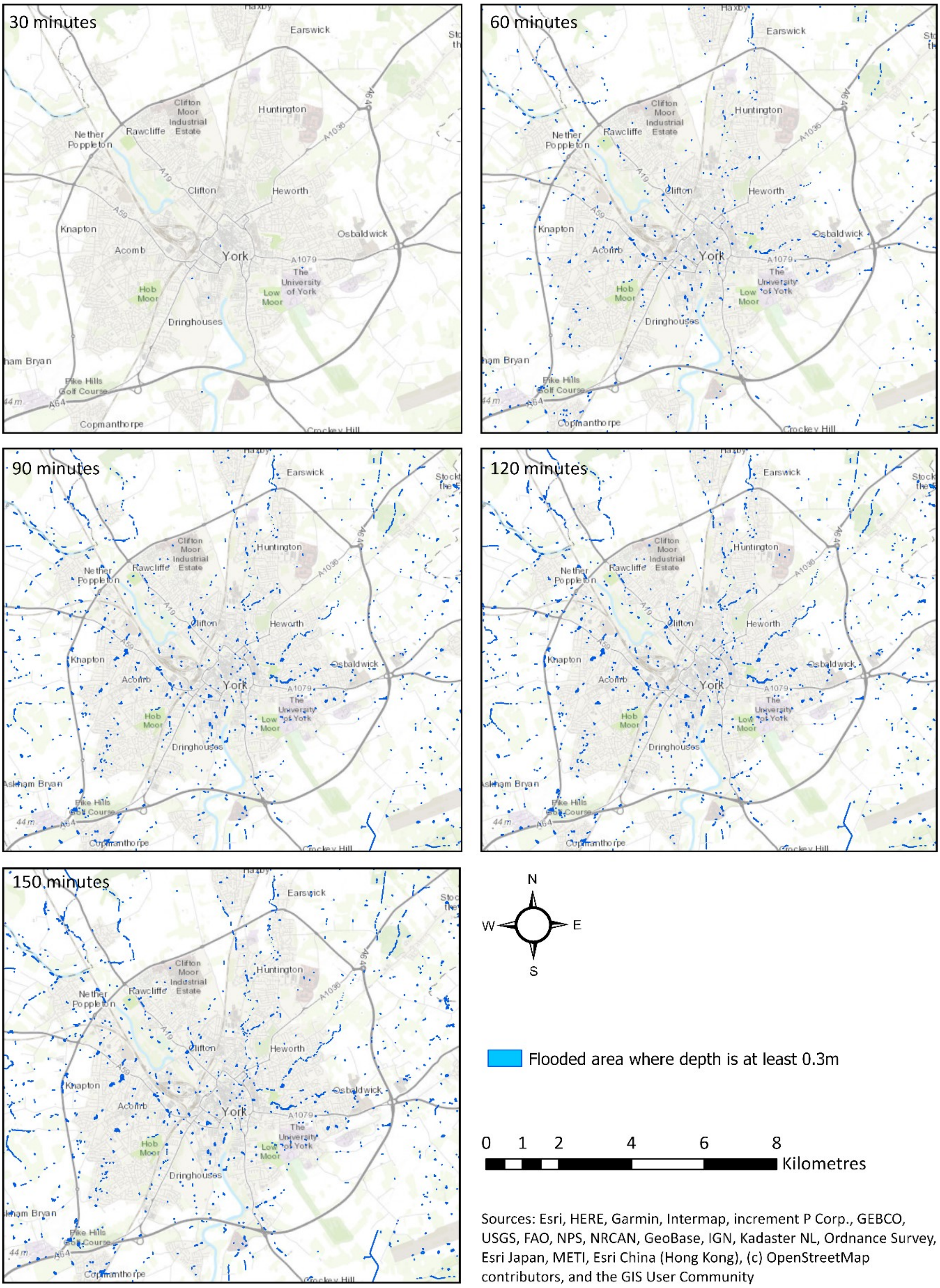


Figure 7.5. 1 in 30yr flood event modelled flood area (area shown is where the flood depth was at least 0.3m with an area of 125m^2, conditions required for there to be an impediment to road travel). Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

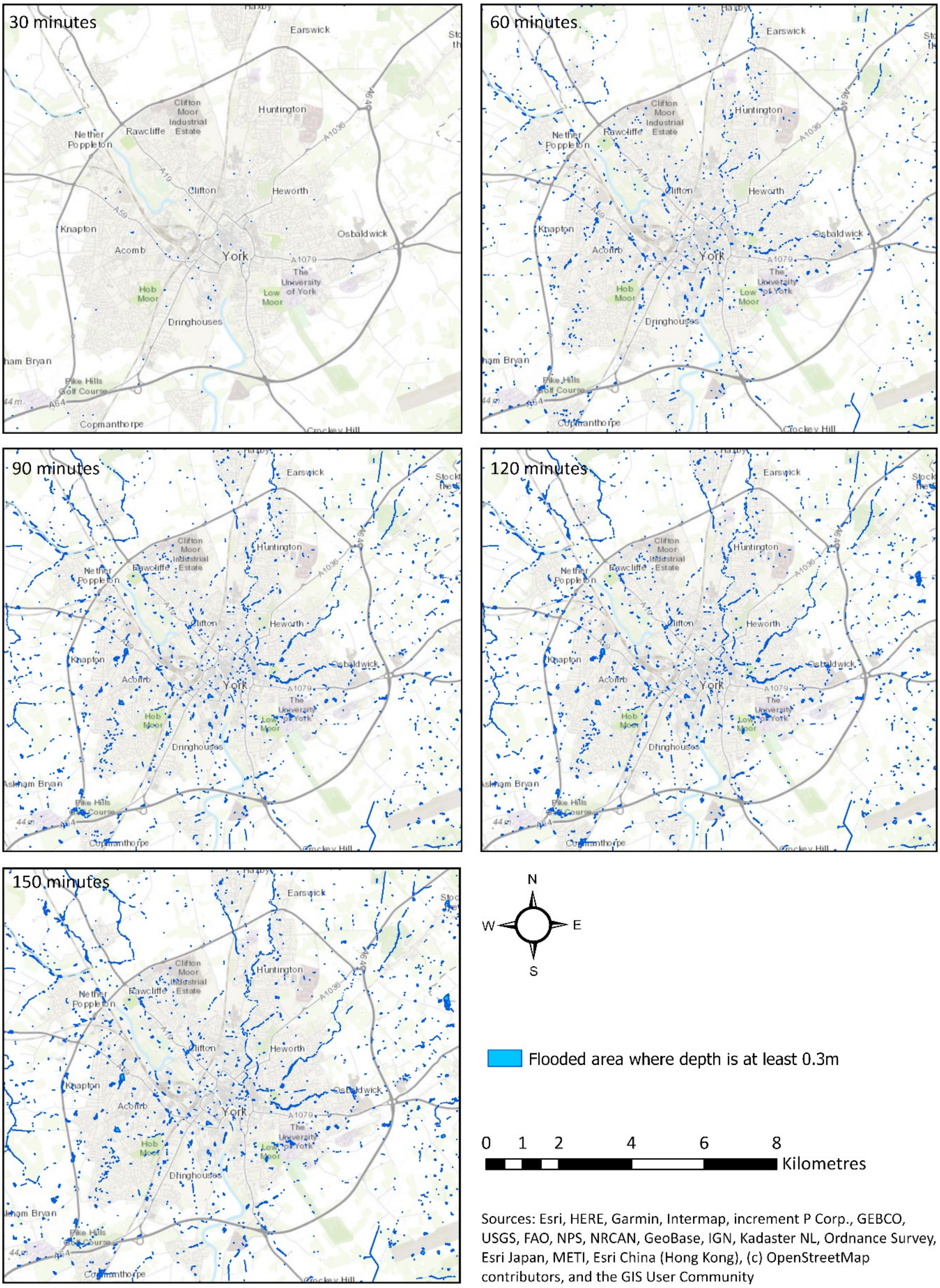


Figure 7.6. 1 in 100yr flood event modelled flood area (area shown is where the flood depth was at least 0.3m with an area of 125m^2, conditions required for there to be an impediment to road travel). Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

Whilst the area susceptible to flooding looks large on the flood maps (Figure 7.5 and Figure 7.6), not all of this flooding is affecting the road network. Figure 7.7 represents the percentage of the road network in the study area covered in disruptive flooding (at least 30 cm deep and 125 m^2 in area), over the course of the three hour flood simulation. It is raining for the first hour of the study (see Figure 4.2). Figure 7.7 reveals that only a small proportion of the road network is directly affected by flooding, a maximum of 0.9% (9.6 km) for the 1 in 100 year magnitude and 0.6% (6.6 km) for the 1 in 30 year event. The 1 in 100 year storm event results in a greater proportion of the road network being flooded, which is unsurprising as a greater area was flooded overall (Figure 7.6). For both magnitudes, the amount of road affected by disruptive flooding increases rapidly from 0.5 hours into the simulation before levelling off from 1.5 hours for the 1 in 30 year storm and 1 hour for the 1 in 100 year storm.

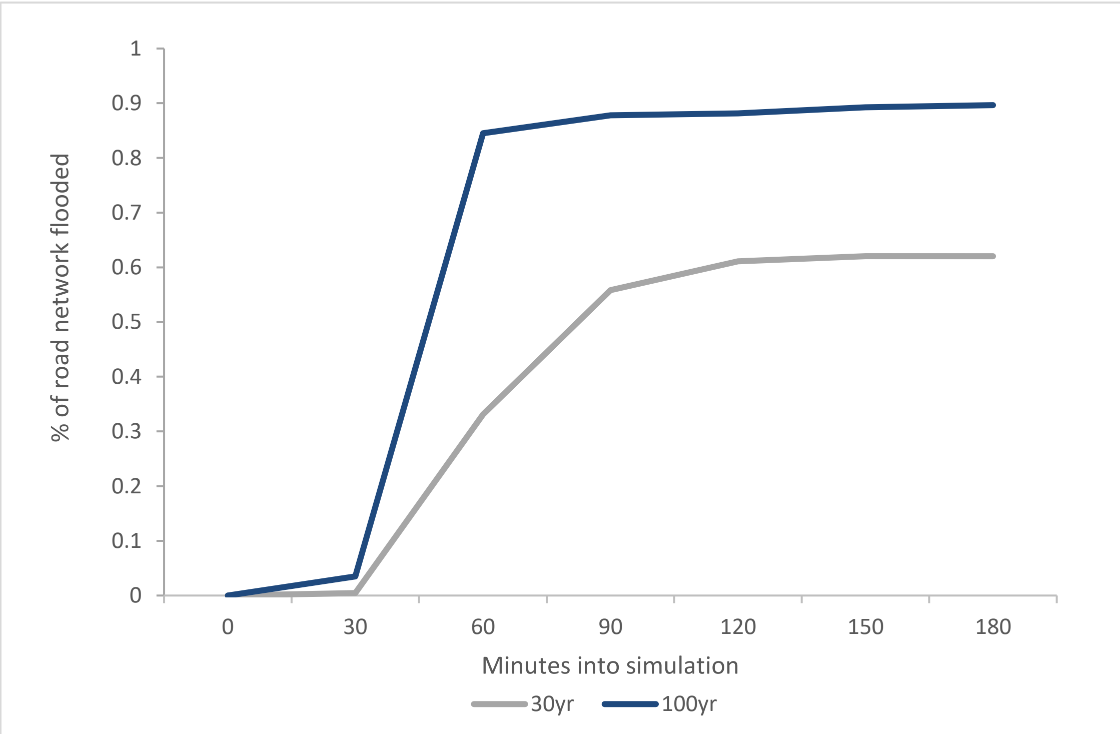


Figure 7.7 Percentage of the road network in the study area which experiences disruptive flooding (at least 0.3 m depth and 125 m^2 area). Contains data which are copyright of Ambiental Risk Analytics, used with permission.

7.3.2 Modelled trips to school and work

The non-flood baseline trips to school and workplaces across York were mapped to see the spatial distribution of people across the network at each time point. The travel profiles used in this chapter (Section 7.2.4) specify the peak road usage is at 08:30 hours for commuters. This is reflected in the maps for commuters (Figure 7.8 to Figure 7.15) with 08:30 hours having the most number of people on the roads and 07:00 hrs the least. Commuters have a more homogenous spatial pattern in the usage of roads, with particularly high volumes of people on the ring road and city centre roads. This is not surprising due to hierarchy function in the network analyst tool (Section 4.2.3) simulating the behaviour of people travelling on major roads. The data for 09:00 hours is interesting as there is more spatial variation in road usage, reflecting that some have finished going to work and others are travelling in later.

The data for school children only have two time points, as primary school children travel between 8-9 am, according to the underlying travel profiles (Figure 7.13 and Figure 7.14). The city centre roads have the highest number of children using them. This likely reflects the fact more primary schools are located in the city centre than on the outskirts and surrounding villages (see Figure 5.1). In contrast to the commuter data, the A road ring road is not utilised much at all. This confirms that journeys to schools are shorter and more local than journeys to workplaces. Some of the roads highlighted as high usage are ones leading up to schools, so it is therefore unsurprising lots of children are using them.

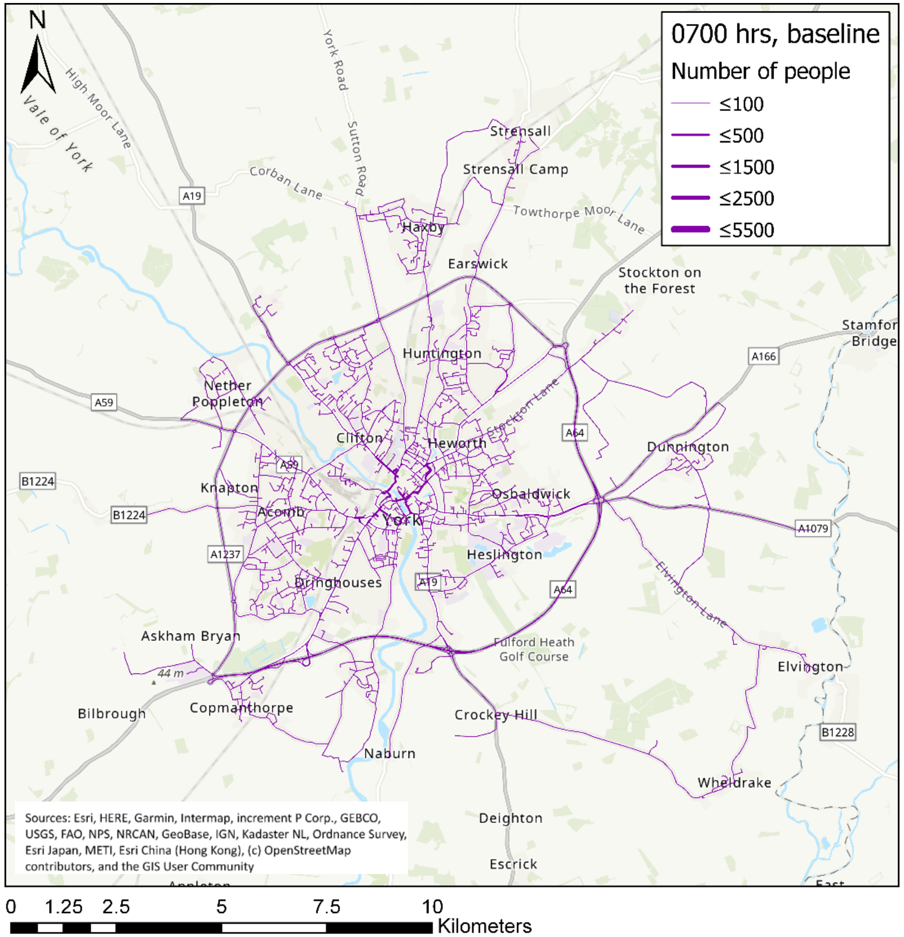


Figure 7.8 The number of commuters on the roads at 0700hrs in non-flood baseline scenario. Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

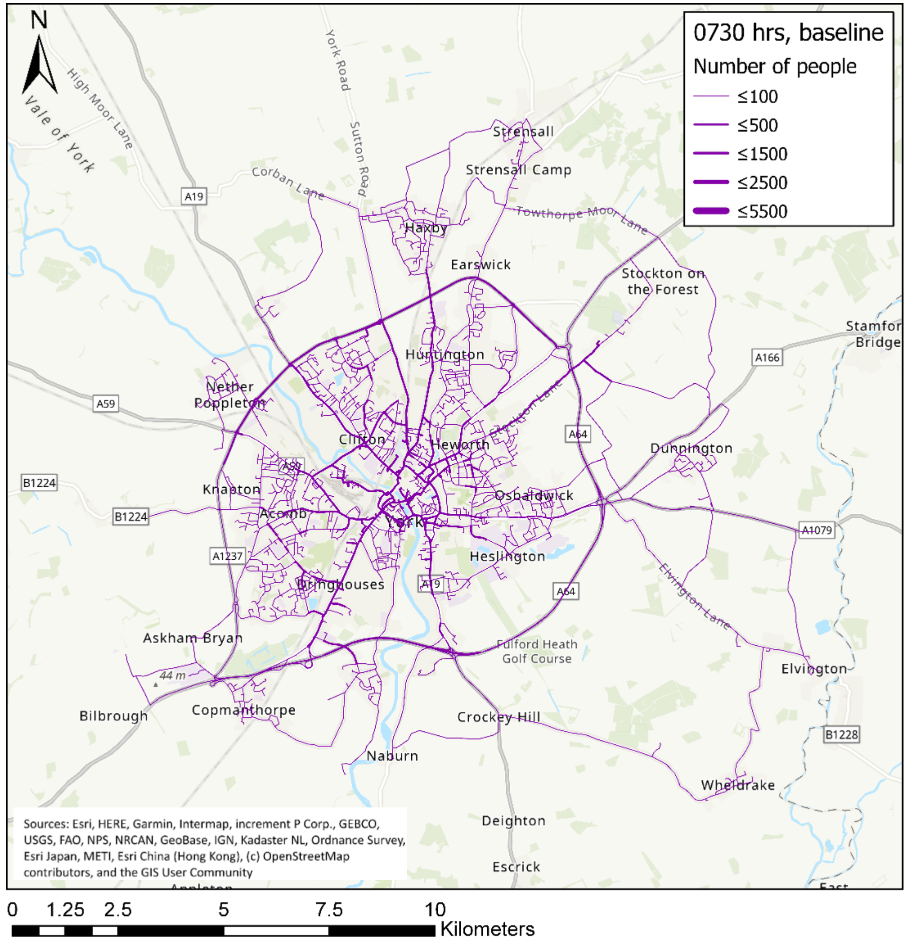


Figure 7.9 The number of commuters on the roads at 0730hrs in non-flood baseline scenarioContains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission

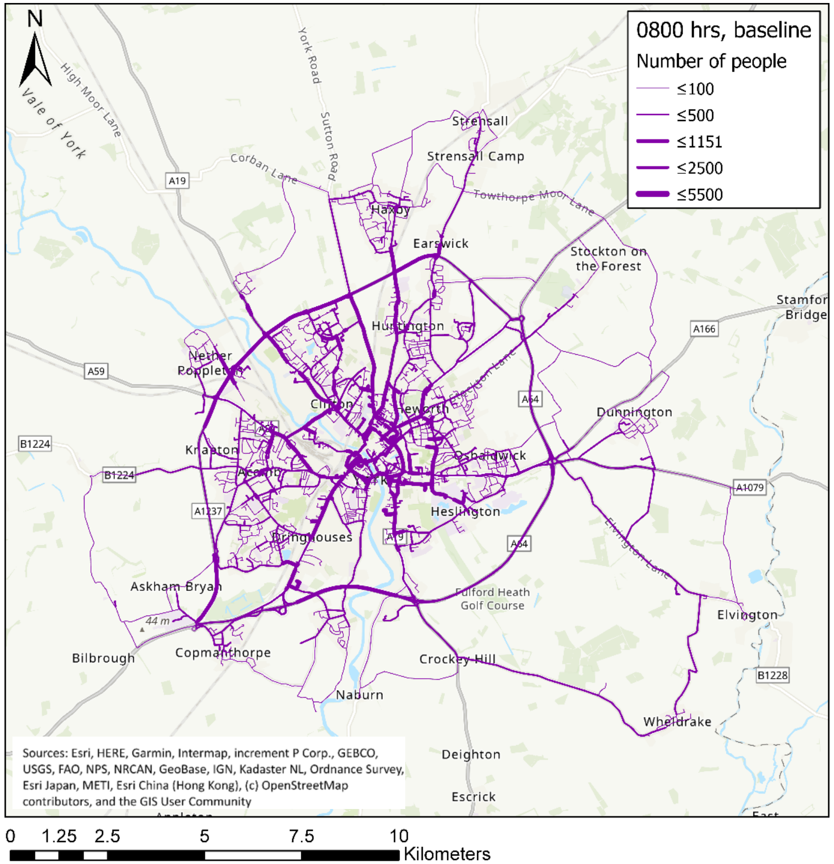


Figure 7.10 The number of commuters on the roads at 0800hrs in non-flood baseline scenario. Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

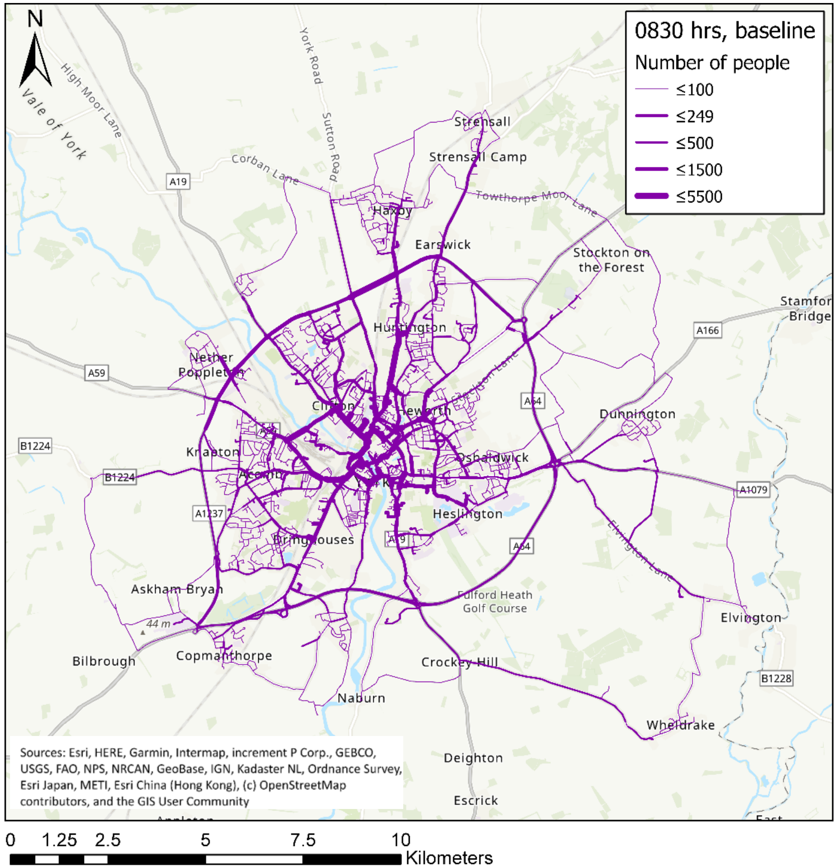


Figure 7.11 The number of commuters on the roads at 0830hrs in non-flood baseline scenario. Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission

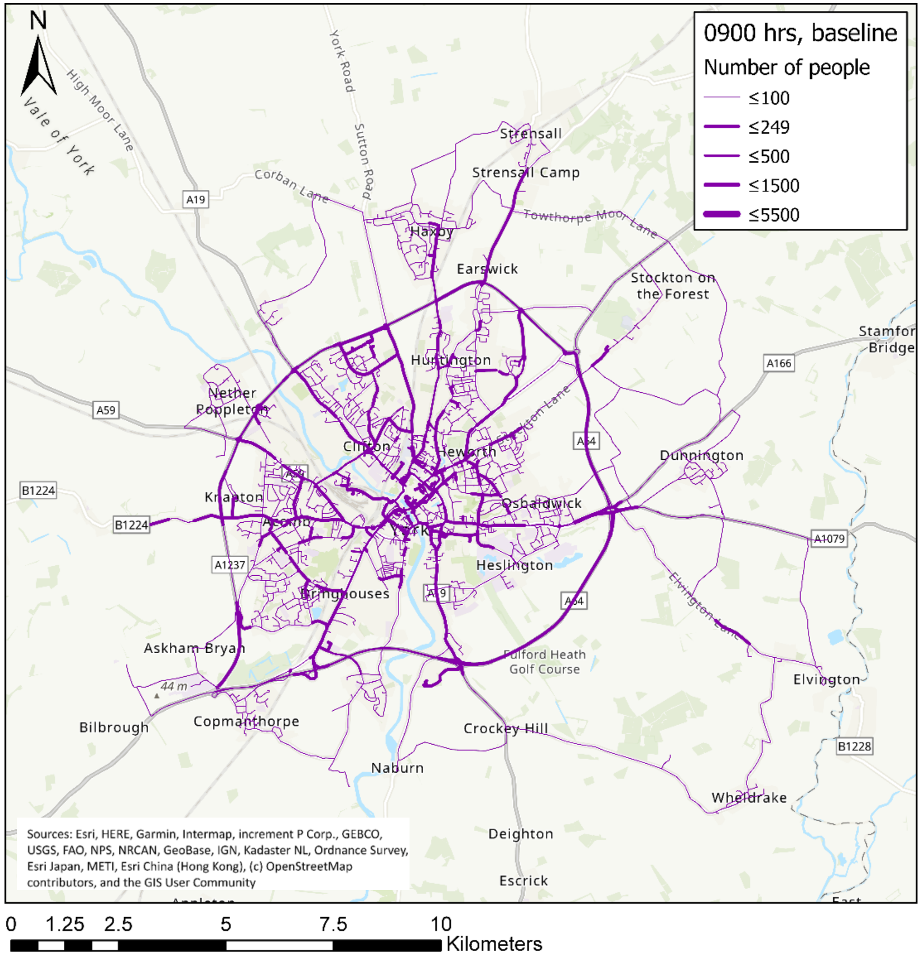


Figure 7.12 The number of commuters on the roads at 0900hrs in non-flood baseline scenario. Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

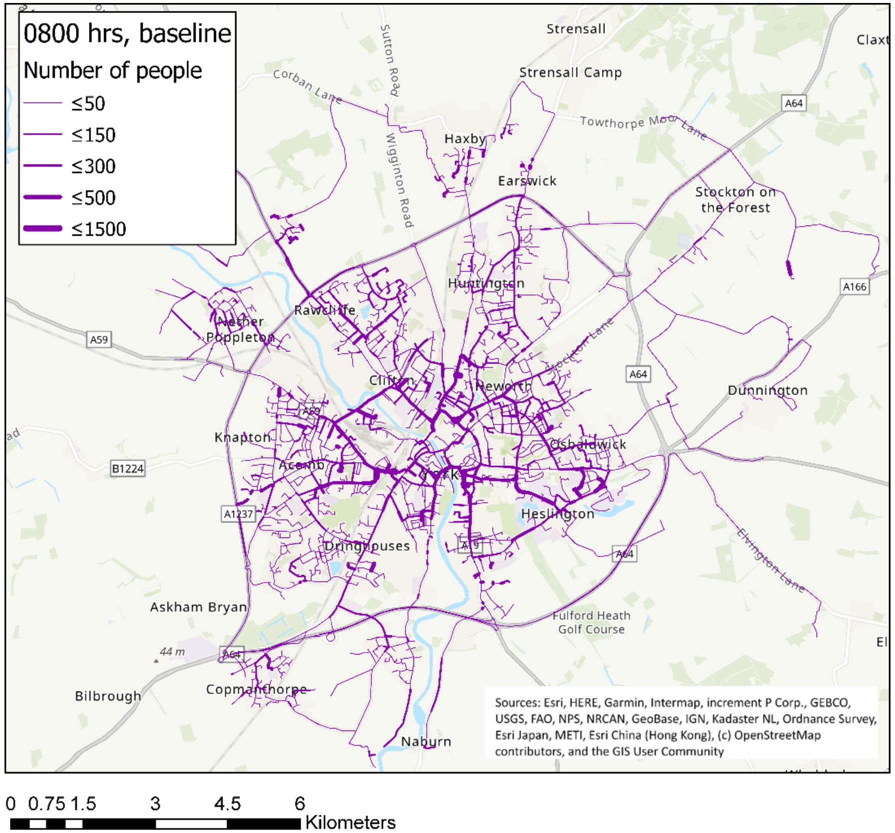


Figure 7.13 The modelled number of primary school children on the roads at 0800hrs in non-flood baseline scenario. Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

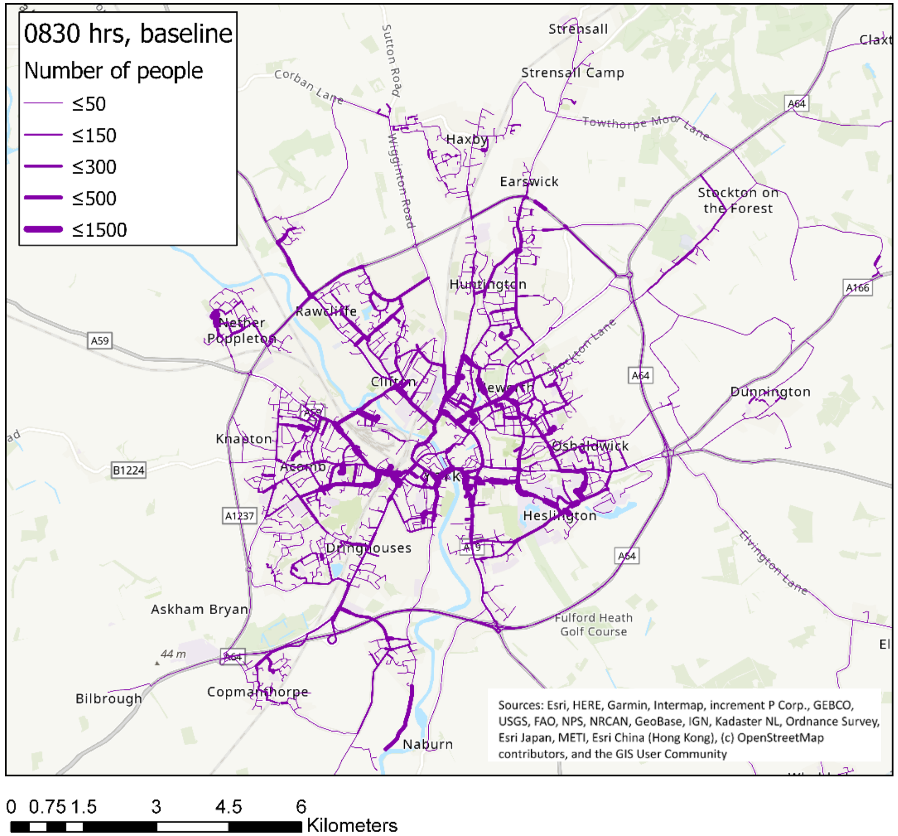


Figure 7.14 The modelled number of primary school children on the roads at 0830hrs in non-flood baseline scenario. Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

7.3.3 Travel disruption by scenario and population group

The first measure of travel disruption presented is the number of destinations that were inaccessible from any origin at each time point across the scenario (Table 7.2). There is an increase in the number of destinations where no origin could reach them between the 30yr and 100yr flood sizes, which is expected as there was a greater percentage of the road network flooded (Figure 7.7). For the first time point where there is sufficient disruptive flooding (07:00 hours FT1, 08:00 hours FT2 and 09:00 hours FT3), the number of destinations where no origins could travel to them is just over double the number in the 30yr equivalent scenarios. For other time points in the scenarios the number of destinations is only slightly higher (1 or 2 more).

Changing the time of flood onset means the number of destinations inaccessible from any origin shifts through the day, which is as expected. The number of destinations with no origins able to travel to them is highest at 0700hrs in FT1. This is due to there being fewer origins where people are travelling from at this time of day (as the number of commuters from the travel survey data is lower), so whilst there is less disruptive flooding, there are fewer places which need to be flooded to prevent travel from these contributing origins. For example, some workplace zones only have one origin where travel is modelled from at 07:00 hours, therefore the risk of this one origin – destination pair being split is higher. These results do not mean any travel to the destination is prevented, as travel from every part of the city is not modelled, just from the origins sampled.

The data were investigated to understand the spatial location of these and the reasons for no route being possible. There were three categories of reasons. Firstly, there was flooding close to the destination preventing access to it from anywhere. This was the case for 6/8 of the destinations in the 1 in 30 year scenarios and 7/21 of the destinations in the 1 in 100 year scenarios. These destinations appeared frequently across all of the modelled scenarios. Secondly, there were destinations where there were a small number of origins, often only one, assigned to that destination at that time point and flooding close to the origin prevented travel to it. This was the case for 2/8 of the destinations in the 1 in 30 year scenarios and 5/21 of the destinations in the 1 in 100 year scenarios. These destinations would only appear once or twice across the scenarios. Finally, the remainder of the destinations were those which were only present in one or two of the scenarios and appeared to be due to flooding blocking travel somewhere between the origin and destination, as there was no obvious flood next to either.

Table 7.2 The number of destinations where travel from the origins to the destination was not possible for commuters. The percentages represent the data as a proportion of the 215 destinations. Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

		30yr			100yr		
Time (hrs)	Baseline	FloodTime1	FloodTime2	FloodTime3	FloodTime1	FloodTime2	FloodTime3
0700	0	7 (3.3%)	0	0	16 (7.4%)	0	0
0730	0	6 (2.8%)	0	0	7 (3.3%)	0	0
0800	0	5 (2.3%)	4 (1.8%)	0	7 (3.3%)	10 (4.6%)	0
0830	0	5 (2.3%)	5 (2.3%)	0	6 (2.8%)	7 (3.3%)	0
0900	0	5 (2.3%)	6 (2.8%)	6 (2.8%)	7 (3.3%)	7 (3.3%)	12 (5.6%)

The average travel time of the origins to each destination were calculated for each scenario a sample are presented in Figure 7.15 (commuters) and Figure 7.16 (school children). For commuters across all destinations, there is a noticeable difference to the baseline average travel times from the modelled flood time scenarios. Generally, in the 1 in 30 year scenarios, FT1 shows an increase in average travel time, compared to the baseline, across the five time points, FT2 an increase from 08:00 hours and FT3 an increase at 09:00 hours. The increase in travel time during FT1 and FT2 generally converges at an average of 8:30am, indicating that despite greater area of flooding the level of disruption remains unchanged. FT3 sees disruption from 09:00am where it diverges from the baseline but does not reach the level of FT1 and FT2 in the 1 in 30 year scenario.

For the 1 in 100 year event, broadly the same patterns are shown as for the 1 in 30 year scenario. Average travel time at 07:00 hours are slightly higher than the baseline and the equivalent in the 30yr flood scenario. Again, data for FT2 diverges from the baseline with greater disruption seen from 7:30am – 8:00am, but it does not reach the same level as the FT1 for all the examples. FT3 shows more divergence from the baseline earlier in the day (08:30 hours), with a greater increase in average travel time than shown in the 30yr scenario. Whilst there are differences, there is not always a large numerical difference in average travel time between the 30yr and 100yr, perhaps one or two minutes, for example Figure 7.15a and b. There are also destinations where there is little or no disruption from flooding.

The data for school children is only measured at two time points, 08:00 and 08:30 hours, and Figure 7.16 displays less variation between the scenarios than for commuters. This is likely due to school children having shorter journeys to complete, reducing the chance of encountering an area of floodwater. It may also indicate that schools are located in areas less at risk from pluvial flooding. The examples for the 1 in 30 year magnitude show that there is a difference in average travel time, however this is mostly less than 1 minute. The only exception is the school with the largest catchment area where there is an increase of two minutes on the average travel time. As the journeys to school are short, a small detour in the route will not have a large effect on the time taken.

A comparison of the number of people using road segments at each time point for each scenario, was created for both commuters and primary school children through maps. Figure 7.17 shows the comparison at 08:30 hours for FT1 between the baseline, 30 year and 100 year flood events. The figure highlights that there are effects of flooding, with sections of road which were used in the baseline scenario no longer being used in the flood scenarios. For example, a part of the A road in the north of the city and the road alongside York station in the centre (Leeman Road). This demonstrates that there is a spatial variation to disruptive flooding for transport links and the places where travellers are diverted to.

For primary school children, there are again noticeable spatial differences in the distribution of travellers across the city (Figure 7.18). As with commuters, the road leading up to York station is affected causing more use of other roads. This appears for both flood magnitudes. There are some spatial changes in the 1 in 100 year magnitude event which are not apparent for the 1 in 30 year event. Firstly there is an area in the North East of the city at Dodsworth Avenue which had previously been a well-used road and does not have anybody using it during the 1 in 100 year storm event. This indicates that there is flooding along this road which prevents travel. Secondly, as was seen in the commuter data, the A road in the North East corner of the ring road has no travel along it indicating a flood barrier, and this is seen for both flood magnitudes. For the 1 in 100 year event there are a greater number of people using rural roads in the North East of the study area which indicate there is flooding which is preventing travel along other roads, leading to detours.

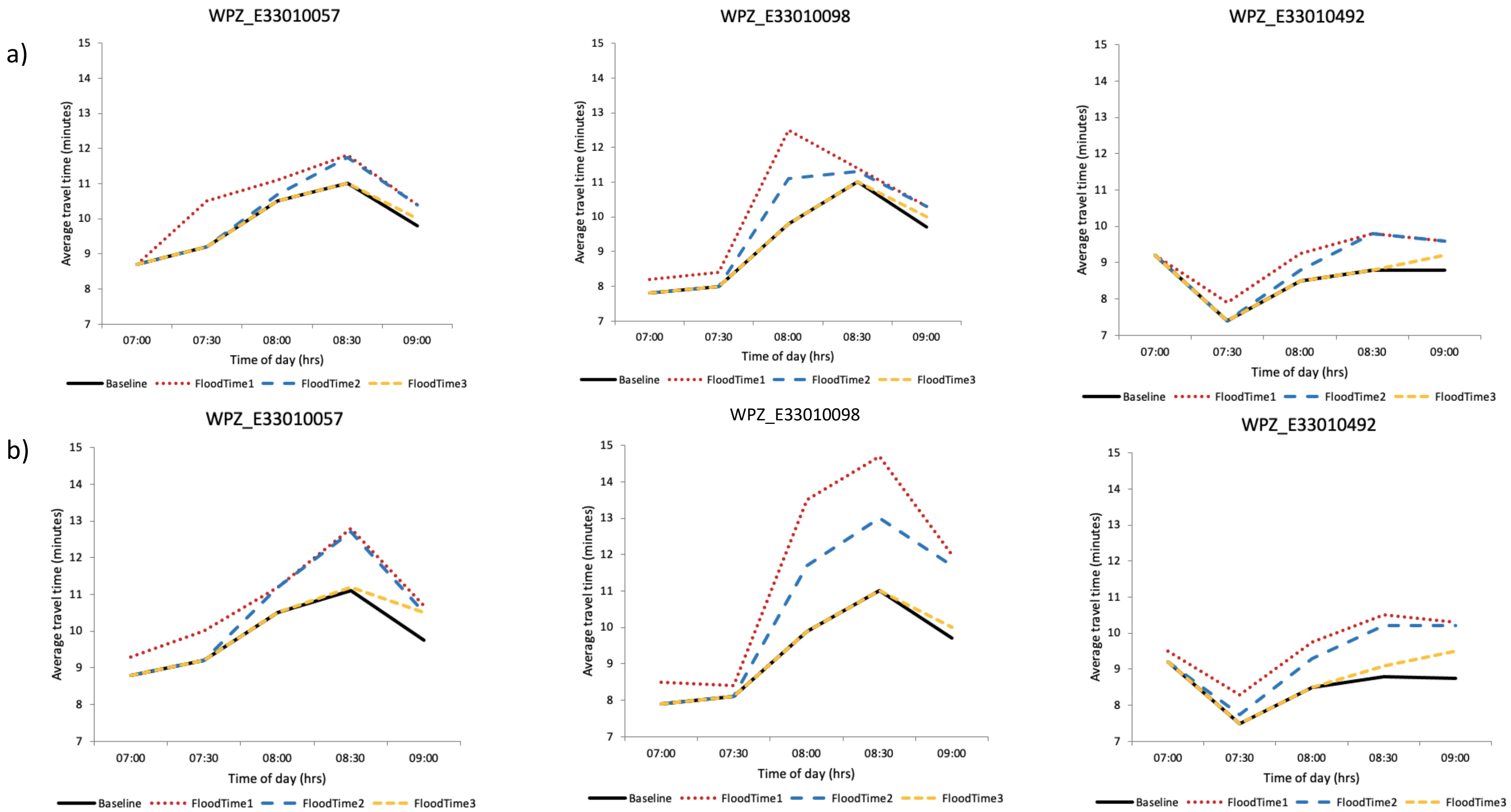


Figure 7.15 Sample of workplace destination graphs comparing the average travel time between scenarios, for a 30yr flood magnitude (a) and 100yr magnitude (b).

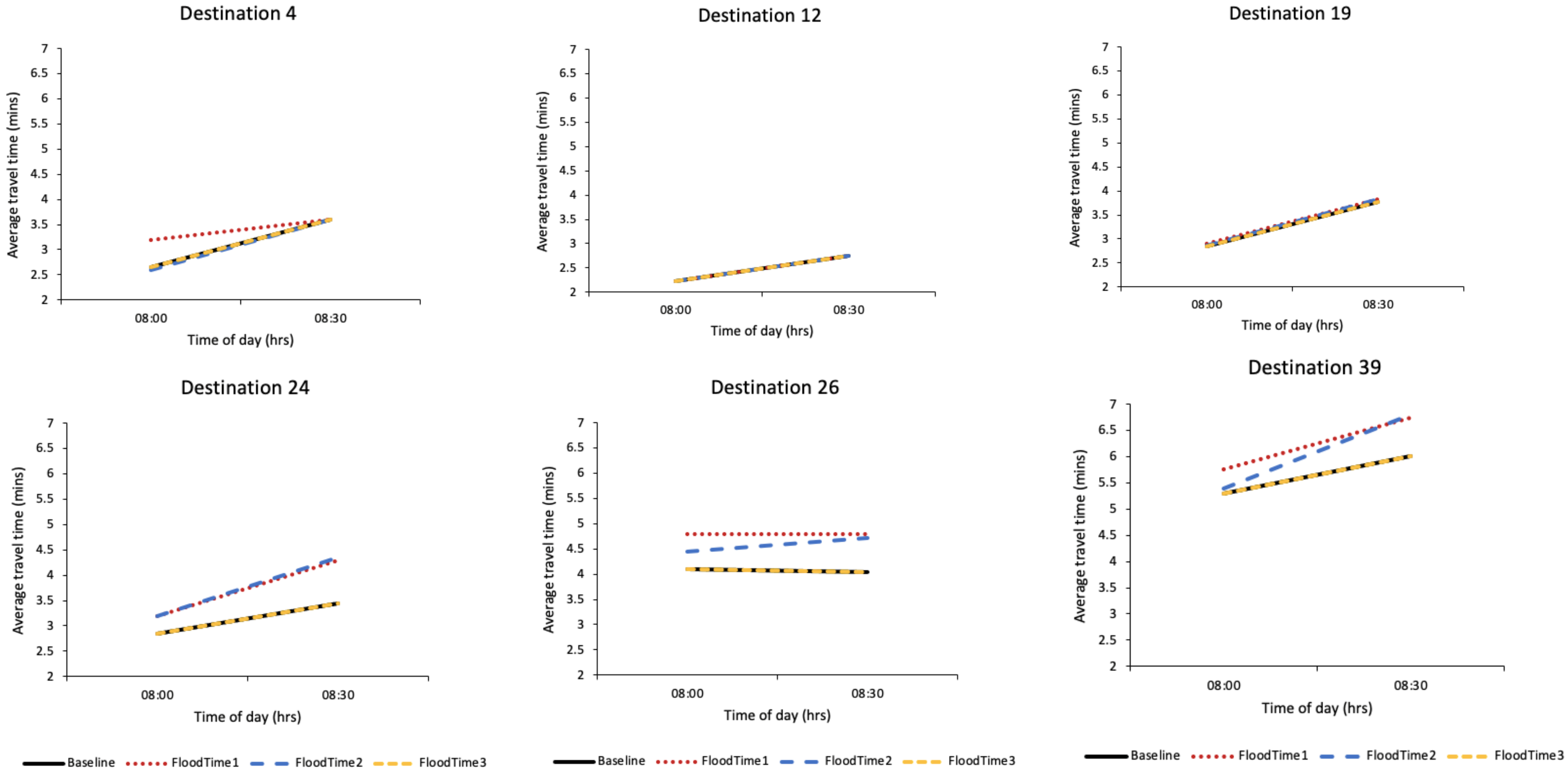


Figure 7.16 Sample of school destination graphs comparing the average travel time between scenarios, for a 30yr flood magnitude. Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

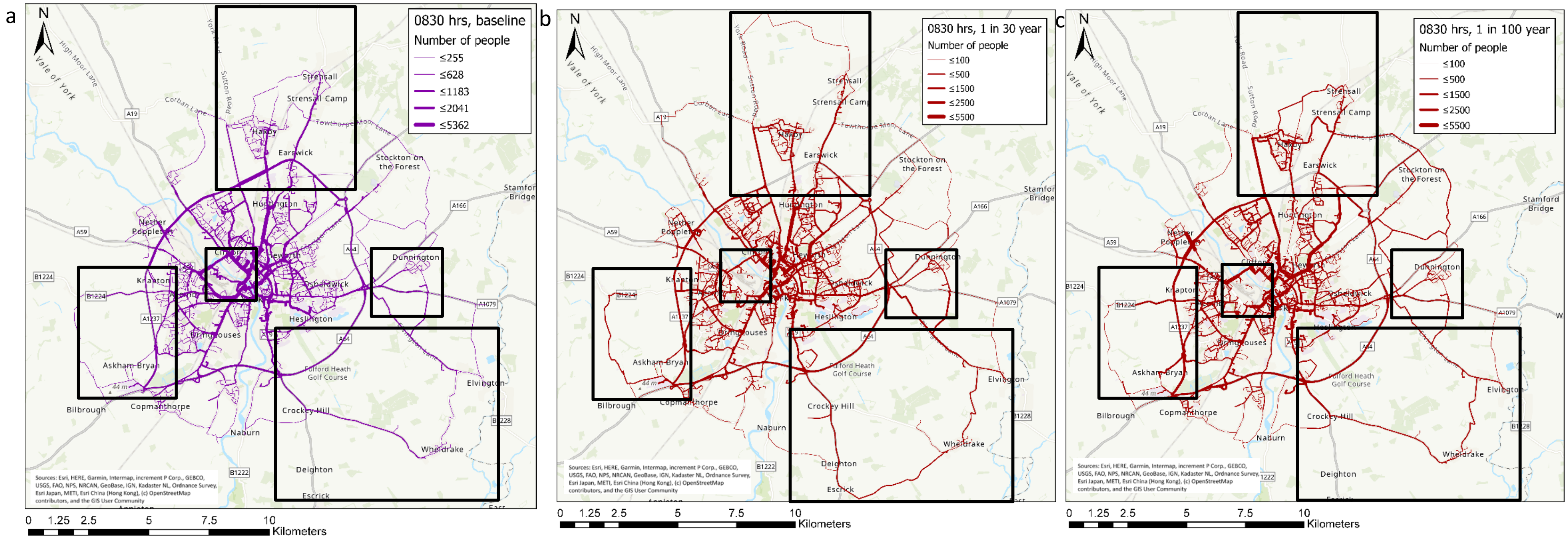


Figure 7.17 The modelled number of commuters on the roads at 0830hrs for a) non-flood baseline scenario b) 1 in 30yr event c) 1 in 100 year event. Black boxes indicate areas where routes noticeably diverge from the baseline scenario. Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

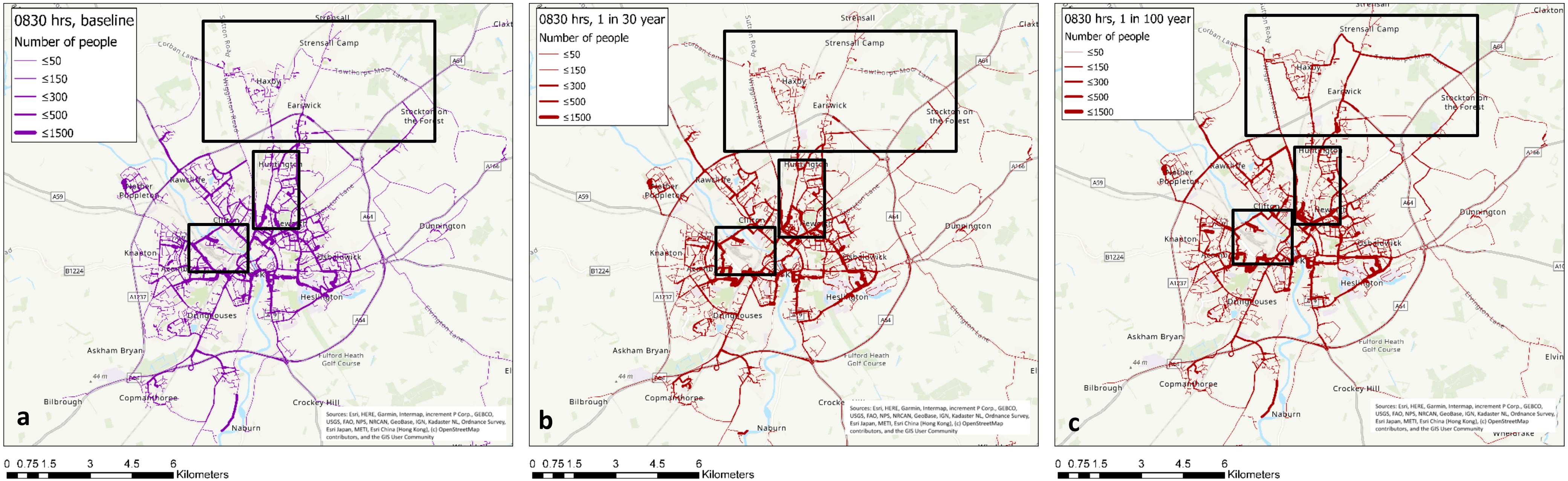


Figure 7.18 The modelled number of primary school children on the roads at 0830hrs for a) non-flood baseline scenario b) 1 in 30yr event c) 1 in 100 year event. Black boxes indicate areas where routes noticeably diverge from the baseline scenario. Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

By visually examining Figure 7.17 and Figure 7.18, roads which were used frequently by travellers in the baseline non-flood scenario, but not in the flood scenarios, were identified. These could be indicators of critical roads for each population group, as the blockage of these well-used roads would cause more traffic to use other roads. Figure 7.19 highlights such roads in York and whether they were important for one or both population groups.

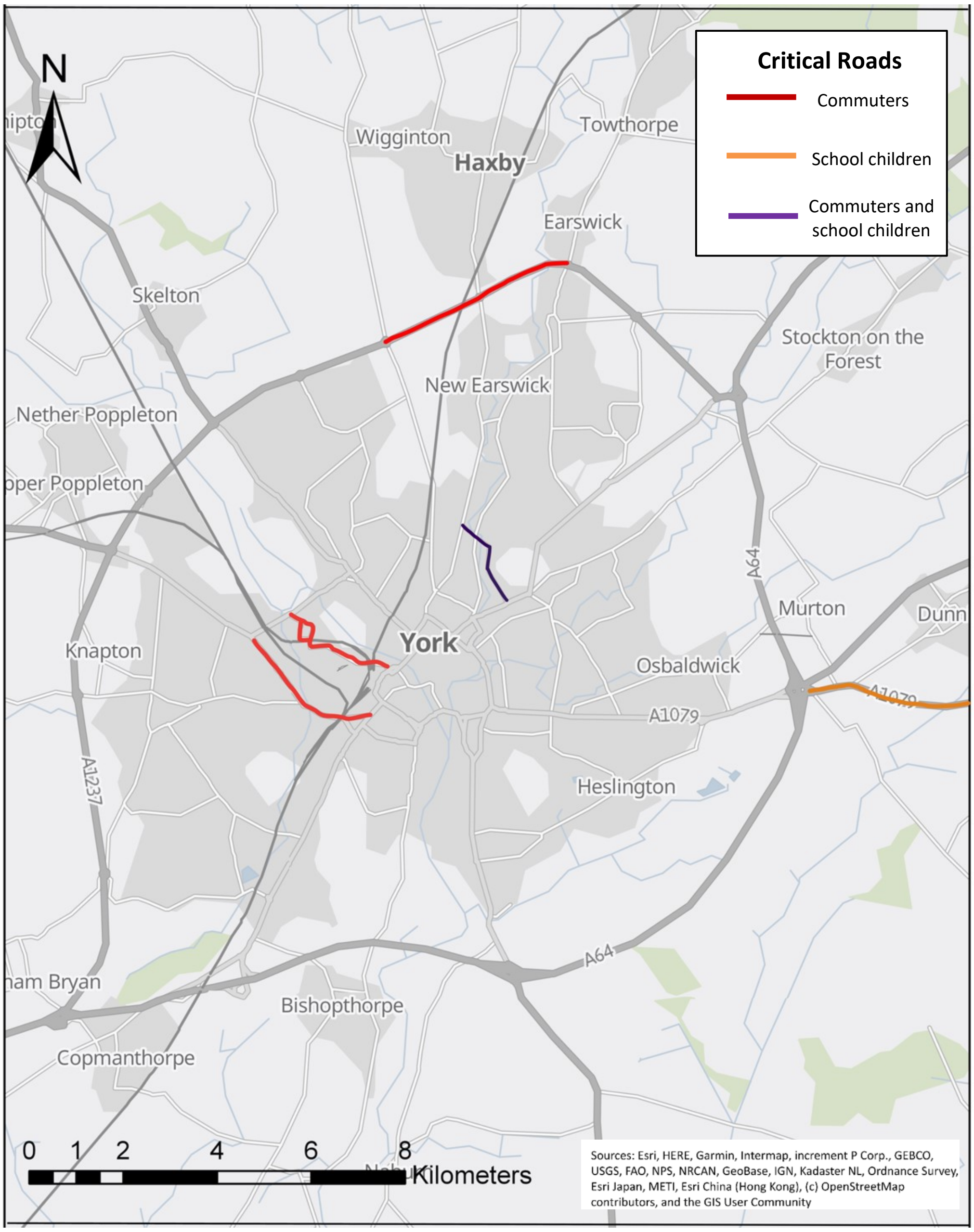


Figure 7.19 Potential critical roads in York based on route results. Red are critical for commuters and school children, purple for school children only and orange for commuters only. Contains public sector information licensed under the Open Government Licence v3.0.

To summarise the travel time data for the whole city and answer the aim of the thesis, the time lost in each flood scenario at each destination were aggregated. Figure 7.20 and Figure 7.21 show that commuters have the higher time lost than school children, which is expected as the journeys to work are longer and there are more commuters than school children so more origins-destination pairs. For both commuters and school children, there is an increase in the amount of time lost at work or school in every scenario. The values for a 1 in 100 year event are greater than a 1 in 30 year event across all scenarios. The time of flood onset does affect the time lost at work or school, most noticeably the difference between FT1 and FT3, which is greater than the difference between flood magnitudes. This indicates that flood onset time is having a larger effect than flood magnitude.

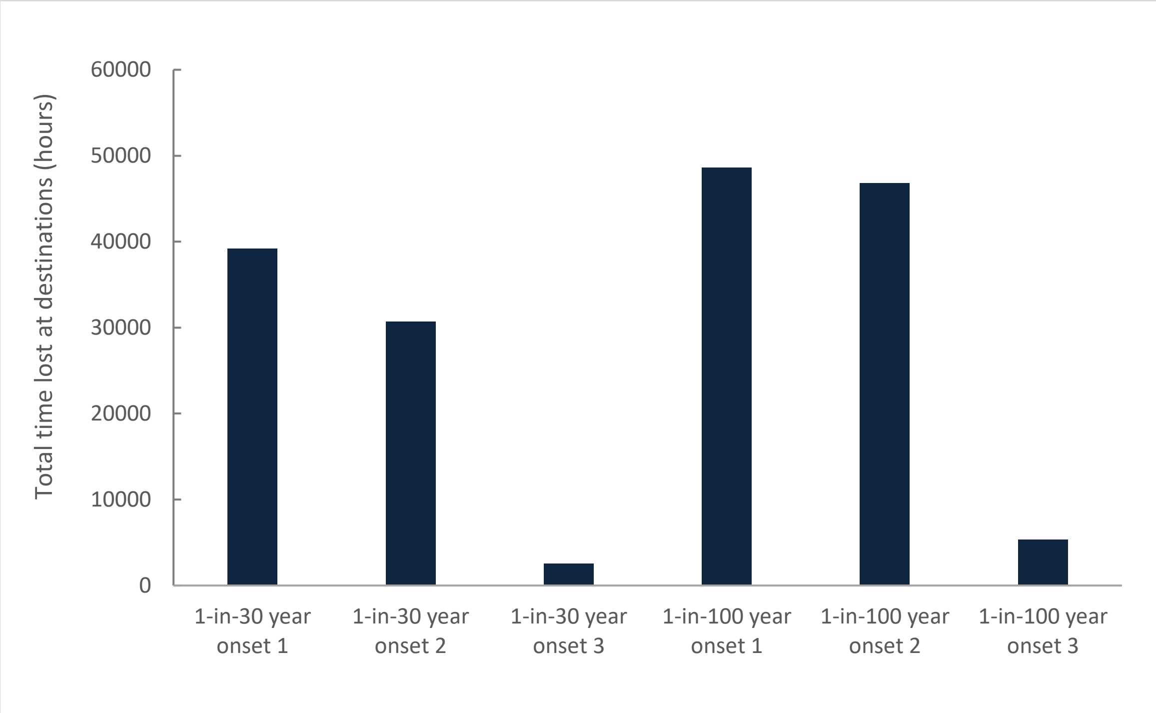


Figure 7.20 Total time lost at all destinations for commuters in each scenario. Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

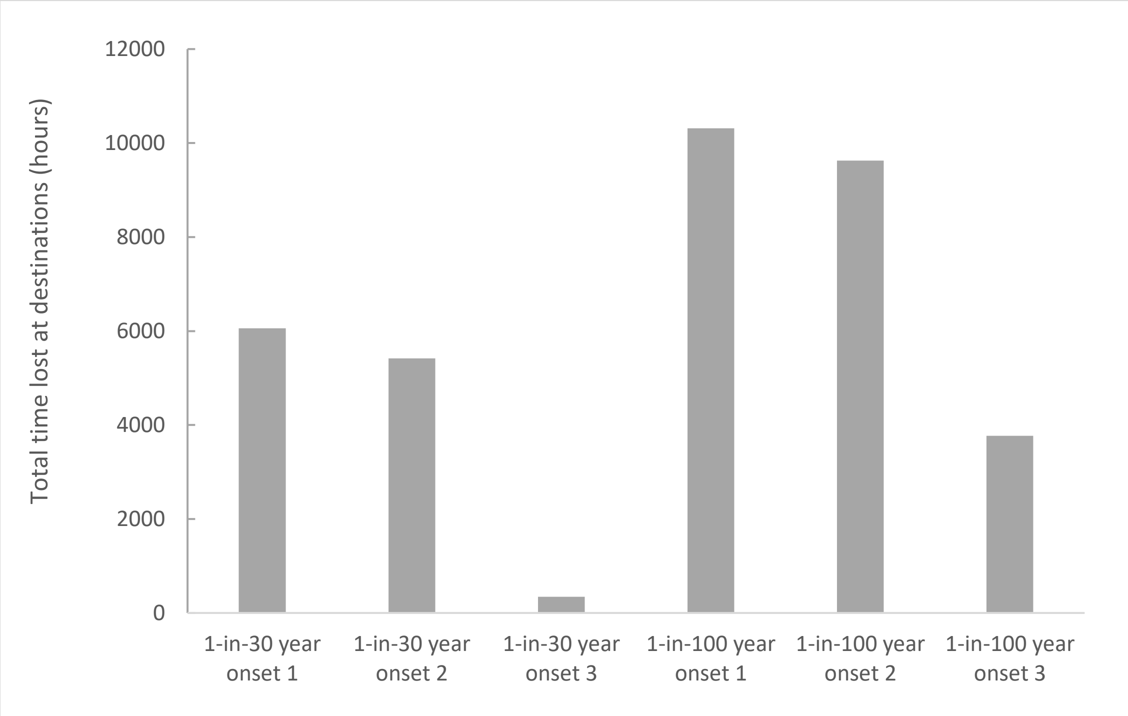


Figure 7.21 Total time lost at all destinations for primary school children in each scenario. Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

Finally, it was investigated if the same destinations across all scenarios were experiencing the greatest travel disruption to them, to see if those most susceptible to disruption could be identified. Firstly for commuters, Table 7.3 presents the correlation coefficients from the Spearman's rank correlation tests and these show that the destination time lost at destinations for the FT1 and FT2 scenarios are statistically similar to each other, for both magnitudes of flood, with high to medium-high agreement. The statistical significance of the correlations is <0.001. FT3 has a lower level of agreement, however still shows correlation to the other flood scenarios, with greatest agreement between FT3 1 in 30 year and FT3 1 in 100 year. Overall, this shows that there is consistency between which destinations have the most time lost at them, therefore are the most susceptible, for FT1 and FT2, FT3 has a different pattern in the most affected destinations.

Table 7.3 Correlation coefficients for commuter travel data for 215 destinations. The columns are labelled as Flood Time (FT), the flood time of onset (1, 2 or 3) and the flood magnitude (30 year or 100 year). Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

	FT1_30yr	FT2_30yr	FT3_30yr	FT1_100yr	FT2_100yr	FT3_100yr
FT1_30yr		.963	.466	.841	.852	.499
FT2_30yr			.521	.813	.839	.539
FT3_30yr				.415	.402	.712
FT1_100yr					.910	.465
FT2_100yr						.520
FT3_100yr						

To view travel disruption by destination spatially, workplace zone centroids were visualised, setting their symbol sizes to be proportional to time lost at work from flood-related travel disruption. As the Spearman's Rank test showed strong correlation between travel disruption for scenarios relating FT1 and FT2 but weaker correlation between these four scenarios and those for FT3, an example of a FT3 scenario is presented here for comparison with a scenario from FT1/FT2. Figure 7.22 demonstrates that there is spatial clustering of the destinations with most time lost at them, around York station in the western centre of the city, with a few smaller clusters on the Northern and Western parts of the city. This appeared to be related to flood water accumulating in these areas. For FT3 (Figure 7.23) there is a more homogenous spread of disruption but with a cluster again around York station clearly defined.

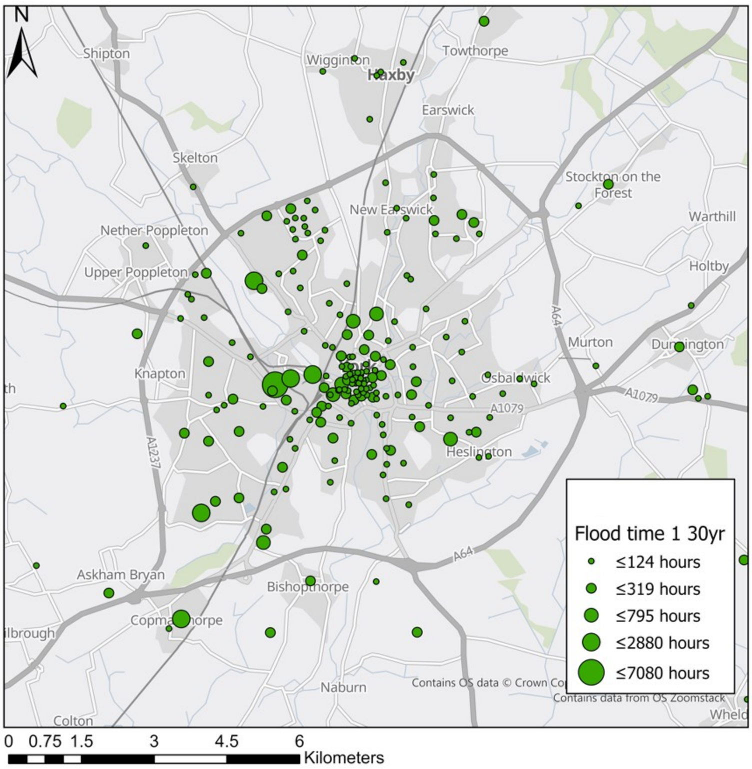


Figure 7.22 Time lost at work due to travel disruption per workplace zone for flood onset time 1 and 2.Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

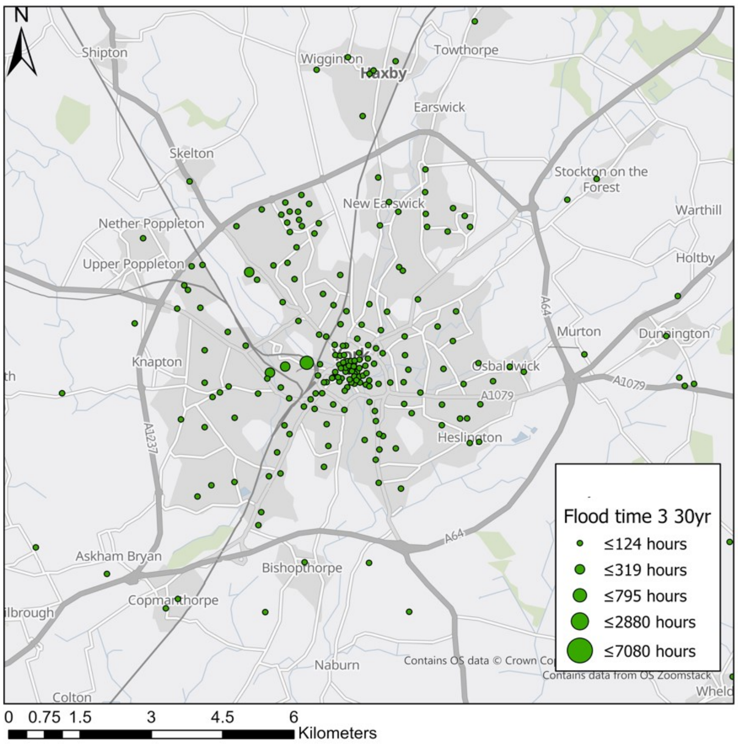


Figure 7.23 Time lost at work due to travel disruption per workplace zone for flood onset time 3.Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

For school children, the Spearman's Rank showed a broadly similar pattern of correlation values, with more agreement between FT1 and FT2 and FT3 not statistically similar to the other data sets. However, there is a greater range of correlation coefficients across the scenarios for commuters. For example, the top row shows that scenario FT1 1 in 30 year is highly correlated to FT2 1 in 30yr, but more weakly agree to FT1 1 in 100 year scenario. FT3, 1 in 30 year scenario has the lowest correlation coefficients, with only one scenario reaching the minimum significance threshold. For FT1 and FT2, the data are significant to >0.001 whilst FT3, particularly for the 30 year magnitude, is not significant. This shows that there is greater variation across the scenarios in terms of which destinations are disrupted more than for commuters. In FT1 and FT2 (Figure 7.24), those schools with the greatest increase in additional travel time were those around the station in the central and north western parts of the city, similar to the workplace zones. Again, for FT3 (Figure 7.25), the data are more homogenous with a small spatial clustering around the station and the destination with the largest additional travel increase central southern part of the city.

Table 7.4 Spearman's rank Correlation coefficients for school travel data. Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

	FT1_30yr	FT2_30yr	FT3_30yr	FT1_100yr	FT2_100yr	FT3_100yr
FT1_30yr		.966	.172	.574	.758	.358
FT2_30yr			.199	.555	.749	.332
FT3_30yr				.083	.165	.387
FT1_100yr					.852	.434
FT2_100yr						.444
FT3_100yr						

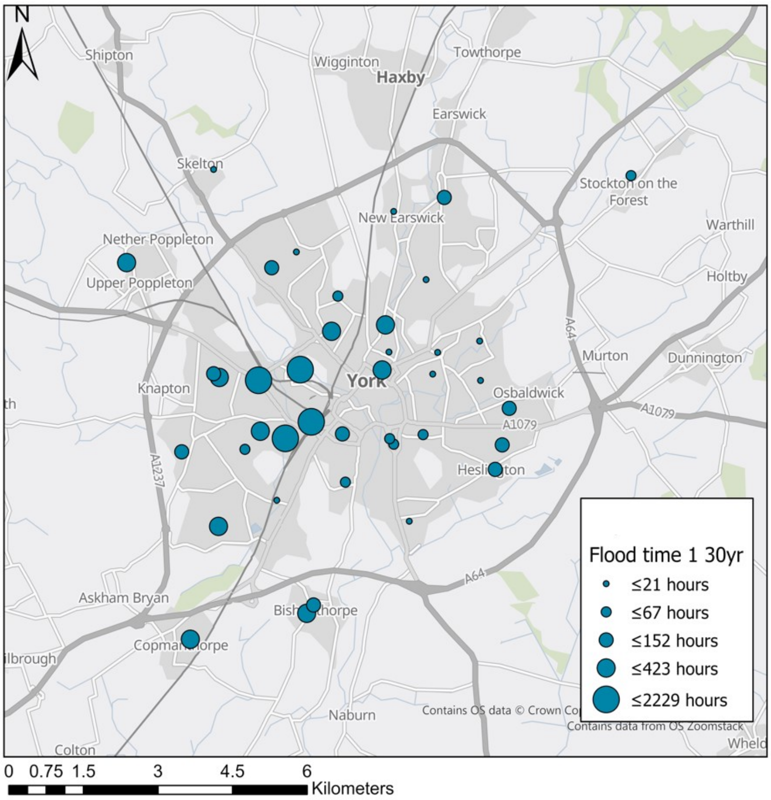


Figure 7.24 Total time lost at school due to travel disruption per school for FT1 and FT2.Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

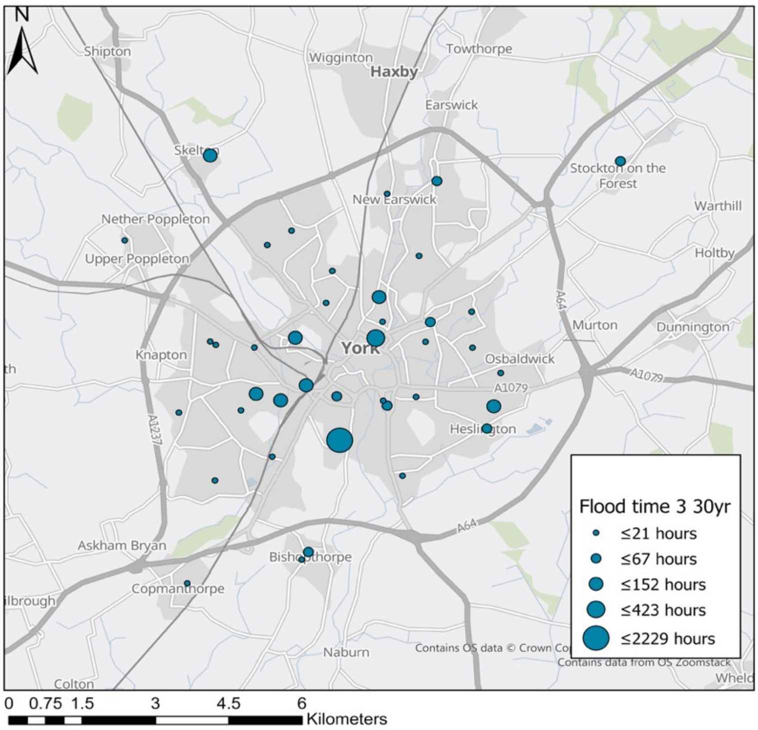


Figure 7.25 Total time lost at school due to travel disruption per school for flood onset time 3.Contains public sector information licensed under the Open Government Licence v3.0. Contains data which are copyright of Ambiental Risk Analytics, used with permission.

7.4 Discussion

7.4.1 Assessment of the results in achieving the chapter research questions

The aim of this chapter was to assess if the time of pluvial flood onset is more important than flood magnitude for disrupting commuter and school travel in York. Three objectives were posed in order to achieve this aim, and each are discussed here in the context of the findings of this chapter.

1. **Analyse the extent to which pluvial flood magnitude and flood onset time affect journey times, due to road network disruption.**

The flood model results show that a greater magnitude storm leads to a larger area of the city being inundated (Figure 7.6) and subsequently a greater amount of the road network experienced disruptive flooding (Figure 7.7). This is expected as more rainfall logically leads to more flooding. However, the translation of this increase in flood water to disruption of the origin-destination journeys is perhaps not as large as expected. For some destinations, there is a difference in the average increase in travel time of a couple of minutes and for some there is no substantial change between magnitude scenarios (Figure 7.15 and Figure 7.16). There was also agreement in which destinations had the largest increases in time lost at destinations for both flood magnitudes in the spearman's rank test (Table 7.3 and Table 7.4). This provides evidence that whilst the size of flood area may change, the same areas are still affected by flooding and the increases in travel time can be slight or the same.

The modelled changes in time of onset were relatively small, with all three time of onsets within a 2-hour window. However, there is a noticeable difference in the average travel time to destinations and time lost at destinations. Figure 7.20 and Figure 7.21 demonstrate that there is a clear reduction in the time lost at destinations between FT1 and FT3, most likely due to more travellers already having reached their destination by the time there is substantial disruptive flooding in FT3. Between FT1 and FT2, there is a relatively small difference in time lost for both flood magnitudes, indicating the flood onset at 06:00 (FT1) and 07:00 (FT2) are similarly disruptive and are both affecting the peak of travellers. When observing changes at the destination level, the Spearman's rank test highlighted the data for FT3 as being weakly correlated with the other scenario outputs, whilst the other scenarios had strong agreement. This demonstrates that the time of day when the flood occurs at seems to have a larger effect that the flood magnitude, supporting the hypothesis.

The temporal properties of risk are often neglected in research (Aubrecht et al., 2013a; Freire & Aubrecht, 2012) despite the time a hazard begins known to be a key factor in its impacts and time dimensions being 'excessively important' in transport systems (Reggiani et al., 2015). The analysis presented in this chapter has highlighted the importance of time when considering risk and has introduced a semi-dynamic method, which differs from other studies of network disruption due to flooding. For example, Kasmalkar et al., (n.d.) used a generic 1-hour window within peak morning commute hours to test the impact of flooding on travel. However, this was not expanded to include more time periods or shift the time of flood onset, like the analysis in this study. Another example is Li et al., (2018) who calculated the delay to journeys from flood events in peak morning travel time, again without examining the effect of changing the time of flood onset around the peak travel period. Therefore, systematically assessing the impact of changing timing of hazard onset represents a novel contribution of the work.

2. **Assess whether commuter's or school children's travel to workplace or school is affected the most.**

Most studies on travel disruption consider transport networks from an engineering point of view, focussing on network connectivity, rather than measuring the impact of road closures on the population which use them (Pfurtscheller & Genovese, 2019). The focus in this chapter on who is travelling, linked to a purpose for the journey (school or work), therefore presents a method which does look at the effect of road closures on the populations which use them. Examining the results presented in this chapter, it seems commuters are more affected than school children, for example the time lost at workplaces is greater than the time lost at schools. Whilst this will partially be explained by there being more commuter journeys happening, the journeys taken to work are also longer than those to school (Figure 7.15 has greater average travel times than Figure 7.16). Therefore, this increases the likelihood of disruption. Debionne et al., (2016) found in their study not all commuters are exposed equally, with those with longer journeys more exposed, similar to the findings in this chapter comparing school and work journeys. Risk theory states that children are more vulnerable than adults to natural hazards (Civil Contingencies Secretariat, 2008). However, in the case of the vulnerability of those in travel to their place of work or school, primary school children, and the adults who travel with them, are less vulnerable partly due to their shorter exposure time on the road network and the location of flooding in the scenarios tested. Therefore, this indicates that more focus could be placed on mitigating the effects for businesses and commuters to future pluvial flood events.

A novel aspect of this research is the comparison between population groups on the impact of flooding on road travel. Other work, including Debionne et al. (2016) examined travel disruption from flooding among commuters in the Gard Region, France; Pregnolato et al., (2017a) examined flood disruption among commuters in Newcastle and Coles et al., (2017) access to emergency services for care homes, but there are few studies that the author is aware of which model the differences in flood-related travel disruption across more than one population group. One similar study is Kunwar et al., (2014), who applied spatiotemporal population data to an activity-based agent model to investigate total evacuation of ten UK cities. They used gridded population data from the Global Urban Rural Mapping Project, version 1 (GRUMPv1) as population data for their agent model, and also included some socioeconomic data on the population to parameterise agent behaviour, to allow investigation of effects on some groups of the population (men, women, workers, students etc.). They found men are more likely to be exposed to flood water while travelling due to their commuting patterns which tended to consist of longer journeys. However, they have not looked at the routes taken by each group or developed the comparison between groups in the same way the analysis in this chapter has done.

3. **Identify schools and workplaces which are most susceptible to disruption from flooding and the effect pluvial flood magnitude and pluvial flood onset time has on identifying them.**

Under the modelled flood scenarios, travel was possible from at least some of the assigned origins to all schools, but some workplace zones became completely inaccessible from their assigned origins. This provides evidence that workplaces are affected more than schools. Upon comparing the identified destinations to the flood area, it became apparent most of the destinations (6/8) with no travel to them during a 1 in 30 year flood scenario were due to flooding at the destination (Table 7.2). For the 1 in 100 year event, a smaller proportion of destinations had flooding which blocked travel to them (7/21) meaning there were more cases where the roads outside the origin points are flooded, or somewhere on the route between the origin and destination experiences flooding. These recurring destinations being cut off in the scenarios indicate these are more susceptible to flooding.

Despite changes in the flood magnitude and time of flood onset, the same destinations were consistently among those experiencing the most disruption in travel to them. As Figure 7.22 and Figure 7.23 highlight, there are spatial patterns for the destinations most affected by flooding, which differ based on the time of day the flood begins. This evidence indicates that time of flood onset is important, and it is affecting the level of disruption more than changes of magnitude. The time of flood onset is important as it relates to the number of people who are in travel, and these results support the claims of risk literature that time of day is an important consideration (Section 2.2). In this research, when a flood event happens even an hour later in the day there is a noticeable effect. To explore this observation further, a Spearman's rank test was conducted to see if those destinations with the most time lost at them were consistent across the scenarios tested. It was found that there was statistical agreement in which destinations were most affected by the flood scenario between flood magnitudes, when comparing the results for the equivalent time of flood onset. This is evidence that the same destinations are most affected, i.e. they are most susceptible to disruption, and are identified even when changing the size of the flood magnitude and time of flood onset. Several studies (Li et al., 2018; Kilgarriff et al., 2019; Pregnolato et al., 2017a) have modelled travel disruption from flooding, but none of these have looked at how such disruption varies across different work and school destinations. The work in this chapter is therefore the first to examine the susceptibility of workplaces or places of education to hazard-related travel disruption, as far as the author is aware.

7.4.2 Implications for flood management in York and England

There are previous instances of pluvial flooding in York causing disruption to travel. For example, the City of York Surface Water Management Plan (City of York Council, 2012) states that flooding on 27th April 2012 at the A19/A1237 roundabout caused ‘major disruption’ to the A1237 ring road around York from 07:00-14:00 adding 1.5 to 2 hours to journey times. Given such road sections therefore have a history of flooding, the modelled flooding on the A road causing lengthy detours is therefore not unreasonable. Those areas which appear most susceptible to flooding in the scenarios modelled are similar to the ‘flood hotspots’ highlighted in Coles et al., (2017), namely the Western part of the city around Acomb and the station, providing confidence to the findings. Based on these findings, the workplace and school destinations which are highlighted as being most affected by flooding could therefore be prioritised by local stakeholders for further investigation of the ingress routes to them.

The semi-dynamic method that has been developed could also be applied to other flood scenarios, times of day, week or year to expand the knowledge base for York. The results have shown that flood events do affect travel to destinations, and overall thousands of hours were lost at work or school. This could be taken forward to estimate the monetary cost to businesses and commuters, as undertaken by Wei et al., (2018) and Kilgarriff et al., (2019), to further understand the cost implications to the local economy of disruption to road networks from flooding. Whilst the data are modelled rather than representative of a real flood event, so cannot be taken prescriptively, the principles could be applied to decision making in York. If better real-time flood data or population travel data were available, these could be incorporated into the workflow.

7.4.3 Application to risk theory

The risk theory presented in Chapter 2 states human exposure to flood events varies throughout the day as people undertake their daily routines (Debionne et al., 2016). The results of this chapter's research support this assertion, as it has been shown that the time of flood onset affects the time lost at work or school and which destinations are affected the most. This work is a proof of concept and has its assumptions and limitations, but it does make a case for considering the population in travel related to the time of flood onset in risk studies. It has been shown to have a slightly greater effect than hazard magnitude, so should be included in flood risk analysis. As there is a lack of short-term forecasting models for pluvial flash flood events (Li et al., 2018), using the framework set out in this chapter could allow scenarios to be tested to identify at risk locations. This would enable the introduction of suitable preparation steps to raise public awareness and reduce disruption. A vulnerable group, primary school children, have been considered as they may require more assistance or rescue during a flood event (Arrighi et al., 2019). However, in these results school children are less affected than commuters, probably due to the short journeys from home to school, so may not require much assistance in a flood event. In a larger area with rural catchments this would likely not be the case as longer journeys increase exposure. Being able to travel around urban areas is not just important for journeys to work and school, but also required for emergency vehicles to reach communities fast (Arrighi et al., 2019). An extension of this research would be to include travel to and from hospitals and the access of emergency services to key sites such as schools. Further research steps are discussed in Section 8.5.

7.4.4 Limitations and assumptions

There are several assumptions and limitations to the work, which will also be discussed within the context of the whole thesis in Section 8.4.3. Beyond using the specific time of day historic traffic data, travel behaviour and congestion during the flood events are not included in the model, although both would have an effect on travel times. This limitation is present in other similar studies (for example Arrighi et al., (2019), Versini et al., (2010)). This is due to a lack of historic traffic data during pluvial flood events in York, and more widely in England. It is also assumed that travellers take the least cost path between locations by using Dijkstra's algorithm, another common assumption in origin-destination studies (Li et al., 2018). Therefore, the travel times presented here are likely to be an underestimation of the disruption caused by a flood event. However, the relative changes between scenarios do provide a useful measure of scale of effect of the changes in flood magnitude and flood onset time. The measure for disruption used in this chapter is time lost at work, which does not directly link to economic activity, which does limit the applicability of the findings to businesses. Quantifying the disruption through a generalised cost function, as Ford et al., (2015a), Davidson & Davidson (n.d.) and Kilgarriff (2019) have done, would increase the applicability to businesses and would be a logical step to advance this framework. Finally, the trips measured in this chapter were single purpose journeys when this is not the case for many travellers. The primary school children would not be travelling to school by car on their own, and their parents may head onwards to work after dropping the children at school. So, the focus on single purpose trips is a limitation. A model which could account for multi-purpose journeys, for example an ABM, could overcome this issue.

7.4.5 Unique contributions

There are several novel contributions presented in this chapter:

- It is an example of one the very few assessments of the importance of hazard magnitude versus timing of onset as determinants of disruption to travellers
- A transferable and scalable methodology has been developed for the assessment of destination susceptibility to travel disruption, rather than flood risk at the site.
- One of very few studies to disaggregate hazard impact on travellers by population sub-group, rather than for population as a whole
- Whilst the method of comparing travel times under hazard scenarios to a business-as-usual scenario is similar to most other studies looking at this issue, for example Mossoux et al., (2019), Wei et al., (2018), Pfurtscheller and Genovese (2019)), the inclusion of historic traffic data and time-specific origin points in the network analysis extends the modelling framework of these earlier studies.

7.5 Conclusion

To conclude, this chapter has presented the analysis from a proof-of-concept methodology, allowing temporally specific analysis of travel disruption to flood events. It has been shown that within the scenarios studied, the time of day of flood onset is as, if not more, important than the magnitude of the flood event. The differences in travel disruption to workplaces and schools has also been explored, and destinations which were most affected by time lost at them when journeys were delayed or prevented spatially identified. Chapter 8 will discuss the findings from this chapter and Chapters 5 and 6 to draw together the main contributions of this thesis.

Chapter 8 Discussion

8.1 Overview

The aim of this thesis was to investigate how daily variation in the population in travel and pluvial flood magnitude interact to cause spatial and temporal variation in disruption to journey times and time spent at destinations across an urban area. Chapter 5 developed a framework for combining spatiotemporal gridded population data with GIS network analysis, with the case study of primary school children in York. The developed framework was successful in creating spatiotemporal origin-destination flow data for primary school children in York and improving the representation of the population in travel. Chapter 6 tested this framework by assessing the differences between using census origin-destination data and spatiotemporal population data to measure the impact of a pluvial flood scenario on commuter travel in York. It was found that the spatially weighted Monte Carlo process produced significantly different origin-destination pairs to those recorded in the census. Finally, Chapter 7 took elements from Chapters 5 and 6 in order to assess if the time of pluvial flood onset is more important than flood magnitude for disrupting commuter and school travel in York. The results indicated that time of flood onset is more important than flood magnitude in sudden onset pluvial flood events. There has been little previous research of the relative importance of changing population exposure versus changes in flood magnitude on travel disruption from flooding. The literature presented in Chapter 2 highlights the importance of time of hazard onset, however there are few empirical studies looking at this. Therefore, this thesis has contributed to risk analysis knowledge.

This chapter will discuss the key contributions of this research within the specific context of the case study and their application to other geographical areas, risk theory and flood policy (Sections 8.2 and 8.3). There will also be a reflection on the limitations and uncertainties of this analysis (Section 8.4), recommendations for further research (Section 8.5) and recommendations for policy makers (Section 8.6).

8.2 Key Contributions

The key contributions of this work have been identified from the analysis conducted in Chapters 5,6 and 7. Arguably, the most significant contributions are those included in Section 8.2.2, which pertain to the flood travel disruption analysis in Chapter 7. This section will state each contribution and present the evidence for them in relation to other studies and wider risk literature.

8.2.1 Methodological contributions

Reconciling administrative and census data from different sources and years can cause data issues when parameterising the SurfaceBuilder247 model.

In Chapter 5, the parameterisation of the SurfaceBuilder247 input data required several iterations, as after the first model run it was discovered there were not enough children in the origin data set to meet the demand at the school destinations (Figure 5.6). Such inconsistencies between input origin and destination data might not be apparent at national scale. However, for the York case study, initially most schools did not have enough pupils and only through close examination of the output grid were schools which were 'missing' due to there being no population for them to select identified. This had not previously been highlighted in Martin et al.'s (2015) paper introducing the modelling framework. One likely cause for this discrepancy is the mismatch between the census data (the origin data source) and the school administrative data (the destination data source). A refinement in school destination data set choice and adjusting the percentage splits in the quantity of each population sub-group which the destinations required, was enough to rectify the discrepancy between origin and destination data sets.

The development of a spatially weighted Monte Carlo simulation tool for modelling origin-destination flows through time.

This thesis' main contribution to spatiotemporal population modelling has been the extension of the framework proposed by Martin et al., (2015), and utilised by Smith et al., (2015) and Renner et al., (2018), to incorporate origin-destination flows. In these previous papers the population in travel were weighted across the study area through the use of a background layer, which denotes where roads are and their relative busyness for a time of day (Section 4.2.2). Martin et al., (2015) state that their framework makes no attempt to directly estimate movements between locations, however this is critical for the purposes of this thesis, assessing the impact of flooding on travel to places. Batista e Silva et al., (2018) state that multi-temporal population grids from the ENACT project, similar to the ones created by the SurfaceBuilder247 model, are useful for assessing exposure to natural hazards and planning and modelling transport. This assertion was therefore tested in Chapters 5 and 6 through the development of a tool for simulating origin-destination flows, given an attending population count for a destination, a catchment polygon for the destination and a gridded population surface.

The aim of Chapter 5 was to develop a method for combining gridded spatiotemporal population data, created using the framework developed by Martin et al., (2015), with GIS network analysis. This required converting the gridded data to a point data set in order to represent spatiotemporal flow data, a challenging task to keep the spatiotemporal dynamic of the population data and to convert a raster grid to specific vector points of origins. This was achieved in Chapter 5 for school children through the creation of a spatially weighted Monte Carlo analysis code, the application specification for which was designed by the author, but the script written by Nieves (2018 pers. comm.), and then applied by the author to the problem scenario. As far as the author is aware, the conversion to spatiotemporal origin-destination flow data in this manner is a novel contribution to the spatiotemporal representation of the population in travel. This conversion was successful as primary school destinations have a well-defined catchment area from which a probability surface could be constructed from. When comparing the representation of the population 'in travel' in Figure 5.8, the output from SurfaceBuilder247 shows a different distribution as the in travel population is weighted across the whole study network in York, for example along the whole outer A road. In contrast, the result of the spatiotemporal flow data and network analysis shows more travel on the smaller local roads and a smaller proportion on the A roads. This is a better reflection of where primary school children in travel are, as most of the distances primary school pupils are travelling are short journeys inside York. The average distance for primary-aged children to school is 1.6 miles (Easton & Ferrari, 2015), so they are unlikely to be distributed as widely as SurfaceBuilder24/7 shows. Therefore, for the purposes of explicitly studying the effect of floods on those in travel, this conversion of the data and subsequent network analysis is beneficial. The success of this conversion is largely due to primary school children being a 'data rich' population group (Harland & Stillwell, 2010), meaning spatiotemporal modelling is easier.

This tool was extended from simple to multiple-band circular catchments when applying the spatially weighted Monte Carlo process to commuters (Chapter 6). The catchment size for workplace zone destinations were based on travel to work data from the census and, unlike schools, this was a series of bands containing weights based on the percentage of commuters who travelled from each distance band. This led to a weighted surface with more probability gradations for the spatially weighted Monte Carlo procedure than was created for school children. This selection method resulted in more commuters chosen from cells closer to the destination than was reflected in the census origin-destination data, with the two data sets proving statistically different (Section 6.3.1). With the current configuration for creating weighted surfaces, the approach of combining census origin-destination pairs with travel survey data provides a more realistic spatial distribution of origins for a destination (as applied in Chapter 7). Other studies have taken this approach. Crols and Malleson (2019) combined census data with travel survey data to define the commuting behaviour of the agents in their model, whilst Martin et al., (2018) used the data for the number of persons travelling from each output area of residence to their primary employment WPZ to calculate aggregate flows.

Although there are developments which could be made to the approach presented in Chapters 5 and 6, the tool for simulating origin-destination flows when only catchment information is known would be useful to analysis beyond the specific research problem of this thesis. Whilst detailed origin-destination data is available for some groups of the population, often it cannot be released to preserve confidentiality. For example, the home origin locations of those using healthcare facilities, like the Demographic and Health Surveys which release distance bands for travel to healthcare facilities, but not origin-destination data (USAID, 2020). Therefore, this tool could be useful for these instances where network analysis is required but the origin data is not available.

8.2.2 Flood travel disruption results

The total amount of road flooded in an urban area does not need to be large for substantial disruption to occur.

Through the analysis conducted in Chapters 6 and 7, it was found that the spatial placement of floodwater is important and even small areas of flooding can have a substantial effect. In York, only a very small percentage of road area was flooded with disruptive flooding i.e. at least 30cm deep and 125m^2 in area (0.9% max for 100 year and 0.6% for 1 in 30 year in Figure 7.7). However, disruption was seen across the city. This has been found in other studies of the effect on flooding on road travel. For example, Sohn (2006) found that disruption to a few road links could lead to whole counties in Maryland being cut off from the rest of the state in the 1 in 100 year flood scenario they tested. It is also widely understood that the risk is not restricted to the geographical area directly impacted by the flood hazard as the effects are felt further afield across the road network (Kasmalkar et al., n.d.).

As roads are built to run at near maximum capacity to minimise costs, this means there is little redundancy in the system to cope with disruptions (Mossoux et al., 2019). Therefore, a small amount of road which flooded can have a much larger effect than disrupting travel just to that section. In Chapter 6, the results show the example of no access into the one-way system in the historic city centre due to flooding located at the junction of Goodramgate and the A1036, leading to a cluster of workplace destinations where no commuters could reach them. In a different study, Rowley et al., (2016) identified that in Worcestershire there are various pinch points, mostly bridges crossing the River Severn, which act as key links to cross-county movement. There are only two bridges across the River Severn in Worcester, one of which is prone to flooding. The nearest other bridge at Upton upon Severn is also often flooded, meaning travel across the county can be impacted from just a small proportion of road being flooded. At a more localised scale, Ellis & Viavatenne (2014) studied a residential area of Coventry to identify ‘critical drainage hotspots’. They found one key pinch point was a flooded area which obstructed a major exit/entrance from a residential site. This is a similar phenomenon to the observed flooding in York, and fits with the nature of pluvial flooding being smaller areas of water than fluvial flooding. These examples in other studies and the results of the analysis in this thesis show that there is a need for tools to enable risk managers to identify sites where small amounts of flooding can cause larger issues. The framework provided in Chapter 7 could be used for this purpose and to allow for retrofitting measures to reduce flooding (Houston et al., 2011). The pinch points in a road network system may be the ‘critical links’ which enable travel to destinations for groups of the population, and the idea of ‘critical links’ is discussed further later in this section.

Destination susceptibility to disruption from flooding can be assessed based on workforce or pupil travel, not just workplace or school site characteristics.

Chapters 6 and 7 provide evidence that despite changes in the flood magnitude and time of flood onset, the same destinations were consistently among those experiencing the most disruption. These destinations can therefore be considered to be the most susceptible to travel disruption from pluvial flooding. In Chapter 6, the network travel time for the two origin datasets to their destinations were compared during a flood scenario. It was found that the five destinations with the highest travel times were the same between the two datasets (Figure 6.8). This is evidence that those destinations most affected by disruption to travel will become apparent even with differences in the origin data sets.

Examining the results spatially, there did appear to be clusters of destinations for schools and workplaces where travel to them was most disrupted, for example in the West of city around York station (Figure 7.22 and Figure 7.25). Investigating the reasons behind the observation that there was no travel from origins to some workplace destinations, it became apparent most of the destinations (6/8) with no travel to them during a 1 in 30 year flood were due to flooding at the destination (Table 7.2). For the 1 in 100 year event, fewer destinations had flooding where travel to them was blocked (7/21) meaning there were more cases where the roads outside the origin points are flooded, or somewhere on the route between is flooded, which prevents travel. This provides some evidence that there are some destinations which could be considered particularly susceptible to disruption from flooding and will be apparent regardless of change in flood magnitude or onset. Similar data could be generated for other sites and destinations to provide the same information to stakeholders for flood preparation measures.

Many studies (for example Li et al., 2018; Arrighi et al., 2019; Pregnolato et al., 2016a; Freiria et al., 2015) have modelled travel disruption from flooding; however, none of these have looked at how such disruption varies across different destinations which the same groups of people attend e.g. workplaces and schools. The work presented in Chapters 6 and 7 is therefore believed to be the first to do this. Commercial flood risk assessments required for planning developments assess the access directly onto the site (Emanuelsson et al., 2014; Ambiental Environmental Assessment, 2020), and national flood maps do not show comparisons of how the effects of flooding vary between places. Therefore, the ideas and framework presented in Chapter 7 present a novel way for viewing flood risk, from the perspective of the destination's susceptibility to disruption from flooding on the road network.

To further explore the observation that the same destinations appeared most susceptible to disruption, a Spearman's rank test was conducted to see if those destinations with greatest increases in travel times to them were consistent across the flood magnitude and onset time scenarios tested. It was found that there was statistical agreement in which destinations were most affected by the flood scenario between flood magnitudes when comparing the results for the equivalent time of flood onset (Table 7.3 and Table 7.4). This is evidence that the same destinations are most affected, i.e. they are most susceptible, and are identified even when changing the size of the flood magnitude and time of flood onset. The exception were the results for FT3 for both flood magnitudes as there was not statistical agreement between these data and the rest of the scenarios, demonstrating that different destinations were more affected. This was seen for both commuters and primary school children, but there were stronger correlations for commuters. These results provide evidence that those destinations most susceptible to the effects of pluvial flooding on travellers will likely be identified whichever flood magnitude is chosen, however changes in the time of flood in relation to commuting patterns can have a bearing on these results. In the context of the risk equation highlighted in Chapter 2 (Equation 2.5) space and time are important for risk. The spatial location is important as flood-prone areas consistently show flooding across these scenarios, leading to the same destinations experiencing difficulty in access to them. Time is also shown to be important here as destinations are flagged as most vulnerable to disruption do change, likely due to later onset of rainfall leading to some destinations not being impacted as much as more of their travellers had already reached them.

Floods disrupt the travel of York's commuters more than travel by its primary school children.

The risk to travellers in flooding situations is an area where an increasing amount of research has been conducted (Liu et al., 2014a), as a high number of fatalities are drivers in their cars and road closures lead to socioeconomic disruption. A novel aspect of this thesis is the comparison between population groups on the impact of flooding on road travel, made possible through the use of spatiotemporal population data. Other work, including Debionne et al. (2016), Pregnolato et al., (2017a), Nyberg and Johansson (2013), Jenelius and Mattson (2012), Rodríguez-Núñez et al., (2014), Sohn (2006), Taylor et al., (2006); Chen et al., (2015) and Balijepalli & Oppong, (2014) and Coles et al., (2017) examine the disruption to the population as a whole or one particular group (for example the elderly). However, there are few examples of studies, that the author is aware of, which model the differences in flood-related travel disruption across more than one population group. One similar study is Kunwar et al., (2014) who applied spatiotemporal population data to an activity-based agent model to investigate total evacuation of ten UK cities and could model the effects on some groups of the population (men, women, workers, students etc.). However, they have not looked at the routes taken by each group or developed the comparison between groups in the same way the analysis in this thesis has.

The evidence from Chapter 7 is that in the scenarios undertaken, commuters were more affected than primary school children in their journeys to their destination. Firstly, there were no school destinations across any of the scenarios tested that became completely unreachable for the origins linked to them. For workplace destinations, as discussed in the contribution above, there were destinations that became unreachable for all commuters across both flood magnitudes and the three times of flood onset (Table 7.2). This provides initial evidence that access to schools was maintained whilst it was not fully maintained to all workplace destinations. In terms of the journey times of those successful origin-destination journeys, Figure 7.15 shows that commuters travelling to destinations have a larger increase in journey times than school children (Figure 7.16) from the same flood event. This is likely due to the shorter journeys that children take to school meaning a slight detour does not add much time, but it also shows that commuters may be coming across floods which cause larger diversions on their routes. Debionne et al., (2016) found in their study that commuters with longer journeys had the greatest risk of being exposed to floodwater, supporting the findings here that the population group with the longest journeys will have the most risk and disruption. The location of some workplaces may also be in areas where there is more likely to be flooding blocking the road leading up to the workplace destination point.

Risk theory states that children are more vulnerable than adults (Civil Contingencies Secretariat, 2008), however, in the case of the vulnerability of those in travel to their place of work or school, primary school children are less vulnerable partly due to their shorter exposure time on the road network and the location of flooding in the scenarios tested. Therefore, this indicates that more focus could be placed on mitigating the effects for businesses and commuters to future pluvial flood events. Business value of time savings can be a significant part of the benefits of transport schemes (Department for Transport, n.d.), so the fact they are more affected than schools could mean that reducing time lost at work through flood prevention measures could have an economic benefit to local businesses. These results cannot be projected to other specific population groups without knowing their journeys around York. However, the principle that those with longer journeys are more likely to encounter pluvial flooding in a storm event is sound.

Critical links can be defined locally by examining which closures lead to the most disruption to the destinations of specific population groups.

Most studies on travel disruption consider transport networks from an engineering point of view, focussing on network connectivity, rather than measuring the impact of road closures on the population which use them (Pfurtscheller & Genovese, 2019). This is evidenced in the definitions and methods for identifying critical road links in a road network. As discussed in Section 3.3.3, other research papers which investigate critical links look at the overall connectivity of the network to determine which links are critical, for example Jenelius (2010a); Jenelius & Mattsson (2012); Rodríguez-Núñez et al., (2014); Sohn (2006); Taylor & D'Este (2003); Taylor et al., (2006); Chen et al., (2007), Chen et al.,(2015); Balijepalli & Oppong, (2014). These studies do not consider the needs of specific groups of travellers, which a stakeholder like a local authority may be interested in, when creating their distinction of a critical road. Assessing critical links from the perspective of the groups who use the roads enhances the exposure and vulnerability components of a risk assessment. The exposure, i.e. number of people using a road link, can be assessed per group of the population, and those groups which are considered vulnerable can have specific roads identified as being critical for them.

'Basic' methods like eliminating a single link at a time to determine the effect on the system and which are most 'critical' (Sohn, 2006; Chandra & Quadrifoglio, 2013; Duan & Lu, 2014), do not consider the geography of the disruption which is key to understanding the effect on travellers (Pregnolato et al., 2016b). Identifying critical links is stated in various research papers as being important for prioritising resources for maintenance (Sohn, 2006; Jaroszweski et al., 2015; Freiria et al., 2015). If the purpose of the analysis is to know which links are critical in order to be targeted for maintenance to keep social and economic functions of a city running, then the analysis should instead be based on the needs of different population groups and how they use the network. Using the framework established in Chapters 5,6 and 7, the effect of flooding on the travel of specific population sub-groups can be calculated and therefore the roads critical to them identified. This is a key contribution of this thesis.

The case for combining the theory of critical road links into flood risk assessments is found in UK government legislation. The Civil Contingencies Act states it is the local government's responsibility to understand the nature of their critical infrastructure and if its failure is likely to cause a local emergency (Cabinet Office, 2011). The methods created in this thesis of integrating specific population sub-groups with travel analysis and hazards could facilitate this identification of critical roads and make it more relevant to the groups who are most vulnerable or economically impacted. By integrating population group data into the disruption analysis, this moves research beyond seeing networks just as physical infrastructure (Demšar et al., 2008) and aligning the concept to the consequences of network failure on people (Muriel-Villegas et al., 2016). This makes these methods useful to local governments who are responsible for pluvial flood management. Whilst most network studies look at travel from the perspective of the whole population, there are other research examples of the effects of road disruption on particular groups of the population. For example, Nyberg and Johansson (2013) and Coles et al., (2017) both look at accessibility to services for elderly residents during hazard events. However, they did not expand their work to identify specific roads which were vital to keep open to give access, i.e. those which could be classed as the critical roads. They also only look at one population group and do not compare which links are critical for different groups within the same study, a unique aspect of this thesis. So, whilst the idea of investigating the effects of natural hazard induced disruption on the travel, or accessibility, of specific population groups is not new, taking the step further to defining critical links in this way and identifying them is novel.

In the specific context of York, the results generated indicate which roads may be critical for primary school children and commuters. The roads which primary school children use to travel to their destinations are mostly smaller local roads (see Figure 5.9 and Figure 5.10), which would probably not be considered critical in 'network scan' approaches (Freiria et al., 2015). However, for this population group they are critical for enabling children's education and those which are most utilised could therefore be targeted for flood mitigation measures. For commuters, the flooding on the A1237 York outer ring road is a clear example of a place where substantial detours can be incurred through its closure. There have been real events where pluvial flooding has occurred on this section of A road leading to its closure for several hours (City of York Council, 2012), demonstrating the reasonableness of the modelled scenario. The road next to York station (Leeman Road) and the roads surrounding the historic city centre (A59, Water End, A1036) are also roads used frequently by commuters, and the flooding near the station is a source of disruption to the surrounding workplaces. These roads are highlighted in Figure 7.19 as potential 'critical roads' for these population groups in York. These were identified based on visual examination of data, but analytical techniques to quantify which are critical could be applied as a further step.

The time of flood onset is more important than flood magnitude for rapid onset pluvial flood events.

A key part of the theoretical underpinning of this thesis was that the time of day when a hazard occurs is very important to the level of risk the surrounding population experiences (see Chapter 2). Whilst this was a common statement in risk theory literature, there are fewer quantitative studies measuring the effect of time of onset on the outcomes for the population. One example is Kasmalkar et al., (n.d.), who used a generic 1-hour window within peak morning commuting hours, to test the impact of flooding on travel. However, this was not expanded to include more time periods like the analysis in this study. Another example is Li et al., (2018), who calculated the delay to journeys from flood events in peak travel time, again without examining the effect of changing the time of flood onset around the peak travel period. They did test the effect of four different magnitudes of flood, capturing the physical hazard element of risk, but not the time of flood onset which has a greater bearing on the socioeconomic aspects of vulnerability and exposure.

One of the aims of Chapter 7 was to assess if pluvial flood onset was more important than flood magnitude in the scenarios tested. The results in Chapter 7 provide evidence of the importance of time of onset. By changing the time of flood onset by two hours it is enough to see differences in the travel time lost at destinations as a result of flood disruption. For example, for the 1 in 30 year flood event there is a decrease of 36,639 lost minutes at work if the time of flood onset changes from 6am to 8am. This large difference is due to a substantial level of flooding not being reached until the majority of commuters have already reached work. Increasing the flood magnitude to 1 in 100 year and keeping the flood onset time the same, therefore just increasing the amount of flood water on the roads, the difference is 9,440 additional minutes at work lost. This is less than the difference between the flood onset being delayed by 2 hours. In Chapter 2, risk was defined as the dynamic function of hazard, vulnerability and exposure (Mechler & Bouwer, 2014) and Chapter 7's results have empirically demonstrated that population movement is a dynamic local factor (Dawson et al., 2011) which is having a larger effect than changes in the physical hazard.

Further evidence for this key finding can be found by examining the numbers of workplace destinations which were unreachable by their assigned origins in the different scenarios (Table 7.2). The scenarios with the most destinations where no origin could reach them (i.e. greatest level of disruption to people) were the two with the earliest rainfall start time (6am), with seven destinations for the 1 in 30 year flood and 16 destinations for 1 in 100 year event. This is likely due to fewer travellers earlier in the day, so there are less origins from which people are travelling from, increasing the chances of all the origins being blocked by flooding. Furthermore, when comparing the statistical agreement between which destinations experienced the greatest time lost at them, the results for FT3 were the only scenarios where there was no statistical agreement. This is evidence that the time of day the event occurred at had a larger effect than flood magnitude, as it interacts with travel patterns. Hence, the time of day is important in sudden onset events like pluvial flooding as it relates to the demand on the road network from travellers, it is not the magnitude of the flood itself which is driving the effect.

The results in Chapter 7 have shown that as the variables change in both time and space, the risk (the effect on travel time) has fluctuated. In risk literature, it is highlighted that the interaction between where people are and the time of hazard onset can substantially change the exposure and vulnerability elements of the risk equation (Freire et al., 2013; Aubrecht et al., 2014; Terti et al., 2015). The influence of time on the risk to those travelling on the road network has been investigated by other researchers. For example, Ruin et al., (2008) investigated a flash flood event in Southern France. Due to the flood beginning on a Sunday evening in September, outside of the normal working week and the tourist season, they state the consequences were less than they could have been at a busier time on the roads. This supports the case that the time of day of sudden onset hazards like pluvial floods is very important. However, they did not empirically test how changing the time of onset would have changed the effect on the population in travel. Spitalar et al., (2014) conducted an analysis to find common factors in flash floods which increased risk to the population, by examining a database of all flash floods which occurred in the US between 2006-2012. They found that short duration events and those which happened in the evening resulted in more fatalities, therefore were riskier to the population. This was due to short duration events often having little pre-warning to enable people to prepare and the darkness in the evenings reducing visibility of flood water. Relating these findings to this study, this supports the notion that for pluvial flood events, which are generally short events with little lead time, the time of day when they happen is important.

The findings specific to York could be applied to other areas which have broadly similar patterns in peak travel times across urban areas, as the time of day is likely to have the same effect in similar settings. The temporally specific analysis conducted in this thesis to assess risk to the population in travel and destination disruption could be incorporated into future modelling for surface water management plans, as it has been shown that the risk will not be static over the course of a day. As far as the author is aware there is not another study which has looked at the effect of time of flood onset and flood magnitude in this way, so this is a novel contribution of the work

8.2.3 Contributions to risk theory

Extending the risk conceptual framework beyond static residential populations to cover travellers and workplaces should be incorporated into transport-based flood risk assessments.

The statement of the research problem (Section 1.2) outlines the focus of this thesis as the effect of flooding on travellers and the resultant disruption to time spent at their destinations. The analysis in this thesis has therefore focussed on the risk posed to travellers, which sets it apart from most traditional flood risk analysis work which is based on the residential population (Balica et al., 2013). For example, Koks et al., (2015) assessed the risk to residential households in Rotterdam; Kubal et al., (2009) also conducted a flood risk assessment for the residential population in Leipzig and Arnell & Gosling (2016) modelled the likely increases in flood risk globally under different climate change scenarios, again for residents of floodplains. UK government requirements for flood risk assessments also view risk from a static population viewpoint. Flood risk reports generally quote the number of houses and businesses affected, but they do not specifically explore the knock-on effect of travel disruption and time lost at work or school. The flood risk maps for England and Wales, created by the EA, are applied to calculate exposure of static populations, either for residential or business sites in FRAs as required for developments. These FRAs look at access and exit from a site during a flood but are not required to consider travel into the wider area (DEFRA & EA., 2019; Ambiental Environmental Assessment, 2020). This therefore fails to capture the dynamic nature of the population between residential and business sites.

The creation of spatiotemporal flow data, maps showing the routes taken by those travelling to destinations and identification of which destinations are most susceptible to disruption in this thesis, provide an example of how temporal population dynamics can be considered in risk analysis. Incorporating the ‘four-step transport model’ into the research approach worked well as it was designed to view transport management from the ‘person perspective’ (McNally, 2007; Ortuzar & Willumsen, 2011; Jones, 2012), therefore spatiotemporal population data fits well within this perspective. Similar work was done by Smith et al., (2015), where spatiotemporal population data enabled risk to the population in residential, workplace and retail areas from flooding to be mapped. However, this thesis is novel as it specifically is focussing on travellers and the disruption posed to them from a flood.

Framing natural hazards research around travellers and the effect on work/school destinations fits within a wider body of work which has focussed on developing methods for incorporating non-resident populations into analyses. Workplace zones, as proposed by Martin et al, (2013), were introduced to allow small-area analysis of populations at work (Berry et al., 2016). These were applied in this thesis as the destination points for the network analysis for commuters. Most recently, Cockings et al., (2020) have developed a geodemographic classification structure based on the characteristics of workers and workplaces, rather than by residential population. This classification has provided new insights into the characters of workers and workplaces at the small area level, and is useful as it can be applied to health, transport and commerce applications. Outside of natural hazard risk, using measures of the ambient population who travel through an area, but do not reside there at night, has been highlighted as a key component of assessing retail patterns (Berry et al., 2016) and assessing crime hotspots (Malleson & Andresen, 2015, 2016). The continuing improvements to the representation of non-residential populations in population data will mean creating risk maps and metrics of transient populations, like those using the road network, will become easier. The framework outlined in Chapter 7 for assessing spatiotemporal risk to travellers and destinations could be expanded and improved on as these data sources become more available.

Including spatiotemporal population data and examining a fine temporal scale of flood onset enhances the temporal aspects of transport-based flood risk analysis.

The main conceptual underpinning of this analysis work was there are spatial and temporal components to all elements of the risk equation. Time is complementary to space, as we move through space over time, so both should be considered in analyses as the challenges and characteristics need to be addressed together (Bogaert, 1996; Pred, 1977). In the case of the applied problem studied in this thesis, it is people moving across a city along the road network and this fluctuates over time as demand changes. Time dimensions are ‘excessively important’ in transport systems and the fluctuations in demand are at a daily to seasonal scale (Reggiani et al., 2015). Therefore, the risks associated with disruption from flooding are time dependent with a high temporal resolution to risk (Cats & Jenelius, 2014; Fuchs et al., 2013; Renner et al., 2018).

As set out in Figure 8.1, the spatiotemporal questions posed in Figure 2.2 can be answered by the results generated in this research. The locations in space and time where the risk of disruption to travellers is greatest have been identified through the results presented in Chapter 7. Including spatiotemporal population data and the fine temporal scale of flood onset has enabled this temporal risk analysis. The framework developed in this thesis could be applied to other locations and hazards.

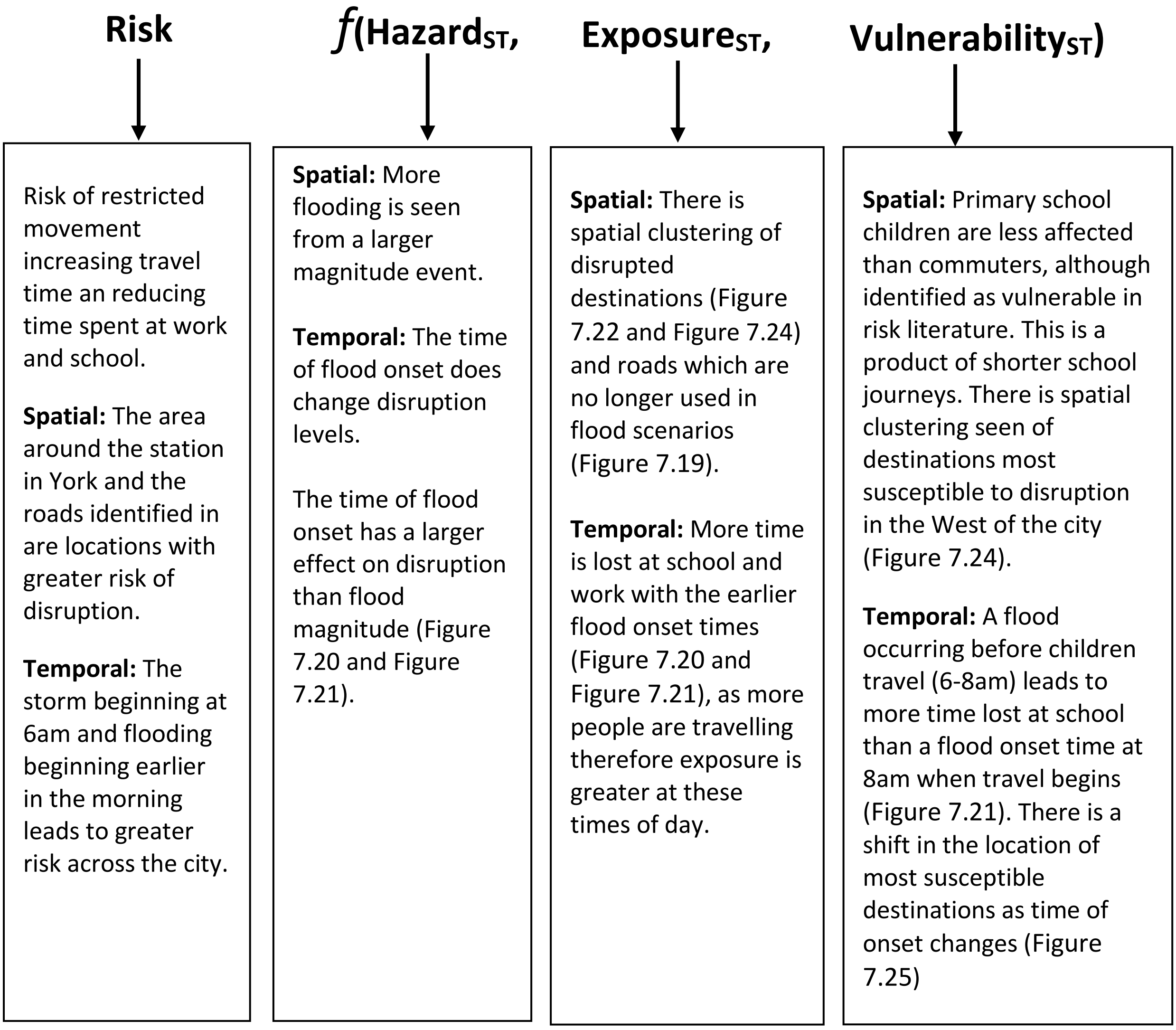


Figure 8.1 The conceptual risk equation and answers to questions set in Chapter 2, categorised by the specific spatial and temporal components of each element.

However, the inclusion of time in an analysis workflow can be difficult, particularly for the social elements of risk. A key part of incorporating time in this thesis has been the inclusion of spatiotemporal population data. Alternative data sets of travel surveys, traffic flow, footfall data and call detail records are available and can have a greater spatial and temporal resolution than census data (Herrera et al., 2010; Deville et al., 2014). However, the lack of demographic data attached to them (Herrera et al., 2010) means specific analysis of population sub-groups is not possible. The usefulness of including spatiotemporal population data in hazard analysis has been demonstrated by other researchers. Smith et al., (2015) applied SurfaceBuilder247 in their study of exposure of different population groups in Cornwall, the inclusion of demographic data allowed the relative vulnerability of people to be considered. Renner et al., (2018) applied spatiotemporal population data created with SurfaceBuilder247 to Bolzano in Italy to also estimate flood exposure. This thesis has expanded on this previous work by considering the specific risks to those in travel, but all these studies show that spatial and temporal aspects of all parts of the risk equation can be included in flood risk analysis with spatiotemporal population data.

There are many examples of research papers where the impact of flooding on travellers is considered, and some include a temporal aspect to their work. For example, Freire et al., (2013) looked at how day and night time population distributions affected hazard exposure and the time taken to evacuate after a tsunami warning in Lisbon. They found night time evacuation was modelled as being quicker, potentially due to less people being in the city centre than during the daytime. However, they did not investigate the effect of smaller scale temporal changes to population which are seen through daily travel patterns or look at the differences in disruption to different population sub-groups. Debionne et al. (2016) and Shabou et al. (2017) in their papers measured the exposure drivers experienced in flood scenarios and ascertained there was a risk amplification effect for drivers, although again they did not test the effect of altering the time of day of the flood event, or investigate if there was differential risk between groups in the population. Li et al, (2018) looked at effect of pluvial flooding on road travellers and tested different magnitudes; however, again they did not incorporate time of day the flood occurs or consider specific groups who are affected. Chang et al., (2010), Pregnolato et al., (2017a, 2017b) and Arrighi et al., (2019) all present methods for measuring disruption to commuters from pluvial rainfall events at the scale of a city. However, their frameworks do not consider time as a component of risk as explicitly as the one in this thesis does and there was not a comparison on how the flood event effects different groups of the population differently.

So, whilst the concept of investigating the risk to travellers from flooding is not new, there are novel elements of the framework designed in this thesis. Firstly, including spatiotemporal population data which allows for the relative risk of specific groups of travellers to be assessed. Secondly, the temporal timescale analysed, the impact of shifting the onset of a flood event by an hour or two, is a finer scale than other studies have investigated but it is important as the dynamics of travel are at this scale. This could open up opportunities for more targeted local risk analysis covering people in travel to assess where mitigation action should take place. UK government policy documents have frequently identified the risk to travellers, and the people who become isolated due to flooding, is an important area to address to build flood resilience (Cabinet Office, 2011, 2016; Pitt, 2008). However, this has not led to the integration of risk from the perspective of travellers into flood risk assessments. Further development of this framework, which can be applied at the LLFA level, could facilitate the inclusion of the risk to travellers in flood risk assessments.

8.3 Transferability and Scalability of Approach

This section describes how the methodology developed in this thesis could be transferred and scaled up to other geographic areas, population groups and hazards. The methodological framework, as shown in Chapter 7, is transferable to other local authorities in the UK, both urban and rural. It could also be conducted at a larger scale than in this research if the computational power is available. The spatiotemporal population model (SurfaceBuilder247) and population data used in this thesis are available for England and Wales, as the majority of data are from census sources and administrative data sets which are available under open government license. The next census is due to take place in 2021, therefore in a few years new data will be available. It is also possible to take the SurfaceBuilder247 model and apply it to other countries, as demonstrated by Renner et al., (2018). With the equivalent data sources, this part of the methodology could be transferred to case studies in other countries. Travel to work data were applied in the SurfaceBuilder247 model and the spatially weighted Monte Carlo analysis to parameterise the selection of origins for destinations. Similar census data is available for countries like Italy (Renner et al., 2018), France (Shabou et al., 2017), Australia (Mees et al., 2008) and many others, increasing the transferability of these methods.

The principles of the analysis, the effect of time of flood onset on disruption, can be assessed from many sources of traffic data and routing tools. The network analysis technique can be conducted at a larger scale and the network data and historic traffic data from ESRI is available for the entire UK and many other countries. Therefore, transferring this part of the methodology to another location in the UK, and many other counties, would be straightforward. If this analysis was conducted over a larger area with more origin-destination journeys, then creating a network data set with time-specific traffic speeds (for example with OS MasterMap Highways Network with speed data) could be more cost-effective than accessing the ESRI HERE data via the server. Alternative network routing tools could be used in place of ESRI's tools if traffic data is available in other countries. As outlined in Section 3.3.4.2, open source tools like OpenTransportMap and OpenRouteService, and in Section 3.3.4.3 commercial tools like Google Routes, exist which could be used within the framework to provide routing and traffic data. Kilgariff et al., (2019) used measured road traffic data for a 17-day period of disruption in Galway and the surrounding area due to flooding, which is different to this thesis and most papers which rely on modelled data. Real event data for the traffic disruption is extremely beneficial for calculating the effects of disruption, however a lot of cities and towns do not have sufficient data for their method to be transferred elsewhere. Also, the testing of the temporal effect of moving the time of flood onset as this thesis has shown is not possible.

Finally, for the flood model component of the methodology it is possible to model flood events for larger areas, and the model used in this thesis has been applied in the UK and abroad with success. As with the previous two computational models, the computing power available may be a restriction on the size of area which could be modelled. There are numerous other flood models available, like LISFLOOD-FP which is often used in UK academic work (for example Sosa et al., 2020; Neal et al., 2011; Rajib et al., 2020; Smith, 2015), so the creation of the flood hazard data for another area is definitely achievable. Although a flood model which capture the dynamics of urban drainage may be most beneficial as pluvial flood events which affect traffic tend to occur in urban areas.

A national data set which detailed destinations, like workplaces and schools, which were at risk of isolation or reduced travel to them during a flood event, would be a useful addition to current flood risk maps. A main challenge in scaling up this methodology to a region or nation would be choosing a flood scenario to incorporate. It is unrealistic to test a scenario where the whole country is flooded at the same time, as pluvial flood hazards tend to be caused by localised heavy rainfall. Although, if the goal was to make a dataset identifying areas at risk of being cut off, this assumption might be needed. Scaling up to the whole country would require experience of national flood modelling in order to determine if there is an appropriate disruption scenarios. A second challenge would be determining which origin-destination pairs should be modelled to gain a national understanding of the impact of disruption. The creation of a national time-specific data set of origins and their associated destinations could be complicated and time consuming, and for groups other than commuters even more difficult. However, it could be possible with enough computing power to iterate through every workplace zone and model the routes from their origins for an area of interest. This could be provided as a data download similar to other products from the Population 247 project (GeoData Institute, 2016).

The methodology presented in this thesis could be transferred to other population groups, natural hazards and types of transport network. Expanding the methodology to the inclusion of other population sub-groups in the flood risk analysis is possible but relies on the availability of sufficient travel data. The main challenge in transferring to other population sub-groups is the availability of mobility data. Smith et al., (2015) demonstrated that there is sufficient data to include tourists and retail populations for flood exposure calculations. However, having the origin-destination data to model journeys in a network analysis is a greater task. Renner et al., (2018) state a limitation of their work is assuming that tourists are static and stay in their study area, as the absence of travel data means their origin-destination journeys cannot be determined. With new administrative and big data available in the next few years, it may become possible to include more groups into this framework.

The flood hazard was included in the network analysis model as a barrier, therefore any hazard which causes a blockage to a road with a known spatial area could be included in the analysis. There are many hazards besides flooding which pose a risk to roads, for example landslides (Postance et al., 2017), Tsunamis (Freire et al., 2013), fallen trees (Nyberg & Johansson, 2013) and earthquakes (Khademi et al., 2015). All of these hazards could easily be incorporated into the framework as shapefiles or points where obstructions have occurred. The potential for including primary hazards (e.g. floods) and secondary hazards (fallen trees) in one analysis is also there.

As stated earlier, the network analysis algorithm can be applied to any road network and the historical traffic data were available for the whole of the UK and many other countries worldwide. The focus here was on road travel by car, however road networks and travel by private vehicle are not the only forms of travel. There have been some studies of transport disruption which have looked at public transport as well (Cats & Jenelius, 2015, 2016; Tympakianaki et al., 2018). So, with the appropriate public transport model, the transference of the key ideas of time being critical in the hazard onset and including population data, may be possible for other forms of transportation.

Whilst census data has been the standard for population research (Aubrecht et al., 2014), the 2021 census may be the last 'traditional' census, as an approach which links administrative data sets to generate the same data are proposed to replace it (Tinsley, 2017). Whilst the spatiotemporal population model used in this thesis relies on census data, alternative data sets and new data sources could be used within the overall risk methodology instead, as they become available. Any data which provides the origin and destination of the journeys of a group of people, and the time of day these are undertaken, could be used as part of the risk analysis methodology in this thesis in place of the spatiotemporal population model. New forms of 'big data' (call detail records, smartphone apps, social media, travel cards, footfall data) currently exist and are continuously being improved. There are many studies which have used this type of data for modelling the ambient population and travel around an area, for example Heppenstall et al., (2016); Crols & Malleson (2019); Malleson et al., (2018); Batista e Silva et al., (2018). At present, these data sets do not contain demographic information which is required for modelling specific groups of the population, so cannot be a substitute for spatiotemporal population modelling.

8.4 Uncertainties and Limitations

Throughout this thesis, the uncertainties and limitations of the methods applied have been highlighted and discussed. In this section, the key points and their likely effect on the work are identified. The first section covers validation of the data created in the analysis, the second section highlights the sensitivity tests conducted and the final section the key assumptions and their effect on the results of the analysis.

8.4.1 Validation

Validation of the research findings, so as to quantify uncertainty, is challenging as the only relevant data sets currently available are those census data and associated administrative data sets which are used in this thesis. Therefore, there is not an independent data set for comparison with the modelled spatiotemporal population data, spatiotemporal flow data or the routes and travel times measured in the scenarios tested.

The validation of the outputs from the SurfaceBuilder247 model are challenging “because we are attempting to estimate time-specific population distributions not directly captured by any other measurement systems” (Martin et al., 2015, p.768). The best available census, and other administrative data, were used in the creation of the spatiotemporal population data, as there is not a data set of the phenomenon of interest already in existence. One way of validating the data would be through direct population counts of areas or buildings at a specific time of day and comparing it to the results from the SurfaceBuilder247 model. However, this would involve a substantial cost and resources (Martin et al., 2015) which are beyond the scope of this thesis.

As already discussed in this chapter, and in Section 3.2.3, alternative data like call detail records, footfall sensors, smart card travel data and automatic number plate recognition exist (Malleson et al., 2018; Heppenstall et al., 2016; Crols & Malleson, 2019) These can provide data on how many people are travelling through an area, but do not have the demographic data on who is travelling (Herrera et al., 2010), so are not necessarily a direct validation data set. However, alternative data can be used to evaluate the output from SurfaceBuilder247. Malleson and Andresen (2016) use data from SurfaceBuilder247 as one representation of ambient population in their study of crime hotspots in London. In their comparisons, they found the data from SurfaceBuilder247 were reasonably strongly correlated to workday (from census) and mobile telephone data, so they were capturing similar populations. This demonstrates one approach which could be used to evaluate the spatiotemporal output from SurfaceBuilder247. However, again this was beyond the scope of this thesis to do a similar evaluation as the main aim of the thesis was on the risk analysis problem. Despite the limitations of validating the data, the approach taken by using SurfaceBuilder247 to create spatiotemporal population data is still sound. The benefits of using this model framework is that it is a volume-preserving approach, meaning all people in the study area are accounted for between time steps, and it is based on known origin and destination flows and travel data from the census (Martin et al., 2015; Smith et al., 2015). Therefore, it can be assumed that the outputs of the model are reliable.

The travel time and routing results presented in this thesis are a product of the spatiotemporal origin-destination flows and the network analysis. There is currently no available data set to directly compare travel time results to, as this would require the observation of pluvial flood events in York and the recording of traffic data for the whole city, beyond the scope of this thesis. Pregnolato et al., (2017b) did use traffic count data to attempt validation of the route and travel time data in their model. However, this only describes volume of traffic and not the key information of who is travelling and making the journey, which was important to this research. Arrighi et al., (2019) suggest validation could be achieved through applying the judgement of experts and sensitivity tests can be used to examine the influence of parameters on the model. They also suggest Google Maps and its function of giving live and historic travel times for routes could be used for validating baseline route modelling. These techniques could be useful in some cases; however, this relies on a pluvial flood event occurring during the time period the research project is active and then modelling the same pluvial flood event for the network analysis model. Qualitative validation based on observations of flood events might strengthen the quantitative findings of risk analysis. For example, qualitative comparisons were made in Section 8.2.2 of modelled flood results in this thesis to previous flood risk assessments done by the council. More widely, the names of places and roads can be used as an indication of their propensity to flood. ‘Water lane’ in York, for example, was a road which experienced a high level of flooding in the modelled flood results and is an access road to a primary school.

8.4.2 Sensitivity analysis

Whilst validation of the results presented in Chapters 5, 6 and 7 is difficult, sensitivity tests can be used to understand how the parameters of the models and methods may be affecting the results. Sensitivity testing was done for two aspects of the overall thesis framework; the SurfaceBuilder247 model in Chapter 5 and the parameters of the network analysis tool in Chapter 6.

In Section 5.3.1, it was found there was unfilled demand at schools, with not enough children in the origin data to fill the school destinations (11% deficit). The input parameter of how many children came from each sub-group for each school destination was adjusted to test the sensitivity of the outcomes. Initially the model parameters were for only children aged 5-10 to attend primary schools, but this is not the case as primary schools have 4 year olds and 11 year olds, and some of the schools were junior schools which are for the youngest children. Adjusting to 13|80|7 split of 0-4|5-10|11-17 year olds based on school count data for York (Department for Education, 2016), rectified this issue. The aim of Chapter 6 was to assess the differences between census origin-destination data and spatiotemporal population data, so this was a sensitivity test for the new data created to see how well it reflected the recorded travel to work data. This was possible for commuters as there were alternative data.

Finally, in Chapter 6 the sensitivity of the Network Analysis tool was tested in relation to the hierarchy property. It was found that this parameter did have an effect when comparing the non-flood to flood scenarios, as when turned on the hierarchy changes how the algorithm assigns routes leading to some instances of distance travelled decreasing in the flood scenario whilst time increases.

8.4.3 Key assumptions

Table 8.1 lists the key assumptions applied during the analysis work and the likely impact on the findings. Overall, the assumptions applied would have led to an underestimation and simplification of the disruption caused by flooding on road travel.

Table 8.1 Key assumptions and their anticipated impact on model outputs

Assumption	Likely impact on findings
There is equal probability of a person travelling from any grid cell in a distance band to the specific destination.	Large underestimation of catchment size of commuters in Chapter 6, due to how the catchment weightings were calculated. A small underestimation of school children catchment size as the catchments covered 95% of children.
Historic traffic data were used and impact of disruption on traffic not modelled.	Large underestimation of traffic levels and therefore travel times, in flood scenarios. Pregnolato et al., (2017b) found travel times may increase by more than 50% when congestion is included.
Catchment areas for workplace and school destinations from administrative data exclude 5% of population making longest journeys (trip distribution).	Moderate underestimation of travel time and disruption
All individuals in a population sub-group act in the same way, so individual behaviour is not included. Bian and Wilmot (2015) call it a key assumption in transport studies.	Moderate underestimation of variation in travel time and disruption. Drivers do not all behave in a rational way like an algorithm does, and perception of risk and purpose of journey can

	have an effect on decisions make whilst travelling (Terti et al., 2015).
Assumed census data correct and adjusted pupil age breakdown to fit school demand profile.	Moderate impact of unknown direction.
Time profiles are generalised for the country and do not allow for multipurpose journeys.	Moderate underestimation of the travel times as multipurpose journeys would be longer, therefore increasing likelihood of exposure.
All drivers have perfect knowledge of the road network and diversion routes.	Moderate underestimation of travel times.
Shortest time path is taken to reach a destination (using Dijkstra's algorithm within network analyst tools), as in similar studies (Postance et al., 2017).	Moderate underestimation of travel time
Each person is travelling in a car on their own, or in Chapter 5 children were walking on their own or by car on their own.	Minor underestimation of the number of people on the roads, especially if parents travelling with their child.
Drainage failures are not modelled.	Small overestimation of travel time increase. The flood model does not include urban drainage systems explicitly, so there may be more flooding than in reality.
That there is a passable/impassable switch in flooding when there is more variation.	Small underestimation in travel time.
Rainfall timing and intensity are constant across the city of York.	Small overestimation in the travel time increases due to disruption.
Origin-destination data for commuters is from the 2011 census.	Direction of effect unknown. It is not known how the home and work locations of the residents of York have changed between 2011 and 2016 (the reference year).
Assumption of disruptive floods being at least 30cm in depth and 125m^2 in area.	Direction of effect unknown. Some vehicles can travel safely through this depth (Pregnolato et al., 2017b), but for the majority of cars this is a sound assumption.

Following on from Table 8.1, there are some further limitations to consider. Firstly, it is only possible to model a small number of flood scenarios out of a great many possible scenarios (Smith et al., 2014a). Previous work on pluvial flooding in York by the city council (2019) found that each real event experienced included flooding in different locations across the city. Therefore, the places found in these modelling scenarios to be prone to flood disruption may not necessarily be seen in a real flood event, and the spatial patterns of disruption seen in this thesis would not be replicated exactly. Choosing the scenarios is an important consideration when designing a study and selecting event which will show a range of flooding, like the 1 in 30 year and 1 in 100 year scenarios in this thesis, is desirable.

Linked to this limitation, an assumption was made on what ‘disruptive flooding’ was. Only two flood magnitudes were tested in this work and more magnitudes between these could have provided some additional information. Pluvial flooding is not just a factor of topography as drainage failures often have a large part to play, and these were not included in the flood model used. The thresholds used for identifying the flood area which would be ‘disruptive’ (30cm depth and over 125m^2 in size) is based on UK environment agency guidelines (Li et al., 2018; Arrighi et al., 2019; Pregnolato et al., 2017b). However, any adjustment to criteria would lead to potentially different results for example if drainage was included. Although the comparison of the effect of the two magnitudes in this study does indicate the effect of magnitude might be small in respect to the amount of disruptive flooding recorded. Ideally, a series of flood barriers where some blocked all travel, for flooded roads above the threshold, and some allowed travel but at a slower rate, for roads with less deep or wide flooding, would be included in the network analysis. The depth categories for this could be based on Pregnolato et al.’s (2017b) depth-disruption function.

Finally, the concept of critical roads has been considered in this thesis and some visually identified for York. However, critical roads are dynamic, not static, and can change as a flood event progresses, as those roads which are identified as critical pre-flood event may not be those which are critical during the flood event. Jenelius (2010a) proposed the idea of secondary critical links, i.e. those which become important when your disruption scenario is introduced, and other roads are no longer accessible. Including this extension of the critical links idea, and with the inclusion of analysis for specific sub-groups from this research, could further how network studies be conducted.

8.5 Future Research Directions

Based on the achievements of this research and its limitations, there are several recommendations for future research.

Options for extending the work presented in this thesis, using the existing framework to model other time periods.

In this thesis, it was only possible to study the morning commute and two population sub-groups (Chapter 7). Repeating the analysis for different day times and types of day is an essential next step in order to build a fuller picture of how the temporal dynamics of population in travel and flood events interact, in order to affirm the finding that time of day has more influence than flood magnitude. It is possible that the interaction of pluvial flood onset time and magnitude may be different for other day types and times of day. For example, the effect may be less at a weekend when the majority of travel is not confined to a few hours in the morning and evening.

Developing alternative approaches for spatiotemporal flow data and calculating disruption from flooding to road travel.

Census data were used in a spatiotemporal population model and adapted to create spatiotemporal flow data of likely origin-destination journeys. There are alternative data sets and methods which could be used for this purpose, as described in Section 8.3. For future research directions, the application of agent-based models for the purpose of determining the disruption to travellers from flooding could be explored. Dawson et al., (2011) did use an ABM for analysing evacuation decisions in Towyn in Wales, but since this most studies have used network analysis methods for analysing disruption at a city level. Since Dawson et al.'s (2011) paper, there has been an exponential increase in the amount of 'big data' sources which have the potential for innovating urban modelling (Heppenstall et al., 2016). Determining which data sets are useful for the specific issue of flood disruption of road travel is a key priority, and these would likely be data sets which provide the origin-destination journeys of the population (perhaps via smartphone apps like Malleson et al., (2018)) but ideally with specific demographic data attached. For a successful ABM there would also need to be data on the underlying behaviours which drive travel decisions both in non-flood and flood conditions. This could be through travel surveys or incorporating social science observations of people's behaviour during natural hazards. For example, Terti et al., (2015) established in their research that travellers would persevere with a homeward journey for longer than an outward journey to work. If an ABM with the right underlying data could be created, this could be a powerful way of capturing spatial and temporal dynamics of individual travellers and modelling congestion as a result of flooding, addressing a common limitation in current studies.

The approach developed in this thesis was semi-dynamic with time steps at which the flood, population and travel elements were synchronously moved. Near-real time models in population modelling, flood modelling and travel either exist or are in development and linking these up could be an interesting exercise. If achieved, the monitoring of events in near-real time to enable response as the event happens could improve flood hazard response. There are examples of other researchers using real traffic and flood data within their travel disruption studies. Coles et al., (2017) recreated real flood events in their flood model to test accessibility of emergency services to care homes. However, they did not have traffic data to see what the effect of the flood was on vehicles at the time, instead relying on network analysis. Kilgarriff et al., (2019) used real traffic disruption data and flood data for a 17-day period of disruption in Galway, Ireland in order to calculate the financial loss due to increased commuting time. Pregnolato et al., (2017b) were also able to use traffic data from a real pluvial flood event and compare it to a non-flood average to assess the impact on traffic count, and with increasing use of Automatic Traffic Counters, CCTV and smart sensors in cities this will become an option for more places. Whilst smart sensors for traffic and pedestrian counts are becoming available, this does not provide information on who is travelling, just the numbers of people or vehicles. Near real time population modelling is still being developed, for example The Population 24/7 Near Real Time project (GeoData Institute, 2018), and integrating this near real time approach with a dynamic flood model could be extremely useful for risk planning.

Include more population sub-groups into the analysis to understand where the focus of mitigation actions should lie.

Only two population sub-groups were included in this thesis, primary school children and commuters. The spatiotemporal methods developed in this thesis could be applied to population sub-groups other than primary school children and commuters providing appropriate temporal data are available. SurfaceBuilder247 has been used to create tourist data (Smith et al., 2015) and applied for analysing the exposure of populations to flooding. If there was sufficient travel data for tourists around an area, then the methods developed in this thesis could be applied to generate spatiotemporal flow data and then analysis of the routes they take. However, the difficulty would lie in determining where people go to and from as tourists have more complicated journeys than school children or commuters. Modelling the risk of people travelling to and from retail sites is potentially more likely than simulating tourists. There are many studies on the travel to and from retail centres, how they are modelled and who the customers are (Newing et al., 2013). These data could therefore be applied to flood risk, as Smith et al., (2015) demonstrated. Other key workers could include those who work on critical infrastructure (for example, communication, civil nuclear, defence, energy and water sites (Centre for the Protection of National Infrastructure, 2020)), school teachers, or those involved in the food supply chain. The author is not aware of any published work which takes this focus. Alternatively, the focus could be on vulnerable groups, the elderly or tourists, if it is a tourist area. The groups which should be prioritised would be determined by the organisation who is conducting the flood analysis and the purpose of the flood risk study.

National and local study sites should be investigated.

The Brown review (Brown, 2014) states that the identification of critical single points of failure in the country's transport system is a key task for flood management going forward. This applies both to local authorities concerned with a single urban area and national governments. At a national level, motorways and main A roads which are prone to flooding could be targeted. In the 2007 floods, the closure and entrapment of 10,000 vehicles on the M5 in Gloucestershire highlighted the large scale disruption caused by the flooding of major roads (Pitt, 2008). This was a singular high impact event, but these types of events could be crucial to national risk assessments. Examining regional flow would also be useful for identifying the potential repercussions on the movement of food and other goods. At a local authority level, the analysis should be done for the whole urban area of interest to identify critical roads and then further work could focus on these areas specifically.

Secondary hazards involved in flood events should be included in future analysis.

The framework developed in this thesis could be applied to other hazards which impact on road travel. This will vary between places which hazards are most important to be modelled, and existing hazard risk assessments can be used as a guide. In England, it is flooding which has the greatest impact on the road network therefore, this should be the primary natural hazard focus (Brown, 2014). Whilst flooding is a primary hazard, there are often multiple secondary hazards which occur as a result of the storm which has caused the flooding. Jaroszweski et al., (2015) highlight that blockages in storms come from flooding and from other sources like fallen trees and collapsed roads. A future extension of this risk analysis framework should aim to include primary and secondary hazards which cause road blockages, to gain a more rounded picture of the disruption caused.

Creating a national database which identifies areas at risk of being isolated in floods or cut off.

Finally, this framework could be utilised to create a data product of places identified as at risk of being cut off in flood events. By running multiple flood scenarios, the results could be used to identify places where travel to and from the wider community is not possible. This is important information to know as current flood risk zones only identify places at risk of direct flood water contact, not areas which are not flooded but are isolated during the event.

8.6 Recommendations for Policy

8.6.1 Flood management in York

Several recommendations can be made for the city of York based on the results of this thesis' analysis. Firstly, pinch points in the road network system should be identified and targeted drainage measures retrofitted to reduce the risk of road closure. Figure 7.19 indicates road which were identified as potentially being critical to the travel of commuters and school children, and more extensive quantitative work could be done to investigate the pluvial flood risk and disruption caused by closures of these roads.

Secondly, if possible, the analysis conducted in this thesis should be extended to focus on retail journeys and tourists, a particularly important industry for York. These are both examples of transient groups which contribute to the local economy, therefore, keeping these groups moving around the city and the places they visit accessible during sudden onset rainstorms would be beneficial economically.

Finally, the city council and flood risk partners should develop a strategy to increase the awareness of pluvial flood resilience within the local population and internal highways design teams. It has been shown that pluvial flooding does disrupt travel and time spent at businesses and should be given consideration alongside fluvial flooding, which is a much more well-known risk in York.

8.6.2 Wider Policy for travel and risk management

Based on the research presented in this thesis, several recommendations can be made for policy and practice. Firstly, local authorities, the level of government responsible for pluvial flood events, could integrate a more temporally dynamic approach to assessing the effect on travel, as demonstrated in the framework created in this thesis. Current surface water (pluvial) management plans states the number of homes at risk of surface water flooding, highlighting the focus on houses not the wider implications of travel disruption (Brown, 2014). This is despite the evidence pointing to more danger posed to people who drive into flood water and the large economic effects of travel disruption. This thesis therefore makes a case that this aspect can be modelled within a flood risk assessment and should be in the future. Spatiotemporal risk assessments for disruption to road networks are worth doing as roads are the lifelines of our society (Chen et al., 2007; Platt 1995; Nyberg & Johansson 2013) and there are a number of hazards which can disrupt them. The explicit function of the wider road network is missing from current flood risk assessments, despite making road networks resilient to extreme pluvial flood events being a key theme of the UK's Surface Water Management Plan (DEFRA, 2018). Key considerations when creating spatiotemporal flood risk assessments of road network effects would be the flood scenarios tested, the temporal properties of the origin-destination flow data and the specific population groups of interest. For sudden onset events, it is vital to model the peak of the hazard with the peak of the road travel. The framework should be used to simulate enough time periods to cover the part of the day of interest, and at least two flood magnitudes would be required. It is not just flood hazards which this framework could be applied to, other hazards like landslides, tree falls, and bridge collapses could also be investigated.

Secondly, population data could be integrated into hazard risk assessments. This research has shown that time of onset is important because of the demands on the road network and the variable location of the in-travel population. The interest in monitoring flood events and other hazards is due to the impact on humans, therefore spatiotemporal population data should be included with flood risk assessments. Scenario planning is already a well-used tool in risk and hazard mitigation, so incorporating temporally specific population data and journeys is achievable. The combination of near-real time population, traffic and flood data are achievable and would create an almost dynamic picture of risk. Temporally specific population in travel data could also be integrated into risk assessments for other hazard assessments, for example landslides and disease vector mapping.

Looking ahead, it seems likely that pluvial flash flood events will become more common, with increasing intensity in the UK (Boeing et al., 2019) without action to reduce the effect of climate change. Therefore, the country will have to adapt and prepare for pluvial flood events reducing travel by targeting resources to pinch points in the road system where flooding causes the largest disruption. With improvements in the data collected in all aspects of life, better datasets and approaches for modelling the spatiotemporal dynamics of the population in travel will be created. The framework developed here can be adapted to accommodate new data as the core issue will remain for our urban areas.

Chapter 9 Conclusions

This thesis has explored the problem of pluvial floods disrupting travel in an urban area and the subsequent effect on travel times and time spent at workplaces and schools. A crucial element of this research has been understanding the interactions of spatial and temporal properties of the hazard, exposure and vulnerability elements which make up the risk. This research has also required an interdisciplinary approach combining theory and methods from population geography, transport planning, and natural hazards science. This interdisciplinary approach has led to novel methodologies being developed to enable spatiotemporal analysis of the effect of flooding on the population in travel.

This chapter summarises the findings of the thesis and their applicability beyond the case study example chosen and is structured as follows. The first section presents the main findings, the second highlights the research contribution before the transferability of this research is summarised. Finally, the third section recommend further research which could be explored.

9.1 Assessment of Thesis Aim and Main Contributions

The overall goal of the thesis was to determine how daily variation in the urban population in travel interacts with the time of pluvial flood onset and magnitude to influence spatial and temporal disruptions to journey times and destinations. This goal has been achieved through the work set out in the three results chapters (5,6 and 7). It has been shown that the time of pluvial flood onset does have a substantial effect on disruption to travel time and time spent at destinations, as there are fluctuations in the number of people in travel.

From the analysis conducted in this thesis, the following contributions have been presented:

- Reconciling administrative and census data from different sources and years can cause issues parameterising the SurfaceBuilder247 model.
- The development of a spatially weighted Monte Carlo simulation tool for modelling origin-destination flows through time.
- The total amount of road flooded in an urban area does not need to be large for substantial disruption to occur.
- Destination susceptibility to disruption from flooding can be assessed based on workforce or pupil travel, not just workplace or school site characteristics.
- Floods disrupt travel of York’s commuters more than travel by its primary school children, due to length of journey time and number of commuters.

- Critical links can be defined locally by examining which closures lead to the most disruption to the destinations of specific population groups.
- The time of flood onset is more important than flood magnitude for rapid onset pluvial flood events.
- Extending the risk conceptual framework beyond static populations to cover travellers should be incorporated into transport-based flood risk assessments
- Including spatiotemporal population data and examining a fine temporal scale of flood onset enhances the temporal aspects of transport-based flood risk analysis

This research has provided knowledge contributions to the three disciplines from which the theory and methods underpinning the analysis were derived. It has also provided an example framework for tackling an applied research problem; travel disruption due to the presence of hazards on a road network.

9.2 Transferability of methods and findings

The methodology developed in this thesis could be transferred and scaled up to other geographic areas, population groups and hazards. The methodological framework, as shown in Chapter 7, is transferable to other local authorities in the UK, both urban and rural, and could be conducted at a larger scale than in this research if the computational power is available. The effect of other natural hazards on road travel could also be examined using the framework outlined, as events like landslides, earthquakes and fallen trees could be incorporated into the network analysis as barriers. Spatiotemporal flow data, as presented in chapters 5 and 6, could be applied to research areas beyond natural hazards. For example, for measuring access to health care services. At present there is a lack of spatiotemporal data for journeys of various different groups of the population, like tourists and the elderly, which restricts the applicability of these methods to the entire population. However, this is likely to improve over the next decade as more data is collected on movement around cities.

9.3 Recommendations for Further Research

Three key recommendations have been identified for future research projects examining the effects of flooding on travel and workplaces. Firstly, there are options for extending the work presented in this thesis by using the existing framework to model other time periods. Useful time periods to explore include the normal evening commute, Saturday and Sunday traffic and school holidays. By evaluating more time periods, a broader understanding of how change in the time of flood onset affects travel time would be achieved.

Second, developing alternative approaches for spatiotemporal flow data and calculating disruption from flooding to road travel. In this thesis, census data were used in a spatiotemporal population model and adapted to create spatiotemporal flow data of likely origin-destination journeys. There are alternative data sets and methods which could be used for this purpose, such as agent-based models, call detail records and smartphone GPS data and open source road network and traffic tools. For future research directions, incorporating agent-based models into the semi-dynamic analysis framework outlined in Chapter 7 for the purpose of determining the disruption to travellers from flooding could be explored.

Finally, future research should aim to include more population sub-groups into the analysis in order to understand where the focus of mitigation actions should lie. Due to time and spatiotemporal data constraints, only commuters and primary school children were able to be included in the analysis conducted for this thesis. With more time to develop the spatiotemporal flow data, other groups of the population could be modelled and the effect of flooding on their travel assessed.

Appendix A Population Data Sources

This appendix describes the data which were used in the SurfaceBuilder247 model for creating spatiotemporal gridded population data for the study site of the City of York. As a 20km buffer area around the study area is required, data for areas beyond the York local authority were included. The area which the study area and the buffer zone encapsulate is referred to as the analysis area in this appendix.

All data sets

Data for the following district authorities in Yorkshire and the Humber were used as inputs into the SurfaceBuilder247 model: East Riding of Yorkshire, Hambleton, Harrogate, Leeds, Ryedale, Selby, Wakefield and York. These are the authorities which fall within the analysis area. Microsoft access was used to extract the output area/workplace zones in the analysis area from the larger census data sets of these 8 district authorities.

Output areas

Population weighted centroids were used to calculate which output areas are within the analysis area. Output area is used as the basis for analysis as this is the finest spatial scale which the required census data is available. The source of these data was the 2011 census.

Residential Population

There are two aspects of the residential population data, the term time and non-term time population, as large numbers of students move around the country between their term and non-term time addresses. Therefore, two population data sets are needed to reflect this internal population shift. York itself is a prominent university city, hence it is particularly important to consider the movements of the student population. The term time population will be dealt with first.

The number of people in each age bracket is a main requirement for running the SurfaceBuilder247 model. The data table QS103EW, age by individual year, from www.nomisweb.co.uk, was chosen to enable the number of people in each population sub-group could be calculated. The population subgroups chosen were: 0-4 years, 5-10 years, 11-17 years, 18-64 years (university students), 18-64 years (non-students), 65-79 years, over 80 year. These population groups equate to the main school age groups, and now 16 and 17 year olds must be in full time education or training, the minimum age of the working population has been increased to age 18. There are two groups for the retired population, as the 'oldest old' are identified as one of the most vulnerable groups in natural hazard events. As it is gathered as single year of age, these data are easy to consolidate into groups of the users choosing. These groups represent mutually exclusive groups whose daily movements can be assumed to follow similar patterns.

These data are for 2011, however they can be extrapolated to the target year of 2016 through using mid-year estimates (MYE). MYE are calculated each year, however this is done for lower super output area level, one spatial level higher than the 2011 single year of age data. Therefore, it is not directly transferable and requires reweighting. The total population for the LSOA MYE was taken and proportionally split amongst the output areas within in it, assuming that the relative proportion of the population of each output area within an LSOA remained unchanged. This gave a new total population for each output area. This then required dividing between each population sub-group, and it was assumed the age structure of the population remained the same between 2011 and 2015. This produced a data set of term time population, divided into age categories, for the year 2016.

Data for out of term time population is provided for output area in five-year age categories, rather than single year of age like term time population. This gives an added complication as the five-year age categories do not match those chosen for analysis, or the English education system. The out of term time data was divided into the analysis age categories by assuming there were equal numbers of each age within a category and reassigning the population appropriately. This provides a reasonable estimate of the number of people in each of the analysis age categories for 2011.

MYE are not available for out of term time populations. In order to produce a reasonable estimate of the population, the percentage change for each output area between term time and out of term time total population was calculated for the 2011 data. This percentage difference was then applied to the term time totals to estimate what the out of term time population would be in 2016.

Immobile population

This is the percentage of people in each output area who can be considered to be immobile due to the communal establishment they live in, for example prisons or long-term care facilities. QS420EW (communal establishment management and type) and DC4210EWla (communal establishment management and type by sex and age) accessed via the Nomis website were used to calculate the percentage of each of the population age categories who were immobile.

The age breakdown in each type of communal establishment is available as an average for a unitary authority. So, the data for York unitary authority was used and it was assumed that this could be applied to each occurrence of an establishment type.

A threshold of 10 people was used as the minimum number of people in a communal establishment who would be considered. This is because any smaller population will blend into the general population and be difficult to distinguish. The types of communal establishments considered were Prisons, detention centres, care homes and secure mental health units.

These data are for 2011 and were not scaled up using MYE as there is not data available for communal establishments outside of the census year. It was assumed that the percentages of immobile population would still be applicable, in the same way the proportional split of the population into age categories remained the same for the general population.

Workplace zone data

Workplace zones were created for the 2011 census to provide a census geography for the working destinations of the population. Workplace zones are used to define workplace destinations within this model. The age breakdown of the workers in each workplace zone were as well as the number of people were required, and this was taken from the sex by single year of age (workplace population) census data. The majority were in the 18-64 category, with a small percentage 64-79. The number of 16/17 year olds for each workplace zone were removed from the total, as those in this age group are now included in the education data set as the minimum age at which full time employment (without training) is permitted is now 18. This gave a total number of people in each industry cluster for each workplace zone.

The following data sets were utilised:

WP605EW – Industry (workplace population)

https://www.nomisweb.co.uk/query/construct/summary.asp?reset=yes&mode=construct&dataset=1314&version=0&anal=1&initsel=

Data capture year: 2011

WP702EW - Sex by single year of age (workplace population)

https://www.nomisweb.co.uk/query/construct/summary.asp?reset=yes&mode=construct&dataset=1320&version=0&anal=1&initsel=

Data capture year: 2011

Appendix A

Time profiles for each industry type were required to specify temporal commuting patterns in the model. Each industry type has different work patterns so require a different time profile. The time profiles used were ones created by Smith et al., (2015) in their work. There are 6 clusters of industries in total, covering all the industries listed in the census, and are categorised by those workplaces with similar working patterns e.g. a hospitability cluster, a retail cluster, an office-based cluster etc.

Travel to work data

Census table WP702EW – Distance travelled to work (workplace population) was used to calculate how far people travel into York's workplace zones. This data was captured in the 2011 census. This is for all usual residents aged 16-74 who work, and it is measured by Euclidean distance. This data is used in the model as a unique wide area dispersion radius for each workplace zone. The majority of workers were in the 18-64 age group. Those workers who were 16 or 17 were discounted as 16- and 17-year olds will no longer be in full time employment.

The distance travelled to work, through the dataset WU02UK – Location of Usual Residence and Place of Work by Age, was also used to understand the proportion of people who lived and worked in York in order to base decisions on the size of buffer area.

Care Homes

The communal establishment data set includes care homes (see immobile population section), and this was used to give an initial estimate of the population in the analysis area residing in them. The data year is 2011. The types of care homes included were those under local authority control with and without nursing, and private care homes with and without nursing. This data was supplemented with York city council data (https://www.york.gov.uk/info/20135/residential_care/1462/residential_care) which lists the five council run care homes in York and the data listed on care homes.co.uk which lists the locations and maximum number of residents a facility can have (https://www.carehome.co.uk/care_search_results.cfm/searchtown/York). Both of these data were captured in 2017.

Hospitals and Health care facilities

For inclusion in the spatiotemporal population model, the number of patients, age breakdown of patients and the location of the hospitals is needed. The number of patients needs to be split into inpatients, outpatients and A&E admissions, as these all have different time profiles, and broken down into an expected daily number. NHS digital has data for hospital episode statistics, which provide the number of patients admitted and the admission category. These data also contain broad age categories. These data are at the resolution of healthcare provider, which can be an NHS hospitals trust, or for some private health care facilities a single hospital. The data for the 2015-16 hospital episode statistics was used. For NHS trusts with multiple hospitals, the number of admitted patients were divided proportionally by the size of the hospital. Hospital size was considered to relate to the number of beds it contains for inpatient care. These data were found from hospital websites and CQC reports. In the analysis area, there is only one large hospital (The York Hospital) so this receives the majority of inpatients. It is also the only A&E facility, so all emergency admissions would be at the York Hospital. To supplement this data, the NHS choices website to search for hospitals, urgent care, minor injury centres and rehabilitation hospitals to get the number of beds at a facility and the services provided. The data capture year was 2017.

Schools

For schools in the buffer area, data from edubase was used. This includes the establishment's details, number of pupils, location etc. This provided destination data for school age children. The age range for each school is given, and the majority fit within the population subgroups chosen for this analysis. However, some Junior schools, infant schools and nursery have an age range which overlaps the 0-4 and 5-11 categories. In these cases, the number of pupil's was assumed to be the same for each age group, and thus divided into the analysis age categories appropriately.

For schools in York, data from York City Council was used to give a more accurate representation of the number of pupils at a school. This data was from the following sources and was for the year 2017:

York City Council

https://www.york.gov.uk/directory/3/schools_directory/category/4

https://www.york.gov.uk/directory/3/schools_directory/category/5

https://www.york.gov.uk/directory/3/schools_directory/category/6

Information provided: Name, Address, Postcode, Age Range, Admissions Policy, Opening Hours, Pupil Count

Data capture year: May 2017

Independent schools were taken from the independent schools council as these are not listed in York council data. This data was also for 2017.

Appendix A

Independent School's Council

https://www.isc.co.uk/schools/england/yorkshire-area-north/york/

Information Provided: Name, Address, Postcode, Age Range, Number of Pupils (borders and day)

Appendix B Conference papers

Short paper abstract submitted to the GISRUK 2019 conference

Integrating spatiotemporal dynamics for modelling disruption to road travel in flood events

Kate Rawlings[11], Jim Wright[21], Alan Smith[32], Sally Brown[43] and Jeremiah Nieves[ǂ1]

[1]School of Geography and Environment, University of Southampton

[2]Geography, Earth and Environmental Sciences, University of Plymouth

[3]School of Engineering, University of Southampton

3 January 2019

KEYWORDS: spatiotemporal population, flood risk, network analysis

Abstract

Flooding can significantly disrupt road networks and the time of day or week in which a flood event happens has a large effect on how disruptive it is. For example, a flood event during the morning rush hour will likely affect more people than one which occurs on a Sunday morning as demand on the road network is greater. Previous GIS-based research has introduced disruptive hazard events to road network models, but have not considered the temporal nature of how risk arises from the interaction of population movement with flooding. This research examines how standard GIS methods compare to an enhanced method for assessing the effect of a storm event on the travel of two population groups, primary school children and commuters, in York, UK.

Figure 1 outlines the methods compared in this work. The standard method (1a) uses school catchment areas and origin-destination travel to work data to determine the output areas associated with a destination (school or workplace zone). These origin-destination pairs are used in ArcGIS

[1] K.Rawlings@soton.ac.uk
[2] j.a.wright@soton.ac.uk
[3] alan.smith@plymouth.ac.uk
[4] sb20@soton.ac.uk
ǂ j.j.nieves@soton.ac.uk

network analyst to generate routes as a measure of travel function. The enhanced method (1b) utilises spatiotemporal population data generated using the Population 24/7 model (Martin et al., 2015), which uses census data and other administrative information to create gridded estimates of the number of people in an area at a given point in time. For example, in this work gridded population estimates are produced for a Tuesday morning in term time and out of term time. Origins are assigned a destination using an R code that statistically calculates the grid cells from which people most likely travel to a destination. Finally, routes are generated using ArcGIS network analyst. Both methods will compare routes generated in a non-flood baseline scenario to the same flood scenario, generated using the Flowroute-i^{TM} flood model from Ambiental Risk Analytics, in order to quantify the disruption to road travel from the flood. It also ensures any variation in results between the methods is due to their differences rather than the flood simulation.

It is expected that the enhanced method's use of temporally specific population estimates will provide more detailed information on the disruption to road travel than the standard method. This data could be used to create flood risk assessments tailored to specific day times of interest. These findings will be useful to those working in the field of spatiotemporal population data, but also those working on assessing disruption to road travel, particularly from natural hazards.

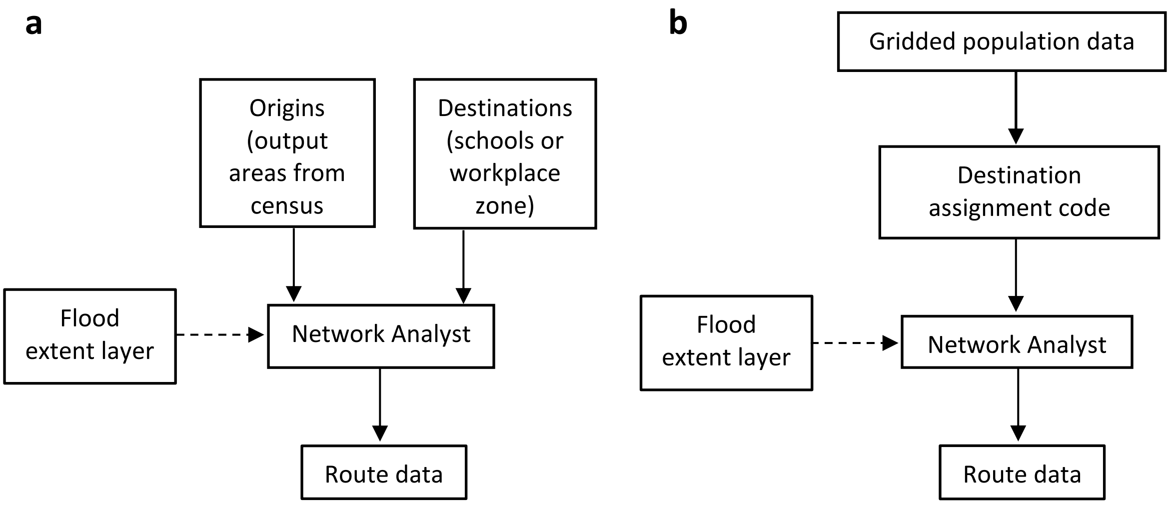


Figure 1: Standard method (a) and enhanced method (b) for assessing the impact of flooding on travel function. Dotted arrow indicates data which is included in the workflow when a flood scenario is considered.

References

Martin, D., Cockings, S., Leung, S, 2015 Developing a Flexible Framework for Spatiotemporal Population Modelling. *Annals of the Association of American Geographers,* 105(4), 754-772

Spatiotemporal modelling of flood-related impacts on daily population movement

Kate Rawlings[5][1], Jim Wright[6][1], Alan Smith[7][2], Sally Brown[8][3] and Jeremiah J Nieves[‡][1]

[1]School of Geography and Environmental Science, University of Southampton
[2]Geography, Earth and Environmental Sciences, University of Plymouth
[3]Department of Life and Environmental Sciences, Bournemouth University

July 2, 2020

Summary

This paper presents research combining spatiotemporal population flow data, flood modelling and network analysis to examine the effect of time of flood onset and flood magnitude on travel across a city for commuters and primary school children. Findings quantify that flood onset time has an effect on the disruption to travel comparable to flood event magnitude.

KEYWORDS: Spatiotemporal modelling, Network Analysis, Risk Analysis

1. Introduction and Background

The majority of fatalities caused by pluvial flash-floods in post-industrial countries concern those travelling in a vehicle (Debionne et al., 2016; Arrighi et al., 2019). The number of people travelling varies greatly over the course of the day, week or year, however mobility aspects are not frequently considered in flood exposure and risk assessments (Debionne et al., 2016; Dawson et al., 2011). This lack of dynamic population consideration means an important aspect of risk is missing when planning for flood events.

Risk is a dynamic phenomenon and varies over both space and time (Mechler & Bouwer, 2014). The ‘risk equation’ is a common conceptualisation of risk but rarely includes time or space (Hu et al., 2017). Therefore, in this paper it is adapted to include both (Equation 1).

$$\text{Risk} = f(\text{Hazard}_{st}, \text{Exposure}_{st}, \text{Vulnerability}_{st}) \qquad (1)$$

When studying the effect of floods on human populations, the time of flood onset is important as the same hazard event could have different effects due to variation in the exposure and vulnerability components of the risk equation (Freire et al., 2013; Dawson et al., 2011). Exposure is highly time dependent (Aubrecht et al., 2012), for example, a flood occurring during the Monday morning rush hour would mean more people are exposed on the roads than a Sunday morning.

---

[5] K.Rawlings@soton.ac.uk
[6] j.a.wright@soton.ac.uk
[7] alan.smith@plymouth.ac.uk
[8] sb20@soton.ac.uk
‡ j.j.nieves@soton.ac.uk

There are also differences in vulnerability as certain groups of the population are less able to react swiftly to floods, for example young children and the elderly (Smith et al., 2015).

This paper presents a proof of concept for combining spatiotemporal population flow data with flood data and network analysis to quantify the effect of time of flood onset versus size of flood hazard. The hypothesis is that the time of flood onset has an effect on travel disruption comparable to flood event magnitude. Commuters and primary school children are the population groups selected for analysis using a case study for York, UK.

2. Methods

Spatiotemporal population origins were produced in two different ways. (i) For commuters, census data describing output area to workplace zone commutes were combined with Labour Force Survey data to create temporal profiles of numbers of people travelling at a given time. (ii) The origin data for schoolchildren were generated through a spatially weighted Monte Carlo process converting spatiotemporal gridded population data into a set of likely origin centroids. The destinations were the population weighted centroids for the workplace zones and the point location of the primary schools. Inundated areas were delineated through pluvial flood modelling of York using the Flowroute-i™ model (Ambiental Risk Analytics). These data were analysed using ESRI's 'closest facility' algorithm and traffic data from HERE to model driving routes for a specific time of day.

A scenario-based approach was taken to test the hypothesis (Table 1). First a 'baseline' set of data were created to be a control comparison to the flood scenarios. Two factors, time of flood onset and size of flood event, were varied in turn. Three times of flood onset were picked to cover the morning commute, 6am (flood time 1, FT1), 7am (flood time 2, FT2) and 8am (flood time 3, FT3). 1 in 30 and 1 in 100 year flood layers formed barriers to travel at each of these onset times and the output routes saved for comparison to each other and the baseline. The network analysis was conducted at set points during the morning commute (7am, 7:30am, 8am, 8:30am and 9am) to capture variation. The analytical overview for this paper is presented in Figure 1. The origin, destination and flood layer (if applicable) were used in the network analysis, producing shapefiles of the routes taken between each origin-destination pair with distance and travel time. This procedure was repeated for each time step in the scenario.

Table 1 Scenarios run during the analysis

Scenario Name	Description
Non-flood Baseline	Network analysis with no flood layer included as a baseline.
Flood time 1, 30yr flood	Network analysis with flood layer for a 1 in 30 year event, flood onset at 6am.
Flood time 2, 30yr flood	Network analysis with flood layer for a 1 in 30 year event, flood onset at 7am.
Flood time 3, 30yr flood	Network analysis with flood layer for a 1 in 30 year event, flood onset at 8am.
Flood time 1, 100yr flood	Network analysis with flood layer for a 1 in 100 year event, flood onset at 6am.
Flood time 2, 100yr flood	Network analysis with flood layer for a 1 in 100 year event, flood onset at 7am.
Flood time 3, 100yr flood	Network analysis with flood layer for a 1 in 100 year event, flood onset at 8am.

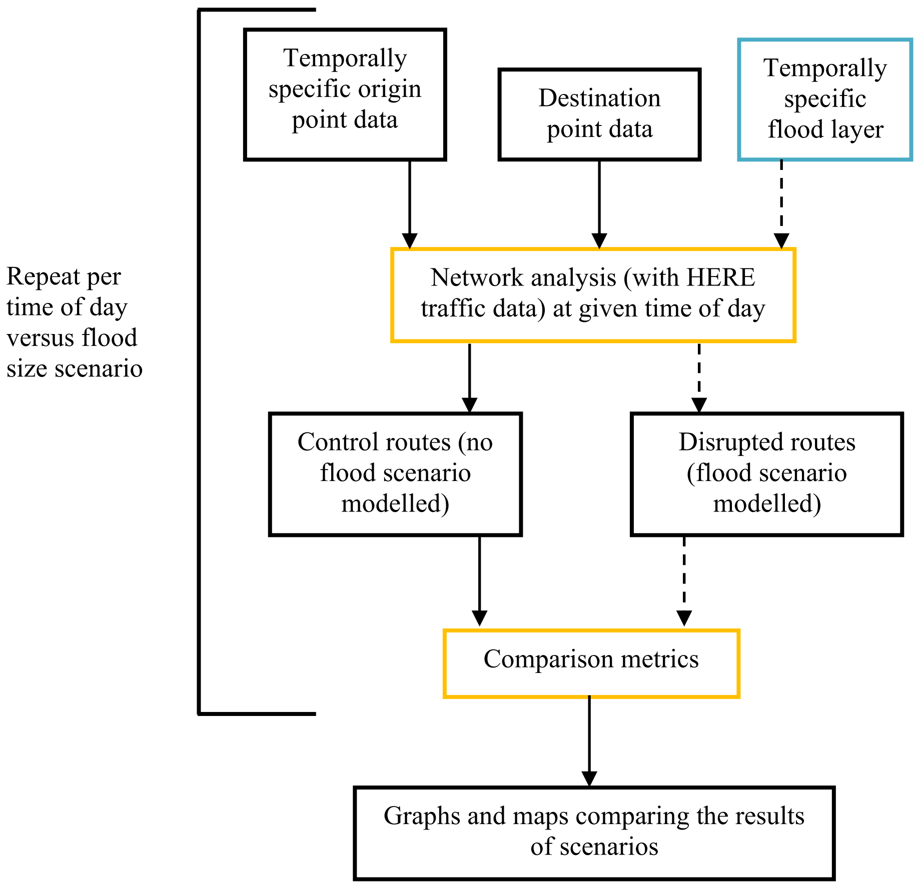


Figure 1 Diagram of analysis workflow. Black boxes indicate inputs and outputs, the blue box optional flood data, the yellow boxes processes and the dashed arrows steps included for flood scenario analysis.

3. Results and Discussion

Three examples of the results from the analysis are given. Firstly, Figure 2 compares road travel between the baseline and a flood scenario, the FT1 1 in 30yr scenario. Whilst there are similarities in road usage volumes, with the city centre and ring road most used, a northern part of the ring road is flooded in Figure 2b and longer routes are taken outside of the city to avoid flooding.

Secondly, Figure 3 compares the average travel time to an example destination between all scenarios. FT1 has the highest average travel times, higher than the baseline at all given time points, with FT2 rising above the baseline after 7:30am. In Figure 3a the increase in travel time for FT2 is similar to FT1 from 8:30am, and in Figure 3b average travel time rises but not enough to reach FT1. In both Figure 3a and 3b, the results for FT3 match the baseline until 9:00am where there is a slight increase from the baseline. The pattern of results is similar between Figure 3a and 3b, but with a greater increase in travel time seen in 3b as the flood magnitude has increased.

Finally Figure 4 is a summary of the effect of each disruption scenario on the city. The time lost was calculated by aggregating the additional travel time per origin-destination pair and, if a journey to a destination was not possible, the time for a full work/school day. This gives the total time lost at workplaces and schools due to flood disruptions. Commuters have more time lost partly due to higher numbers of commuters and workplaces meaning more journeys take place. It shows that the disruption is greater from FT1 for both magnitudes, and the decrease in time lost is greater when changing the time of flood onset than changes in the flood magnitude. This therefore provides evidence to support the hypothesis that flood onset time has a comparable effect to flood magnitude.

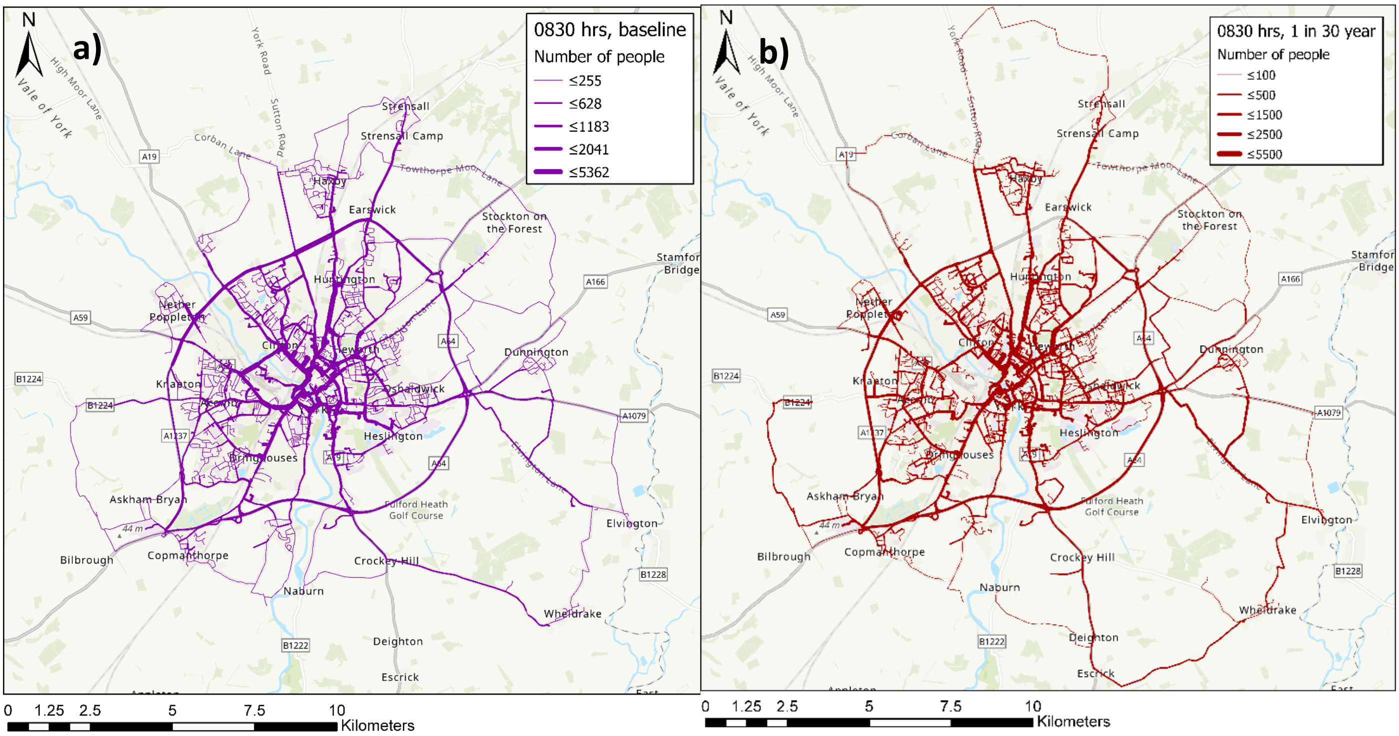


Figure 2 Road usage maps for 8:30am, a) non-flood baseline conditions and b) 1 in 30 year flood at flood onset time 1 (FT1) (6am).

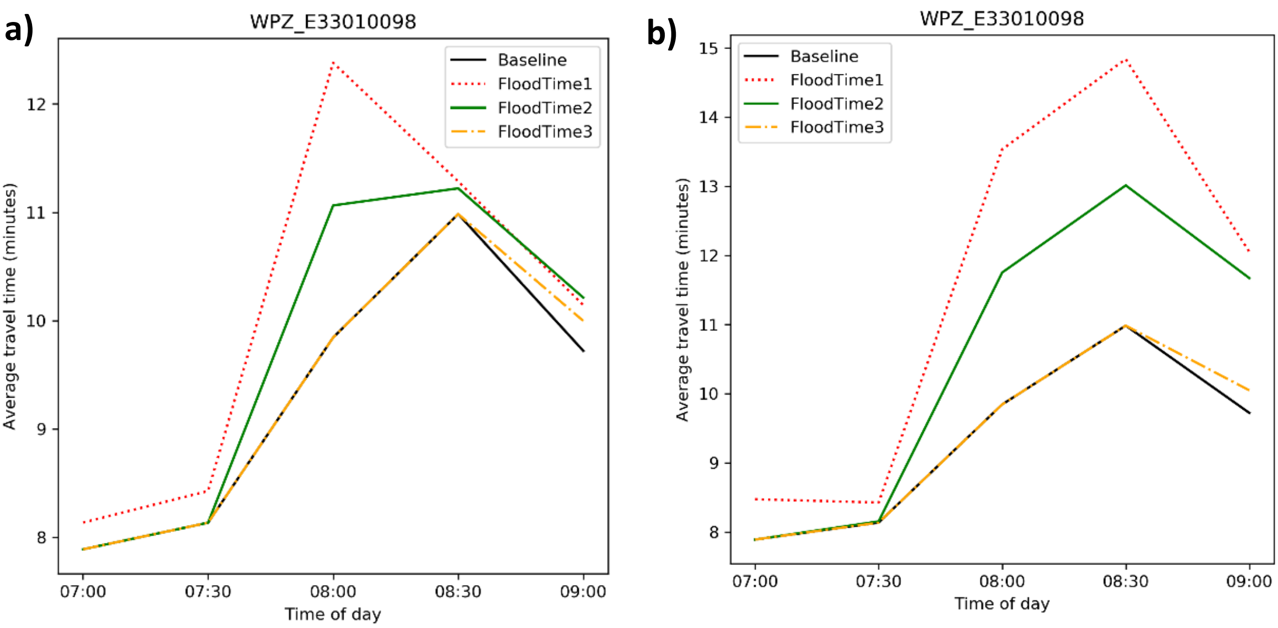


Figure 3 Average travel time of origins to an example destination (Workplace zone E33010098) for each flood magnitude a) 1 in 30 b) 1 in 100. The three flood onset scenarios plus the baseline are shown in each.

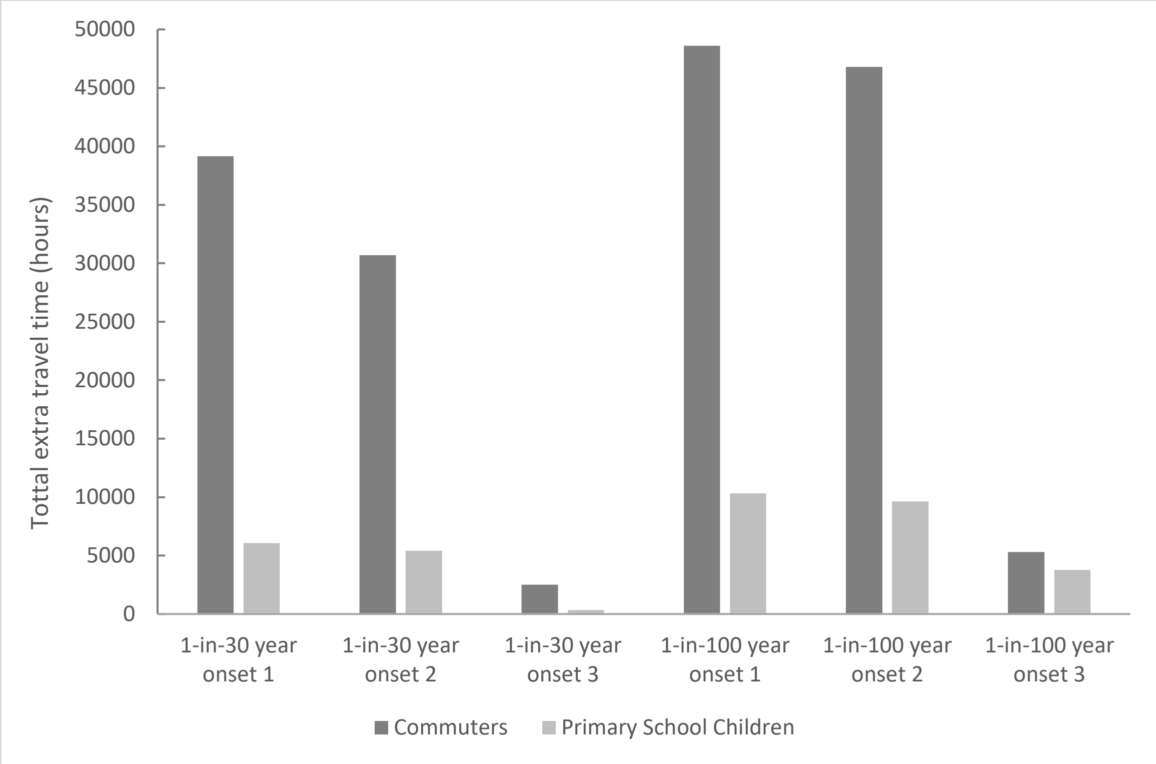


Figure 4 The total time lost (hours) at all destinations in each flood scenario in York. Dark grey represents the time lost by commuters at workplaces and the light grey the time lost by school children at schools.

4. Conclusion

These results show that there are spatial and temporal differences in the impact of flooding on road travel, with time of flood onset and magnitude affecting the average travel time to a destination. This workflow could be applied to other cities and types of hazard for risk assessments including spatiotemporal population data.

5. Acknowledgements

I would like to thank the ESRC's South Coast Doctoral Training Partnership (Grant number ES/J500161/1) for their funding and support of this research. The flood modelling for this research was done with Ambiental Risk Analytics using their Flowroute-i^{TM} model.

6. Biography

Kate Rawlings is a final year PhD student in the School of Geography and Environmental Science at the University of Southampton. Her research interests include risk theory, geospatial network analysis and the impact of natural hazards on travel.

Jim Wright is a Professor in Geographical Information Systems in the School of Geography and Environmental Science at the University of Southampton. His research examines how geospatial data can be used to tackle public health and environmental management problems.

Alan Smith is a Lecturer in Environmental Management in the School of Geography, Earth and Environmental Sciences at the University of Plymouth. His research interests include, spatiotemporal population mapping and modelling, hazard risk assessment and applied GIS.

Sally Brown is Deputy Head of Department for Life and Environmental Sciences at Bournemouth University. She is interested in coastal geomorphology, the impacts of sea-level rise on a range of settings and climate change adaptation at local to global scales, plus the long-term sustainability of coastal zones.

Jeremiah J Nieves focuses on spatio-temporal modelling of populations and urban growth at high-resolution using machine learning methods and remote sensing with applications in planning, public health, sustainability, and disaster risk reduction. He is finishing his PhD and has been a WorldPop researcher since 2014.

7. References

Arrighi, C., Pregnolato, M., Dawson, R.J. & Castelli, F. (2019). Preparedness against mobility disruption by floods. Science of the Total Environment. 654.

Aubrecht, C., Freire, S., Neuhold, C., Curtis, A. & Steinnocher, K. (2012). Introducing a temporal component in spatial vulnerability analysis. Disaster Advances. 5 (2).

Dawson, R.J., Peppe, R. & Wang, M. (2011). An agent-based model for risk-based flood incident management. Natural Hazards. 59 (1). p.pp. 167–189.

Debionne, S., Ruin, I., Shabou, S., Lutoff, C. & Creutin, J.D. (2016). Assessment of commuters' daily exposure to flash flooding over the roads of the Gard region, France. Journal of Hydrology. 541.

Freire, S., Aubrecht, C. & Wegscheider, S. (2013). Advancing tsunami risk assessment by improving spatio-temporal population exposure and evacuation modeling. Natural Hazards. 68 (3). p.pp. 1311–1324.

Hu, K., Yang, X., Zhong, J., Fei, F. & Qi, J. (2017). Spatially Explicit Mapping of Heat Health Risk Utilizing Environmental and Socioeconomic Data. Environmental Science and Technology. 51. p.pp. 1498–1507.

Mechler, R. & Bouwer, L.M. (2014). Understanding trends and projections of disaster losses and climate change: is vulnerability the missing link? Climatic Change. 133 (1). p.pp. 23–35.

Smith, A., Newing, A., Quinn, N., Martin, D., Cockings, S. & Neal, J. (2015). Assessing the Impact of Seasonal Population Fluctuation on Regional Flood Risk Management. ISPRS International Journal of Geo-Information. 4 (3). p.pp. 1118–1141.

Terti, G., Ruin, I., Anquetin, S. & Gourley, J.J. (2017). A situation-based analysis of flash flood fatalities in the united states. American Meteorological Society. p.pp. 333–345.

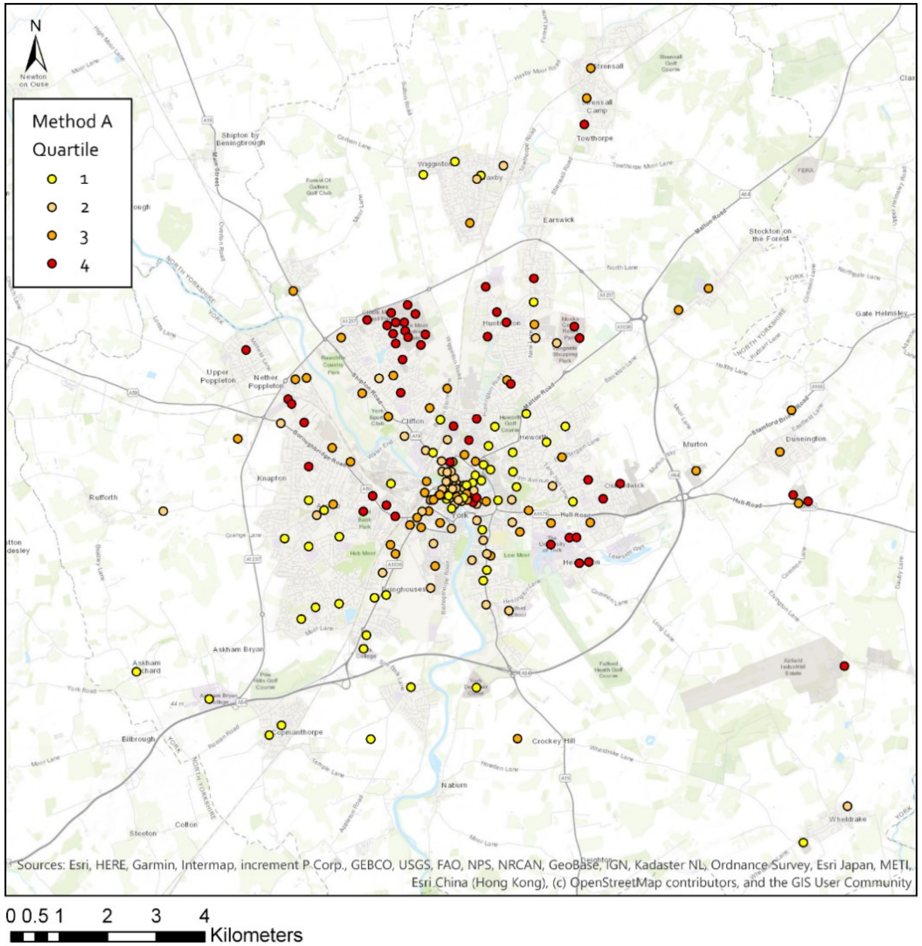
Appendix C Additional Figures for Chapter 6

Figure A.1 presents the 215 workplace zone destinations, categorised into quintiles based on the average travel time increase experienced to reach them during the flood event. The red destinations denote where no travel to this point was possible during the flood event, and these were broadly similar between the two sets of data with spatial clustering primarily in the city centre and a couple of places in the suburbs. In both data sets there is spatial clustering of destinations which were more affected by the flood event and those which were least affected.

Figure A.2. To identify spatial patterns in the flood scenario data, again the average travel time were mapped for each WPZ. This again showed spatial correlation in where the destinations in each quartile were distributed. There were broad similarities between the data for Method A and Method B.

Appendix C

Figure A.1 Method A (a) and Method B (b) average increase in travel time to a destination presented as quintiles. Contains public sector information licensed under the Open Government Licence v3.0



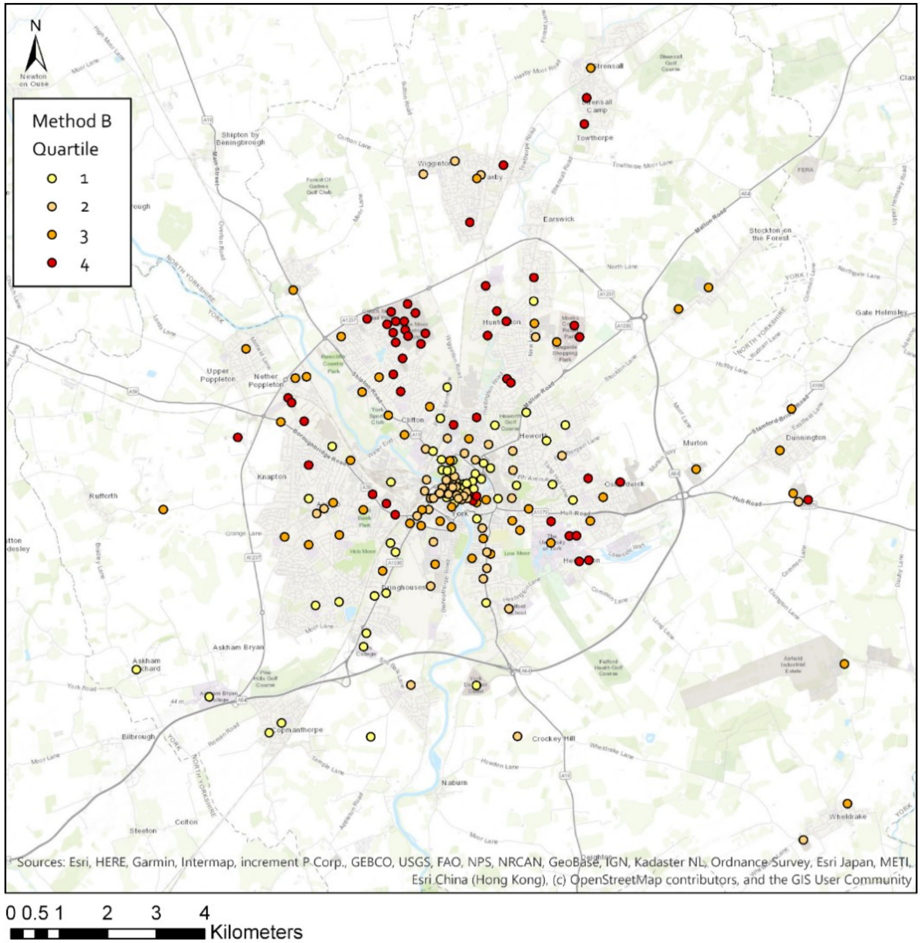


Figure A.2 Quartiles for destinations with flood scenario. Contains public sector information licensed under the Open Government Licence v3

Appendix C

Glossary of Terms

Exposure: The scope of population, infrastructure and other assets located in a potential hazard area (Hirabayashi et al., 2013)

Flood: Flood defined in insurance contracts as "temporary covering of land by water as a result of surface waters escaping from their normal confines or as a result of heavy precipitation" (Kron, 2005)

Fluvial: Water whose source is a river

Hazard: A threatening event which has the potential to cause harm and has a calculated probability of occurrence (Kron, 2005).

Natural Hazard: Any natural process which can result in harm to society e.g. volcanic eruption, earthquake, flood, wild fire, drought

Network: A series of links and nodes which combine to form a network

Network science: The study of network representations of physical, biological and social phenomena leading to predictive models of these phenomena (National Research Council, 2006). Ones involving roads must have the topology of the transport network and the location of people who are evacuating/moving (Brachman & Dragicevic, 2014)

Pluvial: Water accumulated on the surface through rainfall

Risk: Dynamic function of hazard, vulnerability and exposure (Mechler & Bouwer, 2014) which fluctuates over space and time. The function of hazard, vulnerability and exposure over space and time.

Spatiotemporal: Looking at both the spatial and temporal factors/dynamics of a phenomenon

Vulnerability: The degree a person is able to help themselves during a hazard event, influenced by the presence, or lack of, resources to cope with a hazard (Civil Contingencies Secretariat, 2008; McFadden *et al*., 2006) Everyone who is exposed is vulnerable to flooding, but some are more susceptible due to social factors, therefore the risk is greater (Rogelis, 2015).

List of References

Ahmed, S., Hefny, H.A. & Farid Ibrahim, R. (2017). GIS-Based Network Analysis for the Roads Network of the Greater Cairo Area Influence Cascade View project Enzyme Classification and Prediction View project GIS-Based Network Analysis for the Roads Network of the Greater Cairo Area. In: *Proceedings of the International Conference on Applied Research in Computer Science and Engineering*. 2017.

Alaeddine, H., Serrhini, K., Maizia, M. & Néron, E. (2015). A spatiotemporal optimization model for the evacuation of the population exposed to flood hazard. *Natural Hazards and Earth System Sciences*. 15 (3). p.pp. 687–701.

Albano, R., Mancusi, L., Sole, A. & Adamowski, J. (2015). Collaborative Strategies for Sustainable EU Flood Risk Management: FOSS and Geospatial Tools—Challenges and Opportunities for Operative Risk Analysis. *ISPRS International Journal of Geo-Information*. 4 (4). p.pp. 2704–2727.

Alexander, L., Jiang, S., Murga, M. & González, M.C. (2015). Origin–destination trips by purpose and time of day inferred from mobile phone data. *Transportation Research Part C: Emerging Technologies*. 58. p.pp. 240–250.

Ambiental Environmental Assessment (2020). *What is a flood risk assessment? | Ambiental*. [Online]. 2020. Available from: https://www.ambiental.co.uk/what-is-a-flood-risk-assessment/. [Accessed: 28 April 2020].

Anderson, T.K. (2009). Kernel density estimation and K-means clustering to profile road accident hotspots. *Accident Analysis & Prevention*. 41 (3). p.pp. 359–364.

Apparicio, P., Abdelmajid, M., Riva, M. & Shearmur, R. (2008). Comparing alternative approaches to measuring the geographical accessibility of urban health services: Distance types and aggregation-error issues. *International Journal of Health Geographics*. 7 (1). p.p. 7.

Arai, A., Horanont, T., Witayangkurn, A., Shao, X., Kanasugi, H. & Shibasaki, R. (2014). Understanding user attributes from calling behavior: Exploring call detail records through field observations. In: *12th International Conference on Advances in Mobile Computing and Multimedia, MoMM 2014*. 8 December 2014, New York, New York, USA: Association for Computing Machinery, Inc, pp. 95–104.

Arnell, N.W. & Gosling, S.N. (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change*. 134. p.pp. 387–401.

Arrighi, C., Pregnolato, M., Dawson, R.J. & Castelli, F. (2019). Preparedness against mobility disruption by floods. *Science of the Total Environment*. 654.

Aubrecht, C., Freire, S., Neuhold, C., Curtis, A. & Steinnocher, K. (2012). Introducing a temporal component in spatial vulnerability analysis. *Disaster Advances*. 5 (2).

Aubrecht, C., Fuchs, S. & Neuhold, C. (2013a). Spatio-temporal aspects and dimensions in integrated disaster risk management. *Natural Hazards*. 68 (3). p.pp. 1205–1216.

Aubrecht, C., Özceylan, D., Steinnocher, K. & Freire, S. (2013b). Multi-level geospatial modeling of human exposure patterns and vulnerability indicators. *Natural Hazards*. 68 (1). p.pp. 147–163.

Aubrecht, C., Steinnocher, K. & Huber, H. (2014). DynaPop - Population distribution dynamics as basis for social impact evaluation in crisis management. In: *ISCRAM 2014 Conference Proceedings - 11th International Conference on Information Systems for Crisis Response and Management*. 2014, pp. 314–318.

Baggio, R., Scott, N. & Cooper, C. (2010). Network science: A Review Focused on Tourism. *Annals of Tourism Research*. 37 (3). p.pp. 802–827.

Di Baldassarre, G., Viglione, A., Carr, G., Kuil, L., Salinas, J.L. & Blöschl, G. (2013). Socio-hydrology: conceptualising human-flood interactions. *Hydrology and Earth System Sciences*. 17. p.pp. 3295–3303.

Balica, S.F., Popescu, I., Beevers, L. & Wright, N.. (2013). Parametric and Physically Based Modelling Techniques for Flood Risk and Vulnerability Assessment: A Comparison. *Environmental Modelling & Software*. 41. p.pp. 84–92.

Balijepalli, C. & Oppong, O. (2014). Measuring vulnerability of road network considering the extent of serviceability of critical road links in urban areas. *Journal of Transport Geography*. 39. p.pp. 145–155.

Balk, D.L., Deichmann, U., Yetman, G., Pozzi, F., Hay, S.I. & Nelson, A. (2006). Determining Global Population Distribution: Methods, Applications and Data. *Advances in Parasitology*. 62. p.pp. 119–156.

Barnes, J.A. & Harary, F. (1983). Graph theory in network analysis. *Social Networks*. 5 (2). p.pp. 235–244.

Basiri, A., Jackson, M., Amirian, P., Pourabdollah, A., Sester, M., Winstanley, A., Moore, T. & Zhang, L. (2016). Quality assessment of OpenStreetMap data using trajectory mining. *Geo-spatial Information Science*. 19 (1). p.pp. 56–68.

Batista e Silva, F., Rosina, K., Schiavina, M., Herrera, M.A.M., Freire, S., Ziemba, L., Craglia, M. & Lavalle, C. (2018). From place of residence to place of activity: towards spatiotemporal mapping of population density in Europe. In: *AGILE*. 2018, Lund, pp. 1–7.

Beck, U. (2002). The Terrorist Threat World Risk Society Revisited. *Theory, Culture & Society*. 19 (4). p.pp. 39–55.

Berdica, K. (2002). An Introduction to Road Vulnerability: what has been done, is done and should be done. *Transport Policy*. 9. p.pp. 117–127.

Berry, T., Newing, A., Davies, D. & Branch, K. (2016). Using workplace population statistics to understand retail store performance. *The International Review of Retail, Distribution and Consumer Research*. 26 (4). p.pp. 375–395.

Bevan, J. (2018). *Surface water: The biggest flood risk of all - GOV.UK*. 2018. Environment Agency; CIWEM Surface Water conference. Available from: https://www.gov.uk/government/news/surface-water-the-biggest-flood-risk-of-all.

Bhaduri, B., Bright, E., Coleman, P. & Urban, M.L. (2007). LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal*. 69 (1–2). p.pp. 103–117.

Bian, R. & Wilmot, C.G. (2015). Spatiotemporal Population Distribution Method for Emergency Evacuation: Case Study of New Orleans, Louisiana. *Transportation Research Record*. (2532). p.pp. 99–106.

Blanc, J., Hall, J.W., Roche, N., Dawson, R.J., Cesses, Y., Burton, A. & Kilsby, C.G. (2012). Enhanced efficiency of pluvial flood risk estimation in urban areas using spatial-temporal rainfall simulations. *Journal of Flood Risk Management*. 5 (2). p.pp. 143–152

Bocarejo, J.P. & Oviedo, D.R. (2012). Transport accessibility and social inequities: a tool for identification of mobility needs and evaluation of transport investments. *Journal of Transport Geography*. 24. p.pp. 142–154.

Boeing, S., Rabb, B., Birch, C., Taylor, A., Blyth, A., Shelton, K., Hunter, N., Lamb, R., Trigg, M. & Hines, A. (2019). Trialling New Surface Water Flooding Forecasts in Yorkshire, UK. *Geophysical Research Abstracts*. 21.

Bogaert, P. (1996). Comparison of kriging techniques in a space-time context. *Mathematical Geology*. 28 (1). p.pp. 73–86.

Börner, K., Sanyal, S. & Vespignani, A. (2007). Network science. *Annual Review of Information Science and Technology*. 41 (1). p.pp. 537–607.

Borrmann, A., Kneidl, A., Köster, G., Ruzika, S. & Thiemann, M. (2012). Bidirectional coupling of macroscopic and microscopic pedestrian evacuation models. *Safety Science*. 50 (8). p.pp. 1695–1703.

Borruso, G. (2005). Network Density Estimation: Analysis of Point Patterns over a Network. In: O. Gervasi, M. L. Gavrilova, V. Kumar, A. Laganà, H. Pueh, L. Youngsong Mun, & D. Taniar (eds.). *International Conference on Computational Science and Its Applications – ICCSA 2005.* 2005, Springer, Berlin, Heidelberg, pp. 126–132.

Boscoe, F.P., Henry, K.A. & Zdeb, M.S. (2012). A Nationwide Comparison of Driving Distance Versus Straight-Line Distance to Hospitals. *The Professional Geographer*. 64 (2). p.pp. 188–196.

Brachman, M.L. & Dragicevic, S. (2014). A spatially explicit network science model for emergency evacuations in an urban context. *Computers, Environment and Urban Systems*. 44. p.pp. 15–26.

Bracken, I. & Martin, D. (1989). The Generation of Spatial Population Distributions from Census Centroid Data. *Environment and Planning A*. 21 (4). p.pp. 537–543.

Brown, R. (2014). *Transport Resilience Review: A Review of the Resilience of the Transport Network to Extreme Weather Events* D. for Transport (ed.).

Brownlee, J. (2018). *A Gentle Introduction to the Bootstrap Method*. May 2018. Machine Learning Mastery. Available from: https://machinelearningmastery.com/a-gentle-introduction-to-the-bootstrap-method/. [Accessed: 16 June 2020].

Büchele, B., Kreibich, H., Kron, A., Thieken, A., Ihringer, J., Oberle, P., Merz, B. & Nestmann, F. (2006). Flood-risk mapping: contributions towards an enhanced assessment of extreme events and associated risks. *Natural Hazards and Earth System Sciences*. 6 (4). p.pp. 485–503.

Burns, C.M. & Inglis, A.D. (2007). Measuring food access in Melbourne: Access to healthy and fast foods by car, bus and foot in an urban municipality in Melbourne. *Health & Place*. 13 (4). p.pp. 877–885.

Cabinet Office (2011). *Keeping the Country Running: Natural Hazards and Infrastructure*. [Online]. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/61342/natural-hazards-infrastructure.pdf.

Cabinet Office (2016). *National Flood Resilience Review*. [Online]. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/551137/national-flood-resilience-review.pdf.

Cabinet Office (2017). *National Risk Register of Civil Emergencies*.

Cahyanto, I., Pennington-Gray, L., Thapa, B., Srinivasan, S., Villegas, J., Matyas, C. & Kiousis, S. (2014). An empirical evaluation of the determinants of tourist's hurricane evacuation decision making. *Journal of Destination Marketing & Management*. 2 (4). p.pp. 253–265.

Cairns, S., Atkins, S. & Goodwin, P. (2002). Disappearing traffic? The story so far. *Proceedings of the Institution of Civil Engineers*. (1). p.pp. 13–22.

Calabrese, F., Di Lorenzo, G., Liu, L. & Ratti, C. (2011). Estimating Origin-Destination Flows using Opportunistically Collected Mobile Phone Location Data from One Million Users in Boston Metropolitan Area. In: *IEEE Pervasive Computing*. 2011, Institute of Electrical and Electronics Engineers (IEEE).

Carrão, H., Naumann, G. & Barbosa, P. (2016). Mapping global patterns of drought risk: An empirical framework based on sub-national estimates of hazard, exposure and vulnerability. *Global Environmental Change*. 39. p.pp. 108–124.

Carter, P. (2015). UK Infrastructure: Can we Cope with Flooding? *Network: A Technical Journal by WSP*. (79). p.pp. 84–85.

Castellarin, A., Di Baldassarre, G., Bates, P.. & Brath, A. (2009). Optimal Cross-Sectional Spacing in Preissmann Scheme 1D Hydrodynamic Models. *Journal of Hydraulic Engineering*. 298 (1–4). p.pp. 242–266.

Cats, O. & Jenelius, E. (2016). Beyond a complete failure: the impact of partial capacity degradation on public transport network vulnerability. *Transportmetrica B: Transport Dynamics*. p.pp. 1–20.

Cats, O. & Jenelius, E. (2014). Dynamic Vulnerability Analysis of Public Transport Networks: Mitigation Effects of Real-Time Information. *Networks and Spatial Economics*. 14 (3–4). p.pp. 435–463.

Cats, O. & Jenelius, E. (2015). Planning for the unexpected: The value of reserve capacity for public transport network robustness. *Transportation Research Part A: Policy and Practice*. 81. p.pp. 47–61.

Centre for the Protection of National Infrastructure (2020). *Critical National Infrastructure | CPNI | Public Website*. 2020. Available from: https://www.cpni.gov.uk/critical-national-infrastructure-0. [Accessed: 7 July 2020].

Chan, S.C., Kahana, R., Kendon, E.J. & Fowler, H.J. (2018). Projected changes in extreme precipitation over Scotland and Northern England using a high-resolution regional climate model. *Climate Dynamics*. 51 (9–10). p.pp. 3559–3577.

Chandra, S. & Quadrifoglio, L. (2013). Critical street links for demand responsive feeder transit services. *Computers & Industrial Engineering*. 66 (3). p.pp. 584–592.

Chang, H., Lafrenz, M., Jung, I.-W., Figliozzi, M., Platman, D. & Pederson, C. (2010). Potential Impacts of Climate Change on Flood-Induced Travel Disruptions: A Case Study of Portland, Oregon, USA. *Annals of the Association of American Geographers*. 100 (4). p.pp. 938–952.

Charreire, H., Casey, R., Salze, P., Simon, C., Chaix, B., Banos, A., Badariotti, D., Weber, C. & Oppert, J.-M. (2010). Measuring the food environment using geographical information systems: a methodological review. *Public Health Nutrition*. 13 (11). p.pp. 1773–1785.

Chen, A., Yang, C., Kongsomsaksakul, S. & Lee, M. (2007). Network-based Accessibility Measures for Vulnerability Analysis of Degradable Transportation Networks. *Networks and Spatial Economics*. 7 (3). p.pp. 241–256.

Chen, J.-C. (2003). Dijkstra's Shortest Path Algorithm. *JOURNAL OF FORMALIZED MATHEMATICS*. 15.

Chen, X.-Z., Lu, Q.-C., Peng, Z.-R. & Ash, J.E. (2015). Analysis of Transportation Network Vulnerability Under Flooding Disasters. *Transportation Research Record*. 2532. p.pp. 37–44.

City of York Council (2013). *City of York Strategic Flood Risk Assessment*.

City of York Council (2018). *Headline figures for York (Census 2011)*. 2018. Census.

City of York Council (2020). *Primary schools*. [Online]. 2020. Available from: https://www.york.gov.uk/directory/13/schools-in-york/category/19. [Accessed: 4 August 2020].

City of York Council (2012). *Surface Water Management Plan*. [Online]. Available from:

https://www.york.gov.uk/downloads/file/11055/surface_water_management_plan.

City of York Council (2019). *York 13 August 2018 Section 19 Surface Water Flood Investigation*. [Online]. Available from: https://democracy.york.gov.uk/documents/s130816/Annex 1 York 130818 Section 19 Surface Water Flood Investigation March19.pdf. [Accessed: 21 April 2020].

Civil Contingencies Secretariat (2008). *Identifying People Who Are Vulnerable in a Crisis*. [Online]. p.p. 31. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/61228/vulnerable_guidance.pdf.

Cockings, S., Martin, D. & Harfoot, A. (2020). Developing a National Geodemographic Classification of Workplace Zones. *Applied Spatial Analysis and Policy*. p.pp. 1–25.

Coles, D., Yu, D., Wilby, R.L., Green, D. & Herring, Z. (2017). Beyond 'flood hotspots': Modelling emergency service accessibility during flooding in York, UK. *Journal of Hydrology*. 546. p.pp. 419–436.

Comber, A., Brunsdon, C. & Green, E. (2008). Using a GIS-based network analysis to determine urban greenspace accessibility for different ethnic and religious groups. *Landscape and Urban Planning*. 86 (1). p.pp. 103–114.

Conrad, S.H., LeClaire, R.J., O'Reilly, G.P. & Uzunalioglu, H. (2006). Critical national infrastructure reliability modeling and analysis. *Bell Labs Technical Journal*. 11. p.pp. 57–71.

Crichton, D. (2008). Role of Insurance in Reducing Flood Risk. *The Geneva Papers*. 33. p.pp. 117–132.

Crols, T. & Malleson, N. (2019). Quantifying the ambient population using hourly population footfall data and an agent-based model of daily mobility. *GeoInformatica*. 23 (2). p.pp. 201–220.

Currie, G. (2010). Quantifying spatial gaps in public transport supply based on social needs. *Journal of Transport Geography*. 18 (1). p.pp. 31–41.

Cutter, S.L. (1996). Societal Responses to Environmental Hazards. *International Social Science Journal*. 150. p.pp. 525–536.

Cutter, S.L., Barnes, L., Berry, M., Burton, C., Evans, E., Tate, E. & Webb, J. (2008). A place-based model for understanding community resilience to natural disasters. *Global Environmental*

Change. 18. p.pp. 598–606.

D'Este, G.M. & Taylor, M.A.P. (2003). Network Vulnerability: An Approach to Reliability Analysis at the Level of National Strategic Transport Networks. In: *The Network Reliability of Transport*. Emerald Group Publishing Limited, pp. 23–44.

Danese, M., Lazzari, M. & Murgante, B. (2008). Kernel Density Estimation Methods for a Geostatistical Approach in Seismic Risk Analysis: The Case Study of Potenza Hilltop Town (Southern Italy). In: Gervasi O., Murgante B., Laganà A., Taniar D., Mun Y., & Gavrilova M.L. (eds.). *Computational Science and Its Applications – ICCSA 2008*. 2008, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 415–429.

Davidson, P. & Davidson, T. (n.d.). *Transport Modelling - Distribution Models*. [Online]. Available from: http://www.transportmodeller.com/distributionoverview.html. [Accessed: 15 May 2018].

Dawson, R.J., Peppe, R. & Wang, M. (2011). An agent-based model for risk-based flood incident management. *Natural Hazards*. 59 (1). p.pp. 167–189.

Debionne, S., Ruin, I., Shabou, S., Lutoff, C. & Creutin, J.D. (2016). Assessment of commuters' daily exposure to flash flooding over the roads of the Gard region, France. *Journal of Hydrology*. 541.

DEFRA (2018). *Surface Water Management An Action Plan*. Available from: www.nationalarchives.gov.uk/doc/open-government-licence/version/3/oremailPSI@nationalarchives.gsi.gov.ukwww.gov.uk/defra. [Accessed: 28 January 2020].

DEFRA & EA. (2019). *Preparing a flood risk assessment: standing advice - GOV.UK*. [Online]. 1 March 2019. Government Guidance. Available from: https://www.gov.uk/guidance/flood-risk-assessment-standing-advice. [Accessed: 28 April 2020].

Deichmann, U., Balk, D. & Yetman, G. (2001). *Transforming Population Data for Interdisciplinary Usages: From census to grid*. Washington DC. Available from: http://sedac.ciesin.org/gpw-v2/GPWdocumentation.pdf. [Accessed: 26 February 2018].

Demšar, U., Špatenková, O. & Virrantaus, K. (2008). Identifying Critical Locations in a Spatial Network with Graph Theory. *Transactions in GIS*. 12 (1). p.pp. 61–82.

Department for Education (2018). *National pupil database* 2018. Available from: https://www.gov.uk/government/collections/national-pupil-database. [Accessed: 4 August

2020].

Department for Education (2016). *Schools, pupils and their characteristics: January 2016*. 2016. Local authority tables and cross-border movement tables.

Department for Transport (n.d.). *HS2 Outline Business Case: Economic Case*. Available from: https://www.gov.uk/government/publications/rail-alternatives-to-. [Accessed: 25 May 2020].

Department for Transport (2014). *National Travel Survey 2014: Travel to school*. London. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/476635/travel-to-school.pdf.

Department for Transport (2018a). *NTM Future Model Development: NTMv2 recalibration NTMv2R: Demand Model Implementation Department for Transport*. [Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/716487/national-transport-model-implementation-report.pdf. [Accessed: 11 August 2020].

Department for Transport (2018b). *Road traffic statistics - Download data*. 2018. Available from: https://roadtraffic.dft.gov.uk/downloads. [Accessed: 11 August 2020].

Devia, G.K., Ganasri, B.P. & Dwarakish, G.S. (2015). A Review on Hydrological Models. *Aquatic Procedia*. 4. p.pp. 1001–1007.

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F.R., Gaughan, A.E., Blondel, V.D. & Tatem, A.J. (2014). Dynamic population mapping using mobile phone data. *Proceedings of the National Academy of Sciences of the United States of America*. 111 (45). p.pp. 15888–93.

Dijkstra, E.W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*. 1 (1). p.pp. 269–271.

Dingil, A.E., Schweizer, J., Rupi, F. & Stasiskiene, Z. (2018). Transport indicator analysis and comparison of 151 urban areas, based on open source data. *European Transport Research Review*. 10 (2). p.pp. 1–9.

Discovery Software (2020). *STEMgis - Mapping In Time - An Affordable Temporal GIS - Time-Based Mapping Software*. 2020. Available from: https://www.discoverysoftware.co.uk/STEMgis.htm. [Accessed: 4 August 2020].

Donaldson, A., Lane, S., Ward, N. & Whatmore, S. (2013). Overflowing with issues: following the political trajectories of flooding. *Environment and Planning C: Government and Policy*. 31. p.pp. 602–618.

Donner, W. & Rodríguez, H. (2008). Population Composition, Migration and Inequality: The Influence of Demographic Changes on Disaster Risk and Vulnerability. *Source: Social Forces*. 87 (2). p.pp. 1089–1114.

Dottori, F., Di Baldassarre, G. & Todini, E. (2013). Detailed data is welcome, but with a pinch of salt: Accuracy, precision, and uncertainty in flood inundation modelling. *Water Resources Research*. 49. p.pp. 6079–6085.

Duan, Y. & Lu, F. (2014). Robustness of city road networks at different granularities. *Physica A: Statistical Mechanics and its Applications*. 411. p.pp. 21–34.

Ducruet, C. & Beauguitte, L. (2014). Spatial Science and Network Science: Review and Outcomes of a Complex Relationship. *Networks and Spatial Economics*. 14 (3–4). p.pp. 297–316.

Easton, S. & Ferrari, E. (2015). Children’s travel to school—the interaction of individual, neighbourhood and school factors. *Transport Policy*. 44. p.pp. 9–18.

Eiser, R.J., Bostrom, A., Burton, I., Johnston, D.M., McClure, J., Paton, D., van der Pligt, J. & White, M.P. (2012). Risk interpretation and action: A conceptual framework for responses to natural hazards. *International Journal of Disaster Risk Reduction*. 1. p.pp. 5–16.

Ellis, J.B. & Viavattene, C. (2014). Sustainable Urban Drainage System Modeling for Managing Urban Surface Water Flood Risk. *CLEAN - Soil, Air, Water*. 42 (2). p.pp. 153–159.

Emanuelsson, M.A.E., Mcintyre, N., Hunt, C.F., Mawle, R., Kitson, J. & Voulvoulis, N. (2014). Flood risk assessment for infrastructure networks. *Journal of Flood Risk Management*. 7 (1).

Eneh, A. & Arinze, U. (2017). COMPARATIVE ANALYSIS AND IMPLEMENTATION OF DIJKSTRA’S SHORTEST PATH ALGORITHM FOR EMERGENCY RESPONSE AND LOGISTIC PLANNING. *Nigerian Journal of Technology*. 36 (3). p.pp. 876–888.

Environment Agency (2016). *Keeping you informed: Recent Flooding in York* . Available from: https://www.york.gov.uk/downloads/file/2066/foss-barrier-keeping-you-informed-. [Accessed: 28 June 2020].

Environment Agency (2020). *York Flood Alleviation Scheme (FAS) Information page - Environment Agency - Citizen Space*. 2020. Available from: https://consult.environment-agency.gov.uk/yorkshire/yorkfas/. [Accessed: 24 June 2020].

ESRI (2020a). *Algorithms used by the ArcGIS Network Analyst extension—Help | Documentation*. [Online]. 2020. ArcMap Extensions. Available from: https://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/algorithms-used-by-network-analyst.htm#GUID-4F3B8C00-6AD2-4C22-A660-735A61B834DF. [Accessed: 1 July 2020].

ESRI (2020b). *Closest Facility service with asynchronous execution—ArcGIS REST API: Directions and Routing Services | ArcGIS for Developers*. [Online]. 2020. Location-Based Services Documentation. Available from: https://developers.arcgis.com/rest/network/api-reference/closest-facility-asynchronous-service.htm#GUID-1A5A3DA1-C14D-4355-AD72-67217CE1F7CA. [Accessed: 1 July 2020].

ESRI (2018). *Consuming the traffic service using ArcMap—Help | ArcGIS Desktop*. [Online]. 2018. ArcGIS Online Services. Available from: http://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/consuming-the-traffic-service-from-arcmap-10-2-.htm. [Accessed: 17 May 2018].

ESRI (2020c). *Directional Distribution (Standard Deviational Ellipse)—Help | ArcGIS Desktop*. [Online]. 2020. ArcMap Tools. Available from: https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/directional-distribution.htm. [Accessed: 1 July 2020].

ESRI (2020d). *HERE Technologies | Esri Partner*. [Online]. 2020. Available from: https://www.esri.com/partners/here-technologies-a2T70000000TOWiEAO. [Accessed: 4 August 2020].

ESRI (2020e). *Routing and Directions with ArcGIS*. [Online]. 2020. Available from: https://developers.arcgis.com/features/directions/#query-traffic. [Accessed: 4 August 2020].

ESRI (2020f). *Space-time cluster analysis—ArcGIS Pro | Documentation*. [Online]. 2020. Available from: https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/space-time-analysis.htm. [Accessed: 4 August 2020].

ESRI (2019). *World Traffic Service*. [Online]. 2019. Available from: https://www.arcgis.com/home/item.html?id=ff11eb5b930b4fabba15c47feb130de4. [Accessed: 4 August 2020].

Fang, J.Q., Wang, X.F., Zheng, Z.G., Bi, Q., Di, Z.. & Xiang, L. (2007). New Interdisciplinary Science: Network Science. *Progress in Physics-Nanjing*. 27 (3). p.p. 239.

Fekete, A. (2009). Validation of a social vulnerability index in context to river-floods in Germany. *Natural Hazards and Earth System Sciences*. 9 (2). p.pp. 393–403.

FloodRE (2020). *Flood Re - A flood re-insurance scheme*. 2020. Available from: https://www.floodre.co.uk/. [Accessed: 4 August 2020].

Ford, A., Barr, S., Dawson, R. & James, P. (2015a). Transport Accessibility Analysis Using GIS: Assessing Sustainable Transport in London. *ISPRS International Journal of Geo-Information*. 4 (1). p.pp. 124–149.

Ford, A., Jenkins, K., Dawson, R., Pregnolato, M., Barr, S. & Hall, J. (2015b). Simulating Impacts of Extreme Weather Events on Urban Transport Infrastructure in the UK. In: T. Dolan & B. Collins (eds.). *International Symposium for Next Generation Infrastructure Conference Proceedings: 30 September - 1 October 2014*. 2015, Vienna: UCL STEaPP, pp. 233–238.

Fotheringham, A., Brunsdon, C. & Charlton, M. (2000). *Quantitative Geography : Perspectives on Spatial Data Analysis.* Sage Publications.

Freire, S. & Aubrecht, C. (2012). Integrating population dynamics into mapping human exposure to seismic hazard. *Natural Hazards and Earth System Science*. 12 (11). p.pp. 3533–3543.

Freire, S., Aubrecht, C. & Wegscheider, S. (2013). Advancing tsunami risk assessment by improving spatio-temporal population exposure and evacuation modeling. *Natural Hazards*. 68 (3). p.pp. 1311–1324.

Freire, S., Ehrlich, D. & Ferri, S. (2015). Population Exposure and Impacts from Earthquakes: Assessing Spatio-temporal Changes in the XX Century. *Computer Modeling in Engineering & Science*. 109–110 (2). p.pp. 159–182.

Freiria, S., Tavares, A.O. & Pedro Julião, R. (2015). The Multiscale Importance of Road Segments in a Network Disruption Scenario: A Risk-Based Approach. *Risk Analysis*. 35 (3). p.pp. 484–500.

Fuchs, S., Keiler, M., Sokratov, S. & Shnyparkov, A. (2013). Spatiotemporal dynamics: the need for an innovative approach in mountain hazard risk management. *Natural Hazards*. 68. p.pp. 1217–1241.

Gallegos, H.A., Schubert, J.. & Sanders, B.F. (2009). Two-Dimensional, High-Resolution Modelling of Urban Dam-Break Flooding: A Case Study of Baldwin Hills, California. *Advances in Water*

Resources. 32. p.pp. 1323–1335.

Galway, L., Bell, N., SAE, A., Hagopian, A., Burnham, G., Flaxman, A., Weiss, W.M., Rajaratnam, J. & Takaro, T.K. (2012). A two-stage cluster sampling method using gridded population data, a GIS, and Google EarthTM imagery in a population-based mortality survey in Iraq. *International Journal of Health Geographics*. 11 (1). p.p. 12.

Gatrell, A.C., Bailey, T.C., Diggle, P.J. & Rowlingson, B.S. (1996). Spatial Point Pattern Analysis and Its Application in Geographical Epidemiology. *Transaction of the Institute of British Geographer*. 21. p.pp. 256–271.

GeoData Institute (2016). *Pop247*. 2016. Available from: http://pop247.geodata.soton.ac.uk/. [Accessed: 4 August 2020].

GeoData Institute (2018). *POP247NRT*. 2018. Available from: http://pop247nrt.geodata.soton.ac.uk/. [Accessed: 4 August 2020].

Gerber, M.S. (2014). Predicting crime using Twitter and kernel density estimation. *Decision Support Systems*. 61. p.pp. 115–125.

Ghimire, B., Chen, A.S., Guidolin, M., Keedwell, E.C., Djordjević, S. & Savić, D.A. (2013). Formulation of a fast 2D urban pluvial flood model using a cellular automata approach. *Journal of Hydroinformatics*. 15 (3). p.pp. 676–686.

Gil, J. (2015). Building a multimodal urban network model Using OpenStreetMap data for the analysis of sustainable accessibility. *Lecture Notes in Geoinformation and Cartography*. (9783319142791). p.pp. 229–251.

Google (2020). *Google Cloud Routes*. [Online]. 2020. Available from: https://cloud.google.com/maps-platform/routes. [Accessed: 4 August 2020].

Gov.UK (n.d.). *Check the long term flood risk for an area in England - GOV.UK*. Available from: https://flood-warning-information.service.gov.uk/long-term-flood-risk?_ga=2.168278690.267046992.1597324686-1649644163.1580466423. [Accessed: 13 August 2020].

Gross, J. & Yellen, J. (2014). *Handbook of Graph Theory*. J. Gross, J. Yellen, & P. Zhang (eds.). CRC Press.

Gugat, M. & Leugering, G. (2003). Global Boundary Controllability of the De St.Venant Equations Between Steady States. *Annales de l’IHP Analyse non linéaire*. 20 (1). p.pp. 1–11.

Hallegatte, S., Green, C., Nicholls, R.J. & Corfee-Morlot, J. (2013). Future flood losses in major coastal cities. *Nature Climate Change*. 3. p.pp. 802–806.

Hammond, M.J., Chen, A.S., Djordjević, S., Butler, D. & Mark, O. (2015). Urban flood impact assessment: A state-of-the-art review. *Urban Water Journal*. 12 (1). p.pp. 14–29.

Han, G. & Sohn, K. (2016). Activity imputation for trip-chains elicited from smart-card data using a continuous hidden Markov model. *Transportation Research Part B: Methodological*. 83. p.pp. 121–135.

Hankin, B., Craigen, I., Rogers, W., Morphet, J., Bailey, A. & Whitehead, M. (2016). Flood Risk to the Strategic Road Network in England M. Lang, F. Klijn, & P. Samuels (eds.). *E3S Web of Conferences*. 7. p.p. 10001.

Harland, K. & Stillwell, J. (2010). Commuting to School: A New Spatial Interaction Modelling Framework. In: J. Stillwell, O. Duke-Williams, & A. Dennett (eds.). *Technologies for Migration and Commuting Analysis: Spatial Interaction Data Applications*. Hershey, p. 357.

Harman, B.I., Koseoglu, H. & Yigit, C.O. (2016). Performance evaluation of IDW, Kriging and multiquadric interpolation methods in producing noise mapping: A case study at the city of Isparta, Turkey. *Applied Acoustics*. 112. p.pp. 147–157.

Health and Safety Laboratory (2020). *National Population Database*. 2020. Available from: https://data.gov.uk/dataset/72d8bf63-7cc2-48b6-8d97-a85d482bf29c/national-population-database. [Accessed: 4 August 2020].

Hemingway, L. & Priestley, M. (2006). Natural Hazards, Human Vulnerability and Disabling Societies: A Disaster for Disabled People? *The Review of Disability Studies*. 2 (3). p.p. .

Hemson, G., Johson, P., South, A., Kenward, R., Ripley, R. & MacDonald, D. (2005). Are Kernel the Mustard? Data From Global Positioning System (GPS) Collars Suggests Problems for Kernel Home-Range Analyses with Least-Squares Cross-Validation. *Journal of Animal Ecology*. 74. p.pp. 455–463.

Hengl, T., Heuvelink, G.B.M. & Rossiter, D.G. (2007). About regression-kriging: From equations to case studies. *Computers & Geosciences*. 33 (10). p.pp. 1301–1315.

Heppenstall, A., Malleson, N. & Crooks, A. (2016). “Space, the Final Frontier”: How Good are Agent-Based Models at Simulating Individuals and Space in Cities? *Systems*. 4 (1). p.p. 9.

Herrera, J.C., Work, D.B., Herring, R., Ban, X. (Jeff), Jacobson, Q. & Bayen, A.M. (2010). Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transportation Research Part C: Emerging Technologies*. 18 (4). p.pp. 568–583.

Heuvelink, G.B.M., Pebesma, E. & Gräler, B. (2016). Space-Time Geostatistics. In: *Encyclopedia of GIS*. Cham: Springer International Publishing, pp. 1–7.

Heywood, I., Cornelius, S. & Carver, S. (2011). *An Introduction to Geographical Information Systems*. 4th Editio. Harlow: Pearson Education Limited.

Higgs, G., Fry, R. & Langford, M. (2012). Investigating the implications of using alternative GIS-based techniques to measure accessibility to green space. *Environment and Planning B: Planning and Design*. 39. p.pp. 326–343.

Highways England (n.d.). *Highways England – Data.gov.uk – Journey Time and Traffic Flow Data April 2015 onwards – User Guide*. Available from: http://data.highways.gov.uk/ha-journey-times/TRIS+-+User+Guide+r3.pdf.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H. & Kanae, S. (2013). Global flood risk under climate change. *Nature Climate Change*. 3 (9). p.pp. 816–821.

Holt, J.B., Lo, C.P. & Hodler, T.W. (2004). Dasymetric Estimation of Population Density and Areal Interpolation of Census Data. *Cartography and Geographic Information Science*. 31 (2). p.pp. 103–121.

Houston, D., Werritty, A., Bassett, D., Geddes, A., Hoolachan, A. & McMillon, M. (2011). *Pluvial (rain-related) Flooding in Urban Areas: The Invisible Hazard*. Available from: https://www.researchgate.net/publication/266175667_Pluvial_rain-related_flooding_in_urban_areas_the_invisible_hazard.

Hu, K., Yang, X., Zhong, J., Fei, F. & Qi, J. (2017). Spatially Explicit Mapping of Heat Health Risk Utilizing Environmental and Socioeconomic Data. *Environmental Science and Technology*. 51. p.pp. 1498–1507.

Jaroszweski, D., Hooper, E., Baker, C., Chapman, L. & Quinn, A. (2015). The impacts of the 28 June 2012 storms on UK road and rail transport. *Meteorological Applications*. 22 (3). p.pp. 470–476.

Jedlicka, K., Hajek, P., Cada, V., Martolos, J., Stastny, J., Beran, D., Kolovsky, F. & Kozhukh, D. (2016). Open transport map-Routable OpenStreetMap. In: *2016 IST-Africa Conference, IST-Africa 2016*. 2 August 2016, Institute of Electrical and Electronics Engineers Inc.

Jenelius, E. (2009). Network structure and travel patterns: explaining the geographical disparities of road network vulnerability. *Journal of Transport Geography*. 17 (3). p.pp. 234–244.

Jenelius, E. (2010a). Redundancy importance: Links as rerouting alternatives during road network disruptions. *Procedia Engineering*. 3. p.pp. 129–137.

Jenelius, E. (2010b). User Inequity Implications of Road Network Vulnerability. *Journal of Transport and Land Use*. 2 (3/4). p.pp. 57–73.

Jenelius, E. & Mattsson, L.-G. (2015). *Road network vulnerability analysis: Conceptualization, implementation and application*. Available from: http://ac.els-cdn.com/S0198971514000192/1-s2.0-S0198971514000192-main.pdf?_tid=1211fda0-ef84-11e6-b99b-00000aab0f26&acdnat=1486726300_2ef13a74a103e9129b010de922c5d6a5. [Accessed: 10 February 2017].

Jenelius, E. & Mattsson, L.-G. (2012). Road network vulnerability analysis of area-covering disruptions: A grid-based approach with case study. *Transportation Research Part A: Policy and Practice*. 46 (5). p.pp. 746–760.

Jenelius, E., Mattsson, L.-G. & Levinson, D. (2011). Traveler delay costs and value of time with trip chains, flexible activity scheduling and information. *Transportation Research Part B: Methodological*. 45 (5). p.pp. 789–807.

Jenkins, K., Surminski, S., Hall, J. & Crick, F. (2017). Assessing surface water flood risk and management strategies under future climate change: Insights from an Agent-Based Model. *Science of the Total Environment*. 595. p.pp. 159–168.

Jiang, J.R., Huang, H.W., Liao, J.H. & Chen, S.Y. (2014). Extending Dijkstra’s shortest path algorithm for software defined networking. In: *APNOMS 2014 - 16th Asia-Pacific Network Operations and Management Symposium*. 23 December 2014, Institute of Electrical and Electronics Engineers Inc.

Jones, P.M. (2012). *Travel behaviour research in an evolving world: selected papers from the 12th International Conference on Travel Behaviour Research*. 1st Editio. R. . Pendyala & C. . Bhat (eds.). Jaipur: Lulu Publishers.

Jongman, B., Koks, E.E., Husby, T.G. & Ward, P.J. (2014). Increasing flood exposure in the Netherlands: implications for risk financing. *Nat. Hazards Earth Syst. Sci*. 14. p.pp. 1245–1255.

Jurdak, R., Zhao, K., Liu, J., AbouJaoude, M., Cameron, M. & Newth, D. (2015). Understanding

Human Mobility from Twitter Y. Wu (ed.). *PLOS ONE*. 10 (7). p.p. e0131469.

Kasmalkar, I., Serafin, K., Miao, Y., Bick, I.A., Ouang, D., Ortolano, L. & Suckale, J. (n.d.). *When Floods Hit the Road: Resilience to Flood-Induced Commute Disruption in the San Francisco Bay Area and Beyond*.

Kaźmierczak, A. & Cavan, G. (2011). Surface water flooding risk to urban communities: Analysis of vulnerability, hazard and exposure. *Landscape and Urban Planning*. 103 (2). p.pp. 185–197.

Kendon, E.J., Blenkinsop, S., Fowler, H.J., Kendon, E.J., Blenkinsop, S. & Fowler, H.J. (2018). When Will We Detect Changes in Short-Duration Precipitation Extremes? *Journal of Climate*. 31 (7). p.pp. 2945–2964.

Kendon, E.J., Roberts, N.M., Fowler, H.J., Roberts, M.J., Chan, S.C. & Senior, C.A. (2014). Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nature Climate Change*. 4 (7). p.pp. 570–576.

Khademi, N., Balaei, B., Shahri, M., Mirzaei, M., Sarrafi, B., Zahabiun, M. & Mohaymany, A.S. (2015). Transportation network vulnerability analysis for the case of a catastrophic earthquake. *International Journal of Disaster Risk Reduction*. 12. p.pp. 234–254.

Kienberger, S., Blaschke, T. & Zaidi, R.Z. (2013). A framework for spatio-temporal scales and concepts from different disciplines: the 'vulnerability cube'. *Natural Hazards*. 68 (3). p.pp. 1343–1369.

Kilgarriff, P., McDermott, T.K.J., Vega, A., Morrissey, K. & O'Donoghue, C. (2019). The impact of flooding disruption on the spatial distribution of commuter's income. *Journal of Environmental Economics and Policy*. 8 (1). p.pp. 48–64.

Kjeldsen, T.R., Miller, J.D. & Packman, J.C. (2013). Modelling design flood hydrographs in catchments with mixed urban and rural land cover. *Hydrology Research*. 44 (6). p.pp. 1040–1057.

Kjeldsen, T.R., Stewart, E.J., Packman, J.C., Folwell, S. & Bayliss, C., A. (2006). *Revitalisation of the FSR/FEH rainfall-runoff method. Defra R&D Technical Report FD1913/TR*. UK.

Klijn, F., Kreibich, H., De Moel, H. & Penning-Rowsell, E. (2015). Adaptive flood risk management planning based on a comprehensive flood risk conceptualisation. *Mitigation and Adaptation Strategies for Global Change*. 20 (6). p.pp. 845–864.

Koks, E.E., Jongman, B., Husby, T.G. & Botzen, W.J.W. (2015). Combining hazard, exposure and social vulnerability to provide lessons for flood risk management. *Environmental Science and Policy*. 47. p.pp. 42–52.

Kong, F., Yin, H., Nakagoshi, N. & Zong, Y. (2010). Urban green space network development for biodiversity conservation: Identification based on graph theory and gravity modeling. *Landscape and Urban Planning*. 95.

Kreibich, H., Piroth, K., Seifert, I., Maiwald, H., Kunert, U., Schwarz, J., Merz, B. & Thieken, A.H. (2009). Is flow velocity a significant parameter in flood damage modelling? *Natural Hazards and Earth System Sciences*. 9. p.pp. 1679–1692.

Kron, W. (2005). Flood Risk = Hazard . Values . Vulnerability. *Water International*. 30 (1). p.pp. 58–68.

Kubal, C., Haase, D., Meyer, V. & Scheuer, S. (2009). Integrated urban flood risk assessment – adapting a multicriteria approach to a city. *Natural Hazards and Earth System Science*. 9 (6). p.pp. 1881–1895.

Kunwar, B., Simini, F. & Johansson, A. (2014). Large scale pedestrian evacuation modeling framework using volunteered geographical information. *Transportation Research Procedia*. 2. p.pp. 813–818.

Kyriakidis, P.C. & Journel, A.G. (1999). Geostatistical Space–Time Models: A Review. *Mathematical Geology*. 31 (6). p.pp. 651–684.

Lämmel, G., Grether, D. & Nagel, K. (2010). The representation and implementation of time-dependent inundation in large-scale microscopic evacuation simulations. *Transportation Research Part C: Emerging Technologies*. 18 (1). p.pp. 84–98..

Lamond, J.E., Proverbs, D.G. & Hammond, F.N. (2009). Accessibility of Flood Risk Insurance in the UK: Confusion, Competition and Complacency. *Journal of Risk Research*. 12 (6). p.pp. 825–841.

Langford, M. & Higgs, G. (2010). Accessibility and public service provision: evaluating the impacts of the Post Office Network Change Programme in the UK. *Transactions of the Institute of British Geographers*. 35 (4). p.pp. 585–601.

Laversuch, C. (2020). Plan for electronic flood signs to be installed in York | York Press. *The York Press*. 19 February. Available from: https://www.yorkpress.co.uk/news/18244586.plan-electronic-flood-signs-installed-york/https://www.yorkpress.co.uk/news/18244586.plan-electronic-flood-signs-installed-york/. [Accessed: 21 April 2020].

Leandro, J., Chen, A.S., Djordjevic, S. & Savic, D.A. (2009). Simulation, Comparison of 1D/1D and 1D/2D Coupled (Sewer/Surface) Hydraulic Models for Urban Flood. *Journal of Hydraulic Engineering*. 135 (6). p.pp. 495–504.

Leandro, J., Djordjević, S., Chen, A.S., Savić, D.A. & Stanić, M. (2011). Calibration of a 1D/1D urban flood model using 1D/2D model results in the absence of field data. *Water Science and Technology*. 64 (5). p.pp. 1016–1024.

Li, M., Huang, Q., Wang, L., Yin, J. & Wang, J. (2018). Modeling the traffic disruption caused by pluvial flash flood on intra-urban road network. *Transactions in GIS*. 22 (1).

Liu, C., Susilo, Y.O. & Karlström, A. (2014a). Examining the impact of weather variability on non-commuters' daily activity–travel patterns in different regions of Sweden. *Journal of Transport Geography*. 39. p.pp. 36–48.

Liu, J., Wang, S.-Y. & Li, D.-M. (2014b). The Analysis of the Impact of Land-Use Changes on Flood Exposure of Wuhan in Yangtze River Basin, China. *Water Resources Management*. 28 (9). p.pp. 2507–2522.

Liu, S. & Zhu, X. (2004). Accessibility Analyst: An Integrated GIS Tool for Accessibility Analysis in Urban Transportation Planning. *Environment and Planning B: Planning and Design*. 31 (1). p.pp. 105–124.

Local Government Association (2020). *Managing flood risk: roles and responsibilities*. [Online]. 2020. local.gov.uk. Available from: https://www.local.gov.uk/topics/severe-weather/flooding/local-flood-risk-management/managing-flood-risk-roles-and. [Accessed: 23 April 2020].

Locrating (2017). *Locrating*. [Online]. 2017. Available from: https://www.locrating.com/school_catchment_areas.aspx?tab=schools. [Accessed: 4 August 2020].

Longley, P.A., Adnan, M. & Lansley, G. (2015). The Geotemporal Demographics of Twitter Usage. *Environment and Planning A*. 47 (2). p.pp. 465–484.

Lovelace, R. & Ellison, R. (2018). *stplanr: Sustainable Transport Planning*.

Lugeri, N., Kundzewicz, Z.W., Genovese, E., Hochrainer, S. & Radziejewski, M. (2010). River flood risk and adaptation in Europe—assessment of the present status. *Mitigation and Adaptation Strategies for Global Change*. 15 (7). p.pp. 621–639.

Malleson, N. & Andresen, M.A. (2016). Exploring the impact of ambient population measures on London crime hotspots. *Journal of Criminal Justice*. 46. p.pp. 52–63.

Malleson, N. & Andresen, M.A. (2015). Spatio-temporal crime hotspots and the ambient population. *Crime Science*. 4 (1). p.p. 10.

Malleson, N., Vanky, A., Hashemian, B., Santi, P., Verma, S.K., Courtney, T.K. & Ratti, C. (2018). The characteristics of asymmetric pedestrian behavior: A preliminary study using passive smartphone location data. *Transactions in GIS*. 22 (2). p.pp. 616–634.

Martin, D. (1996). An assessment of surface and zonal models of population. *International Journal of Geographical Information Systems*. 10 (8). p.pp. 973–989.

Martin, D. (2017). *Discussion on using SurfaceBuilder247 for spatiotemporal analysis*.

Martin, D. (1989). Mapping population data from zone centroid locations. *Transactions - Institute of British Geographers*. 14 (1). p.pp. 90–97.

Martin, D. (2011). *Surface Builder: User guide*.

Martin, D., Cockings, S. & Harfoot, A. (2013). Development of a geographical framework for census workplace data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 176 (2). p.pp. 585–602.

Martin, D., Cockings, S. & Leung, S. (2015). Developing a Flexible Framework for Spatiotemporal Population Modeling. *Annals of the Association of American Geographers*. 105 (4). p.pp. 754–772.

Martin, D., Cockings, S. & Leung, S. (2009). Population 24/7: building time-specific population grid models . *European Forum for Geostatistics Conference 2009*.

Martin, D., Gale, C., Cockings, S. & Harfoot, A. (2018). Origin-destination geodemographics for analysis of travel to work flows. *Computers, Environment and Urban Systems*. 67. p.pp. 68–79.

McEwen, L., Garde-Hansen, J., Holmes, A., Jones, O. & Krause, F. (2017). Sustainable flood memories, lay knowledges and the development of community resilience to future flood risk. *Transactions of the Institute of British Geographers*. 42 (1). p.pp. 14–28.

McFadden, L., Nicholls, R. & Penning-Rowsell, E. (2006). *Managing coastal vulnerability: an integrated approach*. FHRC Publication.

McNally, M.G. (2007). The Four-Step Model. In: D. A. Hensher & K. J. Button (eds.). *Handbook of Transport Modelling (Handbooks in Transport, Volume 1)*. Emerald, pp. 35–53.

Mechler, R. & Bouwer, L.M. (2014). Understanding trends and projections of disaster losses and climate change: is vulnerability the missing link? *Climatic Change*. 133 (1). p.pp. 23–35.

Mees, P., O'Connell, G. & Stone, J. (2008). Travel to Work in Australian Capital Cities, 1976–2006. *Urban Policy and Research*. 26 (3). p.pp. 363–378.

Melo, N., Santos, B.F. & Leandro, J. (2015). A prototype tool for dynamic pluvial-flood emergency planning. *Urban Water Journal*. 12 (1). p.pp. 79–88.

Mitsakis, E., Stamos, I., Diakakis, M. & Salanova Grau, J.M. (2014). Impacts of high-intensity storms on urban transportation: applying traffic flow control methodologies for quantifying the effects. *International Journal of Environmental Science and Technology*. 11 (8).

Mossoux, S., Kervyn, M. & Canters, F. (2019). Assessing the impact of road segment obstruction on accessibility of critical services in case of a hazard. *Natural Hazards and Earth System Sciences*. 19 (6). p.pp. 1251–1263.

Muriel-Villegas, J.E., Alvarez-Uribe, K.C., Patiño-Rodríguez, C.E. & Villegas, J.G. (2016). Analysis of transportation networks subject to natural hazards - Insights from a Colombian case. *Reliability Engineering and System Safety*. 152.

Nagurney, A. (2011). *Building Resilience into Fragile Transportation Networks in an Era of Increasing Disasters*. Available from: https://supernet.isenberg.umass.edu/visuals/TRB_Panel_Nagurney_Talk.pdf.

National Research Council (2006). *Network Science*. Washington DC: National Academies Press.

Nations Office for Disaster Risk Reduction United Nations (2015). *Sendai Framework for Disaster Risk Reduction 2015 - 2030*. Available from: https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030. [Accessed: 3 August 2020].

Neal, J., Schumann, G., Fewtrell, T., Budimir, M., Bates, P. & Mason, D. (2011). Evaluating a new LISFLOOD-FP formulation with data from the summer 2007 floods in Tewkesbury, UK. *Journal of Flood Risk Management*. 4 (2). p.pp. 88–95.

Néelz, S. & Pender, G. (2013). *Benchmarking the latest generation of 2D hydraulic modelling packages*. Available from: https://consult.environment-agency.gov.uk/engagement/bostonbarriertwao/results/appendix-6---neelz--s.---pender--g.--

2013--benchmarking-the-latest-generation-of-2d-hydraulic-modelling-packages.-bristol_environment-agency.pdf.

Newing, A., Clarke, G. & Clarke, M. (2013). Visitor expenditure estimation for grocery store location planning: a case study of Cornwall. *International Review of Retail, Distribution and Consumer Research*. 23 (3). p.pp. 221–244.

Nieves, J.. (2018). *Spatially Weighted Monte Carlo analysis code*.

Nieves, J.J. (2015). Combining Transportation Network Models with Kernel Density Methods to Measure the Relative Spatial Accessibility of Pediatric Primary Care Services in Jefferson County, Kentucky. *International Journal of Applied Geospatial Research*. 6 (3). p.pp. 34–52.

Noh, S.J., Lee, J.H., Lee, S., Kawaike, K. & Seo, D.J. (2018). Hyper-resolution 1D-2D urban flood modelling using LiDAR data and hybrid parallelization. *Environmental Modelling and Software*. 103. p.pp. 131–145.

Nomis (2020). *Nomis - Official Labour Market Statistics*. 2020. Available from: https://www.nomisweb.co.uk/. [Accessed: 4 August 2020].

NTM; Department for Transport (2009). *NTM Version 4 High Level Overview*. [Online]. Available from: www.dft.gov.uk. [Accessed: 11 August 2020].

Nyberg, R. & Johansson, M. (2013). Indicators of road network vulnerability to storm-felled trees. *Natural Hazards*. 69. p.pp. 185–199.

O'Sullivan, D. & Wong, D.W.S. (2007). A Surface-Based Approach to Measuring Spatial Segregation. *Geographical Analysis*. 39. p.pp. 147–168.

Office for National Statistics (2014). *2011 Census Analysis - Distance Travelled to Work*. 2014. 2011 Census Analysis - Distance Travelled to Work.

Office for National Statistics (2013). *Beyond 2011: Admininstrative Data Sources Report: The English School Census and the Welsh School Census*. Available from: https://www.ons.gov.uk/census/censustransformationprogramme/beyond2011censustransformationprogramme/reportsandpublications.

Office for National Statistics (2016). *Census geography - Office for National Statistics*. 2016. Available from: https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeography. [Accessed: 13 August 2020].

Office for National Statistics (n.d.). *Introduction to Output Areas - the building block of Census geography*. Output areas.

Olsen, A., Zhou, Q., Linde, J. & Arnbjerg-Nielsen, K. (2015). Comparing Methods of Calculating Expected Annual Damage in Urban Pluvial Flood Risk Assessments. *Water*. 7 (12). p.pp. 255–270.

Openshaw, S. (1984). *The Modifiable Areal Unit Problem*. Norwich: Geo Books.

Openshaw, S. & Taylor, P.J. (1979). A Million or so Correlation Coefficients: Three Experiments on the Modifiable Areal Unit Problem. In: N. Wrigley (ed.). *Statistical Applications in Spatial Sciences*. London: Pion, pp. 127–144.

Ortuzar, J. de D. & Willumsen, L.G. (2011). *Modelling Transport*. 4th Editio. Wiley. Available from: http://lib.myilibrary.com/Open.aspx?id=340548. [Accessed: 8 January 2018].

Papathoma-Kohle, M., Kappes, M., Keiler, M. & Glade, T. (2011). Physical vulnerability assessment for alpine hazards: state of the art and future needs. *Natural Hazards*. 58. p.pp. 645–680.

Pebesma, E., Roger Bivand, Achim Zeileis, Michael Sumner & Ping Yang. (2020). *CRAN Task View: Handling and Analyzing Spatio-Temporal Data*. 2020.

Penning-Rowsell, E.C., Johnson, C.L., Tunstall, S.M., Tapsell, S.M., Morris, J., Chatterton, J.B. & Green, C. (2005). *The Benefits of Flood and Coastal Risk Management: A Handbook of Assessment Techniques*.

Penning-Rowsell, E.C., Priest, S. & Johnson, C. (2014). The evolution of UK flood insurance: incremental change over six decades. *International Journal of Water Resources Development*. 30 (4). p.pp. 694–713.

Pfurtscheller, C. & Genovese, E. (2019). The Felbertauern landslide of 2013 in Austria: Impact on transport networks, regional economy and policy decisions. *Case Studies on Transport Policy*. 7 (3). p.pp. 643–654.

Pitt, M. (2008). *The Pitt Review: Learning Lessons from the 2007 Floods* Cabinet Office (ed.).

Platt, R.H. (1995). Lifelines: An Emergency Management Priority for the United States in the 1990s. *Disasters*. 15. p.pp. 172–176.

Postance, B., Hillier, J., Dijkstra, T. & Dixon, N. (2017). Extending natural hazard impacts: an assessment of landslide disruptions on a national road transportation network. *Environmental Research Letters*. 12 (1). p.p. 014010.

Pred, A. (1977). The Choreography of Existence: Comments on Hagerstrand's Time-Geography and Its Usefulness. *Economic Geography*. 53 (2). p.p. 207.

Pregnolato, M., Ford, A. & Dawson, R. (2015). Analysis of the risk of transport infrastructure disruption from extreme rainfall. In: University of British Colombia (ed.). *12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP12*. 1 January 2015, Vancouver: University of British Colombia, pp. 1–8.

Pregnolato, M., Ford, A. & Dawson, R. (2016a). Disruption and adaptation of urban transport networks from flooding. In: *FLOODrisk 2016 - 3rd European Conference on Flood Risk Management*. 2016.

Pregnolato, M., Ford, A., Glenis, V., Wilkinson, S. & Dawson, R. (2017a). Impact of Climate Change on Disruption to Urban Transport Networks from Pluvial Flooding. *Journal of Infrastructure Systems*. 23 (4). p.p. 04017015.

Pregnolato, M., Ford, A., Robson, C., Glenis, V., Barr, S. & Dawson, R. (2016b). Assessing urban strategies for reducing the impacts of extreme weather on infrastructure networks. *Royal Society Open Science*. 3 (160023).

Pregnolato, M., Ford, A., Wilkinson, S.M. & Dawson, R.J. (2017b). The impact of flooding on road transport: A depth-disruption function. *Transportation Research Part D: Transport and Environment*. 55. p.pp. 67–81.

Priestley, S. & Rutherford, T. (2016). *Flood Defence Spending in England*. [Online]. London. Available from: http://researchbriefings.parliament.uk/ResearchBriefing/Summary/CBP-7514#fullreport.

QGIS (n.d.). *QGIS Python Plugins Repository*. [Online]. Available from: https://plugins.qgis.org/plugins/. [Accessed: 13 August 2020].

Rajib, A., Liu, Z., Merwade, V., Tavakoly, A.A. & Follum, M.L. (2020). Towards a large-scale locally relevant flood inundation modeling framework using SWAT and LISFLOOD-FP. *Journal of Hydrology*. 581. p.p. 124406.

Rawlings, K. (2016). *Flood Defence Performance Assessment for Egger UK's Factory in Hexham, Northumberland, for Current and Future Flood Events*. Durham University.

Reggiani, A., Nijkamp, P. & Lanzi, D. (2015). Transport resilience and vulnerability: The role of connectivity. *Transportation Research Part A*. 81. p.pp. 4–15.

Renner, K., Schneiderbauer, S., Pruß, F., Kofler, C., Martin, D. & Cockings, S. (2018). Spatio-temporal population modelling as improved exposure information for risk assessments tested in the Autonomous Province of Bolzano. *International Journal of Disaster Risk Reduction*. 27. p.pp. 470–479.

Rodríguez-Núñez, Eduardo García-Palomares & Carlos, J. (2014). Measuring the vulnerability of public transport networks. *Journal of Transport Geography*. 35. p.pp. 50–63.

Rogelis, M.C. (2015). *Flood Risk in Road Networks*. p.pp. 1–24.

Romieu, E., Welle, T., Schneiderbauer, S., Pelling, M. & Vinchon, C. (2010). Vulnerability assessment within climate change and natural hazard contexts: Revealing gaps and synergies through coastal applications. *Sustainability Science*. 5 (2). p.pp. 159–170.

Rowley, D., Buckler, M. & Barnett, J. (2016). When floods strike close to home. *Proceedings of the Institution of Civil Engineers - Municipal Engineer*. 169 (2). p.pp. 118–123.

Ruin, I., Creutin, J.D., Anquetin, S. & Lutoff, C. (2008). Human exposure to flash floods - Relation between flood parameters and human vulnerability during a storm of September 2002 in Southern France. *Journal of Hydrology*. 361 (1–2). p.pp. 199–213.

Ruktanonchai, C.W., Ruktanonchai, N.W., Nove, A., Lopes, S., Pezzulo, C., Bosco, C., Alegana, V.A., Burgert, C.R., Ayiko, R., Charles, A.S., Lambert, N., Msechu, E., Kathini, E., Matthews, Z. & Tatem, A.J. (2016). Equality in Maternal and Newborn Health: Modelling Geographic Disparities in Utilisation of Care in Five East African Countries M. Ali (ed.). *PLOS ONE*. 11 (8). p.p. e0162006.

Rupi, F., Angelini, S., Bernardi, S., Danesi, A. & Rossi, G. (2015). Ranking Links in a Road Transport Network: A Practical Method for the Calculation of Link Importance. *Transportation Research Procedia*. 5. p.pp. 221–232.

Salonen, M., Toivonen, T., Cohalan, J.-M. & Coomes, O.T. (2012). Critical distances: Comparing measures of spatial accessibility in the riverine landscapes of Peruvian Amazonia. *Applied Geography*. 32 (2). p.pp. 501–513.

Schnebele, E., Cervone, G. & Waters, N. (2014). Road assessment after flood events using non-authoritative data. *Natural Hazards and Earth System Sciences*. 14. p.pp. 1007–1015.

Shabou, S., Ruin, I., Lutoff, C., Debionne, S., Anquetin, S., Creutin, J.D. & Beaufils, X. (2017). MobRISK: A model for assessing the exposure of road users to flash flood events. *Natural Hazards and Earth System Sciences*. 17 (9).

Simini, F., González, M.C., Maritan, A. & Barabási, A.-L. (2012). A universal model for mobility and migration patterns. *Nature*. 484 (7392). p.pp. 96–100.

Singleton, A.D., Longley, P.A., Allen, R. & O'Brien, O. (2011). Estimating secondary school catchment areas and the spatial equity of access. *Computers, Environment and Urban Systems*. 35 (3). p.pp. 241–249.

Slovic, P., Finucane, M.., Peters, E. & MacGregor, D.. (2004). Risk as analysis and risk as feelings: some thoughts about affect, reason, risk, and rationality. *Risk analysis*. 24 (2). p.pp. 311–322.

Smallwood, S. & Lynch, K. (2010). An analysis of patient register data in the Longitudinal Study - what does it tell us about the quality of the data? *Population Trends*. 141 (1). p.pp. 151–169.

Smith, A., Geoff, C. & Panzeri, M. (2014a). *Enhancements in reservoir flood risk mapping: example application for Ulley*.

Smith, A., Martin, D. & Cockings, S. (2014b). Spatio-Temporal Population Modelling for Enhanced Assessment of Urban Exposure to Flood Risk. *Applied Spatial Analysis and Policy*. 9 (2). p.pp. 145–163.

Smith, A., Newing, A., Quinn, N., Martin, D., Cockings, S. & Neal, J. (2015). Assessing the Impact of Seasonal Population Fluctuation on Regional Flood Risk Management. *ISPRS International Journal of Geo-Information*. 4 (3). p.pp. 1118–1141.

Smith, A.D. (2015). *Spatiotemporal Population Modelling to Assess Exposure to Flood Risk*. University of southampton.

Smith, K. (2013). *Environmental hazards : assessing risk and reducing risaster*. Routledge.

Sofianopoulou, E., Rushton, S., Rubin, G. & Pless-Mulloli, T. (2012). Defining GP practice areas based on true service utilisation. *Health & Place*. 18 (6). p.pp. 1248–1254.

Sohn, J. (2006). Evaluating the significance of highway network links under the flood damage: An accessibility approach. *Transportation Research Part A: Policy and Practice*. 40 (6). p.pp. 491–506.

Sosa, J., Sampson, C., Smith, A., Neal, J. & Bates, P. (2020). A toolbox to quickly prepare flood inundation models for LISFLOOD-FP simulations. *Environmental Modelling and Software*. 123. p.p. 104561.

Sperotto, A., Torresan, S., Gallina, V., Coppola, E., Critto, A. & Marcomini, A. (2015). A multidisciplinary approach to evaluate pluvial floods risk under changing climate: The case study of the municipality of Venice (Italy). *Science of the Total Environment*. 562. p.pp. 1031–1043.

Spitalar, M., Gourley, J.J., Lutoff, C., Kirstetter, P.E., Brilly, M. & Carr, N. (2014). Analysis of flash flood parameters and human impacts in the US from 2006 to 2012. *Journal of Hydrology*. 519 (PA). p.pp. 863–870.

Stepanov, A. & Smith, J.M. (2009). Multi-objective evacuation routing in transportation networks. *European Journal of Operational Research*. 198 (2). p.pp. 435–446.

Stevens, A.J., Clarke, D. & Nicholls, R.J. (2016). Trends in reported flooding in the UK: 1884–2013. *Hydrological Sciences Journal*. 61 (1). p.pp. 50–63.

Stevens, F.R., Gaughan, A.E., Linard, C. & Tatem, A.J. (2015). Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data L. A. N. Amaral (ed.). *PLOS ONE*. 10 (2). p.p. e0107042.

Strano, E., Viana, M.P., Sorichetta, A. & Tatem, A.J. (2018). Mapping road network communities for guiding disease surveillance and control strategies. *Scientific Reports*. 8 (1). p.p. 4744.

Tapp, A.F. (2010). Areal Interpolation and Dasymetric Mapping Methods Using Local Ancillary Data Sources. *Cartography and Geographic Information Science*. 37 (3). p.pp. 215–228.

Tatem, A.J. (2017). WorldPop, open data for spatial demography. *Scientific Data*. 4. p.p. 170004.

Tatem, A.J., Huang, Z., Narib, C., Kumar, U., Kandula, D., Pindolia, D.K., Smith, D.L., Cohen, J.M., Graupe, B., Uusiku, P. & Lourenco, C. (2014). Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning . *Malaria Journal*. 13 (52).

Tatem, A.J. & Smith, D.L. (2010). International population movements and regional Plasmodium falciparum malaria elimination strategies. *Proceedings of the National Academy of Sciences of the United States of America*. 107 (27). p.pp. 12222–7.

Taylor, M.A.P. & D’Este, G.M. (2003). Concepts of network vulnerability and applications to the identification of critical elements of transport infrastructure . *Concepts of network vulnerability and applications to the identification of critical elements of transport infrastructure* .

Taylor, M.A.P., Sekhar, S.V.C. & D’Este, G.M. (2006). Application of Accessibility Based Methods for Vulnerability Analysis of Strategic Road Networks. *Networks and Spatial Economics*. 6 (3–4). p.pp. 267–291.

Terti, G., Ruin, I., Anquetin, S. & Gourley, J.J. (2017). A SITUATION-BASED ANALYSIS OF FLASH FLOOD FATALITIES IN THE UNITED STATES. *American Meteorological Society*. p.pp. 333–345.

Terti, G., Ruin, I., Anquetin, S. & Gourley, J.J. (2015). Dynamic vulnerability factors for impact-based flash flood prediction. *Natural Hazards*. 79 (3). p.pp. 1481–1497.

The Heidelberg Institute for Geoinformation Technology (2019). *Openrouteservice*. 2019. Available from: https://openrouteservice.org/. [Accessed: 4 August 2020].

Thiessen, P. (n.d.). *The National Transport Model: Predictions on Trends in Road Transport*. [Online]. Department for Transport. Available from: http://www.iapsc.org.uk/assets/document/0608_P_Thiessen.pdf.

Tinsley, B. (2017). *Beyond 2021 – transforming the Census | National Statistical*. 2017. National Statistical . Available from: https://blog.ons.gov.uk/2017/06/20/an-administrative-data-census-transforming-the-census-for-the-future/. [Accessed: 8 June 2020].

Toma-Danila, D. (2018). A GIS framework for evaluating the implications of urban road network failure due to earthquakes: Bucharest (Romania) case study. *Natural Hazards*. 93 (1). p.pp. 97–111.

Tsekeris, T. & Stathopoulos, A. (2006). Gravity models for dynamic transport planning: Development and implementation in urban networks. *Journal of Transport Geography*. 14 (2). p.pp. 152–160.

Tympakianaki, A., Koutsopoulos, H.N., Jenelius, E. & Cebecauer, M. (2018). Impact analysis of transport network disruptions using multimodal data: A case study for tunnel closures in Stockholm. *Case Studies on Transport Policy*. 6 (2).

UK Centre for Ecology & Hydrology (2020). *FEH Web Service*. [Online]. 2020. Available from: https://fehweb.ceh.ac.uk/. [Accessed: 4 August 2020].

UK Data Service (2020). *UK Data Service*. 2020. Available from: https://www.ukdataservice.ac.uk/. [Accessed: 11 August 2020].

UK Research and Innovation (n.d.). *Population247NRT: Near real-time spatiotemporal population estimates for health, emergency response and national security*. Available from: https://gtr.ukri.org/projects?ref=ES%2FP010768%2F1. [Accessed: 4 August 2020].

University of Southampton (2011). *Research project: Population24/7: Space–Time Specific Population Surface Modelling - Dormant*. [Online]. 2011. Available from:

https://www.southampton.ac.uk/geography/research/projects/space_time.page? [Accessed: 4 August 2020].

USAID (2020). *The DHS Program* . [Online]. 2020. Available from: https://dhsprogram.com/. [Accessed: 4 August 2020].

Versini, P., Gaume, E. & Andrieu, H. (2010). Application of a distributed hydrological model to the design of a road inun-dation warning system for flash flood prone areas. *Natural Hazards and Earth System Sciences*. 10. p.pp. 805–817.

Vickerman, R.W. (1974). Accessibility, Attraction, and Potential: A Review of Some Concepts and Their Use in Determining Mobility. *Environment and Planning A: Economy and Space*. 6 (6). p.pp. 675–691.

Wang, D. & Cheng, T. (2001). A spatio-temporal data model for activity-based transport demand modelling. *int. j. geographical information science*. 15 (6). p.pp. 561–585.

Wang, S.X. (2012). The improved Dijkstra's shortest path algorithm and its application. In: *Procedia Engineering*. 1 January 2012, Elsevier, pp. 1186–1190.

Wang, Y., Correia, G.H. de A., de Romph, E. & Timmermans, H.J.P. (2017). Using metro smart card data to model location choice of after-work activities: An application to Shanghai. *Journal of Transport Geography*. 63. p.pp. 40–47.

Wardrop, N.A., Jochem, W.C., Bird, T.J., Chamberlain, H.R., Clarke, D., Kerr, D., Bengtsson, L., Juran, S., Seaman, V. & Tatem, A.J. (2018). Spatially disaggregated population estimates in the absence of national population and housing census data. *Proceedings of the National Academy of Sciences of the United States of America*. 115 (14). p.pp. 3529–3537.

Weber, E.M., Seaman, V.Y., Stewart, R.N., Bird, T.J., Tatem, A.J., McKee, J.J., Bhaduri, B.L., Moehl, J.J. & Reith, A.E. (2018). Census-independent population mapping in northern Nigeria. *Remote Sensing of Environment*. 204.

Wei, F., Koc, E., Soibelman, L. & Li, N. (2018). Disaster economics and networked transportation infrastructures: Status quo and a multi-disciplinary framework to estimate economic losses. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2018, Springer Verlag, pp. 3–22.

Wilby, R.L. & Quinn, N.W. (2013). Reconstructing multi-decadal variations in fluvial flood risk using atmospheric circulation patterns. *Journal of Hydrology*. 487. p.pp. 109–121.

Wisner, B., Blaikie, P., Cannon, T. & Davis, I. (2004). Key risk theory textbook. *At Risk*. 2nd Ed. Abingdon: Routledge.

Wood, N.J. & Schmidtlein, M.C. (2013). Community variations in population exposure to near-field tsunami hazards as a function of pedestrian travel time to safety. *Natural Hazards*. 65 (3). p.pp. 1603–1628.

Woodward, M., Kapelan, Z. & Gouldby, B. (2014). Adaptive Flood Risk Management Under Climate Change Uncertainty Using Real Options and Optimization. *Risk Analysis*. 34 (1). p.pp. 75–92.

World Pop (2020). *WorldPop*. [Online]. 2020. Available from: https://www.worldpop.org/. [Accessed: 4 August 2020].

Wright, J.K. (1936). A Method of Mapping Densities of Population: With Cape Cod as an Example. *Geographical Review*. 26 (1). p.p. 103.

Wu, S., Qiu, X. & Wang, L. (2005). Population Estimation Methods in GIS and Remote Sensing: A Review. *GIScience & Remote Sensing*. 42 (1). p.pp. 80–96.

Yin, J., Yu, D., Yin, Z., Liu, M. & He, Q. (2016). Evaluating the impact and risk of pluvial flash flood on intra-urban road network: A case study in the city center of Shanghai, China. *Journal of Hydrology*. 537. p.pp. 138–145.

Zhou, Q., Mikkelsen, P.S., Halsnaes, K. & Arnbjerg-Nielsen, K. (2012). Framework for economic pluvial flood risk assessment considering climate change effects and adaptation benefits. *Journal of Hydrology*. 414–415. p.pp. 539–549.