UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL AND HUMAN SCIENCES

Geography and Environment

Spatiotemporal population modelling to assess exposure to flood risk

by

Alan D. Smith

Thesis for the degree of Doctor of Philosophy

March 2015

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL AND HUMAN SCIENCES

Geography and Environment

Thesis for the degree of Doctor of Philosophy

SPATIOTEMPORAL POPULATION MODELLING TO ASSESS EXPOSURE TO FLOOD RISK

Alan Daniel Smith

There is a growing need for high resolution spatiotemporal population estimates which allow accurate assessment of population exposure to natural hazards. Populations vary over range of time scales and cyclical patterns. This has important implications for how researchers and policy makers undertake hazard risk assessments. Traditionally, static population counts aggregated to arbitrary areal units have been used. This thesis shows that these are inadequate for the purposes of hazard risk assessments concerning dynamic populations. This thesis enhances and applies spatiotemporal modelling techniques developed through the Population 24/7 project and integrates the outputs with hydrological models using a loose-coupling approach. This is demonstrated through two case studies to illustrate the impacts for flood risk assessment. These case studies exemplify population fluctuations according to diurnal, weekly and seasonal cycles. The considerably enhanced spatiotemporal population model constructed demonstrates a much wider flexible framework. This thesis establishes that there is a strong requirement to consider time-specific populations for the purposes of flood hazard risk analyses. Population movements have been shown to account for major variations in exposure estimates through the analysis of a range of flood scenarios and population case studies. Significant enhancements can be sought for risk analyses by using spatiotemporal population estimates.

Contents

List of Tables

List of Figures

Declaration of Authorship

I, Alan Smith,

declare that this thesis entitled "*Spatiotemporal population modelling to assess exposure to flood risk*" 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:

Smith, A.D., Martin, D. and Cockings, S. (2014 online first) Spatiotemporal Population Modelling for Enhanced Assessment of Urban Exposure to Flood Risk. *Applied Spatial Analysis and Policy*. (http://dx.doi.org/10.1007/s12061-014-9110-6)

Smith, A. D., Goff, C. A. and Panzeri, M. (2014). Enhancements in reservoir flood risk mapping: example application for Ulley. In A. Pepper (Ed.), *Maintaining the Safety of our Dams and Reservoirs* (pp. 295-306). ICE Publishing, London (ISBN: 9780727760340)

Signed:...

Date: 23 March 2015

Acknowledgements

Firstly, I would like to sincerely thank my academic supervisors Professor David Martin and Dr Samantha Cockings for their unwavering guidance, invaluable support and constant enthusiasm over the past three years. The development of this thesis is testament to their dedicated assistance, intellectual and personal encouragement.

This research was funded by an Economic and Social Research Council (ESRC) Advanced Quantitative Methods Doctoral Training Centre award at the University of Southampton. A further knowledge transfer internship with HR Wallingford Ltd. was funded as part of the Engineering and Physical Sciences Research Council (EPSRC) Impact Acceleration Account through the University of Southampton. Without these the research presented here could not have been facilitated.

I would like to thank all who have provided data, advice and feedback. I am particularly grateful to Dr Andy Newing, Dr Niall Quinn and the warm welcome and support received from staff at HR Wallingford. A special thank you is due to Robin Wilson for his technical advice and supply of python scripts to aid some of data analyses undertaken.

I would like to thank Sophie Madge for her patience and understanding during this process, as well as for proof-reading drafts. Also thank you to my parents whose support has got me this far.

Finally, I would like to thank everyone, including those mentioned above, who I have worked with within Geography and Environment, the Southampton Graduate School and further afield for their intellectual and personal support. Some of us have embarked on numerous intrepid adventures into the Hampshire countryside and beyond, providing some great memories of this process.

Data acknowledgments: the support of the following data providers is gratefully acknowledged and has been instrumental in the production of this thesis. Cornwall seasonal visitor population estimates, Dr Andy Newing, University of Leeds; Geographical Modelling and Planning (GMAP) retail centres, Callcredit Information Group Ltd., Leeds; Seasonal historic property visitor data, English Heritage; Ulley technical drawings, Rotherham Metropolitan Borough Council; Flood map and 2 m LiDAR, Environment Agency; LISFLOOD-FP layers, Dr Niall Quinn, University of Bristol. Annual Business Inquiry/Business Register and Employment Survey (2007&10), Chancellor of the Exchequer's Notice number: NTC/BRES12-P0212; Schools census (2007&10), Department for Education. Ordnance Survey data is Crown copyright and an EDINA/Digimap supplied service. Census output is Crown copyright and is reproduced with the permission of the Controller of HMSO. Contains public sector information licensed under the Open Government Licence v3.0.

Definitions and Abbreviations

σ	Standard Deviation
Δ	Difference
£	British pounds
$	United States dollars (unless specified)
A&E	Accident and Emergency
AADF	Average Annual Daily Flow
ABI	Annual Business Inquiry (see Glossary)
ALARP	As Low As Reasonably Practicable
AOD	Above Ordnance Datum
AV	Area Vulnerability
ASCII	American Standard Code for Information Interchange
AZP	Automated Zoning Procedure
BNG	British National Grid
BRES	Business Register and Employment Survey (see ABI in Glossary)
CASA	Centre for Advanced Spatial Analysis
CBD	Central Business District
CFMP	Catchment Flood Management Plan
CP	Count Point
CRU	Census Research Unit
DEFRA	Department for Environment, Food and Rural Affairs
DfT	Department for Transport
DRI	Disaster Risk Index
DSM	Digital Surface Model
DSSs	Decision Support Systems
DTM	Digital Terrain Model
E&W	England and Wales
EA	Environment Agency (of England and Wales)
ED	Enumeration District
EMBREA	Embankment Breach Assessment Tool
ESDS	Economic and Social Data Service
EU	European Union
FFC	Flood Forecasting Centre
GB	Great Britain (see Glossary)

GIS	Geographic Information System (see Glossary)
GPS	Global Positioning System
GROS	General Register Office Scotland
HES	Hospital Episode Statistics
HM	Her Majesty's
HMSO	Her Majesty's Stationery Office
HSCIC	Health and Social Care Information Centre
IDW	Inverse Distance Weighting
KDE	Kernel Density Estimation
LCF	Living Costs and Food Survey
LFS	Labour Force Survey
LiDAR	Light Detection and Ranging
LLFAs	Lead Local Flood Authorities
LRF	Local Resilience Forums
LSOA	Lower Layer Super Output Area (see Glossary)
MAUP	Modifiable Areal Unit Problem (see Glossary)
MBC	Metropolitan Borough Council
MIU	Minor Injury Unit
MPH	Miles per hour
MSC	Marine Conservation Society
MSTUP	Modifiable Spatiotemporal Unit Problem (see Glossary)
MYE	Mid-year population estimate (see Glossary)
NHP	Natural Hazards Partnership
NPD	National Population Database
NSPD	National Statistics Postcode Directory
NTS	National Travel Survey, Department for Transport, UK.
NZ	New Zealand
OA	Output Area (see Glossary)
ONS	Office for National Statistics
OPCS	Office of Population and Census Surveys
OS	Ordnance Survey of Great Britain
PCT	Primary Care Trust
POI	Point of Interest
PV	People Vulnerability
RE	Recommendation

RMSE	Root Mean Square Error
SEPA	Scottish Environment Protection Agency
SIC	Standard Industrial Classification (see Glossary)
SIMD	Scottish Index of Multiple Deprivation
SMP	Shoreline Management Plan
SNIFFER	Scotland and Northern Ireland Forum for Environmental Research
Sq. Ft.	Square Feet
St	Saint
TUS	Time Use Survey
UK	United Kingdom of Great Britain and Northern Ireland
UKCP09	United Kingdom Climate Projections 2009
UN	United Nations
UNDRO	United Nations Disaster Relief Organisation
UNISDR	United Nations International Strategy for Disaster Reduction
UPC	Unit Postcode (see Glossary)
US	United States of America
WAD	Wide Area Dispersion
WIMD	Welsh Index of Multiple Deprivation

Chapter 1: Introduction

1.1 Population exposure to natural hazards

Natural hazards continue to impose one of the ultimate constraints on human activities (Mitchell 1999b). Humans remain subject to the physical processes of the planet we inhabit. Understanding the human occupancy of hazardous zones, how societies respond to environmental hazards and how to mitigate the risks posed have long been questions for researchers of hazards (Cutter 1996). Natural hazards can occur at any time and over a range of timescales with differing opportunities for predictions or warning. These can be at the timescales of drought or global warming through to earthquakes, volcanoes or flash flooding. Humans can become susceptible to hazards for a variety of reasons such as societal attitudes (Slovic *et al.* 1977), economic constraints (Dobran 2003), public policy (Torrieri *et al.* 2002), collective memory (Slovic 2000) and population growth (Mitchell 1999b).

In the past unforeseen catastrophic natural disasters have destroyed whole towns (e.g. Pompeii in AD 79, Giacomelli *et al.* 2003) or caused the collapse of entire civilisations (e.g. Minoan civilisation, c. 1,500 BC, Antonopoulos 1992). In our modern and increasingly globalised world natural hazards still pose great threats with increasingly global impacts. For example, the temporary impacts of the 2011 Japanese tsunami and resultant nuclear alert on global financial markets (IHS 2011), and the 2010 Icelandic eruption of Eyjafjallajökull that caused an unprecedented shutdown in international air travel (Budd *et al.* 2011). The modern impacts of hazards may differ somewhat from those experienced by our early European Minoan predecessors, but they still provide a reminder that we do not live in a world entirely of our own making and control.

Prominent events since 2011 have refocused attention on natural hazards and human vulnerability both in terms of fatalities and economic impacts. By a considerable margin, 2011 was the mostly costly year to date with natural hazard losses totalling £240 billion and 27,000 fatalities (Munich Re 2012). Flooding continued to dominate global headlines in 2013 with Super Typhoon Haiyan in the Philippines and severe flooding in central Europe. Flooding in Germany, Austria and the Czech Republic saw some areas experience more than 400 mm of rainfall over only several days (Munich Re 2014).

This thesis aims to assess population exposure to flooding, as one exemplar hazard, through detailed local scale analysis of population movements and flood risks in the United Kingdom (UK). Flooding too can occur at a range of scales from rapid onset flash floods to longer-term coastal flood plain inundation. In the UK in February 2014 an exceptional succession of winter storms caused serious coastal damage and widespread persistent flooding (Met Office 2014). It is estimated that 5.2 million properties are at risk from flooding in the UK (DEFRA 2011). Humans are found in flood risk zones for a variety of complex reasons: normally resident as ‘inhabitants’, temporary occupants at places of work, study or leisure, or simply in transit through the region.

In order to better understand the risks posed to humans by hazard events, such as flooding, an improved knowledge of the spatial and temporal distribution of population is required (Bhaduri *et al.* 2007; Fielding 2007; Harper and Mayhew 2012; Aubrecht *et al.* 2012a). Calculating the population exposed is not straightforward as both the hazard and population vary over time (McPherson and Brown 2004). Inadequacies in mapping population have been noted for many decades (e.g. Schmitt 1956). Commonly used official population datasets such as censuses or population registers usually provide only residential ‘night-time’ population counts. Better representations of population distributions that are time-specific are required for improved risk assessment and the development of effective emergency plans.

This thesis implements and further develops a spatiotemporal population modelling approach, known as “Population 24/7” (Martin *et al.* (forthcoming); http://www.esrc.ac.uk/my-esrc/grants/RES-062-23-1811/read) using SurfaceBuilder247 (v. 1.0) software. Most importantly, it integrates the results with environmental models for the first time. The Population 24/7 approach uses a flexible data framework to model the distribution of population in time and space producing a variable gridded output. A detailed model data library containing residential census populations, workplaces, hospitals, educational establishments, leisure locations and retail centres represented as population centroids (georeferenced locations with an associated population count) is created, each with relevant time profile information. Population is then redistributed around these centroids producing bespoke population distributions for each target time.

This thesis demonstrates significant enhancements to the Population 24/7 data library and improvements to the spatial and temporal resolution in population outputs. It integrates spatiotemporal population data with environmental models using a loose-coupling approach (Martin 2009). These previously standalone models are enhanced and their outputs integrated into a compatible format for analysis using GIS.

Three separate population data libraries have been constructed for implementation within the Population 24/7 framework based on three different UK case studies. These case studies are intended to demonstrate how enhanced spatiotemporal population estimates can be used to improve flood risk assessments. The applications include a worked example centred on Southampton to illustrate the methodological framework (Chapter 3) and two substantive case studies. The first of these is centred on Ulley, South Yorkshire (Chapter 4) and evaluates the risk posed to the potentially exposed population following a simulated collapse of the Ulley reservoir embankment, based on a near-catastrophic dam failure that occurred in June 2007. This case study examines the implications of diurnal population cycles for flood risk assessment. The second case study is focused on St Austell, Cornwall (Chapter 5), part of an English coastal county and a popular tourist destination. This case study assesses the impact of daily and seasonal population fluctuations on exposure to a range of fluvial flood risks.

Both of the case studies, Ulley (target year 2007) and St Austell (target year 2010), have reference dates that fall between the 2001 and 2011 censuses in the UK. These dates are either based on the actual event reconstructed (e.g. Ulley dam failure in June 2007) or defined by the availability of required datasets. Both applications employ the 2001 census geography, but with the population data updated or adjusted for the respective target dates. The latest 2011 census data not only post-dates the target dates for these case studies but also had not been fully published at the time this research was undertaken.

1.2 Aims and objectives

The aims of this research are to generate innovative spatiotemporal population estimates by implementing and further enhancing the existing Population 24/7 modelling framework. It is intended that this will demonstrate for the first time the applicability of such models within flood risk assessment. Three overall aims are identified:

Aim 1 Review the existing literature regarding the assessment of population risk to natural hazards and methods for determining exposure.

Aim 2 Examine how population exposure fluctuates spatiotemporally to flood hazard events using census, administrative and survey datasets to construct a spatiotemporal population model applied to illustrative cases studies.

Aim 3 Assess the extent to which spatiotemporal population modelling techniques can be used to provide greater insights for integrated disaster risk management. Discuss to what extent confidence can be placed in their results and outline the challenges for validation.

1.2.1 **Specific objectives**

The following objectives have been formulated to achieve the aims outlined:

I. Review the relevant literature on the risk posed by natural hazard events and over what timescales they occur. Determine to what extent population and hazard data have previously been integrated for successful disaster risk management (Aim 1).

II. Construct a population data library for implementation with SurfaceBuilder247 to examine spatiotemporal cycles within population movements for enhanced spatiotemporal population estimates, in terms of spatial and temporal resolution, with currently available datasets (Aim 2).

III. Examine how population fluctuates in daily and seasonal cycles and demonstrate how this can be incorporated though implementation of the Population 24/7 modelling framework (Aim 2).
IV. Demonstrate and attempt to quantify the difference in spatiotemporal and static population estimates though two case studies to assess hazard exposure integrated with environmental model outputs (Aim 2).
V. Critically evaluate the applicability of this approach for application within natural hazard risk management (Aim 3).

1.3 Thesis structure

This chapter introduces the thesis theme, aims and objectives and structure. It outlines the main techniques that will be used.

Chapter two presents a selected review of the relevant literature within the field of disaster risk management, interpolation of population data and the considerations for spatiotemporality in both natural hazard and dynamic population applications. It contextualises the theme of this thesis within the inter-related fields of hazard risk management, spatial analysis of population data and interpolation techniques. The aim of the chapter is to underpin the arguments for the original contribution of the doctoral research undertaken. It identifies a gap in current knowledge which this thesis fills. The chapter is divided into five main sections. Firstly an overview is provided of human exposure, risk and vulnerability to natural hazard events. Secondly, a more detailed review of specific flood risks within the UK and current international policy agendas is presented. Thirdly, methods and techniques are reviewed for the spatial interpolation of aggregate population data. Fourthly, examples of, and methods for, the interpolation of population data with time are considered. Finally, the Population 24/7 tool, that will be implemented, and concept is introduced.

Chapter three describes the methods and data utilised in this thesis. A version of the chapter has been published as Smith *et al.* (2014a) and is included in this thesis as Appendix A. Firstly, the structure and concepts of the spatiotemporal population tool, Population 24/7, are described. This introduces the components and parameters required to construct a population data library for the model. A worked example is then presented to demonstrate the creation and implementation of a population data library based on Southampton, UK. Example outputs are provided and the implications for population exposure in time and space to coastal and fluvial flooding within the city are explored.

Chapter four presents the first of two case studies detailing application of the Population 24/7 tool and integration with hydraulic flood modelling. A condensed version of this chapter has been published as Smith *et al.* (2014b). The additional hydrological modelling described in this chapter was undertaken by the author facilitated through a knowledge transfer internship at HR Wallingford Ltd. The case study concerns the failure of a dam at Ulley, South Yorkshire, UK that actually occurred during 2007 and the risk to people that it posed.

Chapter five presents the second case study centred on St Austell, Cornwall, UK. The purpose of this case study is to examine how tourism-driven seasonal fluctuations in population affect flood risk assessment. This application employs seasonal tourism estimates provided by Newing (2014) to enhance a population data library used to produce spatiotemporal population estimates that represent daily, weekly and seasonal cycles. These outputs were combined with the UK's national flood risk map and fluvial flood risk scenarios. Bespoke fluvial flood risks were provided by Quinn (2014) created using LISFLOOD-FP, a flood inundation model. The outputs were integrated and analysed, with the effects of population fluctuations on exposure to flood risks examined.

Chapter six discusses the findings of the thesis, and the methodological approach more generally. The discussion considers the integration of spatiotemporal population estimates with environmental models, the enhancements achieved within the Population 24/7 framework, the challenges for model validation and finally the contribution to the field of hazard risk management.

Chapter seven summarises the main findings of the thesis, referring back to the original aims and objectives set out in Section 1.2. It summarises the limitations, but also outlines the significant enhancements achieved to advance our understanding of population exposure to risk. In the final section proposals for further research and extensions are identified.

Finally, a glossary is provided which defines key technical terms used in this thesis.

Chapter 2: Literature review

2.1 Overview

This literature review is organized into seven sections. Following this overview the first concerns natural hazards, their processes and impacts on people. This is followed by the flood risk context in the UK. The third section concerns population data and spatial interpolation. Section 2.5 reviews examples of population interpolation techniques that consider time, while Section 2.6 examines spatiotemporal examples. Finally, all of these components are integrated regarding their application to assess exposure and vulnerability to natural hazards.

2.2 Natural hazards

In this section, the concepts of hazard, vulnerability, susceptibility, resilience and risk will be explored with particular reference to the way in which hazards impact on human populations. UK flooding case study-specific context will then be provided, followed by an analysis of the representation of population data. Cutter (1996) proposed three fundamental questions that have been the focus of hazard researchers considering these issues. These concern: (i) understanding the human occupancy of hazard zones, (ii) how to mitigate the risk and (iii) impact from hazards and societal response. Questions (i) and (ii) raise important considerations that will be examined further in this review, and represent the primary aim of this thesis. Aubrecht *et al.* (2012b) also address the occupation of hazardous zones and risk mitigation. They state that the quality of available data regarding disaster risk management, particularly exposure and impact assessments, in terms of spatial and thematic accuracy, is one of the most important factors. There are numerous instances where spatiotemporal population estimates could be applied to a natural hazard scenario to estimate human occupancy. Bhaduri (2008) suggests these fall into two broad groups: estimating risk from disasters and for the purpose of public health and socioeconomic analysis. The first, focused on population and wellbeing, is considered in this thesis.

Large numbers of people are at risk from natural disasters such as earthquakes, volcanic eruptions, floods, droughts and wildfires. Typically they are mostly predictable (except earthquakes and some volcanic eruptions) and topographically constrained. Although hazards are often unpreventable, their impact can be minimised through effective disaster planning and emergency preparedness (Bhaduri 2008). The population exposed (E) is recognised as a key component when defining risk to natural hazards (H) in addition to their vulnerability (V). This can be empirically represented as some function of risk (Eq. 2.1). Without the population component no risk would exist.

$$Risk = f(EHV) \tag{2.1}$$

Understanding a population's characteristics such as vulnerability and exposure also presents challenges for practitioners through data deficiencies or systems to handle such information. The formula indicates that aspects of population and management account for all contributing factors required for the development of a natural phenomenon into a hazard. This shows that better representations of population are required to make improvements in assessing the impacts of natural hazards and disaster risk management. The variability of vulnerable populations exposed to natural hazards needs to be recognised to make improvements in emergency planning and develop effective procedures (Cutter and Finch 2008). The calculation of the exposed population is not straightforward as populations are not static and shift dramatically over time (McPherson and Brown 2004).

Humans are vulnerable to hazards which can be categorised as either anthropogenic or natural, or considered a factor of the two. These hazards are both numerous and complex. While a multitude of anthropogenic hazards pose risks to life such as smoking or terrorism, only the destructive power of extreme natural phenomena is considered here. Natural events are capable of causing widespread disaster and considerable loss of life. Natural hazards are widely regarded as one of the ultimate constraints on human activity and provide reminders that we do not live in a world of entirely our own making.

Humans are becoming more susceptible to natural hazards, largely as a consequence of population growth and globalization. In the future it is likely that multiple disasters with fatalities exceeding 10,000 will be experienced on an annual basis (Huppert and Sparks 2006). The general trend (Figure

2.1) appears to suggest that the frequency of natural hazards is increasing. However, it is widely argued that this phenomenon probably occurs as a result of increased exposure, improved and more robust detection and recording methods as well as an absence of sufficient data to verify this trend (e.g. Alexander 1993; Tobin and Montz 1997; Hilhorst and Bankoff 2008). The number of reported disasters rose by 93% between 1992-2001. Vulnerability is a more precise measurement of exposure to risk (Hilhorst and Bankoff 2008) rather than just disaster frequencies. Floods, droughts, earthquakes and tropical cyclones accounted for 94% of all natural hazard fatalities recorded between 1980-2006 (Peduzzi *et al.* 2009).

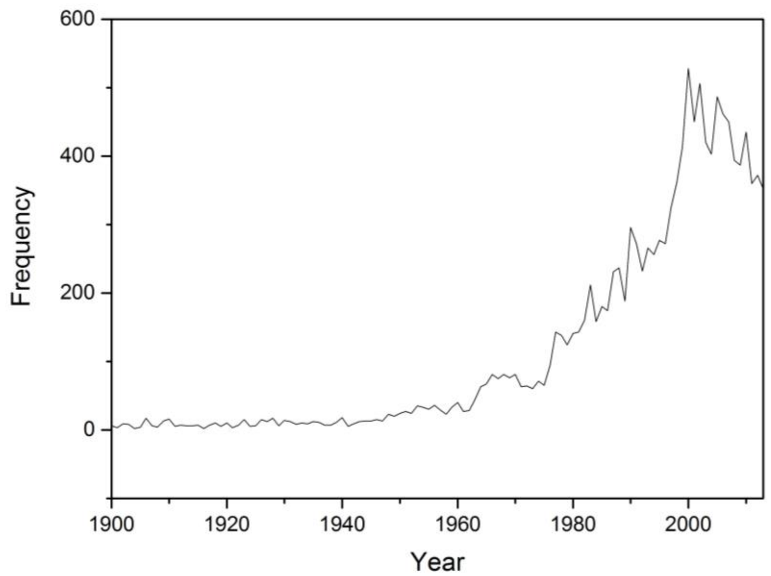


Figure 2.1 Total number of natural disasters by year between 1900-2013. Data source: EM-DAT (2014)

Grasping the variety of natural hazard definitions and concepts (Table 2.1) is important when considering how populations are affected by them. Many definitions contain the same components concerning the physical occurrence of a damaging naturally occurring event, the likelihood of occurrence, and the potential negative impact on humans or the environment.

Natural hazards can be broken down into the following subgroups, which may have an impact on the particular definition derived:

- Meteorological: drought, extreme temperature, lightning, cyclonic depressions, tornado, climate change, geomagnetic storm, wildfires.

- Hydrometeorological: flood, hail, extreme rainfall, blizzard, heavy snow, ice, coastal storm surge.
- Geophysical: seismic activity, volcanic eruption, coastal erosion, landslide, avalanche, lahar, subsidence, tsunami.

Furthermore, aspects of the following events might be considered natural or as a consequence of human interaction combined with other natural triggers (Hewitt 1997):

- Biological: disease pandemics, release of pathogens, escape of radioactive material.
- Environmental: pollution, chemical fires, poisoning, gaseous discharges.
- Technical: infrastructure failure.

A vast literature may be found on each of these subgroups, which in turn may be further subdivided. For example, flooding may be broken down into coastal, pluvial or fluvial with a range of diverse impacts and contributing factors. Some natural hazards such as an avalanche or landslide can be argued to have human triggers or exacerbating factors. Such complicated relationships between humans and the natural environment can often make it difficult to differentiate single or independent causes of an event. Compound hazards occur as the result of another event. Fires, tsunamis and landslides are all potential compound hazards which may occur following an earthquake.

Historically, natural hazards have destroyed entire ecosystems and have been recognised as causal factors in the collapse of civilizations, such as the Minoan empire during the fifteenth century BC. In economic terms, 2011 was the mostly costly year to date regarding natural hazards losses which totalled $380 billion with 27,000 fatalities. The most notable events included the Japanese tsunami and Christchurch, NZ, earthquake. Other high profile events included the occurrence of severe flooding in Thailand, Australia and the USA as well as continued drought in east Africa (Munich Re 2012). The aim of this thesis is to examine how such populations are, and become, exposed to hazards.

2.2.1 Natural hazards and disasters

There is no agreed definition of 'disaster', which can lead to misconceptions in the media and variations in recording data on destructive events (Perry 2007). Consequently, the integrity of databases containing related counts should be carefully analysed within their context. These records are often used to determine the severity of a disaster in terms of economic damage or number of fatalities. The media and everyday language often term as 'disasters' events with considerably different magnitudes and causes. For example, the '1989 Hillsborough Disaster' in the UK resulted in 96 football supporters being fatally crushed, compared to the same application for the 1931 flooding disaster in China which killed 3.7 million people (Smith and Petley 2009). A natural hazard may become a natural disaster once there has been a negative impact on a population, as defined in Table 2.1. In the absence of any population, where no risk to humans exists, the same natural phenomenon does not develop into a hazard (Alexander 1993). The term disaster is often utilised by the media, but without proper definition or calibration of use it may not be that useful in public communication. Many databases will have different definitions and methods governing calculation of fatalities associated with a disaster or natural hazard. The clear definition of a disaster in a dataset is essential when interpreting fatalities that have occurred directly and those that have occurred as result of a subsequent consequence (e.g. EM-DAT 2009b). In addition, disaster fatality datasets may become further complicated when an area affected by a hazard crosses more than one administrative area or national boundary. The United Nations Office for Disaster Risk Reduction's (UNISDR) 'Terminology on Disaster Risk Reduction' provides a concise one sentence definition to provide clarity for the public, authorities and practitioners intended for dissemination to promote better understanding. Definitions taken from the UNISDR and the literature have been summarised in Table 2.1.

Table 2.1 A selection of definitions and concepts regarding natural hazards

Terminology	Definition	Source
Natural Hazard	The probability of occurrence within a specified period of time and within a given area of a potentially damaging natural phenomenon.	UNDRO (1982)
	Threats to humans and what they value.	Harriss *et al.* (1978)
	A naturally occurring or man-made geologic condition or phenomenon that presents a risk or is a potential danger to life or property.	American Geological Institute (1984)
Vulnerability	Being prone to, or susceptible to damage or injury.	Wisner *et al.* (2004 p. 11)
	Characteristics of a person or group and their situation that influences their capacity to anticipate, cope with, resist and recover from the impact of a natural hazard.	Wisner *et al.* (2004 p. 11)
	The degree of loss to each element should a hazard of a given severity occur.	Blaikie *et al.* (1994)
	The characteristics and circumstances of a community, system or asset that make it susceptible to the damaging effects of a hazard.	UNISDR (2009:30)
	Vulnerability = risk amplification measures – risk mitigation measures ± risk perception factors	Alexander (1991)
	The degree of loss (%) resulting from a potentially damaging phenomenon.	EM-DAT (2009a)
	The degree of loss to a given element or set of elements at risk resulting from the occurrence of a natural phenomenon of a given magnitude (expressed 0 to 1, no loss to total loss).	UNDRO (1982)
	The degree to which different social classes are differentially at risk.	Susman *et al.* (1984)
	A possible state that is combined with high risk and an inability to cope.	Cardona (2004)
Exposure	People, property, systems, or other elements present in hazard zones that are thereby subject to potential losses.	UNISDR (2009 p. 15)

Table 2.1 continued

Terminology	Definition	Source
Risk	*Risk = Hazard × Exposed population × Vulnerability*	UNDRO (1979)
	The combination of the probability of an event and its negative consequences.	UNISDR (2009)
	$R_t = ER_s = EHV$**,** where R_t is the total risk, E is the elements at risk, R_s is specific risk, H represents the natural hazard and V is vulnerability.	Alexander (1993)
Environmental Hazard	Extreme geophysical events, biological processes and technical accidents that release concentrations of energy or materials into the environment on a sufficiently large scale to pose major threats to human life and economic assets.	Smith and Petley (2009)
Natural Disaster	A disaster must fulfil at least one of the following criteria: ≥ 10 fatalities or ≥ 100 people affected Declaration of a state of emergency Call for international assistance	EM-DAT (2009b)

2.2.2 **Vulnerability versus exposure**

Some key definitions taken from the literature have been summarised in Table 2.1. This includes an attempt to differentiate between vulnerability, exposure, susceptibility and resilience. This is important when considering risk to population. Vulnerability is described as a combination of factors with varying definitions (Table 2.1). It can be described as a possible state that is combined with high risk and an inability to cope (Cardona 2004). Quantitatively, vulnerability can be considered a function of risk and hazard, or a numeric figure describing the potential degree of loss. A common factor in vulnerability definitions is a population component describing susceptibility, exposure or other characteristics that make them vulnerable when considered with a particular hazard and given level of risk.

Unlike vulnerability, exposure of an element or population might just simply be considered as their physical location within a hazardous zone making them subject to potential losses (UNISDR 2009). It could be argued that exposure only relates to a quantifiable property or physical presence of objects such as population counts, roads and buildings within a hazardous area. An object is exposed as a result of its physical location whereas vulnerability incorporates a greater range of contributing factors that affect the ability to cope. Buildings on an active fault line may be exposed to earthquake hazard, but those buildings built without proper planning regulation to withstand an earthquake will be more vulnerable.

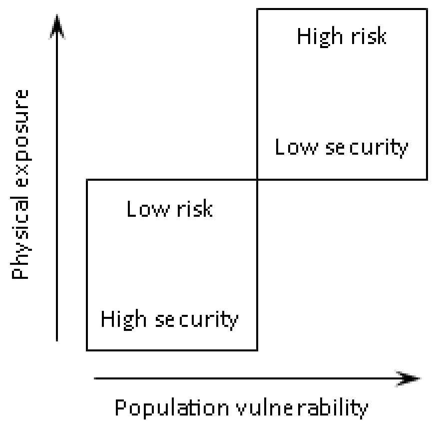


Figure 2.2 A simple vulnerability/exposure matrix after Smith and Petley (2009)

Human sensitivity can be considered a combination of exposure and vulnerability (Figure 2.2). It is clear that the sensitivity of a population to a hazard depends on the degree of physical exposure, such as living on an explosive active volcano. Population vulnerability can also reflect upon preparedness regarding implementation, if any, of an emergency evacuation plan. Populations in less developed countries might experience the same level of exposure for a particular hazard, but have a greater sensitivity through increased vulnerability as a result of inadequate planning or lack of financial resources to adapt. Figure 2.2 highlights the need to understand physical exposure and vulnerability of populations when considering risk and security.

2.2.3 **Susceptibility**

Dow and Downing (1995 p. 4) define vulnerability as 'the differential susceptibility of circumstances contributing to vulnerability'. Consistency in the definition and use of 'vulnerability' is required to advance our understanding of the vulnerability of populations and places to natural hazards (Cutter 1996). It is argued susceptibility can be considered a subset of vulnerability (Dow and Downing 1995), however the two terms are often used interchangeably with little distinction. In the interests of clarity and usability, an attempt has here been made to differentiate these concepts. Although their use is common there is currently little defining literature.

The concept of susceptibility is common in landslide hazard mapping, particularly within the engineering and earth science communities (e.g. Dai *et al.* 2001; Fernández *et al.* 2003; Remondo *et al.* 2003; van Westen *et al.* 2003; Fell *et al.* 2007). These purely concern a probability of an event occurring, where sufficient data are available. Susceptibility may be considered as a spatial probability based on climatic or geological conditions where there is insufficient temporal data (Remondo *et al.* 2003). Unlike vulnerability, which may be defined as a characteristic that causes a potentially harmful situation (Table 2.1), susceptibility suggests a preconditioned detrimental characteristic.

The usage of susceptibility in this context emphasises a probability of occurrence due to a physical pre-conditioning to experience harm rather than the more anthropogenic orientated vulnerability concepts defined in Table 2.1 which consider human characteristics or potential losses. It could be argued that a population living on a volcano is both vulnerable and susceptible, as the consequences of the geologic process of an eruption are independent of human characteristics with a uniform effect on the exposed population. Conversely, it may be argued that such a population is vulnerable due to socioeconomic constraints having a greater influence on its geographic location of residence. In order to assess vulnerability the susceptibility of elements at risk needs to be identified and understood (Birkmann 2006).

2.2.4 **Resilience**

Resilience is a useful concept for hazard risk reduction which has been significantly developed (Walker *et al.* 2004) since its introduction by Holling (1973, 1986). The introduction of resilience theory, particularly that concerning ecological systems, has had a notable impact within the natural and social sciences. Figure 2.3 displays the trajectories of two population's densities through time, suggesting the competition for resources arriving at equilibrium over time. The spiral represents a planar view of the density trajectory with time.

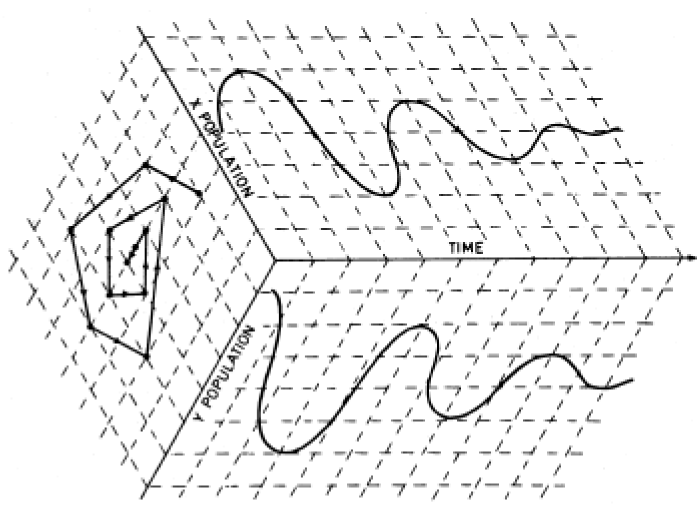


Figure 2.3 Density change in two populations over time (Holling 1973 p.5)

Resilience is often defined as the 'ability to spring back'. In physical terms this can relate to the stress and strain exerted by a load placed on a material. Since Holling's (1973) conceptualization the term has been used in an increasingly metaphorical sense relating to a system's or city's ability to cope with and recover from external shocks (Klein *et al.* 2003a). Resilience can be attributed to promoting sustainability as a mechanism to reduce vulnerability. However, it is perceived that while resilience may be a desirable property, it exists as a poorly defined conceptual framework. Klein *et al.* (2003a) recognise a requirement for better quantification of resilience. The quantitative assessment of populations is important and allows the definition of magnitudes of variation.

Dovers and Handmer (1992) identify three types of societal resilience to environmental risk:

- Type 1 Resistance and maintenance: resistance to change and uncertainty where the *status quo* are maintained. Threats are identified and mitigated but this may lead to inaction to avoid uncertainty. As a result the society may not be fully prepared to deal with unexpected events.
- Type 2 Change at the margins: change that does not challenge the basis of society and serves the interests of the affluent few. Not generally concerned with the general population or environment.
- Type 3 Adaptability: flexibility and the ability to change basic assumptions and institutional structures while effortlessly embracing new ideas.

There needs to be a clear definition of resilience and the factors by which it is determined if it is to be usefully applied to populations in a hazard context. Although often considered a desirable characteristic, the ability of a population to 'spring back' to the same perpetual state of vulnerability following a natural hazard may not in fact be beneficial. Resilience to natural disasters is the capacity of the population to cope with the emergency and learn from the experience by implementing new physical and social structures to become better adapted (Ride and Bretherton 2011). Klein *et al.* (2003a) propose that usage of the term 'resilience' should be reserved to describe specific system properties concerning the amount of disturbance it can absorb while maintaining function and the degree of capability of self-organisation.

2.2.5 **Introducing risk**

There is a level of risk which exists for every activity. A large proportion of this risk is termed acceptable risk. This is the level of potential losses that a society considers acceptable given the existing economic, social, cultural, political, environmental and technological conditions (UNISDR 2009). Our knowledge of natural hazard distribution and historical frequency has advanced, however less is known about risk and its impacts on the population affected (Cutter 2010). Understanding and mapping risk is another fundamental component in assessing the degree of potential harm to humans. Risk (Table 2.1) can be defined as the product of an exposed population, vulnerability and hazard.

2.2.6 **Voluntary and involuntary risk exposure**

In terms of individual risk, two main categories can be considered: involuntary and voluntary risks. Involuntary risks occur without our knowledge or consent such as a lightning strike, earthquake or meteorite impact. Voluntary risks result from activities, or their potential consequences, that individuals choose to undertaken when they consider that the risk involved is acceptable. There are numerous voluntary risks undertaken by individuals daily such as travelling on the road network or smoking tobacco. Living within a flood plain or on an active volcano might be considered voluntary although such decisions may be governed by complex cultural, social or economic constraints outside of an individual's control.

It has been suggested that the observed increase in the frequency of natural disaster events is partly a result of increased exposure. By the end of the first decade of the twenty-first century more people lived in urban rather than rural areas for the first time in human history, resulting in increased population densities (Dye 2008). In Europe this accounts for 70% of the population, while globally the trend in population growth will mainly be urban for the foreseeable future.

Urbanisation extremes are exemplified by megacities which can be defined as urban agglomerations exceeding 10 million inhabitants (UN 2008), characterised by high population density and rapid development (Sekovski *et al.* 2012). Most megacities, 17 out of 21, are located in coastal regions (Nicholls 1995; Sekovski *et al.* 2012; von Glasow *et al.* 2012). Mitchell's (1999b) edited volume poignantly assigns the title 'Crucibles of Hazard' to megacities. Large populations in coastal megacities are vulnerable to natural hydrological and meteorological hazards as well as tsunamis (Klein *et al.* 2003b). Inland cities of high population density are not immune from hazard. The 1989 earthquake in Mexico City, which itself is situated within a volcanic basin, is estimated to have killed 5,000-10,000 people (Mitchell 1999a). It has been suggested that the human preoccupation with achieving economic growth, even in the face of adversity, has spatially concentrated social and economic activities in hazardous areas (Puente 1999). As these cities usually contain the highest concentrations of population and economic activity they can be predisposed to suffer catastrophic natural disasters.

Current populations, or the origins of modern settlements, may have become established because of the benefits brought by successive natural processes. These might include the fertile ground or crossing points on a flood plain, ash rich volcanic soils, mineral deposits or access to the coast for trade and transportation.

Lewis (1989) speculated on a fictional earthquake striking the centre of Tokyo and suggested the impact could devastate trading on Wall Street and lead to global economic recession. Only six years later Japan was struck by the 1995 Kobe earthquake, one of the most powerful ever to affect an urban area. The M 7.2 earthquake centred on Kobe, the world's sixth busiest container port, resulted in c. 600 direct fatalities and made 300,000 people homeless (Horwich 2000). Although economic growth recovered relatively quickly following the Kobe event, the earthquake and mega-tsunami of 2011 has again raised questions about population vulnerability (Hein 2013). The high density construction typical throughout Japan and megacities, such as Tokyo, raises renewed questions on population risk to natural hazards.

Historically, social memory of disaster and crisis would have been communicated by the elders of many indigenous societies. In industrial societies it is unclear who, if anybody, facilitates the passage of collective memory. Increasingly this information may now reside in new media or archives (Berkes 2007). In psychology the *accessibility principle* suggests that the subjective time of an event depends on how much is known about it since it occurred. The more that is known, the more recent it will seem (Brown *et al.* 1985). Events such as extreme volcanic eruptions may only occur on a millennial scale. Therefore it is difficult for humans to retain memories of eruptions from specific volcanoes (Grattan and Torrence 2007).

2.2.7 Calculating risk

There are varying qualitative and quantitative definitions of risk, some of which have been summarised in Table 2.1. Quantitative risk assessments are often not understood by the general public and need to be communicated more accessibly. Uncertainties need to be made clear (Smith and Petley 2009) whilst maintaining trust. In this thesis, calculating differing levels of risk for the hazard scenarios in conjunction with fluctuations in population allows assessments of exposure.

The key elements considered in general risk assessments regarding natural hazards are summarised in Figure 2.4. The evaluation of the direct impact of a hazard as well as any indirect impacts as a result of damage is required alongside an understanding of the population or elements at risk and their vulnerability. In a flood risk assessment this might involve understanding the population at risk within the hazard footprint and an indication of their vulnerability. During an earthquake, increased risk might feedback through severe building damage as a result of construction standards, which in turn might increase vulnerability of the population at risk.

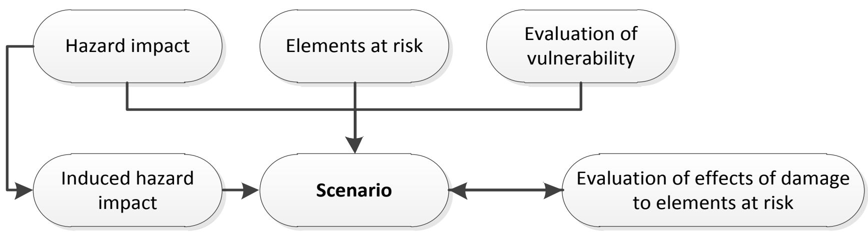


Figure 2.4 Risk assessment for a given scenario after Douglas (2007)

The Disaster Risk Index (DRI) provides a quantitative approach to assessing global exposure and vulnerability to natural hazards, with the ability for comparisons between countries. Global statistical analysis suggests that 11% of the population in less developed countries are exposed to natural hazards and account for 53% of the casualties, whereas 15% of the population in more developed countries are exposed but account for only 1.8% of the fatalities (Peduzzi *et al.* 2009). The formula for calculation is based on the United Nations definition (Table 2.1) that risk contains three components: elements at risk, vulnerability and the hazard occurrence probability (Peduzzi 2006). This includes accessing data on the number of human fatalities per year and the population living in an area exposed (Peduzzi *et al.* 2009) to the potentially damaging effects of a natural hazard. The physical exposure is determined by:

$$Physical\ Exposure = \sum_{i}^{n} F\ Pop_i \qquad (2.2)$$

Where *F* is the annual frequency of events of a given magnitude, Pop_i is the total population of the spatial unit for each event, *i* is the exposed population per event and *n* is the number of events considered. Human vulnerability was considered using a parametric model based on socio-economic variables and the physical exposure (PhExp) (Peduzzi *et al.* 2009):

$$K = C(PhExp)^{\alpha} V_1^{\alpha_1} V_2^{\alpha_2} \dots V_p^{\alpha_p} \qquad (2.3)$$

Where *K* is the number of fatalities for a certain type of hazard, *C* is the multiplicative constant, $V_{1..p}$ are the number of socioeconomic variables and α is an exponent of *V*. This was utilised to define the generalization of the multiplicative to derive the number of fatalities/year by taking the product of physical exposure and vulnerability.

2.2.8 **Risk management and emergency preparedness**

Emergency responders need to be able to estimate who is present at a given time at the start of a hazardous event (McPherson 2006). Perry and Lindell (2003) suggest three critical components of emergency preparedness: planning, training and written communications. The first, planning, will be considered here. Accurate, high-resolution, temporally varying population estimates are required within the planning process (Bhaduri *et al.* 2007). A pre-emptive approach within the planning phase can be to consider the 'business as usual' scenario. This may reflect the actual composition or distribution of potentially affected populations depending on the time of day a hazard may occur. This approach can be suited to both rapid onset events, where there is little time to react, and events with longer lead in times.

Accurate population estimates are still beneficial when considering preparedness for hazard events with slower onset speeds. It may not be possible to account for irrational human behaviour during emergency situations. Therefore, a realistic firm estimate of the population immediately prior to any event is a valuable advantage for targeting plans and resources.

The 'as low as reasonably practicable' (ALARP) principle (Figure 2.5) acknowledges different levels of risk and how tolerable they are considered to be. At the top end of the spectrum, unacceptable risk posed to individuals or society requires mitigation largely irrespective of the financial cost. In demonstrating that a risk is ALARP any cost to further reduce the risk would need to be grossly disproportionate to the benefits that could be achieved (Smith and Petley 2009). The ALARP principle was developed formally in UK policy (Melchers 2001) and popularised through legislation particularly the *Health and Safety at Work etc. Act 1974.*

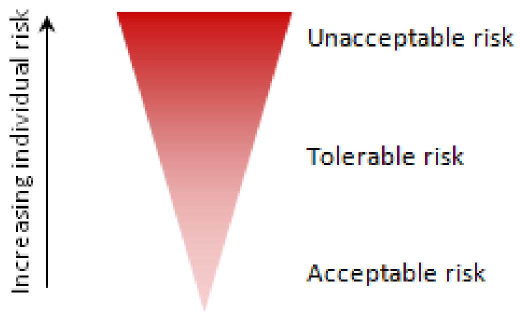


Figure 2.5 The ALARP principle after Crozier (2005)

The uncertainty associated with natural phenomena has been discussed by scientists (Handmer *et al.* 2001). The precautionary principle emerged out of ecology and sustainability discourses where it is argued that it is worth taking preventative actions even if the risk of a disaster is low but the consequences may be high (Smith and Petley 2009). This principle has many similarities within the science of natural hazards and disaster risk reduction.

2.2.9 Risk perception and communication

The international community's response to natural disasters has generally been reactive, with events in the developed world taking excessive precedence in western media (Peduzzi 2006). Severe flooding on the Indian subcontinent killed 2,000 people in August 2004 and attracted 9,000 words in British newspapers. On the same day Hurricane Charley struck Florida with 16 fatalities which was represented by 19,000 words (Adams 2004). The previous section highlighted the importance of human and societal memory, while this section only attempts to briefly acknowledge a wider issue of risk communication and perception and the implications for risk management. Studies have suggested inadequate levels of awareness and understanding within communities of predictable natural hazards such as flooding (King 2000). It has also been noted that the perception of risk declines as any perceived benefit increases (Slovic 2000). This is likely to have a detrimental impact on natural hazard preparedness as populations become complacent or reliant on defences. This is illustrated by the November 2012 flooding in Worcestershire, UK, when newly installed flood defences failed. It was reported that some residents were unprepared or had even been celebrating the launch of the new defence scheme when their properties were flooded (Morris 2012b), although they had previously been victims of flooding.

Communicating risk in a manner understandable to, or likely to have an impact on, the general public is another important consideration for natural hazard risk management and preparedness. Hazard maps can be one method of communication and are considered essential in communicating volcanic risk, although a study focused on Montserrat suggested that people still struggled to interpret traditional maps. It concluded that enhanced three-dimensional maps with perspective photographs aid topographic recognition and orientation (Haynes *et al.* 2007). A UK flood risk poster (Figure 2.6) is an example of a public information campaign for risk communication.

Figure 2.6 An Environment Agency flood risk awareness campaign poster in the UK (EA 2014)

Even with prior warning and risk education, some people still fail to take preventive measures or evacuate if required. A Californian study into flood induced evacuation by Heath *et al.* (2001) found that 19.4% of households failed to evacuate when ordered for a particular event. It found statistically significant trends that people without children or with pets were more likely to remain in their properties. Reasons for this behaviour might have been due to the children's fear or anticipation that they are more likely to be harmed than an adult. It also suggested that pet owners were willing to risk their own lives to remain and look after pets (Heath *et al.* 2001).

Evacuation research has suggested that people are more likely to take precautionary action if they perceive a real threat (Burnside *et al.* 2007). This is indicative of effective risk perception and communication. Research by Burnside *et al.* (2007) on the behaviour of New Orleans residents under evacuation orders during hurricane Katrina offers several reasons for people failing to adhere to evacuation warnings. False alarms are given as the primary cause for residents failing to evacuate as well as not having the financial means to do so, while a subset will just refuse. Others may decide to remain to protect their properties and belongings from loss by the natural hazard or public disorder.

Cole and Fellows (2008) outline four lessons learned from risk communication failure in New Orleans during hurricane Katrina:

1. Effective communication is no good if followed by inadequate crisis messages.
2. Messages should be prepared before the crisis.
3. For a message to be effective it needs to appear reliable to the audience.
4. Risk communication messages must be adapted for the demographic characteristics of the audience such as ethnicity, class and gender.

2.2.10 **Hazard spatiotemporal characteristics**

A hazard's footprint is its manifestation in space and time (BRISK 2010). The footprint of an earthquake might be defined by the extent of damage inflicted or by using interferometric synthetic aperture radar (InSAR) photogrammetry to delineate the extent of ground surface displacement. Presently, estimates of the population exposed to a natural hazard are based on its footprint. Proxies of a hazard's physical footprint such as a watershed or flood plain concerning floodwater inundation are often used (Guha-Sapir *et al.* 2011).

Hazardous zones are occupied for a variety of reasons. As identified, historic development around coastal ports, trading and industry continues to shape urban living today where gentrification and redevelopment have created economic wealth and jobs in areas potentially at risk. There is an evident correlation between poverty and increased vulnerability exemplified by a family which has insufficient income to make property improvements or move to a safer place (Delica-Willison and Willison 2008). Poverty has been identified as one of the largest contributors to vulnerability.

Mount Vesuvius, Italy, is well known to be home to a large population vulnerable to potential destructive eruptions. It is estimated that over 1 million people live on and around it (Carlino *et al.* 2008) making it the world's most populated volcano (Kilburn and McGuire 2001). Population densities around Vesuvius are amongst the highest in Europe (Alexander 1993), while housing quality is some of the poorest (Chester *et al.* 2002). These factors combine to exemplify one of the highest vulnerabilities to natural hazards. It was reported that the Italian authorities would pay £18,000 to property owners living on Vesuvius to leave with the aim of emptying 100,000 properties over 15 years. However, this is considerably below the market value and many people could simply not afford to relocate (Arie 2003).

As well as space, hazards have temporal characteristics. Change in intensity and space over time is an important consideration when examining the human occupancy of hazardous areas. One categorization for hazards types is intensive and pervasive (Kates 1976). Intensive events are often small in areal extent, intense in impact and short in duration. This is often associated with sudden onsets and poor predictability. Examples may include earthquakes, landslides and volcanoes. Conversely, pervasive hazards such as drought, fog and temperature extremes are often widespread, with long durations and gradual onsets. Therefore, they can also be detected and warned against more easily. Some events such as floods can be defined by both extremes from intensive flash to pervasive coastal flooding. This demonstrates the importance of understanding both space and time in hazards to deduce their impacts.

Guidance provided under the European Union Flood Directive (see Section 2.3.5) states that flood hazard maps should contain flood extents with a low, medium (≥ 100 years), and high probability as well as hydrological data such as expected depth and velocity. The specification for flood risk maps, detailing the potential adverse side effects of the hazard should include (Chpt 3(6) 5a&b) an indication of the number of inhabitants potentially affected and economic activity of the area (European Council 2007).

A hazard's 'frequency' can be defined as the number of events of a given magnitude within one unit of time. It follows that a hazard's return period is the reciprocal of its frequency (Alexander 1993). A return period is a unit of time, in which an event of a given magnitude is statistically likely to occur. A return period of a five metre river flood or a magnitude nine earthquake within a specific location, in space, may be quoted in years based on the analysis of disaster records. An improbable combination of factors caused severe flash flooding in the narrow valley catchment village of Boscastle, north Cornwall, UK. The likelihood of these factors occurring together and generating an event on a similar magnitude has been estimated as a 1 in 2000 year event (Bettess 2005; and Murray *et al.* 2012). It may not be possible to predict the onset of some particular types of hazards, such as when and where a sinkhole will develop or lightning bolt will strike.

Statistics show extreme events, of a high magnitude, occur less frequently than lower magnitude events (Alexander 1993). Perhaps large magnitude events might occur in uninhabitable places such as at sea or in regions which they are rendered hostile to human settlement and as a result do not pose a ‘hazard’. Natural hazards can vary greatly in their onset times, durations and spatial extent. An earthquake has a quick onset, whereas the impact felt by a drought is gradual but over a long duration (Peduzzi *et al.* 2009). Disasters can be classified into two distinct phases: the impact and long term effects. The initial impact may be short lived depending on the type of hazard and its onset time. In this phase survivors can do little except wait for the worst to pass. Often disaster survivors are critical of delays in response in providing food, medical treatment and shelter (Alexander 1993). Understanding population vulnerability including more realistic temporal density estimates will help ensure that human crises may be managed better in the future with fewer fatalities.

Two powerful earthquakes occurred in New Zealand causing fatalities and widespread destruction along a previously unknown fault (Elliott *et al.* 2011). The first M 7.1 earthquake centred on Darfield, Canterbury occurred at 04:35 on 4 September 2010 with no direct fatalities. A powerful M 6.2 aftershock centred on Christchurch occurred at 12:51 on 22 February 2011 and caused 181 fatalities (Kaiser *et al.* 2012). Although the aftershock was almost ten times less powerful it had devastating consequences and occurred during the middle of the day. The devastation of the secondary earthquake has been attributed to its proximity to the centre of New Zealand’s second largest city (Kaiser *et al.* 2012), however it also occurred during the middle of the day. It has long been recognised that the number of earthquake fatalities is affected by the time of day in which it occurs (Coburn *et al.* 1992).

Throughout the preceding review the importance of time and space in hazard has been recognised and requires greater analysis in terms of its relationship to population. The processes of disaster development and concepts of risk and vulnerability have important impacts when considering the population exposed.

2.3 Flooding in the United Kingdom

This section examines the background and conceptual context of flood risk in the UK. The background provided exemplifies the case for advances required in natural hazard risk assessment, which is demonstrated in the empirical chapters of this thesis (Chapters 3-5).

In 2010, the UK government's *National Security Strategy* defined international terrorism, cyber-attacks, international military crises and major natural hazard incidents as the highest priority risks affecting the country for 2010-15 (HM Government 2010). In terms of natural hazards, the highest risks, in a broad sense including social disruption and economic harm, are specified as coastal flooding and severe effusive volcanic eruptions (Cabinet Office 2012). The Icelandic 2010 eruption of Eyjafjallajökull caused widespread air travel disruption and is estimated to have cost the aviation industry $250 million a day (Gudmundsson *et al.* 2010). When considering flooding and overseas volcanic eruptions within the conceptual definition of risk (Table 2.1) different concepts may be developed. The footprint of airborne volcanic ash may 'affect' the whole population, or perhaps only those using an airport; whereas the overall vulnerability to an individual may be lower, the economic vulnerability of the country may be greater. Coastal flooding and effusive volcanic eruptions are considered to have the greatest national impact in the UK relative to the occurrence of any other natural hazard, with a probability of occurrence between 0.005 (1:200 years) and 0.0005 (1:2000 years) within the next five years (Cabinet Office 2012).

2.3.1 Types and causes of flooding

There are seven main categories of flooding that have caused notable events in the United Kingdom:

1. Fluvial flooding: this occurs when rivers overtop or burst their banks as channel capacity is exceeded by intense rainfall or snowmelt. The magnitude and response time may be exacerbated when prolonged rainfall has saturated the ground causing water tables to rise.

2. Coastal flooding: occurs when cyclonic conditions combine with a high tide, which can be exacerbated by atmospheric low pressure to form a tidal surge or higher than usual spring tides. There is an inverse relationship between storm surge height and barometric pressure approximately equal to a 1 cm rise per millibar reduction in atmospheric pressure (Welander 1961). This effect is likely to be enhanced by strong winds in an anticyclonic system, resulting in an increase in wave heights. Coastal topography can also amplify this affect where converging estuaries or inlets can act as a funnel to channel storm surges.
3. Pluvial flooding: also referred to as surface water, occurs when rainfall that is usually removed by a drainage system exceeds the capacity of that system. This results in overland flow and accumulation in local topographic depressions. This occurs when intense rainfall falls on impermeable surfaces common in urban areas, saturated or frozen ground. This usually occurs when rainfall rates exceed 20 mm hr^{-1} for up to three hours (Houston *et al.* 2011). It is estimated that 2 million people in urban areas are at risk of pluvial flooding, and this is likely to increase as a result of urbanisation and climate change. Pluvial flood risk accounts for approximately one-third of the UK total flood risk (Houston *et al.* 2011). The majority of pluvial flood events occur during the summer where average precipitation totals are lower, but more intense rainfall events combined with urban ground sealing increases run-off (Smith and Lawson 2012). Intense rainfall in urban areas can result in rapid onset flood events in areas with high spatiotemporal variations in population. This research provides a methodology which aims to quantify this phenomenon.
4. Flash flooding: an extreme flood event generated by intense rainfall over rapidly responding catchments (Brauer *et al.* 2011). In the UK flash floods are considered to peak in under 3 hours in 5-10 km^2 catchments typically characterised by steep basins and thin soils, often associated with impervious underlying geology resulting in rapid run-off (Collier 2007).
5. Infrastructure failure: this can lead to flooding which may be considered to have both anthropogenic and natural causes. The

failure of infrastructure due to natural or human causes has the potential to rapidly release large volumes of water that can result in devastating flash flooding. A damaged water main, coastal levee failure or dam burst (reservoir flood) can all initiate flooding on a variety of spatial extents and magnitudes.

6. Ground water flooding: occurs when the water table rises, or natural springs reactivate, due to prolonged rainfall. This is common on permeable bedrocks or underlying substrates such as chalk, sand and gravel. The onset may occur days or weeks after sustained rainfall with the potential duration to last several weeks. Properties may be flooded through the floor or basements, while water may also emerge from hillsides in the form of springs (Environment Agency 2011).
7. Sewer flooding: this occurs when flood water combined with raw sewage exceeds the capacity of that system and enters properties or emerges from breaches in the street when volumes overwhelm the infrastructure. Parts of the UK sewer system are over 150 years old (e.g. Manchester and London) (Tait *et al.* 2008) and under increasing pressure from climate change and the construction of new housing. Combined sewer overflows discharge excess waste water from intense rainfall into rivers to help prevent flooding in properties and hazards to human health (DEFRA 2012).

While other types of flooding are of significance globally, only the commonest types with predominantly natural causes affecting the UK have been considered above for the purpose of this review and applications within this thesis.

2.3.2 **Climate change and flooding**

Flooding is the most common natural hazard in Europe, and has been linked to consequences of climate change despite greater frequencies in the past (Beven 1993; Wilby *et al.* 2008). In the UK flooding is recognised as one of the most damaging and costly natural hazards (Brown and Damery 2002).

It is estimated that the annual cost of flood damage in the UK is £1.1 billion with 5.2 million properties currently at risk. This is reflected in the high cost of flood risk management, with departmental spending currently at £664 million (DEFRA 2011). Average flood damage costs could rise to £27 billion by 2080. Maintaining existing levels of flood defence, including maintenance of current assets and new construction, will require spending to increase by £1 billion per year by 2035 (UK Parliament 2012).

The UK climate projections (2009) (UKCP09) model the future climate at 30 year intervals and 25 km resolution. In general all parts of the UK are projected to warm by 2080, with the greatest temperature increases in southern England. On the whole there will be little change in average annual precipitation although there will be increased seasonality. Parts of southern England can expect a decrease in summer precipitation by 40%, while western areas of the UK may expect a winter increase of up to 70% (Murphy *et al.* 2009).

This brief summary of the potential implications of future climate change demonstrates the impact this may have on future hazard events and government spending. An increase in winter rainfall is likely to further increase fluvial flood events. Warming summer temperatures may progress the development of convective storm cells, while global sea-level rise will increase coastal flood risk.

2.3.3 **Recent flood history**

A summary of significant contemporary and recent flood events is presented in Table 2.2. These events have been selected for their recent historical importance and to provide a context to British flood severity and policy decisions.

The 1953 North Sea storm surge, or 'The Big Flood', was the worst natural disaster to affect the UK during the twentieth century. As a direct result of exposure or drowning there were 307 fatalities in an already vulnerable post-war population (Baxter 2005). However, the 1947 flood event (Table 2.2) was the most extensive in the twentieth century (Marsh 2008). The widespread damage and disruption caused by the 1947 flood event is considered a benchmark in flood strategy for England and Wales (Horner and Walsh 2000; Marsh 2004).

Contemporary flood events, particularly those during the first decade (2000 and 2007 events) of the twenty-first century have had a profound impetus on British and European flood risk management, policy and legislation (Andryszewski *et al.* 2005).

Table 2.2 A contemporary history of notable flood events in the UK

Area affected	Date	Description	Source(s)
Southern England and Wales	March 1947	Large scale river flooding caused by rapid snowmelt and persistent heavy rain from a frontal system that swept across southern Britain. This followed a severely cold winter, average February 1947 temperature was -3.8 ˚C, with snow depths of up to 1.2 m in lowland England. Damage caused by flooding and heavy snowfall caused coal shortages and electricity blackouts, killed 2 million sheep and left 1 million people without a potable water supply.	Marsh (2004, 2008), RMS (2007b)
Lynmouth, Devon	August 1952	An unprecedented flash flood on 15 August in the steep narrow coastal catchment on the north coast of Devon. Heavy rainfall caused by atmospheric ascent of a slow moving low pressure frontal system promoted by orographic ascent over Exmoor. Rainfall exceeded 200 mm, falling on saturated ground. This resulted in 34 fatalities, damage to 93 properties and vehicles swept out to sea.	McGinnigle (2002)
Eastern Scotland and England	January 1953	A storm surge caused by the combination of a high spring tide and cyclonic system. Atmospheric low pressure and strong winds increased tidal heights destroying flood defences in eastern Britain and northern Europe. This effect was amplified by the funnel-like effect of the southern convergence of the North Sea coastlines. There was major infrastructure failure to telephone lines, gas, water and electricity supplies. Central London narrowly survived inundation, but water levels reached the top of the Victoria and Chelsea embankments, spurring the eventual construction of the Thames Barrier. Number of fatalities: 307.	Baxter (2005)
Midlands and Wales	April 1998	The widespread Easter floods of 1998 were attributed to two frontal systems and a slow moving depression moving south across the UK. While the rainfall levels were not unprecedented the 48 hour maximums were more representative of summer convective storms. Rainfall fell on saturated ground causing losses of £500 million and five fatalities.	Horner and Walsh (2000)

Table 2.2 continued

Area affected	Date	Description	Source(s)
England and Wales	Winter 2000/01	Prolonged heavy rainfall during the winter caused the most extensive fluvial flooding since the 1947 event. Rainfall was exceptional by UK standards caused by the continual passage of frontal systems sustained by several months by south-westerly winds. Some areas experienced constant rainfall for two weeks. Damage totalled £1 billion and 10,000 properties were inundated.	Marsh and Dale (2002)
Boscastle, Cornwall	August 2004	A severe flash flood within the narrow coastal catchment on the north Cornwall coast. Unusual due to a rare combination of factors and highly localised and intense rainfall rather than total levels. Similar to the Lynmouth 1952 event however, remarkably there were no fatalities. This has been attributed to the proximity of an air-sea rescue naval base and dedication of the military and coastguard who navigated rescue helicopters into a convective storm cell, supported by the rapid response of emergency services on the ground.	Rowe (2004); Burt (2005); Lewis (2009); Murray *et al.* (2012)
UK wide	June 2007	Widespread flooding affecting Northern Ireland, Scotland, England and south Wales. The scale of the flood exceeded its predecessors and was unprecedented in recent history, exceeding the 1947 benchmark in some areas. Rainfall totals exceeded 300% of the average in some locations. The event resulted in the iconic image of Tewksbury Abbey surrounded by floodwater, which was partially inundated for the first time in 247 years.	RMS (2007a); Marsh (2008)
UK wide	November 2009	A UK wide flood event, including the Isle of Man, with parts of northern England worst affected. UK wide average rainfall in November 2009 was exceeded by 184%. Cumbria attracted widespread news coverage as six river bridges collapsed, one resulting in the death of a police officer. The army intervened to build a temporary bridge across the River Derwent.	Met Office (2009); MOD (2009)
Southern England and Wales	November 2012	Numerous severe flood warnings were issued by the Environment Agency with wide spread flooding across southern England and north Wales (1,800 properties flooded)	Environment Agency (2012a)

Table 2.2 continued

Area affected	Date	Description	Sources(s)
England and Wales	Winter 2014	During January 2014 a rapid succession of severe winter storms caused widespread flooding in Somerset and on the River Thames. Repeated powerful storm surges destroyed sections of seawall in southern England and Wales and which caused the collapse of two coastal railways. Coastal flooding also caused significant damage to the Victorian promenade at Aberystwyth and necessitated the evacuation of 600 university students from seafront residences.	Gevertz (2014); Met Office (2014)

2.3.4 The Pitt Review

An independent government review was commissioned following major flooding in the UK between June-July 2007. The review, undertaken by Sir Michael Pitt, made 92 recommendations for improvements. The key recommendations (RE) in relation to flood hazard and population are summarised here (Pitt 2008). Three recommendations have been highlighted to provide exemplars of how spatiotemporal population modelling techniques could be directly applied to flood risk management:

> **RE16**: Local authorities should collate and map the main flood risk management and drainage assets.
>
> **RE70**: A programme should be established to encourage individuals to be better prepared and become self-reliant during emergencies. This will allow the authorities to focus on the people most in need.
>
> **RE74**: The impact of flooding on the health and wellbeing of people should be monitored and mitigations put in place to manage these effects.

Recommendations 70 and 74 illustrate the need for a greater understanding of exposed populations. This will aid the development of effective emergency plans or help reduce probability of fatalities or serious injury. Recommendation 16 largely concerns mapping the physical infrastructure, but this provides the opportunity to highlight deficiencies in data concerning population variability in emergency planning.

The UK government and relevant authorities have responded to recommendations contained within the Pitt Review and provided an additional £34.5 million of funding to address these proposals (DEFRA 2008). This has included increasing public risk awareness, publication of online flood hazard maps, development of warning systems and the creation of the Flood Forecasting Centre. The methodology proposed in this thesis will involve producing local scale flood hazard maps that meet these legislative criteria and current policy priorities. While these are always going to be subject to change the requirement for more accurate flood maps that consider population at a greater spatiotemporal resolution is noted.

2.3.5 **Management and mapping**

Flood risk, mapping and management are governed by a hierarchical top-down structure in the United Kingdom comprising of a number of government departments and non-governmental bodies. This is enforced through legislative acts and European Union (EU) directives. The *Civil Contingencies Act 2004* progressed into statute for the purpose of defining a single legislative framework to deal with serious emergencies. This resulted from a governmental review on emergency planning arrangements in times of crisis which included the occurrence of severe flooding in 2000/01 (Table 2.2) (Cabinet Office 2011). While this act covers a multitude of eventualities including major natural hazard incidents, civil unrest, terrorist activity and biological hazards, it focuses on protection at the local level concerning emergency services, utility providers, local authorities and operators of transport infrastructure. Analysing population exposure to hazards at a local scale is a key original contribution from this research, which would be applicable for agencies concerning emergency management.

The EU *Floods Directive (2007/60/EC)* has had implications for the way flood risk is managed in the UK. Parts of the directive (e.g. Section 6) relate more readily to continental Europe concerning cross-border flooding and water courses. However, interagency and international collaboration still remain key components in flood risk management. It is necessary for member states to provide flood hazard maps, and risk maps regarding different flooding scenarios (Section 12). In order to comply, member states must complete flood hazard and risk maps by 22 December 2013, and publish flood risk management plans by 22 December 2015 (European Council 2007). In addition the *Flood Risk Regulations 2009* implement the requirements of the directive in Britain. Under these regulations the Environment Agency is responsible for the production of the hazard and risk maps within the time frame specified. Lead Local Flood Authorities (LLFAs) have been identified to map local flood risk. This is being undertaken using the National Flood Risk Assessment tool, with the consideration to update the national flood map in line with these regulations (Environment Agency 2012b).

The *Flood and Water Management Act 2010* was the UK government's response from urgent recommendations for legislation arising in the Pitt Review (RE 28), and provides a more comprehensive management of population flood risk and protection of infrastructure including drinking water supplies and properties. It also ratifies the EU Floods Directive in British law.

A severe flood warning is the highest alert issued by the Environment Agency in England and Wales, and its counterpart the Scottish Environment Protection Agency (SEPA). This is issued when severe flooding is considered imminent with a significant threat to life. It is preceded by a 'flood alert' and 'flood warning'. Assessing the risk to life requires a prior understanding of the distribution of exposed populations. The new methodologies proposed in this thesis provide a mechanism to achieve this.

2.4 Spatial interpolation of population data

This section provides a review of the methodologies used to estimate population densities which are required to make improved judgments on population exposure to natural hazards. Deriving population data that are contained within flexible geographical referencing systems is a key consideration for assessing the potential risk. Converting data into a common format allows subsequent analysis. Interpolating population data onto a regular grid is an output format common in the methodologies discussed. A flexible geographical referencing system, such as a grid, mitigates issues arising from irregular and incompatible areal units often associated with aggregated demographic datasets. The interpolation techniques contained within this section provide mechanisms to disaggregate population which can be used for the purpose of assessing exposure to natural hazards.

A structure is proposed in Figure 2.7 to divide the two main groups of interpolation methodologies into area and point based, according to the type of input data used, similar to Lam's (1983) spatial interpolation review. However, the present review subsequently sub-divides these categories for the consideration of the use of ancillary data.

2.4.1 **Representing population data**

Population can be considered as a volume spread over a surface occupying a fixed spatial area, such as the distribution of a census count within an output area. This population or 'volume' is fixed. Cartographic techniques for mapping volumetric data traditionally fall into three main categories: choropleth, dasymetric and isarithm (Langford and Unwin 1994). Traditionally the choropleth map has been the conventional tool for displaying population density. This process involves the shading of arbitrary zones, often census areal units or administrative zones. This implies that the areal unit concerned has a uniform population density, with abrupt changes at the boundaries, unlikely to be found in reality. Large areal units tend to show lower population densities due to their comparatively larger area to population ratio. Generalisation of the data increases with the size of the areal unit in a choropleth map. Population density for a typical city may increase towards the centre according to a choropleth map, as expected, but this could also be amplified through decreasing ward sizes (Langford and Unwin 1994). Choropleth maps are a discontinuous method and therefore poorly represent the underlying spatially continuous population (Langford and Unwin 1994). This effect is difficult to quantify as the units are often irregularly shaped.

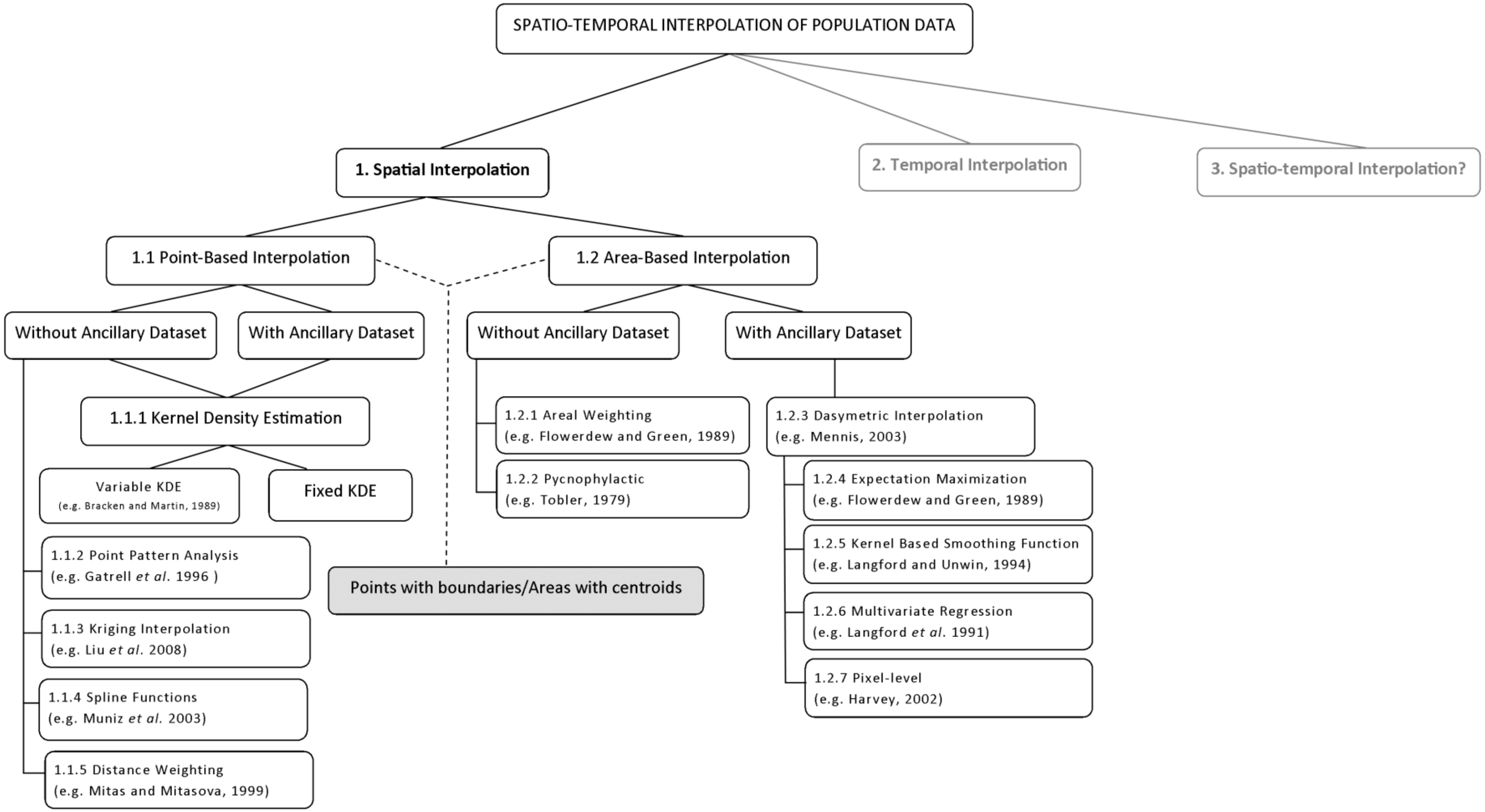


Figure 2.7 A classification of point and area based population interpolation techniques

2.4.2 **Limitations of zonal data**

Zones, such as wards (the lowest areal units for local political representation in England and Wales), are a central problem associated with handling spatial data. Zones where data are available are referred to as source zones. A compilation of various datasets may provide an array of different source zones. For example, population data published in neighbourhood administrative boundaries are not directly comparable to data published within local health or education authority boundaries. The units in which data are required are termed target zones (Mugglin and Carlin 1998). Interpolating data on to a grid, as an example target zone, from an array of different source zones provides a mechanism to represent data using a common structure (Figure 2.8).

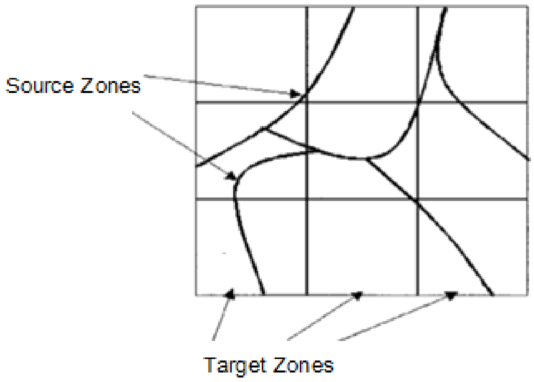


Figure 2.8 Incompatible source and target zones adapted from Gotway and Young (2002)

Zonal representation can become a challenging problem for policy makers and geographers, although for many it may be favoured as a convenient way to map and collate data. The scale threshold is one limiting factor of the zonal approach. The threshold, the scale variance, quantifies the change in scale where the phenomenon observed stops being invariable. Below the scale threshold data representation can be lost. Secondly, the modifiable areal unit problem (MAUP) (Openshaw 1984) is the phenomenon whereby the choice of zonal boundaries for aggregated data can have more of an impact on the output analysis than the real phenomenon observed. The areal units to which census data are aggregated do not remain constant between successive censuses (Openshaw 1984). “The areal units used in many geographical studies are arbitrary, modifiable, and subject to the whims and fancies of whoever is doing, or did the aggregating" (Openshaw 1984 p. 3). A large number of spatial objects can be defined, but few are sets of non-modifiable units (such as people or households in a census). Census output geographies for example have little intrinsic geographical meaning, but it is very likely that future analyses will depend on their definitions (Openshaw 1984). For this reason many methodologies have been utilised (Figure 2.7) to interpolate such data. Areal interpolation of population data has been driven by the demand for small-area population estimates which are often finer than the resolution of data available (Mennis and Hultgren 2006).

In the UK, detailed socioeconomic and census data are not released at a resolution required in point pattern analysis, but aggregated into irregular spatial units. The protection of UK census data is strictly governed by legislation including the *Public Records Act 1958, Census Act 1920* and the *Census (confidentiality) Act 1991* which concern confidentiality and disclosure of personal data. Data can also be aggregated for management and analysis reasons as well as privacy. However, an unintended consequence of fully disaggregated data from large datasets, where available, is that it may make visualisation and analysis unmanageable (Thurstain-Goodwin 2003).

Areal units are often not constant with respect to time, and consequently zonal data in their raw form can become incompatible. The shire is no longer the principal unit for British population data. The decision to disseminate 2001 UK census data in new output areas (OAs), not the enumeration districts (EDs) by which it was collected, resulted in little commonality between the 1991 and 2001 censuses (Langford 2007). Direct comparisons are not possible in areas affected by spatial unit boundary changes.

Spatial interpolation refers to estimation of the value of a variable z (e.g. population) at a location (x, y) given that the variable is known at a number of other data points, which may be randomly scattered (Goodchild and Lam 1980). The ability of basic functions in most GIS applications makes it possible to represent populations as a raster surface which can be highly advantageous. These methods provide a useful tool set for dealing with aggregated outputs.

A common interpolation target zone for these methods is the grid (Figure 2.8). The mapping of populations onto regular geographical grids, as an alternative to irregular spatial units, has a long history, and several advantages. One is stability through time, as data relating to different dates can be compared on a consistent grid (Martin *et al.* 2011). The use of a regular grid also allows integration with other georeferenced datasets from a range of applications covering physical and social characteristics (Martin and Bracken 1993; Martin *et al.* 2011). Population density can be more realistically represented in a uniform grid covering the entire region than conventional choropleth maps (Mennis 2003). The grid also allows for cells to be assigned a zero value where there is no population.

2.4.3 **Point-based interpolation**

Point-based interpolation methods can be categorised as either local or global depending on whether all values are considered at once, or individually within the pre-defined neighbourhood of each point (Wu *et al.* 2005). A number of point based methods have been described, and example applications are provided in Figure 2.7. Point data may be considered without the need for digital boundary data, or when such data is unavailable. The use of point data in publicly available demographic datasets such as censuses can allow easy interpolation into gridded approximations. The desire to transform population data onto a regular grid or into varying spatial units often reflects a desire not to be constrained by the arbitrary or spatially irregular geographic units for which data are initially available.

2.4.4 **Kernel density estimation**

Kernel Density Estimation (KDE) (Figure 2.7) is one method for transforming population point data onto a regular geographical grid. The point based method proposed by Bracken and Martin (1989) is an example of a redistribution algorithm which utilises population centroids, where each centroid corresponds to an irregularly shaped areal unit. Their approach redistributes the total population from the centroid locations with which it is initially associated into the cells of a regular grid. The model uses a distance-decay function to give the most probable distribution surrounding each centroid. A search radius, or kernel size, is also specified. The distance-decay function is used to evaluate the probability of each cell within the kernel receiving some proportion of the centroid's total population. As a result, no cell outside of the kernel will receive a population.

The general form of the model can be described as (Bracken and Martin 1989):

$$\sum_{i}^{s}\sum_{j}^{c} P_i = P_j W_{ij} \tag{2.4}$$

Where P_i, is the population in cell *i* of the output grid with dimension s, P_j is the empirical population of the *j*th centroid, *c* is the number of centroids in the area modelled. W_{ij} is the weighting of cell *i* relative to centroid *j*, which can be defined as:

$$W_{ij} = f_j \left[1 - \frac{d_{ij}}{\sum_{l=1}^{n_j} \bar{d}_{jl}}\right], \text{for } j \neq l \tag{2.5}$$

Where d_{ij} is the distance between cell *i* and centroid *j*. Therefore $\bar{d}_{jl}$ is the mean distance between all centroids (*l*) within the search radius from centroid *j*. Finally, f_j is a distance-decay function relating to the dispersion of the population within the search radius, centred on centroid *j* (Bracken and Martin 1989). Using a model to redistribute a population from a centroid into a raster grid allows for some cells to remain empty, representing an unpopulated area. The model is based upon the assumptions that: i). the centroid is a summary value for the population zone to which it relates and defines a point with an above average population density. ii). the population of a given centroid is redistributed according to a distance-decay function with a finite extent, and iii). regions can exist within the study area where there is no population (Martin 1989). Centroids constrained by their respective boundaries can be termed as points with boundaries in the conceptual framework; or *vice versa* areas with points (Figure 2.7).

Bracken and Martin's (1989) KDE model is one point-based method widely used in the UK with census data. In the UK census population weighted centroids (PWCs), produced by the Office for National Statistics (ONS), act as summary points for the population derived from their respective output geography. Ancillary data such as output geography boundaries, residential areas digitised from maps, or the use of remote sensing can be used to constrain the model output. (Martin 1989). To date, the 1971 census was the only time that population data was aggregated into 1 km grid squares for the whole of Great Britain. The use of a grid has the considerable advantage that the grid squares remain unchanged throughout time (CRC *et al.* 1980).

Currently, aggregate census data with centroids for the UK are openly accessible online. Datasets contain the population count and British National Grid (BNG) reference for centroids within EDs from 1971-1991, and OAs 2001 to present. In many countries including England and Wales census outputs are in the form of irregular zones, except for Northern Ireland where this is in addition to regular grids of cell sizes 100 m and 1 km. Using UK census data it has been possible to validate the surface population model outputs in raster format with a cell size as small as 50 m (Bracken and Martin 1989).

A population model can enable socioeconomic analysis. This analysis between decadal censuses is not straightforward, partly because of incompatible census geographies and the format and nature of the data (Bracken 1995). In 1991 more stringent confidentiality limits, relative to the 1981 census, meant that more EDs were restricted. Data in the 1991 small area statistics (SASs) were restricted therefore EDs were merged with neighbouring areas until a given threshold was met.

There are other sources of population centroids available that can be utilised in a similar way. A further example is unit postcodes (UPCs), which are well suited as an address based geography to map many different datasets at a high resolution. Designed for the efficient delivery of post by Royal Mail, UPCs contain around 14-17 properties on a delivery round. In the case of the CASA Town Centres Project land use indicators were georeferenced according to their UPC and transformed using a KDE (Thurstain-Goodwin 2003). Additionally, business addresses with UPCs from the Annual Business Inquiry (ABI) were georeferenced, and the workforce population represented as a point (Lloyd *et al.* 2003).

A further example of KDE was used in the Town Centres Project to produce a population surface on a regular grid. This spreads out the data from each point across the surrounding area. Town centres are difficult to delineate because of their indeterminate nature. If defined they could be treated as an object in conventional GIS software (Thurstain-Goodwin and Unwin 2000). In the Town Centres project, a kernel of a specified size is passed over the points to smooth the data by allocating a proportion to each grid cell. Cells receive fewer people with increasing distance away from the point. A kernel size of 200-300 m was proposed based upon surveys suggesting that this is as far as people are willing to walk in a town centre (Lloyd *et al.* 2003). The KDE was used to create continuous surface representations of four key factors proposed to characterise a town centre. Surfaces generated for economy, property, diversity of use and visitor attractions were combined to into an ‘intensity of town centredness’ surface. Analysis of peaks on the composite surface were then used to delineate town centres (Thurstain-Goodwin and Unwin 2000).

2.4.5 **Point pattern analysis**

Surface population density estimation can be applied to point data to suit a range of applications. Point pattern analysis (Figure 2.7) provides one method to attempt to transform a series of point data into a continuous density surface. The spacing of points, such as the mean distance to the nearest neighbour is one method of analysis. Secondly, point data can be analysed as a frequency in a regularly defined subgroup such as quadrats. Imposing a fixed grid such as a quadrat is an alternative to imposing a moving kernel. However, the use of quadrats or a fixed grid often produces a less smooth estimation than kernel estimation. These methodologies can broadly be categorised as distance and area-based analyses respectively (Gatrell *et al.* 1996). An advantage to using point pattern analysis is that it is readily available in most GIS software. The process depends on point, or disaggregated data, unlike a KDE that can redistribute data that have already been aggregated to points such as output area centroids or unit postcodes.

In a UK application Gatrell (1994) suggested the use of postcodes representing a number of houses that could collectively be treated as point data with geographical coordinates. Georeferenced individual addresses for Great Britain are now commercially available (Ordance Survey 2012). The current situation allows data based upon an address to be georeferenced and therefore can be treated as a point.

2.4.6 **Kriging**

Kriging is a geostatistical method utilizing the auto-correlation of different attributes (Liu *et al.* 2008) (Figure 2.7). Liu *et al.* (2008) demonstrate area to point kriging interpolation of population data as a methodology to disaggregate census data. Additional data are not required, apart from what are needed for the regression model. The kriging method interpolates the residuals from regression. Accuracy improvements in population densities are accounted for by the location dependence and spatial correlation of the residual population density. Thus the process is two-step. Firstly, regression based estimates need to be derived with corresponding residuals. Secondly, the residuals are interpolated in space by the area-point kriging function (Liu *et al.* 2008).

Kyriakidis (2004) defines a geostatistical framework for area to point interpolation as a special case of kriging. Census tracts and socioeconomic data often need to be downscaled for the purpose of detailed modelling and a methodology is provided under the framework using kriging. Unlike the KDE, point data derived from area-point interpolation need not lie on a regular grid or create a surface. Kriging is a local estimation methodology that provides the best linear unbiased estimator for an unknown quantity, regarding minimum estimation variance (Journel and Huijbregts 1978). However the application of kriging can be more difficult relative to other weighting methods that produce similar results. Furthermore, simple kriging is a linear estimation. It is unlikely that population density against distance away from a point is a linear relationship. If structural information is available regarding a second-order relationship then other non-linear techniques should be considered (Journel and Huijbregts 1978).

2.4.7 **Spline functions**

A spline, a polynomial function, can be applied as an alternative to a standard exponential density function, and is a further example of a point based methodology (Figure 2.7). Exponential functions by definition assume population density decreases with distance away from the town centre. Muniz *et al.* (2003) suggest that 'density craters' in city centres, greenbelts, satellite cities and dense peripheries cannot be replicated by a standard exponential function. A residential density function relates population density with distance from a city, therefore, allowing density to be predicted at a given distance from a city centre. This provides a method to describe the structure and spatial distribution of population (Muniz *et al.* 2003). The density gradient of an exponential function is constant, whereas it is variable in a spline giving the proportion of density variation per unit of length. The spline function requires a fixed distance relative to an origin such as the central business district. This method could be applied to medium sized cities and metropolitan areas using points of population data (Muniz *et al.* 2003). The method might not be appropriate for population centres or small zones outside of urban areas away from the assumed density/distance relationship. An example case could be a large urban conglomeration within distinct satellite settlements. However, a KDE could be applied to all population points in turn, which would preserve the underlying data structure and observed spatial distribution of population.

2.4.8 **Distance-weighting**

Inverse distance weighting (Figure 2.7) is a simple and common approach for population interpolation. Unlike the redistributions methods discussed (e.g. KDE) interpolation assumes that a point value falls between two other known values. Inverse distance weighting (IDW) assumes values for non-sampled locations based upon the value of surrounding points at a specified distance (Mitas and Mitasova 1999). It can be a particularly useful tool for spatially representing population density, as well as many other physical and socioeconomic phenomena on a regular grid. This process is readily available in most GIS packages. Rase (2001) extends this concept for interpolation onto an irregular triangular network (TIN) rather than the traditional orthogonal grid of equidistant lines. He suggests that error is minimised by interpolating irregular lines and points to a TIN rather than regular grid. In this process nodes are calculated to populate a TIN using an inverse distance weighting:

$$zn_i = \frac{\sum_{j=1}^{m} z_j \times d_j^{-p}}{\sum_{j=1}^{m} d_j^{-p}} \tag{2.6}$$

Where zn_i is the new value for point i, z_j is the value of m nearest neighbours, d_j is the distance to m nearest neighbours and p is the exponent of the distance.

However IDW can be criticised for producing local peaks in the data surface that do not conform to the shape implied by the original data (Mitas and Mitasova 1999). In contrast, KDE also has a specified distance of interpolation around a point, determined by the kernel width. However, a KDE applies a specified function to each centroid in order to try and realistically redistribute a point's population. IDW is appropriate for applications concerning point measurements, such as inferring values on a digital elevation model. However, population density is a reference interval function (Nordbeck and Rystedt 1970) which is only measurable with reference to an interval or area. Interpolating population data would artificially increase the density.

2.4.9 **Area-based interpolation**

The second group of methods for spatial interpolation of population data in Figure 2.7 correspond to area-based input data. The following sections outline alternative methods for estimating a population density grid using area-based data.

2.4.10 **Areal weighting**

Areal interpolation, the simplest technique, in its standard form is based on weighting by area and can be utilised where incompatible spatial units are concerned. The weights are calculated by the proportion of the target zone overlapping the source zone, therefore allowing data to be transferred from one reporting zone to another. This methodology provides a solution where ancillary information on distribution is not available. However, the weighted distributions are probably unlikely to be found in reality (Flowerdew and Green 1989).

Goodchild *et al.* (1993) propose a GIS based framework and example application for the areal interpolation of population data, citing applications relevant to census centroids in the USA, Canada and UK. County level population counts from census data for California are compared with major river basin boundaries following watersheds. An areal interpolation method is applied to transfer population, employment and socioeconomic data from county level onto the hydrological boundaries containing data on water consumption and availability.

2.4.11 **Pycnophylactic interpolation**

Tobler (1979) also proposes interpolating values from data given in arbitrary geographical units onto a regular grid, at a specified resolution. The grid resolution needs to be fine enough to have at least one or more points in each geographical unit of the input data, therefore preserving the smallest geography. A population density surface can be smoothed towards the edges of the aggregated boundary. An important attribute of Tobler's (1979) area-based pycnophylactic interpolation methodology for generating a population surface estimation is volume preservation. Pycnophylactic is a derived Greek term for mass preservation. This example can be described as 'pointless' interpolation, which can be used where data are available in aggregated units, not points. This is an alternative to using centroids, or where they may not be available. This is exemplified through the representation of population density by US states (Figure 2.9). Smooth contour maps satisfying volume preservation and non-negativity can be used to convert different output geographies for comparison (Tobler 1979). The volume preservation of population is crucial as this ensures that people cannot be created or destroyed.

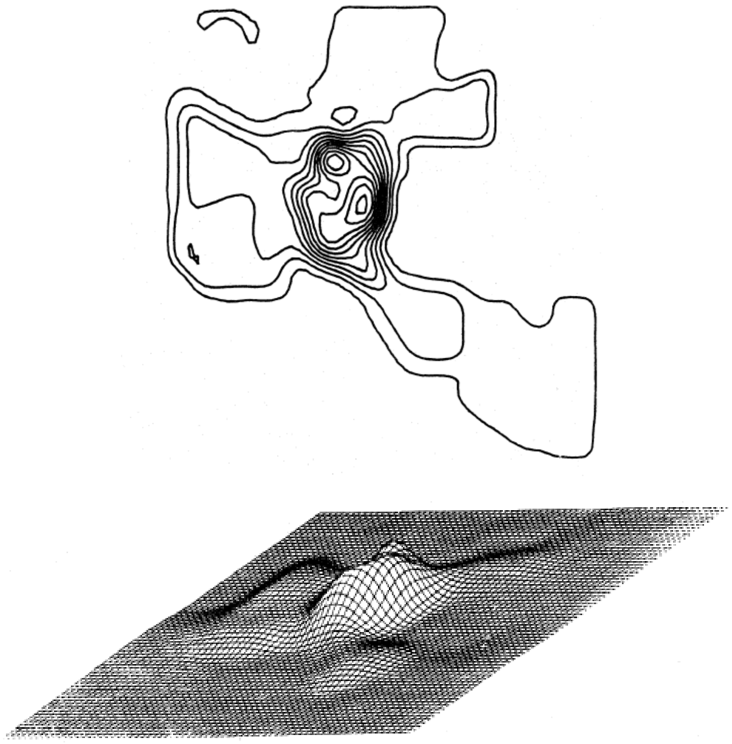


Figure 2.9 Contours and isometric rendering of US state population densities (Tobler 1979)

Tobler *et al.* (1997) report on a pycnophylactic based interpolation methodology to create a raster model of global population, without the need for constraining and changing political and national boundaries. The rationale for the project is based on the availability of satellite imagery for scientific studies concerning impacts on the spread of people in the absence of national borders. A pycnophylactic approach was used to transpose global population data onto five minute longitude/latitude quadrilaterals, approximately a 9.3 km resolution at the equator.

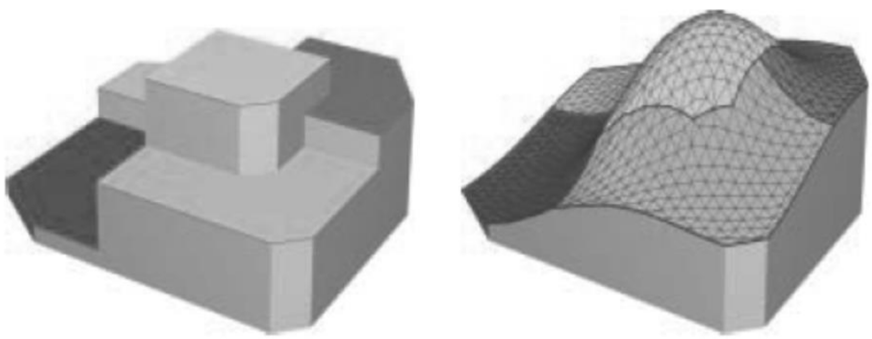


Figure 2.10 Population surface interpolation from areal data (choropleth, left) onto a regular grid (right) (Rase 2001)

Rase (2001) demonstrates a pycnophylactic approach to interpolate a smooth, volume preserving surface from areal, or polygon, data (Figure 2.10). He also raises a critique of point based interpolation methods where centroids are considered without their polygon boundaries, and thus not preserving volume.

2.4.12 **Dasymetric mapping**

Dasymetric mapping is one method which has arisen from the inadequacies noted in earlier examples. Dasymetric mapping as an alternative (Figure 2.8) has been available for many years, but only more recently used with the rise of geographic information improvements in computing power (Langford and Unwin 1994). Wright (1936) popularised dasymetric mapping in the United States but is often wrongly accredited as its inventor, as an earlier reference can be made to the Russian cartographer Semenov Tian-Shansky in 1922 (Mennis and Hultgren 2006).

Dasymetric methodology involves disaggregating spatial data to finer units, often using ancillary datasets. As a result the process produces areas of homogeneity in the mapped data that more closely resemble the scenario being modelled (Maantay *et al.* 2007). Wright (1936) contrasted the dasymetric approach against conventional choropleth mapping by reference to Cape Cod, Massachusetts, USA. Conventional choropleth maps assigned a single population density to large expanses of the cape that were uninhabitable, and only occasionally frequented by backpackers, wardens or hunters. Defining uninhabited is not straightforward. On a high resolution map, gardens may well be considered uninhabited. However, if regions that people sometimes cross or visit should be considered inhabited, then this would apply to the remote sand dunes and moraines of Cape Cod that are occasionally frequented. The basic principle of dasymetric mapping involves distributing populations to the 'inhabited' areas. Unlike choropleth maps, vast areas of water or uninhabited countryside therefore will not receive a population density. In fact the permanent inhabitants of Cape Cod resided in a small compact village near the harbour. The choropleth approach assigned population densities of between 3 and 77 people per km^2 to vast 'uninhabited' areas of the Cape according to Wright's analysis of 1930s data.

More recently, Wright's technique has been applied to dasymetric population interpolation utilising remote sensing data. Langford (2007) postulates a straightforward methodology based upon pixel values: the colours used in Ordnance Survey (OS) maps. In general the OS use the same 8-bit colour palette for all raster pixel maps. This enables the process to be consistently repeated for neighbouring tiles, until the desired area is covered. For example the value of pixels visually containing the brown shade representing buildings could be extracted and used as a mask in a dasymetric interpolation.

Similarly to pixel-level interpolation for population density estimation, an expectation maximisation algorithm can be applied to pixels to estimate population. The estimate can be derived by iteratively regressing pixel values based on spectral properties. This approach can be used to compute values where data are missing or restricted. The algorithm consists of two steps (Flowerdew and Green 1989):

- Expectation – Values are computed based upon the conditional expectation of the given dataset.
- Maximization – Fit the derived data to the model as the maximum likelihood that the values are from real observations.

Harvey (2002) proposes expectation maximisation in the absence of ground reference data for pixel population values in a remote sensing application. Firstly, pixels are classified as residential and non-residential. Initial estimates of residential pixel populations are then iteratively refined. Although results indicate a high level of accuracy, new population estimation would require extensive calibrating and ground referencing. Without adaptation the iterative model cannot be uniformly applied over time and space, even on a national scale (Harvey 2002). The addition of ancillary layers such as building height have been used to estimate population at different levels. For example Aubrecht *et al.* (2009) use airborne laser scanning data to estimate building heights in part of Austria. This in turn is combined with population data to produce a highly detailed 3D building model that correctly estimates residential population density within multi-storey buildings.

An alternative method of population density estimation utilises satellite imagery and the spectral properties of individual pixels. The LandScan Global Population project combined census data, land cover, night-time illumination and information on topography to produce a global population estimate at 30 arc-sec (c. 1 km at the equator) resolution (Dobson *et al.* 2000).

The amount of change in a dependent variable for a given change in an independent variable can be indicated by a regression equation (Johnston 1978). Multivariate regression estimates a single regression model with more than one outcome variable. For example, Langford *et al.* (1991) compress Landsat TM land cover classifications into five categories for analysis. Pixel counts for digitised UK census wards for each of the five classifications were completed using GIS.

A dasymetric application to census data can use a 'grid three class' method to weight areas of census output depending on land use: urban, agricultural and forested in order to receive a proportion of the aggregated population (Mennis 2003). A potential disadvantage of the dasymetric approach using remotely sensed data is the misclassification of industrial areas as residential due to their similar appearance or spectral properties (Mennis 2003). However, dasymetric mapping can be implemented in most raster-based GIS packages without the need for programming and is adaptable to suit a number of user requirements (Mennis 2003).

The growth of this methodology has been sustained as most publicly available demographic datasets such as censuses are aggregated to areal units. Data are commonly aggregated into areal units defined by statutory or administrative boundaries (Mennis 2003). However, problems arise in the display of such aggregated demographic data such as the display of highly detailed spatial attributes contained within a census.

A kernel based smoothing function (Figure 2.7) can be combined with a dasymetric methodology to utilise additional data (Wu *et al.* 2005). A close approximation to a continuous population grid can be achieved using a moving kernel. The aggregate population is divided by the area of the kernel and applied to the central cell to estimate the new density. This process is repeated until every pixel has been covered (KDE, Section 2.4.4). The effect of changing the search radius or kernel size alters the final appearance. The larger the radius the smoother the output appears. Although the process is similar to the KDE the method is applied to area-based dasymetric data that has been pre-processed rather than the raw population centroids.

Langford and Unwin (1994) apply the kernel function to a dasymetric population density estimation. The dasymetric estimation is derived by defining residential housing in UK census units using Landsat satellite imagery integrated within GIS. Land classification was used to designate all other land use types unoccupied except the pixels containing some form of residential housing. Global classification accuracy is estimated at 85-95%. However, whether a residential pixel remains unoccupied remains uncertain (Langford and Unwin 1994).

If little attention is given to the effect of areal units when presenting data on population density large discrepancies in understanding can occur. The average population density of UK cities ranges from 1000 to 6000 people per km^2. When uninhabited locations such as parks, gardens and industrial sites are taken into account the density increase is much larger. A dasymetric method produces more realistic population densities for parts of Glasgow of up to 50000 km^2 (Langford and Unwin 1994). Populations can be assigned to the residential area of wards instead of being uniformly distributed across the entire area based on ancillary datasets such as a land use classification. This technique can also be used with a KDE. A close approximation to a continuous population surface can be derived by using the 'floating window/grid' procedure (Langford and Unwin 1994). The aggregated population is divided by the area of the kernel and the result assigned to the centre cell, the process is then repeated, shifted by one pixel each time. This technique could be carried out using the neighbourhood or spatial filtering functions in a raster based GIS.

2.4.13 **Alternative representation of population data**

This subsection introduces alternative and emerging techniques for novel and non-census based population data representation. It provides an overview on the application of data that concern building levels, location based social media and geolocated mobile telephone records.

Wu *et al.* (2008) disaggregate US Census data onto individual buildings. This example of building level data representation demonstrates a great improvement from simple aggregate census data. Their approach requires the use of ancillary building footprint data, from which a volume is created within a GIS. This is coupled with occupancy data in a method to re-weight census population onto buildings within their respective areal units. Similar to the Aubrecht *et al.* (2009) approach (see Section 2.4.12) building height data is derived from LiDAR surveys. However, Aubrecht *et al.* (2009) specifically create a 3D building model with a European implementation. The creation of accurate building level population data is identified in both examples as having direct relevance for natural hazard planning. This is due to the spatial improvements achieved compared to traditional aggregate datasets.

The advancement of population data representation is continuing with emerging novel methods. One example is the use of location-based data from the social media site Twitter. Since its inception in 2007 half a billion tweets (images or 140 character messages, often georeferenced) are sent per day, and 80% of these originate from mobile telephones (Twitter 2015). Mapping the density and distribution of georeferenced tweets has been demonstrated for London by Hudson-Smith (2014). One of the features is the ability to resolve major tourist sites and transportation hubs (e.g. Heathrow Airport and railway termini).

Longley *et al.* (2015) also identify the significance of using Twitter social media data to move away from traditional 'night-time' geodemographic datasets. Location-based tweets have a high spatiotemporal resolution in near real time as each contain a set of coordinates and timestamp. They demonstrate the ability to resolve patterns of very short term population movements at a high spatiotemporal granularity. However, two limitations of their approach identified were the lack of uniform population coverage and the restricted nature of user demographic data.

The WorldPop project (www.worldpop.org.uk), expanding original implementations for the East Africa region, created national scale 100 m resolution static population estimates. These were constructed on the basis of reweighting census data taken at the smallest output units onto settlements locations. Settlement locations were derived from a combination of Landsat satellite imagery and land cover data (Tatem *et al.* 2007). This provides an example of a valuable, large coverage, high resolution disaggregation technique, albeit with a static population.

Finally, the use of anonymised mobile telephone handset records provide an opportunity to map temporally varying populations without census data. In some applications mobile telephone location data provide the opportunity to map actually observed population distributions without the need for further modelling or disaggregation. Every mobile handset has a unique identifier that is recorded within a database to specific cells. Cell registration data can be used to triangulate the location of individual handsets. Successful novel studies include applications in Portugal, France (Deville *et al.* 2014) and Estonia (Ahas *et al.* 2010). However, mobile telephone data is subject to the same limitations on user demographic information and coverage of the whole population as social media Twitter data. Mobile telephone data is not currently widely available, or restricted to sensitive commercial retail analytical applications (e.g. Smart Steps, dynamicinsights.telefonica.com).

2.4.14 **Relating population mapping and risk exposure**

Like hazards, population also varies considerably in space and time which has a large impact on sunsequent risk analyses. Natural hazard and exposure data are combined using a dasymetric approach to assess populations at risk by Chen *et al.* (2004). Hazard attributes such as intensity and footprint are usually available in a raster layer. However, exposure data such as population, business and dwellings are usually only available in aggregated census units.

Similar to the critique of choropleth maps, the aggregated census output can be incorrectly assumed to be uniformly at risk. Figure 2.11 identifies the spatial extent of a given hazard within a postcode. Chen *et al.* (2004) argue that using the hazard intensity at the centroid could be misleading. If the centroid is determined by postcode boundaries in an irregular shape it may not reflect the actual intensity of the hazard experienced. For example, applying the hazard intensity at the centroid for the scenario in Figure 2.11D suggests that it is outside the hazard footprint, although the residential area is not. An improvement in assessing the spatial distribution of risk is omitting the unoccupied areas within the postcode, or areal unit. Chen *et al.* (2004) used a dasymetric approach to try and differentiate residential areas. Firstly, urban road networks in the form of lines within a GIS based map were buffered by 100 m assuming the surrounding area would be inhabited. Data in layers containing features such as parks and water bodies were also used to define the limits of residential areas.

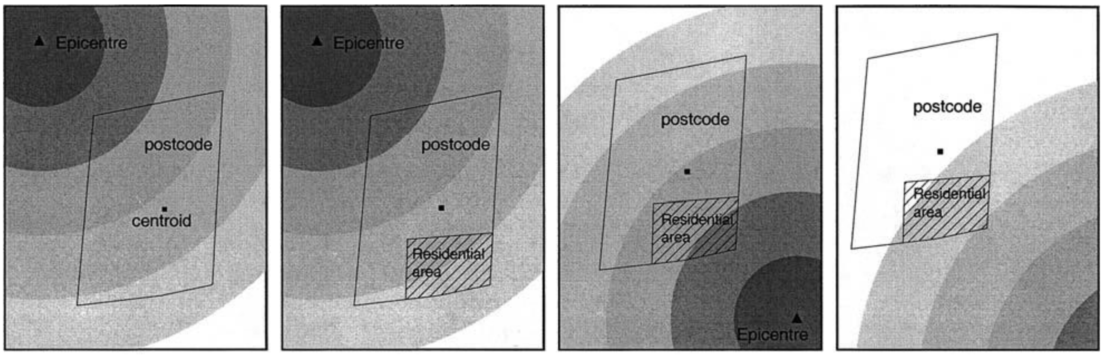


Figure 2.11 Illustration of hazard intensity to an exposed population aggregated into postcode units (left to right: A-D) (Chen *et al.* 2004)

Improvements suggested by this method, likely to be welcomed by insurance industries, include more accurate loss estimation by better representation of populations at risk. The dasymetric approach suggested by Chen *et al.* (2004) aims to outline the area exposed to a particular hazard. Simply using aggregated data in areal units could overestimate exposure. In a case study for Sydney, Australia only 22% of the region covering 8879 km^2 has been identified as residential using the approach suggested by Chen *et al.* (2004). The schematic representation in Figures 2.11B and C shows how hazard exposure would be underestimated if census output units or postcode districts were relied upon. There is still some concern by selecting a somewhat arbitrary 100 m buffer for the road network where residential buildings are assumed to be. However, it does appear to be a far closer representation of reality than the aggregated postcode districts.

Overall, it appears that the dasymetric approach to mapping offers advantages over conventional choropleth maps. In particular it provides more realistic demographic analysis as populations are not distributed to uninhabited regions. It also allows improvements in representing aggregate census data. Conversely, uncertainties can arise from land cover classification, in particular what can be regarded as ‘uninhabited’. Similar to critiques of the arbitrary nature of some census output units, there is a degree of arbitrary decision making involved in some of the dasymetric methodologies discussed.

For risk mitigation accurate and timely population maps are required for exposure assessments. Mobile phone data have been used to estimate population movements in space and time for strategic malaria elimination planning in Namibia by Tatem *et al.* (2014). This dynamic approach identified key settlements with higher than average travelling populations in specific at risk.

Day and night-time human populations vary widely, and often most in metropolitan areas. Freire (2010) notes that accounting for the spatiotemporal distribution of population at the local scale for risk analyses is fundamental and often misrepresented by census data alone. For example human susceptibility to contemporary tsunami risk in Lisbon, Portugal is modelled by Freire *et al.* (2011). They use a model to create high-resolution (25 m) daytime population estimates. A dasymetic approach is also adopted to disaggregate census (including commuting statistics) and employment data with road network and land use classifications. They apply three scenarios representing a toxic plume, earthquake and shopping terrorist attack. It is observed that traditional census datasets under-represent exposed populations in some densely populated metropolitan areas. This occurs where there is a large increase in daytime population density. The simulation of daytime population density improves the assessment of exposure and can contribute to better emergency planning (Freire *et al.* 2011).

It has also been demonstrated that projected future population estimates can be applied to hazards with the potential to occur over much wider (national) scales. For example heat stress associated with climate variability, can affect vulnerable populations over large areas. Aubrecht *et al.* (2012a) create 1 km spatiotemporal population estimates using census and ancillary datasets. These cover a north-south European transect with population age. The effect of heat waves on an ageing population, projected to 2030, who are more vulnerable to heat-related illness was considered. Aubrecht *et al.* (2012a) maintain that a central objective of vulnerability assessment is to provide indications where, when and how people may be affected by a specific impact.

Despite censuses typically being accurate, geographically refined and current, they only represent night-time residential population counts. Therefore, censuses can be misleading when trying to locate and quantify population exposure to a daytime disaster (Garb *et al.* 2007). Population data utilised for the purposes of hazard risk assessment can be prepared in advanced and kept updated.

2.5 Interpolation of population data with time

Human population distribution is a function of both space and time, but the view that better temporal population estimates are required has been around for some time (e.g. Wright 1936; Foley 1954; Schmitt 1956). However, until relatively recently interpolation methods, as discussed in the previous sections and illustrated in Figure 2.7, have largely focused only on the spatial component. Figure 2.12 expands on the integration of the temporal component in the interpolation of population data. The interpolation of such data is recognised in a wide range of applications including security, emergency planning, transportation and healthcare provision. It is recognised that current risk models do not sufficiently consider the temporal variation in population distribution, whereby the spatial distribution of populations at risk varies greatly by time of day, day of week, time of year. The time of occurrence of a rapid onset event will have a different outcome for those exposed depending on diurnal and seasonal population variations. The risk to population sub-groups is also dependant on demographic characteristics. Further analysis of this data is required to examine this relationship. For example, students in higher education are a highly spatially mobile section of the population, and at time clustered at locations, with strategic importance (King and Ruiz-Gelices 2003). In 2011/12 the UK had 2.5 million students registered in higher education (HESA 2013) comprising around 4% of the total population.

Risk models can be developed when the causes and consequences of a given event are known, and can be used for emergency preparedness planning, such as accounting for the number of inhabitants requiring evacuation (Ahola *et al.* 2007). This section outlines representation of populations in time-space and the role of time-geography. The concept and examples of the spatiotemporal interpolation of population data are provided in Section 2.6, however few practical examples exist.

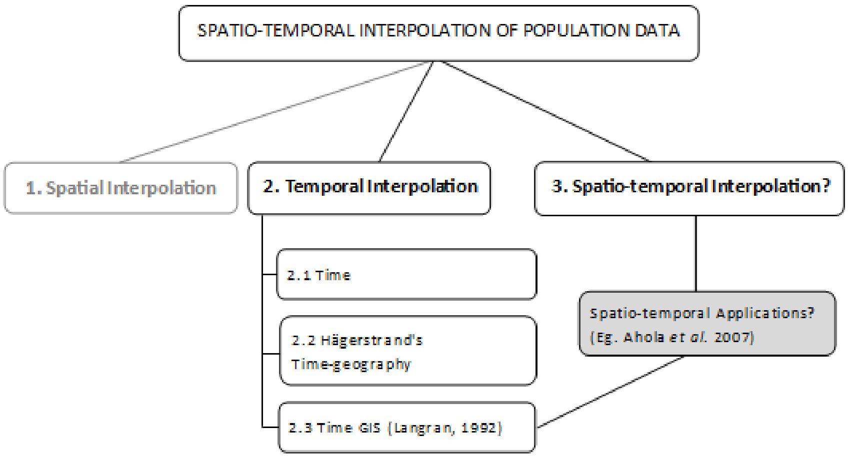


Figure 2.12 Proposed structure to evaluate population interpolation methods

2.5.1 Time

In order to understand the temporal interpolation of population data and the reasons and methodologies behind them it is important to take time to understand the brief concept of temporality itself. The section only attempts to briefly highlight the history and perception of time and the resultant consequences for human reality. A large literature already exists on the fundamentals and quantum physics of time which is not revisited in detail for this summary (Denbigh 1981; e.g. Davies 1995; and Hawkin 1995)

Time and space create and constrain the very fabric of human interaction (Raper 2000). However, the integration of time into our representation of population becomes difficult as we can only 'see' change when motion or rapid movement occur (Raper *et al.* 2005). The concept of motion is not possible without considering time. Today's culture views time as a line with no end points extending infinitely into the past and future (Langran 1992). The opinion that time can be considered an illusion is an alternative argument in philosophy. Objects including humans form a static pattern in a four dimensional space-time block. It is argued that time only feels as if it is passing because our memories retain the past, as opposed to remembering the future and the present. What is actually remembered at the present is what has just passed (Langran 1992).

Time is always expressed relative to something. There are a number of anthropogenic ways in which time and duration can be expressed and referenced. For example, in Madagascar ‘rice cooking’ and ‘the frying of a locust’ traditionally represented half an hour and a moment respectively (Sorokin and Merton 1937). The ancient Mayan civilisation calendar consisting of 13 *b’ak’tun*, periods of 144,000 days began in 3,114 BC, and along with the ancient Egyptians monitoring the rise and fall of the Nile, these are some of the earliest known examples of time reference systems.

There are many temporal constructions used in the English language, also believed to be universal across all languages. Examples include ‘the weekend is *coming*’, ‘I am *going to* get up early tomorrow’ and ‘May is *before* June’. However research suggests that tribal Amondawa speakers of the remote Amazon living in the Uru-eu-wau-wau reservation now located in Brazil, lack a lexicon containing any temporal definition (Shina *et al.* 2011). The research suggests that it is possible for no numeric or calendar system to exist.

Before official contact in 1986, the Amondawa population was believed to be around 160. The population rapidly decreased on contact to just 45 by 1991. The rapid decrease was caused by the spread of colds and other viruses in which the indigenous tribe had no immunity (Shina *et al.* 2011). This rare example of self-sufficiency and isolation perhaps demonstrates an example of space without time. However, time continues to constrain activities whether we are conscious of the fact or not. A more widespread time reference system is the current Gregorian calendar based on the motion of lunar and solar cycles, and encompassing the time units used today in Western societies.

2.5.2 **Time-geography**

The tradition of time-geography in human geography treats time and space as resources that enter directly into the realms of social life (Gregory 2000). The basic ideas were coined by the Swedish geographer Torsten Hägerstrand. Time-geography prominently came to attention during the 1970s, but references can be traced back into the 1960s. This was an era of self-review and realignment of the discipline of human geography. The success of time geography has been associated with providing an alternative to the ‘spatial science’ which resultantly promoted the study of individuals

(Per Olof 1991). Until then first-hand geographic observation of individuals and groups moving in their environment was rare (Hägerstrand 1973).

The notion of time geography partly evolved from the desire to understand more about society's interactions with nature. Hägerstrand described society as 'a peculiar mix of mental and physical structures, confusing and unpredictable, which we cannot survive without' (Hägerstrand 1976 p. 329). It was argued that geography could contribute in an ever more positive way if given the opportunity to contribute to prediction and planning. Hägerstrand defined bounded areas as a set of populations comprised of individuals who are described by continuous trajectories through time. Humans and their society are just a pattern in the big tapestry of nature, that history is weaving (Hägerstrand 1976). The human mosaic or tapestry can be the outcome of collateral processes, which cannot unfold freely as they have to accommodate the pressures and opportunities that lead from their common coexistence in terrestrial space and time.

Hägerstrand's geographical framework recognises that an individual's participation in an activity has both spatial and temporal dimensions. The space-time relationship emerges as a key definition for an object, individual, or population in order to attribute spatial context to a location, path or trajectory (Hägerstrand 1973). Examples can be drawn where nature and humans may struggle to coexist in the same space without mutual distortion, such as road networks across a fluvial plain liable to flood. If an area or space is 'mapped' over a period of time, the complete set of unbroken space-time paths, produced by all members of the population, are contained within a domain or pass through it. Humans are always seeking to reach goals, which can be considered as destinations. Common goals can be grouped into bundles to form steps to the final outcome. A project can be defined as a total cluster of activities, individuals and items which must participate in the reaching of some defined goal (Hägerstrand 1973) (Figure 2.13).

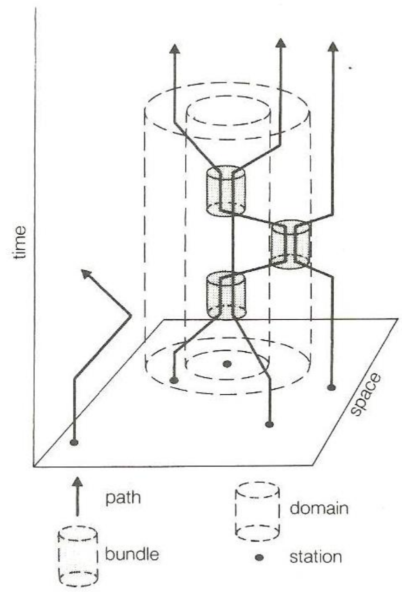


Figure 2.13 Hägerstrand's time-geography model (Cloke 1991)

The basic framework can be conceptualised across four propositions (Gregory 2000):

1. Space and time are resources upon which 'populations' have to act to reach projects.
2. Realization of a project by an individual is constrained by:
    a. Capability constraints: the limitation of individuals due to their own physical ability, or the ability to command the required facilities e.g. time-space paths flowing through accessible stations such as shops and schools.
    b. Coupling constraints (time-space bundles): governing how long individuals have to join other individuals or materials in order to produce, transect or consume.
    c. Authority constraints: steering mechanisms that impose conditions of access or movement through the time-space domain
3. The constraints are interactive and mark out the possible paths for individuals
4. Within the evolved structure, competition between projects for free paths is a central problem for analysis.

A very large number of trajectories exist across the environment, by which they are influenced (Hägerstrand 1975). 'Bundles' of trajectories give up individual degrees of freedom to allow others to be kept under control. A number of rules or institutions aim to manage conflict where capacity in the space makes the overlap of planned trajectories unavoidable. Traffic regulations and laws governing the ownership of land, tools and buildings exist in the world of barriers and prevent trajectories making certain turns, and let them move freely in one direction. Hägerstrand's Lund research group simply mapped urban areas as supply points such as shops and work places (Thrift 1977), which in turn were described according to opening hours and their location.

It is argued by Pred (1977), a pioneer of Hägerstrand's time-geography, that the concept has the potential to 'spill' into the other social and life sciences. This coincided with an identity crisis of human geography relying on models from other disciplines. The physical existence, life paths of individuals, goods, materials and other non-human populations can be traced in time. Time-geography can specify the necessary conditions for virtually all interactions between humans and the natural environment. The typical daily path of a Boston merchant trader is conceptualised by Pred (1984) (Figure 2.14). The time-space context is given and individual paths and shared interactions are represented in the form of meetings at the coffee and club houses (Pred 1984). The path taken through time and space is represented in Figure 2.14.

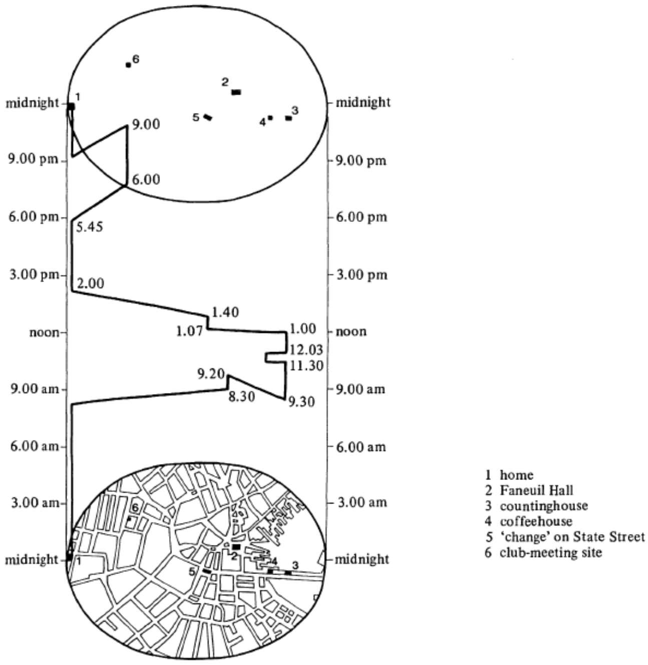


Figure 2.14 Representation of a Boston merchant trader's movement in time and space for a given day (Pred 1984)

Pred (1977) was concerned that the application of time-geography may only be considered for planning purposes. Several other applications within the discipline were highlighted. One such application concerned the study of regions and landscape evolution. The framework allows the classification of landscape components, that form part of an independent region, as a whole (Pred 1977). Innovation provides another application for time-geography. This use involves the building up of time-space trajectories of several people and inanimate objects, as well as information and energy, for example, agricultural or manufacturing processes. Time-geography may be applied to migration and urban growth. Insights into the mean migration-distance variation between different occupational groups can be considered by the spatiotemporal characteristics of activities and analysis of small bounded regions (Pred 1977).

Structuration theory generally recognises that social activities take the form of concrete interaction in space and time as actions and events making up an individual's temporal and spatial elements. This idea can be conceptualised diagrammatically as an unbroken path through space-time (Pred 1977) (Figure 2.14). There are a number of time-geographic realities that impact on the life content of human individuals (Pred 1978): a human is indivisible and cannot be in more than one location at the same physical point in time. Neither can a human, measured by tracing out trajectories, simultaneously partake in activities spatially separated at the same time. A person's space-time path through both dimensions makes up an individual's existence. Activities occur at a precise location in space, for a limited time period. An individual can sacrifice time for space to move to an activity occurring at a difference 'place'. Many activities such as working at the office are fixed in space. The faster someone or something travels the lower the gradient of their space-time path (Figure 2.14). This occurs as less time is sacrificed for more space (Miller 2005). Understanding these concepts is important as they provide key constraints in handling population data over time in a GIS or distribution models. An individual is not divisible, cannot be created or destroyed and can only be in one place at a 'time'. The notion that mobile populations are always moving between goals or destinations also must be upheld. The temporality is a larger concern when the time of day is considered, relative to the direction and movement of others within the same space-time block. For example, during 'rush hour' periods many people are competing at *a time* for the same path through space. This is evident in many urban areas as arterial road network or town centres. As a result Hägerstrand's notion of constraints is apparent in dealing with the conflict. Furthermore, this delays or restricts some people's movements. Thus, time-space paths, and their gradients are an important consideration for any time GIS or model.

Every task an individual can undertake is time demanding and therefore uses up time as a resource. Movement between two points spatially separated can only occur with the sacrifice of time. Every facility that provides a space that can be occupied, such as a piece of land has a limited packing capacity. This occurs as no two physical objects can occupy the same space at the same time (Pred 1981). Therefore a path and individual consume the space in which they exist. As a result of this concept and the notion that a finite space-time resource is indivisible it forces alterations in individual movements (Pred 1981). Activity bundles form where the paths of two or more individuals combine and their convergence is the 'station' or 'domain' (Figure 2.14). The concept of the space-time prism is similar to that of the path. Attending something such as an 'event' takes up time, a finite resource during a day, in a bounded area containing a population. However, an object's path is constrained by time. For example, if a person leaves and returns to work at two fixed times in a given interval, there is only so far in space that their path is permitted to go and this is termed the potential path space.

It is accepted that Hägerstrand's ideology is a useful descriptor of how the daily life of individuals unfolds in space and time, however it does not explain how domains and stations are produced or how meaning becomes associated with history, time, place and space (Harvey 1989). Following this critique, the argument has arisen that time-geography is too physical and mechanical and treats the individual as an object without thought, experience, feelings or expectations for the future (Lenntorp 1999).

An objection regarding time-geography is that it does not consider an individual as an acting subject. Five main critiques have been delineated (Per Olof 1991). Firstly, it is argued that the concept represents physicalism, the philosophy that only physical objects exist. Therefore it is argued that time-geography may be more suited to the physical rather than social sciences. This is a philosophical argument that things do not exist beyond their physical properties, therefore non-physical things do not exist. Secondly, it is argued that little consideration is given to social interaction, and individuals are considered as mere objects. Thirdly, humans or objects are not considered as acting components. Fourthly, human activity and social processes are not considered. Finally, time-geography is undeveloped and has an unproblematic perception of time. However, concepts of this approach are contained within the proposed methodology in this thesis. Population data are considered with respect to corresponding demographic information. Secondly population is considered as highly mobile and interactive. Here it is appropriate to treat population as a physical, socially complex asset in order to ascertain physical exposure estimates for hazard applications within the natural sciences.

2.5.3 **Time GIS**

There is a clear shift in the spatial distribution of population between night and day as people migrate between their usual residence and places of work and other activities. Currently, most available population datasets have no temporal component (McPherson and Brown 2004). Large errors in population data occur in daytime estimates as they do not consider travel to work, education, shopping and other actives between which people may migrate during the day. McPherson and Brown (2004) note that there are two primary methods for building population datasets: labour intensive demographic counts or employing the use of a GIS or remote sensing. Langran (1993) noted several practical issues which need to be considered while attempting to create a spatiotemporal database using a TGIS concerning representation, updates and longevity. While this section focuses on the temporal component of a GIS there is inherently some spatiotemporal discussion as traditionally GIS are spatial in nature and this section attempts to address the addition of a temporal component. Whether there are currently any interpolation applications that are truly spatiotemporal in nature rather than just 'space plus time' is reflected upon in Section 2.6.

People's lives consist of activities undertaken in both space and time. All of these activities such as work, study, travel, shopping and leisure occur at a geographic location and for a specified time of a given duration. Societies devote a large amount of resources into easing or trying to overcome spatial or temporal constraints (Miller 2005). One of the main challenges facing the geo-representation of data in GIS is the problem of 'timeless space' based on the conventional two or three dimensional approaches. The argument for geo-representation into the future is based on the inclusion of time with space in order to explore the dynamic phenomena experienced (Raper 2000).

The addition of time as the fourth dimension into a GIS is required for a spatiotemporal representation. Most methods represent space with reference to time, such as the traditional time-slice noted by Massey (1999). However, systems need to shift from just organising space *over* time to representing real world phenomena *in* space and time (Wachowicz 1999). Many people consider maps a snap-shot in time, rather than of time, often pre-empted by the ideology of publication dates or the timespan over which data were collected (Wood and Fels 1986). Geographic data have been described as having three components; theme, location and time (Sinton

1978). In order to measure a component another has to be controlled and one fixed. Often in the case of cartography time is fixed (Langran and Chrisman 1988), an historical theme of space without time traditionally considered in GIS.

Similarly to Hägerstrand and Pred, Wood and Fels (1986) represent a spatiotemporal representation of movement (of a bus) through space. This is represented in a three-dimensional diagram where the bus is always moving up, along an axis, in time. A planar representation is given, demonstrating the two-dimensional 'traditional map' depiction. The temporal dimension has been flattened to zero thickness and space emerges as the product of temporal flattening and the closure of movement.

The proliferation, and seemingly rapid advance, of ICT and telecommunications has changed the ways in which people act and how they communicate, altering the spatial and temporal distribution of human activities (Yu and Shaw 2008). Of the attempts made to create a truly temporal GIS, Hägerstrand's space-time prism concept is projected as a 2D surface, usually adding time as attribute data. This approach highlights the limitations in early GIS designed for handling spatial and cartographic data.

The increasing use of dynamic information in GIS can allow some temporal properties to be applied to objects, such as individuals in a population. Dynamic information concerns the data stored about an object that can change in a short period of time. This can be subdivided into real time, near real time and time stamped data (Yu 2006). The delay in entering real time data into a GIS may occur due to transfer times or for data processing to take place. Therefore, often near-real time is currently the most realistic approximation to real time. Time stamped data may often utilise temporal information added as an attribute. Such attributes may contain when the event occurred, the duration or when the data was transferred.

The differences in time described are not in the physical properties of time itself, but in the conceptual models used. Models are utilised to simplify reality and many are temporally unconscious. Langran (1992) summarises three aspects of cartographic time, distinguishing database time relative to real world time, where database time is equal to, less or greater than world time. These situations are based on the statements of what we currently know, how something appeared in the past and how it may appear in the future respectively. The traditional use of time in GIS has been to use a timeline.

A common theme in the traditional use of GIS and many time models is to use a point in time to describe when an event occurred, but to have no duration. This can be referred to as ordinal time. This technique is common in most GIS packages and it additionally allows the calculation of a duration between two points (Frank 1998). The use of integers or real numbers on a timeline can exclude temporal information that is available but not at the precise level of detail required.

The traditional method to organise time is on a globally recognised fixed scale such as days, months and years. A fixed time interval might also be chosen by the user to reflect the temporal resolution they require or have available. Although the fixed time interval is relatively straightforward to implement, an issue arises concerning the arbitrary choice of temporal boundaries (Shaw *et al.* 2008).

The cyclical measurement of time is widely accepted, often as a result of the astronomical processes first used to measure time. Many natural processes are influenced by these such as tidal change or diurnal animal migrations (Frank 1998). The order of cyclical time points is meaningless. For example morning is *before* evening which is *before* the following morning. Time measurement systems such as the 24-hour clock or dividing the day into two periods, *post* and *ante meridiem* allow a relative order, but is midday still before midnight? The same argument can also be highlighted by attempting to define the order of the seasons.

GIS software continues to be updated to develop new methods for handling and visualising temporal data. Examples include time series animation in ESRI's ArcGIS package, the earth trends modeller in the IDRISI software system and the TimeManager QGIS plugin to animate vector features.

In summary, time has been recognised in GIS for nearly as long as the history of GIS itself, however the ability to handle temporal information has been a point of discussion as explored in this review. The argument has been proposed that although time may have been taken into consideration, the complexity of handling temporal information and a primary focus on spatial data has often resulted in atemporal uses of GIS. The application examples that follow demonstrate analysis of population with respect to time. Better integration and recognition of time as a fourth dimension in contemporary, more computationally efficient, systems may highlight the requirement for more thought to be given to future analysis concerning time and space. There is still a discussion on whether current systems can truly be considered four-dimensional or spatiotemporal in their approach (Section 2.6).

The importance of the temporality of population data has been recognised by Bhaduri *et al.* (2007) in the novel LandScan USA application. LandScan USA has created a high resolution (90 m) population density map covering the USA using a dasymetric model and applied a temporal profile. As already noted, population data, commonly published as census outputs, are spatially constrained and atemporal. Censuses are usually only representative of who stayed where on one night, usually once in a decade. This is often referred to as the 'night-time population'. The mobility of the population results in the temporary relocation to other, often daytime, locations which are different to those represented by the census counts. Bhaduri *et al.* (2007) conceptually describe day and night-time population as:

$$\begin{aligned} Night\ time = {} & Residential\ Population + Night\ time\ Workforce + Visitor \\ & + Immobile\ Population \end{aligned}$$

$$\begin{aligned} Daytime = {} & Workforce + Students\ \&\ Pupils + Visitors \\ & + Residual\ Night\ time\ Population + Immobile\ Population \end{aligned}$$

This example highlights the use of the diurnal cyclical time scale. As already noted precise definitions of time can pose complex challenges, and the same applies to the notion of 'daytime' which is variable in terms of length of daylight hours. Bhaduri (2008) suggested the use of 'normal business hours' according to the US census definition. The UK census definition for daytime population counts the population aged 16-74 who do not work, but are resident within an area as well as all people who are working in the area (National Statistics 2004).

LandScan USA uses a dasymetric interpolation model to create national gridded population density estimates at three arc seconds (c. 90 m). The temporal GIS component, concerning the daytime distribution, was obtained using student and workforce location datasets. The LandScan global model exemplifies diurnal change in population distribution due to employment location. It uses the assumption that most people leave their night-time address and work during the daytime at their employer's location. The travel to work time is estimated, but does not take into account the time of day and the effect on the journey. Therefore, it represents an average over a 24 hour period (Bhaduri *et al.* 2007).

In one dataset example, the LandScan USA model uses a national school dataset to identify an important location of dense and vulnerable population. Following calibration spatial and temporal inconsistencies were discovered within the national dataset (Patterson *et al.* 2007), highlighting Langran's (1993) key spatiotemporal database issues concerning representation, update and longevity. The Population 24/7 method proposed (Section 2.7.1) can also handle a range of datasets, particularly those concerning school and workplace population. Its structure also permits updates for longevity.

Daytime population estimates contain two distinct properties, the locations of daytime activities such as workplaces, and the distribution of the population at those locations. The first property concerning physical geographical locations is usually easy to obtain and static. However, it is extremely difficult to obtain information concerning the movement and magnitudes of population change during the day. Although some isolated datasets exist, these are rarely on a national scale, sufficient to estimate the displacement of night-time population counts captured in a census (Bhaduri 2008).

2.5.4 **Analytical and simulation tools**

Microsimulation models (MSMs) provide an alternative methodology to represent population change over time and space. They capture interactions such as those between people and policy decisions at the level of the individual decision making units (Orcutt 1957). Lee (1973) critiqued such models noting the lack of transparency for policy makers and the failure of large scale urban models to reach set goals. At the time, large scale was a reference to the costly computer resources required for models that were then computationally intensive.

In contrast, the Modelling and Simulation for e-Social Sciences programme (MoSeS) aims to demonstrate a scenario in which the capabilities of gridded computing are utilised to develop tools in population modelling and simulation which surpass previous attempts in terms of power and capability (Birkin *et al.* 2007a). This concerns the development of a UK national demographic model and simulation of the population on an individual scale (Birkin *et al.* 2007b). MoSeS is based around microsimulation, with a similar requirement to that of SurfaceBuilder247 in relation to establishing a complete and detailed representation of a base population (Birkin *et al.* 2009). The MoSeS programme model attempts to simulate future population scenarios based on individual agents. Unlike the Population 24/7 model to be discussed in Sections 2.7.1 and Chapter 3, which employs a distance decay function based on temporally observed patterns, MoSeS relies on simulations to inform future patterns of a population's behavioural characteristics rather than focusing on an estimated location. An advantage of coupling microsimulation with agents based models permits the use of traditional mathematic modelling with the addition of individual behavioural information (Birkin and Wu 2012).

Ahola *et al.* (2007) present a convincing spatiotemporal interpolation of population data for a city centre to improve risk assessment and decision support analysis to support the emergency services and military. Building data were used to classify areas based on their use and as a spatial reference for a population. This includes data such the locations of leisure facilities, schools and shops. Population counts were obtained from the number of children in schools, people in a shopping centre and data relating to traffic volumes. Yuan (1996) employs a three domain model which treats attributes, spatial and temporal data in separate linked domains. For example an attribute might be a building with description and number of population contained. The spatial component is the physical location and temporal information concerns the number of inhabitants expected at different intervals. Temporal information relating to population variations were divided into ten categories broadly concerning sub-population types by age or occupation.

Tools to simulate population movements are also widely used by logistics and transport modellers. The TransCAD transportation planning software (http://www.caliper.com) integrates demand modelling with GIS. It permits network analysis, distance/travel time estimates and destination-origin flows. These tools could potentially be integrated with the Population 24/7 method currently being developed. They provide exciting practical developments in the ability of GIS systems to handle spatial and temporal data. Presently, the datasets for use with Population 24/7 also hold information on population origins and destinations, but it currently is not possible to define the exact route of flows between such sites.

Within these sections a range of time scales have been discussed, how population varies within these and methods to represent this data. Traditional GIS software has handled population data without time, although spatiotemporal concepts have been around for some time (e.g. Hägerstrand 1976; Langran and Chrisman 1988). The consideration of time in addition to the spatial distribution of population and how this changes has been demonstrated to be an important characteristic. This thesis builds on these concepts to develop a tool, and it's applications regarding hazard, to handle spatiotemporal data.

2.5.5 **Space-time kriging**

Space-time kriging has been used for the geostatistical analysis of environmental data (e.g. Bogaert 1996; Heuvelink and Griffith 2010). They have been applied to human population data analysis in terms of diseases modelling (e.g. Gething *et al.* 2007). Rouhani and Myers (1990) suggest limitations in the application of this technique on some geographic datasets particularly where there is a high degree of dependence in either the spatial or temporal domain. It is acknowledged that kriging applications can be used with spatiotemporal data. However, this thesis is focused on the spatial analysis of population data considering the spatial and temporal dimensions rather than a geostatictical exercise. A spatiotemporal KDE is the preferred method of choice best suited to the population datasets available for this study. One advantage is the ability to use known local dispersion parameters for specific population sites.

2.5.6 **National Population Database**

The problems associated with static population estimates for emergency responses are recognized by the Health and Safety Laboratory (HSL), an agency of the UK's Health and Safety Executive. They identify that national, robust and detailed population data are required to address the risks posed to people from hazardous events.

The National Population Database (NPD) (HSL 2014) is a novel, powerful GIS tool created by the HSL. Its coverage extends across the whole of Great Britain (GB). It is formed of a geodatabase containing mixed-use population centroids. The population groups included within the NPD tool are summarised in Table 2.3. It was created by applying generic population multipliers to buildings, transport and land use types to produce a database of population density.

Table 2.3 NPD population components (Smith *et al.* 2005)

Population type	**Example/temporal properties**
Residential	Day and night-time, daytime tem-time
Sensitive	Hospitals, schools and care homes
Transport	Motorway and A-road average/peak populations
Workplace	Daytime working population
Communal establishments	Prisons, campsites, stadia, leisure facilities.
Retail	Average/maximum daytime population

The tool is built on a wealth of datasets including the postcode directory, local authority information, government administrative datasets and OS data. The database is restricted to public service users as a collection of population centroids or a 100 m grid of points. The database is queried though a standard GIS interface. It provides some great advantages to standard aggregate population data. The Population 24/7 approach that will be discussed and implemented within this thesis (see Chapters 3-5) uses a similar categorised system of population centroids. However, unlike the NPD they are modelled onto a variable raster grid. This increases processing time, but allows higher resolution temporal estimates to be produced. This permits the estimation of hourly population outputs compared to singular day/night-time estimates currently available within the NPD.

2.6 Spatiotemporal interpolation of population data

Applications of the spatiotemporal interpolation of population data have been highlighted in Section 2.5 with examples. This section outlines the need to move from traditional space and time representations to truly spatiotemporal interpolations, although very few practical examples can be found in the literature (e.g. Ahola *et al.* 2007 and Bhaduri 2008).

There is a large expanse of literature concerning time-space, space-time and timespace (e.g. Hägerstrand 1975; and Thrift 1977; Pred 1984; Massey 1999). A literature also exists on the place for space-time within the geography discipline (Merriman 2012) which suggests a paradigm shift away from 'science' and views space and time as a dominant view in western epistemology that clouds other judgement.

Time-space snap shots mimic the effect of a slow motion camera as the best approximation to spatiotemporality in a GIS (Langran 1992). Snap shots represent change, an important feature of time. They only represent states, the change from one time to another, but not the events that cause the change. Spatiotemporal representation has a huge potential, but it is highly challenging (Raper *et al.* 2005). Traditional GIS and spatial data models such as raster and vector models limit the way in which dynamic data can be handled. A core element of spatiotemporal modelling concerns how to measure change over time and analyse the results in GIS. Ahola *et al.* (2007) suggest a suitable data model is the second core element for spatiotemporal modelling. Such data models can be categorised according to their organisation as space, time, feature based or a combination. Methods have been used to incorporate the spatiotemporal transformation including 4D to 2D, 2D or 1D plus time (Raper *et al.* 2005), where time is considered the fourth dimension in the space-time continuum or a two-dimensional approximation is considered or a simple one-dimensional snap-shot in time. There is a requirement to develop tools, structures and visualizations to fully make use of spatiotemporal information (Raper *et al.* 2005).

Conventionally, GIS facilitates the representation and interpretation of spatial data. Therefore it is useful for analysing the spatial patterns of activity (Wang and Cheng 2001). The functions of GIS have been hugely benefited by advances in computing power and availability, but still require further development to describe spatial changes with time. This builds on the argument put forward by Massey (1999) that ‘space’ is open and dynamic, it is not just a classic slice through time. ‘Space’ has a necessary temporal component with which it should be paired.

The static representation of most map views in GIS often favours description. The consideration of movement and mobility often leads to the problem of sorting time-dependent attributes. GIS must support a number of other data types for the purpose of modelling complex behaviour. These include flow matrices that concern the movement of an object at a given point in time, between an origin and destination (Goodchild 2000). Therefore modelling movements often needs representation and methodologies not currently present in most GIS. Most GIS software is based on handling spatial data for mapping and navigational purposes. Digital representations are often static descriptions because it is difficult to build operations that process change over time. Combining space and time raises the issue of the representational problem (Fisher and Unwin 2005).

Disaster risk cannot be fully assessed without taking into account spatiotemporal variations in population. Several novel high-resolution spatiotemporal population modelling techniques are emerging that readily contribute to the field of disaster risk management. Day and night-time population estimates for Lisbon to assess population exposure to seismic hazards have been constructed by Freire and Aubrecht (2012) with a spatial resolution of 25 m. Furthermore, Aubrecht *et al.* (2014) have developed DynaPop, a spatiotemporal population dynamics model, which further increase temporal granularity. Population outputs in raster format from DynaPop at 100-500 m have an hourly temporal resolution. This example has the potential to make significant advances on simple previous day and night-time estimates. This work draws on useful parallels with the Population 24/7 approach, the latter will be the focus of this thesis. In contrast to DynaPop, the Population 24/7 technique has been tested and suited for UK data applications. However, as spatiotemporal modelling techniques continue to evolve future cross-overs may become a focus of further work.

Finer temporal resolution population density estimates are still required for natural and technological hazard assessment (Bhaduri *et al.* 2007). A temporal GIS can be defined by the representation of location, attribute and time without fixing any component. Often one dimension is fixed, and therefore the GIS cannot be considered truly spatiotemporal (Langran 1992). Previous research has shown that diurnal shifts in population can be estimated (McPherson and Brown 2004) as has been demonstrated in the LandScan USA programme (Bhaduri 2008) and by Ahola *et al.* (2007), a starting point for the recognition of the temporal variation of population over space.

2.7 Space and time in populations exposed to natural hazards

Population is a function of both space and time. Limitations in mapping population have been acknowledged for some time (e.g. Schmitt 1956). Methods and caveats in mapping, storing and obtaining spatiotemporal information have been explored throughout this review and are summarised here in relation to the assessment of risk and vulnerability to natural hazard scenarios. Requirement for a better understanding of the population affected by natural hazards has been noted (Bhaduri *et al.* 2007; Aubrecht *et al.* 2012b). Large numbers of people are at risk from natural disasters such as earthquakes, volcanic eruptions, floods, droughts and wildfires. Typically they are mostly predictable (except earthquakes and some volcanic eruptions) and topographically constrained. Although hazards often cannot be prevented their impact can be minimised through effective disaster planning and emergency preparedness (Bhaduri 2008).

The spatial and temporal components of both population distributions and natural hazard footprints are complex and likely to be the focus for interdisciplinary research between both the social and natural sciences to improve methodologies. The hazard component itself also has complex spatiotemporal characteristics, independent of population and administrative boundaries, which may be represented by simple static risk maps or dynamic models. In order to make accurate risk assessments, not only are adequate population data required but also the likely onset time and extent of a hazard scenario. In the study of natural hazards it has been recognised that improvements are required in the spatial and temporal detail of population density (Cutter and Finch 2008) to better plan emergency response, provisions and impacts.

Spatial population interpolation techniques lend themselves well to natural hazard risk evaluation and provide an alternative to traditional methods. A number of examples such as kriging, pycnophylactic and masked gridded representations have been discussed within this review. Each provides different advantages such as volume preservation, surface smoothing, or the ability to handle zero population densities. Such modelled surfaces provide an alternative to the conventional choropleth map and associated and unrealistic uniform population densities.

An attempt to combine population, space and hazard is exemplified through Houston *et al*'s (2011) study into pluvial flooding of urban areas. They use census OAs, the smallest census output geography in England and Wales, to demonstrate the population potentially at risk by recognising the need to estimate numbers of people at risk rather than just properties, common in existing risk assessments. While this study recognises a decadal scale, it contains little temporal information. It might be argued that it is impossible to tell exactly when a hazard will occur, but the time of day it might strike will have enormous implications for the population.

The size, location and demographic characteristics of a population are all drivers for the impacts of a natural hazard (Cutter 2010). Three components provide the intellectual basis for analysing population vulnerability and resilience to natural hazards (Cutter 2003):

1. Physical processes: an understanding of the science of natural hazards.
2. Human systems: interactions between society and the natural environment which contribute to vulnerability, such as occupation of hazardous areas.
3. Local geography: an understanding of specific characteristics of a place such as the landscape, history, demographics and economics.

Estimating population risk and vulnerability to natural hazards is complicated in part, often due to data availability. In a coastal flood inundation scenario this estimation is further complicated in areas of rapid population growth and development which continuously change the factors influencing risk and vulnerability (Chakraborty *et al.* 2005). This also contributes to spatial variations in the populations risk and vulnerability. Spatiotemporal population interpolation estimation techniques have been developed and lend themselves for application to hazard risk management where vulnerability and population are defined as key characteristics (Table 2.1). This requirement has been reflected in both the literature and legislation resulting from natural hazard events.

Decadal censuses in many countries provide accurate and comprehensive aggregate population counts. However, these pose limitations for application to natural hazards, particular in the event of cross boundary impacts. Areas affected by natural hazards are not constrained within arbitrary administrative boundaries.

Gridded population surfaces provide a universal structure with increased stability through time which also promotes greater compatibility with other environmental datasets required in the hazard management process. Population data are often aggregated to administrative zones or units before publication. These vector data structures and arbitrary units often relate little to the local topography and are incompatible with traditionally raster based environmental datasets (Martin and Bracken 1993). Where gridded population datasets are available these have been applied to make better predictions about the population affected.

The shift from displaying population within modifiable areal units to a universally compatible grid provides the first step towards enhanced assessment for potential hazard exposure. Where population data are not published as a grid, spatial interpolation methods such as those discussed in Section 2.4 by Tobler (1979) and Martin (1989) provide a method for disaggregation. Tobler's pycnophylactic approach produces a 'smoothed' population surface. This removes abrupt changes at the boundaries of data contained within irregular units. However, Martin's (1989) approach allows for locations of zero population density. This is particularly important when concerning uninhabited locations such as water, parks or woodland. Therefore this can provide a more realistic location for populations potentially at harm during a natural hazard event. While the footprint of some hazards may be unpredictable such as earthquakes or meteorite impacts, others may form predictable extents such as the path of a hurricane or water inundation in low lying floodplains. High resolution disaggregated population counts permit the micro-assessment of impact as well as understanding the wider implications.

Although grids provide greater flexibility for comparison to hazard data they still only provide a static representation of the population at a given point in time. Censuses provide a 'night-time' residential population count to produce a decadal time slice. The size, location and demographic characteristics of a population are all drivers for the impacts of a natural hazard (Cutter 2010). These all have an important implication for natural hazard impact assessments. The predicted or actual time of the manifestation of a natural hazard occurring is going to expose greater numbers of a potentially vulnerable population if for example it occurs at peak work rush-hour or during educational term time.

Resilient societies may have developed mechanisms to cope with high frequency, low magnitude events. Developing and combining realistic temporal population estimates and natural hazard scenarios of variable magnitudes allows the analysis of a worst case situation and measure of risk. This may be in the form of an unexpected or high magnitude event that a population is unprepared for or vulnerable to.

Gridded population estimations (Section 2.4.4) have been demonstrated to provide a more flexible and realistic distribution of population. A critical application for accurately estimating the population at risk during a hazard scenario includes assessing the population exposed or injured and emergency preparedness (Bhaduri *et al.* 2007).

To conclude, a number of questions and considerations in the application of spatiotemporal population modelling to natural hazards have arisen following this review:

- Develop a practical method for spatiotemporal population modelling
- Overcome challenges of data processing and representation
- Proper integration between population and natural hazard risk models which both change over time

2.7.1 **Population 24/7: spatiotemporal modelling concepts**

A new approach is required and a possible method is provided by Population 24/7 (Martin *et al.* 2009) which will be introduced in the empirical chapter of this thesis. Like the original SurfaceBuilder software (Section 2.4.4) (Bracken and Martin 1989) the Population 24/7 approach also redistributes centroid populations onto a variable grid. A population may be associated with different types of locations representing spaces of human activity such as places of residence and work. These are represented by centroids. A centroid is a geographically referenced point with an associated population count. Population centroids weighted to the most likely place of habitation have been released in recent UK census datasets. However, custom population centroids can also be created for use with the Population 24/7 method to represent population activity types. Centroids can be divided into two categories: population origins and destinations. Origin centroids, which supply the population to be redistributed onto a defined grid, might be comprised of the aforementioned census centroids. Destination centroids provide locations to receive the given population at a stated capacity such as a business, hospital or school. These are the locations where population may be present at different times during the day. Importantly, temporal profiles can be associated with each of the destination centroids, a difference from the NPD (Section 2.5.6). The temporal profile governs the proportion of the total capacity for a destination centroid to be occupied for a given time. At any moment in time, the whole population is either at one of the locations represented by a centroid or travelling between them. These can be termed the 'on-site' and 'in travel' population. The on-site spread describes the population present at or immediately surrounding a centroid, either origin or destination, location. The in travel count describes the population in transit between centroids. The proportion of an origin centroid's immobile population can be user specified or based on available data. A background layer may also be created to weight the in travel population onto road networks or to prevent placement in uninhabited areas such as water.

The Population 24/7 approach has been chosen to underpin the empirical work within this thesis because of its ability to facilitate the production of gridded population distributions for a specific time and date. Its flexible data structure also allows the richness of population attributes to be increased and ensures data currency. These are fundamental considerations when assessing human risk to natural hazards which are both space and time specific. The methodology provides a mechanism to estimate temporary and transient populations that are not accounted for by traditional means. For example, this includes information on the number of employees or shoppers in retail locations on a high street. This would need to consider both the temporal and spatial attributes. The combination of available datasets with the ability to store temporal information produces spatiotemporal gridded representations of populations accounting for local variation. The variable gridded method allows such micro trends to be resolved.

The Population 24/7 approach also facilitates analysis of the vulnerability of populations through the ability to handle socioeconomic attributes for sub-groups where data can be assembled. As already noted, vulnerability is a key contributing factor in the development of risk. The capacity to handle age information may inform management choices for groups such as the young or elderly who may face greater risks and require additional support during a natural hazard event. It also increases the accurate placement of populations within space and time, such as school age children at school locations.

Chapter 3: Methods and Data

3.1 Overview

The work presented here combines the use of a spatiotemporal gridded population model to estimate time-specific variations in population with natural hazard exposure estimates in the form of flood inundation data. It has been exemplified through an application centred on Southampton (UK) using Environment Agency flood map inundation data. The proposed methodology is summarised in the analytical overview provided in Figure 3.1. This can be considered as three subcomponents. These will be outlined in turn within this section. Figure 3.1A covers the spatiotemporal modelling process, Figure 3.1B the hazard component and Figure 3.1C brings these GIS based datasets together for analysis. Often natural hazard and population models are not linked. The proposed methodology aims to demonstrate the improvements to risk analysis when these are considered together. This chapter begins with a description of the Population 24/7 methods and creation of datasets, then the study site and flood hazard data are introduced. Finally the data analysis is described and results presented.

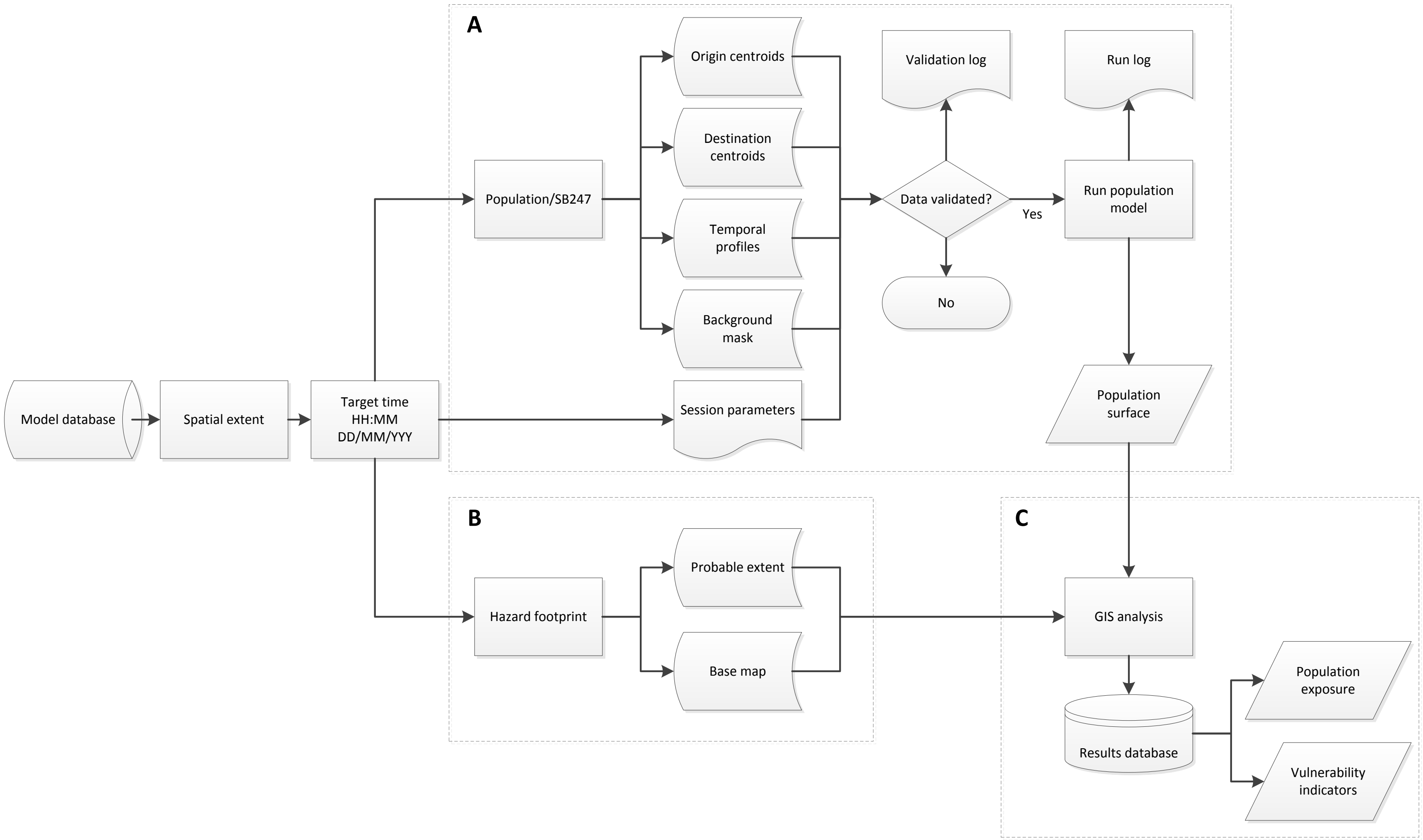


Figure 3.1 Overview of analytical operations to assess natural hazard exposure.

3.2 Modelling Population, Population 24/7 and data structure

The spatiotemporal population modelling component (Figure 3.1A) follows the Population 24/7 methodology developed by Martin *et al.* (forthcoming). For this application population data are modelled onto a regular grid with 200 x 200 m cell size, which is appropriate to the resolution of the available input aggregate data sources. The technique employs a variable kernel density estimation technique to redistribute population from origin to potential destination locations according to a distance decay function depending on the time of day. A dataset of destination locations was initially collated for the Population 24/7 project, which includes places of work, education, and health care. This research has extended these methods and datasets by building on this original 'library' of potential population locations. The library is extensible, thus allowing users to develop their own datasets, particularly for non-residential locations.

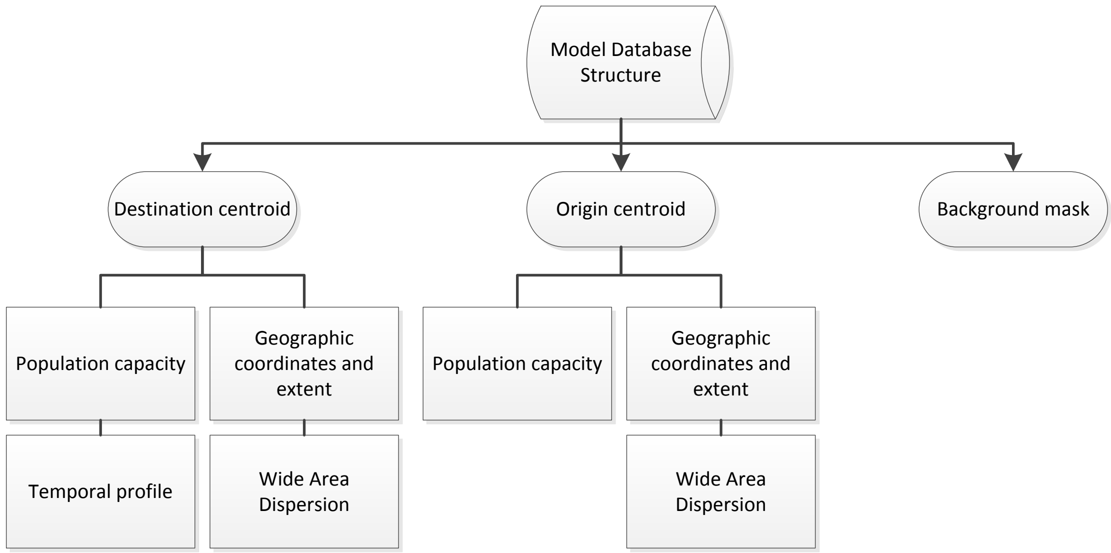


Figure 3.2 Population model data structure

This section provides an overview of the SurfaceBuilder247 software and the method for compiling the required datasets on which the model runs. The revised algorithm and updated software currently packaged as SurfaceBuilder247, the successor to the original SurfaceBuilder has several additional features. The spatiotemporal, '24/7', version of the model has been developed to accept destination centroid locations as well as an associated temporal profile, in addition to the original origin centroids (Martin *et al.* 2009) (Figure 3.2). This has the effect of describing a population's presence at different times in a more realistic spatial distribution. These are described in turn in the subsequent sections.

Following a review of the literature it is observed that the methods and criticisms in the interpolation of population data have been around for some time. It has been noted that population data are commonly aggregated into a wide range varying of spatial units and reporting zones. This is primarily driven by data management and privacy constraints. Innovative spatiotemporal population modelling methods developed through the Population 24/7 project will be developed in this thesis. It is intended that these developments will permit significant improvements in the assessment of populations exposed to natural hazards.

3.2.1 **Origin centroids**

The origin centroids provide the source, or supply pool, of population that can be utilised and distributed by the model. An example would be census centroids. However, origin centroids could be created by the user to georeference an associated known point of population. Locations of immobile populations such as prisons can be specified. An attributed population count specified as immobile will not be redistributed when the model runs. The combined population of the origin centroids defines the model's input population.

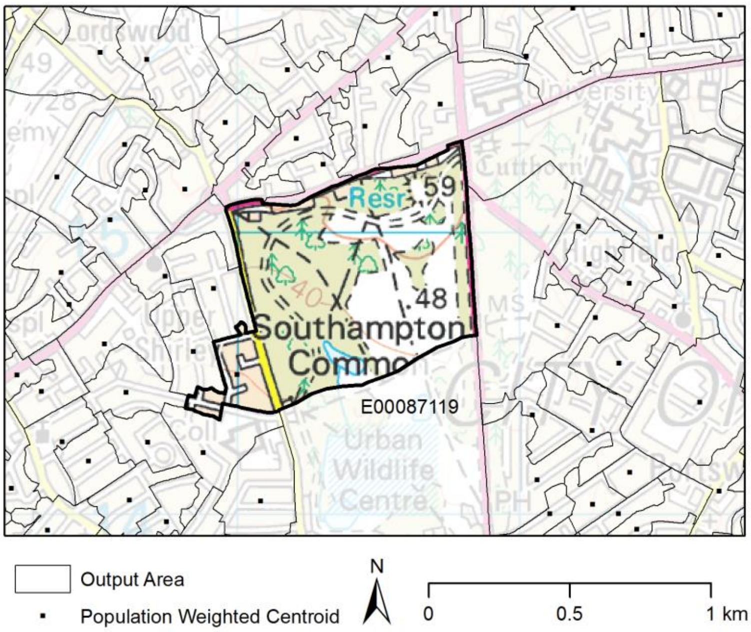


Figure 3.3 Census output areas and population weighted centroids for Southampton, UK.

Population weighted census centroids are a common feature of many census datasets. The centroid represents how the population was spatially distributed within that reporting zone. Figure 3.3 provides an example for a mainly uninhabited, and therefore comparably greater spatial extent than a more densely populated census OA[1] in Southampton. In this example the centroid is located within the area of highest population density as expected. Figure 3.3 also represents the highest resolution input population data currently available for this study.

3.2.2 **Destination centroids**

Destination centroids represent all possible locations where a population may move throughout the day. Common examples include places of work, leisure or education. The user can compile any number of destination centroids of choice. The purpose of this feature allows a pool of available population, attributed to residential origin centroids, such as 'night-time' census counts, to be allocated to more temporally relevant locations. Each destination centroid is assigned a capacity to receive population, which may

[1] At the time of writing small area statistics for the 2011 UK census had not yet been fully released. Therefore, 2006 MYEs and 2001 output geography have been used for the current study.

vary depending on the time of day and day of week, and is governed by an associated temporal profile. The purpose of the introduction of the destination category of centroids is to provide more realistic daytime population density estimations.

The population of the destination centroid is assumed to be drawn from the surrounding origin locations. The area of influence is the distance over which population is reallocated from origin centroids to the destination in question. This is equivalent to a catchment area for the destination and can be described by a wide area dispersion (WAD) function for each individual destination centroid. Careful data collection or analysis may be required to justify the WAD function. In some examples such as a school the destination's population is likely to be constrained to origin centroids falling within the radius of the school's catchment area. Future, more sophisticated versions of the model could be adapted to integrate spatial interaction models to inform the WAD. If the data were available, real-time information on population flows could be incorporated in the future.

In order to describe a destination centroid location the following essential factors need to be considered (Martin 2011):

- Geographical location: this provides the spatial information on exactly where a given population will be, such as the coordinates of the building or venue represented.
- Population capacity: the population capacity for each destination. For a school this may be derived as the number of pupils and staff. Additional datasets and analysis are required when attempting to estimate the population of a destination whose visitor population is less clearly defined, such as a shopping centre (Section 3.3).
- Time profile: information is required at each destination centroid to describe the temporal pattern over which population is expected to be present (see Section 3.3.2).
- Spatial extent (or Local Dispersion): by definition a centroid represented as a single (x,y) coordinate pair is a precise 'pinpointed' location. In reality almost any feature such as a shopping centre or school is going to occupy a larger spatial extent. For example a

primary school site may cover an area with radius of 200 m so would be assigned an extent of 200 m to reflect this.

- Wide Area Dispersion: The distance the population travels to reach the destination centroid location. As it is extremely unlikely that everybody travels the same distance, a wide area dispersion function can be specified in various ways (see Section 3.2.4).

3.2.3 **Temporal profiles**

The temporal profile is only required for destination centroids (Figure 3.4), referenced in the corresponding destination dataset. An exact time profile could be constructed from raw count data if such data are available at the required spatiotemporal resolution. An example would be the school age population allocated to a school during opening hours. A profile is not required for the origin centroids which represent residential locations; although a part of the population may be declared as 'immobile' if they are considered unable to travel to other destinations. Examples would include prisoners and a proportion of the elderly or hospital inpatients.

A time profile comprises two parts (Figure 3.4). The first part concerns the 'in travel' population associated with the destination. The second defines the proportion of the destination's population capacity that is actually present, 'on site' at a specified time. A defined population capacity for the destination is paired with a time stamped interval, t_1, t_2 … t_n. The population capacity represents the notional population of the destination when it is fully occupied, for example when all children are at school or all employees at work. The time profile defines the destination's expected population as a proportion of this capacity during each time period.

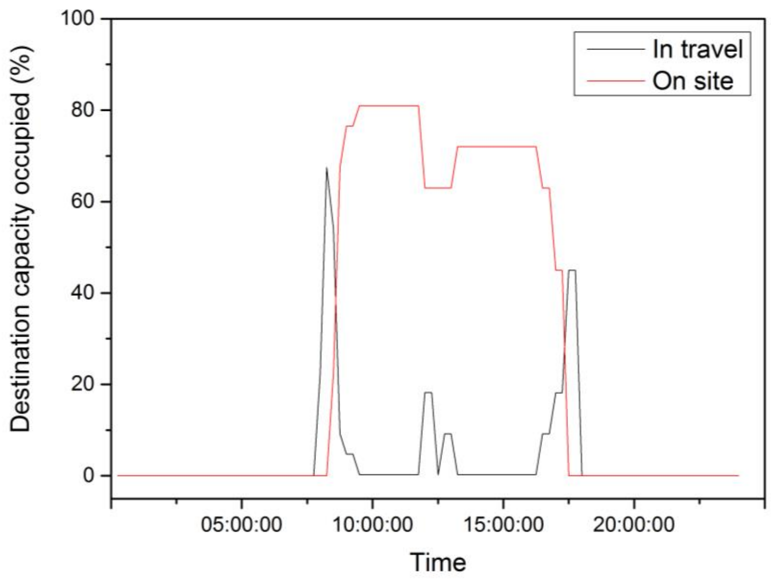


Figure 3.4 A simple temporal profile illustrating the proportion of a destination centroid capacity occupied (on site) and the population in travel.

In this example (Figure 3.4) none of the destination's capacity population are in travel at 4 am, and by 10 am, 81% are on site. A time profile can be constructed in any number of intervals covering a 24-hour day. Different time profiles may be assigned to different days of the week, term time/holiday time or seasons. A specific target time is selected for each run of the model, and any destination centroids within the analysis area will be processed according to their time profile. Therefore, if the model is run for 10 am in this example (Figure 3.4), 81% of the capacity population is expected to be present. Thus, if the declared capacity is 100 people, then 81 people will be allocated from surrounding origin centroids, which fall within the area of influence to the destination.

3.2.4 **Wide Area Dispersion (WAD)**

The wide area dispersion (WAD) provides important information utilised by the model concerning the supply of people to an area. The WAD describes the population demand of the destination centroid at the target time which is transferred from population origins within the area of influence. Conventionally this may be informed through spatial interaction models (e.g. Diansheng 2009), however the WAD in effect records data on the distances travelled from known locations or centroids. Population counts are only supplied from origins falling within this area of influence.

For the purpose of this model the WAD comprises bands which describe the proportion of population travelling from a specified distance. In the example represented (Figure 3.5) a proportion of a destination's population is supplied from within 100 m, 2 km and 3 km. A worked example is provided in Section 3.3.6.

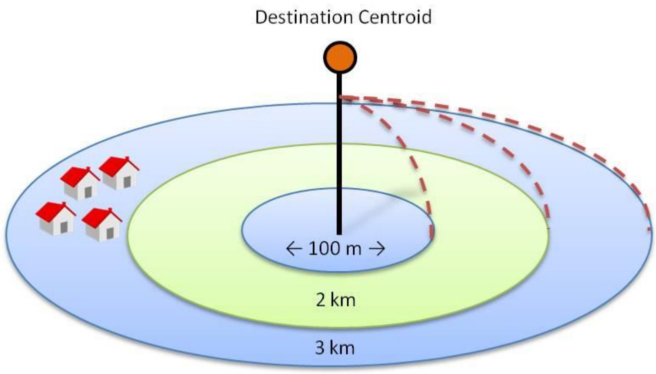


Figure 3.5 Wide Area Dispersion for a destination centroid

If there is insufficient population to supply the demand in any band, the search radius will be increased (Martin 2011). In some cases the WAD may be clearly defined such as travel constrained within a catchment area, for example in relation to a school.

3.2.5 Background masking layer

A background layer can be used to represent land use and transportation. This provides constraints on the locations of any population that is not allocated to a centroid location (e.g. the population 'in travel'). The background layer should prohibit placement of population in uninhabitable space such as open expanses of unoccupied countryside, large water bodies and oceans. For the current version of SurfaceBuilder247, the dataset may be prepared in a GIS and exported as a raster data file to be read by the model.

Using a dasymetric approach (e.g. Langford 2007) a land use classification may be applied in to the masking layer to constrain population to inhabited locations. This can be applied to the population not present around any of the centroids at the specified time (such as those 'in travel'). However, this may have limitations within the current study as traditionally 'uninhabited' locations such as shipping docks or areas of industry can receive large temporary populations when passengers and the workforce are present.

To generate a background mask it is necessary to combine the relevant datasets depending on the study area or user's requirements using GIS. Features can be weighted, if possible, to where they are likely to receive more population than others. The GIS dataset is rasterised at the required output resolution based on the weighting values to be read by the model.

3.3 Applied example: Southampton, UK

This section provides a worked example for the creation and application of a retail customer dataset covering Southampton, UK for use in SurfaceBuilder247. An exemplar retail dataset has been created because the retail population is an important spatiotemporal phenomenon as many city centres and shopping districts see a large increase in population number during the day that is not apparent in traditional census datasets. This dataset has been developed for implementation in the current model for this thesis. This section describes the creation of origin census centroids, a retail temporal profile, retail destination capacity and finally a retail WAD. This is just one of many other leisure datasets that could be created to represent population activity in addition to work and education. While this is a specific retail example the principles of dataset creation are universal.

A 25 x 25 km study area centred on Southampton has been selected to demonstrate dataset construction (Figure 3.6). A further 25 km buffer zone surrounds the study area in order to supply population to destinations towards the edges, helping to relieve any ‘edge effect’ on the study area in question. An edge effect occurs where centroids on the boundary will in reality be interacting with populations outside the defined study area as there is often no finite line where population movements and interactions cease. However this case may be argued along international borders or physical geographic barriers.

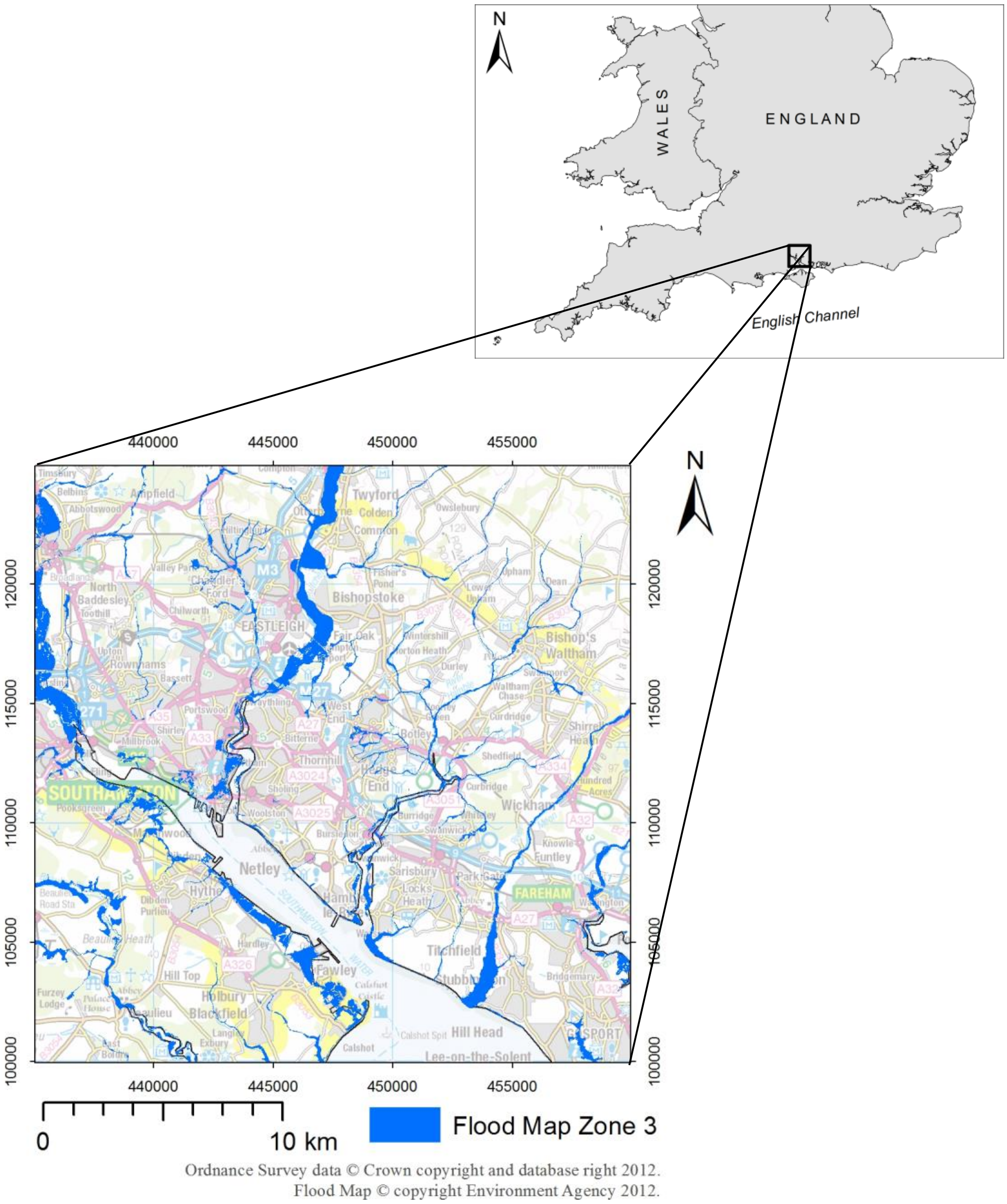


Figure 3.6 Southampton study area and flood risk

3.3.1 United Kingdom census origin centroids

Census population weighted OA centroids are used for the origin centroids. These provide a comprehensive population count at the current highest spatial resolution for the UK census. The majority of OAs contain between 110-139 households (ONS 2012a). The centroid locations are population weighted, providing a best fit by means of a single point of the spatial distribution of population from each respective OA (ONS 2012b).

3.3.2 Temporal profile creation

The retail dataset is just one of several datasets required to set up the population model used in this thesis. Its purpose is to map the times and locations of people doing their shopping that contribute to large temporary population variations. However, these data are not readily available so the necessary parameters outlined have been derived from broad national datasets for specific locations. In a retail example electronic point of sale (EPOS) data available commercially or collected by large retailers could instead be used as a proxy for the retail time-specific population providing a high temporal resolution. Where true count data are unavailable, numerous approaches could be devised in an attempt to create a temporal profile. Temporal profiles are specific to particular destination centroid type. The following discussion will help to demonstrate one method to construct such profiles. This discussion concerns the two components required for constructing a valid temporal profile, the on-site and in-travel populations (Figure 3.7). Some additional information is required: the average occupancy duration, the average travel time to the destination (which may also be required for WAD calculations) and the proportion of the total population capacity expected to be present at a given time.

Figure 3.7 demonstrates that there are three principal population subgroups to consider in relation to destination locations. These are the population movements to and from a destination and the population present. This population can be taken from any origin centroid from within the destination's catchment.

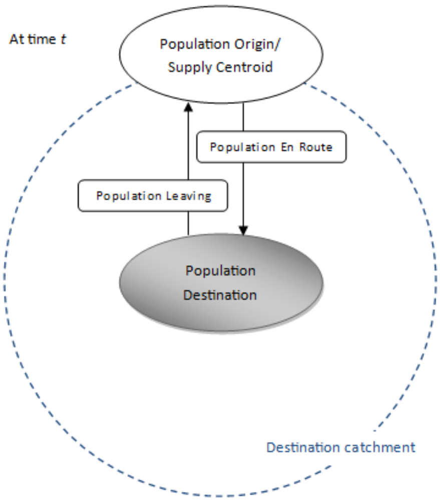


Figure 3.7 Modelled population movements and locations

One method for calculating a temporal profile begins by considering the proportion of people expected at a given start time, the first entry on the time profile (t_1) (Figure 3.4). The following calculation is proposed to estimate the on-site population, the population expected to be present at the destination site for the given time:

The population on-site (*O*) at centroid *j* at for a specific time interval:

$$O_j = C_j p_j \tag{3.1}$$

Where, *C* is the population capacity of destination centroid *j* and *P* is the proportion of the centroid's population expected to present (Figure 3.7).

The in-travel population is determined through the sum of the following formulae (Eqs. 3.2 to 3.4) which concern three principal movement states. The first is associated with the in-travel population travelling to a destination centroid. This is calculated by detecting a population increase between two time intervals representing the population en route. If this value is zero, then nobody is in effect currently en route and thus it will not contribute a population count towards the overall in-travel population:

$$If\ (P_{t_2} - P_{t_1}) + (P_{t_3} - P_{t_2}) > 0, \qquad then \tag{3.2}$$

$$Population = \ (P_{t_2} - P_{t_1}) + (P_{t_3} - P_{t_2})$$

$$Otherwise, \qquad Population = 0$$

The second component of the in-travel population estimate outlines the site's capacity and its population demand at a given time. This governs how much of the surrounding population is drawn to the centroid. Equation 3.3 is formed of two parts. The first is a constant derived relevant to the particular destination centroid. This is the total number of visitors (n) minus the destination centroid's capacity (C_j) divided by the total mean duration ($\bar{t}$) of individuals' visits to the location. This is then multiplied by the population expected to be present at centroid j (at time, t_1).

$$\frac{n-C_j}{\bar{t}}p_j \tag{3.3}$$

As Equation 3.2, the population leaving can be considered the reverse of those arriving:

$$If\ \left(P_{jt_2} - P_{jt_1}\right) + \left(P_{t_1} - P_{t_2}\right) > 0, \qquad then \tag{3.4}$$

$$Population = \left(P_{t_2} - P_{t_3}\right) + \left(P_{t_1} - P_{t_2}\right)$$

$$Otherwise, \qquad Population = 0$$

The on-site and in-travel population calculations need to be performed, and repeated until t_n, where n is the final time interval.

For example, the proportion of population present at centroid *j* (P_j) for time one (t_1) could = 10%, t_2 = 20% and t_3 = 30%. This formula determines that the proportion present at t_2 is increasing because it is greater than it was in the previous time step (t_1) but smaller than the next (t_2).

3.3.3 **Time Profile for UK retail destinations**

The development of a time profile for a retail destination dataset has been based upon the UK Time Use Survey (TUS) 2000 (Ipsos-RSL and ONS 2000), in order to estimate shopping habits in the UK following the profile calculation method proposed.[2]

The survey comprises self-completed diaries of a nationally representative sample of UK householders. Diary episode data is available at 10 minute

[2] The 2000 Time Use Survey is currently the most comprehensive in the UK and has not yet been repeated.

intervals covering seven days throughout a week. Time use has been allocated to a number of coded activities which include categories for shopping (Table 3.1). These will be used to estimate a temporal profile for retail activity.

Table 3.1 United Kingdom 2000 Time Use Survey: 36. Shopping and Services (Ipsos-RSL and ONS, 2000)

Diary Code	Description
3600	Unspecified shopping and services
3610	Unspecified shopping
3611	Shopping mainly for food
3612	Shopping mainly for clothing
3613	Shopping mainly related to accommodation
3614	Shopping or browsing at car boot sales or antique fairs
3615	Window shopping or other shopping as leisure
3619	Other specified shopping
3620	Commercial and administrative services
3630	Personal services
3690	Other specified shopping and services

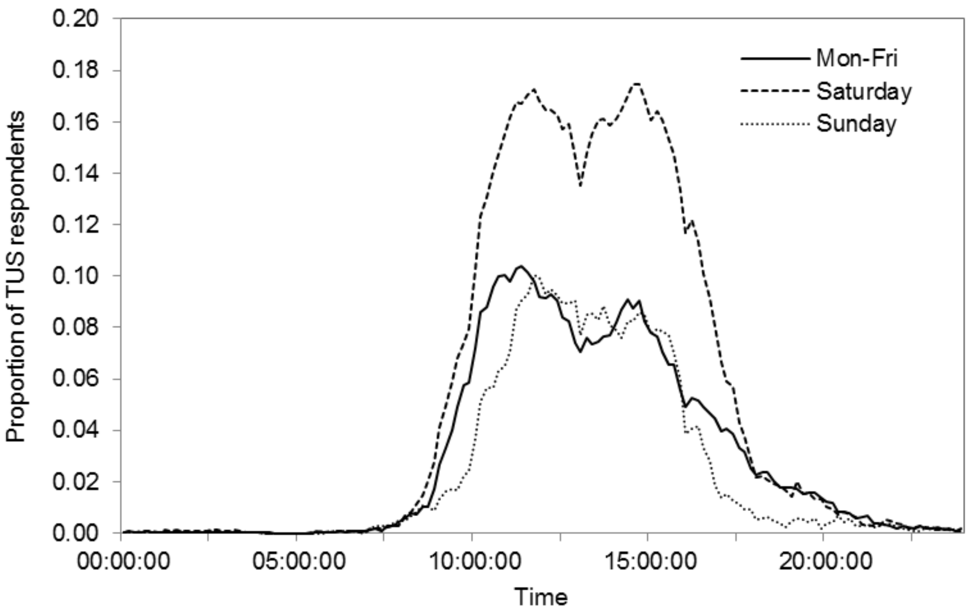
A total of 20,981 UK respondents aged eight and over kept diaries in 144 x 10 minute intervals representing a given 24-hour period, recording a primary and secondary activity. The retail time profile has been based upon the number of respondents who stated shopping related activities (Table 3.1) as their primary activity. The data were analysed for three separate time periods: Monday-Friday, Saturday and Sunday. The total number of people and proportion who had a primary activity code beginning 36xx was calculated for all days at each 10 minute interval.

The TUS also provides a coded diary entry concerning '9360 – Travel related to shopping'. Along with the above data this has been used to create the average time per day spent shopping and travelling for shopping (Table 3.2).

Table 3.2 Average shopping travel and duration derived from the TUS 2000 for respondents who shopped

Time Period	Average travel time for Shopping (min)	Number of respondents	Average shopping duration (min)	Number of respondents
Mon-Fri	38.36	3691	65.40	4287
Sat	45.33	2707	88.88	2886
Sun	38.97	1682	63.23	1630
All days	35.09	8080	72.67	8803

The proportion of TUS respondents who stated that they were carrying out a shopping related activity at a given time is displayed in Figure 3.8. A large peak in Saturday shopping can be observed, as expected, a traditional busy period when many people are off work and able to shop. A restriction on



Sunday shopping is observed and explained in the next paragraph.

Figure 3.8 Proportion of shoppers by time according to TUS 2000 diary data

The time profile suggested by the TUS data appears to be consistent with expectations with a large surge on Saturdays and restricted opening hours on a Sunday. The Sunday Trading Act 1994 restricts Sunday trading of stores in England and Wales with a floor space > 3000 sq. ft. to a maximum of six hours permitted between 10 am and 6 pm. This legislation does not apply to Scotland, although many major British retailers adhere to these restrictions UK-wide (e.g. John Lewis and Marks and Spencer). However, many small (< 3000 sq. ft.) convenience stores and larger food stores in Scotland have extended Sunday opening times as this restriction does not apply to them (UK Parliament 1994). Some large 24-hour supermarkets in Scotland are open all day, seven days a week, including Sunday. These factors are likely to account for the Sunday evening shopping activity represented in the UK TUS data.

Three separate time profiles were created for Monday-Friday, Saturday and Sunday using the given methodology in Section 3.3.2, as three distinctive shopping patterns can be detected in the data available (Figure 3.8). The average shopping duration and travel time from Table 3.2, were derived from the TUS 2000 data.

3.3.4 Estimating destination centroid capacity

This is the first step in estimating a population count for retail dataset to complement the other database library. The following approach has been devised in the absence of shopping data but required to locate shopping activity in time and space. Firstly, the number of retail employees derived from the Annual Business Inquiry dataset (now Business Register and Employment Survey, BRES) (ONS 2006a) are considered. The ABI employee numbers published at LSOA level were re-weighted to OAs. A multiplier is created to allocate shopper numbers in proportion to the daily duration of shopping sustained by each retail employee for OAs:

$$\frac{UK\ Population}{Number\ of\ TUS\ respondents} = 2900 \tag{3.5}$$

Where the 2006 MYE population of the UK is 60,584,300 (ONS 2011) and the number of Time Use Survey respondents is 20,891.

Therefore, based on the above calculation, each respondent is representative of 2900 people in the UK. However, not everybody shops every day. In fact only 42% (8803) of TUS respondents stated that they carried out a form of shopping as a primary activity for the day that was sampled (Table 3.2). In total these TUS respondents carried out 10,662 shopping hours per day, derived from individual diary analysis of the dataset, averaged over Monday to Sunday. This was performed using the total number of diary entries (10 minute intervals) where a shopping related activity was taking place.

Table 3.3 Analysis of TUS respondents engaged in shopping (Ipsos-RSL and ONS 2000)

TUS Respondents who stated their day contained:	Count	%
Shopping as a primary activity	8803	42
Did not undertake any form of shopping	12178	58
Total	*20981*	*100*
Travel for shopping as a primary activity	8141	39
Did not travel for shopping	12840	61
Total	*20981*	*100*

Although 42% of TUS respondents stated that they engaged in some type of shopping as a primary activity, only 39% travelled for shopping (Table 3.3). The discrepancy that the number of shoppers is apparently greater than those who travelled for shopping, may be explained by some people who travelled for other purposes such as work, who then engaged in shopping or whose trips were of very short duration.

The assumption has been made that the TUS is nationally representative of the UK. According to the calculation each respondent equates to 2900 people of the UK's total population. Therefore, on an average day the population of the UK generates 30,920,000 (4sf) shopping hours (10662 TUS shopping hours x 2900 representative UK population).

The following shopping hours have been calculated by scaling-up the hours generated by the TUS for the whole UK population. As the TUS is a UK wide study, the shopping hours are adjusted for England and Wales, which is the scope of the current ABI retail employee data used in this example. The 2006 MYE population of England and Wales was 53,725,800 (ONS 2011), which comprises a proportion of 0.89 of the total UK population. The total UK shopping hours are multiplied by this proportion to represent those generated by the population of England and Wales to give 27,520,000 (4sf) (30,920,000 x 0.89).

$$\frac{E\&W\ Shopping\ hours\ per\ day\ (27{,}520{,}000)}{No.of\ E\&W\ retail\ employees\ (2{,}375{,}140)} = 11.6\ hours\ day^{-1}\ employee^{-1} \quad (3.6)$$

According to the ABI the 2006 retail workforce in England and Wales was 2,375,140 (ONS 2006a). Therefore, on average each employee generates 11.6 shopping hours per day. This was based on the number of employees at businesses that fall within the retail standard industrial classification (SIC) 2003-2007 code.[3] This is a wide category that includes many obvious and unusual sectors such as (amongst others): food, fruit, tobacco, fuel, art, floor coverings, textiles, clothing, cosmetics, pharmaceuticals, household appliances, medical goods, hardware, books, jewellery, sports goods and mobile telephones (ONS 2009). Therefore, further refinements would be possible if retail specific TUS data could be matched to the workforce data in more detail.

A retail employee multiplier (Table 3.4) has been calculated to be applied to the number of employees in each England and Wales OA to estimate the expected number of shoppers based on information contained within the TUS.[4]

Table 3.4 Average shopping duration and ratio to ABI employee

Day of week	Mean shopping duration (Hrs)	Retail employee multiplier
Weekday (Mean)	0.45	2.27
Saturday	2.19	3.12
Sunday	1.72	1.77
Week (Mean)	0.88	2.32

[3] UK SIC codes were revised in 2007 to meet EU regulations for uniformity.
[4] An alternative methodology based on retail floor space is described in Chapter 5.

3.3.5 Locating destination centroids

Georeferenced large user and small business UPCs were used as the destination centroid locations for retail population to represent plausible business locations where shopping is likely to take place. The relevant postcode districts were downloaded from the National Statistics Postcode Directory (NSPD) for August 2006 (ONS 2006b) (Figure 3.9). The study area including the buffer is relatively extensive and covers six postal districts: Southampton (SO), Portsmouth (PO), Bournemouth (BH), Salisbury (SP), Guildford (GU) and Reading (RG). The information required from within the NSPD is summarised in Table 3.5.

Table 3.5 Information obtained from the NSPD (ONS 2006b)

Field in NSPD	Description
UPC	Unit postcode
OACODE	Output area code for which the UPC is situated
OSEAST/NORTH	British National Grid coordinates of each UPC
USERTYPE	User type: "1" (large user, >25 pieces of mail) and "0" (small user)
SMLBUSCT	Small business count: the number of small businesses at each postcode
DOINTR	Date of postcode introduction
DOTERM	Date of postcode termination

Firstly, only current UPCs at the reference date of the library are of interest, so all terminated UPCs were omitted from the dataset which was clipped to the geographic extent of the study area and its buffer. Secondly, all large user UPCs and those with a small business count were retained while the rest were discarded. However, this approach will inherently also include offices and business closed to the public, but currently this method provides a plausible location. The UPC spread largely correlates with the shopping and business districts expected in Southampton. Clusters are evident in the known business districts within the city (Figure 3.9).

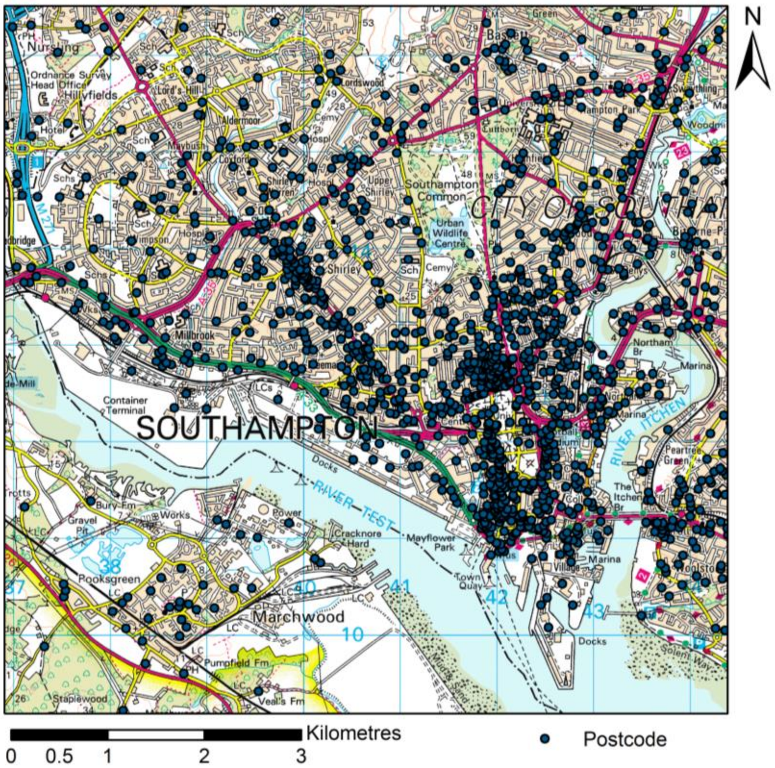


Figure 3.9 Large user and business postcodes from the National Statistics Postcode Directory. Used as destination locations to receive a mobile retail population for Southampton, UK.

The retail customer count calculated for each OA was divided equally by the number of business UPCs present within their respective OA. For this purpose a small number of non-geographic PO Boxes used by some organisations had to be disregarded. As expected, some OAs did not contain any business UPCs and therefore the original OA centroid locations were retained.

The importance of improved destination locations can be seen in Figure 3.10. The estimate of retail footfall derived from the multiplier discussed was distributed onto business UPCs. This provided a more accurate spatial distribution compared to using the original OA centroids. The multiplier to the retail workforce at population weighted residential OA centroids instead of UPCs posed some spatial anomalies. The proposed method provides shopping estimates within each OA. This could be refined if detailed location, retail classification, time use or employee data was available.

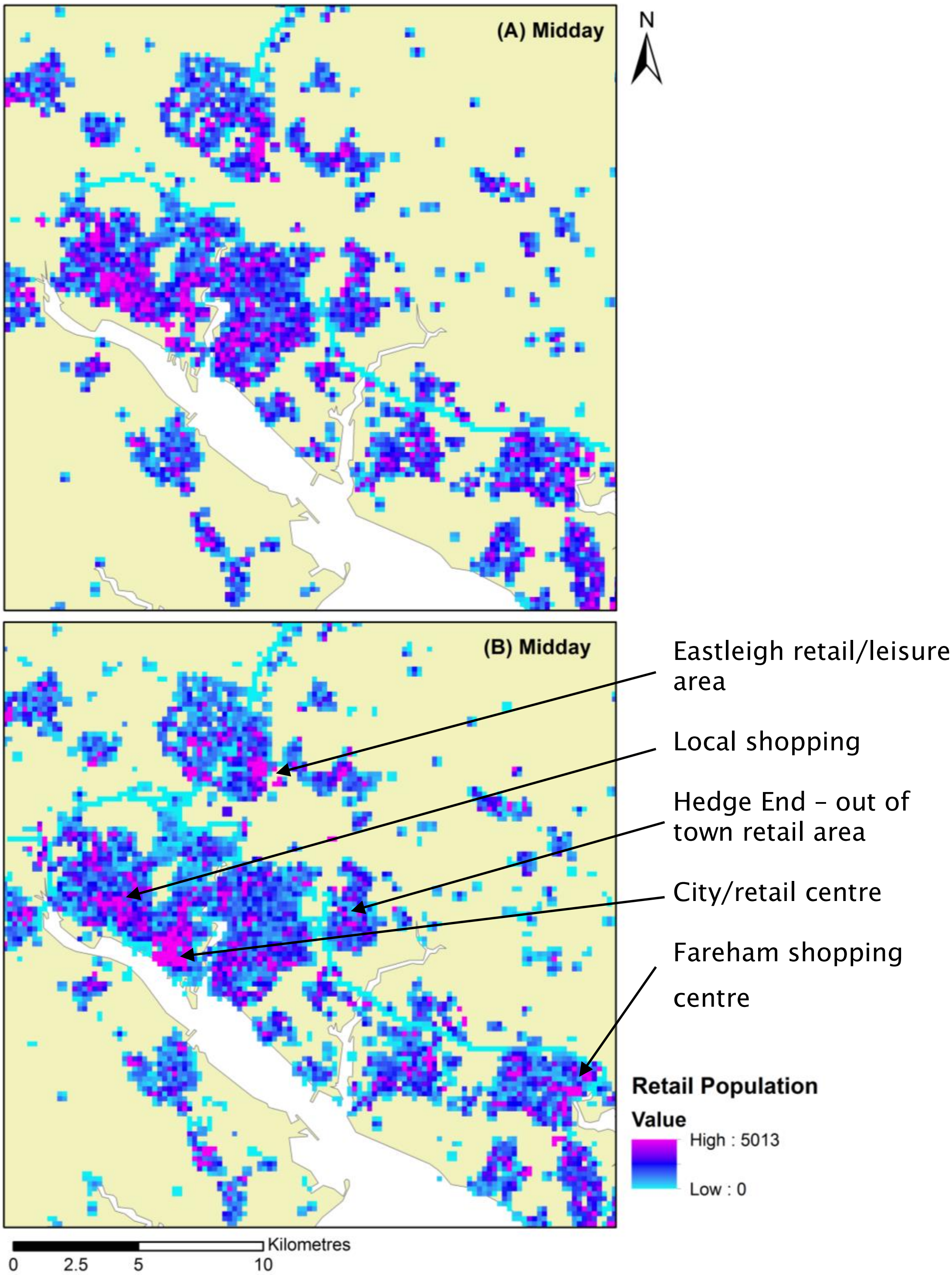


Figure 3.10 SurfaceBuilder247 results for an estimated Saturday temporary retail population at 200 m for the Southampton study area based on (A) OA centroid locations and (B) Business UPC locations. Including the population travelling to and from a retail activity.

3.3.6 Retail Wide Area Dispersion (WAD)

The WAD represents how far people travel to a destination location. There are several commercial datasets and localised studies that give some indication of the average distance travelled to the nearest food store in the UK. For a Bristol study, Sustrans (2006) estimated that 12% of shoppers travelled less than ½ mile, and 40% > 2 miles. The National Travel Survey 2010 states that the average (1995-2010) trip length for shopping in Great Britain was 4.3 miles (DfT 2010).

Time spent travelling for shopping at a 10 minute resolution for the UK is available from the TUS 2000. Travel time has been converted into distance in order to formulate a WAD for this example by assigning an average travel speed. This is likely to impact differentially on estimated distance travelled for those who walk to a shop rather than drive.

A nationally representative vehicle speed estimate of 17 mph has been used based on the data sources detailed in Table 3.6 to estimate distance. Data concerning 'A' roads and motorways in England is available in the DfT's Congestion and Reliability statistics (DfT 2011a). The data are supplied to the DfT by Trafficmaster. The data is obtained from GPS tagged subscribers for traffic information. They feedback live data supplemented by sensors (Trafficmaster 2012). The example WAD created is assumed to be an average representation. However, for small scale modelling an appropriate speed revision by local authority in England could be considered.

Table 3.6 Average urban vehicle speed on 'A' roads during the morning peak (7-10am) 2010/11 (DfT 2011a)

Unitary Authority	Average speed (mph)
City of London	9.2
Bristol	15.5
Manchester	15.8
Southampton	17.0
Liverpool	17.3
Birmingham	18.5
Newcastle upon Tyne	19.1

According to the National Travel Survey 22% of people walk to the shops (DfT 2010). The Accessibility Dataset uses the assumptions that an able person is willing to walk up to 1.2 miles (1.9 km) at a speed of 3 mph (4.8 km/h) to reach a shop (DfT 2011c).

Travelling 1.2 miles at 3 mph would take 24 minutes. Based on the assumption that people are prepared to walk for up to 24 minutes, 22% (The NTS proportion likely to have walked) of the TUS respondents who spent 10-30 minutes travelling to the shops are assumed to be walkers, and have their travel distance reduced accordingly. As 10-30 minutes of walking at 3 mph will produce a shorter, more realistic distance travelled over the same time as those driving or using public transport at 17 mph. Therefore, 22% of the respondents represented in the 10, 20 and 30 minute bins are assumed to have walked at 3 mph. This distance has been adjusted accordingly (Figure 3.11).

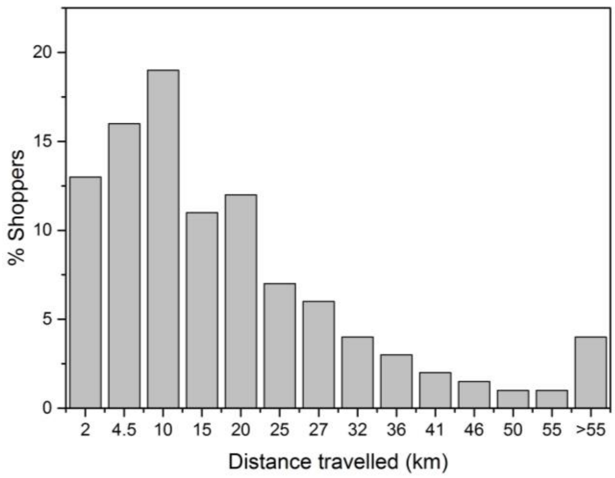


Figure 3.11 Histogram of distance travelled for shopping derived from TUS 2000 shopping behaviour data

The percentages used represent the proportion of the number of people who stated that they travelled for shopping in the TUS on the sample day in question. This consisted of 38.8% of all TUS respondents who travelled for shopping (Table 3.3). Respondents who travelled greater than 23 km were grouped into the final bin as at this point 78% of the respondents had already been accounted for, and this is to improve computational efficiency.

The model has been designed to read population data and produce results for different subgroups. Currently age bands have been used, based on the fact that different age groups are going to have varying behaviours. Furthermore, it would be possible for the user to define their own bands based on subgroups relevant to their application such as gender divides etc. The retail age bands used in this example were taken from the information on the age and employment status of the TUS respondents. As a category existed for ‘full time student over 16’ it was possible to determine the number of students who shopped.

The above approach has been proposed as an option in the absence of available commercial datasets. The assumption has been made that there is a direct relationship between the number of store employees relative to the attracted number of customers. The approach is based on the assumption that a large busy store, such as a supermarket, will employ more people than a small newsagent. However, it currently does not take into account the differing ratio of retail employee to customer for the different types of retail activity. A major retailer might have much more accurate information about a sample of their own customers or a spatial interaction model to estimate regional flows but the exact dataset is not available. The steps outlined above propose an attempt for a retail activity model based on data currently available. This approach resolves the main locations and daily trends, although it could be improved with more detailed information, still provides a more realistic estimate than assuming the population remains at residential locations all day.

The proposed retail dataset is very general and based on some fairly large assumptions; however it provides a plausible estimate where currently there is little data for widespread uniform daytime town centre population counts. There are a number of ways in which this dataset could be further refined if desired for a specific purpose or location. This dataset very broadly concerns ‘shopping and retail’, but it would be possible to further differentiate by retail type. This is important as there is likely to be a significant difference in any assumed employee to customer ratio. A large supermarket is likely to have many customers per employee compared to a small specialist store. Secondly, the time spent engaged in different types of retail activity may be different. People may be more likely to spend longer in a large supermarket, department or home improvement store compared to a small convenience store or newsagent. There may also be a considerably different temporal profile depending on the different types of retail activity mentioned.

A further limitation may be considered in the age of the time use survey which was completed in 2000. The rise of online shopping and changing opening hours in the form of the 24-hour mega-store are likely to have had an impact on customer behaviour since the TUS. A future step to attempt to quantify any behavioural shift might begin with a more in depth look into contemporary retail studies.

The number of retail employees by LSOA according to the ABI dataset assumes that the given number of employees is present every day, not taking into account higher staffed busy periods or shift patterns. Some of the ABI classified retail employees may also take 'head office' roles inflating the number of employees working in a 'shop'. It also assumes that part-time employees are evenly distributed throughout the week. In this implementation employees are redistributed over time in the same distribution as the retail customers.

This approach estimates the number of potential customers to an LSOA by scaling-up shopping hours generated in the UK. These were then divided by the average shopping duration and converted into population counts. It provides one example of an additional activity type. Although the full dataset library of the model has not been described in detail it also contains workplace data, education data by category of institution and health data by type of patient. Datasets produced as part of the Population 24/7 project have been used for these. This can be amended or extended as required. Compared to static day or night-time models the results become very detailed when different temporal profiles and activity measures are included. The case studies presented in Chapters 4 and 5 demonstrate the comprehensive construction of new population data libraries.

3.3.7 **Study area background mask**

A rasterized population masking layer has been constructed using GIS for GB. This is formed of the GB outline preventing population placement offshore (Figure 3.1A). This is particularly important for this coastal study area, and a noticeable improvement from local authority boundaries that extend to the centre line of watercourses. The background mask also includes lakes as uninhabitable water bodies. The road network is represented including motorways and category 'A' roads. The road network is weighted to receive a varying proportion of the in-travel population. The relative weight depending on the date and time is derived from the Department for Transport's National Transport Model statistics (DfT 2010). Therefore a range of background layers have been compiled that adjust the transportation weighting accordingly for the day and time. A Monday at 08:00 will have a greater weighting taking into account commuters than a Sunday at 08:00.

A background mask has been created using national datasets and clipped to the relevant areas of interest. Firstly, a coastal outline, water bodies, motorways and the principal road network (OS Meridian 2) polygons/lines were processed using a GIS. The landmass was rasterised to the required output grid resolution (200 and 100 m) and assigned a cell value of zero. Water bodies and the surrounding ocean was assigned a no data value. This provides the most basic layer on which population are constrained to the landmass for use in SurfaceBuilder247. The model facilitates this by excluding the placement of population in locations containing no data values.

The layer is enhanced with DfT road traffic statistics in the form of count data. Average Annual Daily Flows (AAFD) from the Great Britain Road Traffic Survey (DfT 2013) provide a vehicle count by type, georeferenced location on the road network (count point) and road type. The AADF is the total daily traffic count at a specific point on the road network averaged across a year. These are available for rural and urban principal (category A) and major (motorway) roads. Raw average daily vehicle counts by vehicle type are converted to population estimates based on average vehicle occupancy data (Table 3.7).

Table 3.7 Average vehicle occupancy by vehicle type and time of day (DfT 2011b)

Vehicle type	Weekday (by time of day)				Average		
	0700to 1000	1000to 1600	1600to 1900	1900to 0700	Week-day	Week-end	All week
Pedal cycle[1]	1	1	1	1	1	1	1
Motorcycle[1]	1	1	1	1	1	1	1
Cars and taxis	1.46	1.59	1.53	1.54	1.54	1.88	1.58
Buses and coaches	13.2	13.2	13.2	13.2	13.2	13.2	13.2
Light goods vehicle	1.23	1.23	1.23	1.23	1.23	1.35	1.25
All HGVs	1	1	1	1	1	1	1

[1]Approximatly = 1

The National Transport Model (NTM) (DfT 2005) estimates the distribution of the AADF on the road network depending on the time of day and day of week. This is given in 19 time periods (Table 3.8). This allows the temporal distribution of traffic counts and resultant population estimated to be spread accordingly on the relevant road type.

Table 3.8 NTM time periods used for traffic flow distribution (DfT 2005)

Time Period	Day	Time	Description	Duration (hours)
1	Mon-Fri	00:00-06:00	Off peak	6
2	Mon-Fri	06:00-07:00	AM peak	1
3	Mon-Fri	07:00-08:00	AM peak	1
4	Mon-Fri	08:00-09:00	AM peak	1
5	Mon-Fri	09:00-10:00	AM peak	1
6	Mon-Fri	10:00-16:00	Inter peak	6
7	Mon-Fri	16:00-17:00	PM peak	1
8	Mon-Fri	17:00-18:00	PM peak	1
9	Mon-Fri	18:00-19:00	PM peak	1
10	Mon-Fri	19:00-22:00	Shoulders	3
11	Mon-Fri	22:00-00:00	Off peak	2
12	Saturday	00:00-09:00	Saturday night	9
13	Saturday	09:00-14:00	Saturday Day	5
14	Saturday	14:00-20:00	Saturday Day	6
15	Saturday	20:00-00:00	Saturday night	4
16	Sunday	00:00-10:00	Sunday Night	10
17	Sunday	10:00-15:00	Sunday Day	5
18	Sunday	15:00-20:00	Sunday Day	5
19	Sunday	20:00-00:00	Sunday Night	4

The method followed to produce the background mask layer in a GIS is outlined below:

1. Rasterise land mass (cell value = 0, no data value = -9999) and lake polygons (cell value = no data) to user required resolution (200-100 m).
2. Import vector polyline major and principal road network.
3. Import Annual Average Daily Flow (AADF) counts by georeferenced count point (CP) and vehicle type for all major roads (Motorway and A roads).
4. CP attribute added to indicate whether rural or urban.
5. Proportion of daily flow adjusted according to NTM time periods (DfT 2005).

6. Population estimate based on average occupancy data (Table 3.7) depending on vehicle type and time of day/day of week for all CPs.
7. Population count extracted by CP for target day and hour based on average hourly percentage distribution of daily flow (DfT Statistics).
8. Interpolation tool (e.g. ArcGIS: IDW, Spatial Analyst) used to distribute population count onto the road network (mask). This should be undertaken in two stages for motorways and principal roads. This is in order to prevent motorway counts being spread onto the surrounding road network and vice versa.

Under the UK road classification system category B to unclassified roads have not been considered in this application. This is partly due to the availability and coverage of sufficient count data. Many minor routes are not monitored (manually or automatically). Therefore it is not currently possible to make an informed decision on the traffic flow of minor roads at present. In 2010 the majority (64.1%) of traffic travelled on the major and principal road network (DfT 2011b).

This layer helps to determine the distribution of temporary in-transit locations of population, ensuring that people travelling by foot, car or public transport will be constrained by the road network and not unrealistically spread over empty space in the background layer. A refinement could consider other aspects of the transportation network such as rail that are not currently included.

3.3.8 **Retail data opportunities**

Precise customer footfall data can be purchased and compared with EPOS information. However, difficulties can still arise when attempting to allocate regional footfall data to individual sites. Furthermore, access to such data may also be constrained by expense and geographical coverage. Some UK shopping centres utilise *FootPath*, an intelligence technology that tracks signals from a consumer's mobile telephone using discreet censors. Data are fed back to a central processing unit for analysis that can pinpoint people to within a few metres (Path Intelligence 2010). However, the technology has attracted criticism from civil rights campaigners concerned that such technology constitutes a breach of privacy, although individuals remain anonymous (Morris 2012a).

An alternative to using the number of retail employees could be allocating footfall in proportion to the floor space of the store (see Chapter 5). The Department for Transport has released food store locations with coordinates and floor space for England under the Accessibility Destinations Dataset (DfT 2011c). Similarly, this concept is based on the same assumptions that larger stores, with assumed higher costs, are sustained by higher footfall.

Additionally, or where data restrictions occur, free point of interest (POI) data are available from some commercial satellite navigation companies. Many major store chains release free POI downloads for users containing the locations and names of their stores. These can also contain additional information such as facilities and telephone numbers. Free data from commercial websites can be converted (e.g. using GPSBable: www.gpsbabel.org) into computer readable universal file formats and imported as decimal degrees into a GIS.

3.3.9 **Flood hazard component**

In the preceding sections the population model components (Figure 3.1A) have been discussed with an illustrated example of a dataset creation. This section briefly examines the natural hazard data input and study site context (Figure 3.1B) with analysis of the results (Figure 3.1C).

The study area (Figure 3.6) has been chosen for an example flood risk application. The Solent separates the Isle of Wight from southern England and provides a natural deep water channel for large shipping vessels. Approximately 24,000 properties are considered to be within the tidal flood plain of a 1 in 200 year flood in the Solent (NFDC 2009). Historical record analysis by Ruocco *et al.* (2011) discovered up to 20 flood events in Southampton since 1935. The region’s industrial and shipping success has been attributed to the complex tidal system, resulting in double high-tide each day. However, under storm surge conditions the Solent can experience an increase in sea level of up to 1 m (Ruocco *et al.* 2011). The combination of the region’s topography, location and tidal system has the potential to dramatically increase the flood risk within this area. Southampton Water is a narrow funnel-like channel leading from the Solent to the Port of Southampton, vulnerable to storm surges driven by low pressure systems, or north sea surges that propagate through the English Channel (Wadey *et al.* 2012). When combined with high spring tides these events pose a heightened flood risk and exert pressure on existing defences. This region and accompanying low lying areas contain major coastal transportation links, population centres and commercial and military ports. These activities have contrasting spatiotemporal patterns, making this study area of particular interest for spatiotemporal population modelling.

There are large fluctuations in urban populations over a range of timescales. For example, the two universities within this study site, the University of Southampton and Southampton Solent University, had a combined 2011/12 student population of 38,885 (HESA 2012). Likewise large events or sporting fixtures also contribute to these changes. A major contributor is the Port of Southampton which is the UK's busiest cruise ship terminal and second largest container dock (ABP 2013). In celebration of the 175th anniversary of the P&O company in July 2012, the operator's seven cruise ships left Southampton Docks in formation with an estimated 40,000 passengers on board. Analysis of the Population 24/7 model results suggests that the average weekday population size of this study area is approximately 522,000. This is greatly influenced by people commuting to work and shoppers in key retail locations. WestQuay is a prominent regional shopping centre with an average weekly footfall of c. 300,000 (WestQuay 2011).

Figure 3.1B relates to the incorporation of natural hazard datasets or models. The modelled population density grids (Figure 3.1A) are combined with Environment Agency flood map data (for July, 2012). The Environment Agency is the public body responsible for issuing flood warnings and maintaining flood defence infrastructure within England and Wales. The flood map is the result of probabilistic and scenario-led hydraulic modelling. The most likely scenario under the 'zone three' (high probability) (Environment Agency 2012c) extent (Figure 3.6) has been utilised: this models inundation caused by fluvial and tidal flooding with a 1% and 0.5% annual probability of occurrence respectively. Within the study area, approximately 78 km^2 is at risk of tidal or fluvial flood inundations under the zone three scenario. The modelled population surfaces and flood extents have been analysed using GIS and are stored in a results database with the modelled population data (Figure 3.1C). Potential exposure estimates to the flood hazard are reported. This has been conducted for a range of time slices and age bands. The methodology developed allows the integration of additional hazard maps or more sophisticated hazard models in Figure 3.1B. These results are reported in Section 3.3.11.

3.3.10 **Potential enhancements for spatial resolution**

The example application constructed has shown the appropriate use of census population weighted centroids for use as origin centroids within the modelling process outlined. They represent a realistic distribution of residential population (e.g. Figure 3.3). This has permitted the use of the 200 m output model resolution based on the spatial distribution of these points. By their design PWCs related to residential census populations. To increase model output spatial resolution further (e.g. Chapters 4 and 5) the use of georeferenced UPCs is suggested.

Georeferenced UPCs can be used as one method to increase the spatial resolution of data within the remit of what is currently available. For the residential population, census OA counts can be redistributed onto UPCs that are contained within the respective OA. Full UPCs are available coded to the OA in which they are situated. Census OA counts can been re-weighted onto UPCs in proportion to the address count. The address count is the number of postal delivery locations (assumed to be approximately equal to the number of households) within the UPC. Typically a residential postcode may represent around 15 properties. The proportion of the population allocated to each UPC is derived from the total OA population according to the share of address counts each UPC contains. For example a UPC with address count of 25 within an OA with a total address count of 100 would be allocated a quarter of the OA's population. This technique provides an enhanced spatial distribution of population density at a sub OA level without exceeding the smallest scale in which data are available. Figure 3.12 illustrates an example LSOA and its constituent OAs each with a PWC alongside the residential postcode distribution. This example LSOA, located in Southampton, covers an urban and city parkland area. It comprises of five OAs indicated by the number of PWCs (as each OA only has one). This includes the same OA illustrated earlier in Figure 3.3. Both the distribution of PWCs and UPCs within this LSOA show an accurate reflection of the likely places occupied by the population. They are all concentrated within the urban portion of the LSOA rather than the parkland.

The increased density of UPCs compared to PWCs provides greater detail in the LSOAs population distribution. The same procedure can be followed for non-residential UPCs. The main difference is that businesses usually have their own unique postcode making the redistribution of employees reported by LSOA onto UPCs more straightforward. Where non-residential UPCs contain a small business count it can be taken into account within the overall redistribution weighting.

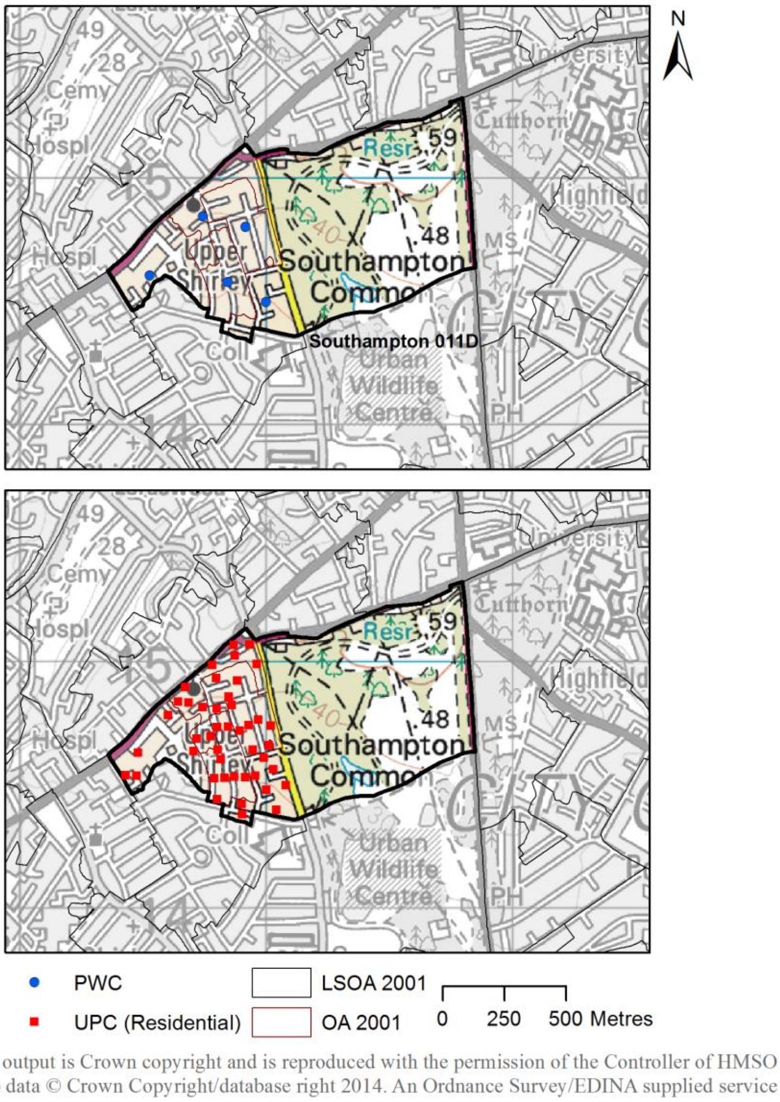


Figure 3.12 Comparison of PWC and residential UPC distribution and density for an example LSOA in Southampton.

Enhancements based on this method have been implemented following the lessons learnt from this applied example to increase the model output resolution from 200 to 100 m for the two case studies (Chapters 4 and 5).

3.3.11 **Southampton applied example: results**

Figure 3.13A shows total population for rasterised 2001 census output areas and the population modelled in 200 m grid cells for three different times of day (Figure 3.13B-D). There is a stark difference between the conventional area-based population model in Figure 3.13A and the gridded representation in Figure 3.13B-D which much more accurately indicates the higher central densities and extensive unpopulated areas, even in this relatively urbanized region. A large daily variation in population occurs. During the working day (Figure 3.13B), population becomes highly concentrated in specific areas such as the city centre and in local clusters such as schools and colleges, as employees and students travel to, and temporally remain at, places of work and study. Population also increases in the non-residential areas of the city centre due to people travelling to engage in other activities as Southampton is a major retail and transportation hub. The 08:00 model differs from the 20:00 model with more people in the transportation network at 08:00 – mostly on their way to work or school (Figures 3.13B and D).

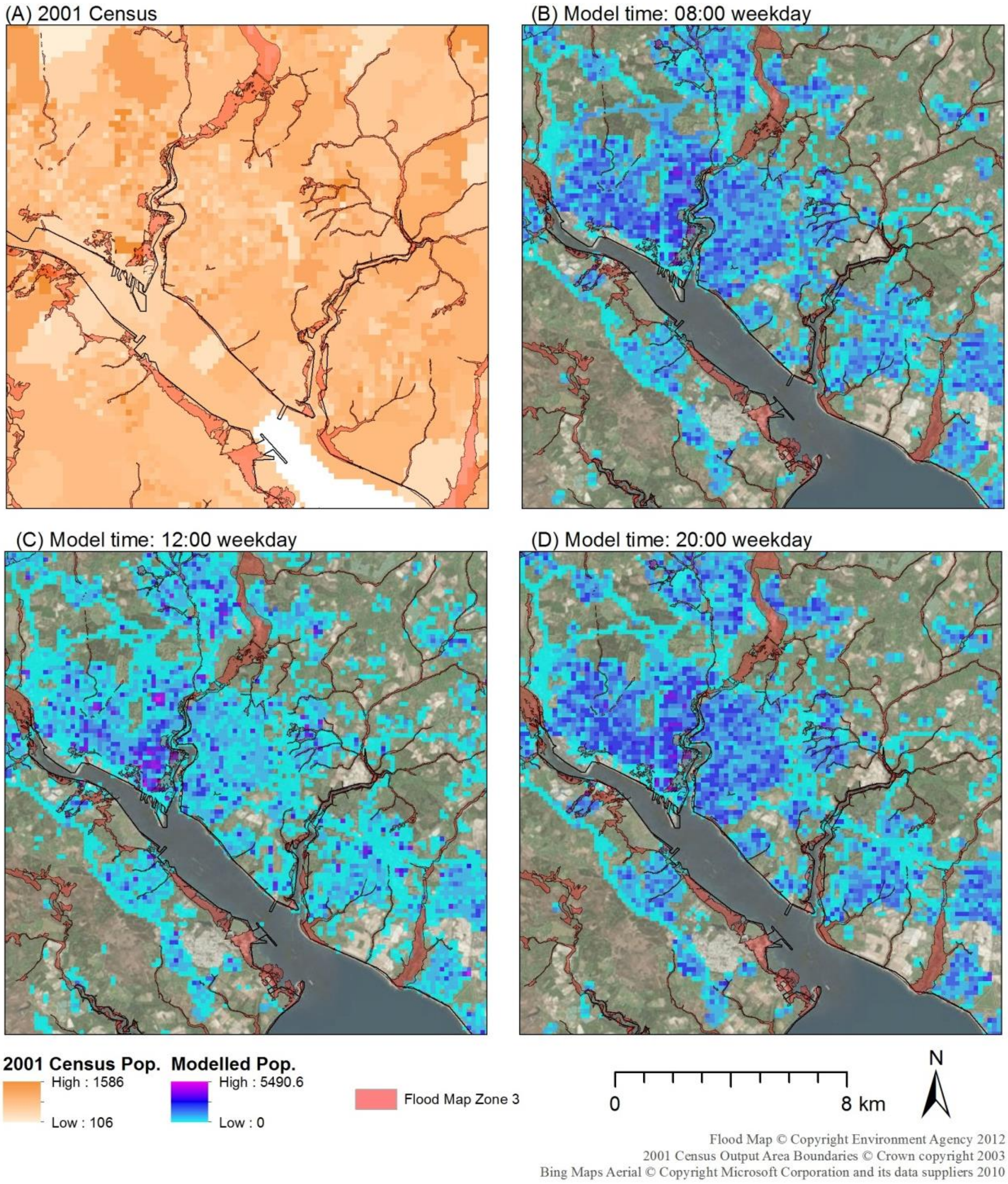


Figure 3.13 (A) Rasterised 2001 census output area total population counts. (B)-(D) Spatiotemporal model outputs showing total population for three time slices, together with Environment Agency flood map data. All maps at 200 m resolution for a 'typical' weekday.

There are spatial (Figure 3.13) and temporal (Figure 3.14) variations in populations potentially exposed during the day. Preliminary analysis suggests that the total population exposed to the Flood Map Zone 3 flood risk peaks towards the end of the typical working day (Figure 3.14). However, differentiating the flood risk components (Figure 3.14) highlights an interesting phenomenon within the Southampton study area demonstrating the power of spatiotemporal population estimates. Throughout the day, exposure to fluvial flood risk closely resembles the reciprocal of tidal exposure with a symmetry approximately centred during standard working hours (08:30-17:00). Tidal patterns have not been accounted for. The pattern observed may be attributable to the region's coastal concentration of industry and commercial activity. As employees commute to the coastal regions during the day their tidal flood exposure increases; when they return home to residential locations further inland in the evening fluvial flood risk becomes the dominating factor.

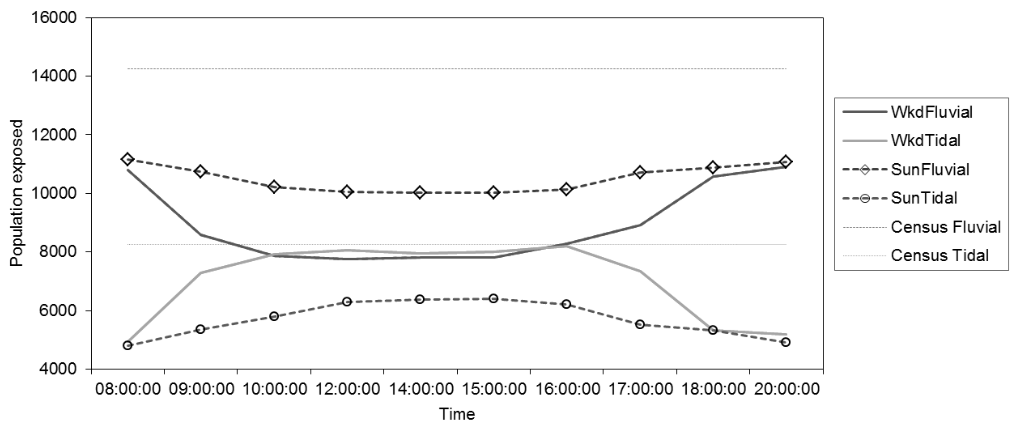


Figure 3.14 Estimated total population exposure by time of day for flood map zone three comparing the static census, a modelled 'typical' weekday and a typical Sunday. Where the annual fluvial flood risk is 1% and tidal 0.5%.

Figure 3.15 indicates the percentage composition by age groups of the total population exposed at 12:00 and 20:00 and also the changes in the size of the exposed population over time. The reversal in exposure between tidal and fluvial flooding is most notable within the working aged population (16-64), where there is a large decrease in tidal exposure in the evening when comparing 16-64 tidal exposure between 12:00 and 20:00. It can also be seen from the heights of the bars that the total population exposed fluctuates over time.

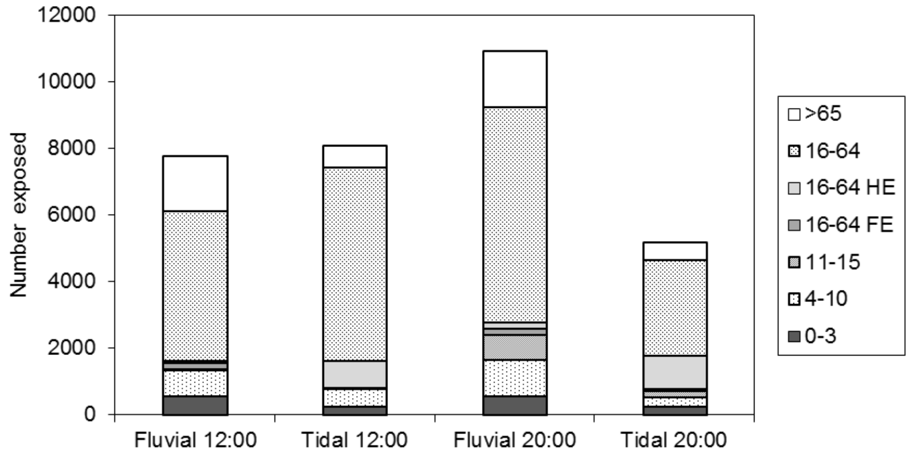


Figure 3.15 Age composition of population potentially exposed to fluvial and tidal flooding during the working day (midday) and evening (20:00). HE: Higher Education; FE: Further Education

Further analysis of the population engaged in different activities has been conducted for the working age population for representative times of 08:00 and 12:00 (Figure 3.16A-C). The figure shows the population in these groups who are travelling in the transportation network or at destination (i.e. non-residential) sites. These were selected to examine the spread of population during the morning commute and middle of the working day. At 08:00 (Figure 3.16B) it can be seen that the working age population at non-residential destinations is relatively low, compared to Figure 3.16C which reflects the concentration of this group at workplaces and other destinations during the day. As expected, when more people migrate into a hazardous zone their exposure is dramatically increased within this age range during the day (Figure 3.16D).

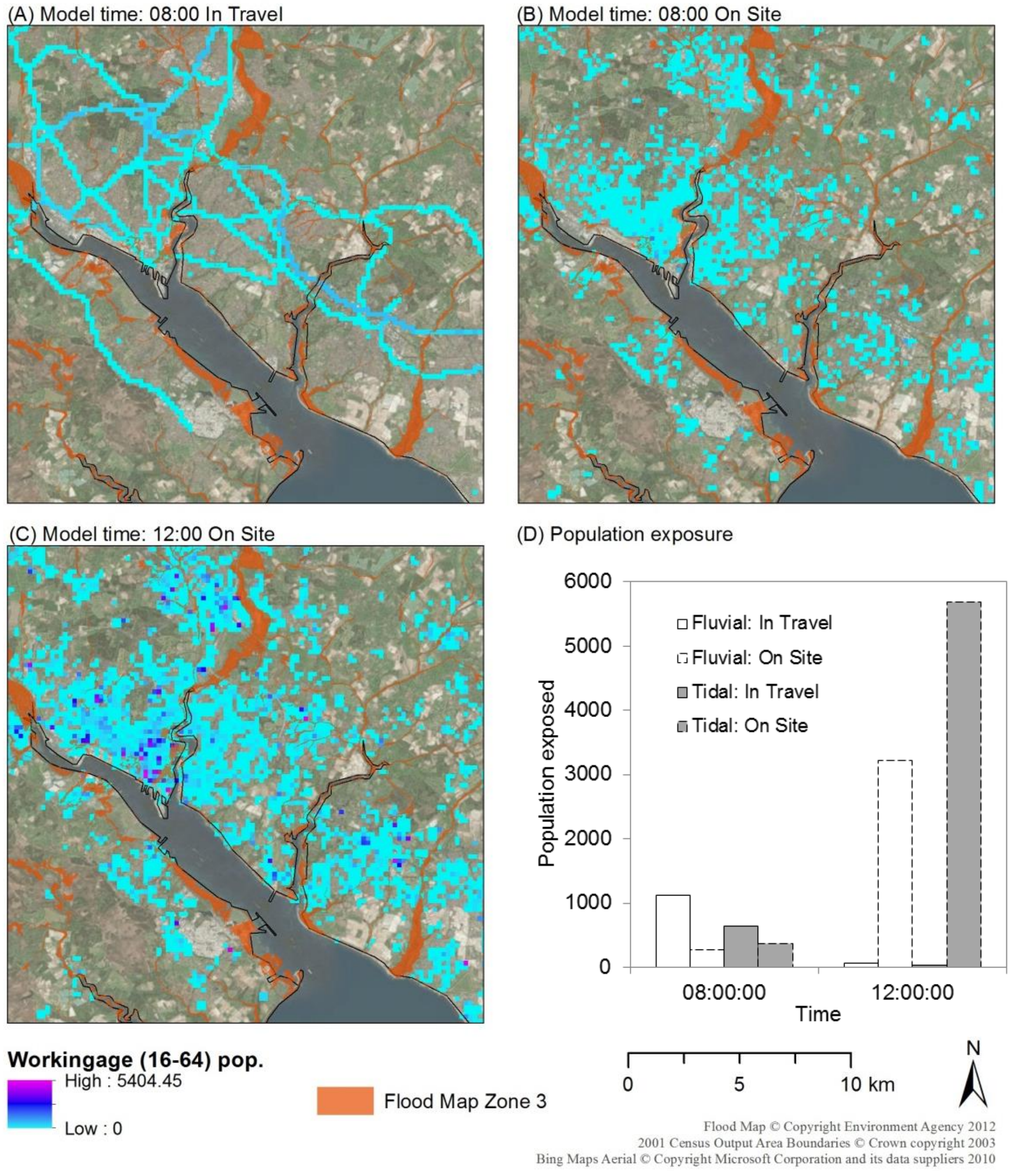


Figure 3.16 Modelled results for (A) the working aged (16-64) population in travel, (B) on-site at 08:00, (C) on-site at 12:00 and (D) the flood risk exposure to the working age population for a typical weekday (cell size: 200 m).

Similarly, the spread of the university student population has been examined for the same typical term time weekday (Figure 3.17). During the day it can be observed that this population is concentrated on the city's two universities and spreads back into the student residential areas during the evening.

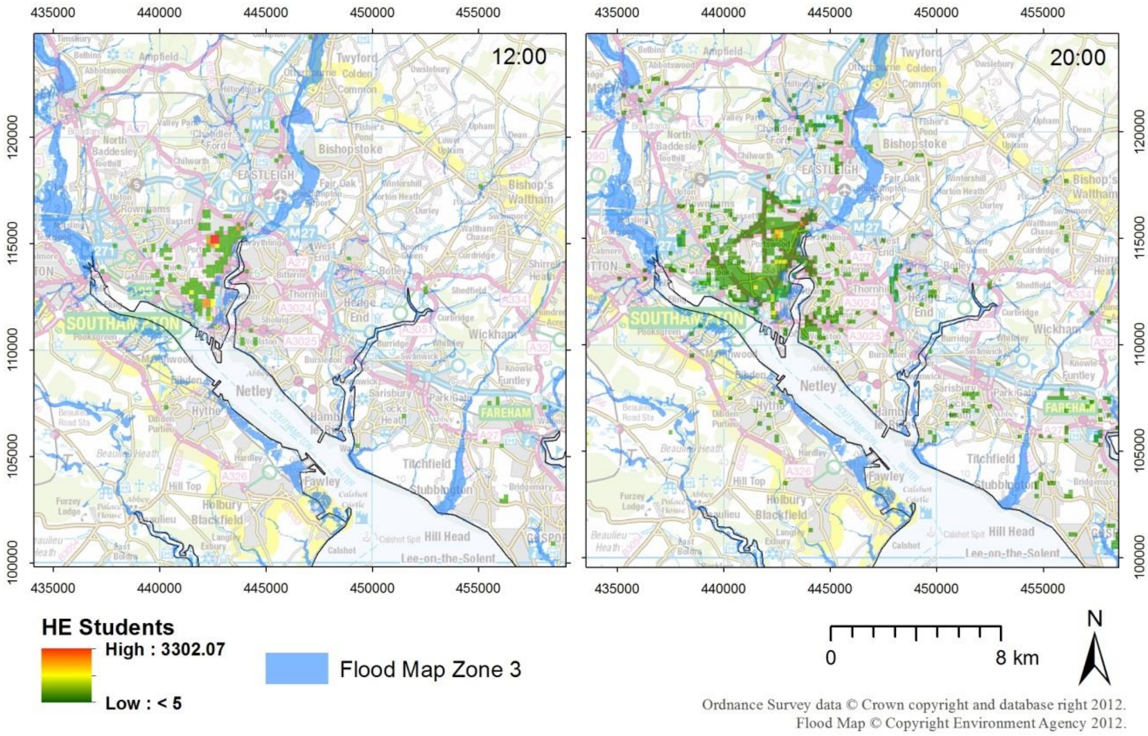


Figure 3.17 Modelled representation of the higher education (HE) student population for a typical term time weekday at 12:00 and 20:00 (cell size: 200 m).

3.3.12 **Observations from the Southampton example**

The modelled outputs contrast starkly with the coarse, static 'night-time' population density coverage given by the 2001 census output area map. The difference between the maximum cell values in the census and modelled outputs highlights the concentrated nature of population density during the working day, which is excluded from traditional census maps. The modelled outputs also provide a more realistic distribution with zero population densities for uninhabited areas.

The selected study site has a high proportion of coastal industry and university students. The diurnal trends in the modelled data (Figure 3.14) may have been predicted, but can be isolated and quantified using the methodology proposed. During 'typical' weekday working hours there is a shift from fluvial to tidal flood risk. It was possible to narrow this change down to a particular population subgroup, those of working age population (Figure 3.15). Population exposure was further analysed for representative times of midday and 20:00. This provided insight into two contrasting points within the usual daily cycle. The reversal from tidal to fluvial exposure occurs in the 16-64 working age category. At midday, tidal risk is the predominant risk to this population subgroup, however by 20:00 this again becomes fluvial. One suggested explanation for this pattern is employees travelling to work in the more exposed coastal locations. On return home to the predominantly inland residential locations fluvial risk becomes the main factor. It can also be observed that there is a notable tidal flood risk to higher education students throughout the day. While there is not such a large reversal of this trend in the evening it does decrease. Possible reasons for this may be the locations of student halls of residence or sites within the institutions. The spread of students in evening residential locations (Figure 3.17) shows that a large portion are within the central area of the city with greater exposure to risks. For vulnerability assessment and the purposes of emergency preparedness this spatiotemporal technique indicates potential improvements over traditional static hazard maps.

The calculated fluvial and tidal population exposure according to the 2001 census is represented by the baselines on Figure 3.14. These can be contrasted with the dynamic modelled results. It can be observed that the fluvial exposure according to the census appears to greatly overestimate flood exposures. However, even for emergency management this provides little insight due to the unrealistic spread of population densities into uninhabited areas which may be at risk of flooding. The modelled results suggest that large portions of the time-specific population at risk are highly concentrated in specific areas depending on the time of day and population age range (e.g. Figures 3.13, 3.16 and 3.17). The spread and concentration of this population at a given time is of more relevance to emergency planners for targeting resources and emergency plans.

The modelled results provide innovative opportunities to make enhanced assessments of subgroups of the population and their activities. It can be used to simulate peak travel times according to the population in travel. This case study exemplifies that different age groups are more at risk at different times. The richness of this detail cannot be interpreted from static census estimates alone.

High resolution disaggregated population counts permit the detailed assessment of impact as well as understanding the wider implications. The flood scenarios described in the Solent study area may not be considered rapid onset events with modern prediction and forecasting tools, but flooding does have the potential to occur rapidly and without prior warning (Murray *et al.* 2012).

In addition to the spatial grid, a structure has been tested that has the ability to handle the storage and representation of temporal information regarding population distributions. Static grids can be readily updated with the inclusion of ancillary datasets and temporal information. The methodology provides a mechanism to estimate temporary and transient populations that are not accounted for by traditional means. The combination of available datasets with the ability to store temporal information produces spatiotemporal gridded representations of populations accounting for local variation (Figure 3.14). The gridded method allows such detailed trends to be resolved (e.g. Figure 3.17).

The Population 24/7 methodology also provides scope to analyse the vulnerability of populations through the ability to model any population sub-groups for which relevant data can be assembled. As already noted, vulnerability is a key contributing factor in the development of risk. The capacity to handle age information may inform management choices for groups such as the young or elderly who may require additional support during an emergency situation. It also improves the accurate allocation of populations in space and time, such as school age children at school locations. This information could be used to target emergency response and health care provision.

This approach cannot provide insights into human behaviour in an emergency situation or provide real time information. It essentially describes predictable population redistribution over time based on data availability and computational efficiency. There are very many ways in which the data sources and detailed distributional models could be enhanced within the current modelling framework. However, the approach could be utilised to inform emergency plans for known risks under a range of scenarios and temporal scales. While it is not possible to know exact future population movements this approach allows the computation of probable distributions. The temporal characteristics of the natural hazard have not been directly addressed in this case study. For example, tidal information indicating a high spring tide under storm surge conditions in the middle of the working day is going to have a greater effect on some of the population subgroups identified in this study.

3.3.13 **Applied example: lessons learned**

This empirical chapter demonstrates the enhanced insights and improvements in the accuracy of exposure estimates to hazards to be gained by combining innovative spatiotemporal population modelling techniques and GIS based layers concerning natural hazard extents. The potential utility of such models for flood risk management has been demonstrated and provides a wealth of detailed data for analysis where the census alone is not suitable for this particular task. Pending further dataset development and validation, this technique has direct application to natural hazard scenarios both within the UK and globally. The data structure of the model allows the user to readily refine or supplement the input datasets. Web data mining and the rise of open-source data are likely to make compiling time-referenced population datasets easier and more accurate in the future. One example of their use could be to refine temporal signals in population movement. An additional step in this methodology would be the integration of a spatial interaction model to enhance the catchment areas of destination locations. The results in this report provide a window of opportunity to further refine this methodology for policy makers and emergency planners and address the key aims outlined at the start of this thesis.

Chapter 4: Case study I - Ulley

4.1 Overview

This chapter introduces the first of two case studies designed to demonstrate the application of the population modelling techniques outlined in Chapter 3. The primary focus of this case study is around the modelling of population exposed to a rapid onset dam failure scenario at Ulley in South Yorkshire, UK, on 25 June 2007 and is described in further detail below. The emphasis concerns the intersection of spatiotemporal population and environmental models. The analysis of the effect of a sudden onset dam failure flood event is presented, while considering a temporally varying diurnal population trend. Hydrological modelling has been undertaken to simulate the dam failure and analysed with results from the population modelling. The additional modelling is required to provide a dynamic insight on rapidly evolving events, where a static risk map alone may not be sufficient.

The rest of this chapter is structured as follows: the first section provides the background and rationale for the choice of case study. The risk of the dam breach at Ulley occurred while the UK was experiencing severe nationwide flooding. The second section describes additional case specific methods and data overview. These concern the creation of a Population 24/7 data library and the construction of embankment breach and flood spreading models. The third section presents the modelled outputs and integrates these with flood hazard data. The fourth section provides some comparison with census data in an attempt to validate model results. A condensed version of this chapter has been published in the proceedings of the British Dams Society (Smith *et al.* 2014b).

4.2 Flood context and background

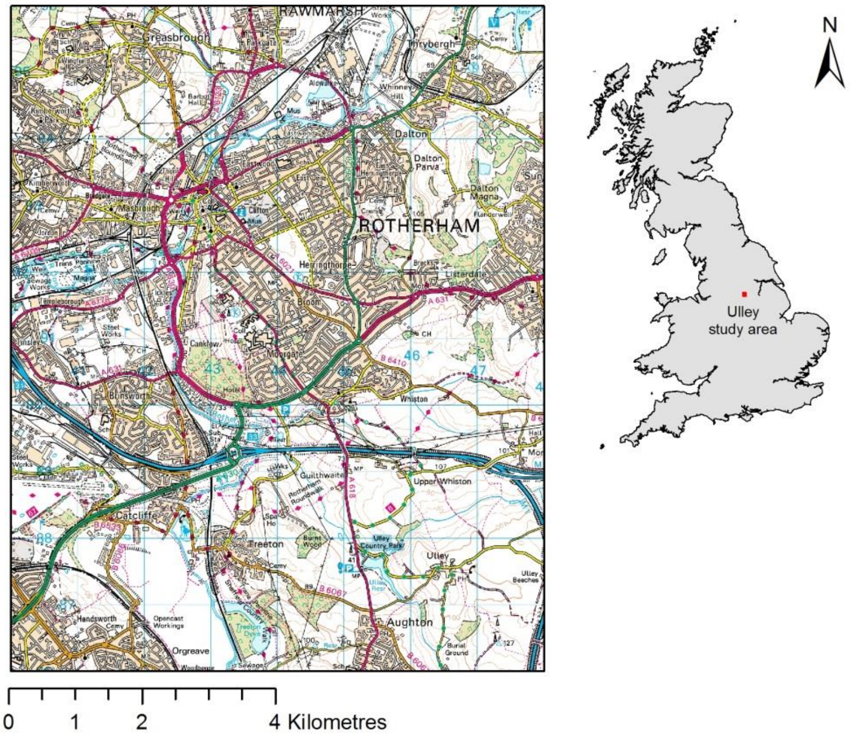
Dam failure events can occur with little or no warning with rapid onset times. This may result in devastating catastrophes in downstream areas (He *et al.* 2008) with little chance to respond. The risk from such events remains high in locations with significant potential for severe losses, although the frequency is fortunately low. Human susceptibility and key infrastructural assets heighten vulnerability and the risk posed from sudden dam failures.

The remainder of this chapter sets out to demonstrate this scenario by re-examining a notable dam failure that actually occurred in the UK in 2007. The failure at Ulley Reservoir in South Yorkshire, a nineteenth century clay-earth embankment dam that was left it in a critical condition, posing a real threat of collapse and narrowly avoiding disaster (Bissell 2010). The near-miss situation became a focus of a period of severe nationwide flooding experienced in the UK that year. It featured prominently in Pitt's (2008) independent review of the flooding experienced. Following the incident at Ulley and Pitt's recommendations subsequent safeguards have been adopted and legislated for. This included a review of reservoir inspection and risk assessment procedures (Porter 2012).

4.2.1 Ulley reservoir and case study location

Ulley reservoir is located three miles south-east of Rotherham and five miles east of Sheffield, Yorkshire, UK. It is presently a country park, owned by Rotherham Metropolitan Borough Council (MBC). Construction of the earth embankment dam was completed in 1874. The supply of drinking water from Ulley ceased in 1986 when it was taken over by MBC as a recreational facility. During exceptional widespread flooding experienced in the UK during summer 2007 the dam was destabilised. A study area has been focused (Figure 4.1) on Ulley reservoir and the surrounding locality.

Increasing industrialisation and population growth within the Yorkshire region during the nineteenth century increased the demand for an adequate and clean water supply. This was driven by the increase in the cotton and steel industries and concerns over healthcare and access to safe drinking water. Poor health and intermittent water supply caused by shortages prompted the construction of the reservoir at Ulley to alleviate these concerns.



OS Map data © Crown Copyright/database right 2014. An Ordnance Survey/EDINA supplied service

Figure 4.1 Ulley study area extent with Great Britain insert

The features described here are typical of many dam constructions. These are basic construction components which are important to understand in order to inform the modelling undertaken and the lead up to the 2007 Ulley flood event. Ulley dam was constructed by Messrs Lawson and Mausergh of Westminster between 1871 and 1874 (Arup 2008). Like many dams of similar design within this region it consists of a single earth embankment with puddle clay core. ‘Puddle’ is a watertight clay-based material. The puddle core forms an impermeable barrier at the centre of the embankment with the base excavated well into the ground to prevent water seepage underneath the core. At Ulley the top of the core has a width of 6 feet above the water level tapering out to a thickness of 15 feet at its base (original technical drawings provided by MBC). The core is supported by a surrounding layer of ‘select material’, usually a sand-based mixture or loam. This is intended to deter burrowing animals that may otherwise excavate into the core and compromise the integrity of the dam. Finally, this is surrounded by earth as the outermost and protective layer to form a traditional embankment as seen in Figure 4.2A. The reservoir is relatively small at approximately one mile in length with a capacity of 580,000 m^3 (Arup 2008).

Until 2007 Ulley had three spillways (two being disused) (Figures 4.2A and E). Spillways are channels that facilitate the controlled release of water from a reservoir (Figure 4.2C). These prevent dams from overtopping and allow the response to the variability of water levels based on rainfall and other inputs. The original 1874 spillways were in the form of a symmetrical pair of stepped masonry channels, masonry blocks joined (pointed) with mortar. These were disused, but would have channelled flow across the embankment from either side (see Figure 4.2A). These were superseded by a concrete stepped spillway constructed in 1943 which channels overflow away from the embankment. In addition to the spillway, like most dams, Ulley also has a scour pipe. This is a pipe or tunnel outlet that runs through the bottom of the dam and embankment which can be used to lower the reservoir water level quickly in an emergency (BDS 2010), such as during the threat of imminent dam collapse. This allows water to be removed from the reservoir and piped downstream of the embankment.

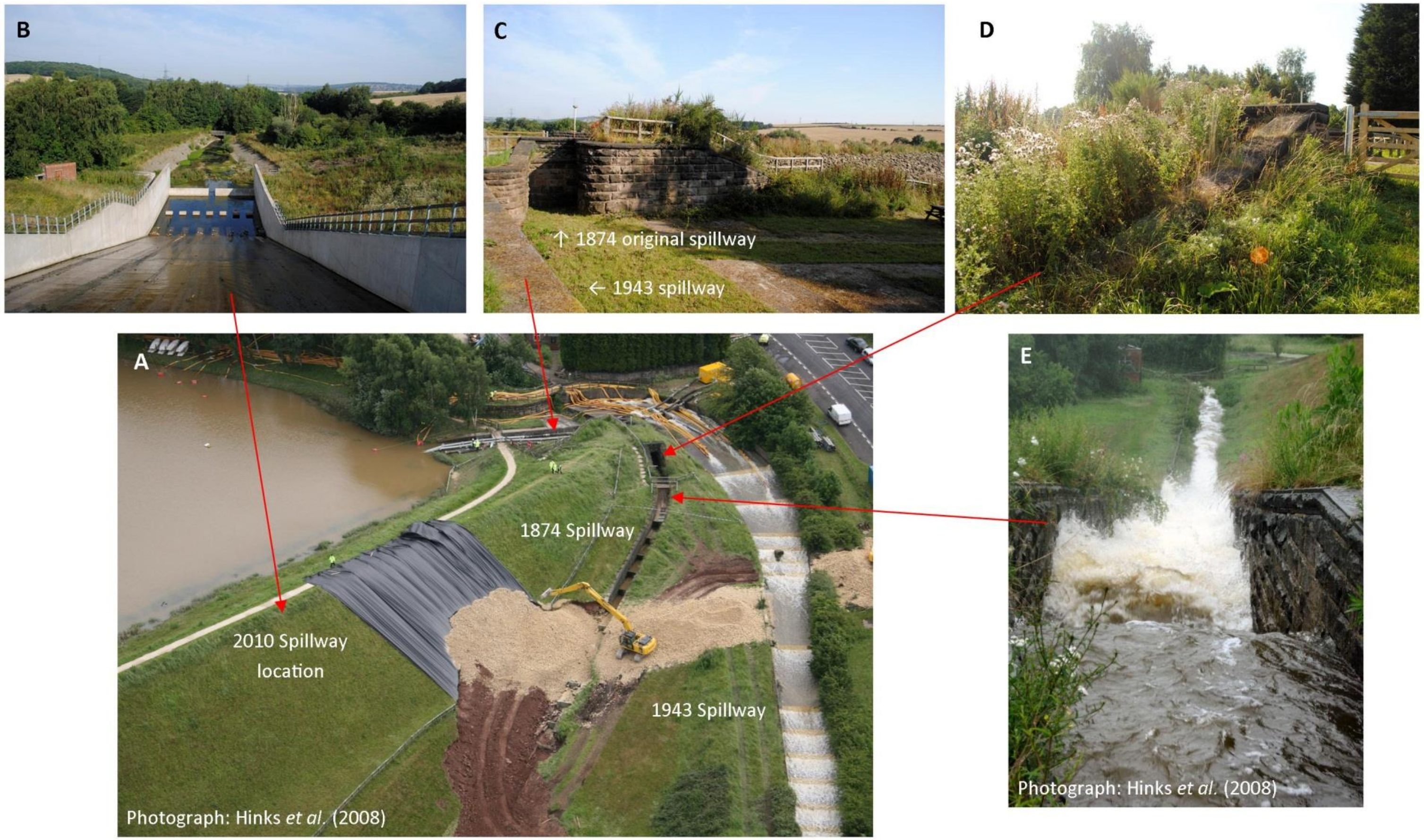


Figure 4.2 The anatomy of Ulley dam (A) Emergency stabilisation June 2007, limestone filled scour hole, (B) Replacement 2010 concrete spillway, (C) Remaining channel entrances to the original 1874 masonry stepped and superseding 1943 concrete spillways (now above maximum water level), (D) Entrance ramparts comprising the only remains of the 1874 spillway (the rest was removed post-2007), (E) June 2007 storm overflow in the 'unused' 1874 spillway. Photographs: Alan Smith, August 2014 (unless cited).

4.2.2 June 2007 incident

On 25 June 2007 a slow moving depression bought prolonged heavy rainfall to northern and central England, with more than 90 mm of rain falling in 18 hours (Environment Agency 2007b; Met Office 2011). June 2007 was the wettest in England and Wales since 1860 (Marsh and Hannaford 2007). Intense slow moving frontal rainfall on the 25 June fell on saturated ground with some rivers already exceeding capacity and reservoir levels high. It is estimated that the rainfall levels that led to this event had an annual probability of occurrence of 1% (Warren and Stewart 2008). The prolonged rainfall had already caused widespread flooding in this region. The impact of the collapse of the Ulley embankment would have been exacerbated by significant volumes of standing floodwater already within the catchment (Section 4.2.3, Figure 4.3).

The mechanics of the events leading to the risk of destabilisation at Ulley have been well documented (e.g. Hinks *et al.* 2008; Mason and Hinks 2008, 2009). Despite a larger concrete spillway having been constructed in 1943, flood water reverted to the original masonry stepped spillway which intercepted the main earth embankment. This in itself represented an original design flaw as it exposed the vulnerable front face of the embankment to an unnecessary risk of erosion (Porter 2012) and subsequent destabilisation.

High outflows caused by prolonged rainfall on saturated ground overtopped the 1943 spillway and caused the reactivation of the original 1874 masonry stepped spillway that it had replaced. The masonry spillway, which was unused but remained in situ, suffered deterioration of its channel after reactivation due to overspill from the 1943 spillway (Figure 4.2E). The hydraulic pressure of the overspill flow, the force exerted by the flow, that had managed to reactivate the original spillway, exceeded the retaining wall threshold (its resistance to spillway flow pressure) causing it to collapse and facilitating the erosion of the dam embankment material (Warren and Stewart 2008). This caused a 20 x 6 m scour hole in the face of the embankment (Oliver and Owen 2007), putting the stability of the entire dam at risk (Figure 4.2A). During the flood, peak flow on the failed spillway was estimated at 6.1 ms^{-1} (Figure 4.2E) (Horrocks 2010). Rotherham MBC was advised to take immediate emergency action to prevent major flooding downstream following destabilisation of the dam (Environment Agency 2007b).

The larger 1943 concrete spillway was constructed following an earlier inspection finding that the original 1874 spillway lacked sufficient capacity (Arup 2008). Following the 2007 incident the previous spillways (1874 and 1943) were removed and replaced with a single central concrete spillway, which was completed in 2010 (Figure 4.2B). The latest concrete spillway permits a further increase for peak flow capacity on the 1943 rate as well as greater resilience to turbulent flows compared to masonry spillways, particularly if the mortar pointing between blocks has been allowed to deteriorate. The capacity of the scour pipe was increased to twice the original capacity. It can now drain 40,000 m^3 day^{-1}, enough to lower the reservoir water level by 1 m per day (Horrocks 2010).

Approximately 1000 people were evacuated in downstream areas of the dam from the villages of Catcliffe, Whiston and Treeton. he M1 motorway was closed northbound between junctions 32 and 34, and southbound between junctions 34 and 36 (Sturcke *et al.* 2007) for 40 hours at an estimated cost of £2.3 million (Environment Agency 2007b). In addition to the population exposure there was also a substantial risk to critical infrastructure and assets. These included a high pressure gas main, high voltage electricity pylons, a regional substation, telecommunication towers, highways, water treatment works and the M1 motorway.

Emergency work to re-stabilise the dam and reduce water levels continued before the motorway was reopened. The initial remedial action involved packing the scour hole with 2,500 tonnes of crushed limestone and pumping water from the reservoir into the 1943 spillway to lower the reservoir level. Repair of the dam cost £3.8 million and resulted in the construction of improved scour pipe capacity, and a new reinforced concrete spillway in the centre of the dam (Horrocks 2010).

4.2.3 Antecedent downstream conditions

The collapse of Ulley dam would have added the reservoir's volume to an already inundated downstream catchment (Figure 4.3). Measurements referenced from aerial imagery were used to confirm the pre-existing water level at 31 mAOD (Above Ordnance Datum), which was subtracted from a 2 m LiDAR DTM to estimate depth.

Figure 4.3 Aerial photograph (08:23 26/07/2007) showing antecedent flood extent on Ulley Brook and the River Rother (M1 junction 33 image centre). Red flood water discolouration caused by suspended sediment from Ulley embankment erosion (upstream) Photograph: Hinks and Mason (2007)

The downstream catchment was revisited and present-day photographs taken against static features used to further estimate flood depth. These were compared to photographs taken at the time of the incident by Lomas (2007) (Figure 4.4). The photographs taken in Catcliffe show an antecedent floodwater depth at this location of 1.2-1.5 m. The red telephone box (located on the B6066 Orgreave Road roundabout, Catliffe, BNG: 442532, 388467, Figure 4.4D) is the same one that is visible in Figure 4.4A-C. The modelled flood water depth at the same location 1.46 m (see Section 4.4) and can be corroborated by these observations.

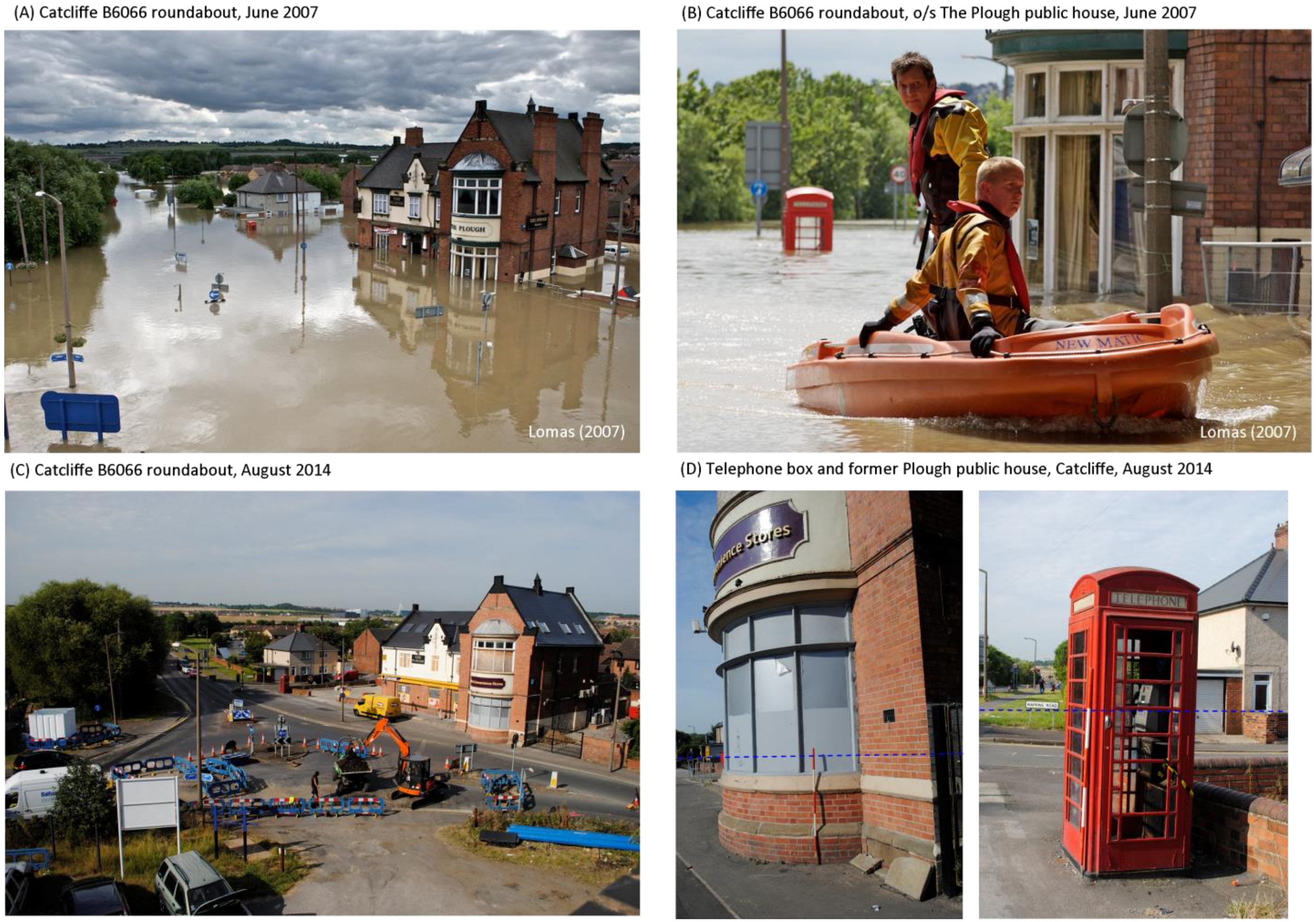


Figure 4.4 Estimating the June 2007 floodwater depth downstream from Ulley (survey pole 1.5 m). Blue dashed line: estimated water level (D). Photographs (C and D): Alan Smith (August 2014).

4.2.4 Yorkshire dam failures

Unprecedented, catastrophic dam failures are not unknown in the region surrounding Ulley. The complete collapse of Dale Dike Dam in 1864 (13 miles west of Ulley) caused the Great Sheffield Flood, resulting in considerable downstream destruction and 244 fatalities. The dam collapsed under severe storm conditions while being filled for the first time. The breach resulted in the discharge of c. 3 million m^3 of water (Amey 1974) into the narrow catchment below. The embankment was of the same earth/clay construction type as Ulley, which was to be constructed less than ten years later. In June 2007, under the same conditions that contributed to the Ulley incident, flooding on the River Don at Sheffield was also at its worst extent since the 1864 collapse of Dale Dike dam (Environment Agency 2007a).

On 19 June 2005 the same mechanism that was to be repeated at Ulley only two years later resulted in a dam failure at Boltby Reservoir, North Yorkshire (60 miles north of Ulley). Boltby, also a nineteenth century earth embankment dam, had a considerably smaller capacity than Ulley at 130,000 m^3 (Porter 2012). However, the reactivation and failure of a masonry stepped spillway under a storm flow eroded the dam's embankment material to expose the puddle clay core and risk collapse. The use of the reservoir as a drinking water supply had already ceased in 2003, prior to the incident. Therefore, a decision was made to permanently drain the reservoir to negligible levels rather than attempt a costly repair. A 'v' shaped notch was cut into the embankment to prevent the reservoir from re-filling (Walker 2008). These incidents highlight the important nature of reservoir risk management and the potential for catastrophic impact on the exposed human populations that live downstream. In all three examples at Dale Dyke, Boltby and Ulley the requirement for a rapid response with little (or no) warning to intervene and prevent a disaster is clear. This small snapshot of dam failures in this specific locality demonstrates that any subsequent disaster has the potential to have a high impact in both human and economic costs. The sensitivity of the time at which an emergency might occur is important for the potentially exposed population. Estimating populations in time and space, such as the application of the Population 24/7 techniques, provides a method to assess and analyse these sensitivities.

4.3 Case specific method and data

This section introduces further methods and data specific to this case study which are in addition to what has already been presented in Chapter 3. The additional hydrological modelling outlined has been supported through collaboration with HR Wallingford Ltd. Figure 4.5 provides an analytical overview of the methodology. This is split into three categories: (A) modelling the embankment breach for Ulley under the conditions of the 2007 incident, (B) modelling the resultant flood inundation and (C) bespoke considerations for the population modelling component. These will be discussed in turn within this section. The embankment breach modelling using EMBREA (A) provides an estimate of the dam's outflow hydrograph should it have failed. This is a required input parameter for the flood inundation modelling (B). This major flow input from the dam breach influences the extent of flood spreading and downstream flow velocities and depths. The output from the inundation modelling using TELEMAC-2D consists of flood extent, depth and velocity. This is combined with 100 m gridded spatiotemporal population estimates using the Population 24/7 approach (C). The combination of the estimate of potentially exposed populations that vary by time of day, flood flow velocities and depths facilitates the analysis of the risk posed to people. This is described in Section 4.3.5 in the evaluation of a flood hazard rating.

The ability of the Population 24/7 approach to produce population estimates for varying time intervals allows the observation of how the risk posed to people by a rapid onset dam failure is time sensitive. Furthermore, the gridded population output permits the integration with such environmental models which could not be achieved with static population data reported in arbitrary, variable spatial units. An 8 x 10 km study area (Figure 4.1, BNG origin: 440000, 386000 m) centred on Rotherham encompassing Ulley Reservoir to the southeast has been selected. This covers the population immediately downstream of the reservoir that could be exposed to a dam failure event.

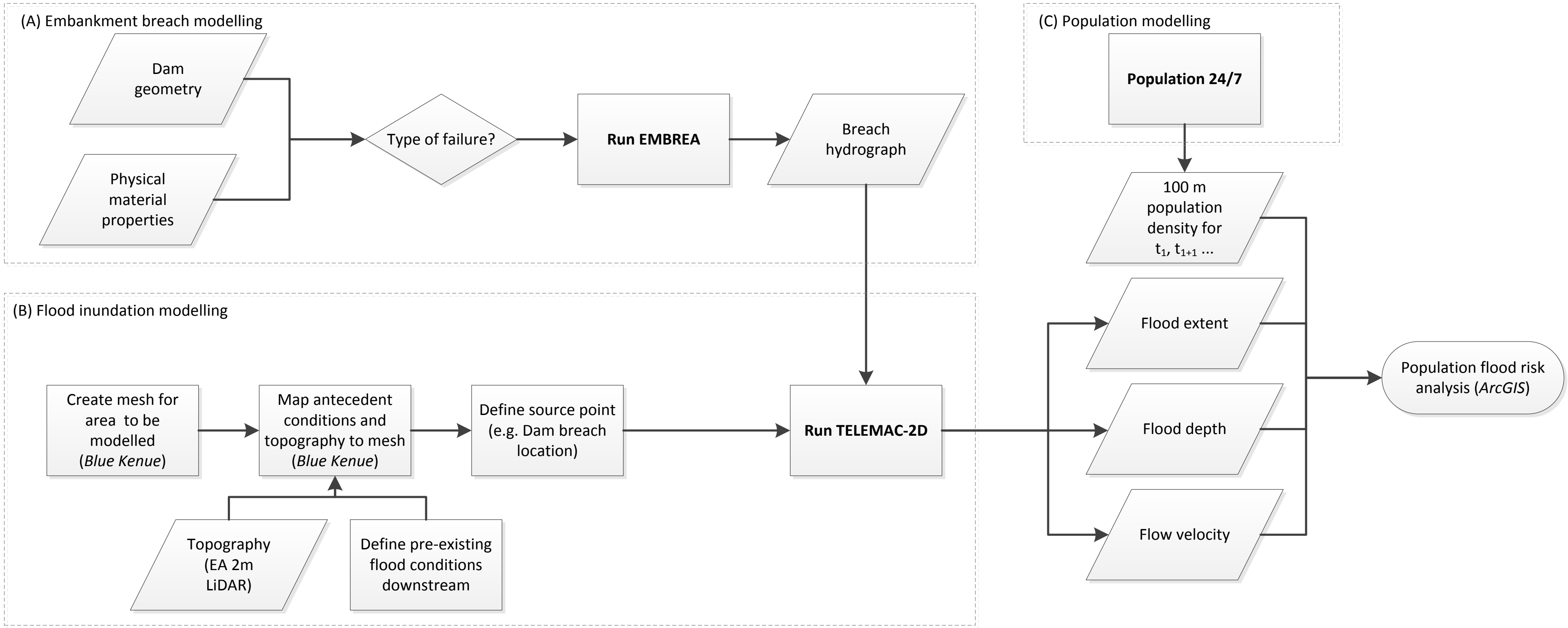


Figure 4.5 Analytical overview of (A) flood risk modelling components and analysis in terms of embankment breach modelling, (B) flood inundation modelling and (C) the population component. Model: bold type, specific software: (italicised in brackets).

4.3.1 Slope stability analysis

Before any breach analysis for Ulley was conducted the warnings of imminent collapse were evaluated to confirm whether a breach (i) could be likely and (ii) if so, how it would occur. This step is required before any modelling is commenced. The outcome of the slope stability analysis will be used to inform the type of failure parameter in the embankment breach modelling (Figure 4.5A). There is a general acceptance that the dam almost certainly would have been breached if the failure of the embankment had been allowed to continue unchecked or unless immediate remedial action was not taken (e.g. Oliver and Owen 2007; Mason and Hinks 2008; Horrocks 2010). This was required to prevent further damage and relieve pressure on the structure, exerted by the water contained behind the embankment.

The physical condition of the embankment as surveyed post-incident by the geotechnical engineering firm Arup revealed a number of weaknesses within the embankment structure (Arup 2008). The clay core was found to contain higher than expected amounts of silt, making it more susceptible to erosion. There was also the possibility that cracks had formed in the core allowing water to permeate through the embankment. The height of the clay core was increased in 1969 with a concrete extension, however this was found to have deteriorated and was judged to be life extinct. The report also found that the select material surrounding the core was mostly indistinguishable from the rest of the embankment material. Finally, there was evidence of seepage at the back of the scour hole indicating that water had indeed penetrated the core and saturated the embankment, decreasing its stability.

An analysis of the embankment's slope stability under these conditions was conducted using the software SLOPE/W (Figure 4.6). This resolves both moment and force equilibrium equations to calculate a factor of safety (FOS). In this context the FOS is defined as the ratio of total available shear stress of the soil within the embankment to the shear stress required to maintain equilibrium along a potential slip surface (USDOI 2011). Thus it is a ratio of stabilising and destabilising forces that gives an indication to the probability of failure of an embankment depending on a set number of conditions.

The geometry of the embankment at Ulley (derived from original technical drawings) was created in SLOPE/W and represented as a homogenous earth embankment with puddle clay core (Figure 4.6). The select material was not treated separately because it had been deemed indistinguishable in the post-incident geotechnical report (Arup 2008). The geometry was altered to represent the removal of the supporting toe material, earth supporting the base of the embankment, by the scour hole. As the geotechnical report also indicated clear evidence of seepage the cohesion, resistive property, of the embankment material was reduced to simulate full saturation (Figure 4.6). This occurs when water has penetrated the puddle core and fills the pore spaces within the embankment material. This acts to reduce the supporting material’s cohesive strength and increases the embankment’s likelihood of slope failure under gravity.

The SLOPE/W analysis suggests that the supporting embankment material would have been liable to slipping following erosion of the toe material under the 2007 incident conditions if remedial action had not been taken. This would have exposed the core and the horizontal force equilibrium would be lost causing it to be overcome by destabilising forces without the support of the downstream embankment structure. The Ulley embankment was given an FOS of 0.718 in the scenario where undercutting had been permitted to continue (Figure 4.6). Removal of the toe material is likely to have destabilised the embankment. An FOS >1.5 is considered satisfactory for dams while <1.0 is unsafe (FERC 2005). Undercutting of the embankment material was initiated during the 2007 flood event but was fortunately prevented from worsening following emergency remedial work. Should the breach have continued, these preliminary core stability calculations suggest that the core would have failed.

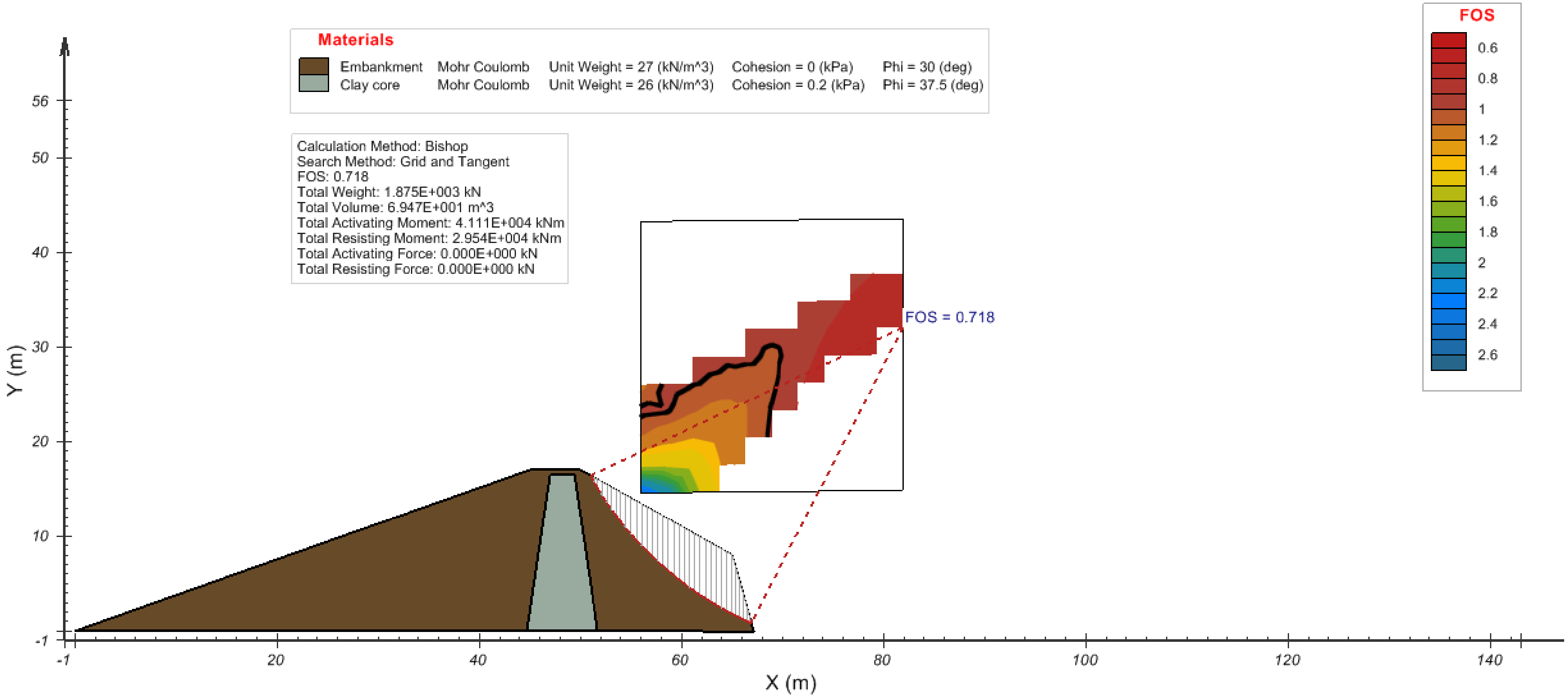


Figure 4.6 Ulley dam cross-section and slope stability analysis using SLOPE/W

4.3.2 Embankment breach modelling (EMBREA)

The EMBREA (EMbankment BREAch Assessment) complex model, developed by HR Wallingford Ltd., was used to undertake the breach analysis for the embankment at Ulley. EMBREA is a research and development model commercially available as HR Breach (FLOODsite 2009). It was used to simulate the failure mechanism of the dam and derive the resultant outflow hydrograph. This is necessary to predict the impact of the dam failure and it is a key parameter for the resultant flood spread modelling (Figure 4.5A) (see next Section 4.3.3).

There are several reviews of dam breach models (e.g. Mohamed 2002; Mohamed *et al.* 2002; Wahl 2004) which evidence an established history of breach modelling. However, these types of models, and therefore their outputs, are often subject to large uncertainties (Froehlich 2008). Given the varied nature of dam and embankment construction designs, materials and types there is not a 'one size fits all' modelling approach. EMBREA was selected for this case study as it is most appropriate for earth embankment dam breach simulations. This is in part due to its ability to handle structures with multiple earth layers such as a clay core, select and embankment material. It also permits a range of failure options for layered embankments. The most likely failure scenario for Ulley (Figure 4.5A parameter) is embankment overtopping. This is based on the evidence of geotechnical reports and slope stability analysis (Section 4.3.1). Overtopping occurs when the embankment core fails, resulting in the release of the water contained behind the dam. The slope stability analysis suggested that the core would fail when the supporting embankment material was removed through scour and slope failure (Figure 4.7).

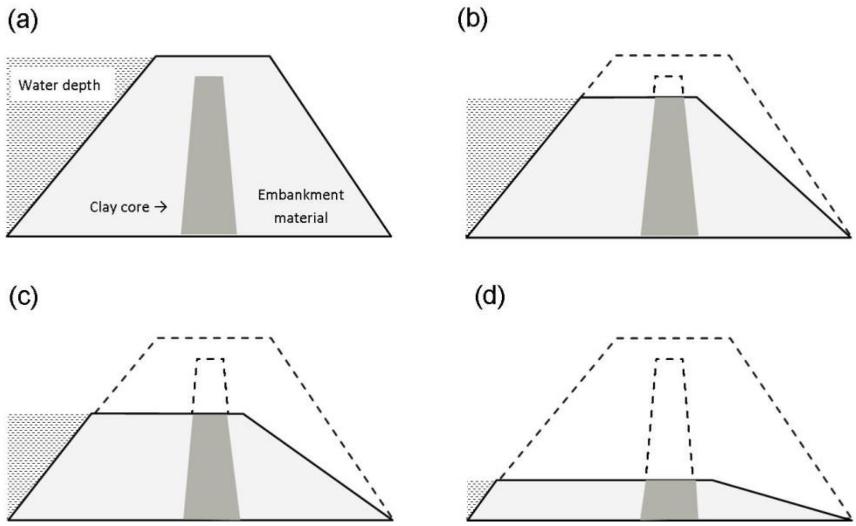


Figure 4.7 Cross-sectional schematic of the embankment failure process on an earth embankment dam. Where (a) is the initial condition and (b-d) represent the first to third stage failures of the embankment.

A block failure occurs (Figure 4.7B) when an exposed section of the core is destabilised due to the initial slip of supporting embankment material. This results in a breach flow, a sudden release of dammed water which causes further removal of embankment material supporting the core. With increasing exposure of the core, stresses in the core increase and subsequently give rise to further block failures (Figure 4.7C and D) as the embankment material is eroded.

The required input parameters used in this study for EMBREA (Figure 4.5A) are summarised in Tables 4.1 and 4.2. The output is a breach hydrograph (Figure 4.8).

Table 4.1 Input parameters for EMBREA

Parameter	Description	Source(s)
Upstream condition	Reservoir inflow	Not used
Volume stage curve	The volume of water within the reservoir	Arup technical report
Dam geometry	Dimensions of core and layers	Original drawings
Embankment material properties	Physical material properties (see Table 4.2)	Rotherham MBC geotechnical report

The upstream condition parameter (Table 4.1) was not used as the inflow to the reservoir was not deemed significant given the enormous magnitude of a potential breach. Furthermore, on un-gauged inlets it would be arbitrary to quantify a highly variable estimate based on current and previous conditions within the catchment. Like the 2007 incident, the model scenario assumes the reservoir is already full and at overtopping capacity. Therefore any minor additional input is inconsequential during the initial stages of slope failure considered in this case.

The values defined for the physical properties of governing the dam breach modelling parameters (Figure 4.5A) in Table 4.2 have been based upon specific information from within the geotechnical survey (Arup 2008). Where these are unknown expected and default values have used under the guidance of consultant dam engineers at HR Wallingford. For example the cohesion of the clay core is much greater (by intended design) than the surrounding saturated select material (as anticipated under the June 2007 conditions). The friction angle, dry unit weight, erodability, plasticity index and Manning's n (a coefficient representing channel friction) have been assigned standard values based on the composition of materials used in earth/clay dam embankments.

Table 4.2 Dam physical material properties

Property	**Clay core**	**Select material**	**Source(s)**
Median particle size (D_{50}) (mm)	0.05	0.2	Arup (2008)
Porosity (%)	0.37	0.37	Arup (2008)
Dry unit weight (kN m^{-3})	24	24	
Friction angle (deg)	30	30	
Tensile strength (kN m^{-2})	0.01	0.01	
Cohesion (kN m^{-2})	100	0.1	
Erodability coefficient (cm^3 Ns^{-1})	0.01	5	
Plasticity index	0	0	Default
Manning's n	0.025	0.025	Default

A selection of model parameters based on the dam's physical properties are summarised in Table 4.2. The dam has a crest height of 16 m and a volume of 580,000 m^3 (Hinks *et al.* 2008). Whilst there are inherent uncertainties in any form of modelling there still remains a known volume of water within the reservoir that will cause significant flooding downstream following a breach.

Figure 4.8 is the breach hydrograph derived from EMBREA following successful modelling of the embankment breach through overtopping. This is a direct input into the next stage of flood spread modelling required for the risk analysis (Figure 4.5B). The hydrograph shows breach discharge against time from the initial failure. Three distinct peaks can be observed at approximately 0, 300 and 850 seconds of 250 and 600 m^3s^{-1} respectively. It represents an instantaneous onset with a very short lived duration (c. 200 seconds). The peaks (b) to (d) correspond with the series of block failures schematically represented in Figure 4.7. The sudden failure and removal of embankment material and release of stored water through these block failure events produces the instantaneous discharges shown in this hydrograph.

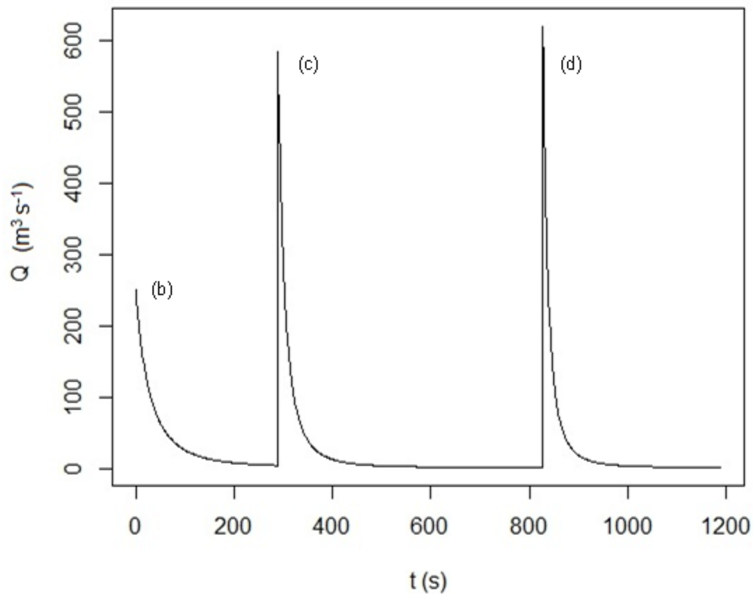


Figure 4.8 Ulley breach discharge hydrograph

4.3.3 **Flood spreading and inundation modelling (TELEMAC-2D)**

The extent of a potential inundation following a breach at Ulley was simulated using the open source TELEMAC-2D hydraulic model (www.opentelemac.org) (Figure 4.5B). The tool has been widely used for flood and breach applications (e.g. Malcherek 2000; Fernandes *et al.* 2001; Cooper *et al.* 2013). It models free surface flows based on nodes joined in an irregular triangular mesh. This allows refinement for specific areas of interest such as rapidly changing topography. TELEMAC-2D was employed for this case study because this mesh provides better description of topographic features that interfere with the inundation process (e.g. road embankments) (Di Baldassarre *et al.* 2009a) which are a key feature of the study area. TELEMAC-2D is a two-dimensional model that solves the 2D shallow water equations, enabling it to represent water movements between the channel and flood plain (Di Baldassarre *et al.* 2009b). The complex topography in this case results in flow prediction requirements from the embankment breach as well as constraining topographic features. For this reason TELEMAC-2D was appropriate to model the flood inundation from Ulley Reservoir.

The mesh is created using Blue Kenue which is a pre/post processing hydraulic tool developed by the Canadian Hydraulics Centre of the National Research Council Canada (NRCC). The tool is currently available under a free use licence (NRCC 2014). It has been used to prepare model data input for this study as it permits the integration of geospatial data with model input and results data for a number of hydrological models including TELEMAC-2D.

For this study a notional impermeable barrier bounded the selected study area. This is insignificant for the modelling as the flood spreading was naturally constrained by topography within the area of interest. Topographic information was contained within the mesh which was generated from the creation of nodes. These represent spot heights derived from a 2 m LiDAR digital surface model (DSM) (Environment Agency 2013). Two-metre resolution LiDAR is currently the finest available at the scale required for this flood analysis. Natural (e.g. Ulley's situation within the Rother Valley) and constructed (e.g. dam and motorway embankments) topographic features within the study area have a critical influence on flood water spreading (Figure 4.9). The 2 m resolution DSM is required to resolve features such as raised road and rail embankments for the purpose of the flood spread modelling in this example.

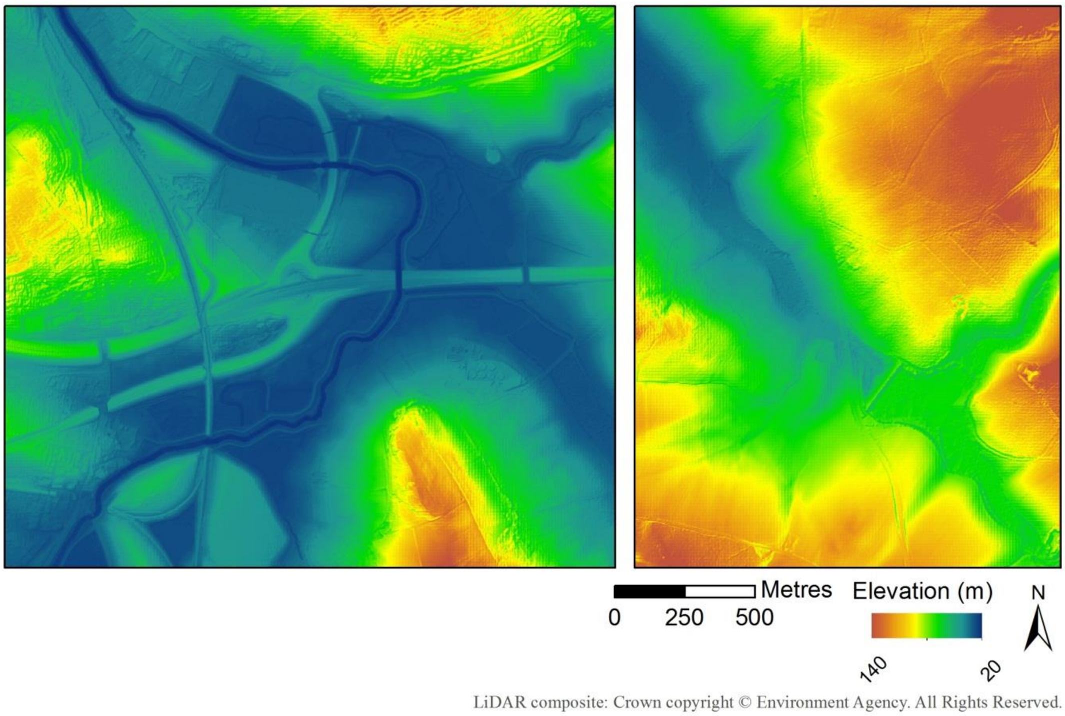


Figure 4.9 LiDAR digital surface model (resolution: 2 m, November 2012) showing M1 junction 33 (left) and the reservoir embankment (right).

The extensive pre-existing flood extent downstream has been accounted for as initial conditions within the models input parameters. The depth of standing flood water is mapped onto the input mesh for TELEMAC-2D (Figure 4.5B). If the dam had failed it would have been in addition to inundation on the flood plain immediately downstream. Depths for the existing downstream flood extent were estimated from aerial photographs (Figure 4.3) taken during the emergency response. The image was used to define boundary height forming the edges of the flood extent. This was achieved using Ordnance Survey spot heights on a 1:25k scale map. The height the antecedent flood inundation reached (31 mAOD) was extracted as an isoline from the topography mesh using Blue Kenue. The difference between the isoline and the topography was calculated to estimate the flood depth across the flood plain.

The EMBREA breach hydrograph was used for the reservoir discharge parameter within TELEMAC-2D (Figure 4.5A). The discharge source location is assigned to a node’s coordinates. A location in the centre of the reservoir embankment was chosen. Culverts through notable barriers downstream such as the railway and motorway embankments are accounted for by breaks in the DSM (Figure 4.9). While the capability to account for friction variation based on land use classification remains it was deemed insignificant for this study. This was attributed to most of the existing and output flood extent being within similar land use zones. The effect of friction was still accounted for by applying a representative default value across all areas. The modelled outputs concerning the spreading of a breach event for water depth and velocity were exported from the model's mesh as xyz data. The data were subsequently rasterised to a 15 m resolution for analysis using ArcGIS. This resolution is the finest achievable based on the density of nodes within the model's mesh, which is in turn derived from variation within the topography.

4.3.4 Spatiotemporal population modelling

The Population 24/7 tool that has been described in detail in Chapter 3 has been used to produce spatiotemporal estimates for the Ulley study area (Figure 4.1). The population estimates (Figure 4.5C) are analysed with the hydrological model outputs using ArcGIS for an assessment of the flood risk to people. An 8 x 10 km grid of 100 m cells (n cells = 8000) of population density for hourly time intervals during a 'typical' term-time weekday (reference date: June 2007) has been produced. This permits the population potentially exposed to the flood risk of an Ulley dam failure to be analysed within a diurnal cycle. The reference date for the data used to construct the model has been selected to be representative of the conditions at the time of the Ulley 2007 incident.

Table 4.3 outlines the sources used to construct the Ulley data library for the Population 24/7 model. The model for Ulley only considers a term-time population base-line as the dam failure occurred on a June weekday. The workplace and education locations facilitate the known movement of people to their respective activities following a predictable daily trend.

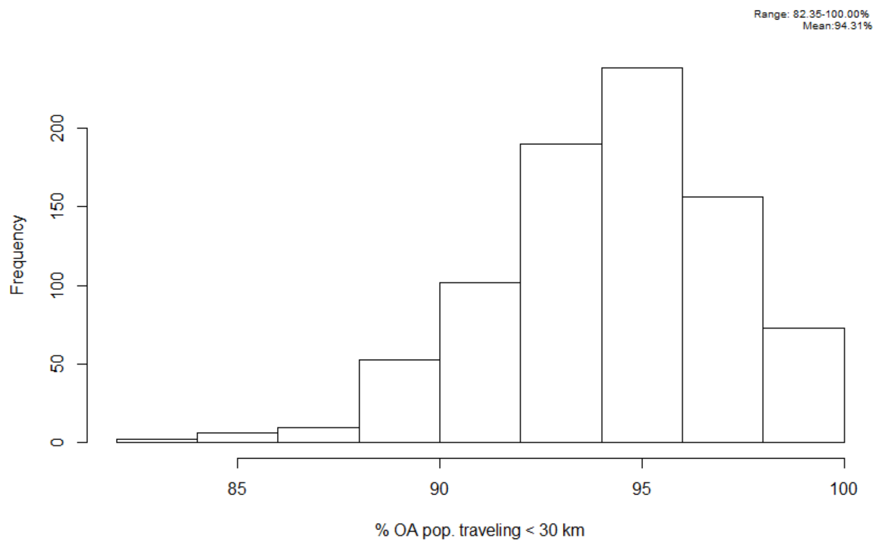


Figure 4.10 Histogram of percentage of Rotherham OA population travelling less than 30 km to work

For the purposes of the population modelling component a 30 km buffer zone has been applied to the study area to mitigate against edge effects caused by peripheral population movements. This was derived through analysing travel to work data from the 2001 census (Table UV80). This was the highest resolution and most comprehensive dataset available at the time of writing. 94.3% of the working population within OAs in the Rotherham district travelled < 30 km to work, while 91.9% travelled < 20 km (Figure 4.10). These data concern the usual resident population aged 16-64 in employment (excluding those with no fixed place of work, working offshore or outside of the UK).

Table 4.3 Ulley Population 24/7 data library composition and sources

Name	Description	Type	Data source(s)			
			Population	Location geography	Temporal profile	Wide Area Dispersion
Ulley usually resident population	Baseline 'term time' population	Origin	2007 MYE	NSPD (May 2007) *residential UPCs*	N/A	N/A
Education	Count of pupils and students in full time education (school, college and university)	Destination	Schools Census (DfE 2007) Independent Schools Census (2007) HESA (2007)	NSPD (May 2007) *Georeferenced UPC for institutions site(s)*	Population 24/7 project	Population 24/7 project
Education workforce	Workplace population counts (SIC: P)	Destination	ABI (2007)	NSPD (2007) *Business UPCs*	LFS (2007)	Census (2001, Table UV80)
Health service workforce	Workplace population counts (SIC: Q)	Destination	ABI (2007)	NSPD (2007) *Business UPCs*	LFS (2007)	Census (2001, Table UV80)
Manufacturing and construction workforce	Workplace population counts (SIC: C, F)	Destination	ABI (2007)	NSPD (2007) *Business UPCs*	LFS (2007)	Census (2001, Table UV80)
Wholesale and retail workforce	Workplace population counts (SIC: G)	Destination	ABI (2007)	NSPD (2007) *Business UPCs*	LFS (2007)	Census (2001, Table UV80)

Table 4.3 continued

Name	Description	Type	Data source(s) Population	Location geography	Temporal profile	Wide Area Dispersion
Office based workforce	Workplace population counts (SIC: J, K, L, M, N, O, R, S, T, U)	Destination	ABI (2007)	NSPD (2007) *Business UPCs*	LFS (2007)	Census (2001, Table UV80)
Transportation workforce	Workplace population counts (SIC: H)	Destination	ABI (2007)	NSPD (2007) *Business UPCs*	LFS (2007)	Census (2001, Table UV80)
Catering workforce	Workplace population counts (SIC: I)	Destination	ABI (2007)	NSPD (2007) *Business UPCs*	LFS (2007)	Census (2001, Table UV80)

Notes: SIC (Standard Industrial Classification, Annual Business Inquiry (ABI) 2007)

4.3.5 **Flood hazard rating**

A flood hazard rating methodology has been developed by Penning-Rowsell *et al.* (2005) to assess the risk of death or serious injury to people exposed to flooding. The flood hazard rating aims to quantify the flood hazard in terms of the flood characteristics (depth and velocity) (DEFRA 2006). This is one method to compare one flood event to another against the same criteria. While the methodology outlined aims to quantify population exposure to a given hazard by time of day, the use of a flood hazard rating considers the magnitude of the hazard. Like population the magnitude of a hazard also changes throughout space and time. This method of quantification in conjunction with detailed hazard maps, such as those outlined in this chapter, permits the severity of the flood hazard to be analysed in space and time. This rating has been adopted for this study as a tool that can be applied nationally within the UK (e.g. DEFRA 2006). The hazard rating can be combined with assessments of the area vulnerability (Eq. 4.2) and the people vulnerability (Eq. 4.3) to estimate the number of injuries and fatalities (Eqs. 4.4 and 4.5).

The flood hazard rating (HR) is formulated:

$$HR = d(v + 0.5) + DF \tag{4.1}$$

Where, d = depth, v = velocity and DF = debris factor (0, 0.5 or 1). The three stage debris factor score depends on the probability that debris will lead to a significantly greater hazard. The debris factor has been set to 0.5 for this example. The area immediately downstream of the dam is arable pasture with few sources of debris. The flood modelling has identified a number of crucial culverts. Therefore, although the sources of debris may be low, the effect of blocked culverts would be high. The hazard rating is assessed according to the criteria in Table 4.4.

Table 4.4 Flood hazard rating and risk to people after Priest *et al.* (2007)

Hazard rating	Flood hazard	Description
< 0.75	Low	Flood zone with shallow water or deep standing water.
0.75-1.25	Moderate	Flood zone with deep water or high velocities. Danger for some.
1.25-2.00	Significant	Flood zone with deep fast flowing water. Danger for most.
> 2	Extreme	Flood zone with deep fast flowing water. Extreme danger for all.

Area Vulnerability (AV) considers the amount of flood warning, flood onset speed and the nature of the area. Each parameter for the AV is given a score based on set criteria (Table 4.5) which is then summed (Eq. 4.2). This gives a range of 3 (least vulnerable) to 9 (most vulnerable) depending on the area's characteristics.

$$AV = Speed\ of\ onset + Nature\ of\ area + Flood\ warning \qquad (4.2)$$

Table 4.5 Area vulnerability score matrix after Penning-Rowsell *et al.* (2005)

Parameter	Score		
	1 (low risk)	**2 (medium risk)**	**3 (high risk)**
Speed of onset	Slow (many hours)	Gradual (≈ hour)	Rapid
Nature of area	Multi-storey apartments	'Typical' residential	Mobile homes, schools, busy roads, bungalows etc
Flood warning	Effective with emergency plans	Warning system present, but limited	No warning system

Overall area vulnerability for the Ulley study area has been evaluated using Equation 4.2. This produced a score of 9 (maximum), indicating a very high area vulnerability. The flood warning and speed of onset parameters are given the highest score (3). The onset of breach events is often instantaneous (rapid) leaving little time to warn even if a general flood warning system is in place. The nature of the area includes major transportation infrastructure as well as a mix of typical residential dwellings in addition to schools and commercial units (score = 3).

People Vulnerability (PV) (Eq. 4.3) assesses the population characteristics and their propensity to experience harm. The PV is given a percentage score of 10, 25 or 50% based on criteria relating to the populations age and health (Table 4.6).

$$PV = \%_{\text{residents}}\ longterm\ illness + \%_{\text{residents}}\ aged > 75 \tag{4.3}$$

Table 4.6 People vulnerability score matrix after Penning-Rowsell *et al.* (2005)

Parameter	**Score (%)**		
	10 (low risk)	**25 (medium risk)**	**50 (high risk)**
% Pop. Aged > 75 years	Above national average	Around national average	Above national average
% Pop. long-term sick/disabled	Below national average	Around national average	Above national average

This information has been referenced from the 2001 census because 2007 is an inter-censual year and this level of detail is only available for 2001. The population aged over 75 years was obtained from Table KS002: Age Structure. The percentage of the population living with a limiting long-term illness was obtained from Table KS008: Health and Provision of Unpaid Care. The values for the Ulley study area are summarised in Table 4.7. These have been determined from the required age and long term illness characteristics reported for parishes and local authorities within the study area that are at risk from the dam failure inundation. They represent the population across the whole reporting zone as published without further interpolation.

Table 4.7 Population characteristics by parish/local authority (Census 2001, Tables KS002 and KS008) for locations within the study area compared to national average (England).

Location	Pop. aged over 75 (%)	Difference from national average†	Pop. Long-term illness (%)	Difference from national average‡
Rotherham	7.09	-0.54	22.41	+4.48
Catcllife	6.22	-1.32	24.63	+6.70
Treeton	6.85	-0.69	22.00	+4.07
Study area mean	6.72	-0.82	23.01	+5.08
England (Average)	7.54†		17.93‡	

Based on the data in Table 4.6 a population vulnerability of 75% (25% + 50%) for the study area has been calculated. The proportion of population aged over 75 is around national average (score = 25%, Table 4.6). However, the population with a limiting long-term illness is much greater than average (score = 50%, Table 4.6).

Finally, the preceding components are combined to give an estimate on the number of injuries and fatalities for a given flood event.

$$N(I) = 2N_z \frac{HR \times AV}{100} \times PV \tag{4.4}$$

Where, N(I) = number of injuries, N_z = population living in the flood plain, HR = hazard rating, AV = area vulnerability and PV = people vulnerability

$$Fatalities = 2N(I)\frac{HR}{100} \tag{4.5}$$

The AV and PV have been determined based on the area as a whole. However the population exposed, flood depth and velocity varied by cell and is the output of the modelling described within this chapter. The number of injuries, and therefore fatalities can be calculated on a cellular level. The problem of spatial units still exists by means of differing resolutions between these gridded flood inundation (15 m) and population (100 m) outputs. This is much easier to overcome in a grid than with varying arbitrary spatial zones. The 100 m population dataset has been resampled and adjusted to 15 m resolution using a scale factor of 44 $(100 \div 15)^2$. This is provided with the caveat that the actual variation of population within the 100 m output cell cannot be determined based on the resolution of currently available input data. Therefore an equal distribution has been assumed. It is important to note that this thesis does not intend or attempt to claim potential injuries or fatalities to the nearest 15 m. Instead this method allows the assessment of the overall flood event while considering local variations in flow parameters and population. Furthermore this is approach is repeated for population data for different times of the day to examine any spatiotemporal variability.

A hypothetical worked example is provided in Table 4.8. The same process is repeated for all cells (n = 9613) within the flood polygon.

Table 4.8 A hypothetical worked example for a single 100 m cell for time *t*.

Property	Value	Notes
Flood depth	0.5 m	Cellular value determined from TELEMAC-2D depth output
Flood velocity	2 ms^{-1}	Cellular value determined from TELEMAC-2D velocity output
Debris factor	0	Unlikely (range: 0-1), value applied to all cells
Population exposed	35	Cellular value determined from Population 24/7 for time t
Area vulnerability	9	3 + 3 + 3 (Table 4.5, Eq. 4.2)
People vulnerability	75%	25 + 50 (Table 4.6, Eq. 4.3)
Hazard rating	1.25	Eq. 4.1
Number of injuries	6	Eq. 4.4 (rounded up to 0 dp)
Number of fatalities	0.13	Eq. 4.5

4.4 Results

Results following the integration the spatiotemporal population and environmental modelling outputs outlined in the previous section are presented here. They will be discussed in the following order: firstly, the outputs from the hydraulic modelling undertaken. This presents estimates for the flood extent, depth and velocities derived from the combination of the flood breach and spread modelling (Figure 4.5A and B). Secondly, a flood hazard rating has been calculated based on the flood depth and velocity variables according to Equation 4.1. Velocity and water level (height AOD) time series have additionally been extracted from the TELEMAC-2D inundation outputs at the motorway embankment. Thirdly, the population distribution has been described and spatiotemporal variation in flood risk exposure illustrated. The spatiotemporal population estimates have been compared to static census counts. Finally, a fatality estimate has been calculated and spatially represented.

4.4.1 Hydraulic modelling (TELEMAC-2D)

Water depth and velocity results derived from TELEMAC-2D for the post-breach inundation extent for Ulley Reservoir are shown in Figure 4.11. These represent the flood level 45 minutes after the start of the breach. The maximum flood depth and velocity do not necessarily occur at the same time however this extent is a close approximation. The output extent recognises the antecedent flood conditions. The greatest depths occur in river channels, while increased velocity occurs from the initial breach and through culverts (see also Figure 4.12). The highest flows of 5.5 m s^{-1} occur immediately downstream of the embankment. The channelization effect of the culverts creates localised intensification in velocity. The maximum depth (up to 6 m) outside of the river channels is to the south of Catcliffe.

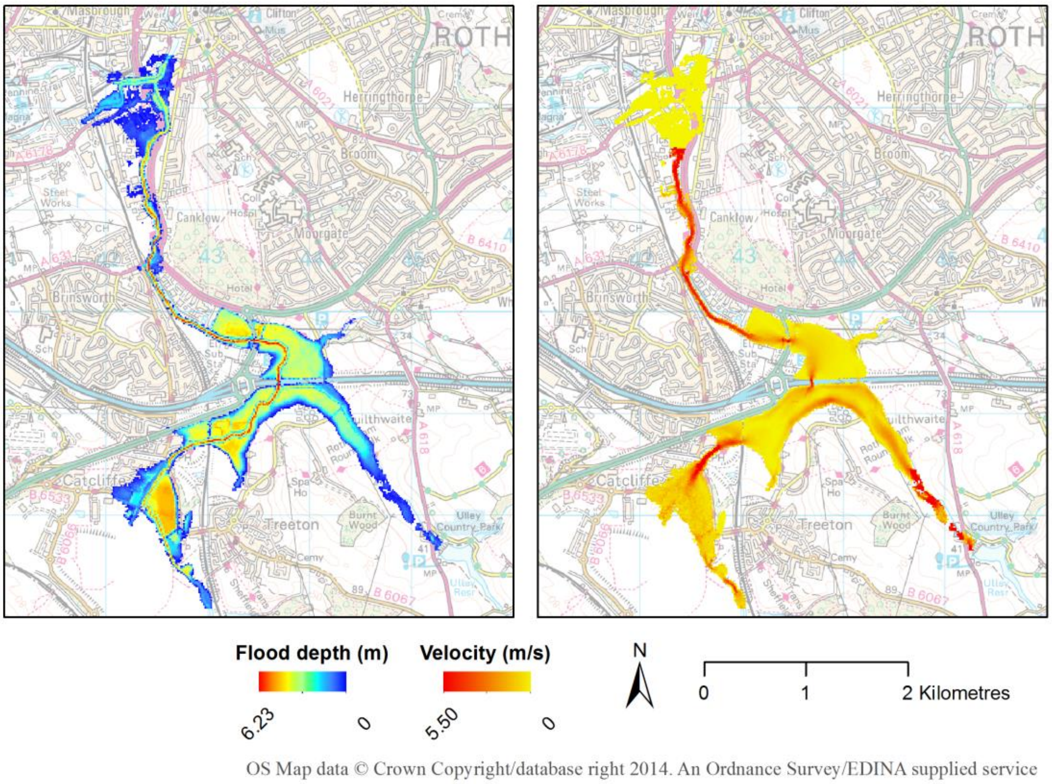
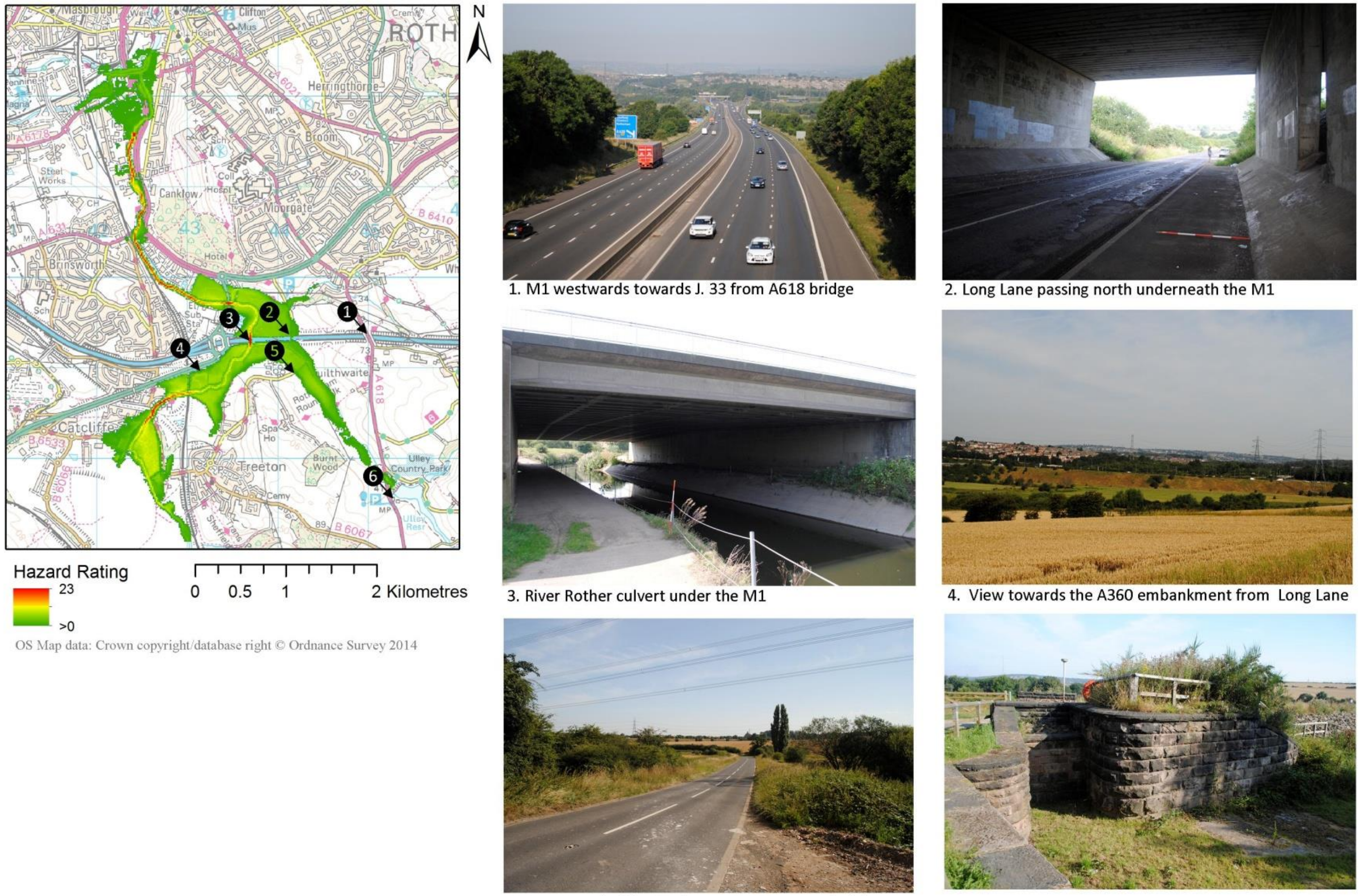


Figure 4.11 Flood inundation results for water depth (left) and velocity (right)

4.4.2 Flood hazard rating evaluation

The flood hazard rating (Figure 4.12) identifies the locations of greatest depth and velocity. The HR for each cell has been calculated (Eq. 4.1) using the flood modelling outputs (Figure 4.11). The highest rating occurs within the original channels. The majority of the centre of the southern half of the flood polygon has a hazard rating >5. This is well into the extreme category and would pose extreme danger to all exposed (Table 4.4). While we may not expect people to be present in the river channels where the hazard is rated the greatest (HR > 20), people do still enter or get into difficulties in the high flow of such channels which ultimately results in fatalities. Multiple fatalities within this region in June 2007 illustrate the extreme risk to people. In Humberside a man died in flood water after getting his foot stuck in a manhole grate as firefighters tried to free him. In Sheffield a 68 year old man and 14 year old boy were separately swept into swollen river channels resulting in two further fatalities (Williams and Glendinning 2007). Contained within the same report, 1000 people could not return home or were evacuated in Sheffield alone by the emergency services and Royal Air Force.



1. M1 westwards towards J. 33 from A618 bridge

2. Long Lane passing north underneath the M1

3. River Rother culvert under the M1

4. View towards the A360 embankment from Long Lane

5. Long Lane looking NE from c. OS 31 m spot height

6. Pre 2007 (1943) storm overflow at Ulley reservoir

Figure 4.12 Flood hazard rating for the Ulley breach scenario (left). Ground observation of features of the flood model output (right) (Survey pole: 1.5 m). Photographs: Alan Smith (August 2014)

Photographs relating to the flood inundation characteristics observed in the TELEMAC-2D model results were taken on a visit to the site in August 2014 (Figure 4.12):

Photograph 1: View west along the M1. This is a major national arterial route that transects the centre of the flood inundation polygon. This is the centre of the section closed because of fears that a breach at Ulley could overtop the carriageway here.

Photograph 2: The bridge where Long Lane passes beneath the M1 motorway is one of two breaks in the motorway embankment that acted as a culvert during the antecedent flooding immediately downstream from the reservoir in 2007.

Photograph 3: The second culvert where the River Rother flows beneath the M1 and junction 33 entrance/exit slip roads.

Photograph 4: A view of the A360 road embankment, similar to raised section of M1 motorway. The study area is characterised by these raised earth embankments that have a notable impact on the flood inundation modelling (Figure 4.11)

Photograph 5: View of Long Lane at 31 mAOD. This section of road was inundated by at least 3 m of standing flood water due to the flooding already experienced in the area (Figure 4.8) during June 2007. It is also immediately downstream of the reservoir embankment.

Photograph 6: Channel entrance to the original 1874 spillway (now sealed and the spillway removed) where flow initiated the destabilisation of the embankment at Ulley in June 2007.

A velocity time-series was taken at the motorway embankment immediately downstream from the reservoir (Figure 4.13) with a profile closely aligned to the initial hydrograph. It commences from the initial embankment failure with three peaks in velocity at approximately 2500, 7500 and 17500 seconds. This was extracted from the closest corresponding node within the model's mesh using Blue Kenue. It approximately corresponds to the left of the Long Lane motorway underpass (Figure 4.12(2)). The three peaks of around 2 m s^{-1} are a response to block failures in the reservoir embankment. Although flood flows have slowed at this point velocities around 2 ms^{-1} are still significant. This is within the velocity threshold for masonry and concrete great enough to cause structural damage (Priest *et al.* 2007). Therefore it is possible that the integrity of the motorway structure could be compromised, particularly at the locations of culverts and bridges where localised peaks in velocity occur.

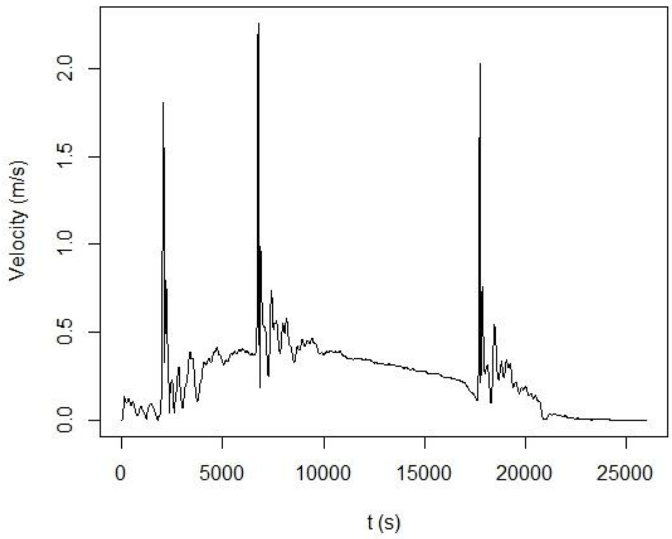


Figure 4.13 Velocity profile at motorway embankment

A time-series for flood water depth normalised to height above ordnance datum (AOD) was also extracted from the model results using Blue Kenue (Figure 4.14). It was taken from the lowest position of the motorway embankment at the River Rother culvert (Figure 4.11(3)). The comparison of the potential water depths adjacent to the motorway and the lowest elevation of the carriageway surface (31.7 mAOD) according to the LiDAR data suggest that the motorway embankment may not have been overtopped. However this is based on the assumption that the culverts (or road underpasses acting as temporary culverts) are unobstructed. Nevertheless, the level of the maximum water depth for the scenario modelled indicates that the water level could have come within 0.70 m of overtopping the carriageway as during the first two peaks. It is estimated to have reached at least the 31 m contour (Figure 4.14).

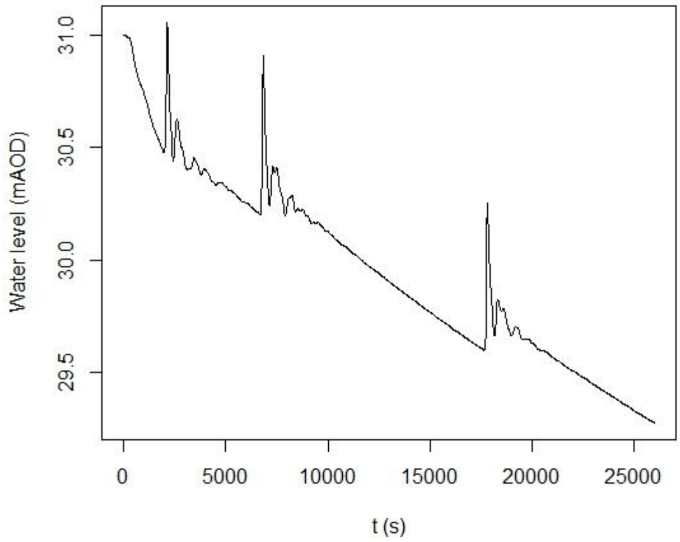


Figure 4.14 Flood level at the motorway embankment

4.4.3 **Spatiotemporal population outputs**

Outputs from Population 24/7 show variations in population density at 100 m resolution in both space and time. Features of a 'night-time' modelled population layer for the study area are identified in Figure 4.15. Low and high density variations in residential living spaces can be detected. Figure 4.15 also illustrates zero densities which are exemplified by an area of housing adjacent to a large recreation area. The features are present because of the use of residential postcode centroids (Chapter 3) which correctly allocate population to their respective residential locations. A mix of land uses, including industrial is also evident.

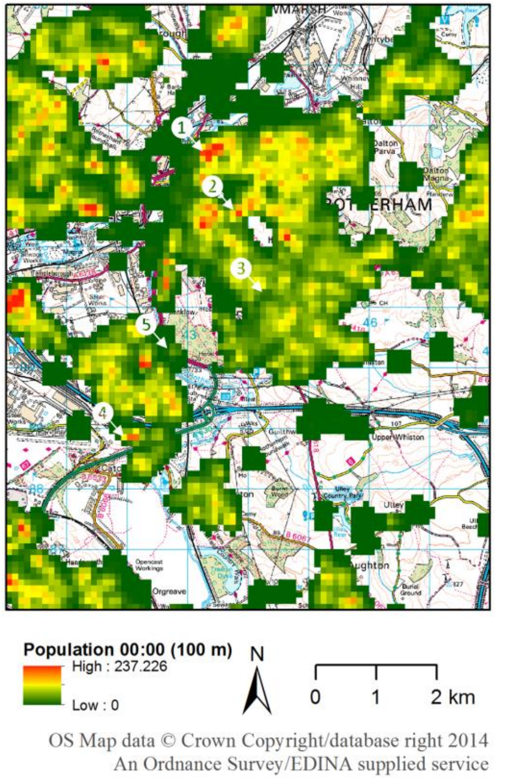
1. High density residential (Hatherly Rd)

2. Medium/High density adjacent to playing field

3. Outer suburb low density residential (Lymister Ave)

4. Suburban new build residential (Oliver's Way)

5. Canklow industrial estate



Ulley study area 'night-time' (00:00) population (100 m)

Figure 4.15 Corroborating modelled population density output using Population 24/7 with ground observations. Photographs: Alan Smith (August 2014)

The population potentially exposed within the flood hazard polygon (Figure 4.12) has been calculated in ArcGIS at hourly intervals through the day using both the original census data and spatiotemporal estimates (Figure 4.16). The two census datasets (LSOA and OA) give a single static exposure estimate. For the highest resolution census data at at OA level for 2001 the daytime population count (Table UV037) was used. This is defined as the population aged 16-74 resident in the area who do not work, plus all people who are working in the area (ONS 2004). The outputs presented from Population 24/7 show high population densities (up to 700 people per ha) emerging on highly concentrated sites by mid-day (e.g. Figure 4.17E). The highest of these densities are secondary school sites with up to 1500 pupils. By the end of the school day (16:00) fewer sites are occupied (Figure 4.17F). These can be explained by the relatively large number of school staff still present at the same sites and the workforce located at other places of work such as high-density office workspaces. The population ‘in travel’ is also distinguishable on the road network and evident around peak travelling times (e.g. Figure 4.17D and E).

Figure 4.16 shows that the magnitudes to which the estimates differ is so great that there is an element of meaningless. For example it would not be logistically nor financially viable to always plan for 250% of the resources that ‘may’ actually be required if following the precautionary approach.

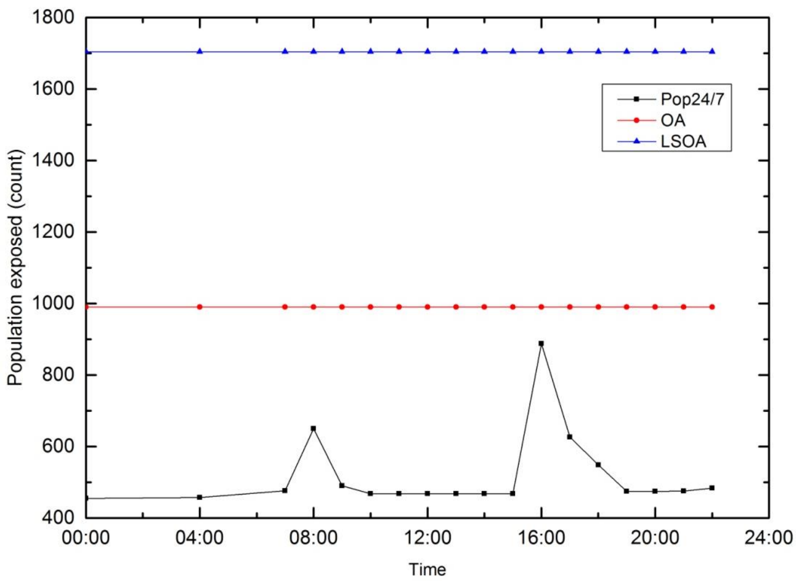


Figure 4.16 Population exposure estimate within the flood extent polygon

The Population 24/7 exposure estimate shows a temporal trend within this study area. The estimated population exposed to a breach of Ulley reservoir increases between 400-600 people during the peak travelling times. The population density during typical working hours (10:00-15:00) remains constant and slightly below the night-time residential level (21:00-04:00). This would be as expected for the nature of the study area. The flood polygon encompasses mainly residential areas and major arterial routes. Therefore, there is a large peak in the exposed population during the typical commuter times and population decreases in between (as people leave the residential areas to work elsewhere during the day).

Figure 4.16 illustrated the spatial variation in population density for a 'night-time' usually resident population. The population within the study area has been modelled at hourly time-slices throughout a weekday representative of 25 June 2007. A selection of outputs for 00:00, 08:00, 12:00, 16:00 and 20:00 are illustrated in Figure 4.17. This also contains a comparison with census datasets (Figure 4.17A and B) overlaid by the predicted modelled flood extent.

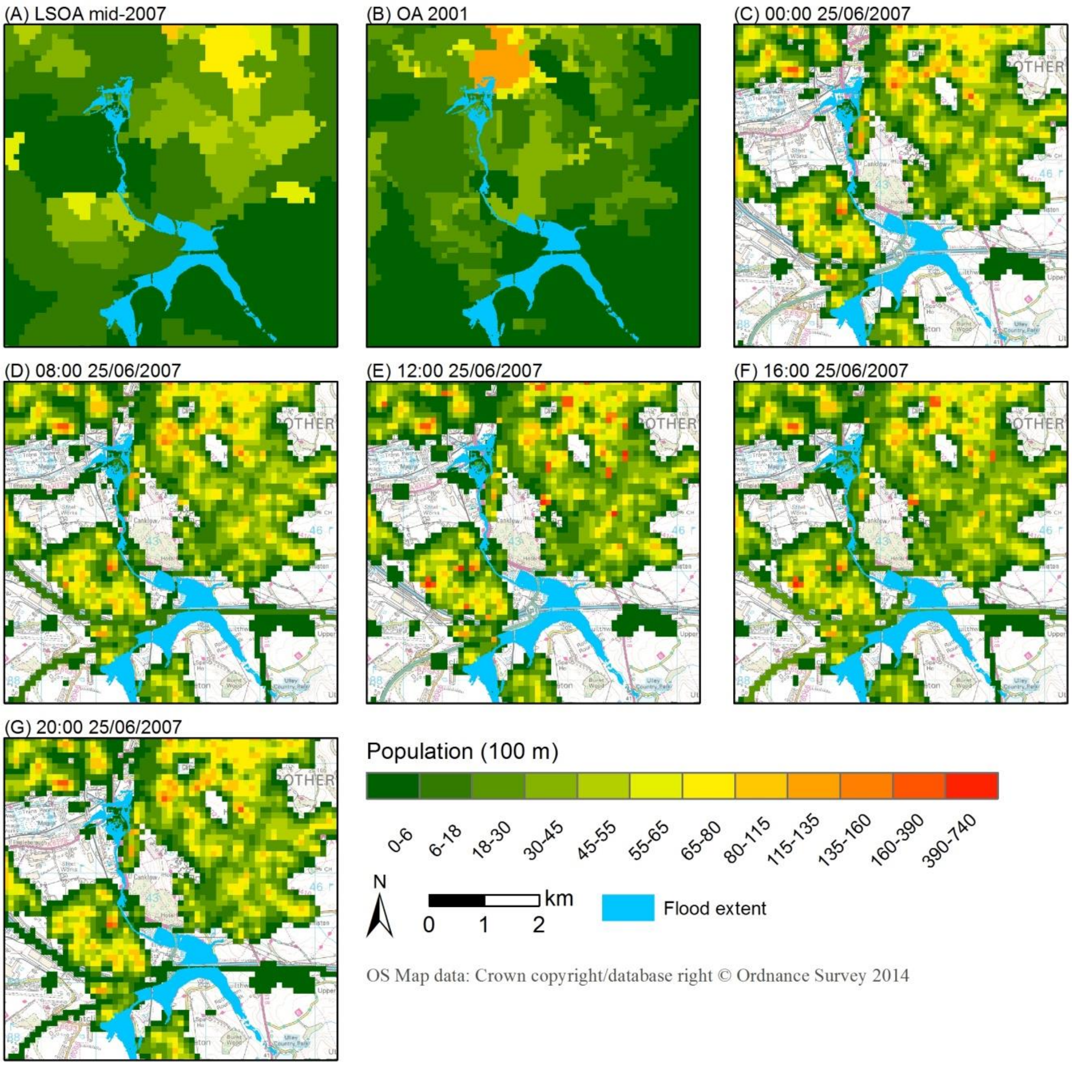


Figure 4.17 Gridded Population 24/7 (100 m) spatiotemporal population estimates for the Ulley region compared to census datasets overlaid with TELEMAC-2D flood inundation polygon.

Figure 4.17A shows rasterised LSOAs with a 2007 MYE. This is the best available standard published population data for the target date (2007). The 2001 census OA population count is also illustrated (Figure 4.17B). While this only represents the population at the year of the 2001 census it is the highest resolution population data available prior to the target date without the need for further modelling (e.g. that undertaken using Population 24/7). In stark contrast to the spatiotemporal estimates (Figure 4.17C-G) they only provide a static count of uniform population density across large areas. The census representation of population is inadequate for the example described here. It is intended to act as an illustrated example for choosing the appropriate dataset or method suited to the task for which it is required.

However, in current flood risk assessments it is this (Figure 4.17A-B) kind of static population data that are often utilised (e.g. Arrighi *et al.* 2013; Taylor *et al.* 2013). The approach outlined here aims to demonstrate improvements to this practice when considering spatiotemporal variation in population density. It is known that populations are not static in time and this is shown in the Population 24/7 result. The census over-estimation results from its geographic coverage as a set of contiguous zones with uniform densities. Therefore when the flood polygon is used to extract the underlying population potentially exposed all cells contain a count. However, in reality much of the flood polygon (arable pasture) is uninhabited. If adopting the precautionary principle approach an over-estimate may seem desirable.

4.4.4 **Population fatality estimates**

The number of potential fatalities has been estimated from the Population 24/7 results using the approach outlined in Section 4.3.5. These were calculated for each cell containing a population count that is contained within the breach flood risk estimate. Five sample hourly intervals have been chosen for a 'typical' weekday with a target date of June 2007. A break-down is provided in Table 4.9.

Table 4.9 Ulley fatality estimates

Time of day	Fatalities
00:00	6
08:00	8
12:00	5
16:00	12
20:00	6

This distribution of fatalities for 12:00 and 16:00 has been mapped in Figure 4.18. In line with the methodology these generally correspond with where the hazard rating is greatest. This occurs in the river channel where depths and velocities are the highest. Further refinement is likely to be required due to the nature of the area in this case study. The two main areas with the greatest fatalities predicted are located where the river channels are bounded by inhabited areas. It is not unreasonable to assume deaths can occur within the channels, as has already been reported for this area. These channels (sub 100 m) cannot be resolved in the population modelling at the resolution achievable with currently available data. It may be likely that the application of a mask post-modelling may be required to constrain population from the channel locations. However, this raises questions for population volume preservation.

These results show that population exposure and fatalities are sensitive to the time of day that the hazard occurs. The detailed spatiotemporal modelling approach illustrated has made it possible to identified high-risk areas with the potential to cause population fatalities. This could not be achieved using census data alone. It confirms that this process is sensitive to population and hazard fluctuations. This is a reflection on reality where dynamic populations inevitably intersect hazardous zones.

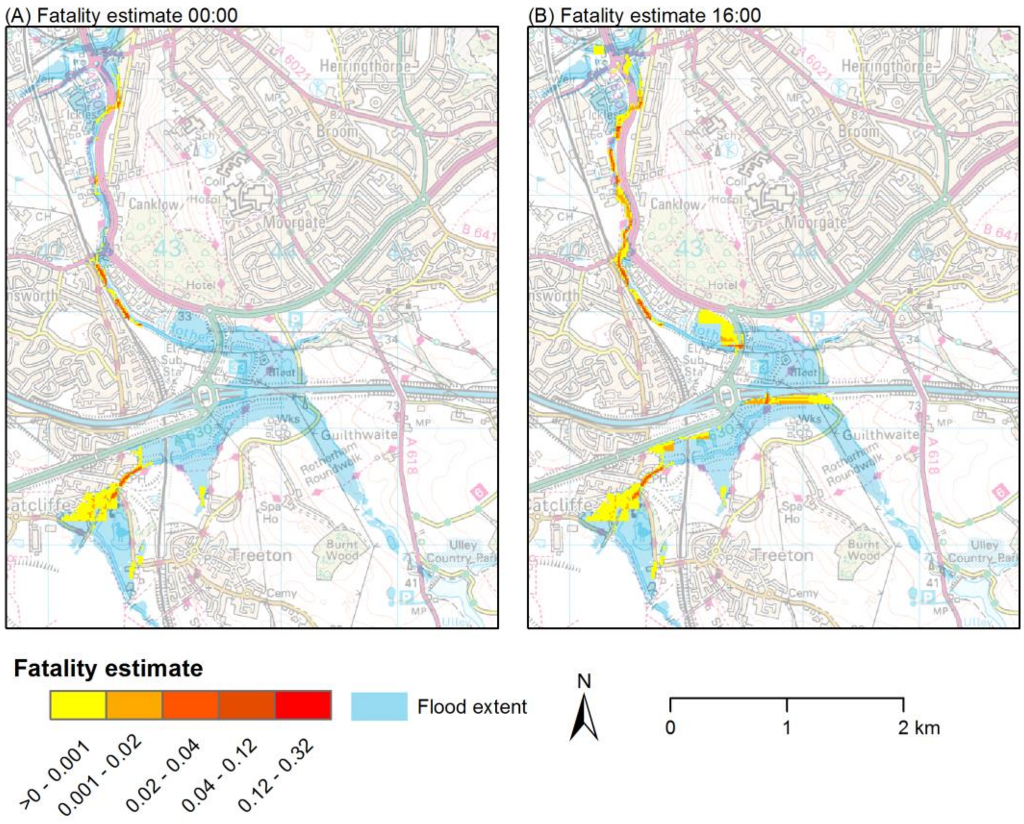


Figure 4.18 Distribution of potential fatalities at midnight and 16:00 for a typical weekday (June 2007)

4.5 Comparison with census estimates

The differences between the Population 24/7 estimates and census datasets have been examined. These should be treated with caution as they are not like-for-like comparisons, but rather the closest possible alternative datasets (e.g. highest resolution census daytime estimate only available for 2001, and 2007 estimate only available at LSOA level). The comparison with census OA level represents population counts for 2001 (census year). These are the highest resolution population data available for the study area, but six years before the target date. For this reason MYEs are produced for inter-censual years. However, these are only available at the highest resolution for LSOA level. While this provides the population estimate required for the target date (2007) the spatial resolution is diminished.

The difference (Δ) for all raster 100 m cells (n= 8000) within the study area of selected Population 24/7 outputs (12:00, 16:00 daytime and 00:00 night-time) between rasterised census 2001 OA daytime and 2007 LSOA night-time counts has been evaluated in Figure 4.19 and summarised in Table 4.10. The standard deviation (σ) provides a comparative measure for the spread of values. This is not a test where any one dataset should resemble the other as they are both different measures although the closest available comparison.

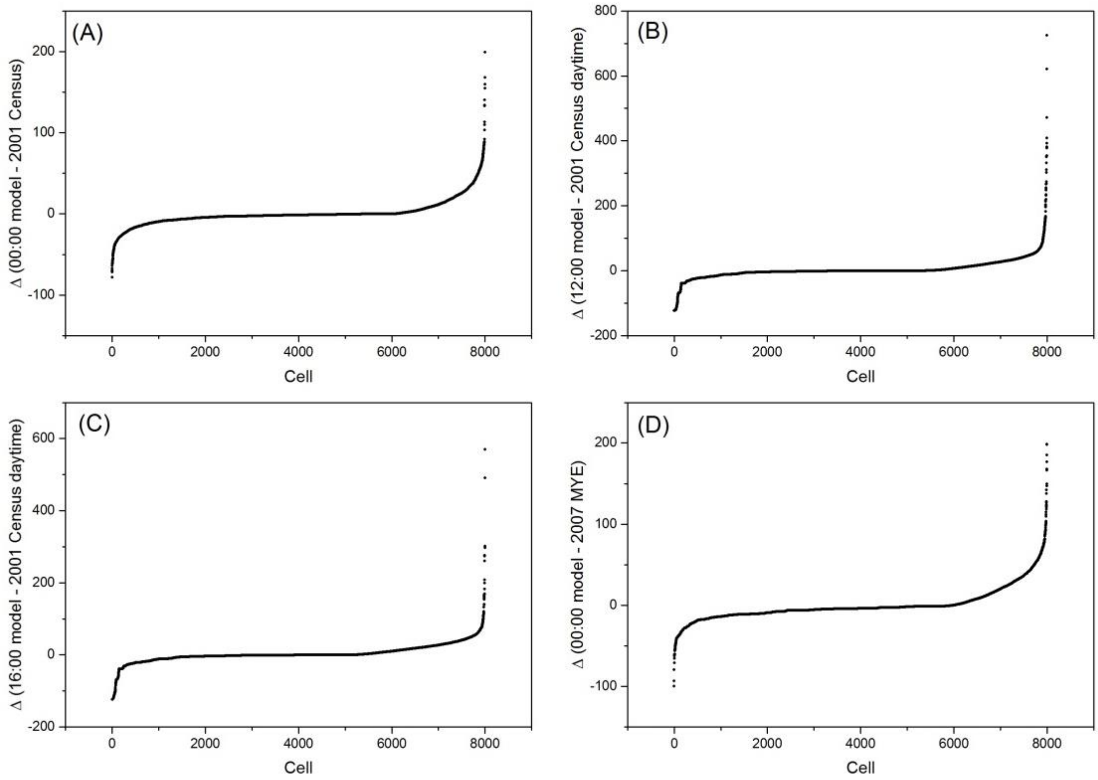


Figure 4.19 Difference between 00:00 Population 24/7 estimate and 2007 (LSOA) night-time census population count. (B) Difference between 12:00 Population 24/7 estimate and 2001 (OA) daytime census population count.

In terms of night-time (00:00) population comparisons there is a greater variance in the values for the Δ between night-time Population 24/7 2007 estimate and the night-time LSOA 2007 count (Figure 4.19A and Table 4.10, σ = 20.36) although these are closest in terms of population reference date (i.e. both 2007). The least amount of variance is between Δ of the Population 24/7 daytime population estimate and 2001 census daytime count (Figure 4.19B and Table 4.10, σ = 15.81). Although the difference between the two is smaller they are still datasets with reference dates six years apart. This measure suggests that greater resolution in the spatial distribution of population is more important than target date in this case. The difference for all of the model and census comparisons for the times selected is positive. This shows that these cells in the Population 24/7 layers are assigned higher values than their corresponding counterpart in census layers. This would be expected because the Population 24/7 approach concentrates population to locations expected to be occupied (e.g. Figure 4.17) rather than the uniform across zone approach represent in the census comparison.

Table 4.10 Ulley study area model difference analysis

Values		Graph	Measure		
Observed (model)	**Predicted (census)**	**Figure 4.19**	**SD**	**Mean**	**RMSE**
Ulley 00:00	2001 Census	A	15.81	0.25	15.81
Ulley 12:00	2001 Daytime	B	32.52	3.51	32.71
Ulley 16:00	2001 Daytime	C	26.74	0.09	26.95
Ulley 00:00	2007 MYE	D	20.36	0.09	20.36

The largest difference between the model estimates and census counts occurs in the daytime model comparisons (12:00 and 16:00). The greatest variation in values is observed for the Δ between the census and the 12:00 layer (σ = 32.52). This is reduced when comparing the 2001 census daytime population count with the 16:00 layer (σ = 26.74). The 12:00 layer also has much greater mean difference (+3.51) compared to the rest of the layers evaluated in Table 4.10. The difference can also be observed in Figure 4.19B where this layer has the greatest spread and magnitude of cells with a positive difference from the daytime census estimate. The concentration of very high populations on relatively small sites such as schools is the driver of this trend. The difference has reduced by 16:00 when the pupils have left these sites however high concentrations of staff still remain. These differences occur for the same reason discussed in relation to Figure 4.17. The Population 24/7 modelling technique concentrates population to workplace locations using georeferenced business or actual workplace postcodes. This results in the spatial concentration of population density at these sites rather than a uniform across zone distribution. The root mean square error (RMSE) (Table 4.10) calculated for these differences closely corresponds to the values and pattern observed in the standard deviation of the difference.

The spatial distribution of the actual cellular difference for the modelled 00:00 (Figure 4.19A) and 12:00 (Figure 4.19B) results and census estimates is illustrated in Figure 4.20. The green colour represents a negative difference (i.e. the model appears to underestimate census counts) and the yellow-red positive differences. The lightest green colour shows where the census represents a small non-zero uniform density estimate but the modelled outputs represent this as zero in unoccupied locations. This treatment of unoccupied cells accounts for the green background colouring illustrated. The largest differences between 350 to 725 people per 100 m^2 (Figure 4.20B) occur within the daytime estimates. These typically represent concentrated populations at schools.

The spatial distribution for the selected daytime difference (Figure 4.20B) also shows some large decreases within population density of up to 125 people per 100 m^2 represented by the dark green colour. This primarily occurs within Rotherham town centre. This can be explained by the larger increase in the census daytime estimate, caused by representing people at their nominated place of work, increasing the density within these town centre OAs. Therefore the difference from the model is greatest outside of specific cells where employees have been concentrated at workplace locations rather than spread across the whole OA. These phenomena highlighted demonstrated the importance of accurate population distribution for realistic representations. The degree to which the census differs varies by time with the greatest changes occurring during daytime hours. This daytime difference is caused by the inadequate representation of daytime population clusters within traditional census outputs.

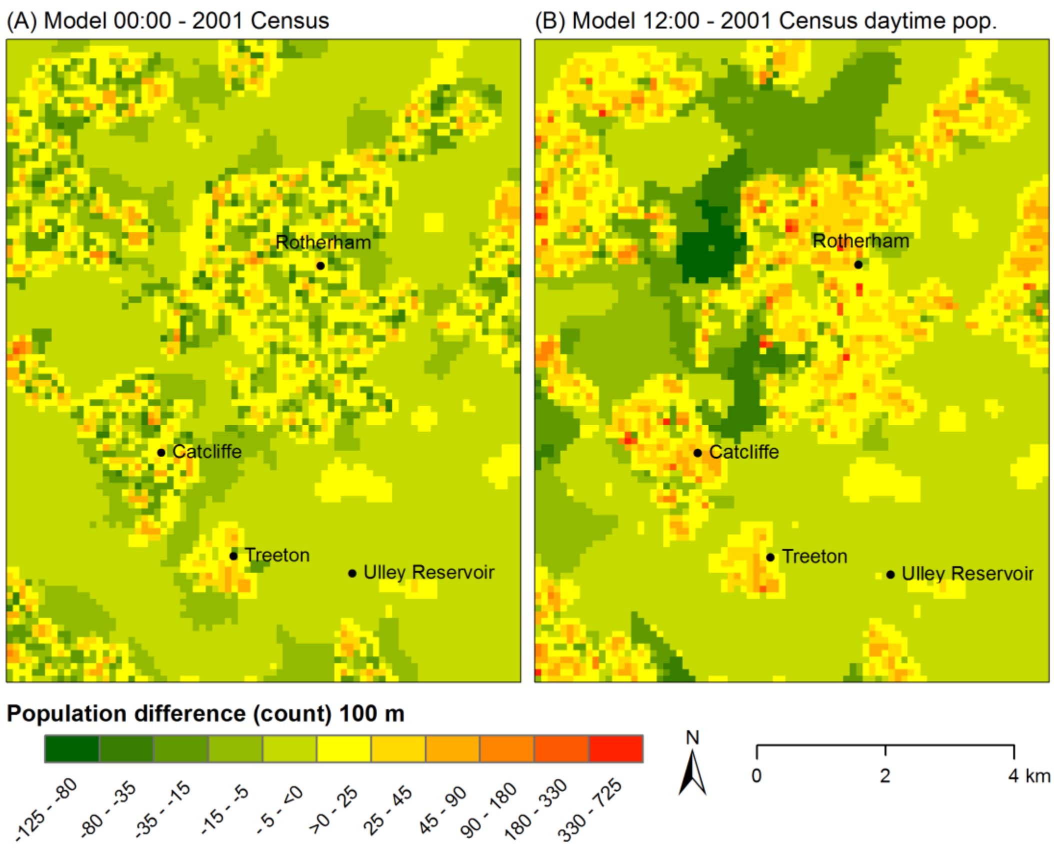


Figure 4.20 Cellular (100 m) population difference between the (A) 00:00 modelled outputs and the 'night-time' 2001 OA census estimate and (B) 12:00 modelled output and the daytime 2001 OA census estimate.

4.6 Summary

The chapter has presented one potential flood output from a reservoir breach event that actually nearly occurred during June 2007. It is acknowledged that this does not represent the only possible flood event from this reservoir but was constructed using the best available data. However, the modelled results can be corroborated with reasonable flood depth estimates from photographs taken at the time of the event (e.g. Figure 4.4). The rapid onset of this and documented similar events, and the volume of water within the reservoir to be released can be determined to a high degree of certainty. The primary focus of this chapter was to assess how vulnerable, exposed, populations fluctuate with time to a given hazard event.

This case study has demonstrated strong diurnal cycles in population exposure that includes population not resident within the flood zone (e.g.

those working or travelling through the area). These spatiotemporal movements are not represented using census data alone. The static census estimates have been shown to dramatically overestimate population exposures (e.g. Figure 4.16). This lacks the sufficient detail to begin targeting specific locations and times when populations will be most vulnerable. The flood risk analyses undertaken are sensitive to population fluctuations, as would be expected in reality. The spatiotemporal method applied does not attempt to give the final answer on Ulley's flood risk that is subject to external conditions and changes in physical characteristics. However, this example does show that time-specific populations are required to make improved assessments of hazard risk.

Chapter 5: Case study II – St Austell

5.1 Overview

This chapter outlines the second Population 24/7 application case study focused around St Austell, Cornwall, UK and considers the population exposed to fluvial and coastal flooding. It offers innovations beyond the standard implementation of the spatiotemporal population modelling framework offered in Martin *et al.* (forthcoming) through expansion of the population data library. This location has been chosen because it experiences large seasonal population fluctuations driven by tourism and overnight visitors. These spatiotemporal properties have been evaluated within the framework of Population 24/7. The previous examples (Ulley, Chapter 4 and Southampton, Chapter 3) have demonstrated that daily cycles can be resolved. This chapter aims to enhance this by examining how population varies in both space and time with seasonal influences with a 100 m resolution.

This chapter is structured as follows: it begins with the background to the St Austell study area and flood risk context (Section 5.2). Section 5.3 introduces case study specific flood inundation data. Bespoke flood modelling has been completed by Quinn (2014) and provided for use in this case study. This section also outlines the construction of a flood hazard rating and population fatality estimate (following the method introduced in Chapter 4). Section 5.4 offers case specific population enhancements for use within the Population 24/7 framework. This centres on creating a new seasonally varying population origin classification concerning tourist populations. This is based on seasonal overnight visitor data provided by Newing (2014). Further non-term time origin population dataset is also outlined. The creation of new destination datasets is introduced. These include the development of retail, healthcare and leisure destinations. The results following the integration of the flood and population modelling are presented in Section 5.5. Finally, a difference analysis and model evaluation is provided in Section 5.6.

5.2 Study area and flood context

A 15 × 20 km study area named after its most populous town, St Austell, has been selected in Cornwall, UK (BNG origin x: 200000, y: 49000 m). Cornwall is an English coastal county and local authority area which comprises most of the south-westerly peninsula of Great Britain (Figure 5.1). This area has been selected because of the large seasonal fluctuations in population that it experiences as a result of its status as a major domestic tourism destination. Preliminary analysis of tourism data (see Section 5.4.1) suggested that a peak increase of greater than 10,000 temporary overnight visitors occurs over the year from the low to peak tourism season. A recent 2011 Census report, published since this research has been undertaken, on coastal communities has identified locations such as St Austell has having unique and notable characteristics (ONS 2014c). According to this report these are often identified has having higher than average populations aged over 65 (an important consideration for flood risk vulnerability) and greater than usual occupation by people other than ‘usual residents’.

The range of tourist attractions, rural landscapes and attractive coastline makes the county a desirable visitor destination. This area was chosen to examine seasonally dependant temporal signals in population change using the SurfaceBuilder247 software. So far examples presented here (Southampton and Ulley) have demonstrated strong daily population cycles in exposure to flood risk. However, the St Austell application provides an additional perspective on how these cycles also vary by season and day of week. There are two factors that this case study specifically considers: the influx of seasonal population and the changes in the student population which is discussed in detail in Section 5.4.1.

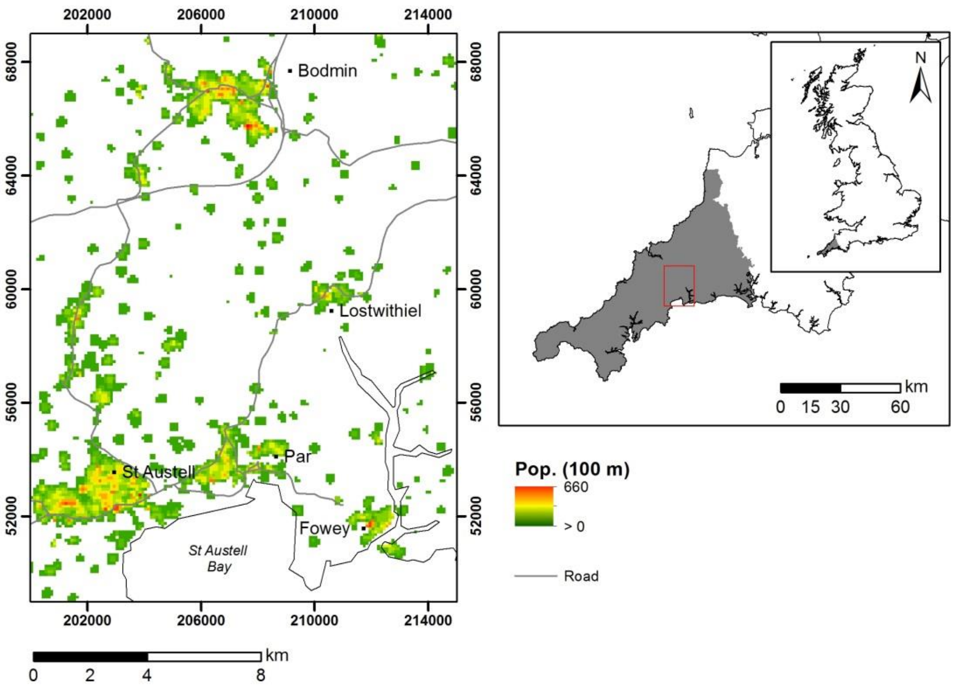


Figure 5.1 St Austell study area outlined in red, showing location within Cornwall (shaded grey) and Great Britain insets. An example 100 m gridded population distribution provided for contextual purposes.

The study area consists of the primary locations: St Austell, the largest town in Cornwall by population (19,958 2011 Census), Bodmin, Par, Lostwithiel and Fowey. The A30 (principal trunk road) and Great Western mainline railway (London to Penzance, including stations at Lostwithiel, Par and St Austell) are important infrastructural assets that intersect the study area. The south of the study area is bounded by the coast along St Austell Bay and the Fowey estuary. Relatively small settlements are dispersed throughout pastoral farmland with the large expanse of Bodmin Moor to the northeast.

The following paragraphs concern the flood context to the study area and events of significance within the wider region in which it is situated. The West and East Cornwall Catchment Flood Management Plans (EA 2012b, 2012a) provide an overview of the flood risk within the study area. Catchment and Shoreline Management Plans (CFMPs and SMPs) assess the risk from tidal and inland flooding.

The St Austell Bay area including Par and St Blazey is subject to fluvial, tidal and surface water flooding. Exposed critical infrastructure includes three electricity substations, fire and police station (St Blazey), telephone exchanges and a care home (Par). St Austell is not served by a flood warning system. The warning system on the River Par provides less than two hours' notice of flooding (EA 2012b). Although St Austell's elevated position is at little risk from tidal flooding the increase in settlements in surrounding low lying coastal areas such as Par and St Blazey are of concern. Furthermore, the 'tide-locking' of local watercourses, high tides preventing drainage at coastal outlets, poses an additional risk of fluvial flooding. Tidal flood risk dominates the east of the study area (Par, Fowey and Lostwithiel) (EA 2012a).

The Par area (including the neighbouring village of St Blazey) contains the highest number of properties at risk from current and predicted future flooding (2100 projection for 1% annual probability) in the whole of Cornwall, even when considering existing flood defences.

Table 5.1 Number of properties at risk from current and future flooding (1% annual probability of occurrence) in selected principal locations within the study area for the years 2010 and 2100

Location	Properties presently at risk (2010)	Projected increase (2100)	Flood hazards
St Austell	230		White River
Par/St Blazey	630	} 75	River Par, St Blazey Stream
Lostwithiel/Fowey	350	70	Fowey estuary and tidal flooding
Camelford/Bodmin	214	30	River Camel
Total	1424	175	

Major flood events in Cornwall can occur at any time of the year (Figure 5.2), however the probability based on a record of major historic flood events (Cornwall Council 2011) varies by month. A general cycle in events appears to correlate with unsettled autumn weather and winter depressions (October to January) and summer convective storms (June to August). Major flood events within the study area are summarised in Table 5.2.

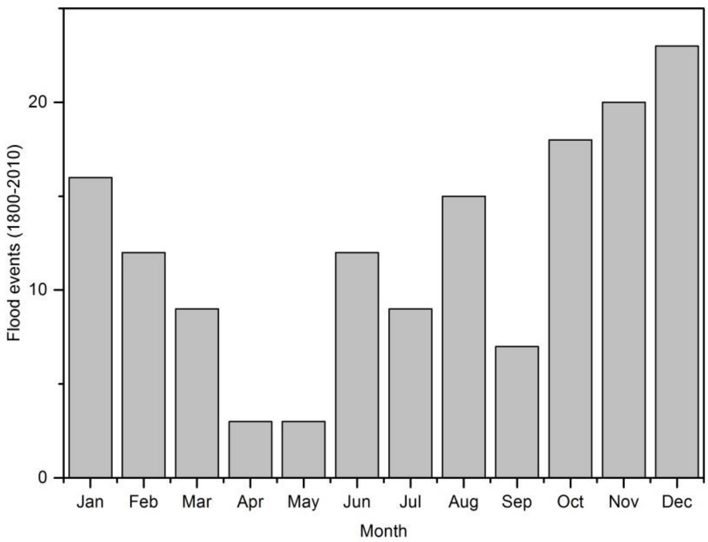


Figure 5.2 Annual distribution of major flood events in Cornwall (1800-2010) (Data: Cornwall Council 2011)

Table 5.2 Selected major flood events (1800-2010) from within the study area (Cornwall Council 2011)

Date	Location	Description
November 1852	Par and St Blazey	Serious flooding
November 1954	Lostwithiel	80 properties flooded
August 1959	St Blazey	Flooding
February 1974	Par and St Blazey	50 properties flooded
July & September 1975	St Austell	Flooding following leat breach
March 1976	Par and St Blazey	Extensive flooding
September 1976	St Austell, Par and St Blazey, Lostwithiel, Bodmin	Widespread flooding following severe storms
December 1979	Lostwithiel	60 properties flooded
October 1981	Fowey	Storm surge and high tides
December 1981	St Austell, Par and St Blazey	Fluvial and surface water
May 1996	Lostwithiel	15 properties flooded

More recently a succession of south-westerly winter depressions and associated storm surges during January-February 2014 caused severe tidal flooding in southern England and Wales which also affected parts of Cornwall. Coastal railways in Gwynedd (Wales) and Devon (England) collapsed into the sea following the destruction of supporting embankments. In Aberystwyth, Wales, 600 students were evacuated from seafront residences following the direct threat of severe coastal flooding (Gevertz 2014). The general synopsis of the 2014 flooding conditions has been described here for the region in which the St Austell study area is situated because of their impacts on understanding how people react during emergency situations. The village of Moorland in Somerset was particularly badly affected and isolated by floodwater. A severe flood warning indicating imminent danger to life was issued by the Environment Agency. A police helicopter's public address system was used to broadcast overhead evacuation warnings of imminent flood danger (BBC 2014a). According to this media report the police intervention caused panic to some local residents who continued to ignore evacuation directions despite rapidly rising flood waters.

A 100 m section of the Great Western mainline railway at Dawlish, 75 miles east from Par station (National Rail Timetable 2014), collapsed into the sea following a storm surge on 5 February 2014 (Figure 5.3). This exhibited unprecedented levels of damage and repeated storm surges. As a result the only rail connection to the southwest of England, including the St Austell study area was lost for two months while the sea wall was repaired.

Figure 5.3 Collapse of the railway at Dawlish, Devon in February 2014 following a succession of storm surges. Photograph: www.networkrail.co.uk

5.3 Case specific flood risk mapping

Environment Agency flood risk mapping has been obtained for the study area and follows the method described in Chapter 3 and Smith *et al.* (2014a). It is used to determine the potential exposure of population to the flood risk within the study area. The Environment Agency flood map zone three (April 2014) (Figure 5.4) represents the extent of the annual probability of occurrence for flood risk from rivers and the sea of ≥ 1% and ≥ 0.5% respectively. Additionally an ~ 8 × 4 km subsection of the study area has been modelled using LISFLOOD-FP, a raster based flood inundation model, by Quinn (2014) (dashed outline Figure 5.4) (see next Section 5.3.1). This area was selected because it covers the main population centres within the study area as well as the greatest flood risk identified by the Environment Agency flood map. The seasonal variation of population within this zone and potential exposure to flood risk has been performed in additional detail.

Figure 5.4 shows that the study area has a dispersed range of leisure and tourism attractions which are labelled in Figure 5.4. These are of significance in terms of the justification for the choice of study area and impact of seasonal population change within an area of flood risk. This is examined in more detail within Section 5.4 which addresses the population characteristics to be modelled.

Figure 5.4 Flood risk within the St Austell study area

5.3.1 **LISFLOOD-FP flood inundation model**

In contrast to the static Environment Agency flood map (Figure 5.4), designed for national coverage, bespoke modelling of the individual characteristics of the precise area concerned provides the potential to support more detailed analysis and scrutiny. For this study area LISFLOOD-FP was used. The LISFLOOD-FP flood inundation model predicts channel and floodplain flows. Although there are a wide range of hydraulic models suited to different tasks (a selection have already been discussed or implemented in Chapter 4, e.g. TELEMAC-2D) LISFLOOD-FP provides a simplified alternative.

Three LISFLOOD-FP scenarios at a 5 m resolution have been created and provided for this case study by Quinn (2014). It has been suggested that at least a 5 m resolution is essential for modelling flow dynamics within urban areas to account for small scale variations (Mark *et al.* 2004; Fewtrell *et al.* 2008). The layers provided represent return periods (R) of 100, 250 and 500 years (Figure 5.5) for an extreme rainfall event of 11 hours duration. The LISFLOOD-FP layers explicitly account for flood defences, topographic features and blocking due to buildings (derived from 2m LiDAR). The risk posed by rainfall intensity for each return period was calculated using the Flood Estimation Handbook software (CEH 2014) by Quinn (2014). This estimates rainfall frequency for the UK which relates rainfall depths to a given probability of occurrence.

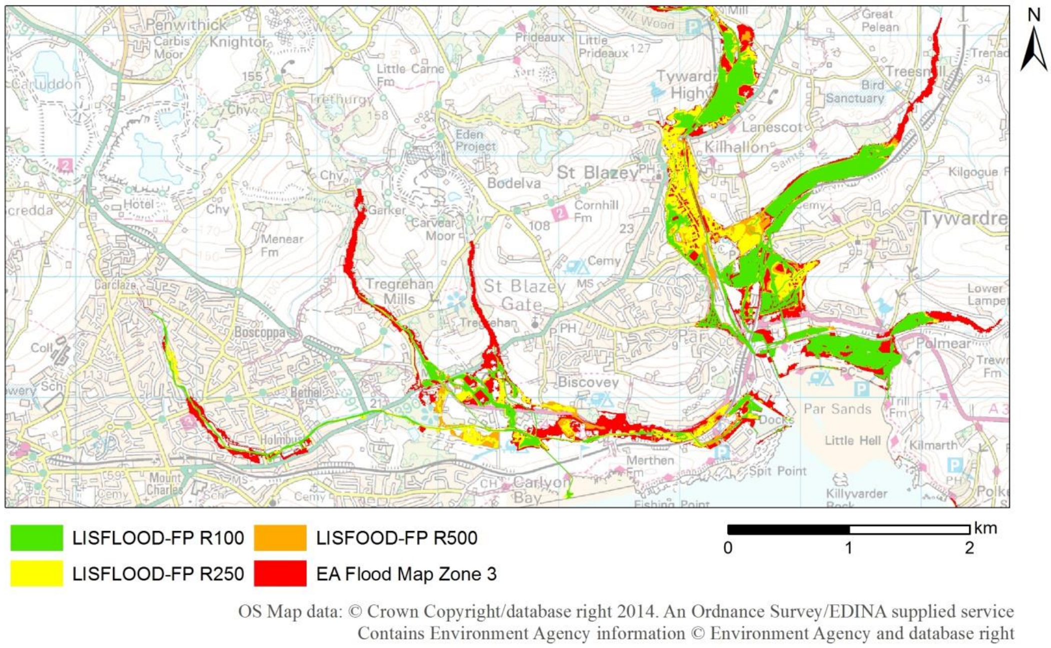


Figure 5.5 Comparison of LISFLOOD-FP and Environment Agency flood inundation for the selected area covering St Austell and Par within the study area.

LISFLOOD-FP is an appropriate model for this application because it is both computationally efficient at high resolution (1-10 m) and the code can be run on the latest high performance computing technology (Neal *et al.* 2011). The 5 m (minimum) resolution is required for this case study to accurately predict flow hydraulics within the urban centres of the study area and to model flow around buildings and coastal defences. This information is required to produce flow velocity and depth estimates. These were combined by Quinn (2014) to calculate a flood hazard rating. This follows the same method presented in Chapter 4. Three raster layers with a flood hazard rating for each 5 m cell corresponding to the three return periods of

100, 250 and 500 years with a flood hazard rating were provided for analysis with population outputs from SurfaceBuilder247.

In comparison, the Environment Agency's flood map zone three represents the amalgamation of return periods of 100 (fluvial) and 200 (tidal) years for flood risk and does not account for the presence of any defences. While it is possible to obtain an extract of nationally consistent coverage, the LISFLOOD-FP outputs provide a more realistic scenario accounting for existing defences. The EA flood map provides an inundation polygon but does not estimate flow depth or velocities. Without these data, which are contained within LISFLOOD-FP outputs, it is not possible to evaluate a flood hazard rating or fatality estimate.

The Environment Agency flood map shows a much greater extent of inland flooding (Figure 5.5), however it also includes coastal flooding and assumes that there are no defences whereas the bespoke LISFLOOD-FP extracts specifically account for these. Furthermore, in November 2014 it was reported that the EA is reassessing English coastal flood risk as some maps may underestimate these risks (BBC 2014b). This follows a review of the widespread coastal flooding experienced at the start of 2014, which was summarised at the start of this chapter. All layers have been combined with seasonally varying population estimates to analyse the effect of spatiotemporal cycles. The EA Flood Map has also been included here because it is the currently accepted national flood risk assessment used by planners and local authorities.

5.3.2 Flood hazard rating and fatality estimates

This subsection reintroduces the method for calculating a flood hazard rating, the first of a series of subsequent stages to obtain a population fatality estimate. The hazard rating aims to quantify the flood hazard characteristics (depth and velocity). It follows the same method as previously demonstrated in the Ulley example (Section 4.3.5 and Penning-Rowsell *et al.* 2005). Following the calculation of a flood hazard rating, area and people vulnerability ratings need to be determined to finally estimate the number of injuries and derive a fatality prediction.

For this case study the flood hazard rating has been calculated by Quinn (2014) using the depth and debris outputs from the LISFLOOD-FP inundation modelling and provided as a raster layer. This has been conducted following the same widely accepted hazard rating formula (Eq. 4.1) used in the Ulley example.

An Area Vulnerability has been calculated for the St Austell study area as 6 (1 + 3 + 2). This is based on the score matrix provided previously in Table 4.4 where speed of onset = 1 (low risk, many hours – based on a model scenario for an 11 hour rainfall duration), nature of area = 3 (high risk, mobile homes, bungalows and busy roads) and flood warning = 2 (medium risk, warning system present but limited). There is a flood warning system within the area but not on all rivers (see Section 5.2).

To determine the study area's people vulnerability (PV) score (see Chapter 4, Eq. 4.3, Table 4.6) population data on age composition and long-term illness was taken from the 2001 Census (Table 5.3). Overall the population within the study area aged over 75 and the proportion of those with a long-term illness are above the average for England. Therefore a PV score of 100% was assigned (50% + 50%).

Table 5.3 Population characteristics within the St Austell study area by parish (Census 2001, Tables KS002 and KS008) compared to the national average (England).

Location	Pop. aged over 75 (%)	Difference from national average†	Pop. Long-term illness (%)	Difference from national average‡
St Austell Bay	10.91	+3.37	19.01	+1.08
Tywardreath and Par	12.85	+5.31	23.98	+6.05
St Blazey	6.61	-0.93	21.23	+3.30
Lostwithiel	10.96	+3.42	20.34	+2.41
Fowey	13.45	+5.91	23.41	+5.48
Study area mean	10.95	+3.42	21.59	+3.66
England (Average)	7.54†		17.93‡	

The hazard rating, area and people vulnerabilities were combined according to Eqs. 4.4 and 4.5 to estimate the number of fatalities for each of the

LISFLOOD-FP scenarios (see 5.4 Results, Table 5.11). This was evaluated for each cell of the spatiotemporal population outputs. These were evaluated using raster based calculations of these equations in ArcGIS. The spatiotemporal population output was resampled from 100 m to 5 m (resolution of LISFLOOD-FP cells) for the purpose of the raster based calculations using ArcGIS.

5.4 Case specific population method and data

This section outlines specific population modelling using SurfaceBuilder247 (Figure 5.6) and data sources for the St Austell study area. These are in addition to the broad overview described within the main methodological section (Chapter 3). Specific enhancements have been identified to ensure the best representation of the study area's population within the Population 24/7 framework. While the main methods overview is still applicable a number of additional enhancements have been developed which have been identified as specific to the requirements of this study area. These enhancements are discussed in turn within the following subsections: accounting for seasonal population variations, creating a population leisure destination dataset, enhancing retail footfall estimates and producing a healthcare dataset and finally determining a weekday/weekend workplace capacity estimate.

For the purpose of the spatiotemporal population modelling a 25 km buffer has been assigned to the study area based on analysis of average distances travelled to work for each OA (2001 Census, Table UV35) to eliminate any potential edge affects. However, the study area and buffer region is largely geographically constrained in this example to the north and south by the coastline (Figure 5.6). Furthermore, the northeast of the study area is constrained by Bodmin Moor, 150 square miles of moorland which also contains Cornwall's highest point at 1,368 ft AOD (Visit Cornwall 2014). Within the model these features are represented through the background masking layer (Figure 5.6) which prevents population placement off-shore. They are also represented within the postcode centroid dataset where there is a natural lower density of origin and destinations locations on the moor and surrounding rural landscape which are geographically constrained to the physical vicinity of residential addresses.

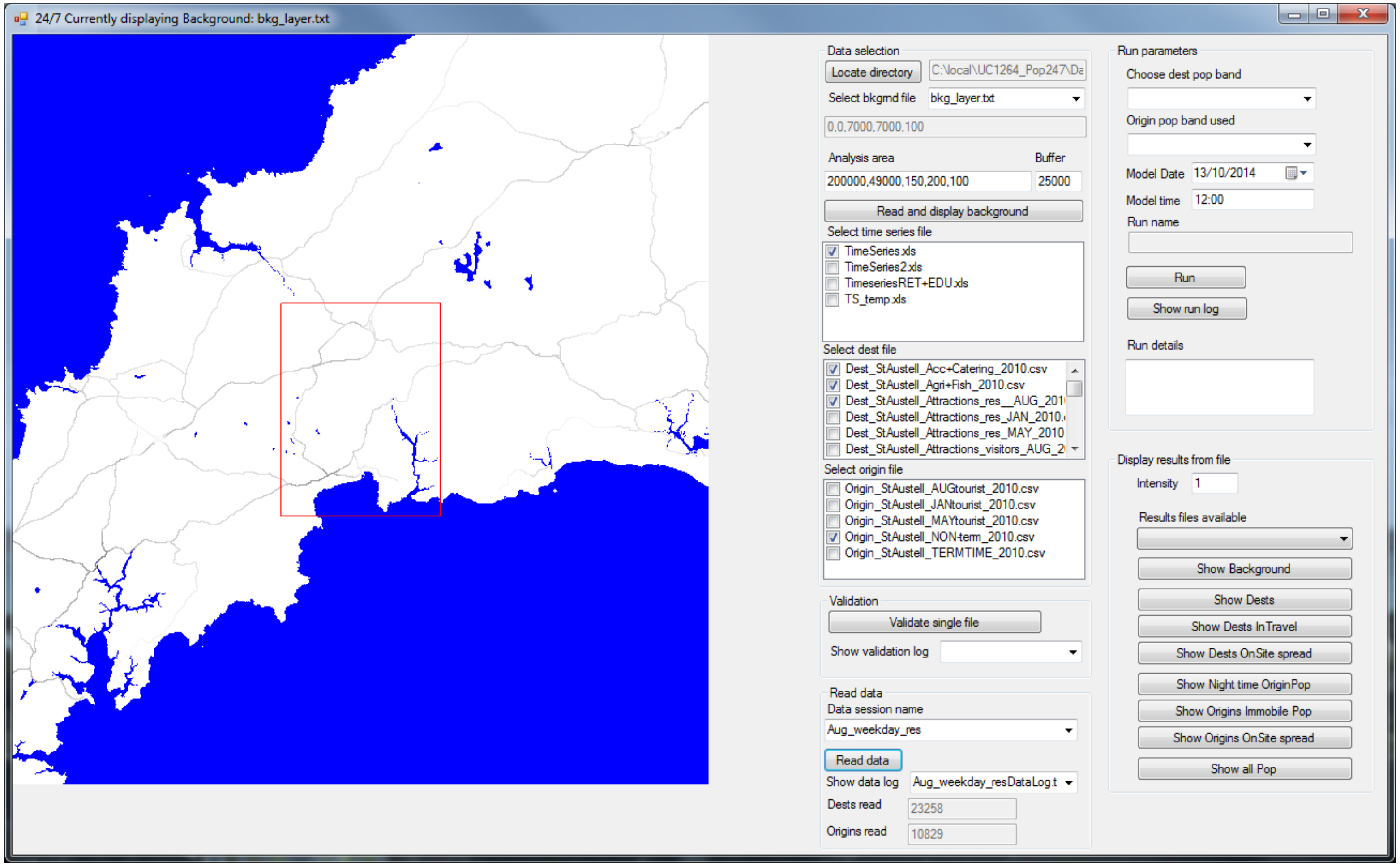


Figure 5.6 SurfaceBuilder247 software interface showing St Austell study area (red rectangle 15 × 20 km, surrounding buffer 25 km) session parameters

A diagrammatic overview of population origin and destination datasets and their seasonal variation is provided in Figure 5.7. It is constructed in two rows and three columns. Each column represents the seasonal scenario modelled. The first row concerns population origins and shows two different classes of origin which contain the usually resident and overnight visitor populations separately. The second row contains the destination locations. The connecting arrows show that the different origin classes populate different (or different proportions of) destination locations. For example the visitor population does not populate workplace or school destinations but is assigned to leisure destinations. Table 5.4 summarises the data sources used to construct the model's data library. These are referred to in the following sections regarding the construction of case specific population data libraries to undertake the spatiotemporal modelling.

In keeping with the previous applied examples the population has been modelled in seven age subgroups. The population is subdivided into the following age subgroups: 0 to 3 years, 4 to 10 years (primary school pupils), 11 to 15 years (secondary school pupils), 16 to 64 years (further education students), 16 to 64 years (higher education students), 16 to 64 working aged and those aged over 65. These have been chosen for their unique spatiotemporal characteristics.

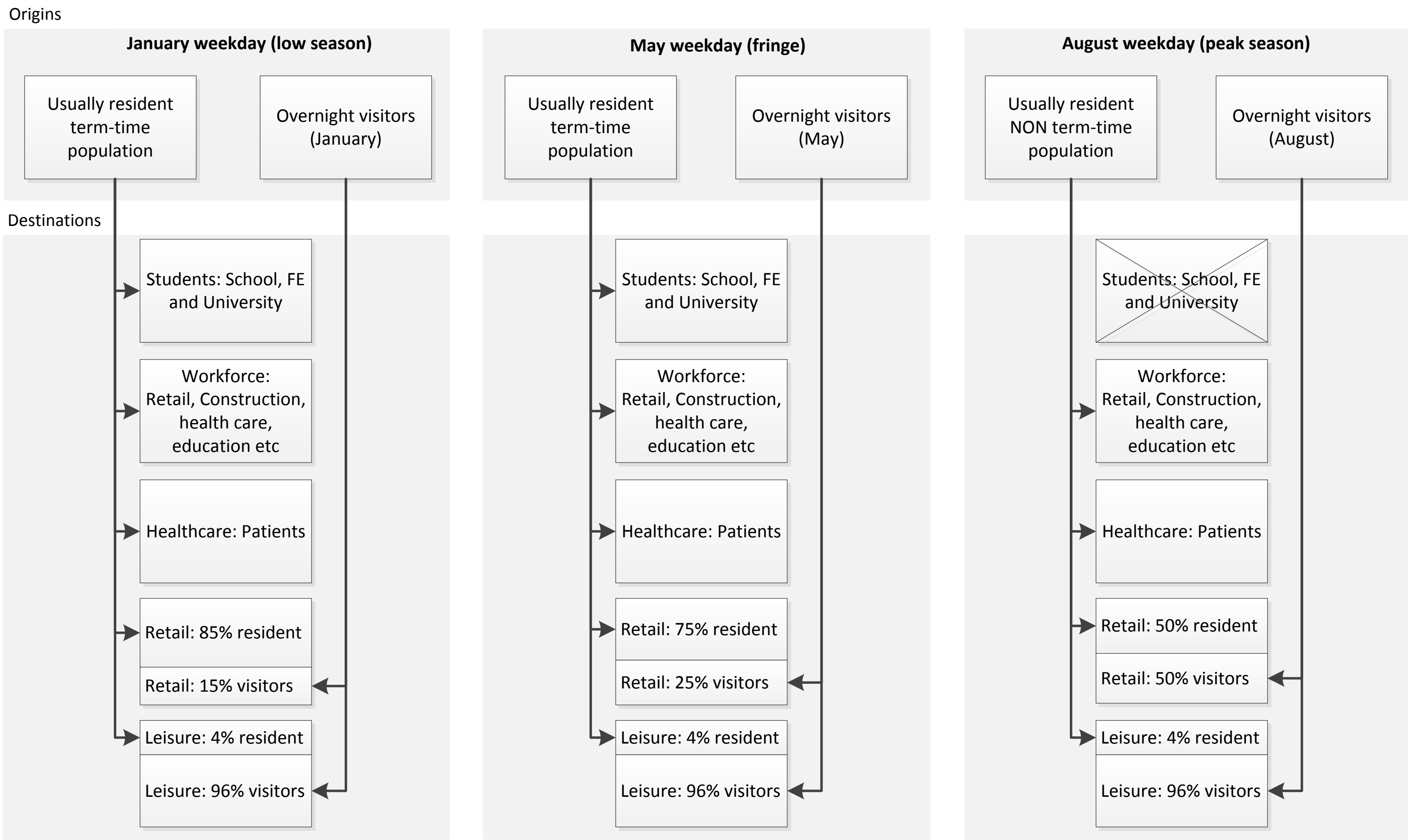


Figure 5.7 Diagrammatic overview of population origin and destination datasets used to construct the St Austell case study model

Table 5.4 Summary of content and sources for the St Austell study area population data library constructed for SurfaceBuilder247

File Name	File Description	File Type	Data source(s)			
			Population	**Location geography**	**Temporal profile**	**Wide Area Dispersion**
St Austell usually resident population	Baseline 'term time' population	Origin	2010 MYE	NSPD (Feb 2010) *residential UPCs*	N/A	N/A
St Austell non-term time resident population	Baseline non- 'term time' population	Origin	2010 MYE; Census (2001, Tables UV05, KS13)	NSPD (Feb 2010) *residential UPCs*	N/A	N/A
St Austell JAN	Resident overnight visitors in January	Origin	Newing (2014); VisitEngland (2010)	NSPD (Feb 2010) *residential UPCs*	N/A	N/A
St Austell MAY	Resident overnight visitors in May	Origin	Newing (2014); VisitEngland (2010)	NSPD (Feb 2010) *residential UPCs*	N/A	N/A
St Austell AUG	Resident overnight visitors in August	Origin	Newing (2014); VisitEngland (2010)	NSPD (Feb 2010) *residential UPCs*	N/A	N/A
Education	Count of pupils and students in full time education (school, college and university)	Destination	Schools Census (DfE 2010); Independent Schools Census, HESA (2010)	NSPD (Feb 2010) *Georeferenced UPC*	Population 24/7 project	Population 24/7 project

Table 5.4 continued

File Name	File Description	File Type	Data source(s)			
			Population	**Location geography**	**Temporal profile**	**Wide Area Dispersion**
Retail visits	Retail outlets and centres	Destination	See section 5.4.4	GMAP (2014)	TUS (2001)	NTS (2010)
Visitor attractions	E.g. historic houses, gardens, places of worship	Destination	Visit England (2010); English Heritage (2010)	Site location (UPC)	TUS (2001)	Visit England (2010)
Healthcare	Healthcare patients	Destination	HES (2010); local websites and FOIs.	Site location (UPC)	HES (2010) arrival times	Roberts *et al.* (2014)
Education workforce	Workplace population counts (SIC: P)	Destination	BRES (2010)	NSPD (2010) *Business UPCs*	LFS (2010)	Census (2001, Table UV80)
Health service workforce	Workplace population counts (SIC: Q)	Destination	BRES (2010)	NSPD (2010) *Business UPCs*	LFS (2010)	Census (2001, Table UV80)
Manufacturing and construction workforce	Workplace population counts (SIC: C, F)	Destination	BRES (2010)	NSPD (2010) *Business UPCs*	LFS (2010)	Census (2001, Table UV80)
Wholesale and retail workforce	Workplace population counts (SIC: G)	Destination	BRES (2010)	NSPD (2010) *Business UPCs*	LFS (2010)	Census (2001, Table UV80)

Table 5.4 continued

File Name	File Description	File Type	Data source(s)			
			Population	**Location geography**	**Temporal profile**	**Wide Area Dispersion**
Transportation workforce	Workplace population counts (SIC: H)	Destination	BRES (2010)	NSPD (2010) *Business UPCs*	LFS (2010)	Census (2001, Table UV80)
Catering workforce	Workplace population counts (SIC: I)	Destination	BRES (2010)	NSPD (2010) *Business UPCs*	LFS (2010)	Census (2001, Table UV80)

Notes: SIC (Standard Industrial Classification, BRES 2010)

5.4.1 Seasonal visitor population

This subsection outlines seasonality within the student and overnight domestic visitor populations and discusses how and why they are addressed separately for the purpose of the spatiotemporal population modelling for this case study. Cornwall attracts more than 4.4 million annual domestic visits (2006-2009 average) which makes it the second most popular English county destination outside of London (Visit England 2010a). Tourism accounts for 25% of all employment in Cornwall and tourism is estimated to contribute £1.8 billion to the UK's economy (Visit Cornwall 2011). This large fluctuation in population, as well as daytime movements, is not represented in traditional census counts or one single dataset alone. In order to determine this seasonal, as well as spatial cycle the Population 24/7 approach is employed (Chapter 3). This requires the compilation of an extensive population data library (Table 5.4).

In comparison to the previous case study and worked example (Ulley and Southampton) this study area is predominately rural with urban clusters of population. Seasonal fluctuations comprise a large proportion of the usually resident population due to its tourism industry. A usual resident is defined as someone who spends the majority of their time residing at that address (National Statistics 2004, p. 17). It does not include students living away from home or people temporarily present at the location while on holiday or visiting friends. Therefore additional information is required to account for overnight visitors who are not enumerated within the usually resident census population. Three seasonal scenarios have been chosen for January, May and August (target year 2010) to represent the low, fringe and peak tourism seasons respectively.

Seasonal overnight visitor counts in Cornwall for January, May and August 2010 have been provided by Newing (2014) aggregated to postcode level with population age breakdown. These have been constructed at the level of individually georeferenced commercial accommodation sites (e.g. a hotel, guesthouse, campsite or holiday park or commercial accommodation chain). These counts also include all overnight visitors staying with family or friends or within a second/holiday home, with these visits being distributed across the existing residential housing stock.

The methodology is based on bed space and occupancy data which is described fully in Newing *et al.* (2013a). They derive the location of tourism accommodation providers and the total number of bed spaces available for each individual site from data collected by South West Tourism. In addition the locations of second homes are derived from the 2001 Census. Population at second homes is not included in the usually resident census count. Having established accommodation provision utilised by visitors, occupancy rates are then applied. Occupancy rates were derived from data published by South West Tourism and Visit Cornwall and show a strong seasonal cycle. In August (peak season) occupancy rates were high (94%) falling to just 9% in January (Newing *et al.* 2013a).

These data have been directly used as provided except for the required additional formatting for use as SurfaceBuilder247 population origin datasets. The location of accommodation provision remains constant but the occupancy rate varies by season so three separate origin datasets were created to represent overnight visitors in January, May and August (Figure 5.7). They represent the night-time temporary address of domestic overnight visitors and are not included within the usually resident census population. The change in overnight visitor estimates in each of these seasons for the St Austell study area is summarised in Table 5.5.

Table 5.5 Overnight visitor estimates within the St Austell study area. Original data source: Newing (2014)

Month (season)	Overnight visitor estimate
January 2010 (Low)	1,049
May 2010 (Fringe)	6,269
August 2010 (Peak)	12,389

5.4.2 Non-term time population

Seasonality within traditional census counts is limited, as already highlighted in the previous section. However, it is still possible to make an important distinction between the educational term and non-term time populations. The movement of student populations, particularly those in higher education moving away from their census non-term time ‘home’ addresses to places of study, results in a seasonal change. For the purpose of the census students are enumerated at their term-time address (e.g.

college/university halls of residence or other private temporary accommodation). Basic characteristics are reported separately at their non-term time OA (National Statistics 2004). Therefore using the following method it is possible to use these two reporting outputs to derive the term and non-term time student counts for each census OA (2001 census table references in brackets):

Term-time population = *Usually resident population (KS001)* (5.1)

Used as reported as this can be considered the term-time count with students enumerated at their term-time address.

Non-term time population: *KS001 + UV05 – KS13* (5.2)

Where 2001 Census Table KS001 is the *Total usually resident population*, Census Table UV05 represents the *Schoolchildren and Students in Full-time Education Living Away From Home During Term-time* and Census Table KS13 is the *Total number of full-time students and schoolchildren: Aged 18-74 (Qualifications and Students)*.

To derive the non-term time population base, this basic adjustment returns students to their respective OAs that have lost their student population to term time residences (2001 Census Table UV05) and removes those who were there during term-time (i.e. who would have returned to their respective non-term time OAs) (2001 Census Table KS13). The greatest magnitude in change is likely to be experienced within university towns and cities. This is where the term-time student population would be expected to be highly concentrated within an accessible distance to higher education institutions. In contrast the non-term time population is very dispersed.

The caveat in the method applied here is the absence of the recognition of international students, who if ‘resident’ within the UK on the census night would have been enumerated at their British term time address. However, they would be removed from their respective OAs during vacation periods following this method. In reality it is not certain whether these students remain at their term-time location, return to their country of origin or reside elsewhere. Secondly, Scottish and Northern Irish students retuning to a home address outside of England and Wales are lost. These are important limitations to consider, but it is not deemed a matter of significance within the study area selected for the reasons outlined in the following paragraph. The difference in the term and non-term time population following this method for the St Austell study area is summarised in Table 5.6.

Table 5.6 Usually resident population within the St Austell study area (2001 Census adjusted for 2010 MYE).

Origin file	Population
Term time	67,599
Non-term time	67,830

As expected for this study area the usually resident population summary (Table 5.6) following the approach outlined (Eq. 5.2) shows little term to non-time population change, with a negligible increase during non-term time periods (+0.34%). This study area does not contain any higher education institutions (HEIs) and therefore it does not attract large numbers of term-time only student populations. The rise in the non-term time approach can be explained by those students who have studied at HEIs external to the study area having then returned home for vacation periods. The distinction has been made for this study area in order to demonstrate this conceptual component of the Population 24/7 modelling approach. This illustrates that the model can handle, and that it is important to consider, multiple input population origin datasets to best represent the baseline population required for the time, location and scenario modelled.

The approach outlined within these subsections can be viewed in context within the overview diagram (Figure 5.7). It reveals that each seasonal scenario requires a pair of population origin files, the usually resident census population and temporary visitors. It has been highlighted that both of these change throughout the year. For example, the August (peak season) scenario has the greatest number of overnight visitors (Table 5.5) as well as a non-term time baseline census population as August is a vacation period, whereas January and May receive their respective portion of overnight visitors but use the same baseline census term-time resident population origin because these months fall within academic term-time. Following the method outlined in Chapter 3 all census population origin data have been re-weighted from census reporting zones (OA and LSOAs) onto georeferenced residential postcode centroids weighted according to their address count. This is with the exception of the overnight visitor data that has already been constructed at this level. The usually resident 2001 Census baseline populations have been adjusted to mid-2010 counts (target year). The data match the target year but use the 2001 Census geography (as 2010 is an inter-censual year).

For the purpose of the non-term time August scenario the education (students and pupils) destination dataset has not been used. Separate usually resident and overnight visitor origins ensure that destinations receive their population from the correct base (Figure 5.7). For example only the usually resident population is drawn by the workplace and education (term time only) destinations. The non-working overnight visitor population populates their share of the retail and leisure destination locations (as well as their own origin centroids).

Since this research has been undertaken an alternative out of term time population base has been created for England and Wales in the 2011 Census (ONS 2014b). This recognises that society is increasingly complicated and mobile. This out of term population base includes students and school children at their term time address (if no out of term address was provided), students and school children at the 'home' address (if a non-term time address was provided and was located within England and Wales) and all other usual residents at their usual residence.

5.4.3 **Leisure destinations**

This subsection outlines the creation of a leisure destination dataset within the flexible framework of the Population 24/7 approach. This dataset comprises the sites that overnight visitors and the usually resident population visit for leisure purposes as shown in the diagrammatic overview (Figure 5.7) for each month. This has been identified as a requirement in addition to Martin *et al.*'s (forthcoming) implementation to reflect visitor population movements within the model's data library.

The creation of this leisure destination dataset is described in the following order within this subsection: firstly the locations associated with leisure activities are defined and, where possible, associated with annual visitor numbers (currently only routinely published at the annual level). However, a seasonally varying daily estimate is required for the population data library. A method is proposed to derive daily proxies at the spatiotemporal resolution required. Secondly, the geographic location of leisure sites is required and is obtained based on georeferenced UPCs within the respective postal addresses of these sites. Thirdly, the age distribution of visitors to these leisure sites is obtained from demographic visitor data. This is required to divide the total population for each site into population subgroups according to age. Fourthly, the catchment area of each leisure site has been derived using national travel statistics. This governs the spatial extent of population origins that populate each leisure destination site. Finally, the ratio of usual residents to overnight visitors using these sites is estimated.

Within the Population 24/7 framework this leisure destination dataset will be populated by two classes of origin datasets concerning the usually resident and visitor populations as shown in Figure 5.7. Leisure attractions are not mutually exclusive to visitors or usual residents, however data shows that ratio of visitor to resident footfall to these sites is not equal and varies with seasonal cycles. Therefore three leisure destination datasets have been created for each seasonal scenario (Jan, May and Aug). They all contain the same attractions but seasonally varying population capacities. Each of these destination datasets is further subdivided to reflect the proportion occupied by visitors and residents. These are drawn from the respective origin class for each season represented by the links in Figure 5.7.

It has already been noted in the previous section (5.4.1) that temporarily resident visitors, at times, comprise a significant proportion of the study area's population. During the August peak the influx of seasonal overnight visitors represents an 18% increase on the usually resident non-term time baseline population (Tables 5.5 and 5.6). As expected for an area within a tourism-driven economy, the study area and buffer contain a large number of prominent visitor attractions. To account for these within the spatiotemporal modelling a leisure destination dataset has been constructed (analytical overview, Figure 5.7). These contain a range of locations such as historic houses, castles, gardens, botanic sites and places of worship. These are locations with known associated visitor numbers. These data are supplied by Visit England (2010b) in the *Annual Survey of Visits to Visitor Attractions*.

Thirty attractions with reported annual visitor numbers are located within the study area and buffer zone. These have been collated to produce a leisure destination dataset. A selection have been summarised in Table 5.7. The geographical location of each destination is assigned using postcode data. The postcodes for each site were determined from their respective websites. These were converted into national grid references using GeoConvert (UKDS 2014) to provide the destination centroid location. This facilitates the creation of a destination centroid dataset comprising a georeferenced location which can be assigned a population capacity for each attraction.

Table 5.7 Example attractions and 2010 visitor numbers within the St Austell study area and buffer (Visit England 2010b)

Attraction	Type	Visitor numbers
The Eden Project	Botanic site	1,000,511
Tintagel Castle	Historic property	190,246
Lanhydrock House	Historic property	210,362
Truro Cathedral	Place of worship	170,000
Trelissick Gardens	Garden	128,671

The Visit England survey shows that in 2010 the Eden Project, a major attraction comprising artificial biomes, was the most visited site in Cornwall and second in southwest England (after Stonehenge) with over a million visitors (Table 5.7). The Eden Project is also within the centre of focus within the study area located 3 miles NE of St Austell. The local tourism economy within this coastal area supports many large attractions. Sites of this size that attract major footfalls are of very high significance and therefore need to be addressed as destination locations within the spatiotemporal modelling undertaken for this study area.

The annual visitor count requires subdivision into monthly estimates which are not routinely published for each site. Monthly visitor data for 2010 for a typical tourist attraction within the area, Tintagel Castle, have been provided by English Heritage (2010) (Figure 5.8). Following personal communication footfall counts were made available for each month during 2010 by English

Heritage who manages Tintagel Castle on behalf of the UK government. The annual distribution has been used to derive monthly, and then daily, estimates for attractions within the study area for January, May and August. These months receive a 0.65%, 9.93% and 27.59% share of the 2010 annual visitor footfall respectively which is calculated from the Tintagel footfall distribution. The daily footfall was estimated by dividing monthly counts by the number of days (thirty-one) within each month chosen. This is a trade-off because daily footfalls are currently not available at the required spatiotemporal resolution. Typical for a tourism-driven region August received the highest proportion of annual visitors at 27% (Figure 5.8).

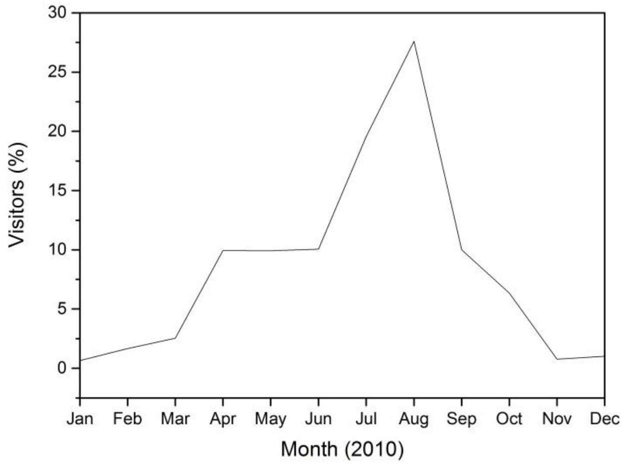


Figure 5.8 Monthly distribution of annual visitors at Tintagel Castle, Cornwall. Data: English Heritage (2010)

The temporal distribution of daily population capacities is governed by a temporal profile (previously described in Chapter 3). If hourly footfall figures were available this would inform the profile. However, a proxy has been constructed using the TUS (Ipsos-RSL and ONS 2000), episodic diary data detailing daily activities, because data at this level of granularity are currently not available. The diary entry codes for visiting a botanic, historic site and historic house reported at 10 minute intervals were used to construct the temporal profile. These are from a nationally representative survey which is the most recent comprehensive survey for time use in the UK.

Following the establishment of a daily footfall estimate, further information is required on the age characteristics of the destination population capacity for this case study (Table 5.4). This is required to maintain continuity in the seven age subgroups modelled as population subgroups within this example. Limited demographic information on visitors to attractions is provided at a regional level by Visit England. This can be used to inform the main age subgroups required. These data generally only contain adult respondents which limits the information available for visitors less than 18 years of age. The missing information of child demographics has been substituted by analysis of data collected for the *Taking Part in Heritage* report (National Statistics 2014). This details adult and child visits to heritage sites. This distinction is of importance for two reasons: (i) child population subgroups have been modelled because of their unique spatiotemporal characteristics, and (ii) heritage sites are often frequented through school trips and community organisations by these subgroups.

A further requirement for creating a destination dataset for the modelling undertaken is to define the destination centroid's catchment, the origins from which to draw population. Again, due to the granularity of currently available data, an appropriate substitute has had to be derived. The National Travel Survey (total respondents = 17169) records the distance travelled for a daytrip. The 2001-2012 average (Figure 5.9) has been calculated based on this variable ("LDJDistance", DfT 2012b) where participants specified the distance (miles) travelled for this purpose (daytrip travellers = 15058, mean distance = 87 miles). Within this survey a trip is defined as a one way course of travel having a single main purpose (DfT 2012a), appropriate for destination catchment calculations. While this method does not allow the determination of the exact activity undertaken for day trips it does give an appropriate quantification for leisure travel. These data have been computed into five bins for those who travelled less than or equal to 50 miles up to a distance greater than 150 miles.

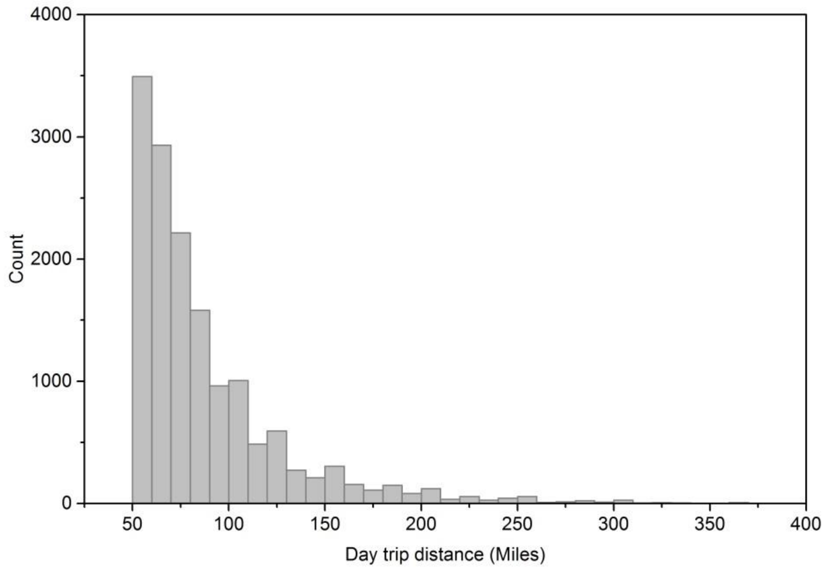


Figure 5.9 Distance travelled for day trips (NTS 2001-2012 average)

Finally, an adjustment is required specific to the nature of this study area. It has already been noted that two population origins classes (which also vary by season) are required for the three seasonal scenarios modelled (analytical overview, Figure 5.7). The leisure destination file has been split on the ratio of visitor (94%) to usually resident population (6%) footfall. This is illustrated in Figure 5.7 by the connection of the population origin class types to their respective share of the leisure destination capacity. This is an important consideration to represent the correct proportion of attraction footfall to visitors. The ratio has been informed using Visit England data providing the origin of visit by region. This has been used as a proxy to estimate the proportion of visits external to the study area that can be deemed as visitors for the purpose of this case study. Visits that originate within the same region have been designated as residential footfall. This distinction ensures that residential visits are drawn from the usually resident population origin dataset and external visitors from the overnight visitor origins. While the resident to visitor footfall ratio has been fixed the destination capacity does not remain constant by month. This varies in line with the seasonal cycle actually measured within the region (Figure 5.8). Six separate leisure destination datasets have been created based on this information (Figure 5.7): two leisure destination datasets (populated by the usually resident origin population and visitors respectively) for each of the three months modelled. The capacity of each month's respective resident and visitor populated destinations reflects variation within seasonal footfall.

5.4.4 **Retail destinations**

Retail locations are another major category of destinations which populations temporarily occupy. A retail destination dataset specific to this study area has been created as part of the modelling data library (Figure 5.7). This section introduces significant improvements regarding the creation of a retail destination dataset from what has already been described in Chapter 3. The difficulty in making such estimates was the focus of the work presented in Chapter 3. The rest of this section is structured as follows: the use of commercial retail centre data to locate retail destination centroids; the estimation of footfall based on sales density, floor space and average transaction values to assign a capacity to retail destination centroids, and finally, the analysis of the ratio by season of footfall generated by visitors and usual residents.

Detailed retail footfall data and innovative products are commercially available (e.g. Experian Goad, Telefonica Dynamic Insights). However for the purpose of this thesis demonstrating model development and application these are not financially viable for the scale required. It does not mean that these site specific counts could not be used, as they can be implemented following the same framework described for the creation of an alternative approach. The hybrid approach for estimating retail footfall for the study area within this section uses a combination of commercially available and public data.

Commercially available retail centres with associated national grid reference and floor space (GMAP 2014) were used as retail destination centroid locations. These are locations that mark the location of a retail 'centre', such as a parade of shops, a large stand-alone store (e.g. supermarket) or a retail park (Figure 5.10). Each is available with a total floor area subdivided into grocery and comparison retail space. For the purpose and enhancement of the population modelling undertaken these have some advantages. Using the retail centre centroids reduces point density for improved computational efficiency, rather than attempting to represent individual stores within a centre as separate ‘destinations’. The use of building level retail data is a finer resolution than it is currently possible to resolve within the model’s 100 m output cells. The model output cell size, although flexible, is currently constrained by the resolution of currently available population data. Therefore this resolution at the sub-cellular 100 m output level would be lost, but still result in an increase in computation time with no change in final results (Figure 5.10). Furthermore, consumers will likely visit multiple stores within any one centre. The spatial footprint around the retail centre centroid is still conserved as it is the sum of the constituent floor space, which is used to define the centroid’s local dispersion for the modelling process.

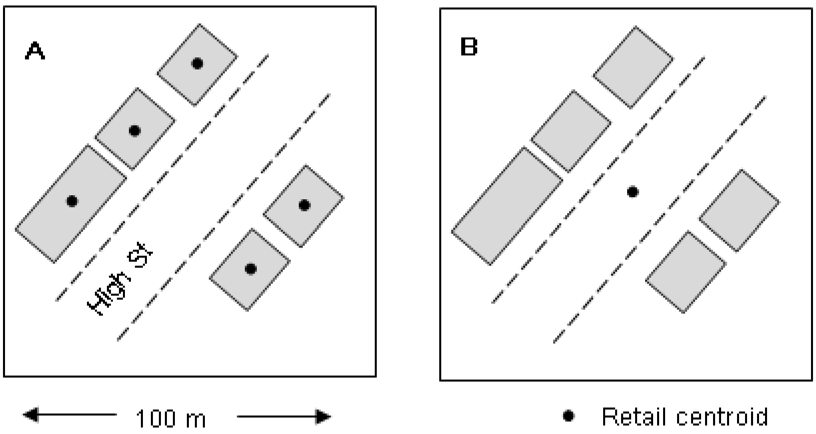


Figure 5.10 Hypothetical retail units within a 100 m model output cell (A) individual retail outlet location centroids, (B) a single centroid representing the retail centre.

In the absence of further and costly commercial footfall estimates a method has been devised to assign a population capacity to each centroid using publicly available data. It must be noted that this approach is just one option that only attempts to estimate retail consumer footfall. It is acknowledged that the commercial datasets mentioned could provide closer approximations but due to the required costs are not available for this demonstrative case study. The approach uses average sales density, income per unit floor area (e.g. £/sq. ft), divided by average transaction values per person (Eq. 5.3). This first portion of this equation estimates the customer footfall generated per unit area of retail floor space. The second part of Eq. 5.3 adjusts this accordingly depending on the size, in terms of the total floor space (square feet) of the particular retail centroid. Therefore, larger retail centroids support a greater footfall.

$$\text{Footfall} \approx \frac{Sales\ density\ (£\ per\ sq.ft)}{Average\ transaction\ value\ (£)} \times Centroid\ floorspace\ (sq.ft) \qquad (5.3)$$

Sales density, and sometimes average transaction value, is often reported by major retailers in their annual financial report for shareholders (e.g. John Lewis Partnership 2010; Morrisons 2013; Sainsbury's 2014; Tesco 2014). A further report by the GLA (2005) used commercial data and convenience store databases representing 11,000 retailers in and around London to provide a comprehensive sales density average. These sales density values were used to produce a comprehensive average for grocery and comparison shopping. Equation 5.3 was applied individually to the comparison and grocery floor space for all retail centroids, adjusting the average transaction value accordingly. The footfall generated for the grocery and comparison retail space for each centroid was summed to give a total estimate. This distinction was made due to variations in average transaction values for grocery and comparison spend which would have affected footfall estimates. Values have been adjusted to account for inflation up to 2010 (target year). The average transaction for value for food and non-food was derived from the Living Costs and Food Survey (ONS 2012c).

Like visits to leisure attractions (Section 5.4.3) there is not an even ratio of the usually resident population to overnight visitor population who shop at these retail locations. This is because of the study area's unique characteristics influenced by the seasonal change in overnight visitors due to tourism which are important considerations for spatiotemporal modelling. Analysis of retail transactions based on loyalty card data for the same region by Newing *et al.* (2013b) shows large seasonal variations in footfall and the resident to non-resident sales ratio. Newing *et al.*'s (2013b) analysis of weekly loyalty card transactions indicates the number of sales within a store's catchment (local residents) or external (visiting population). This varies greatly by month and can be used to allocate the correct portion of residents and non-residents from the respective origin files. The ratio used for the three months chosen for this case study is summarised in Table 5.8.

Table 5.8 Ratio of residents to visitors by month for study area retail sites. Data after Newing *et al.* (2013b)

Month	Residents	Overnight visitors
January	85%	15%
May	75%	25%
August	50%	50%

These data are used to create copies of each retail destination dataset with the footfall divided in the ratio identified in Table 5.8, for the season concerned, between visitors and usual residents. This allows the correct population origin category to populate each site (represented by the retail destination links in Figure 5.7).

Newing *et al.* (2013b) make specific mention of the St Austell Tesco supermarket highlighting that it is a unique store with highly seasonal demand, particularly from summer visitors on holiday nearby and visiting the Eden Project. To alleviate the shortage of space and overtrading (when a larger store would not be sustainable during quieter winter months) they describe that such stores locate a temporary sales marquee in the car park during the summer peak. This is illustrated in the image of Tesco's St Austell store (Figure 5.11). This additional space addresses the seasonal fluctuations in the footfall experienced.

Figure 5.11 A temporary marquee in the store car park at Tesco, St Austell. (Imagery: Google Maps, June 2014)

In conclusion there are a number of improvements since the creation of a retail dataset for the Southampton worked example (Chapter 3, Smith *et al.* 2014a). The location of retail destination centroids has been improved and made more computationally efficient due to the availability of the GMAP (2014) dataset for this case study. This is in contrast to using business postcodes as a retail location proxy in the Southampton example. The GMAP retail centres ensure greater spatial accuracy for retail locations. Enhanced footfall estimates have been derived based on the actual retail floor space data and averages for sales density. Analysis of industry retail data by Newing *et al.* (2013b) relevant to this study area can directly inform seasonality within retail destination footfall, which is a unique temporal consideration for this study area. It is acknowledged that this model could be improved further with the application of actual sales or footfall data but to date it is not financially viable to access these data for the area required. The example presented here provides a hybrid commercial/public comprise suitable for demonstration of the unique spatiotemporal concepts within this thesis.

5.4.5 Healthcare destinations

Another case study specific enhancement for the spatiotemporal population modelling approach is the development of a healthcare destination dataset. An initial review of healthcare statistics indicates that only small scale providers (e.g. community hospital and treatment centres) operate within the study area. Nevertheless, for the interests of completeness a healthcare dataset has been constructed. Healthcare provision has important spatiotemporal characteristics. For example, a large hospital will have admitted immobile populations who are present on site for substantial or long-term periods, while emergency and outpatients will be present for much shorter timescales. When considering spatiotemporal variations in populations these are important destinations that can often have highly concentrated temporary populations occupying the site. This subsection addresses the creation of a healthcare destination dataset which is a component of this case study's population data library (Figure 5.7).

In the UK a list of public healthcare providers by institution is available, organised within primary care trusts (PCTs) during 2010. A PCT may contain just one institution such as a major accident and emergency or university hospital, or multiple smaller community hospitals and providers. Annual Hospital Episode Statistics (HES) (HSCIC 2010) provide the number of consultation episodes and accident and emergency admissions by PCT. Where a PCT contains more than one care provider the annual hospital episodes have been allocated weighting them by the number of bed spaces (as an indication of size). Publicly available sources of the number of bed spaces within a provider vary widely and a combination of data sources need to be used. These range from data contained within existing Freedom of Information Requests, PCT websites and hospital review sites. Reweighting these data by bed spaces means that providers with the greatest number of bed spaces get the greatest proportional allocation of annual patient episodes accordingly. The annual allocation is divided by 365 to get an average daily estimate. HES data at a daily resolution is currently not available.

Within the HES data basic patient demographic information is recorded and the data on the age of patients are used to inform the population subgroup split by age for the modelling undertaken. Average arrival by hour at accident and emergency (A&E) departments and minor injury units (MIUs) is available at provider level. This has been analysed for providers within the study area to construct temporal profiles for attendance (Figure 5.12). MIUs provide care for less serious injuries that do not need immediate emergency treatment and have restricted opening hours that can be observed within the temporal profile created (Figure 5.12). They are typically open between 08:00-22:00, compared to the 24 hour service at A&E departments.

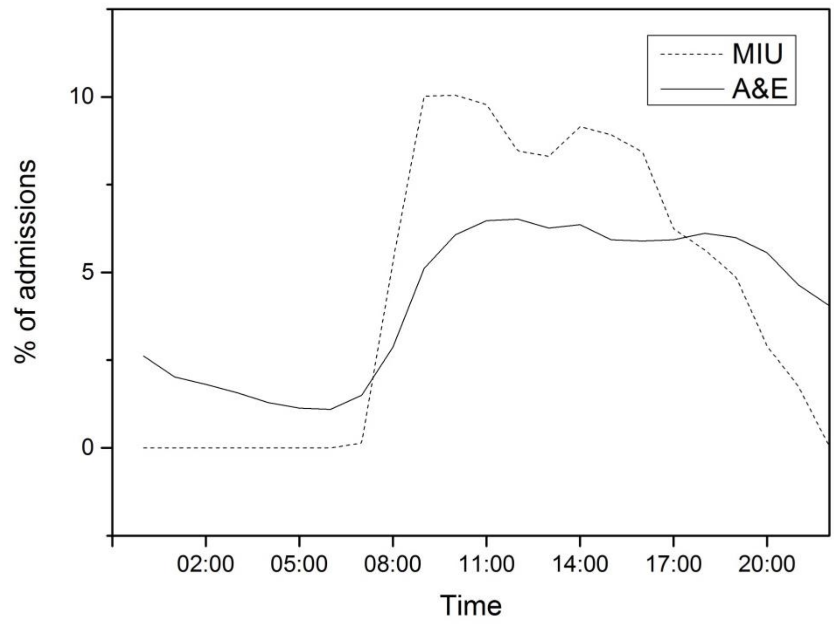


Figure 5.12 Accident and emergency (A&E) and minor injury unit (MIU) temporal profiles for the St Austell study area. Data: HES (HSCIC 2010)

Finally, the healthcare destination's catchment has been constructed using the data reported by Roberts *et al.* (2014). They analysed HES data reporting a patient's home location (by LSOA) and the postcode of the hospital where treatment was received to derive distances travelled. These HES data are not routinely published and require an application for a bespoke extract, therefore the information within the report by Roberts *et al.* (2014) has been used for this case study. There is a cost associated with the extraction of HES data which also require access approval. Their approach could be followed for the construction of a population library for a stakeholder however sufficient detail has been maintained for the purpose of a demonstration within this case study.

5.4.6 **Weekday versus weekend workplace capacities**

The final enhancement created for this case study is the estimation of the change in workplace capacities for a typical 'weekday' versus a weekend. In the previous examples only a daily cycle has been analysed. So far this chapter has discussed seasonal variation within the St Austell study area. However, there is also a weekly change that can be accounted for within the model's framework. Workplace data have been constructed using ABI/BRES (2010) labour force statistics following the process outlined in Chapter 3 (see Table 5.4 for breakdown summary). However, it is known that not everyone works routine set hours seven days a week, whereas, some employees will work shifts including, or exclusively, weekends. Analysis of the Labour Force Survey (LFS) for the broad SICs modelled has been undertaken for the usual days worked by each sector. Respondents are asked whether they usually work on the day in questions for all days within the week. This has been summarised in Table 5.9. Separate workplace destination datasets have been created for weekdays and weekends and have been allocated the correct proportion of the workforce accordingly. This is based on the percentage distribution by day worked of all respondents within each SIC category. These have been aggregated to represent weekdays (Monday to Friday) and weekends (Saturday and Sunday). This gives the proportion of the total number of employees within each SIC who would be expected to be working on any given weekday or weekend. For example the proportion calculated to be present and working on weekends will include employees who stated they exclusively worked weekends within the LFS or whose usual working days also included a weekend day.

Table 5.9 Employees by SIC for the St Austell study area and the breakdown for those who usually work a weekday or weekend. Data: LFS 2010

Workforce	**Percentage distribution of employees (%)**	**Work weekdays (%)**	**Work weekends (%)**
Office based	23	94	6
Retail	20	84	16
Healthcare	14	93	7
Accommodation and catering	14	82	18
Manufacturing and construction	13	94	6
Education	10	93	7
Transport	4	90	10
Agriculture and fishing	2	79	21
Total	*100*		

Subsequent modelling has been carried out that for weekdays and weekends based on the distribution presented in Table 5.9.

5.5 Results

This section provides the results of analysis from differing flood risks associated with a seasonally changing population for the St Austell study area. It has already been noted that results at this spatiotemporal scale are not achievable using static population or hazard data alone. In contrast to the two previous examples (Ulley and Southampton) the discussion in this chapter has highlighted that this case study demonstrates some unique characteristics in space and time. Its location is geographically constrained by the coastline while the area is largely rural. However, the study area notably experiences a large flux of visitors on a seasonal scale. For this reason (masking effect), and in part due to the rural nature, daily commuter flows do not dominate this example. Instead different temporal factors on a larger seasonal scale are the primary influences on this region.

5.5.1 SurfaceBuilder247 data library construction

The distribution of origin and destination centroids, constructed for the SurfaceBuilder247 data library, within the St Austell study area is shown in Figure 5.13. The origin centroids are comprised of residential UPCs and show clusters of residential locations (e.g. along the southern coast and Bodmin) interspersed with rural settlements. The origin centroids are populated by the usually resident and overnight visitor population. In contrast the destination centroids are more tightly clustered to the main residential concentrations. This shows the distribution of leisure locations (attractions), workplaces (business UPCs), education establishments, hospitals and retail centres. These all receive temporary, and usually daytime, population counts such as employees at work and children at school from the surrounding origin centroids. Origin and destination sites are each referenced by a single point (e.g. UPC) but population is dispersed into the immediate surrounding area governed by a local dispersion parameter that reflects the size of the site. This size of the site is estimated based on its type or referenced from known sources (e.g. retail floor space).

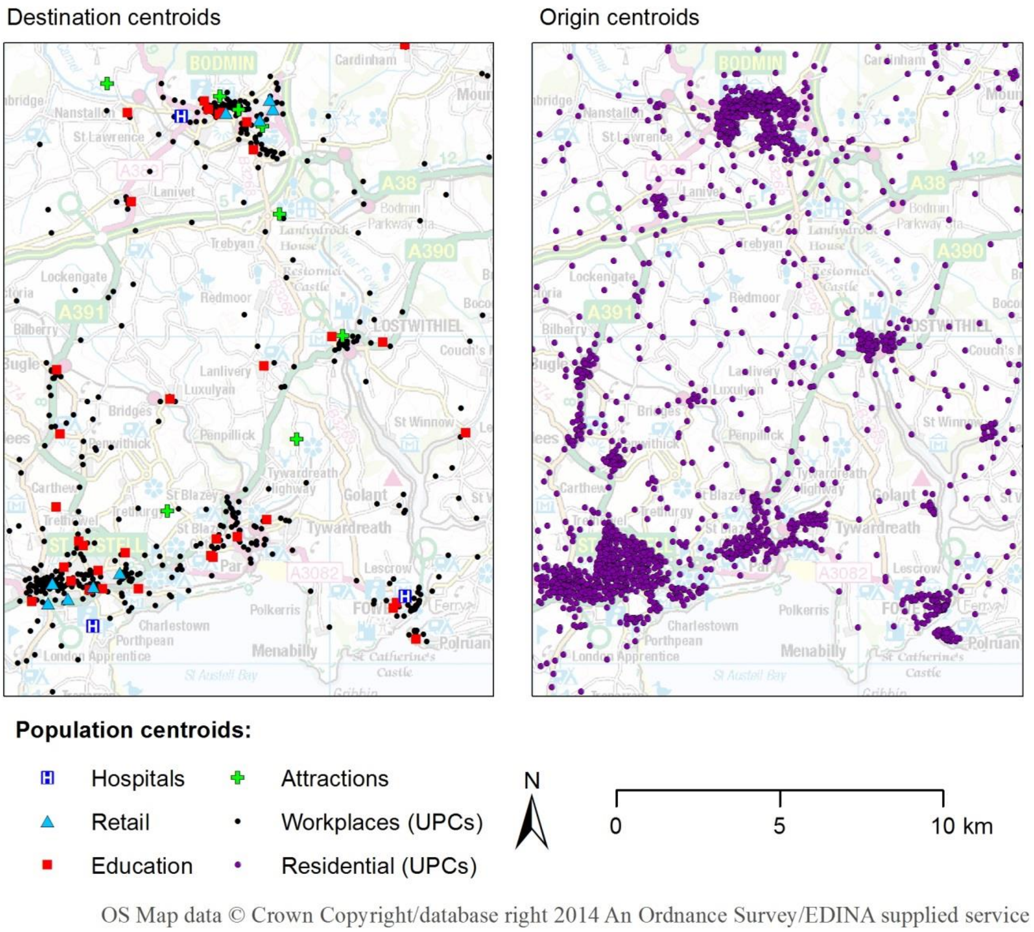


Figure 5.13 Origin and destination centroids within the St Austell study area

5.5.2 Spatiotemporal population distribution

Hourly spatiotemporal population estimates have been produced for the study area representing a 'typical' working weekday within each month for three seasonal scenarios (January – low, May – fringe and August – peak). January and May represent school and university term-time and therefore use the same term-time usually resident population origin base (Figure 5.7). August is a school and university summer vacation period in England and Wales. Therefore it is represented by a non-term time population origin base (Figure 5.7). The population exposure to the EA's flood map zone three has been calculated for the whole study area (Figure 5.4) and is presented in Figure 5.14. It has been compared with static exposure estimates from rasterised census outputs representing: the baseline 2001 Census population at OA level (highest resolution available), 2001 Census daytime population at OA (only released for 2001) and the 2010 mid-year estimate (closest to target date but only available at LSOA level).

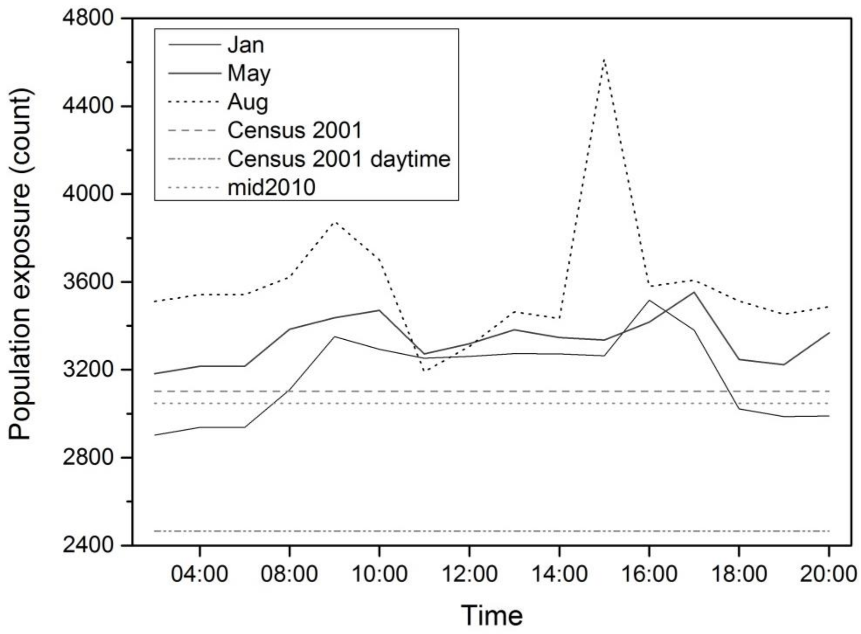


Figure 5.14 Flood exposure estimates from the EA Flood Map Zone 3 for the St Austell study area using seasonal spatiotemporal model outputs at hourly intervals for a 'typical' weekday.

The St Austell case study differs in some important respects from the previously reported Ulley and Southampton examples. Firstly, the census estimates underestimate exposure in this case study compared to large overestimations in the previous examples. This is likely to be for a number of reasons: the modelled outputs have accounted for seasonal visitor influxes in addition to the census population. On average the visitors increase the usually resident population by +6,570 in May and by greater than 12,000 at the peak in August (Table 5.5). The rural characteristics of the study area result in geographically larger OAs and LSOAs (to meet minimum population confidentiality levels), producing lower population densities compared to an urban region. In contrast smaller OAs (by area) in urban locations produce higher population densities although all OAs contain roughly similar numbers of households. Secondly, there are no very distinct daily cycles in the modelled outputs for St Austell. It is possible to identify what appears to be a familiar daily cycle in the January output characterised by a morning and evening travelling peak (around 08:00 and 16:00, Figure 5.14) which is also evident in the previous examples. However, as larger numbers of overnight visitors are accounted for in the May and August examples the daily cycle in the baseline residential population (closely attuned to the January example with few visitors) appears to be masked by their movements. There is a clear seasonal cycle in overall exposure although the daily signal appears to become more variable. However, it is the temporal variation at a seasonal scale that is the main feature of this particular case study.

A spatial comparison of the rasterised (and density adjusted) census datasets and SurfaceBuilder247 model output is illustrated in Figure 5.15. The 2001 Census (OA) is the highest resolution data available prior to the target date (Figure 5.15A). A 2001 daytime census population at OA level (Figure 5.15B) is also available. This is defined as the count of people aged 16-74 who do not work plus those who work in the area (National Statistics 2004). The population estimate closest to the target year (but at lower resolution) is the 2010 mid-year estimate at LSOA level (Figure 5.15C). These are all static with no seasonal variation or areas of zero population density due to the census structure of contiguous zones. In contrast, the model output for a May weekday ‘night-time’ population illustrates the enhancement in spatial resolution (Figure 5.15D). Population is concentrated in inhabited areas with surrounding unoccupied farm and moorland receiving a zero population density. The area is characterised by dispersed groups of rural dwellings which receive a low but non-zero population producing a speckled effect in the modelled results which corresponds with the population origin centroids (Figure 5.13). This is another contrast to modelled outputs for the Ulley and Southampton case studies where the population had previously been concentrated in large urban zones, rather than the dispersed rural settlement pattern observed here.

The seasonal spatiotemporal variation has been illustrated for a weekday day (12:00) and night-time population (00:00) estimate for each of the three seasons modelled. The usually resident and overnight visitors have been displayed separately. These selected spatiotemporal outputs are displayed in Figure 5.16 (January – low), Figure 5.17 (May – fringe) and Figure 5.18 (August – peak). In all three examples a general concentration in the usually resident day-time (12:00) population occurs from the night-time (00:00) locations (Figures 5.16-18A and B). There is increased clustering at the main population centres in the day-time examples (St Austell to the south and Bodmin in the north) and a greater population in travel on the road network. The concentration is observed in the main population centres, analogous with the main workplace locations showing a rural to urban commute from the surrounding areas. For example, for the May usually resident population there is a peak in concentration in the St Austell town centre at 12:00 (Figure 5.17B) of 1,400 people/100 m^2, compared to just 54 people/100 m^2 at 00:00. This is presented at the group of orange coloured cells in the south-western corner of the study area in Figure 5.17B. It also highlights the known phenomenon that town centres are predominantly only populated during the daytime as they host a range of retail, leisure and workplace locations but at night have very few usually resident people.

Similarly, there is a concentration in the overnight visitor population from the night time locations they occupy to concentrated locations of daytime activity (Figures 5.16-18C and D). These daytime concentrations, most notable in August with the visitor peak, occur in the main town centres (e.g. St Austell, Bodmin and Lostwithiel). There is a large increase in the overnight visitor population of greater than 12,000 people between January and August (Table 5.5). Another clear observation is that the distribution as well as concentration and number of estimated overnight visitors increase between January and August. Most notable is the August night-time concentration of visitors in the coastal areas south of St Austell (Figure 5.18C). Secondly, the central area of the study area's extent receives a greater share of overnight visitors. This is attributed to the location of rural guesthouses, campsites and caravan parks which are not populated with the additional population in traditional census datasets.

An increase in the concentration of the daytime (12:00) usually resident population is observed between January and May (Figure 5.16B and 5.17B) although they come from the same term-time population origin base. This change can be explained by seasonal variation in different sets of destination sites for the two months. Footfall at leisure attractions and retail locations increases between January and May and this is attributed to the change observed.

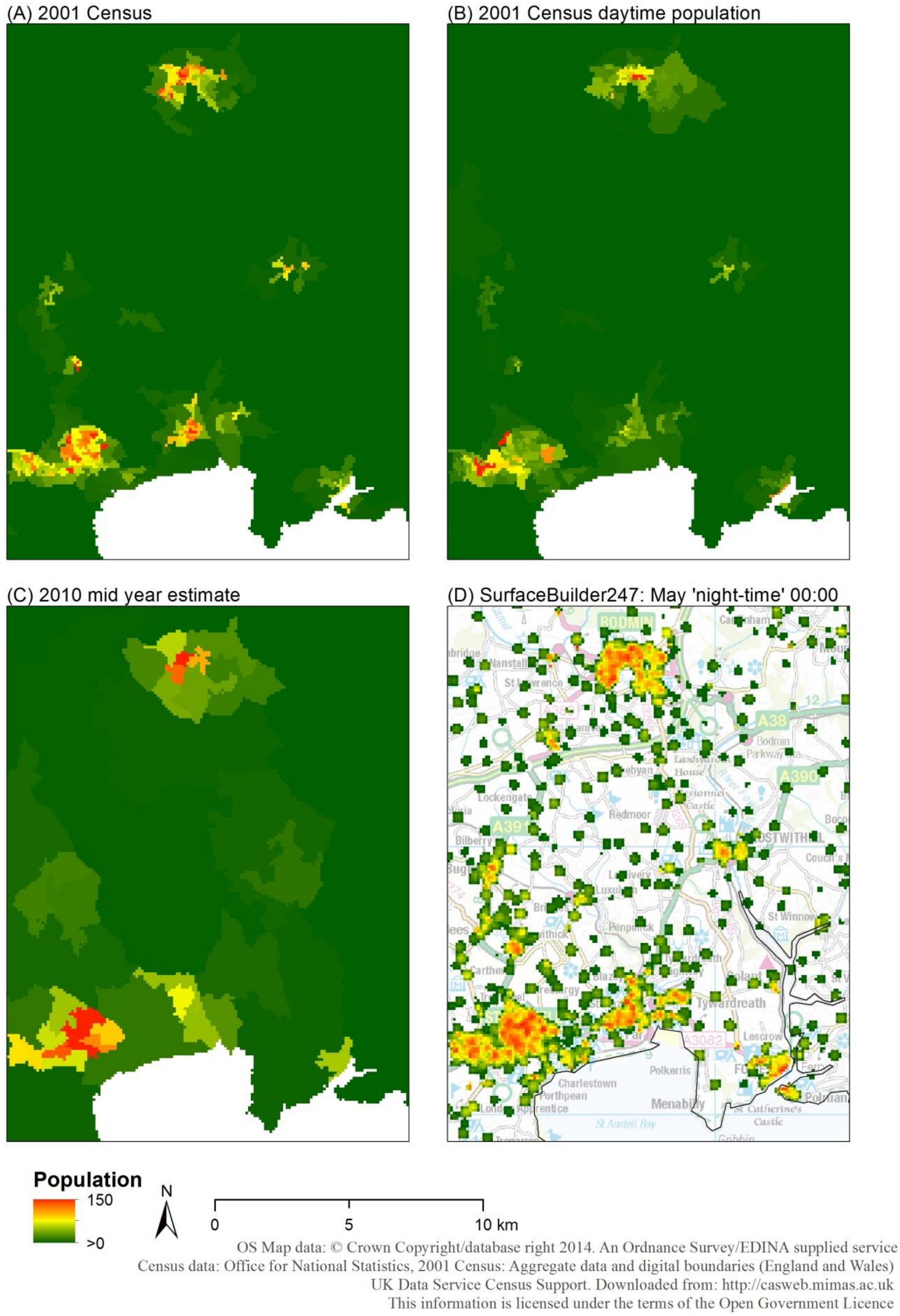


Figure 5.15 Comparison of rasterised (100 m) census datasets for the St Austell study area for (A) 2001 OA counts, (B) 2001 OA daytime counts, (C) mid-2010 LSOA counts and (D) example model results for May 00:00.

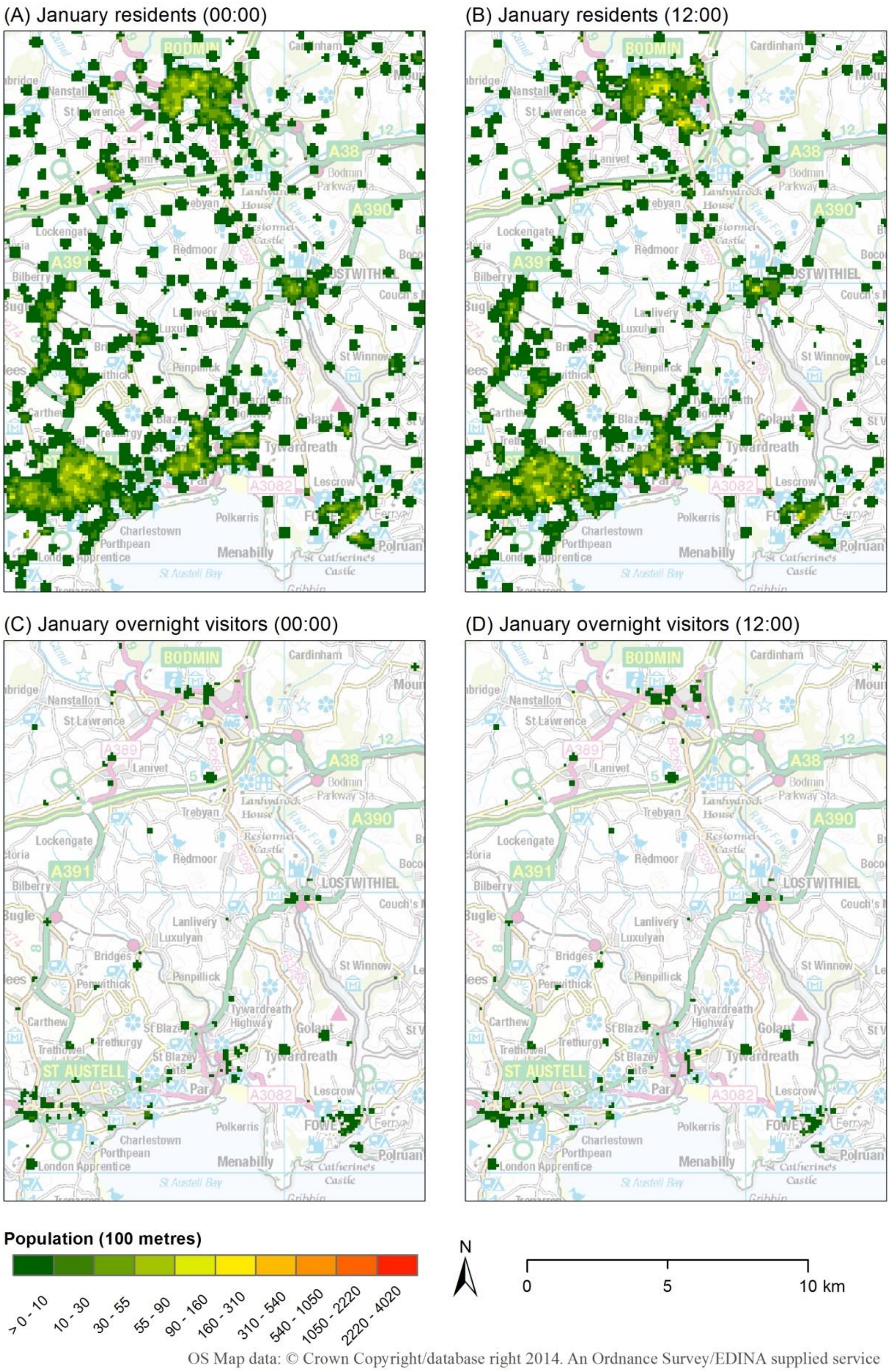


Figure 5.16 Modelled seasonal population outputs (100 m) for the St Austell study area for a January weekday. (A) Usually resident night-time (00:00) population, (B) Usually resident daytime (12:00) population, (C) Overnight visitor night-time (00:00) population and (D) Overnight visitor daytime (12:00) population.

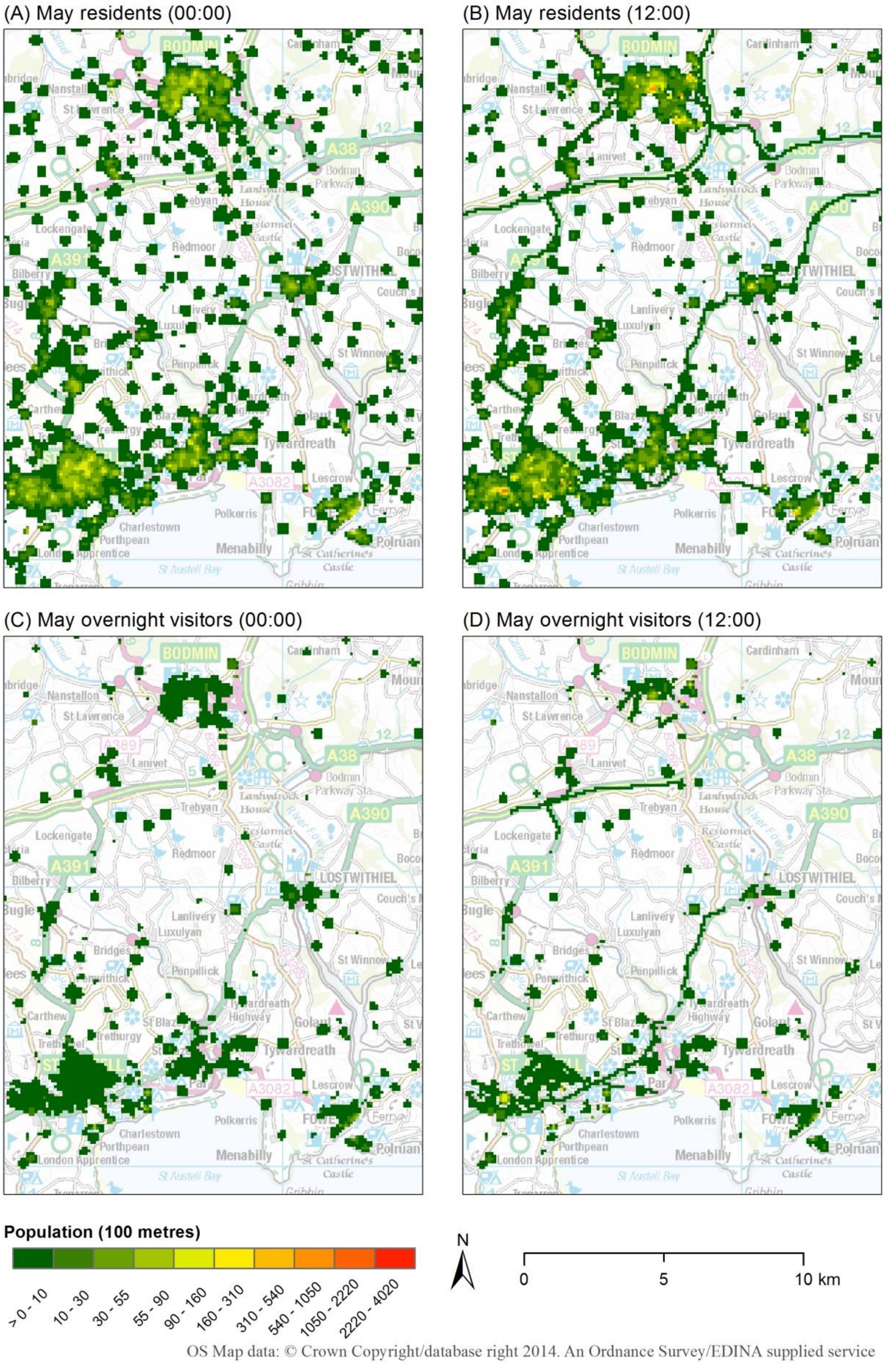


Figure 5.17 Modelled seasonal population outputs (100 m) for the St Austell study area for a May weekday. (A) Usually resident night-time (00:00) population, (B) Usually resident daytime (12:00) population, (C) Overnight visitor night-time (00:00) population and (D) Overnight visitor daytime (12:00) population.

Figure 5.18 Modelled seasonal population outputs (100 m) for the St Austell study area for an August weekday. (A) Usually resident night-time (00:00) population, (B) Usually resident daytime (12:00) population, (C) Overnight visitor night-time (00:00) population and (D) Overnight visitor daytime (12:00) population.

To further exemplify the spatiotemporal population outputs achieved using SurfaceBuilder247 comparative extracts for a week day at 10:00 and a weekend at 12:00 are shown in Figure 5.19. These represent the total population (i.e. visitors and usual residents combined). Population has been modelled at hourly intervals for three seasonal scenarios (e.g. Figure 5.14). Until now only representative examples for a weekday daytime (12:00) and night-time (00:00) population have been spatially illustrated. The comparison of the population at 10:00 for a weekday in January and August (Figure 5.19A and B) shows greater clustering and in-travel populations in the August example. A notable difference, in terms of the in travel population, can be observed between the January weekday 10:00 scenario (Figure 5.19A) and the January weekday 12:00 scenario (Figure 5.16B and D). In this example there is a greater in travel population at 12:00 compared to 10:00, a change observed in just a two hour period.

A similar feature of the in travel population is also observed in the total population for a May weekday at 12:00 (Figure 5.19C) and May weekend at 12:00 (Figure 5.19D). This represents the same total population scenario but for different days (weekday vs. weekend). The weekday example has a considerably greater in travel population for the same time (12:00) on a weekday compared to a weekend. However, the weekend population is more concentrated, corresponding with retail centre locations (Figure 5.13).

If all scenarios were to be illustrated concerning the two population classifications (visitors and residents), six seasonal scenarios (January, May and August for a weekday and weekend) and three LISFLOOD-FP extents over 24 hours, there would be 864 possible unique combinations ($2 \times 6 \times 3 \times 24$) in this example alone. The temporal aspect of the modelling undertaken allows much greater insights into possible population cycles. If sufficient data were available to further increase the temporal resolution to 15 minute intervals ($24 \times 4 = 96$) and consider weekdays individually (3 seasons $\times$ 7 days = 21) the number of possible combinations would increase to 12,096 ($2 \times 21 \times 3 \times 96$). Introducing temporal granularity within the environmental modelling would increase this even further. In addition to this the seven population subgroups for the two population origin classes chosen could also be considered separately.

In contrast the census output, even with a sophisticated static areal interpolation, with the three flood scenarios would still only give six possible combinations (2 × 3), using the census residential and daytime population counts. However, the difference in many of these 864 possible combinations is likely to be negligible and therefore it is important to choose appropriate and contrasting snapshots for analysis and static visualisation. The SurfaceBuilder247 approach adopted facilitates detailed evaluations for population exposure to flood risk while considering changes in season and time of day. In any final assessment there is the potential for large variations in the outcome depending on the combination of events chosen, as exemplified in the St Austell application illustrated in this chapter.

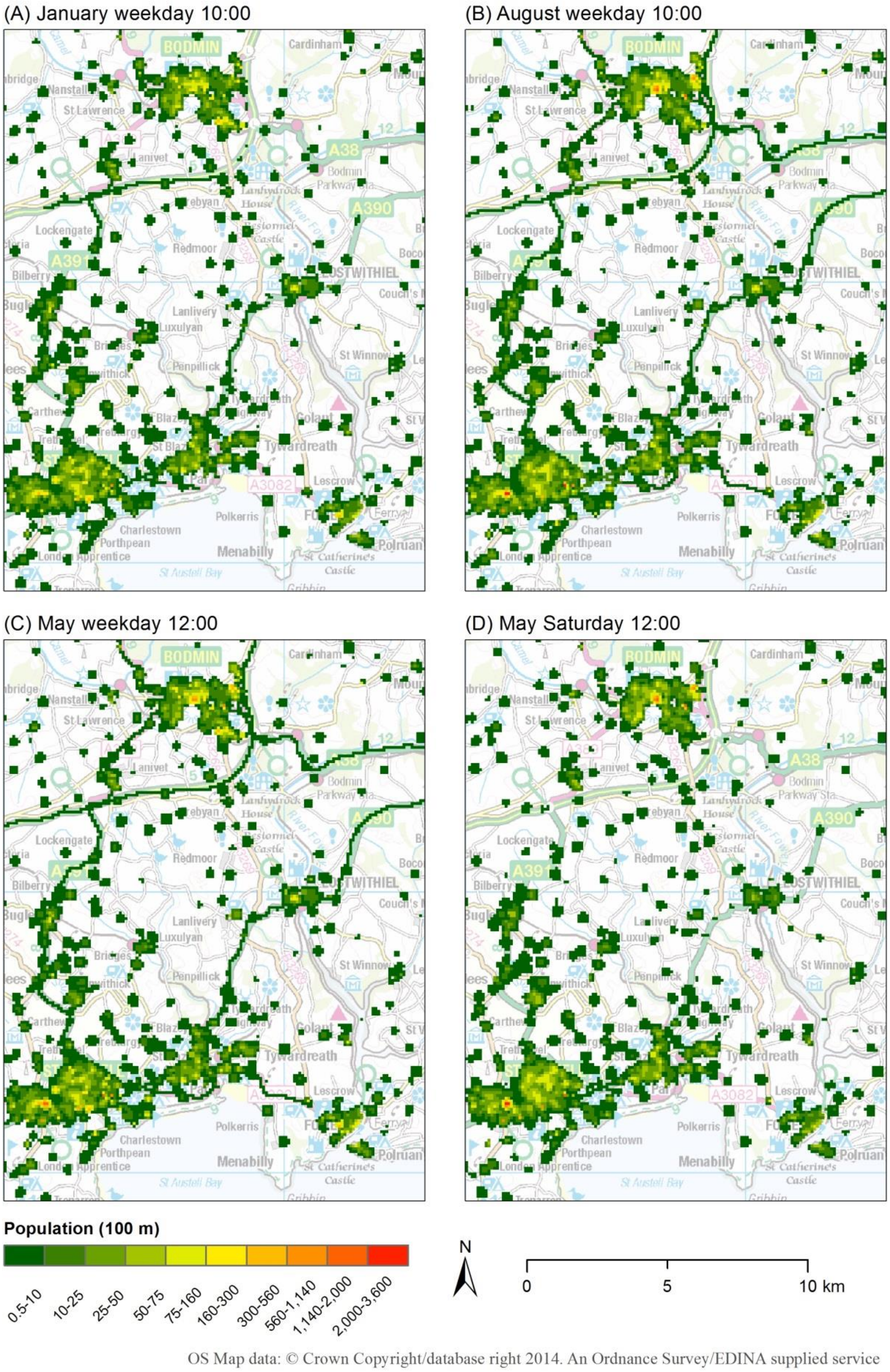


Figure 5.19 Population comparisons of (A) January (term-time, low season) weekday 10:00 and (B) August (non-term time, peak season) weekday at 10:00; (C) May weekday at 12:00 and (D) May weekend 12:00. All represent total populations (residents and visitors).

A detailed comparison has been conducted using selected 1 km national grid square extracts from the St Austell study area model results (Figure 5.20). Two have been selected from August daytime (A and B) scenario for the total population and two for an August night-time (C and D). All represent the total population (usually resident and visitors combined). The distribution of 100 m output cells is clearly visible within the detailed 1 km square extracts. The modelled results have been compared to Ordnance Survey base mapping and aerial imagery for the same location and scale.

The OS map extract for the first example (Figure 5.20A) shows a part of Fowey, with a range of tourist attractions (indicated by the blue map symbology). The August daytime model results for the same area show that the population is appropriately constrained to the land mass (due to the background masking layer) and concentrated on the coastal locations of the amenities outlined on the map extract.

The second extract is focused on one of the highest concentrations within the study area (Figure 5.20B), showing part of St Austell town centre. The August daytime population concentration exceeds 1000 people per 100 m^2. The Holmbush area is a retail district which includes the St Austell Tesco supermarket (Figure 5.11). St Austell has the highest floor space in terms of retail within the study area which is informed by the retail destination datasets that have been created for this case study. Comparison with the aerial imagery shows close model alignment with the populated areas.

The third extract is the first of two August night-time examples. Figure 5.20C shows the location of a large static caravan site immediately behind Par Beach. Population densities within the model cells correspond with the caravan site, summing to approximately 150 people. The aerial imagery provides the detail which is just shown as a series of tracks on the OS background mapping (as caravans are not permanent structures and therefore not mapped). This area also corresponds to high levels of flood risk under all of the inundation scenarios (Figure 5.5).

Finally the fourth extract (Figure 5.20D) shows the night-time population estimate for what appears to be an uninhabited area, but which is clearly designated as a campsite in the OS mapping and discernible within aerial imagery. The population density corresponding to the campsite area shows moderate population densities of up to 50 people per 100 m^2. The small settlement of Lower Penhale is represented by an area of low non-zero population densities. This appears to be a slight overspread, but still demonstrates a refinement based on the census zonal data alone. Furthermore it would not be possible to resolve the August peak in population at this campsite (which is simply an unoccupied field at other times of the year) relying on the census data alone. This example tests the limits of the current spatial resolution of the model using currently available population data for this case study; however they are still significant improvements. Reasons for this overspread are likely to be caused by the underling population origin centroid density. As residential postcodes were used the rural locations identified on the map, Lower Penhale and Polgassick Farm (Figure 5.20D) are likely to share a postcode which may not be georeferenced directly on one particular site. The dispersed nature of rural properties sharing a rural postcode is greater than in concentrated streets in urban areas. This is another important factor to consider the application of models using only postcode centroids where spatial accuracy and density can vary.

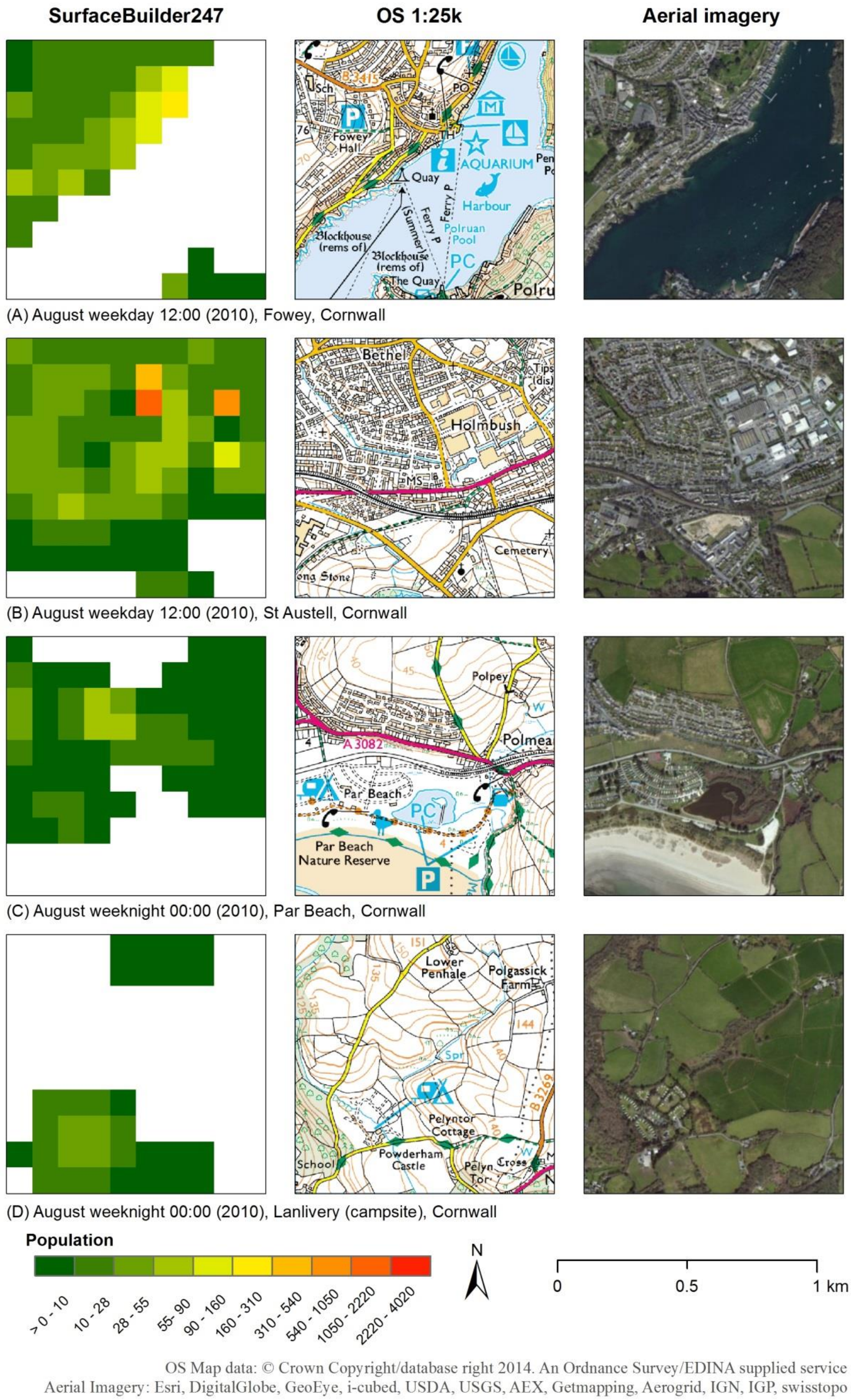


Figure 5.20 A detailed comparison of SurfaceBuilder247 (100 m resolution) results within the St Austell study area with 1:25000 scale Ordnance Survey (OS) background mapping and aerial imagery for selected 1 km national grid squares. (A) and (B): August weekday ‘daytime’ population. (C) and (D): August weekday ‘night-time’ population.

Figure 5.21 (A) Par beach looking southwest from the low level sand dunes (B) protecting the caravan site behind from coastal inundation. Photographs: Alan Smith (May 2014)

A large caravan site is situated immediately behind Par Beach (Figure 5.20C and 5.21). This location is only protected by small natural sand dunes and all flood inundation scenarios used in this case study identify this site as high risk (Figure 5.5). McEwen *et al.* (2002) highlight the susceptibility of caravan sites often located in flood prone locations (usually because permanent structures may not be permitted) and note the vulnerability of residents and their associated high exposure to flood risks (such as this Par example). They recommend that there is strong evidence to treat such residents as a distinct vulnerable group with specific requirements for the purpose of flood warning and emergency planning.

5.5.3 Population exposure to flood risk

An assessment of the population's exposure to flood risk has been carried out for the sub-area of the study area identified in Figure 5.4 using the EA and LISFLOOD-FP inundation scenarios (Figure 5.5). This has been undertaken for daytime population estimates (12:00) for the three illustrative seasons selected. Within each season the visitor and usually resident population has been analysed separately (Table 5.10).

Table 5.10 Daytime usually resident and visitor population exposure to three LISFLOOD-FP inundation scenarios (R = return period) and EA flood map zone three for January, May and August (increasing levels of inundation left to right).

Population	**LISFLOOD R100**	**LISFLOOD R250**	**LISFLOOD R500**	**EA Flood Map**
Residents 12:00 Jan	542	939	1069	1725
Visitors 12:00 Jan	2	5	7	15
Total	*544*	*944*	*1076*	*1740*
Residents 12:00 May	546	994	1139	1729
Visitors 12:00 May	34	108	131	114
Total	*580*	*1102*	*1270*	*1843*
Residents 12:00 Aug	498	1019	1178	1741
Visitors 12:00 Aug	65	206	249	212
Total	*563*	*1225*	*1427*	*1953*

An interesting phenomenon observed in the seasonal flood map analysis (Table 5.10) is actually a decrease in the August 12:00 exposure to the LISFLOOD-FP R100 flood risk, compared to the May 12:00 exposure total to the same LISFLOOD-FP R100 flood extent. Total population exposure for a weekday at 12:00 under the LISFLOOD-FP R100 scenario decreases from 580 to 563 from May to August. This is driven by the usually resident population. It is the reverse of the cycle observed in all of the other scenarios modelled where the August 12:00 residential population exposure increases relative to the respective January and May levels. The midday January and May exposure of the usual residents for the LISFLOOD-FP R100 scenario remain similar. This could be expected as they are derived from the same term time census population base. The August usually resident population base is different to account for non-term time changes but nonetheless this is still an increase in population (Table 5.6) so not a cause for the exposure decrease. The variability between the January and May usually resident population for the other scenarios modelled (progressively larger polygons, Figure 5.5) is likely to occur due to more seasonally varying destination locations. Visitor exposure has increased in line with expectations between January and August for all inundation scenarios, following the cycle in the tourism season. It was expected that the overall increase in visitor numbers would inevitably also lead to an increase in flood exposure by season.

To examine the unexpected decline (as all other examples increase with the seasonal cycle) in flood risk exposure for the August and May LISFLOOD-FP R100 scenario (Table 5.10) the population has been further analysed at the population subgroup level for seven age subgroups for both usual residents and visitors (Figure 5.22). It can be observed that the largest contribution in the decline in exposure to flood risk between May and August (at 12:00 for LISFLOOD-FP R100) is the 16-64 working aged population. Exposure in this group decreases from 389 to 281 between May and August.

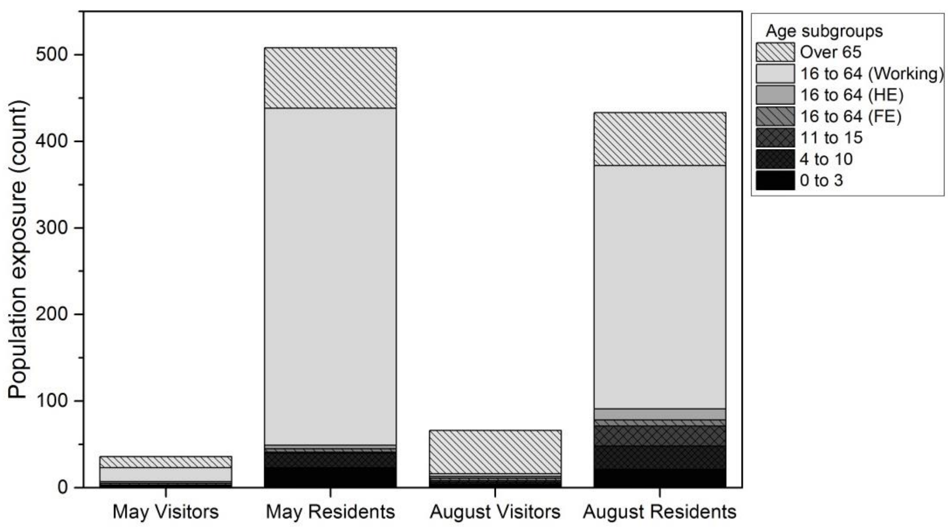


Figure 5.22 Comparison of daytime (12:00) population LISFLOOD-FP R100 exposure estimates broken down into age subgroups for visitors and residents in May and August

Figure 5.23 illustrates the movement in the working aged (16-64) population subgroup between May and August. It shows greater concentration within the urban centres of the study area. This corresponds to the locations of attractions and retail centres (Figure 5.13). Known seasonal variation in capacity at these sites means that a greater number of the 16-64 subgroup occupies these locations in August. Although this is the ‘working’ aged population it does not necessarily mean that they only populate workplace locations. This would suggest the greater August concentration of this population subgroup at inland locations is perhaps reducing exposure to flood risk following the movement of the population from the surrounding areas into relatively safer locations.

Although the total population exposure to flood risk in August for LISFLOOD-FP R100 decreases compared to May, the number of the elderly (>65 years) potentially exposed increases (Figure 5.22). This increase of 385% (May to August) is derived from the influx of overnight visitors. While overall it would appear that flood risk is lower, there is actually a large increase in the elderly population exposed to flooding in the R100 August weekday 12:00 scenario. This does not mean that overall elderly visitors dominate the whole study area’s August tourist population (also dominated by family holidays) but just the flood polygon analysed. This insight could not be achieved looking at the total population alone or without modelling exposure at population subgroup level.

Population subgroups from the August residential population base that actually increase in LISFLOOD-FP R100 12:00 exposure (although overall there is a net decrease) those containing the school aged population (population subgroups 4 – 10 and 11 – 15, Figure 5.22). The spatial distribution at 12:00 for May and August for one of these subgroups, 11 – 15 (secondary school aged) has also been illustrated in Figure 5.23C and D. It is clear that in May (midday term-time weekday) this subgroup is highly concentrated at school sites. During August, non-term time (i.e. school vacation period), this population is dispersed throughout the surrounding residential locations. In this example these school sites appear to be in places of relative safety according to the inundation scenarios modelled. However, the more dispersed non-term location has increased the chances of exposure to potential flood risk through greater spatial distribution, including within potentially hazardous zones.

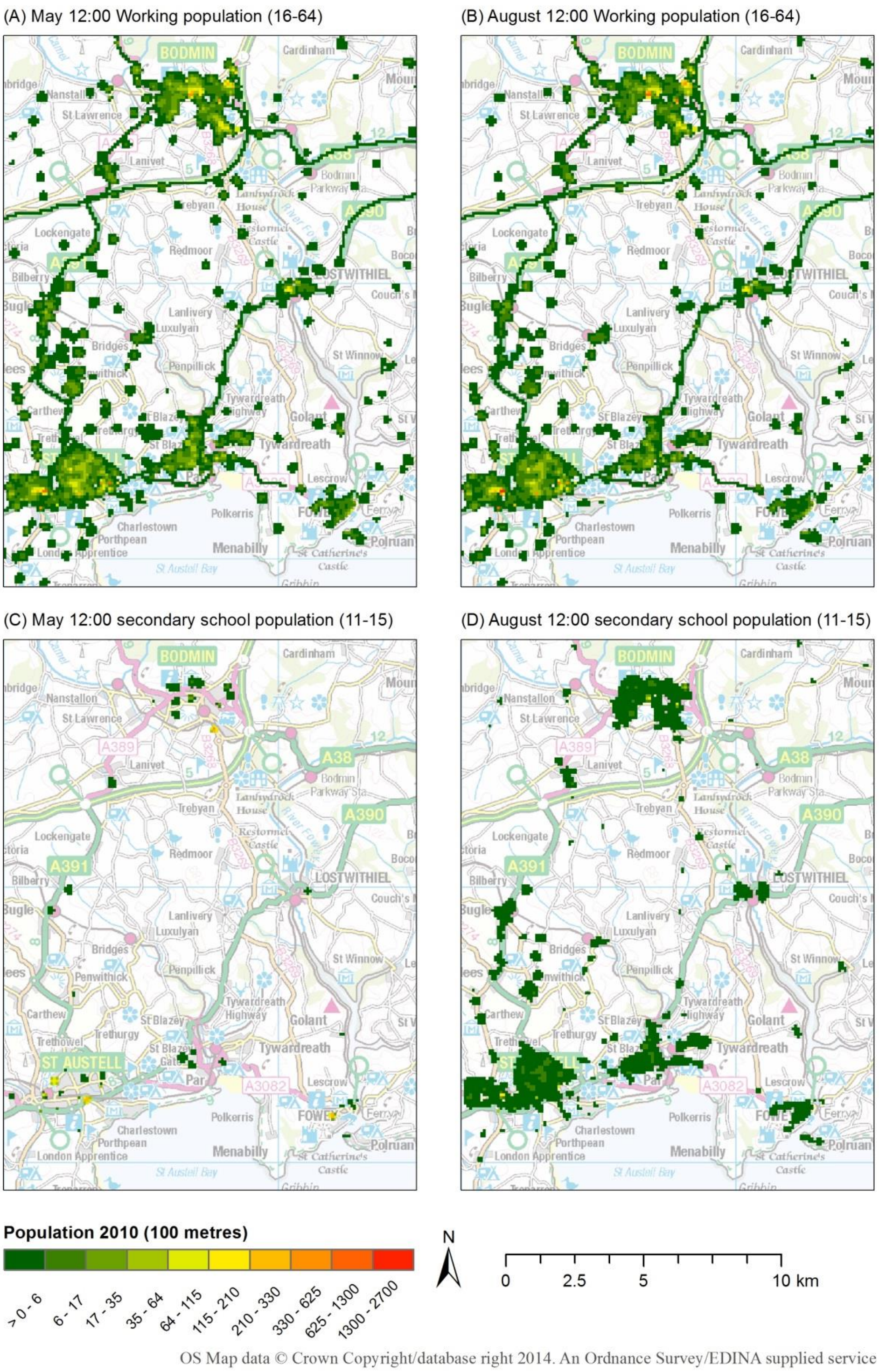


Figure 5.23 Spatial distribution of the usually resident 12:00 weekday population subgroups (Working aged 16 – 64 (A and B) and secondary school aged 11 – 15 (C and D)) for May and August within the St Austell study area.

5.5.4 Population fatality estimates

Population fatality estimates have been calculated using the spatiotemporal population distributions generated and the LISFLOOD-FP layers provided. This follows the calculation of a flood hazard rating and method introduced in Chapter 4.3.5 and summarised previously in Section 5.3.2. Fatality estimates have been calculated for the extent of the LISFLOOD-FP model (Figure 5.4) for a weekday at midday in January and August 2010 broken down into the usually resident and visitor population (Table 5.11). The methodological steps outlined to estimate the number of injuries, flood hazard rating (raster LISFLOOD-FP layer provided by Quinn 2014) and potential fatalities were combined with the spatiotemporal population outputs and evaluated using raster based calculations in ArcGIS. The months chosen intend to represent the difference between the low (January, term-time) and peak (August, non-term time) tourism season.

Table 5.11 Total population and fatality estimate for the LISFLOOD-FP model extent within the St Austell study area for January and August 2010

Weekday (12:00)	Total population	Fatality estimate for return period:		
		100 years	250 years	500 years
January Visitors	350	0.3	0.35	0.42
January Residents	28,887	61	79	100
August Visitors	4,945	13	16	17
August Residents	38,288	67	85	105

It is evident that there is a large increase in the visitor population from 350 to c. 5,000 between January and August within this subsection of the St Austell study area (Table 5.11). In August this 8 × 4 km subsection (Figure 5.4) accounts for almost half of the entire study area's visitor population (Table 5.5). Consequently exposure and risk of fatalities to the visitor population also increases from the low to peak season. In the worst case scenario presented (500 year return period) for this subsection the January to August influx of additional visitors accounts for an estimated increase of fatalities of up to 17 people. In contrast the relatively low visitor population in January (e.g. Figure 5.16D) is insignificant however the increased visitor population in August is of great concern. The spatiotemporal characteristics of seasonally varying populations outlined in this application demonstrate their importance for consideration in risk from hazards. The estimated number of all fatalities increases with event magnitude.

Another notable observation in Table 5.11 is the c. 9,400 person increase in the baseline usually resident population from January (term-time) to August (non-term time) for this study area subsection. This is generated by non-overnight visitors drawn from the neighbouring regions to the leisure destinations. It has already been determined that there is little change in the study area’s overall non-term time population (Table 5.6). However, it is possible the slight non-term time increase is concentred within this subsection (also the main population centre) and does also contributed towards the increase observed. The most likely explanation is the usually resident population’s increase in daytime visits to attractions and leisure locations within this study area. Like the overnight visitor population their leisure activity also follows seasonal cycles and concentrations. The non-term time summer vacation period in August is also peak season for the usually resident population’s leisure activities. The spatial distribution of fatality estimates for the 250 year scenario contained within Table 5.11 is shown in Figure 5.24. The January and August visitor fatality (Figure 5.24B and D) estimates are concentrated around Tywardreath Highway and adjacent to the River Par south of Par railway station. In addition, the fatality potential for the usually resident population (Figure 5.24A and C) also includes a greater area expanded around the River Par including St Blazey, as well adjacent to drainage channels in Holmbush, St Austell

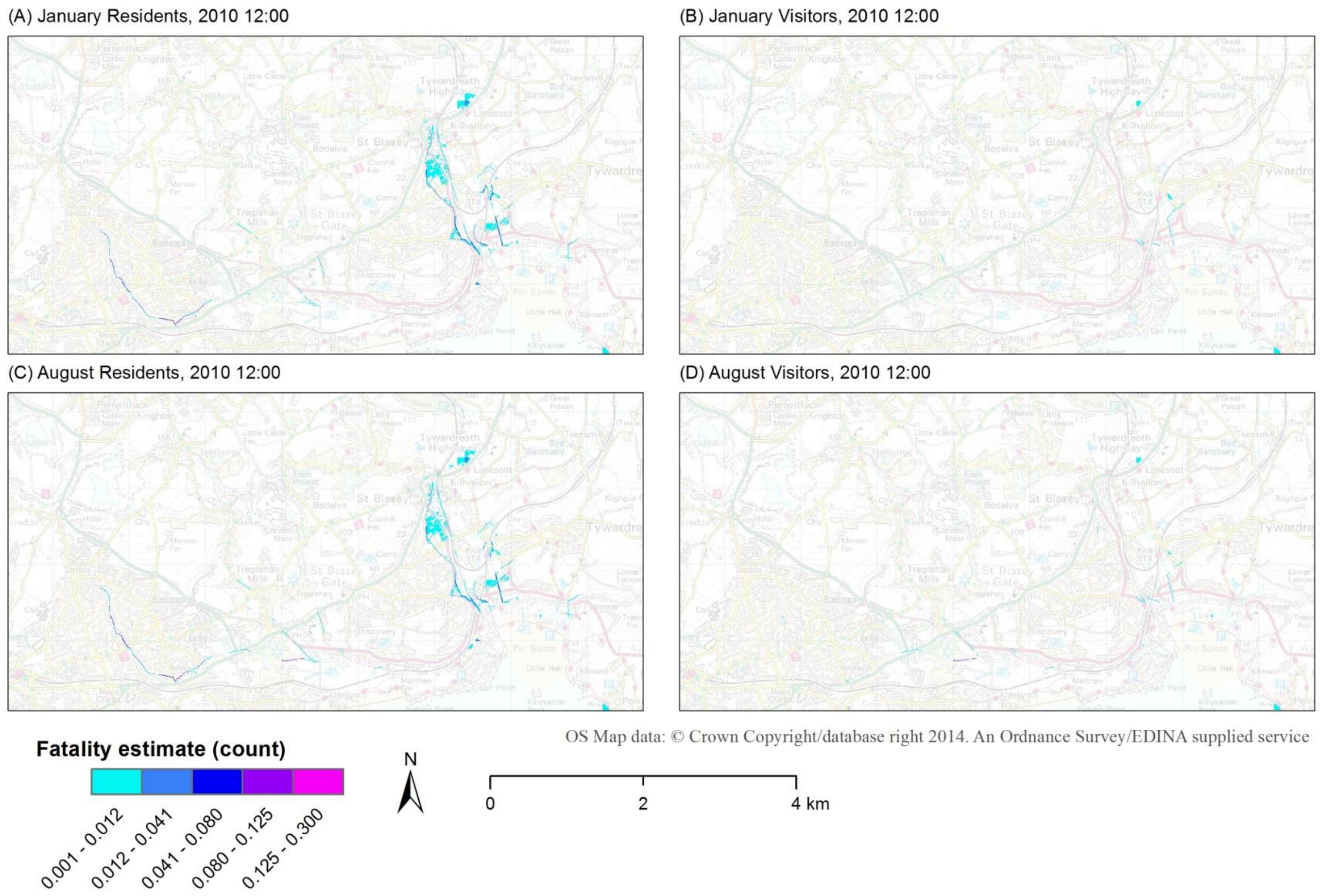


Figure 5.24 Population fatality estimates under the LISFLOOD-FP 1 in 250 year event scenario for the visitor and usually resident populations in January and August at 12:00.

5.6 Comparison with census estimates

A final analysis has been undertaken to compare modelled results for January with rasterised census datasets for the St Austell study area at a cellular level (100 m). The difference (Δ) has been calculated between the January night-time (00:00) model output and the 2001 Census (at OA level) (Figure 5.25A), January day-time (12:00) model output and the 2001 Census daytime count (at OA level) (Figure 5.25B) and the January night-time (00:00) model output and the 2010 mid-year population estimate (at LSOA level) (Figure 5.25C). All graphs show that overall the positive difference is greater (i.e. the model output produces greater values than the corresponding rasterised census cells). This can be visually explained through reviewing the spatial distribution of population in the modelled results (e.g. Figures 5.16-18, 5.23). It can be observed, and it is an intention of this model, to appropriately concentrate population onto the actual locations where they are likely to be present depending on the season and time of day. This also more accurately reflects areas of zero population density. In contrast to the uniform density across contiguous zones census structure this will result in the model cells containing higher population counts in occupied areas and zero in between.

The greatest difference occurs in the daytime comparison of the model and census daytime estimate (Figure 5.25B) where the standard deviation (σ) = 11.16. This comparison also exhibits the greatest positive difference. This is expected, as in terms of the spatiotemporal modelled outputs this is when population within this study area will be the most concentrated at school and workplace destination cells. This highlights an issue with the underestimation of population occupying daytime locations within the census data.

Figure 5.25C shows the least variance within the three examples (σ = 7.42) for the model difference from the mid-2010 LSOAs. However, there is in fact little overall difference between the model compared to the 2001 OA and 2010 LSOA (Figures 5.25A and C). There is a slight improvement with fit using the newer 2010 adjusted LSOA counts. This is the same target year (2010) for the population data modelled. It also has the lowest mean difference by cell (0.07) and lowest RMSE (Table 5.12)

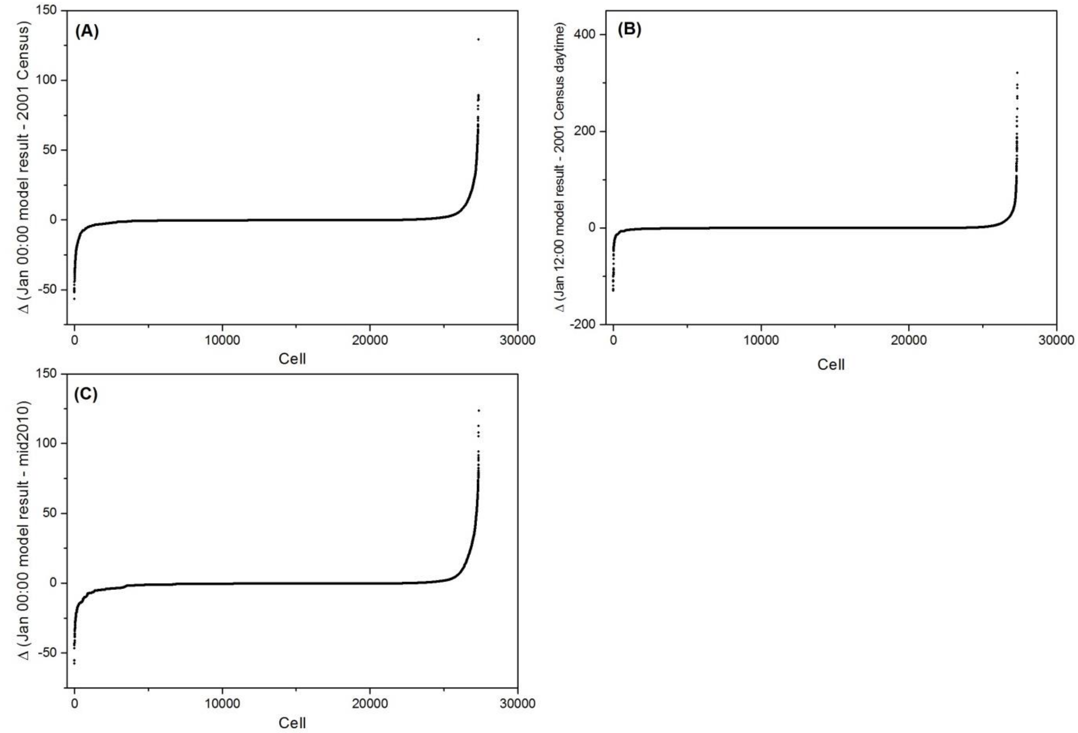


Figure 5.25 Difference at a cellular level (100 m) between day and night-time model results for January and rasterised census data concerning the 2001 OA population (A), 2001 OA daytime population (B) and 2010 LSOA mid-year estimate (C) across the St Austell study area.

In addition to the mean difference and standard deviation a mean percentage error (MPE) and root mean square error (RMSE) have been calculated (Table 5.12). The modelled result is taken as the observed value and the census the predicted. The closest model scenario that corresponds to the census count has been chosen. The census baseline and MYE does not account for seasonal variation and is considered a night-time count. Therefore, the January 00:00 model output was chosen. A January daytime model estimate (12:00) was chosen to correspond with the census daytime count.

Table 5.12 St Austell study area model difference analysis

Values		Measure			
Observed (model)	**Predicted (census)**	**SD**	**Mean**	**MPE**	**RMSE**
January 00:00	2001 Census	8.09	0.66	34.31%	8.12
January 12:00	2001 Daytime	11.16	0.71	52.21%	11.16
January 00:00	2010 MYE	7.42	0.07	13.34%	7.42

Where Δ is the difference between the observed and predicted values

The RMSE closely corresponds with the standard deviation, it represents the average ‘error’ (or difference) for each observation (in this example 30,000 100 m cells). On ‘average’ the model observations overestimate the census values, by up to more than 50% (daytime scenario). For greater clarity the actual difference at a cellular level (100 m) is displayed spatially in Figure 5.25. This shows the model difference from the census day and night-time value (first two rows of Table 5.12). The percentage ‘error’ or difference from the night-time model (January 00:00) results and the 2001 Census (considered a night-time count) is shown in Figure 5.26A. The actual difference in population count between the daytime model (January 12:00) and the 2001 Census daytime count is shown in Figure 5.26B. The greatest positive difference occurs outside of the main population centres (St Austell and Bodmin). In the night-time evaluation (Figure 5.26A) the model clusters population to actual residential locations. Lanivet, Lostwihiel, Lerryn and Golant are labelled as examples. These are predominantly rural settlements, and in contrast the census distributes this population over geographically larger OAs. This results in lower census population densities. Where the model concentrates this population to inhabited residential locations large differences occur from the census count that approximately correspond with the outline of these settlements. This occurs to a lesser extent in the main population centres for the converse of this argument. For example, OAs in more densely populated areas such as St Austell are geographically smaller in comparison and result in higher census population densities. In these locations there is less of a difference between the modelled and census outputs.

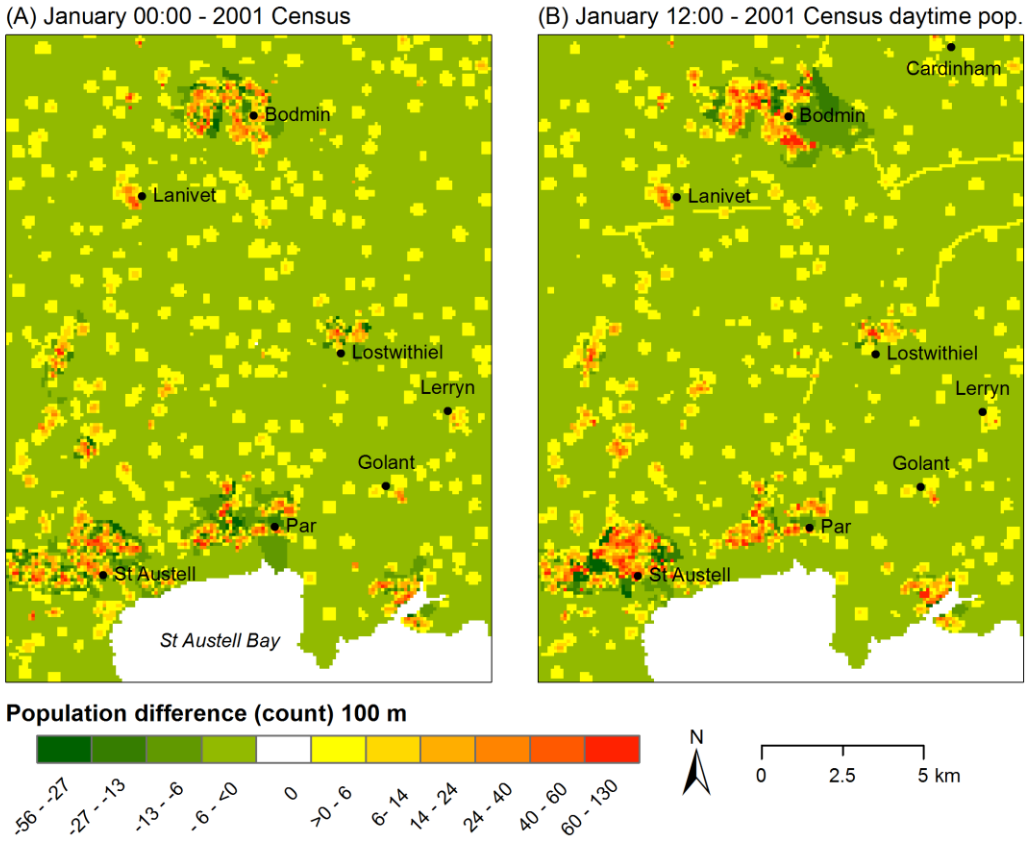


Figure 5.26 Cellular (100 m) difference between spatiotemporal population model results and rasterised day and night-time 2001 Census counts

The contiguous zonal structure of the census does not permit areas of zero population density. Where the model does assign a zero density (e.g. unoccupied countryside between settlements) it 'appears' to be underestimating the census counts by up 5 people per 100 m cell. This is indicated by the light green background colour in Figure 5.26. However, in reality these locations are not 'inhabited' by people and the model reflects this. The largest difference of +130 people occurs in the small hamlet of Cardinham, situated on the extreme northern edge of the study area northwest of Bodmin (Figure 5.26B). The corresponding cell has a daytime census population density of 0.12 people/100 m^2. However, this location contains a small primary school (Cardinham School) with a weekday term-time population of 78 pupils; its centroid is visible in Figure 5.13. This increases further with the inclusion of additional staff. In reality the census estimate simply does not account for localised daytime concentrations in population, including school sites. Administrative datasets (e.g. School Census) used to construct this model's data library confirm that it is known that there is a population associated with that school site for the time modelled.

Parts of the road network can be discerned in yellow in Figure 5.26B. This indicates areas where the daytime model output seems to overestimate the census baseline by up to 6 people per cell. This is expected because the census does not account for the population on the road network where this model specifically does.

5.7 Summary

This chapter has evaluated a case study within a rural coastal setting that is driven by large temporary fluctuations in the seasonal tourism population. At times this additional population has accounted for a significant proportion of the census baseline estimate. Seasonal overnight visitor data have been combined with a usually resident term and non-term time population base to illustrate variation throughout the tourism season, exemplified for three distinct months.

In conclusion the spatiotemporal modelling approach adopted has facilitated the inclusion of a highly influential seasonally varying tourism population. This has advanced current insights on high resolution variation in space and time to produce a range of realistic population estimates. For example, the comparison of a night-time peak season modelled output (seasonally adjusted to account for overnight visitors) with OS background mapping (Figure 5.19) has allowed the identification of temporally occupied locations. In this example an empty field becomes a busy campsite during the peak season.

These insights are simply not possible using static or traditional datasets in isolation. Inclusion of peak summer tourism population has been shown to increase the residential baseline population by up to 18%. Results have demonstrated large changes in population exposure to flood risk by time of day within the same and varying flood inundation scenarios as well as seasonal increase in potentially vulnerable populations. Furthermore it highlights the season specific requirements for vulnerable groups such as occupants of highly exposed caravan sites.

Validation and error analysis is demonstrably difficult for the model results obtained. However, assessment of the differences between the model and census counts supported with administrative datasets show that the model more accurately portrays locations of population at a sub-OA scale. There is a high confidence that the modelled outputs, although contrasting to the census, do depict accurate population clusters. For example, the case of children being present on a school site which is not portrayed in the census is confirmed by administrative datasets. It is acknowledged that the actual magnitudes can be highly variable and subject to external influences such as the weather, tourism economy and social trends. Therefore it is difficult to define a precise value. Nonetheless, predicted clusters appear to correspond with concentrated centres of population.

This application has demonstrated what Martin *et al.* (forthcoming) term the modifiable spatiotemporal areal unit problem where even the most detailed spatial data may be inadequate to support time-sensitive analyses. In this case study population exposure outcome is highly dependent on the time of day, season of the year and varying extent of flood inundation polygons. Just a few of a potentially exponential range of scenarios have been demonstrated and these show large variations in the results obtained. For this reason, an analysis using static data without the temporal insights achieved here will only provide one result which has been demonstrated to have strong spatiotemporal sensitivities.

Chapter 6: Discussion

6.1 Overview

This thesis has sought to develop appropriate methods for the integration of spatiotemporal population estimates with existing environmental models and to demonstrate their practical implementation. Applied examples in Southampton (Chapter 3), Ulley (Chapter 4) and St Austell (Chapter 5) have demonstrated the implementation of spatiotemporal population estimates created for this thesis. The creation of a population data library within a flexible framework has been presented to better represent population movements in space and time. Population movements have been explored for daily, weekly and seasonal time scales.

Following the review of population exposure to natural hazards and techniques for the interpolation of population data (Chapter 2), the importance of spatiotemporal population estimates at appropriate resolutions for the purpose of flood risk analysis was identified and illustrated using appropriate case studies (Aim 1, Specific objectives I, II, IV). The construction and evaluation of two datasets and a population model data library for use with SurfaceBuilder247 was completed (Aim 2, Specific objectives III and V). This chapter discusses the final aim, to assess the extent to which spatiotemporal population modelling techniques can be used to provide greater insights for integrated disaster risk management, and examine to what extent confidence can be placed in their results (Aim 3).

To achieve this aim the focus has been to enhance and apply the newly developed Population 24/7 modelling tool, generating spatiotemporal gridded population estimates beyond the original implementation by Martin *et al.* (forthcoming). New data libraries have been created for this thesis. Relevant extensions of the work of others have been identified and, where appropriate, integrated into the examples presented. This includes small area seasonal visitor estimates (Newing 2014) and bespoke LISFLOOD-FP flood inundation layers (Quinn 2014).

This chapter is divided into five main sections. The first discusses the extent to which model integration has been achieved and important considerations for population exposure that have arisen throughout this process. Section

6.3 discusses the enhancements made to the Population 24/7 modelling framework. This evaluates the creation of a new expanded Population 24/7 data library to account for variation in output resolution and population movements for different temporal scales. It deals with the integration of new data types and creation of new parameter estimates which have not been produced before. Section 6.4 explores model validation against census estimates and examines the challenges involved in validating spatiotemporal population estimates. Finally, Section 6.5 outlines the potential contributions of this research to the field of dynamic population modelling and disaster risk management.

6.2 Integration of spatiotemporal population estimates and environmental models

This section discusses the suitability of the method developed for the creation of spatiotemporal population estimates for integration with environmental models. It compares the gridded spatiotemporal population estimates with other current ambient population estimates available. The integration of population and hazard models that vary in space and time raises questions about application scale (Section 6.2.3). Practical considerations regarding the integration of population and environment models will be discussed in Section 6.2.4. This draws on the evidence proposed in the two case studies (Chapters 3 and 4). Finally, the potential contribution to the improvement of natural hazard risk assessment and further application contexts is discussed in Section 6.2.5.

6.2.1 **Spatiotemporal gridded population outputs**

The representation of population data was the focus of the literature review in Section 3.3. Chapters 3 to 5 have demonstrated the advantageous nature of gridded population estimates through applied examples and case studies. Generating gridded population representations has proved to be beneficial, particularly for the accurate delineation of occupied areas.

The use of a grid as a universal format has many advantages, including the ease of integration of population and environmental model outputs or datasets. This is discussed in more detail in Section 6.2.4. A second advantage is the simple data structure of a grid (e.g. ASCII grid format) for subsequent analysis in GIS, editing outputs, cellular level comparisons and software compatibility. The gridded format provides the stability through time, alignment with national grid reference systems and data refresh options. For the purposes of flood risk analysis it can be beneficial to update population datasets to represent different target dates or scenarios. The grid provides a stable platform of consistent units which allows direct like for like comparisons of different output data from a range of sources.

6.2.2 Spatiotemporal and ambient population estimates

The results and analysis (Chapters 4 and 5) have demonstrated high-resolution spatiotemporal population estimates, at hourly intervals for 100 m cells. It has been shown that variations depending on time of day (e.g. Figures 4.16 and 5.14), day of week (e.g. Figure 5.19) and time of the year (e.g. Figures 5.16-18) can be resolved in the spatiotemporal population modelling approach that has been followed.

It has been observed that static daytime population tables are available from 2001 and 2011 censuses, which provide a comparison to residential night-time counts. The global LandScan database is an example of an ambient population estimate, a population average over a 24-hour period. LandScan USA increases the temporal resolution further with separate day and night-time population estimates (Bhaduri *et al.* 2007). However, the Population 24/7 approach adopted here has produced time-specific population estimates at a much greater temporal (hourly) resolution, which is extensible even further if required. Ambient population estimates are simply not sufficient for most hazard modelling or emergency planning scenarios and also suffer from a lack of agreement over what constitutes 'ambient'. Therefore, for the purpose of this thesis the day/night-time and ambient population datasets also available are not considered spatiotemporal in the same sense at the Population 24/7 outputs demonstrated. This is directly beneficial for emergency planners for the assessment of the risk posed to those affected with a range of scenarios. This is one example of producing greater insights on population movements, valuable for integrated disaster risk management (Aim 3).

Novel evolving applications already noted such as hourly population estimates (e.g. DynaPop, Section 2.6, Aubrecht *et al.* 2014) show promising advances towards truly spatiotemporal population estimates for the purposes of hazard risk reduction. High spatial resolution day and night-time population estimates have also been noted at resolutions of 250 m (McPherson and Brown 2004), 250 m (LandScan USA, Ahola *et al.* 2007) and 25 m (Freire and Aubrecht 2012). There are merits in terms of increased spatial resolution demonstrated within the ambient population examples. However, DynaPop and Population 24/7 further advance temporal granularity, minimising the ambiguity associated with ambient estimates, which is an important consideration within some disaster risk reduction contexts.

6.2.3 **Spatiotemporal scales of hazard**

It has already been identified that hazards occur at a range of spatiotemporal scales in the literature review (Section 2.2.10). However, analysis of the case study results has identified important considerations for the assessment of risk for potentially exposed population.

Types of hazard vary widely in terms of physical characteristics, warning periods and onset speeds. This thesis has undertaken a detailed analysis for two hydrological events, deliberately chosen to represent hazards with very different onset times. In the case of Ulley, a dam failure, the onset is near instantaneous, with total event duration of around 40 minutes (Chapter 4). This potentially results in little or no time for any effective warning. Therefore, there is a compelling case for high-resolution spatiotemporal population estimates to assess the risk from specific events such as this. The approach presented set out to examine how the risk posed from a sudden onset and worst-case scenario flood event varies according to time of day. It has shown that risk assessments at this scale are feasible.

In contrast, the St Austell example (Chapter 5) explores an event with a much longer onset time. It involves varying levels of flood inundation based on an 11-hour rainfall event. In reality, this would also have been combined with prior weather warnings. In this scenario it is acknowledged that such weather events and warnings could have an impact on the population present within flood risk zones. The argument is maintained that it is not unreasonable to again consider the worst-case scenario with the maximum population present. Brown and Damery's (2002) suggestion for enhanced flood risk management requires long-term risk strategies to be grounded in an understanding of exposure to flood hazards and patterns of vulnerability. This is required as a starting position for the development of targeted warning systems, emergency plans and future policies. Equally, in the absence of relevant data, it is difficult to substantiate assumptions as to who may, or may not, be present or who could, or could not, have responded to prior warnings. The Population 24/7 approach demonstrated in this thesis could be used to prioritise warnings and to make informed decisions about effective measures to be targeted at the most vulnerable population subgroups and areas.

The Population 24/7 modelling approach estimates population distributions based on quantifiable cycles and known administrative data counts associated with specific locations. The implementation in this thesis does not attempt to simulate human behaviour during a hazard scenario. For this type of application an agent based model or microsimulation may be more appropriate, but these would require a different type of modelling that has not been the focus of this thesis. However, the Population 24/7 process could be used to seed such models (see Section 7.6). To some extent the Population 24/7 data library can be modified to better represent population during a hazard scenario, but this is not the same as a dynamic simulation. The background layer can be modified to reflect severe congestion or closures on major arterial traffic routes if appropriate data are available. In Chapter 5, seasonal adjustments were made to the population baseline to reflect tourist cycles.

6.2.4 **Integration of existing models and datasets**

This research has successfully demonstrated the ability to integrate the population results with those from hydrological models (TELEMAC-2D, LISFLOOD-FP) (Specific objective IV). Unlike zonal census or administrative data both the hydrological models used and SurfaceBuilder247 produce a raster gridded output. These have all produced grids at different resolutions (Table 6.1). A feature of all three models is the capability to produce outputs with a variable grid resolution. The SurfaceBuilder247 output is governed by the resolution of available input aggregate data sources to avoid overspreading (Martin *et al.* 2000). An increase in resolution from Martin *et al.*'s (forthcoming) original implementation and Fielding's (2007) static application (using the original non-temporal SurfaceBuilder) of 200 to 100 m has been achieved by re-weighting aggregate census data onto georeferenced postcodes. While it is technically possible to produce gridded population estimates to such high resolution as the hydrological models listed, it is not defensible to do so due to current data limitations. The

Population 24/7 approach provides a framework that is adaptable to multiple input data sources allowing the spatiotemporal resolution to be increased in the future should appropriate datasets become available. The LISFLOOD-FP layers were produced by Quinn (2014) at an output resolution within the range considered appropriate to resolve flow characteristics (Neal *et al.* 2011). The TELEMAC-2D output was converted from an irregular triangular mesh at 15 m to preserve all data points.

Table 6.1 Model output resolutions

Model	**Output resolution (metres)**	**Application example**
SurfaceBuilder	200	Fielding (2007)
SurfaceBuilder247	200	Martin *et al.* (forthcoming)
SurfaceBuilder247	100	This thesis, Chapters 4 & 5
TELEMAC-2D	15	This thesis; Smith *et al.* (2014b)
LISFLOOD-FP	5	Quinn (2014)

Although in this case all model outputs have different resolutions, a benefit of a regular grid means that they all can be aligned to the same national grid system to undertake spatial analysis. The population grid was resampled to the same common resolution as the hydrological outputs. The ability to compare corresponding output cells of each model permits the calculation of exposure estimates and evaluation of Penning-Rowsell's (2005) flood hazard methodology with specific population exposure values at a cellular level. The advantage of achieving a common data structure is that this further analysis can be undertaken in standard GIS or statistical software package without additional specialist knowledge. Datasets produced for alternative purposes or by different organisations are often available at different levels of aggregation. The approach presented in this thesis permits the creation of outputs that could provide valuable enhancements in flood risk management, by allowing spatiotemporal population estimates to be combined with environmental models where there is no single ideal population dataset for the purpose.

Through application case studies it has been demonstrated that recasting all the data into a common format, in this case a regular grid aligned to the national grid reference system, can be used to successfully integrate environmental and population models. However, this is just one mechanism that demonstrates the real advantages for the handling, alignment and integration of temporal data cycles. These examples have been evaluated for UK applications; however the underlying concept for transferability remains the same. Where data are sufficiently available for the resolution required there is no reason why the model framework and integration method presented cannot be created for other contexts and in other countries (See Section 6.3).

The seasonal and hourly population patterns modelled represent 'typical' baseline conditions. In the example of the Ulley case study (Chapter 4) only one static flood map was incorporated. This is justified on the need to understand the baseline population in advance, and keep it updated, prior to a hazard event occurring. This is a key factor for hazard risk reduction (McPherson and Brown 2004; Freire *et al.* 2011). The extensible framework and data library construction for Population 24/7 permits data to be refreshed or modified. As demonstrated (Chapters 3-5), methodologically it is possible to combine raster based hazard and population data using simple GIS functions. However, changes in human behaviour and spatiotemporal distribution are not the focus of the Population 24/7 data library constructed. Therefore, it is not appropriate to use the same population data library for multiple hazard time-slices without the ability to account for changes in population behaviour as the hazard unfolds.

6.2.5 **Improvements for population risk assessment**

Results from the case studies have shown dynamic population fluctuations, which in comparison to static census outputs give valuable time dependent assessments of population exposure to hazards. Chen *et al.* (2004) (Chapter 2.4.13) illustrated the problem of aggregate population units in population hazard risk assessments. The approach demonstrated improves the spatial accuracy in mapping population placement (e.g. Figure 5.20). However, these distributions vary both spatially and in intensity with time which can be related to quantifiable population cycles. The analysis undertaken demonstrates this cyclical nature such as the increasing density of the usually resident population with seasonally dependant retail, leisure and work activity (e.g. Figures 5.16B and 5.18B). These insights are only achieved by also considering the temporal resolution of such data, and the Population 24/7 framework provides a mechanism to handle this in addition to the spatial dimension. There is an increase in the spatial distribution and concentration of the temporary overnight visitor population driven by the tourist seasons (e.g. Figures 5.16D and 5.18D) which is not accounted for in existing datasets.

Aubrecht *et al.* (2013) also discuss the integration of spatiotemporal population characteristics for disaster risk management. They highlight that the requirement to understand the development of risk over time is crucial and population changes in both space and time are often not sufficiently studied. The Population 24/7 development and applications presented have shown that it can produce high resolution population estimates in space and time suitable for hazard risk assessments.

One crucial population characteristic resolved using the Population 24/7 approach was the temporary occupation of caravan sites and holiday parks by seasonal visitors. In the St Austell example caravans placed directly behind sand dunes on Par beach are within the flood risk and former intertidal zone (Figures 5.20 and 5.21). McEwen *et al.* (2002) highlights the particularly high susceptibility of caravan sites to flood risks due to their structural integrity and often precarious placement. This Population 24/7 approach at the 100 m scale facilitates these observations. Caravan and campsites are not occupied all year around and do not contain usually resident census populations. This approach not only identifies these population features that are otherwise omitted, but also shows how they vary seasonally with the tourism cycle.

The sensitivity of aggregate data analysis, particularly prevalent in population data, depends spatially on the MAUP (Openshaw 1984) and temporally on the modifiable temporal unit problem (Çöltekin *et al.* 2011). Subsequent data analysis can be restricted by imposed temporal constraints such as the representative time slices considered. Martin *et al.* (forthcoming) term this the modifiable spatiotemporal unit problem (MSTUP), which combines the two phenomena. They propose that data required for analysis need to be sufficiently detailed in both the spatial and temporal dimensions. This limits the effect of distortion in data analysis. Appropriate resolutions are dependent on specific application requirements. These phenomena can potentially have significant impacts on risk analyses generated. This is not a barrier for spatiotemporal flood risk assessment, but does draw attention to the requirement to consider the spatial and temporal granularity of input data used. These phenomena highlight the problems of using static, aggregate population data for risk analyses to derive single values for highly variable characteristics.

In the examples demonstrated in this thesis the finest spatial resolution currently achievable, 100 m, is appropriate for the application scale and data limitations and has been implemented within the Population 24/7 framework. This is based on the availability of input data sources with variable resolutions. The aim of this research was to create estimates with the finest spatial and temporal resolution based on the data currently available to limit the effect of the modifiable spatiotemporal unit problem outlined by Martin *et al.* (forthcoming). Overspreading of population data occurs with the current data library for output resolutions < 100 m. The current output resolutions is appropriate for the centroid density and variation between urban and rural locations. If it was possible to increase the spatial resolution further, for example to 75 m, this may pose limitations for integration with some environmental datasets. For example, data sets aligned to 1 km national grid squares where 75 m is not wholly divisible. Aside from available input data resolution, it is not currently possible to accurately predict high spatial resolution (e.g. 25 m) population distributions for the desired temporal granularity with current data.

A high degree of variation in terms of population exposure can be observed within the hourly temporal scale used (e.g. Figures 4.16 and 5.14). This demonstrates one instance of the modifiable temporal unit problem where consideration of just one time-slice (e.g. using a static population base or considering just one hourly interval) can have large impacts on the results obtained. Greater temporal granularity, such as that beginning to be shown in the results presented within the two case studies, has important implications for analysis where risk to people is concerned. The method proposed captures these important variations and demonstrates that static representation is inadequate in these examples.

6.3 Enhancements to the Population 24/7 framework

Martin *et al.*'s (forthcoming) objective through the Population 24/7 project was to produce a framework for spatiotemporal population modelling. This was intended to be extensible and the initial implementation did not explore all the necessary details. This thesis has tested the extensible nature of this framework through the creation of two brand new population data libraries for each case study presented. One benefit of constructing a population data library within the flexible modelling framework is the ability to account for specific population characteristics.

The creation of new population data libraries for use with the Population 24/7 tool and application to two case studies has demonstrated a range of enhancements beyond the original implementation by Martin *et al.* (forthcoming). These enhancements are evaluated and discussed in turn within this section. This section begins by discussing creation and use of multiple population origin datasets and their importance in applications for disaster risk management. Section 6.3.2 discusses the enhancements to hazard risk assessment achievable when considering population subgroups. Section 6.3.3 discusses the enhancements for hazard risk analysis through more realistic representation of population using destination centroids. Section 6.3.4 evaluates improvements achieved in the spatial resolution of population distributions using UPC data for population centroids. Finally, Section 6.3.5 comments on the challenges associated with the static representation of dynamic datasets.

6.3.1 **Multiple population origin classes**

The St Austell case study (Chapter 5) demonstrated the advantages of creating multiple categories of population origins within the model's data library. These were temporary overnight visitors by season, the usually resident term time population and usually resident non-term time population. The non-term time population was created based on a secondary analysis of census data following the method proposed in Chapter 5.3.2.

Since this research was undertaken, ONS have released a specific non-term time census dataset in addition to the usually resident census baseline population (considered a term-time count, National Statistics 2004). The ONS (2014b) report on the methodology is largely consistent with the methods proposed in this thesis. This new 2011 Census release could be used in the same way to construct a non-term time population origin dataset within future applications for desired target dates since 2011. For consistency with the data available, the St Austell case study, completed before the 2011 census release, corresponds to a target year of 2010. The ONS (2014b) report on out of term time populations recognises the effect that the unique spatiotemporal characteristics of the student population has on the 'usually' resident population. The non-term time origin dataset was created for the St Austell (Chapter 5) application because analysis of the peak tourism season in August also corresponded with non-term time for the usually resident population. Although subsequent analysis showed little difference from the term-time estimate (Table 5.6) for this study area location it is a matter of wider significance. For example, according to the 2011 Census, ONS (2014b) report large university cities such as Leeds, Manchester and Sheffield had a non-term time population of at least 20,000 people lower than their usually resident populations. A non-term time origin dataset was not created for the Ulley case study (Chapter 4) because its focus was on the representation of the June 2007 (term time) flood risk situation. Secondly the temporal scale of this case study focuses on diurnal rather than seasonal population cycles.

A second significant enhancement to the Population 24/7 framework was the creation of a seasonally varying population origin class comprised of temporary overnight visitors. This was based on visitor data constructed by Newing (2014). The analysis demonstrates large fluctuations within the St Austell study area of greater than 12,000 people (Table 5.5). These are based on known population flows informed by bed space occupancy surveys, census data on second home ownership (where no 'usually' resident population is enumerated) and the pull factors of some of southwest England's largest tourist attractions. These produce insights on quantifiable population fluctuations that simply cannot be achieved from census data alone and account for potentially significant exposed population groups.

One example of the significance of temporary overnight tourists is illustrated in Figures 5.20C and D. The comparisons with Ordnance Survey background mapping for selected 1 km national grid squares, they show the level of accuracy achieved in the georeferenced placement of tourist populations. In these examples it is clear where the model has assigned a population density that corresponded with caravan parks and campsites. These would be considered uninhabited by a usually resident census population, but become temporally occupied by overnight visitors. The modelling approach adopted allows the identification of this known phenomenon that otherwise is not possible using the census alone.

It has been noted (Section 5.2.2, Figure 5.20D) that a slight overspreading of lower population densities is observed, particularly around Lower Penhale in the example illustrated. This was attributed to the wider dispersion of rural households around a single postcode. However, it is encouraging that the spatial distribution of model outputs does correspond to local features. This in itself demonstrates significant improvements over standard choropleth census density maps. It is worth reiterating that the modelling undertaken does not claim to predict individual human behaviour. While this data library has been constructed to the highest geographical resolution currently achievable, deliberating on the micro scale fluctuations within what may happen in an additional 50-100 m of a centroid is not within the scope of this application. This application is focused on the wider impacts of larger scale flood risk on spatiotemporally varying populations. Population counts at individual household level are currently not publicly available in the UK. It is possible that future research (see Section 7.6) could further improve the accuracy in the representation of rural dispersed dwellings. An advantage of the Population 24/7 framework and creation of individual population origin classes is the ability to update the data library when new data become available.

The enhancements discussed in this section demonstrated modest improvements for disaster risk management. They have all been implemented within the existing extensible Population 24/7 framework. It demonstrates that the data library constructed for use with the Population 24/7 modelling tool can be adapted for straightforward applications, as proposed in Martin *et al.* (forthcoming), or account for more specific characteristics (e.g. Chapter 4 and 5).

6.3.2 Modelling population subgroups

Modelling population subgroups to assess exposure to flood hazard has provided significant enhancements to understanding risk in the examples presented. The insights gained from accounting for vulnerable population subgroup exposure to hazards based on age or other demographic characteristics (for gridded population outputs) have been demonstrated before. Fielding's (2007) environmental injustice paper on the disproportionate exposure to flood risk of those from lower socioeconomic backgrounds in England and Wales (200 m grid), measures exposure by socioeconomic classification. Secondly, Aubrecht *et al.* (2012a) consider the population aged over 60 years at 1 km resolution to assess heat-related vulnerability and exposure. The applications constructed here also build on these developments in the area of risk management.

In both Martin *et al.*'s (forthcoming) implementation and the applications presented in this thesis seven population subgroups defined by age were modelled. These were chosen because of the unique spatiotemporal characteristics they exhibit. For example school aged children spend term-time weekdays highly concentrated on school sites, and higher education students have already been observed as being highly mobile and a significant proportion of the population (e.g. Section 5.4.2; Figure 5.23; Smith *et al.* 2014a). Age subgroups can also be used to infer the resilience of populations typically considered more vulnerable (e.g. the very young and elderly). Age subgroups also lend themselves well to the data library structure created. The appropriately aged populations can be drawn from origin centroids for the most appropriate destination categories. For example, only school age children populate school destination centroids (together with the associated staff) and the 16-64 working aged population occupy their respective workplace destination centroids. Fielding (2007) created a static implementation using population subgroups based on socioeconomic status. This exemplifies a possible alternative to age.

However, for a spatiotemporal implementation, it becomes more difficult to differentiate population subgroups by socioeconomic group and decide relevant occupancy of destination locations. Although age is not the only method for defining population subgroups within the Population 24/7 framework, the inclusion of alternative demographic classifications would require further careful consideration on how allocation to destination sites might work and whether suitable data were available to inform such allocation.

The Southampton worked example (Chapter 3; Smith *et al.* 2014a) demonstrated a diurnal cycle in the weekday dominance of fluvial and tidal flood risk exposure (Figure 3.14). Based on the population subgroup breakdown by age (Figure 3.15) it was shown that the working aged population became more exposed to tidal flooding during the working day. This was explained by the nature of the coastal concentration of industry and commerce within the Southampton example. It was only possible to achieve this insight based on subsequent analysis of the population subgroups modelled. This illustrates two important contributions by this thesis to traditional flood risk management: (i) flood risk to people is not static and varies by both time of day and flood hazard type, and (ii) modelling of age subgroups allows a more precise estimate of the population sector that is most at risk. It would not have been possible to derive these assessments using only static population data or without the consideration of population subgroups. The 100 m output resolution has also been shown to be appropriate to resolve these features on a city level scale compared to alternative gridded resolutions (e.g. 1 km Gridded Population of the World).

The St Austell case study (Chapter 5) shows another advantage of age based population subgroups. This example illustrated the requirement to account for the effects of large seasonal population fluctuations. In one flood scenario for August the exposed population decreased despite the fact that the study area total increased (Table 5.10, LISFLOOD-FP R100). The suggested cause of the decrease in this scenario was the increased clustering of population in lower risk areas. However, investigating the age breakdown (Figure 5.22) showed a disproportionate increase in the retirement aged population (over 65 years) within the August 12:00

LISFLOOD-FP R100 exposure estimate. This increase was primarily comprised of the August visitor population. The exposure of the population aged over 65 increased by 385% between May and August for the same flood scenario (R100). It is possible that the overall result for the total population exposure (all ages, residents and visitors combined) for the R100 flood scenario (Table 5.10) may have been disregarded on the basis of an overall decrease in population exposure. However, the analysis of the population subgroups for the same scenario revealed a significant increase in the older, and more vulnerable, population exposed. This is potentially of high significance to emergency planners where vulnerability assessment and risk mitigation is concerned, particularly within less resilient population subgroups.

Two recent examples have been chosen to support the spatiotemporal population estimates produced for risk evaluation within this thesis. A potential critique is that this approach does not account for human behaviour, but this is not the intention of the work present here. Nonetheless, in the example of 2014 flooding in Moorland, Somerset (Section 5.1; BBC 2014a) some residents still refused to evacuate despite the village being isolated by floodwater and immediate evacuation requests from overhead police helicopters. Secondly, Cole and Fellows (2008) identified that a number of population subgroups failed to evacuate New Orleans during Hurricane Katrina. They proposed a variety of reasons from concerns for security of property, caring for pets not permitted inside evacuation shelters and those with reduced mobility and their carers. These examples show that even the best intentions can struggle to account for unpredictable human behaviour. Considering a ‘normal’ population baseline prior to any hazard event permits estimation of the worst case scenario and identification of potentially vulnerable population subgroups. This allows the development of effective and targeted risk mitigation strategies without waiting until it is too late.

6.3.3 **Development of destination datasets**

The development of destination centroids for the SurfaceBuilder247 data library allows the more accurate placement of population at sites with known population capacities and temporal signals. Static census counts provide a reliable residential population count. However, they do not account for population movements away from residential locations.

The mechanism to create destination centroids based on known population locations and capacities within the flexible Population 24/7 framework is a powerful tool. Where sufficient data exist this allows for appropriate representation of the temporal characteristics of the chosen population to be modelled. For this thesis, new destination datasets were created to represent education, workplaces, healthcare, retail and leisure attractions.

One improvement is shown in the development of a retail destination dataset created for the Southampton applied example (Chapter 3) and the St Austell case study (Chapter 5). The approach outlined in Chapter 3 used business postcodes as a proxy for retail locations that excluded residential address, and devised a footfall estimate based on the number of retail employees. Actual retail footfall data are beyond the financial resources for this demonstration of the Population 24/7 modelling techniques, and often restricted due to commercial sensitivities (Newing *et al.* 2013a). Instead, feasible alternatives using available data have been used for case study demonstration purposes. An advantage of the Population 24/7 approach is the framework that allows the adaptation to new or evolving datasets. The retail example demonstrated here could be enhanced by a research team with access to commercial datasets.

Retail estimates for the St Austell case study were improved using commercially available GMAP (2014) retail centres. These replaced business postcode locations and only represent retail locations, as opposed to other businesses, for greater spatial accuracy. Another advantage was the inclusion of an estimate of retail floor space for each centre. This permitted footfall to be estimated based on average sales density (income per floor area) and average transaction values obtained from publicly available financial reports. This resulted in greater spatial accuracy by using a hybrid of commercial and open data. Retail is an important destination category because town centres are often associated with high retail footfalls within the commercial districts. Accounting for high concentrations of temporally varying population is of high relevance within disaster risk management and for emergency preparedness. The Population 24/7 framework adopted has enhanced the original implementation because it can accommodate different types of data relating to destination activities. The detail and range of destination categories (e.g. retail and leisure) are dependent on the study area selection. The flexible modelling framework has been demonstrated to handle these location specific characteristics.

6.3.4 Use of unit postcode (UPC) data for centroids

The previous subsection identified some disadvantages associated with using business UPCs to identify retail centres because retail cannot be distinguished from other business types. However, their use for the construction of other origin and destination centroids has permitted an increase in model output resolution to 100 m. The use of residential UPCs for reweighting census OA counts is shown to increase resolution and accuracy of residential population densities (e.g. Section 3.3.10, Figure 3.12). Although this results in a greater centroid density and associated computational time, it is a major improvement on using single OA centroids or PWCs.

Similarly, workforce counts reported at LSOA (the finest available areal unit for which they are published) from the ABI/BRES were re-weighted onto business UPCs. There is a significant difference in the spatial distribution of residential and business UPCs (e.g. Figure 5.13) which more accurately reflects these locations compared to using the same census centroids for both. UPCs, although originally created for the delivery of mail, are often used as reporting zones in some administrative and survey datasets which can support the collation of datasets otherwise reported in varying spatial units.

6.3.5 Static visualisation of dynamic data

A challenge associated with reporting and communicating spatiotemporal data effectively is visualisation. This can be difficult in traditional media such as print. This thesis has described dynamic population examples, but so far has had to rely on static time-slices for illustration. The examples shown are part of a greater temporally varying dataset that has been created. A 3D visualisation has been developed for display in Google Earth.

The example provided in Figure 6.1 is a screenshot captured from Google Earth showing the total St Austell population at midday for a weekday in August. The vertical bars signify population density. This dynamic method for illustration allows a much more interactive approach to data communication.

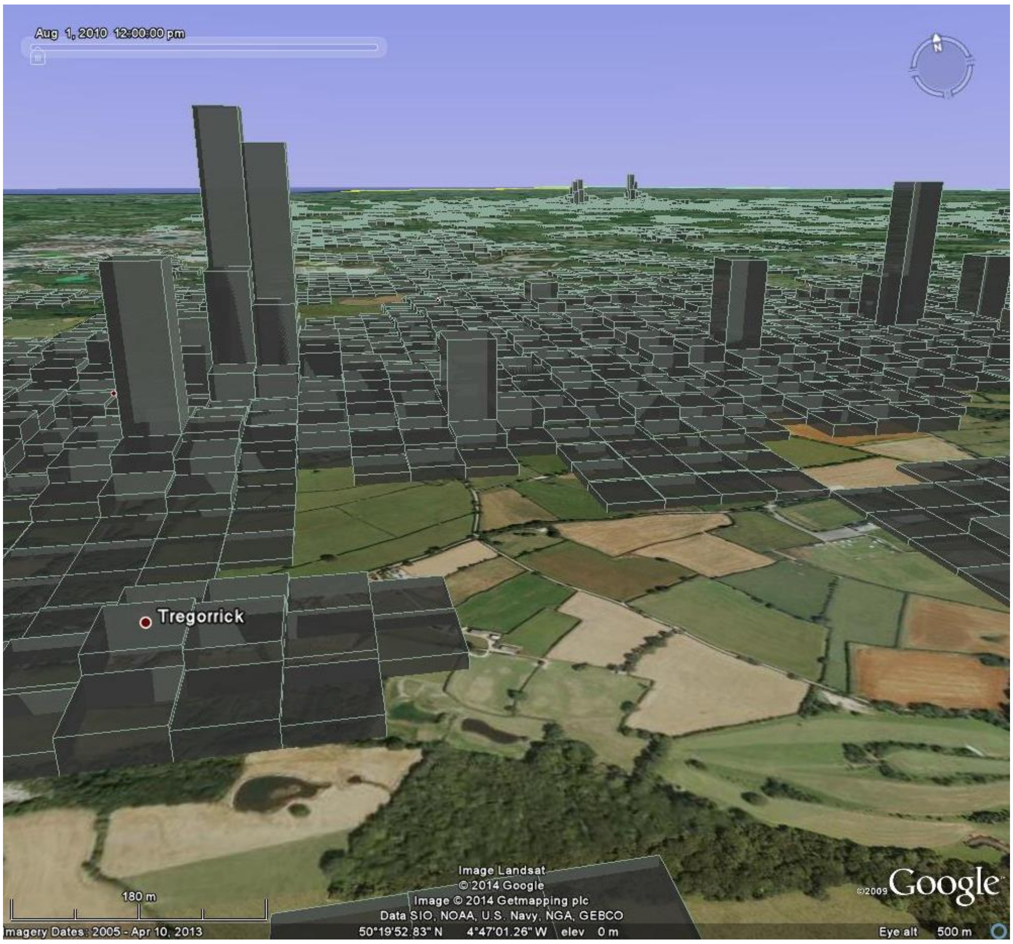


Figure 6.1 Static image from an interactive 3D visualisation of SurfaceBuilder247 results using Google Earth. Vertical bar height represents population density (cell size 100 m).

The ability to visualise spatiotemporal population estimates interactively is an additional powerful toolset. It makes use of standard, freely available, software and existing functionality such as the time slider, zoom and panning tools. Complicated population datasets can be difficult to display and interpret, even for practitioners. This is further complicated by the addition of temporal components. The use of familiar interactive software can be used to aid risk communication with non-specialist audiences or engage policy makers.

6.4 Challenges and model validation

This section discusses the significant challenges associated with spatiotemporal model validation. It is broken down into the following subsections: firstly, known limitations in the general application of census data and then for risk applications are examined. Secondly, the extent to which it is possible to validate the spatiotemporal estimates produced in this thesis is discussed. Finally the options are explored for future validation.

6.4.1 **Known limitations of traditional datasets**

Censuses in developed countries are often considered the gold standard in population data. This is usually because of their high levels of accuracy and universal coverage. The EU Census Hub project (Eurostat 2014) aims to create comparable census statistics EU-wide. The value of small area census statistics across environment, health and commercial sectors is summarised in RGS (2014) with example case studies. While censuses provide valuable information for governments and policy makers for the provision of services and funding they are not universally suitable for all hazard risk analysis applications. Known limitations associated with unrealistic uniform daytime population densities across census zones have been accepted for some time (e.g. Wright 1936). Openshaw (1984)'s MAUP highlights the impact of the spatial units chosen for analysis and the effect on any results obtained. Furthermore, Langford and Unwin (1994) have described how traditional choropleth maps, such as those commonly derived from census, grossly distort and mask underlying population densities. Based on the known limitations of traditional census data, for applications such as risk analysis where detailed population estimates are required, this thesis presents one alternative which also varies with time.

The gridded spatiotemporal population estimates generated reduce the MAUP by allowing the comparison and analysis of data on a high resolution stable uniform grid. Unlike static census data, and addressing Wright's (1936) reservations, the Population 24/7 approach concentrates daytime population to known occupied locations depending on the time of day. For the purposes of flood risk assessment this contributes enhanced dynamic insights into population movement that are not captured in static uniform census densities. The origin centroids for the modelling undertaken were constructed using census data, but the argument proposed intends to highlight the benefits achieved using a model to combine multiple relevant datasets into a stable output format that accounts for temporal variation. Static areal interpolations (e.g. Martin 1989) address some of the limitations discussed but still do not account for population movements in time as well as space.

For confidentially purposes published census data are aggregated to areal units. In England and Wales it is likely that census data will continue to be released for just a single aggregation geography (to prevent unlawful disclosure of census data). Duke-Williams and Rees (1998) suggest that it could be possible to safely publish census data at a 5 km grid in addition to the standard areal units. The approach presented in the two case studies demonstrates the current ability to achieve much greater resolutions down to 100 m. A 5 km census grid may address some of the data analysis and integration limitations for England and Wales. However, some censuses are already routinely published as a high resolution grid. E.g. 100 m in Northern Ireland (NISRA 2015) and from 100 m in Austria (Statistics Austria 2013). This may provide further refinement opportunities for Population 24/7 as exemplified with the static SurfaceBuilder implementation against 100 m gridded Northern Ireland census data (Martin *et al.* 2011). Future implementations or advances in input data resolution could allow the spatial resolution of Population 24/7 to be increased further.

Census data alone are still static and the spatiotemporal estimates produced show great potential to resolve cyclical trends at higher resolutions. Population variation has been demonstrated within 1 km national grid cells using the 100 m outputs created (e.g. Figure 5.20). For the application scale presented these are valuable insights for risk management that would not be resolved using lower resolution data.

6.4.2 **Model comparison with census estimates**

SurfaceBuilder247 model results have been contrasted with equivalent (as far as possible) census data in a comparison exercise (e.g. Sections 4.5 and

5.6). These results (e.g. Figure 5.26) demonstrate the extent to which census data in contiguous areal units do not represent observed population distributions. In reality, as predicted by the model, populations are concentrated at residential or other occupied known locations. They are not universally distributed as the census reporting mechanism suggests. Another observed census limitation (Section 5.5) is that it does not account for the daytime concentration of the population at places of work or study. Where this is accounted for in the model large discrepancies from the census baseline appear to occur (Figure 5.26). These estimates can be corroborated with known population counts from administrative data sources.

The comparison attempt has highlighted that the census is fundamentally at odds with known, corroborated, day and night-time population concentrations. This raises a wider question for further validation returning to the original aim that there is no single dataset currently capable of providing these insights. Future validation possibilities have been considered and are proposed as possible extensions to this research in Section 7.6. Such possibilities could include the use of ‘big’ data sources to enhance or calibrate the existing model. The framework has already been demonstrated as capable of being adapted to include a range of data. There are potentially opportunities to utilise real-time data feeds to inform dynamic population movements. These types of enhancement may become more widely accepted by academics, practitioners and policy-makers in the future as supplementary material to traditional censuses.

The comparison with census data suggests that the least amount of variance occurs in the daytime model result (January 2010, 12:00) compared to the 2001 Census OA daytime population count in the Ulley example (Figure 4.17B). However, for the St Austell case study it is the night-time model results (January 2010, 00:00) compared to the 2010 LSOA MYE (as considered a night-time residential count) that show the least variance. In both examples an appropriate model extract representing day or night-time population estimates was compared to the corresponding day or night census estimates. The January seasonal population scenario was used because census estimates represent term-time counts without temporary visitors. January is both term-time and has the lowest number of visitors in the Population 24/7 data library.

In the case of St Austell it might be reasonable to expect that the 2010 LSOA MYE more closely resemble the same 2010 target date for the model data library rather than the 2001 census releases (although at LSOA level, the 2010 MYE is at a lower resolution than 2001 Census OA counts). The range of formal ‘error’ quantification conducted (Table 5.12) supports this expectation. The St Austell model shows least variance from the 2010 MYE compared to the higher resolution 2001 census OA estimates.

In contrast to St Austell, the Ulley model (for target year 2007) appears to show greater correspondence with the 2001 daytime census estimate than the LSOA estimate. This could be because the difference from the census year to the application target date is less (e.g. 2001-2007 for Ulley rather than 2001-2010 for St Austell). This contradicts the perception (and St Austell example) that comparing population estimate datasets of the same location and target date would have the least variance. In the Ulley case this suggests that data resolution rather than the target date is a greater contributing factor. 2001 OA estimates are at a much greater resolution (c. 300 people) compared to the target year 2007 LSOAs (c. 1,500 people) (e.g. OA-LSOA comparison Figure 3.12). The Ulley example includes Rotherham, a large metropolitan area. In this location more OAs cover smaller geographical areas (compared to rural St Austell) and still comply with the minimum population thresholds for census confidentiality (due to higher urban population densities). Therefore, when rasterised they too result in higher population densities. A feature observed in the Population 24/7 model outputs is the concentration of populations to known occupied locations. Smaller OAs tend to result in higher population densities compared geographically larger LSOAs. Therefore when higher urban OA densities are compared with the model outputs the difference between the highest concentrations reduces. Analysis on these study areas suggests that spatial resolution can have a greater effect than reference date in urban areas. The analyses indicated that there is no universal best fit or validation technique.

The model comparison with small area census statistics has raised some important issues. The traditionally accepted view that the census provides a 'true' population value is questionable, given the static nature of the census and the limitations outlined (Section 6.4.1), when trying to validate spatiotemporal population outputs. Indeed a decennial census can be considered as giving a true and highly accurate reflection of the residential population on the census night (except for the representation of uniform population densities) but when these values are used out of context for risk assessment and spatiotemporal model calibration, differences arise that are not necessarily indicative of failings in the modelling undertaken.

The modelling undertaken here has used administrative and government datasets that give a robust estimate of the population occupying the destination locations modelled. Census counts have been used to inform the residential origin population. Model outputs will differ from the census because it shows only a uniform density night-time residential population count.

6.4.3 Handling population subgroups

One of the strengths of the Population 24/7 approached demonstrated is the ability to handle population subgroups and observe how these also change in space and time. However, incorporating population subgroups while constructing the population data library is challenging. One challenge is based on the availability of input data and the ability to resolve the desired subgroups. Uniformity in subgroupings also needs to be maintained throughout all constituent library datasets (e.g. origin and destination datasets). In the examples presented age was chosen. Age permits spatiotemporal characteristic judgements to be made (e.g. working 'aged' population, school 'aged'). Constructing age subgroups is increasing challenging even with the wealth of new data becoming available. For example the new 2011 Census releases report age at a range of spatial resolutions and groupings (single year – 5 year age bands) for the usually resident population, out of term time population, student population and workplace zone population. Finding commonality within just one dataset can govern the subgroups (e.g. age) that it is possible to define as opposed to those that are desired.

Fielding (2007) implements an interesting static gridded population using socioeconomic subgroups. Further work is required to examine how or if alternative subgroups such as socioeconomic classifications can be aligned to a temporal profile in the same way as age for spatiotemporal applications. Defining population subgroups (e.g. age or gender), if desired, is an important consideration during the initial stages of data library construction for the Population 24/7 approach.

6.5 Contribution of enhancements for hazard applications

Finally, this section discusses how the work demonstrated through the two case studies presented can make a tangible contribution to practices and conceptual frameworks within the field of natural hazard risk management. This is divided into two sections outlining the enhancements to risk assessment practices and the wider applicability of the research presented here.

6.5.1 **Enhancements to risk assessment practices**

The research presented demonstrates the innovative development and application of a new spatiotemporal population modelling technique and shows modest enhancements to current risk management practices. The need for better spatiotemporal population estimates to assess vulnerability and exposure to natural hazards, where aggregate census data alone are insufficient has been widely documented (e.g. Ahola *et al.* 2007; Bhaduri *et al.* 2007; Cutter and Finch 2008; Zevenbergen *et al.* 2008; Aubrecht *et al.* 2013). Producing truly spatiotemporal high-resolution (100 m) population estimates as opposed to 'ambient' representations is the most significant contribution of this research. It acknowledges that there are varying circumstances where high or low resolution estimates are most appropriate but this intends to enhance what others have done considering the former. It has successfully integrated spatiotemporal population estimates with the outputs from established environmental models (e.g. LISFLOOD-FP and TELEMAC-2D) using a loose-coupling approach. Previously, this integration has had to rely on inadequate population data in lieu of an appropriate alternative. The method presented provides an example of such an alternative.

Integration with environmental models has demonstrated the ability to resolve population fluctuations in terms of hazard exposure and potential fatalities estimates. The output resolution achieved has permitted these cycles to be distinguished at the city or town level. These observations are not possible using static census data alone. Furthermore, modelling population based on age subgroups has proven to be of relevance for emergency planners. The case studies have demonstrated that it is possible to differentiate certain subgroups that are more vulnerable based on their age and spatiotemporal characteristics.

In terms of policy impact this research directly addresses current national and international agendas. The applications of the approach described within this thesis are well aligned with the UK Government's National Security Strategy (HM Government 2010). Natural hazards are identified as high priority risks. Following the Pitt review (2008) of severe flooding experienced within the UK this research has the ability to directly address at least three key recommendations contained within this report (Section 2.3). These concern the mapping of flood risks at local authority level, equipping individuals to be better prepared and aware of flood risks and the monitoring and mitigating the impacts of flooding on population health and wellbeing. Within a national perspective it also has the potential to contribute to the UK's obligations under the EU Floods Directive (Section 2.3.5). This requires member states to assess flood risk for differing flood scenarios. However, the methodological development and application of the flexible framework is of greater relevance than to just a single county's identified priorities. Hazards do continue to affect humans, but this method provides an option to better account for these impacts more generally. This thesis has demonstrated that flood risk is not static because hazards and populations both vary spatiotemporally and in their composition.

6.5.2 Applications for spatiotemporal population data

The spatiotemporal population modelling techniques demonstrated through the example case studies have focused on population exposure to natural hazards. However, this is only a small part of a much wider range of potential applications for such data. These applications have been demonstrated for, but are not limited to, flood risk examples. The collaboration developed to integrate Newing's (2014) seasonal visitor estimates utilised a dataset originally created to estimate store level retail demand. Spatiotemporal population estimates have been identified as appropriate for many examples. These are some examples: retail store planning and commercial interests (e.g. Newing *et al.* 2013b); assessing population exposure to radiological hazards (e.g. Martin *et al.* 2014); healthcare provision; transportation modelling; national security (e.g. Bhaduri *et al.* 2007) and crime analysis (Malleson *et al.* 2010; Ceccato and Uittenbogaard 2013). As new sources of complex and big data continue to emerge it is likely interest in spatiotemporal population estimates will only get greater.

Chapter 7: Conclusions

7.1 Overview

This chapter summarises the findings of the thesis. It is structured in five sections. The first section provides a summary of the main research findings. This is followed by a reconsideration of the research aims set out in the introduction. The third section reviews the limitations and applicability of the work undertaken. The fourth section presents conclusions on the overall contribution of the research undertaken and the enhancements this may offer for future hazard risk assessment and spatiotemporal population estimates. The final section identifies and recommends areas for further research.

7.2 Main findings

Findings from the creation and application of time-specific population estimates for flood risk assessments are summarised below:

- Population exposure to flood risk fluctuates with time as populations move within zones of flood risk. Population exposure has been shown to vary depending on the time of day, day of week and season of the year.
- Spatiotemporal population fluctuations have a large impact on hazard risk analysis once it has been accepted that population is dynamic and constantly varying in both space and time.
- Baseline population levels (e.g. census estimates) can experience significant fluctuations. This is exemplified in the St Austell case study (Chapter 5) where the change observed is driven by temporary overnight visitors which peaks in August, aligned to the tourist season. In a fluvial flood risk analysis for an estimated return period 1 in 100 years (LISFLOOD-FP R100) the proportion of the population aged over 65 years exposed increases nearly four-fold between May and August. In this single flood risk assessment example it is demonstrated that population subgroups can be underestimated by up to 400% depending on whether a static census baseline is used or not. This has major impacts on any resultant flood risk assessment.

- The construction of a highly detailed population data library and spatiotemporal population outputs provides insights that are not available from using a single dataset alone.
- The daily transition of the 16-64 aged working population in Southampton (Chapter 3) towards the coast generates an increase in exposure to tidal flood risk during ‘typical’ working hours on weekdays, overnight fluvial flood risks dominate. This effect is less pronounced over the same period on weekends. It shows the different population subgroups experience different flood risks at different times.
- Flood inundation modelling concerning an embankment breach at Ulley reservoir (Chapter 4) suggest that the flood depths would have been dangerously close to overtopping the carriageway of the M1 motorway. Baseline population exposure to a breach at Ulley increases by up to 100% due to employment related circulation during a typical working weekday. Where vulnerable populations are under direct threat the effect of the population baseline used to adequately account for population exposure will have a large impact on subsequent risk analysis.
- Re-weighting the finest resolution UK census data (OA level) onto georeferenced postcodes to represent residential locations and workplaces permits an increase in previous spatial resolutions using SurfaceBuilder247 to 100 m.
- Using SurfaceBuilder247 to produce gridded outputs allows the integration of datasets published for different geographies and produces a time dependant gridded output. It has been shown that this can be effectively loosely-coupled with environmental datasets and models.

7.3 Evaluation of research aims

Three research aims, and five specific objectives were outlined at the beginning of this thesis (Section 1.2). Following the work presented they will be reviewed and evaluated in turn within this section:

> **Aim 1** Review the existing literature regarding the assessment of population risk to natural hazards and methods for determining exposure.

The first aim was to review the relevant existing literature regarding population exposure to natural hazards. The hazard and population data interpolation literature is a broad and well established field. The ultimate focus of this literature review was to examine the risks posed by hazards to people by considering spatiotemporal population estimates. Its purpose was not to deal with specific physical characteristics of hazards, although inevitably some of these were subsequently discussed in relation to the flood inundation modelling undertaken within the case study chapters (Chapters 4 and 5).

The review of existing literature confirms that the consideration of population distributions in space and time for improved vulnerability assessments is crucial (e.g. Cutter and Finch 2008). The limitations of using static population data, especially when it is constrained within arbitrary contiguous reporting zones, have been long established (e.g. Schmitt 1956; Openshaw 1984; Langford and Unwin 1994). These limitations have been cited by others in examples such as the LandScan USA project (Bhaduri *et al.* 2007) and by Aubrecht *et al.* (2013) where better representations of population distributions in time and space have been sought.

The literature review identifies an emerging field for the creation and application of spatiotemporal population estimates. The application for natural hazard risk and vulnerability assessments has been identified as one key area where significant enhancements can be brought. The work shown in this thesis contributes to this area through the bottom-up construction of high-resolution, in time and space, spatiotemporal population estimates to complement hazard risk analysis.

There are very few examples of true spatiotemporal population tools being developed around the world none of which are fully mature or fit within existing accepted frameworks. There are still larger conceptual and data issues to be addressed. The Population 24/7 approach has been developed within this thesis as one example of an extensible spatiotemporal framework that was deemed to be best suited to the UK flood hazard case studies presented.

> **Aim 2** Examine how population exposure fluctuates spatiotemporally to flood hazard events using census, administrative and survey datasets to construct a spatiotemporal population model applied to illustrative cases studies.

In the work presented within this thesis, the Southampton example (Chapter 3), Ulley (Chapter 4) and St Austell (Chapter 5) case studies have demonstrated that population exposure to flood hazard fluctuates considerably with time. A range of publicly available and subscription datasets such as census counts, education and employment registers, visitor surveys, healthcare statistics, retail data and tourism estimates were used to construct a data library to produce detailed spatiotemporal population estimates. A substantial amount of preparatory data modelling and analysis was required to gain the most from currently available datasets. Data linkages were also created to combine census and business employment register counts with UPCs to increase the modelling output resolution to its current potential within the example illustrated. The Population 24/7 model has been developed in various new ways to account for tourist populations, retail footfalls and leisure activities.

The application examples have demonstrated population variability at a range of spatial and temporal scales. Spatially, predominantly urban (e.g. Southampton and Ulley) applications have been contrasted with the rural St Austell example. Through enhancement of the Population 24/7 approach improvements in spatial gridded resolution from 200 to 100 m have been achieved compared to Martin *et al.* (forthcoming) and Feilding's (2007) (non-temporal) original implementations. The range of case studies selected fulfilled their purpose to demonstrate population fluctuations across differing temporal scales. The applications have shown temporal variation in population exposure to flood hazards at daily, weekly and seasonal scales.

Despite being a predominantly rural study area with urban centres, the St Austell case study shows the greatest population variation in terms of net fluctuation and magnitude change from the usually resident baseline. Compared to the urban example presented, where flood risk assessments often dominate, St Austell perhaps shows the most interesting impacts of spatiotemporal population estimates for hazard risk analysis.

The precise numeric values of the exposure estimates produced are subject to the same scrutiny, assumptions and interpretation as any other example. They do not attempt to profess 'the' definitive answer, if there even is one at all. However, the data presented have been based on the best information available for this study with rigorous interrogation. The purpose of these examples, using numeric values as a guide, is to demonstrate a credible framework to account for population variation in time and space for hazard based applications.

> **Aim 3** Assess the extent to which spatiotemporal population modelling techniques can be used to provide greater insights for integrated disaster risk management. Discuss to what extent confidence can be placed in their results and outline the challenges for validation.

The illustrative examples presented are part of a much wider class of spatiotemporal population modelling approach that demonstrates the enhancements achievable within disaster risk management. This thesis has shown the development of Martin *et al.*'s (forthcoming) framework within the spatiotemporal hazard concepts identified by others and included within the literature review.

There are many advantages associated with a greater understanding of time-specific population distributions in often time critical hazard scenarios. Firstly, the basic requirement to understand the potentially exposed population is a fundamental emergency preparedness principle. The examples and techniques developed within this thesis show spatiotemporal changes in population and hazard intensity all have the potential to radically change how hazard risk assessments are conducted. It has allowed the identification of specific locations or population subgroups that are particularly vulnerable such as the elderly. The technique shows that the location of such subgroups fluctuates in spatial distribution and intensity with time. These dynamic insights are simply not available when using aggregate static population data alone. This spatiotemporal modelling technique permits the combination of crucial datasets that are relevant to the study area or hazard scenario and the gridded output produces a useable universal format. It equips the user with a method to combine multiple datasets that are often produced in incompatible spatial units in their original publication. This is often the case because datasets are usually produced to serve a certain purpose and conform to the original producer's desired reporting zones (usually also associated with data confidentially concerns). However, these datasets can often serve valuable alternative purposes. This research has identified relevant datasets and brought them together to produce spatiotemporal population estimates where no single dataset has sufficiently done so before.

There are instances where advanced planning for the worst case scenario or taking the 'business as usual' population approach for a hazard event of a given magnitude is useful. The evacuation of 600 students from seafront residences in Aberystwyth, UK, in January 2014 (Gevertz 2014) due to fears of imminent coastal flooding demonstrates this principle well. In this example Aberystwyth is an isolated university town where students comprise a significant proportion of the population and alternative accommodation is severely limited. University policies to restrict student's cars from being brought into the town and the coastal railway as the only main transport connection increase vulnerability under these circumstances. In terms of health and physical ability, younger students may be considered one of the least vulnerable population subgroups. However, under this unique combination of circumstances that actually occurred, they suddenly become susceptible. Emergency preparedness accounting for spatiotemporal population variations, such as the Population 24/7 approach implemented in this thesis, provides a powerful tool for risk analysis by producing advance insights. The specific consideration of population subgroups, daily and seasonal cycles (e.g. term vs. non-term time when students would not be expected to be present) that have been demonstrated within this thesis have great impacts for any risk analyses undertaken.

During the first applications of innovative techniques validation can be challenging as already discussed in Section 6.4. It is intended in the future that newly emerging datasets (see Section 7.4) will continue to strengthen and provide additional validation opportunities for these complex spatiotemporal questions. However, the estimates provided within this thesis are based on known robust administrative, survey and census counts of population occupancy for specific locations which adhere to predictable cycles. It is acknowledged that these estimates are subject to variation based on external factors such as the weather, economy and local decision making. A high degree of confidence is placed in the primary temporal cycles observed and affirmed by the underlying data on which they have been constructed. It is the wider ability to resolve these cycles for the purposes of emergency planning that is deemed of significance, rather than the micro level variation in the values presented within the illustrative case studies contained within this thesis.

7.4 Limitations and applicability

The potential limitations of the technique proposed have been openly addressed, particularly around validation. Like any method, in the current form presented here, it may be more appropriately suited to some applications more than others. This is the same for a physical tool or device that is designed to fulfil a particular task. All current attempts to model time-specific populations (e.g. LandScan USA) are subject to the same validation difficulties. However, this alone should not limit the valuable applications achievable through developing such models. Potential enhancements and important questions for flood risk analysis relating to population exposure is just one example. This section summarises two areas which concern application scale and data availability where additional consideration for application may necessary.

7.4.1 Scale of application

Following the completion of the original Population 24/7 project 200 m population estimates covering the whole of England and Wales were made available online (http://pop247.mimas.ac.uk) for selected temporal intervals. This demonstrated an example of national coverage. It would be very time-consuming to replicate the work presented here in its current format at the same scale for hazard applications. It is proposed that the scale of application achieved by the work presented here is an intermediate compromise. In order to scale this application up to the national level, assumptions already made, or decisions that reflect the unique characteristics to one particular region concerned, could potentially become inflated. For example it might not be possible to justify the decision that the retail or workplace behaviour of residents in St Austell is the same as that for a small Welsh market town or for the centre of London. Larger funded studies may be able draw on government, commercial and transportation datasets that could not be used in this thesis in order to gain empirical evidence and reduce assumptions.

At the other end of the spectrum the 100 m resolution achieved is based on the detail of available input datasets. A feature of the flexible structure of the Population 24/7 framework is the variable grid. Any grid size could be selected, but this decision is ultimately driven by the availability of suitable input data. It has been clearly stated that even at the 100 m output resolution, it was not the aim of this research to predict individual human behaviour. Therefore, it was not intended to produce an isolated street-level estimate, but rather to combine cellular uncertainties within a much more robust city level approach. This has the effect of mitigating for small area (sub or inter grid cell fluctuations) when decisions are made at higher scales.

7.4.2 **User access to data**

Where possible the applications presented within this thesis have sought to use open data. However, there have been occasions where commercial or datasets requiring institutional subscriptions have been required. Where these have been used it has been to prevent compromises that may undermine the credibility or output resolution achievable through these examples. One constraint of this approach is therefore the access to appropriate datasets to derive the output detail required. This remains a challenge for applications within countries with limited data provision or quality, although as demonstrated the modelling framework exists. The recommendations for further research (Section 7.6) suggest that future prevalence of innovative datasets may offer alternatives and new opportunities for spatiotemporal model construction.

7.5 Contribution

This research has made a tangible contribution to the development of spatiotemporal population estimates and demonstrated relevant enhancements within the hazard risk assessment field. It has produced spatiotemporal estimates that are arguably amongst the first of their kind, certainly for the locations modelled. It has acknowledged the associated challenges but continues to pave the way for future developments. This research has significantly enhanced the Population 24/7 approach beyond the first implementation and work completed for the original project. Expansion of the population data library and improvements in resolution are just two examples contributed by this work. However, the greatest contribution is the enhancement of the method to handle data for time-specific population applications.

The spatiotemporal methods implemented within this study can be extended and further developed to build appropriately calibrated time-space population models from the wealth of data available. This includes data that might become accessible in the future. The Population 24/7 technique implemented provides a method to handle data with a fully extensible framework. It is the ability to handle spatiotemporal data to produce meaningful results that is unique. New datasets will continue to arise and almost certainly will eventually be surpassed, but it is the flexible framework to capture and adapt to such information that is important and demonstrated here.

The research presented has demonstrated a challenging first step of combining time-specific population data with environmental model outputs using a loose-coupling approach. This has developed a pathway aiming towards truly integrated dynamic population and environmental models. This research has tested the boundaries for the integration of population and environmental data by successfully tackling the fundamental issues of data resolution (spatial and temporal) and format.

This research has contributed, through the publication of original research papers (e.g. Appendix A), to the development of this field. It has tackled current UK and EU priority issues on population exposure to hazards and flood risk assessment, but these are of greater significance beyond national boundaries. This research contributes to a greater conceptual issue on representing dynamic populations. It has identified relevant work being undertaken by others, sometimes for entirely different purposes, and successfully integrated them to enhance the spatiotemporal population estimates presented. Finally, it has sought to introduce a practical method that may change how populations exposed to hazards are considered.

7.6 Recommendations for further research

Future advances in the release of new administrative datasets, crowd sourced geographic information (e.g. Open Street Map), the open data movement (e.g. EU Open Data Portal, data.gov.uk) and prevalence of big data (e.g. Zhong *et al.* 2014) will inevitably provide opportunities to enhance this research in the future. Some of these datasets are already emerging but they are often focused on large cities and urban areas (e.g. London Datastore; Dublin Dashboard; Batty 2013).

Work undertaken by others has already shown the potential in emerging data feeds such as public transport travel-card data, cycle hire scheme usage statistics and analysis of georeferenced tweets (e.g. http://www.bartlett.ucl.ac.uk/casa). Tafazolli (2014) proposes that development of the fifth-generation (5G) of mobile data networks will enable properly connected smart cities in the future. This is likely to produce even greater opportunities for innovative or real-time data that may inform population movements.

Work currently being undertaken by Deville *et al.* (2014) has already demonstrated the ability to undertake dynamic population analysis using mobile telephone data with proven accuracy. This has shown an alternative method capable of mapping population movements in space and time. It is believed that this type of analysis could go some way towards validating the spatiotemporal population estimates presented within this thesis, although does still only represent the population with a phone that is in their possession. This information could also be used to enhance existing spatiotemporal models by integrating the intelligence gained from temporal signals within the dataset.

The new UK ESRC Administrative Data Research Network and Consumer Data Research Centre may provide future opportunities for streamlined access to linked or previously unavailable data. These opportunities and those already discussed within this section can potentially further develop data libraries for spatiotemporal population modelling or provide validation opportunities.

Options have been considered to advance the research presented here further. These explore the possibility to utilise UK mobile telephone data as a validation technique and working with non-academic partner organisations on the role of spatiotemporal population estimates in their hazard analysis. There is an interest to develop the model's software interface to harness real-time and 'big' data feeds. In relation to flood hazard, these could be in the form of reports that provide local area updates such as evacuated areas. There are many applications for accurate spatiotemporal population estimates, beyond the flood risk examples presented here. Potential sectors for further applicability could include retail, resource allocation, public health and network and transport planning. The research that has been presented demonstrates the wide potential benefits for detailed time-specific population estimates.

Appendix A Published research paper

Appendix A Published research paper

Accepted research paper in Applied Spatial Analysis and Policy:

Smith, A.D., Martin, D. and Cockings, S. (2014 online first) 24/7 Population modelling for enhanced assessment of exposure to natural hazards. *Applied Spatial Analysis and Policy*.

Full text open access: http://dx.doi.org/10.1007/s12061-014-9110-6

Appl. Spatial Analysis
DOI 10.1007/s12061-014-9110-6

Spatio-Temporal Population Modelling for Enhanced Assessment of Urban Exposure to Flood Risk

Alan Smith · David Martin · Samantha Cockings

Received: 23 October 2013 / Accepted: 2 July 2014

Abstract There is a growing need for high resolution spatio-temporal population estimates which allow accurate assessment of population exposure to natural hazards. Current approaches to population estimation are usually limited either by the use of arbitrary administrative boundaries or insufficient resolution in the temporal dimension. The innovative approach proposed here combines the use of a spatio-temporal gridded population model with flood inundation data to estimate time-specific variations in population exposed to natural hazards. The approach is exemplified through an application centred on Southampton (UK) using Environment Agency flood map inundation data. Results demonstrate that large fluctuations occur over time in the population distribution within flood risk zones. Variations in the spatio-temporal distribution of population subgroups are explored. Analysis using GIS indicates a diurnal shift in exposure between fluvial and tidal flooding, particularly attributable to the movement of the working age population. This illustrates the improvements achievable to flood risk management as well as potential application to other natural hazard scenarios both within the UK and globally.

Keywords Spatio-temporal modelling · Population surface modelling · Natural hazards · Vulnerability · Urban exposure

Introduction

Natural hazards are one of the ultimate constraints on human activities (Mitchell 1999) and occur on a variety of time scales from sudden onset events such as earthquakes to gradual processes such as drought. A natural hazard has the potential to develop into a disaster as soon as it poses a risk to humans. Increasing incidence and frequency of natural hazards events during the last century have been attributed to better recording and increased habitation of hazardous areas (Alexander 1993; Tobin and Montz 1997; Hilhorst and Bankoff 2008). Humans are becoming more susceptible to hazards as a

A. Smith (✉) · D. Martin · S. Cockings
Geography and Environment, University of Southampton, Southampton SO17 1BJ, UK
e-mail: Alan.Smith@soton.ac.uk

Published online: 01 August 2014

consequence of population growth and urbanisation (Huppert and Sparks 2006). In particular, flood losses are increasing in major coastal cities (Hallegatte et al. 2013). Prominent events since 2011 have refocused attention on natural hazards and human vulnerability both in terms of fatalities and economic impacts. By a considerable margin, 2011 was the mostly costly year to date with natural hazards losses totalling US$380 billion and 27,000 fatalities (Munich Re 2012). The most notable events included the Japanese tsunami and Christchurch earthquake as well as severe flooding in Thailand, Australia and the USA and continued drought in east Africa.

To better understand the effects of natural hazards and develop more robust emergency plans, an improved knowledge of the spatial and temporal distribution of population is required (Bhaduri et al. 2007; Fielding 2007; Aubrecht et al. 2012a; Harper and Mayhew 2012). Calculating the population exposed is not straightforward as both the hazard and population vary over time (McPherson and Brown 2004). Inadequacies in mapping population have been noted for many decades (e.g. Schmitt 1956). Commonly used official population datasets such as census or population registers usually provide only residential 'night-time' population counts. In the United Kingdom (UK) it has been estimated that 5.2 million properties are at risk from flooding (DEFRA 2011), but there is little understanding of the numbers of people at risk. This paper highlights the significant improvements that can be achieved in assessment of population exposure to natural hazards by the integration of innovative spatio-temporal population modelling methods with maps of hazard extents. It demonstrates the importance of high-resolution space- and time-specific population data and illustrates their utility in a UK flood risk case study. The analysis is implemented using the SurfaceBuilder247 software (Martin 2011) and Geographic Information Systems (GIS).

The remainder of the paper is structured as follows. The next section provides an overview of flooding in the UK and the proposed spatio-temporal population modelling approach. The third section describes the source datasets and modelling methodology. Section four introduces the study site used for the combination of flood hazard and spatio-temporal population modelling. This is followed by the results and discussion in sections five and six. The final section presents some conclusions and suggested further work.

Background

For the first time in human history more people now live in urban than rural areas (United Nations 2008). Rapid urbanisation and the concentration of population in flood prone areas has increased human susceptibility to flood-related hazards. New building and redevelopment in urban areas has increased ground surface sealing and encroachment into low lying land and areas at risk of flooding. A reconsideration of risk management policies is required in order to account for population at varying spatio-temporal scales and reverse the trend of increasing urban flood risk (Zevenbergen et al. 2008). Northern Europe, including the UK, has a long history of severe coastal flooding (Ruocco et al. 2011). Qualitative descriptions of these events and their impacts on local populations are contained within numerous historical records. The most catastrophic sudden onset event to occur in the UK during the last 500 years was the 20 January 1607 flood in the Bristol Channel (Horsburgh and Horritt 2006; RMS 2007) which led to many fatalities in Southwest England and South Wales. The UK continues to have a record of significant destructive fluvial (Marsh 2008; 2004), tidal (Baxter 2005) and flash (Bettess 2005; Burt 2005; Murray et al. 2012; Rowe 2004; McGinnigle 2002) flood events.

Prominent flood events over the last decade in the UK and Europe have had a significant impact on legislation and policy. Major natural hazard incidents have been recognised by the UK government as one of the highest priority risks facing the country (HM Government 2010). The Pitt review (Pitt 2008), which was commissioned following substantial UK-wide flooding in June and July 2007, resulted in 92 recommendations. Key recommendations concerning flood risk assessment and management included: calls for local authorities to collate and map the main flood risk management and drainage assets (Recommendation 16); encouragement for individuals to be better prepared and become self-reliant during emergencies allowing the authorities to focus on the people most in need (70); and the impact of flooding on the health and wellbeing of people to be monitored and mitigations put in place to manage these effects (74). Recommendations 70 and 74 illustrate the need for a greater understanding of exposed populations, while recommendation 16 concerns mapping the physical infrastructure. Improved mapping and modelling of both the hazard and the population exposed are required for effective emergency planning and management.

The European Union Floods directive of the European Parliament (OJ 2007) has also had implications for the way flood risk is managed in the UK. Inter-agency and international collaboration remain key components in flood risk management. Member states were to have provided flood hazard and risk maps by 22 December 2013 and flood risk management plans by 22 December 2015 (OJ 2007). The Flood Risk Regulations 2009 implement the requirements of the directive in the UK. The UK Environment Agency is responsible for producing the relevant hazard and risk maps for England and Wales within the specified timescale and lead Local Flood Authorities have been identified to map local flood risk.

The spatial and temporal distributions of population and natural hazards are complex and occur at a range of scales. The population exposed is a key component when defining risk to natural hazards, as is vulnerability of specific population subgroups. Variation in the location and magnitude of vulnerable populations exposed to natural hazards at different times needs to be recognised if emergency planning is to be improved (Cutter and Finch 2008). Lack of data or inadequate sophistication in either, or both, of the spatial and temporal aspects of population estimation have seriously hampered previous efforts to understand such characteristics. Some researchers have quantified population exposure based on properties exposed and average occupancy rates (e.g. Hunter et al. 2008; Hall et al. 2005; Wadey et al. 2012). The methodology presented here extends population exposure estimation into the temporal dimension. The paper is not concerned with the definition and investigation of vulnerability and resilience per se, nor does it attempt to make predictions of perceptions of risk or human behaviour during emergencies, all of which are also important; rather, it demonstrates how enhanced spatio-temporal estimates of population can facilitate the identification of vulnerable populations exposed to natural hazards. Such estimates can then inform improved risk and emergency management plans.

The hazard component also has complex spatio-temporal characteristics, which may be represented by simple static risk maps or dynamic models. In order to make accurate risk assessments, knowledge of the likely onset time, severity, duration and geographical extent of a hazard scenario are required. For example, while a storm surge may have a long lead-time a tsunami may not.

The detailed spatial distribution of population is often obscured by the convention of displaying uniform densities in contiguous zones, such as census areas, that cover the entire surface (Martin and Bracken 1993). Similarly, the spatial distribution of natural hazards does not conform to administrative boundaries. Any analysis using geographically aggregated data is prone to the modifiable areal unit problem (MAUP) Openshaw (1984), although the effects of MAUP can be minimised by using data at the finest possible spatial resolution. Dasymetric methods can improve the spatial accuracy of interpolation via the use of ancillary datasets and have been successfully employed in population surface estimation (e.g. Mennis 2003).

Interpolating data onto a grid is a further method which can offer some resilience to the problems associated with irregular zonal representations. Benefits include the inclusion of unpopulated cells; the standard size and shape of individual cells; the invariant nature of the cell boundaries over time and the relative ease of incorporating data from statistical models of physical phenomena. Examples of gridded population models include the 1 km disaggregated gridded population for Europe (Aubrecht et al. 2012b) and gridded population of the world (CIESIN 2013).

Although gridded population models provide many advantages for estimating population exposure to natural hazards they are still only a static representation of population at a given point in time. The method proposed in this paper offers flexibility and granularity in the temporal dimension while building on the inherent advantages of the gridded model, but a gridded representation is not a pre-requisite for its implementation.

Methods and Data

The methodology proposed here contains three subcomponents, shown in Fig. 1. Figure 1a covers the spatio-temporal population modelling process, 1b the hazard component and 1c brings these GIS-based datasets together for analysis. Often natural hazard and population models are not linked. The proposed methodology aims to demonstrate the improvements to risk analysis when these subcomponents are more fully integrated.

The spatio-temporal population modelling component (Fig. 1a) is illustrated schematically in Fig. 2, which shows a small area containing three locations a, b and c at two times, T_1 and T_2. For each time, the left hand diagram shows population density on the vertical axis and space on the horizontal, while the right hand diagram shows the same population distribution in map view. In this highly simplified example, T_1 represents night time, when population is present at locations a and b, which are the centroids of residential areas of differing spatial extent, but no population is present at location c which is a workplace, only occupied during the day. In the absence of exact areal extents for every possible population location, it is necessary to employ a distance decay function around points of high information such as population-weighted centroids. We here use the Cressman (1959) distance decay function, established in this context (Martin 1996). T_2 represents day time, when lower population densities are present at residential locations a and b, but the workplace c displays a high population

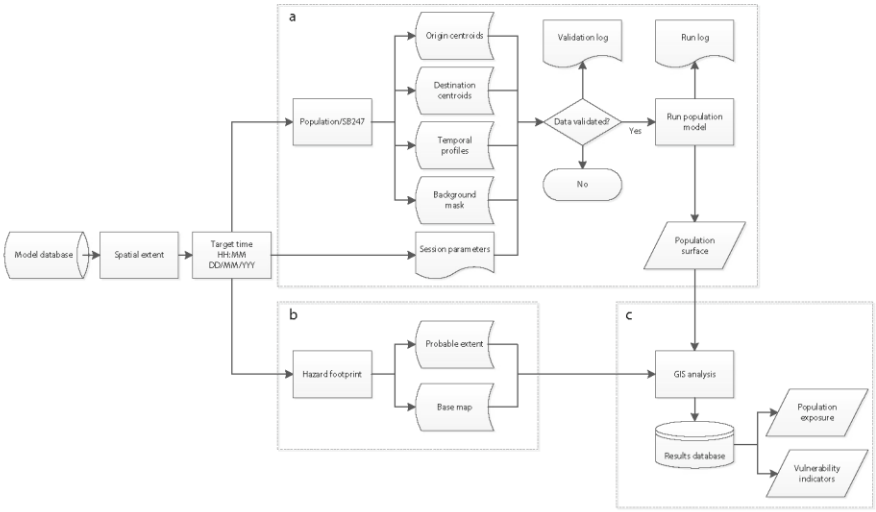


Fig. 1 Overview of analytical operations to assess natural hazard exposure

density over a relatively small spatial extent. Between these times, the study area may have exchanged commuters with neighbouring areas and the total population contained at times T_1 and T_2 therefore need not be the same. At either time period, the population present may be aggregated to any relevant mapping units, such as the cells of a regular grid. (Martin 1989, 1996) presented a surface modelling algorithm which mapped population from population-weighted centroid locations onto a regular grid by means of adaptive kernel density estimation, with the local density of centroids determining the spatial extent over which population is spread. In common with most conventional census mapping, this approach effectively reproduced only the night time residential situation shown here at time $T_{1,}$ without any explicit consideration of reference times.

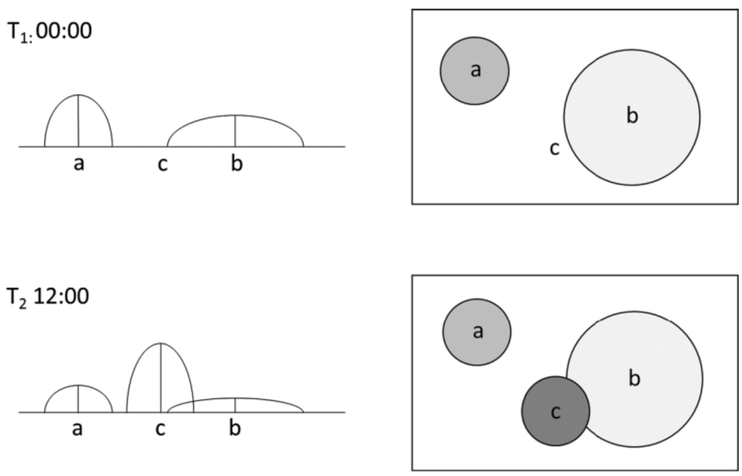


Fig. 2 Schematic of population centroid density represented by vertical density curves (left) and spread (right) at two given times of the day

Appendix A

In this paper, we employ an extension of this approach using the SurfaceBuilder247 software described in Martin (2011), which is a development of the original model and which uses a time profile of human activity at each centroid rather than a single value. The same algorithm is employed for redistribution of population across space, but is here undertaken with reference to the populations anticipated to be present at each centroid at a specified target time. In the SurfaceBuilder247 implementation, centroids are designated as either origins or destinations. The total population to be modelled is determined by the sum of the populations at the origin centroids within the study area, such as a and b in Fig. 2.

Processing begins with the establishment of a full dataset of origin and destination locations for the chosen study area and a surrounding region to reflect anticipated commuting flows. Detailed file specifications and data field definitions are provided in Martin (2011). Censuses typically provide an accurate and comprehensive small area population count and the smallest spatial units for which detailed census data are available in England and Wales are output areas (OAs), with a mean population size of 300. Population weighted OA centroids from the 2001 census, updated with data from 2006 mid-year estimates, have been used as the origin centroids in this study. It should be noted that the full range of equivalent data from the 2011 census had not been published at the time this research was undertaken.

Secondly, destination centroids represent the range of possible non-residential locations of population. A population capacity is assigned to each destination centroid, such as the number of staff and students at a school or patients and staff at a hospital, obtained from administrative data sources. Any number of destination centroids may be introduced to represent additional locations of human activity which do not contribute to the residential population base. The spatial extents and catchment areas of destinations are specified, which indicate, respectively, the size of the site and the geographical range from which population should be drawn. Thus (for example) a primary school could be represented as a destination location, with a time profile in which all children on the roll are present during the school day, allocated to a small site of extent 100 m and drawn from a catchment radius of 3 km. A dataset of destination locations was initially collated for the Population 24/7 project,[1] which includes places of work, education, and health care. The example presented here extends these datasets by building on this original library of potential population locations. It is a feature of the model's structure (Fig. 1a) that users can develop their own datasets, particularly for non-residential locations. Time profiles, the third component in Fig. 1a, have been developed for each destination centroid, based on a range of data sources describing the opening hours of different workplaces and services. The time profile derived from ancillary information provides the proportion of a site's capacity population which is present at any specific time.

The development of retail destinations is provided here as an example of how additional activity types can be added to the Population 24/7 destination datasets. This example serves to exemplify the process of dataset creation and is just one of many destination types that could be represented. Retail was chosen as it is another major activity type which attracts significant non-residential populations and was not explicitly included in the original Population 24/7 estimates. This example uses national time

[1] Population24/7 project details: http://www.esrc.ac.uk/my-esrc/grants/RES-062-23-1811/read

use survey data to estimate the total number of shopping trips and then allocates these to destinations based on the recorded size of the retail workforce in each area. Equivalent retail data of higher accuracy could be derived from commercial sources as part of a large-scale implementation but were neither available nor necessary for the purposes of this example.

A temporal profile for shopping activity was created as a simple estimate of retail activity (Fig. 3) by analysing episodic diary data from the Time Use Survey (TUS) (Ipsos-RSL, and ONS 2000). Respondents (n: 20,981) kept a coded diary entry for 10 min intervals throughout the day. Analysis of the retail components within the TUS data identified average shopping durations and travel times. These were both used to inform travel to catchment parameters and length of retail trips by individuals. In line with expectations, it can be observed that people shop for longer on a Saturday and travel further. These data produced estimates of the total numbers of shoppers and time spent shopping.

The number of retail employees at the Lower Layer Super Output Area (LSOA) geography was derived from the Annual Business Inquiry (ABI) dataset (ONS 2006a). The size of the retail workforce was then used to allocate total retail activity from the TUS to approximate retail locations based on business postcodes (ONS 2006b). LSOAs are the next smallest statistical units in England and Wales after census OAs and typically contain 1,500 residents. They are the smallest units for which ABI data are available. Postcodes are georeferenced locations designed to facilitate the delivery of mail and businesses receiving large amounts of mail will generally be allocated their own postcode. It is acknowledged that large user postcodes will also include non-retail offices and business locations, but this simple approach nevertheless permits the reallocation of large population volumes from residential to retail locations during shopping hours. The total shopping time from the time use survey and retail employees from ABI were combined to estimate the number of shopping hours supported by each retail employee, taking account of the demand pattern for each day and the variable duration of shopping activities. Estimated population numbers engaged in retail activity are thus allocated to each business postcode by time of day.

A background raster GIS data layer constrains the population distribution, delineating areas of uninhabitable space such as the sea and other water bodies, and incorporating the principal road network weighted according to UK Department for Transport

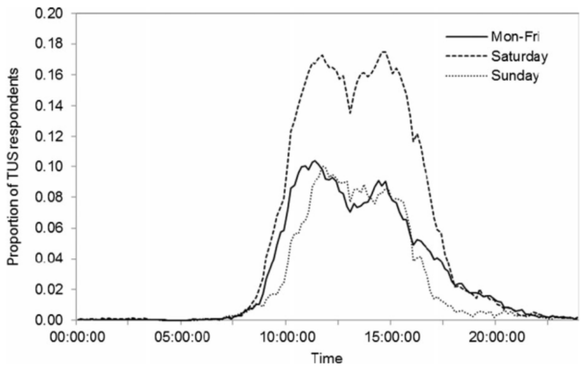


Fig. 3 Retail temporal profiles according to TUS (2000) respondents who engaged in retail activity

traffic count data (DfT 2013). Average annual daily flows for each DfT survey point by vehicle type are redistributed according to time of day, day of week and road category (motorway and principal urban/rural roads) according to the National Transport Model (DfT 2005). This reproduces observed temporal variations in traffic flow throughout the week. The traffic count by vehicle type is subsequently converted to a population estimate for each rasterized road cell based on specific vehicle occupancy rates (DfT 2009). The population estimated to be travelling at any target time is assigned onto this weighted background layer.

For this study, population data are modelled onto a regular spatial grid with 200 m cell size which is appropriate to the resolution of the available input data sources. The Population 24/7 methodology has been designed to allow subgroups of the population to be modelled separately. In the application described here, modelling has been undertaken for seven age groups: pre-school aged children, primary school age, secondary school age, further education students, higher education students, remaining working age population and those retired. These subgroups have distinctive temporal patterns such as typical school hours and term times. The modelling technique is volume preserving. Population counts derived from the origin locations are redistributed across origin and destination locations and the background layer, ensuring that the overall population total is preserved within each modelled age group. Known in- and out- flows to the system, such as long-distance commuting, can be incorporated into the model but have not been applied in this case study.

Figure 1b relates to the natural hazard modelling component. In the example application presented here (Environment Agency flood map (for July 2012)) (EA 2012) are employed to delineate the spatial footprint of the flood hazard. The flood map is the result of probabilistic and scenario-led hydraulic modelling. The most likely scenario under the ‘zone three’ extent (high probability) (EA 2012) has been utilised. Its spatial extent is shown in Fig. 4. This scenario models inundation caused by fluvial and tidal flooding with a 1 % and 0.5 % annual probability of occurrence respectively. The modelled population surfaces (Fig. 1a) and flood extents (Fig. 1b) have then been integrated using GIS (Fig. 1c).

The modelling framework allows exploration of variation in population exposure to flooding over a range of temporal scales e.g. hourly, daily or seasonally, broken down by different subgroups of the population. In this paper, the populations for a typical working weekday and a Sunday during school term time for 2007 are modelled and population exposure estimates within the potential flood extent are calculated for two-hourly intervals between 08:00 and 20:00. A static baseline exposure estimate using rasterised 2001 census data has also been calculated for comparative analysis.

Example Application: Study Area

A 25 × 25 km study area centred on Southampton Water (Fig. 4) has been chosen for an example flood risk application. The Solent separates the Isle of Wight from southern England and provides a natural deep water channel for large shipping vessels. The region’s industrial and shipping success has been attributed to the complex tidal system, which produces a double high-tide, but the combination of the region’s topography, location and tidal system also has the potential to dramatically increase the flood risk

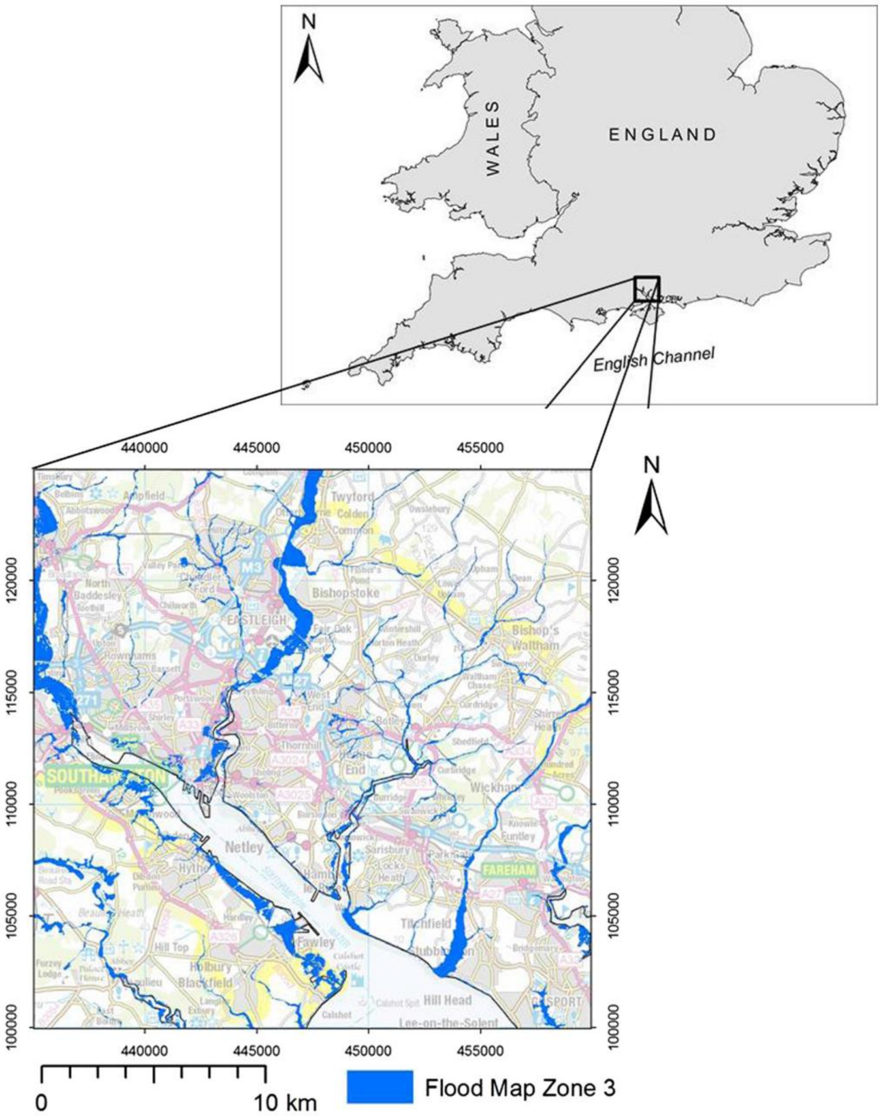


Fig. 4 Application study area centred on Southampton (UK) showing potential flood inundation

within the area. Southampton Water is a narrow funnel-like channel leading from the Solent to the Port of Southampton, vulnerable to storm surges driven by low pressure systems which can lead to increases in sea level of up to 1 m (Ruocco et al. 2011) or north sea surges that propagate through the English Channel (Wadey et al. 2012). When combined with high spring tides these events pose a heightened flood risk and exert pressure on existing defences. Historical record analysis by Ruocco et al. (2011) discovered up to 20 flood events in Southampton since 1935. Approximately 24,000 properties are considered to be within the tidal flood plain of a 1 in 200 year flood event in the Solent region (NFDC 2009). Within the study area, approximately 78 km^2 is at risk of tidal or fluvial flood inundations under the zone three scenario (shown in blue in Fig. 4).

The study area and its low lying urban areas contain major coastal transportation links, population centres and commercial and military ports. These activities have contrasting spatio-temporal patterns, making the area of particular interest for spatio-temporal population modelling. A major employer is the Port of Southampton which is the country's busiest cruise ship terminal and second largest container dock (ABP 2013). On a busy summer weekend, the central area population may be swelled for example by tens of thousands of passengers and crew on-board cruise ships visiting the port (not modelled here) and by visitors to the region's major shopping centre which has an average weekly footfall of c. 300,000 (WestQuay 2011).

Estimates of potential exposure to the hazard are calculated for various subgroups of the population for a range of time slices. Students in higher education are recognised as a key spatially mobile section of the population with strategic importance (King and Ruiz-Gelices 2003). The predominant local authority (Southampton) modelled in the study area for this paper had a 2011 residential population count of 236,882, including 32,111 (13.6 %) full-time students in higher education. The two universities within the study area (University of Southampton and Southampton Solent University) had a combined 2011/12 student population of 38,885 (HESA 2012). Students account for a significant proportion of population flows and seasonal variation. Notably, two evacuations of the student population in Aberystwyth, UK, during severe coastal flood events in January 2014 (Gevertz 2014) has demonstrated their vulnerability and strategic importance, even if they are generally considered to be a relatively resilient sector of society.

Results

Figure 5a shows total population for rasterised 2001 census OAs and the population modelled in 200 m grid cells for three different times of day (Fig. 5b to d). There is a stark difference between the conventional static census based population model in Fig. 5a and the gridded representations in 5b-5d which much more accurately indicate the higher central densities and extensive unpopulated areas, even in this relatively urbanized region. A large daily variation in population occurs. During the working day (Fig. 5c), population becomes highly concentrated in specific areas such as the city centre and in local clusters such as schools and colleges, as employees and students travel to, and temporally remain at, places of work and study. Population also increases in other non-residential areas of the city associated with activities such as retail and transport. The 08:00 model differs from the 20:00 model with more people in the transportation network at 08:00 – mostly on their way to work or school (Fig. 5b and d).

There are spatial (Fig. 5) and temporal (Fig. 6) variations in populations potentially exposed to flood hazards during the day. Analysis of the results suggests that the total population exposed peaks towards the end of the typical working day (Fig. 6). Differentiating the flood risk components (Fig. 6) reveals a further interesting phenomenon within the Southampton study area: throughout the day, exposure to fluvial flood risk

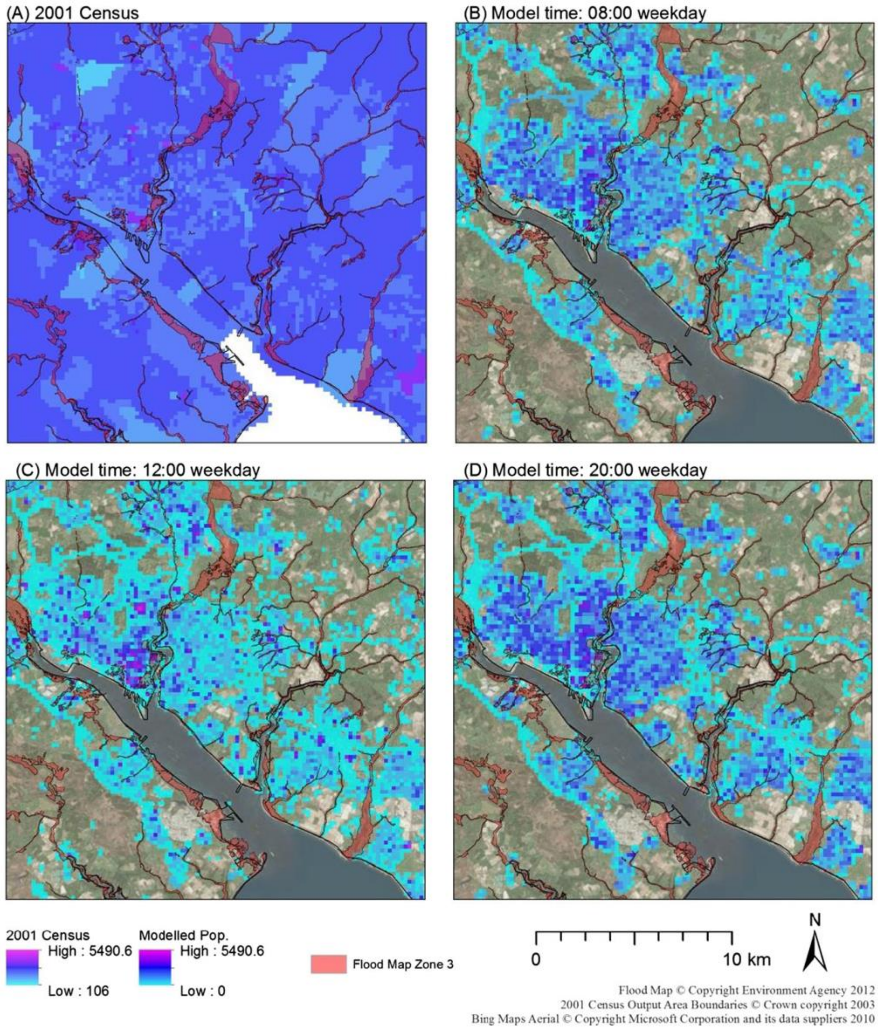


Fig. 5 Rasterised 2001 census output area total population counts (**a**). Spatio-temporal model outputs showing total population for three time slices, together with Environment Agency flood map data (**b** to **d**). All maps at 200 m resolution for a 'typical' weekday

closely resembles the reciprocal of tidal exposure with a symmetry approximately aligned to typical working hours (09:00–17:00). This is attributable to the concentration of industry and commercial activity in coastal locations compared to large residential areas inland close to major rivers (Itchen and Test). As employees commute to the coastal regions during the day their tidal flood exposure increases; when they return home to residential locations further inland in the evening fluvial flood risk becomes the dominating factor. This spatio-temporal variation contrasts strongly with the static representation of population exposure produced by the traditional census (also shown in Fig. 6), which generally tends to overestimate exposure.

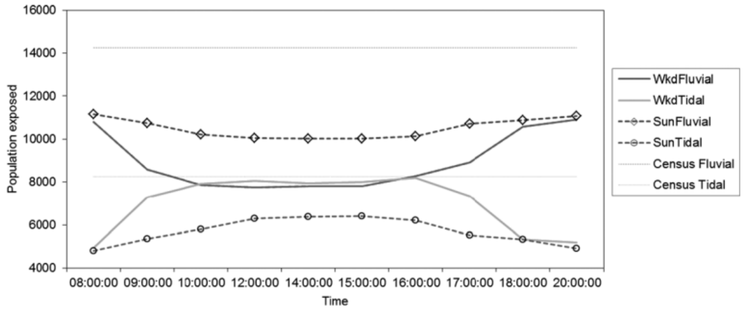


Fig. 6 Estimated total population exposure by time of day for flood map zone three comparing the static census, a modelled 'typical' weekday and a typical Sunday. Where the annual fluvial flood risk is 1 % and tidal 0.5 %

Figure 7 indicates the total population exposed to fluvial and tidal flood risk at 12:00 and 20:00, showing the composition by age group. The differences in exposure between tidal and fluvial flooding at different times of day are most notable within the working aged population (16–64). It can also be seen from the heights of the bars that the total population exposed fluctuates over time with tidal exposure at 20:00 being the lowest.

Further analysis of the population engaged in different activities at representative times of 08:00 and 12:00 has been conducted for the working age population (Fig. 8a to c). The figure shows the population in these groups who are travelling in the transportation network (*in travel*) or at a non-residential destination site (*on site*). These were selected to demonstrate the utility of the method for analysing the distribution of population during a typical morning commute and middle of working day. At 08:00 (Fig. 8b) it can be seen that the working age population density at non-residential destinations is relatively low, compared to Fig. 8c which reflects the concentration of this subgroup at workplaces and other destinations during the day. As expected, when more people migrate into a hazardous zone, exposure is dramatically increased within this age range during the day. Figure 8d shows a large increase in the working age population in both potential exposure to fluvial and tidal flood risk at midday when this population is typically *on site* at a place of work. The distribution of the university student population has also been examined for the same typical term time weekday (Fig. 9). During the day, the student population is concentrated on the city's two universities and spreads back into the student residential areas during the evening.

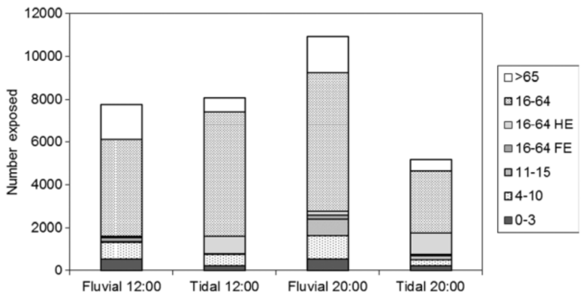


Fig. 7 Age composition of population potentially exposed to fluvial and tidal flooding during the working day (midday) and evening (20:00) HE: Higher Education; FE: Further Education

The daily variation of the student population that only occurs during term time demonstrates how their potential susceptibility is time dependent (Fig. 6).

The diurnal trends in the modelled data (Fig. 6) might have been predicted, but can be isolated and quantified using the methodology proposed. During 'typical' weekday working hours there is a shift from fluvial to tidal flood risk. It was possible to narrow this change down to the working aged population (Fig. 7). Population exposure was further analysed for representative times of midday and 20:00. This provided insight into two contrasting points in the usual daily cycle. The reversal from tidal to fluvial exposure occurs in the 16–64 working age population. At midday tidal risk is the predominant risk to this population subgroup, however by 20:00 this again becomes fluvial. As the working population returns to primarily residential locations further

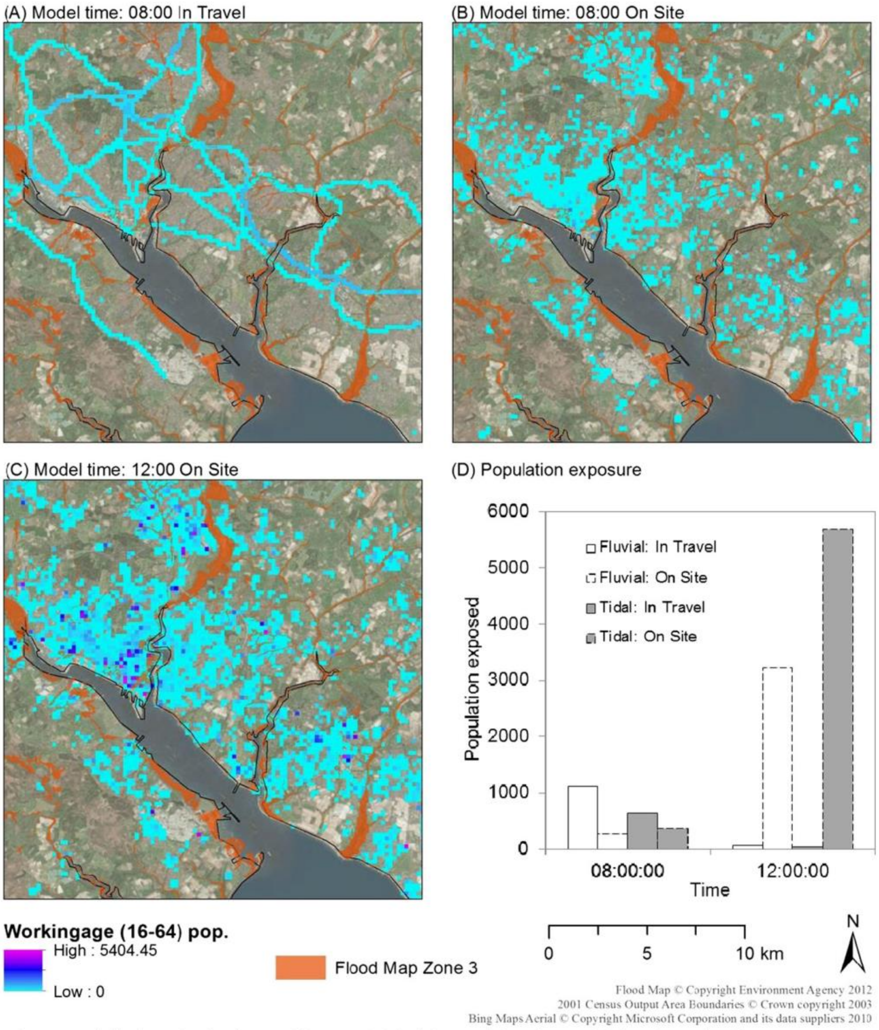


Fig. 8 Modelled results for the working aged (16–64) population in travel (**a**) and on site (**b**) at 08:00 and at 12:00 (**c**). The flood risk exposure to the working age population (**d**)

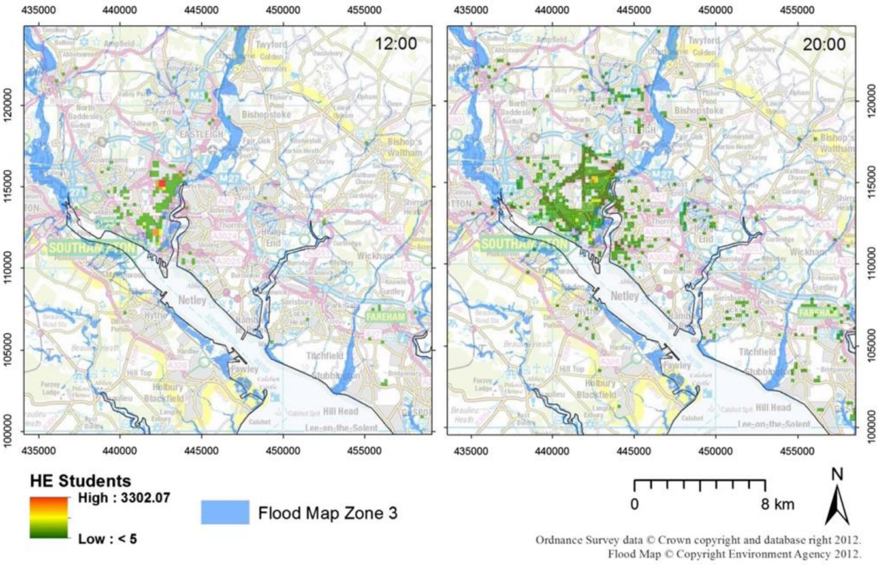


Fig. 9 Modelled representation of the higher education (HE) student population for a typical term time weekday at 12:00 and 20:00

inland, fluvial risk becomes the main factor. It can also be observed that there is a notable tidal flood risk to higher education students throughout the day. While there is not such a large reversal of this trend in the evening it does decrease reflecting the locations of student halls of residence or sites within the institutions. The spread of students in evening residential locations (Fig. 9) shows that a large portion are within the central area of the city with more prominent flood risks. For vulnerability assessment and the purposes of emergency preparedness this technique provides valuable improvements over traditional static hazard maps.

Figure 6 shows that calculating fluvial and tidal population exposure using the 2001 census residential locations tends to overestimate flood exposure, particularly for fluvial flooding in this study area. As well as providing better differentiation of which areas are inhabited or not at specific times the modelled results suggest that large portions of the time-specific population at risk are highly concentrated in specific areas depending on the time of day and age range (e.g. Figs. 5, 8 and 9): the spread and concentration of population at given times is of critical importance to emergency planners for targeting resources and developing plans.

Discussion

The modelled outputs, even in this relatively simple example, contrast starkly with the coarse, static 'night-time' population density coverage given by the 2001 census OA map. The modelled outputs reveal the ebb and flow in population density at different locations during the working day, including areas which are uninhabited. Despite their widespread use for exposure assessment, such variations in population density are not captured by traditional census maps. The proposed methodology allows the integration of additional hazard maps or more sophisticated hazard models in Fig. 1b. Urban areas and large conurbations experience a diverse range of population interactions and fluctuations. While the hazard footprints used in this case study appear static they are in fact based on complex hydrological modelling undertaken by the Environment Agency and their partners. Important trends can be resolved in the static hazard data and the techniques could be equally applied to more dynamic models.

Figure 3 illustrates an example temporal profile for a specific population activity. Acquiring data of sufficient resolution to inform temporal distributions of population for a given activity is one challenge for this approach. In this example, time use data from a large scale nationally representative survey were employed. The assumed constant ratio of customers to retail employees is used as a proxy for store size and resultant footfall in the absence of uniform data coverage. This approach encompasses numerous simplifying assumptions but nevertheless achieves transfer of population into major retail centres following a realistic temporal profile. This can be amended without fundamentally altering the overall modelling approach. In terms of retail activity it would be possible to improve this profile and destination capacities using detailed site specific commercial datasets. Indeed, any or all of the input datasets could be enhanced or updated without fundamentally altering the overall modelling approach.

The temporal characteristics of the natural hazard have not been directly addressed in this case study. Tidal time and height are not accounted for here but vary continuously as part of a predictable pattern. Tidal information indicating a high spring tide under storm surge conditions in the middle of the working day would potentially have a greater effect on some of the population subgroups identified in this study. Such temporally-specific hazard information could be easily implemented to advance the modelling further.

The modelled results provide enhanced assessments of subgroups of the population and their activities. They can be used to simulate peak travel times and populations in transit. This case study confirms that specific age groups are more at risk at different times. The richness of this detail cannot be derived from static census estimates alone, although the availability of accurate census small area statistics remains an important data source. The enhancements seen here come partly from the non-space-filling nature of the gridded model used, but importantly also reflect the continuous temporal redistribution within the course of a day. Alternative spatial models include dasymetric approaches (e.g. Mennis 2003) and other areal interpolation techniques (e.g. Goodchild and Lam 1980; Xie 1995). These all provide means of representing more realistic spatial population distributions compared to the basic census zone approach but none of them provide a method for modelling population in both space and time: the spatio-temporal gridded method adopted in this paper provides one possible solution. While a raster output may not be desired in all situations it does facilitate integration with the raster outputs of flood inundation models. High resolution disaggregated spatio-temporal population counts permit detailed assessment of potential hazard exposure and impact as well as understanding of wider implications. The flood scenarios described in the example study area may not be considered rapid onset events with modern prediction and forecasting tools, but flooding does have the potential to occur rapidly and without prior warning (Murray et al. 2012).

The key contribution of this study has been to test a modelling approach which combines a non-space-filling representation with the ability to store and analyse the temporal redistribution of the population. The methodology provides a mechanism to estimate temporary and transient populations not previously accounted for in many flood risk applications.

The Population 24/7 method also provides scope to analyse the vulnerability of entire populations in more detail through the ability to study population subgroups with differing vulnerability characteristics. Vulnerability is a key contributing factor in the development of a hazard: the capacity to estimate vulnerability may inform management choices for groups such as the young or elderly who may require additional support during an emergency situation. The prominence of the student subgroup in the Southampton example serves to illustrate the importance of daily movements by large subgroups and the sensitivity of analysis to other cyclical timescales, in this case educational term dates. The approach followed here also improves the allocation of population subgroups to specific location and activities, such as school age children at school locations during school time. This type of information is of importance when targeting emergency response and health care provision.

Like hazards, the global population is unevenly distributed, often being concentrated on coastlines, along rivers and active seismic or volcanic regions. The size, location and demographic characteristics of a population are all drivers for the impacts of a natural hazard (Cutter 2010). Spatio-temporal population estimates can highlight cyclical patterns in population distribution. These may be on seasonal scales such as tourism, seasonal shoppers or students arriving at a university town during term time. Notable diurnal cycles have also been observed such as children attending school or employees arriving at work. These all have important implications for natural hazard impact assessments. The timing of the manifestation of a natural hazard will affect the population numbers and vulnerability of those exposed.

This approach does not attempt to provide insights into human behaviour in an emergency situation or provide real time information. It essentially describes predictable population redistribution over time. There are many ways in which the data sources and detailed distributional models could be enhanced within the current modelling framework. For the purpose of this simple demonstration a static flood hazard map has been utilised. Integrating this model with dynamic hazard models is a logical further step that would greatly enhance the understanding of spatio-temporal variations in risk to flooding. The approach can be utilised to inform emergency plans for known risks under a range of scenarios and temporal scales and, while it is not possible to know exact future population movements, it allows for the computation of probable distributions.

Validation of this approach and its modelled outputs is difficult. Where data are available and accessible at the required scale, there are some potential routes for validation. Approaches to consider may be the use of geo-referenced social media interactions, mobile telephony data or individual membership and visitor datasets. These suggestions may allow validation based on certain subgroups of the population, specific geographic areas or specific times, but it is not possible to validate all temporal and spatial scales at once. Currently, there is still work to be done on the availability, coverage, and release of such potentially sensitive data. Further evaluation needs to be undertaken of the sensitivity of modelled outputs to variation in input parameters and the production of measures of uncertainty. It is an inherent feature of this modelling approach that it cannot allocate more population to a location (e.g. pupils to a school) than are in fact recorded in the relevant data source.

The methodology proposed in this paper can be used to estimate population at risk to a given hazard event and can be adapted to a range of spatial and temporal scales. This has the potential to aid policy makers in managing risk such as defining local flood hazard under the EU Flood Directive which meets legislative criteria. The directive's preliminary flood risk assessment stipulates a requirement to assess the potential adverse consequences of future flood events for human health, including the position of populated areas. While this technique may not be appropriate at the micro level because of uncertainties within population movements it does provide richness of information at a regional level to better inform flood risk in populated areas. The main advantage in any technique with temporal variation is to acknowledge that the extents and populations of inhabited areas are not static. A major development of this methodology would be the integration of a spatial interaction model to enhance the catchment areas of destination locations, and the location of populations in the transportation system more broadly.

Conclusion

This paper demonstrates improvements in the accuracy of estimated population exposure to hazards to be gained via the adoption of an innovative integrated modelling approach which takes explicit account of both space and time. This is of particular importance when assessing population exposure to natural hazards which are themselves subject to continuous spatio-temporal variation. The potential utility of such models in flood risk management has been demonstrated, providing a wealth of detailed data for further analysis which could not be achieved using conventional means of processing key sources such as census data. Pending further dataset development and validation, this technique has direct application to natural hazard scenarios both within the UK and globally. The data structure of the model allows the user to readily refine or supplement the input datasets. Web data mining and the rise of open-source data are likely to make compiling time-referenced population datasets easier, albeit with increased challenges for data interpretation and validation. The results in this paper provide significant opportunities to further refine this methodology for policy makers and emergency planners.

Acknowledgments Funded by the Economic and Social Research Council. The authors gratefully acknowledge the research assistance of Samuel Leung.

References

ABP (2013). ABP Southampton Port Information. http://www.southamptonvts.co.uk/Port_Information/. Accessed 22/03/2013.

Alexander, D. (1993). *Natural disasters*. London: UCL Press.

Aubrecht, C., Özceylan, D., Steinnocher, K., & Freire, S. (2012a). Multi-level geospatial modeling of human exposure patterns and vulnerability indicators. *Natural Hazards*, 1–17 (DOI: 10.1007/s11069-11012-10389-11069), doi:10.1007/s11069-012-0389-9.

Aubrecht, C., Steinnocher, K., Köstl, M., Züger, J., & Loibl, W. (2012b). Long-term spatio-temporal social vulnerability variation considering health-related climate change parameters particularly affecting elderly. *Natural Hazards*, 1–14, doi:10.1007/s11069-012-0324-0.

Baxter, P. J. (2005). The east coast Big Flood, 31 January–1 February 1953: a summary of the human disaster. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 363*(1831), 1293–1312. doi:10.1098/rsta.2005.1569.

Bettess, R. (2005). *Flooding in boscastle and north Cornwall, august 2004* (*Phase 2 studies report* (pp. 170)). Environment Agency: HR Wallingford.

Bhaduri, B., Bright, E., Coleman, P., & Urban, M. (2007). LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal, 69*(1), 103–117. doi:10.1007/s10708-007-9105-9.

Burt, S. (2005). Cloudburst upon hendraburnick down: the boscastle storm of 16 august 2004. *Weather, 60*(8), 219–227. doi:10.1256/wea.26.05.

CIESIN (2013). Gridded Population of the World (GPWv3). http://sedac.ciesin.columbia.edu/data/collection/gpw-v3/sets/browse. Accessed 18/01/2013.

Cressman, G. P. (1959). An operational objective analysis system. *Monthly Weather Review, 87*(10), 367–374. doi:10.1175/1520-0493(1959)087<0367:aooas>2.0.co;2.

Cutter, S. L. (2010). Social science perspectives on hazards and vulnerability science. In T. Beer (Ed.), *Geophysical hazards* (pp. 17–30). Netherlands: Springer.

Cutter, S. L., & Finch, C. (2008). Temporal and spatial changes in social vulnerability to natural hazards. *Proceedings of the National Academy of Sciences, 105*(7), 2301–2306. doi:10.1073/pnas.0710375105.

DEFRA. (2011). *Flood risk management in England*. London: National Audit Office, TSO.

DfT (2005). National Transport Model: FORGE The Road Capacity & Costs Model. http://webarchive.nationalarchives.gov.uk/20110202223628/http://www.dft.gov.uk/pgr/economics/ntm/etheroadcapacityandcosts3031.pdf. Accessed 18/03/2014.

DfT (2009). Values of Time and Operating Costs: TAG Unit 3.5.6. http://www.dft.gov.uk/webtag/documents/archive/1104/unit3.5.6.pdf. Accessed 18/03/2014.

DfT (2013). Great Britain Road Traffic Survey. http://www.dft.gov.uk/traffic-counts/. Accessed 27/03/2013.

Environment Agency (2012). England and Wales Zone 3 Flood Map. http://www.environment-agency.gov.uk/homeandleisure/37837.aspx. Accessed 22/03/2013.

European Council (2007). Council Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks. *Official Journal of the European Union L 288/27, 6.11.2007*, 27–34.

Fielding, J. (2007). Environmental injustice or just the Lie of the land: an investigation of the socio-economic class of those at risk from flooding in England and Wales. *Sociological Research Online, 12*(4), 4.

Gevertz, J. (2014). Aberystwyth University evacuates 600 students following extreme weather reports. The Independent 30/01/2014. http://www.independent.co.uk/student/news/aberystwyth-university-evacuates-600-students-following-extreme-weather-reports-9097255.html. Accessed 17/03/2014.

Goodchild, M. F., & Lam, N. S. (1980). A real interpolation: a variant of the traditional spatial problem. *Geo-Processing, 1*, 297–312.

Hall, J., Sayers, P., & Dawson, R. (2005). National-scale assessment of current and future flood risk in England and Wales. *Natural Hazards, 36*(1–2), 147–164. doi:10.1007/s11069-004-4546-7.

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

Harper, G., & Mayhew, L. (2012). Applications of population counts based on administrative data at local level. *Applied Spatial Analysis and Policy, 5*(3), 183–209. doi:10.1007/s12061-011-9062-z.

HESA (2012). *All students by HE institution 2011/12*. Higher Education Statistics Agency. https://www.hesa.ac.uk/. Accessed 10/07/2014.

Hilhorst, D., & Bankoff, G. (2008). Introduction: mapping vulnerability. In G. Bankoff, G. Frerks, & D. Hilhorst (Eds.), *Mapping vulnerability: disasters, development and people* (pp. 145–158). London: Earthscan.

HM Government. (2010). *A strong Britain in an Age of uncertainty: the national security strategy*. London: HMSO.

Horsburgh, K., & Horritt, M. (2006). The Bristol channel floods of 1607 –reconstruction and analysis. *Weather, 61*(10), 272–277. doi:10.1256/wea.133.05.

Hunter, N. M., Bates, P. D., Neelz, S., Pender, G., Villanueva, I., Wright, N. G., et al. (2008). Benchmarking 2D hydraulic models for urban flooding. *Proceedings of the ICE - Water Management, 161*(1), 13–30.

Huppert, H. E., & Sparks, R. S. J. (2006). Extreme natural hazards: population growth, globalization and environmental change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 364*(1845), 1875–1888. doi:10.1098/rsta.2006.1803.
Ipsos-RSL, & ONS. (2000). *UK time Use survey 2000*. Colchester, Essex: UK Data Archive.
King, R., & Ruiz-Gelices, E. (2003). International student migration and the European 'year Abroad': effects on European identity and subsequent migration behaviour. *International Journal of Population Geography, 9*(3), 229–252. doi:10.1002/ijpg.280.
Marsh, T. J. (2004). The january 2003 flood on the Thames. *Weather, 59*(3), 59–62. doi:10.1256/wea.212.03.
Marsh, T. J. (2008). A hydrological overview of the summer 2007 floods in England and Wales. *Weather, 63*(9), 274–279. doi:10.1002/wea.305.
Martin, D. (1989) Mapping population data from zone centroid locations Transactions of the Institute of *British Geographers NS, 14*, 90–97.
Martin, D. (1996). *Geographic information systems: socioeconomic applications*. London: Routledge.
Martin, D. (2011). SurfaceBuilder247: User Guide. http://www.esrc.ac.uk/my-esrc/grants/RES-062-23-1811/outputs/Read/ece508b5-6438-4e96-99eb-8fd68f1d3b99. Accessed 30/01/2014.
Martin, D., & Bracken, I. (1993). The integration of socioeconomic and physical resource data for applied land management information systems. *Applied Geography, 13*(1), 45–53. doi:10.1016/0143-6228(93)90079-G.
McGinnigle, J. B. (2002). The 1952 Lynmouth floods revisited. *Weather, 57*(7), 235–242. doi:10.1256/004316502760195894.
McPherson, T. N., & Brown, M. J. (2004). *Estimating daytime and nighttime population distributions in US cities for emergency response activities*. Paper presented at the Symposium on Planning, Nowcasting and Forecasting in the Urban Zone, 84th AMS Annual Meeting, Seattle.
Mennis, J. (2003). Generating surface models of population using dasymetric mapping. *Professional Geographer, 55*(1), 31–42.
Mitchell, J. K. (1999). Natural disasters in the context of megacities. In J. K. Mitchell (Ed.), *Crucibles of hazard: mega-cities and disasters in transition* (p. 450). Tokyo: United Nations University Press.
Munich Re. (2012). *Natural catastrophes 2011. Analyses, assessments, positions*. München, Germany: Topics Geo.
Murray, S. J., Smith, A. D., & Phillips, J. C. (2012). A modified flood severity assessment for enhanced decision-support: application to the Boscastle flood of 2004. *Weather and Forecasting, 27*(5), 1290–1297. doi:10.1175/waf-d-12-00033.1.
NFDC (2009). Cabinet 06/07/09 Report A: North Solent Shoreline Management Plan. http://www.newforest.gov.uk/committeedocs/cab/CDR05054.pdf. Accessed 22/03/2013.
OJ (2007). Official Journal of the European Union L 288, 6.11.2007, (p. 27–34). Special edition in croatian Chapter 15 Volume 001 p. 186–193.
ONS (2006a). Annual Business Inquiry. Office for National Statistics.
ONS (2006b). National Statistics Postcode Directory Office for National Statistics.
Openshaw, S. (1984). *The modifiable a real unit problem*. Norwich: Geo Books.
Pitt, M. (2008). *The Pitt review: learning lessons from the 2007 floods*. London: Cabinet Office.
RMS (2007). *1607 Bristol Channel Floods: 400-year retrospective*: Risk Management Solutions Report.
Rowe, D. (2004). *Boscastle: 16th august 2004 - the day of the flood*. Cornwall, UK.: Truran.
Ruocco, A., Nicholls, R., Haigh, I., & Wadey, M. (2011). Reconstructing coastal flood occurrence combining sea level and media sources: a case study of the Solent, UK since 1935. *Natural Hazards, 59*(3), 1773–1796. doi:10.1007/s11069-011-9868-7.
Schmitt, R. C. (1956). Estimating daytime populations. *Journal of the American Institute of Planners, 22*(2), 83–85.
Tobin, G. A., & Montz, B. E. (1997). *Natural hazards: explanation and integration*. New York: The Guilford Press.
United Nations. (2008). *World urbanization prospects: the 2007 revision*. New York: United Nations, Department of Economic and Social Affairs.
Wadey, M. P., Nicholls, R. J., & Hutton, C. (2012). Coastal flooding in the Solent: an integrated analysis of defences and inundation. *Water, 4*(2), 430–459.
WestQuay (2011). Media Pack. http://www.west-quay.co.uk/assets/Documents/Media%20Pack%202011.pdf. Accessed 12/10/2012.
Xie, Y. (1995). The overlaid network algorithms for a real interpolation problem. *Computers, Environment and Urban Systems, 19*(4), 287–306.
Zevenbergen, C., Veerbeek, W., Gersonius, B., & Van Herk, S. (2008). Challenges in urban flood management: travelling across spatial and temporal scales. *Journal of Flood Risk Management, 1*(2), 81–88. doi:10.1111/j.1753-318X.2008.00010.x.

Glossary

Annual Business Inquiry (ABI): The ABI is an ONS survey that contains employment and financial information from businesses. One variable produced is the number of employees, grouped according to their SIC (see Glossary). Access can be requested for these data published at LSOA or postcode sector level. In 2009 the Business Register Employment Survey (BRES) replaced the ABI and the Business Register Survey.

Areal interpolation: the re-aggregation of data from one set of zones (source) to another (target).

ASCII Grid: A raster format that can be used with GIS to transfer information within cell based systems.

Dasymetric Mapping: A type of areal interpolation used in population mapping that utilises ancillary datasets to distribute population that reflects more probable locations of placement. It builds on critiques arising from conventional choropleth mapping, where population densities are uniformly distributed (e.g. Wright 1936; Mennis 2003).

Geographic Information System (GIS): a data management system to store, retrieve, analyse, manage and display geographic information.

Great Britain (GB): The island that consists of England, Wales and Scotland. As opposed to the United Kingdom which includes Great Britain and Northern Ireland.

Lower Layer Super Output Area (LSOA): are the second smallest areal unit for which census estimates are published in England and Wales (after output areas). They typically contain 1,500 people and approximately five OAs. They are also the smallest areal unit for which MYEs are published.

Mid-year Population Estimates (MYEs): Mid-year population estimates provide an updated estimate on the population between census years. It refers to the population that is usually resident on 30 June for the reference year. For England and Wales MYEs provide population estimates by sex and single year of age down to LSOA level (ONS 2013a).

Modifiable Areal Unit Problem (MAUP): A phenomenon where the choice of zonal boundaries in which statistics are aggregated has a greater effect on the output analysis than the real distribution of the phenomenon being observed (Thurstain-Goodwin 2003). The issue was first proposed by Gehlke and Biehl (1934) and described by Openshaw (1984).

Modifiable Spatiotemporal Unit Problem (MSTUP): In an extension to the MAUP, the SMAUP is a term used by Martin *et al.* (forthcoming) that states data are required in the finest spatial and temporal units to mitigate the effects of spatiotemporal aggregation and data distortion.

Output Area (OA): The output area is the smallest areal until for the census in England, Wales and Northern Ireland. OAs must contain a minimum of 100 people and 40 households, but typically represent around 300 people. All residential addresses receive a census form that is allocated a grid reference to a resolution of 1 m. This is used to aggregate households into OAs for census outputs to protect census confidentiality thresholds (National Statistics 2004).

Population weighted centroid (PWC): PWCs are a summary reference point for the population within every census OA, LSOA or Medium Super Output Area (MSOA) in England and Wales. They represent the spatial distribution of population, based on the location of the highest density (ONS 2013b).

Pycnophylactic: A Greek term coined by Tobler (1979) defined as mass preservation. A process that describes the reallocation of aggregated data, without points to form a population surface estimation.

Source Zones: The areal units in which spatial data are available

Standard Industrial Classification (SIC): a classification of business based on the type of economic activity in which they are engaged.

Target Zones: The areal units in which spatial data are needed, often incompatible or different from *Source Zones.*

Unit Postcode (UPC): A UPC is an alphanumeric code (e.g. SO17 1BJ representing the University of Southampton) with an associated national grid reference. UPCs are UK wide and maintained by Royal Mail to identify mail delivery addresses. UPCs have two main categories for large (receiving over 1,000 pieces of mail per day) and small users. Small user postcodes typically represent 15 adjacent addresses (ONS 2014a).

Usually resident population: Where this term has been used it refers to the 2001 census definition for England and Wales. This is defined as someone who spends the majority of their time residing at that address.

List of References

ABP (2013). ABP Southampton Port Information. http://www.southamptonvts.co.uk/Port_Information/. Accessed 22/03/2013.

Adams, R. (2004). Where death really counts, 20 August 2004, The Guardian.

Ahas, R., Silm, S., Järv, O., Saluveer, E., & Tiru, M. (2010). Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones. *Journal of Urban Technology, 17*(1), 3-27, doi:10.1080/10630731003597306.

Ahola, T., Virrantaus, K., Krisp, J. M., & Hunter, G. J. (2007). A spatio-temporal population model to support risk assessment and damage analysis for decision-making. *International Journal of Geographical Information Science, 21*(8), 935-953, doi:10.1080/13658810701349078.

Alexander, D. (1991). Natural Disasters: A Framework for Research and Teaching. *Disasters, 15*(3), 209-226, doi:10.1111/j.1467-7717.1991.tb00455.x.

Alexander, D. (1993). *Natural Disasters*. London: UCL Press.

American Geological Institute (1984). *Glossary of geology*. Virginia: American Geological Institute.

Amey, G. (1974). *The Collapse of Dale Dyke Dam 1864*. London: Cassell.

Andryszewski, A., Evans, K., Haggett, C., Mitchell, B., Whitfield, D., & Harrison, T. (2005). *Levels of service approach to flood forecasting and warning.* Paper presented at the International conference on innovation advances and implementation of flood forecasting technology, Tromsø, Norway, 17 to 19 October 2005.

Antonopoulos, J. (1992). The great Minoan eruption of Thera volcano and the ensuing tsunami in the Greek Archipelago. *Natural Hazards, 5*(2), 153-168, doi:10.1007/bf00127003.

Arie, S. (2003). Italy ready to pay to clear slopes of volcano. http://www.guardian.co.uk/world/2003/jun/05/italy.sophiearie#article_continue. Accessed 11/12/2012.

Arrighi, C., Brugioni, M., Castelli, F., Franceschini, S., & Mazzanti, B. (2013). Urban micro-scale flood risk estimation with parsimonious hydraulic modelling and census data. *Natural Hazards and Earth System Science, 13*(5), 1375-1391.

Arup (2008). *Ulley Reservoir Geotechnical Report*. Leeds: Ove Arup & Partners Ltd.

Aubrecht, C., Steinnocher, K., Hollaus, M., & Wagner, W. (2009). Integrating earth observation and GIScience for high resolution spatial and functional modeling of urban land use. *Computers, Environment and Urban Systems, 33*(1), 15-25, doi:http://dx.doi.org/10.1016/j.compenvurbsys.2008.09.007.

Aubrecht, C., Steinnocher, K., Köstl, M., Züger, J., & Loibl, W. (2012a). Long-term spatio-temporal social vulnerability variation considering health-related climate change parameters particularly affecting elderly. *Natural Hazards*, 1-14, doi:10.1007/s11069-012-0324-0.

Aubrecht, C., Özceylan, D., Steinnocher, K., & Freire, S. (2012b). Multi-level geospatial modeling of human exposure patterns and vulnerability

indicators. *Natural Hazards*, 1-17 (DOI: 10.1007/s11069-11012-10389-11069), doi:10.1007/s11069-012-0389-9.
Aubrecht, C., Fuchs, S., & Neuhold, C. (2013). Spatio-temporal aspects and dimensions in integrated disaster risk management. *Natural Hazards, 68*(3), 1205-1216, doi:10.1007/s11069-013-0619-9.
Aubrecht, C., Steinnocher, K., & Huber, H. (2014). DynaPop–Population distribution dynamics as basis for social impact evaluation in crisis management. *ISCRAM*.
Batty, M. (2013). The future cities agenda. Editorial. *Environment and Planning B: Planning and Design, 40*, 191-194.
Baxter, P. J. (2005). The east coast Big Flood, 31 January–1 February 1953: a summary of the human disaster. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 363*(1831), 1293-1312, doi:10.1098/rsta.2005.1569.
BBC (2014a). Police helicopter warns residents to evacuate 5 February 2014. http://www.bbc.co.uk/news/uk-26056528. Accessed 24/09/2014.
BBC (2014b). English coastline flood risk from waves reassessed. BBC News, 12 November 2014. http://www.bbc.co.uk/news/uk-england-30013195. Accessed 19/11/2014.
BDS (2010). Spillways & Outlets - Scour. http://britishdams.org/about_dams/scour.htm. Accessed 27/08/2014.
Berkes, F. (2007). Understanding uncertainty and reducing vulnerability: lessons from resilience thinking. *Natural Hazards, 41*(2), 283-295, doi:10.1007/s11069-006-9036-7.
Bettess, R. (2005). Flooding in Boscastle and North Cornwall, August 2004. *Phase 2 Studies Report* (pp. 170). Wallingford: HR Wallingford/Environment Agency.
Beven, K. (1993). Riverine Flooding in a Warmer Britain. *The Geographical Journal, 159*(2), 157-161, doi:10.2307/3451405.
Bhaduri, B., Bright, E., Coleman, P., & Urban, M. (2007). LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal, 69*(1), 103-117, doi:10.1007/s10708-007-9105-9.
Bhaduri, B. (2008). Population distribution during the day In S. Shekhar, & H. Xiong (Eds.), *Encyclopedia of GIS*. New York: Springer.
Birkin, M., Chen, H., Clarke, M., Dew, P., Keen, J., Rees, P., & Xu, J. (2007a). MOSES: Modelling and Simulation for e-Social Science. *CASS Delegation Visit, University of Manchester*.
Birkin, M., Turner, A., Wu, B., Townend, P., & Xu, J. (2007b). An architecture for social simulation models to support spatial planning. *International Conference on e-Social Science, Ann Arbor, MI*.
Birkin, M., Turner, A., Wu, B., Townend, P., Arshad, J., & Xu, J. (2009). MoSeS: A Grid-Enabled Spatial Decision Support System. *Social Science Computer Review, 27*(4), 493-508, doi:10.1177/0894439309332295.
Birkin, M., & Wu, B. (2012). A Review of Microsimulation and Hybrid Agent-Based Approaches In A. J. Heppenstall, A. T. Crooks, L. M. See, & M. Batty (Eds.), *Agent-Based Models of Geographical Systems*. New York: Springer.
Birkmann, J. (2006). *Measuring vulnerability to natural hazards: towards disaster resilient societies*. Tokyo: United Nations University Press.
Bissell, J. J. (2010). Resilience of UK infrastructure. *Post note (UK Parliament), 362*, 1-4.

Blaikie, P., Cannon, T., Davis, I., & Wisner, B. (1994). *At Risk: Natural Hazards, Peoples Vulnerability, and Disasters*. London: Routledge.
Bogaert, P. (1996). Comparison of kriging techniques in a space-time context. *Mathematical Geology, 28*(1), 73-86, doi:10.1007/bf02273524.
Bracken, I., & Martin, D. (1989). The generation of spatial population distributions from census centroid data. *Environment and Planning A, 21*, 537-543.
Bracken, I., Martin, D. (1995). Linkage of the 1981 and 1991 UK Censuses using surface modelling concepts. *Environment and Planning A, 27*, 379-390.
Brauer, C. C., Teuling, A. J., Overeem, A., van der Velde, Y., Hazenberg, P., Warmerdam, P. M. M., & Uijlenhoet, R. (2011). Soil buffer limits flash flood response to extraordinary rainfall in a Dutch lowland catchment. *Hydrol. Earth Syst. Sci. Discuss., 8*(1), 111-150, doi:10.5194/hessd-8-111-2011.
BRISK (2010). Bristol Environmental Risk Research Centre (BRISK) Glossary. http://www.bris.ac.uk/brisk/sappur/glossary/. Accessed 04/10/2012.
Brown, J. D., & Damery, S. L. (2002). Managing flood risk in the UK: towards an integration of social and technical perspectives. *Transactions of the Institute of British Geographers, 27*(4), 412-426, doi:10.1111/1475-5661.00063.
Brown, N. R., Rips, L. J., & Shevell, S. K. (1985). The subjective dates of natural events in very-long-term memory. *Cognitive Psychology, 17*(2), 139-177, doi:http://dx.doi.org/10.1016/0010-0285(85)90006-4.
Budd, L., Griggs, S., Howarth, D., & Ison, S. (2011). A fiasco of volcanic proportions? Eyjafjallajökull and the closure of European airspace. *Mobilities, 6*(1), 31-40.
Burnside, R., Miller, D. S., & Rivera, J. D. (2007). The impact of information and risk perception on the hurricane evacuation decision-making of greater New Orleans residents. *Sociological Spectrum, 27*(6), 727-740, doi:10.1080/02732170701534226.
Burt, S. (2005). Cloudburst upon Hendraburnick Down: The Boscastle storm of 16 August 2004. *Weather, 60*(8), 219-227, doi:10.1256/wea.26.05.
Cabinet Office (2011). *Civil Contingencies Act: a short guide (Revised)*. London: The Stationary Office.
Cabinet Office (2012). *National Risk Register of Civil Emergencies (2012 edition)*. London: Cabinet Office.
Cardona, O. D. (2004). The need for rethinking the concepts of vulnerability and risk from a holistic perspective: a necessary review and criticism for effective risk management In G. Bankoff, G. Frerks, & D. Hilhorst (Eds.), *Mapping Vulnerability, Disasters, Development and People*. London: Earthscan.
Carlino, S., Somma, R., & Mayberry, G. C. (2008). Volcanic risk perception of young people in the urban areas of Vesuvius: Comparisons with other volcanic areas and implications for emergency management. *Journal of Volcanology and Geothermal Research, 172*, 229-243.
Ceccato, V., & Uittenbogaard, A. C. (2013). Space–Time Dynamics of Crime in Transport Nodes. *Annals of the Association of American Geographers, 104*(1), 131-150, doi:10.1080/00045608.2013.846150.

References

CEH (2014). Flood Estimation Handbook, Centre for Ecology and Hydrology. http://www.ceh.ac.uk/feh2/insidethefloodestimationhandbook.html. Accessed 19/11/2014.

Chakraborty, J., Tobin, G., & Montz, B. (2005). Population Evacuation: Assessing Spatial Variability in Geophysical Risk and Social Vulnerability to Natural Hazards. *Natural Hazards Review, 6*(1), 23-33, doi:10.1061/(ASCE)1527-6988(2005)6:1(23).

Chen, K., McAneney, J., Blong, R., Leigh, R., Hunter, L., & Magill, C. (2004). Defining area at risk and its effect in catastrophe loss estimation: a dasymetric mapping approach. *Applied Geography, 24*(2), 97-117, doi:10.1016/j.apgeog.2004.03.005.

Chester, D. K., Dibben, C. J. L., & Duncan, A. M. (2002). Volcanic hazard assessment in western Europe. *Journal of Volcanology and Geothermal Research, 115*(3–4), 411-435, doi:http://dx.doi.org/10.1016/S0377-0273(02)00210-X.

Cloke, P., Philo, C. and Sadler, D. (1991). *Approaching Human Geography*. London: Paul Chapman Ltd.

Coburn, A. W., Spence, R. J. S., & Pomonis, A. (1992). Factors determining human casualty levels in earthquakes: Mortality prediction in building collapse. In *Proceedings of the tenth world conference on earthquake engineering, Madrid, Spain, 19-24 July 1992*. Netherlands: A.A. Balkema.

Cole, T. W., & Fellows, K. L. (2008). Risk Communication Failure: A Case Study of New Orleans and Hurricane Katrina. *Southern Communication Journal, 73*(3), 211-228, doi:10.1080/10417940802219702.

Collier, C. G. (2007). Flash flood forecasting: What are the limits of predictability? *Quarterly Journal of the Royal Meteorological Society, 133*(622), 3-23, doi:10.1002/qj.29.

Çöltekin, A., De Sabbata, S., Willi, C., Vontobel, I., Pfister, S., Kuhn, M., & Lacayo, M. (2011). *Modifiable temporal unit problem.* Paper presented at the ISPRS/ICA workshop: “Persistent problems in geographic visualization”(ICC2011), Paris, France, 2011.

Cooper, A., Turnbull, M., Grey, S., & Loilier, P. (2013). *Tropical Cyclone Modelling with TELEMAC-2D.* Paper presented at the XXth TELEMAC-MASCARET User Conference, Karlsruhe, 16-18 October 2013.

Cornwall Council (2011). Preliminary Flood Risk Assessment: ANNEX 5 – Chronology of Major Flood Events in Cornwall. http://www.cornwall.gov.uk/media/6959245/Cornwall-PFRA-Annex5-June-2011.pdf. Accessed 24/09/2014.

CRC, OPCS, & GROS (1980). *People in Britain: A Census Atlas*. London: HMSO.

Crozier, M. J. (2005). Management Frameworks for Landslide Hazard and Risk: Issues and Options In T. Glade, M. Anderson, & M. J. Crozier (Eds.), *Landslide Hazard and Risk*. Chichester, UK: John Wiley & Sons Ltd.

Cutter, S. L. (1996). Vulnerability to environmental hazards. *Progress in Human Geography, 20*(4), 529-539.

Cutter, S. L. (2003). The Vulnerability of Science and the Science of Vulnerability. *Annals of the Association of American Geographers, 93*(1), 1-12.

Cutter, S. L., & Finch, C. (2008). Temporal and spatial changes in social vulnerability to natural hazards. *Proceedings of the National Academy of Sciences, 105*(7), 2301-2306, doi:10.1073/pnas.0710375105.

Cutter, S. L. (2010). Social Science Perspectives on Hazards and Vulnerability Science In T. Beer (Ed.), *Geophysical Hazards*. Netherlands: Springer
Dai, F. C., Lee, C. F., Li, J., & Xu, Z. W. (2001). Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environmental Geology, 40*(3), 381-391, doi:10.1007/s002540000163.
Davies, P. (1995). *About Time: Einstein's unfinished revolution*. New York: Simon and Schuster.
DEFRA (2006). *Flood Risks to People: Phase 2 Project Record FD2321/PR*. London: Department for Environment, Food and Rural Affairs.
DEFRA (2008). *The Government's Response to Sir Michael Pitt's Review of the Summer 2007 Floods*. London: Department for Environment, Food and Rural Affairs.
DEFRA (2011). *Flood Risk Management in England*. London: National Audit Office, TSO.
DEFRA (2012). *Waste water treatment in the United Kingdom – 2012 Implementation of the European Union Urban Waste Water Treatment Directive – 91/271/EEC*. London: Department for Environment, Food and Rural Affairs.
Delica-Willison, Z., & Willison, R. (2008). Vulnerability Reduction: a task for a vulnerable people themselves In G. Bankoff, G. Frerks, & D. Hilhorst (Eds.), *Mapping vulnerability: disasters,development and people*. London: Earthscan.
Denbigh, K. G. (1981). *Three concepts of time*. Berlin: Springer-Verlag.
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, 111*(45), 15888-15893, doi:10.1073/pnas.1408439111.
DfT (2005). National Transport Model: FORGE The Road Capacity & Costs Model. http://webarchive.nationalarchives.gov.uk/20110202223628/http://www.dft.gov.uk/pgr/economics/ntm/etheroadcapacityandcosts3031.pdf. Accessed 18/03/2014.
DfT (2010). National Travel Survey Colchester, Essex: UK Data Archive.
DfT (2011a). Congestion and Reliability Statistics. http://www.dft.gov.uk/statistics/series/congestion-and-reliability/. Accessed 15/04/2012.
DfT (2011b). *Road traffic statistics: annual road traffic estimates 2010*. London: Department for Transport.
DfT (2011c). Accessibility Destination Datasets. London: Department for Transport.
DfT (2012a). *National Travel Survey: Data Extract User Guide, 2002-2012, Key Definitions*. London: Department for Transport.
DfT (2012b). National Travel Survey. Colchester, Essex: UK Data Archive.
DfT (2013). Great Britain Road Traffic Survey. http://www.dft.gov.uk/traffic-counts/. Accessed 27/03/2013.
Di Baldassarre, G., Castellarin, A., Montanari, A., & Brath, A. (2009a). Probability-weighted hazard maps for comparing different flood risk management strategies: a case study. *Natural Hazards, 50*(3), 479-496, doi:10.1007/s11069-009-9355-6.
Di Baldassarre, G., Castellarin, A., & Brath, A. (2009b). Analysis of the effects of levee heightening on flood propagation: example of the River Po, Italy. *Hydrological Sciences Journal, 54*(6), 1007-1017, doi:10.1623/hysj.54.6.1007.

References

Diansheng, G. (2009). Flow Mapping and Multivariate Visualization of Large Spatial Interaction Data. *Visualization and Computer Graphics, IEEE Transactions on, 15*(6), 1041-1048, doi:10.1109/tvcg.2009.143.

Dobran, F. (2003). VESUVIUS 2000: A multidisciplinary project for the Vesuvius area. In *EGS-AGU-EUG Joint Assembly, Nice, France, 6-11 April 2003 2003* (Vol. 1, pp. 4569). Nice: EGS-AGU-EUG.

Dobson, J., Bright, E., Coleman, P., Durfee, R., & Worley, B. (2000). LandScan: A Global Population Database for Estimating Populations at Risk *Photogrammetric Engineering and Remote Sensing, 66*(7), 849-857.

Douglas, J. (2007). Physical vulnerability modelling in natural hazard risk assessment. *Natural Hazards and Earth System Sciences, 7*, 283-288.

Dovers, S. R., & Handmer, J. W. (1992). Uncertainty, sustainability and change. *Global Environmental Change, 2*(4), 262-276, doi:10.1016/0959-3780(92)90044-8.

Dow, K., & Downing, T. E. (1995). Vulnerability research: where things stand. *Human Dimensions Quarterly, 1*, 3-5.

Duke-Williams, O., & Rees, P. (1998). Can Census Offices publish statistics for more than one small area geography? An analysis of the differencing problem in statistical disclosure. *International Journal of Geographical Information Science, 12*(6), 579-605, doi:10.1080/136588198241680.

Dye, C. (2008). Health and Urban Living. *Science, 319*(5864), 766-769, doi:10.1126/science.1150198.

EA (2012a). *East Cornwall Catchment Flood Management Plan*. Exeter: Environment Agency.

EA (2012b). *West Cornwall Catchment Flood Management Plan*. Exeter: Environment Agency.

EA (2014). Floods destroy. Be prepared. https://www.gov.uk/floodsdestroy. Accessed 01/12/2014.

Elliott, J. R., Nissen, E., England, P. C., Jackson, J. A., Lamb, S., Li, Z., Oehlers, M., & Parsons, B. E. (2011). *Slip in the 2010-2011 Canterbury Earthquakes, New Zealand and implications for future seismic hazard in Christchurch.* Paper presented at the American Geophysical Union, San Francisco, 5-9th December 2011.

EM-DAT (2009a). The EM-DAT Glossary. http://www.emdat.be/glossary/9#lettere. Accessed 21/09/2012.

EM-DAT (2009b). Criteria and Definition. http://www.emdat.be/criteria-and-definition. Accessed 21/09/2012.

EM-DAT (2014). The international disaster database. Brussels, Belgium: Centre for Research on the Epidemiology of Disasters.

English Heritage (2010). Tintagel Castle monthly visitor numbers 2010 (Personal communication).

Environment Agency (2007a). 2007 summer floods - Environment Agency - A table showing the likelihood of the 2007 summer floods occurring at places where we measure river flows and levels. http://www.environment-agency.gov.uk/static/documents/Research/returnperiods_1918541.pdf. Accessed 13/11/2013.

Environment Agency (2007b). Case study 2007 summer floods: reservoir safety – learning from Ulley. http://www.environment-agency.gov.uk/static/documents/Research/reservoirscasestudy_1917484.pdf. Accessed 03/04/2013.

Environment Agency (2011). *Flooding from groundwater*. Bristol: Environment Agency.

Environment Agency (2012a). Flood update: Thursday 29 November – 13:30. http://www.environment-agency.gov.uk/news/144522.aspx?coverage=&persona=&sector=&frompanel=1. Accessed 29/11/ 2012.
Environment Agency (2012b). Flood hazard and flood risk maps. http://www.environment-agency.gov.uk/research/planning/135518.aspx. Accessed 30/11/ 2012.
Environment Agency (2012c). Flood map. http://www.environment-agency.gov.uk/homeandleisure/37837.aspx. Accessed 14/12/2014.
Environment Agency (2013). 2m LiDAR Data Acquisition. http://www.environment-agency.gov.uk Accessed August 2014.
European Council (2007). Council Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks. *Official Journal of the European Union L 288/27, 6.11.2007*, 27-34.
Eurostat (2014). European Union Census Hub. http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/. Accessed 27/11/2014.
Fell, R., Whitt, G., Miner, A., & Flentje, P. N. (2007). Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. *Australian Geomechanics Journal, 42*(1), 13-36.
FERC (2005). Engineering Guidelines for the Evaluation of Hydropower Projects - Chapter 4 Embankment Dams. https://www.ferc.gov/industries/hydropower/safety/guidelines/eng-guide/chap4.PDF. Accessed 01/09/2014.
Fernandes, E. H., Dyer, K. R., & Niencheski, L. F. H. (2001). Calibration and validation of the TELEMAC-2D model to the Patos Lagoon (Brazil). *Journal of Coastal Research*, 470-488.
Fernández, T., Irigaray, C., El Hamdouni, R., & Chacón, J. (2003). Methodology for Landslide Susceptibility Mapping by Means of a GIS. Application to the Contraviesa Area (Granada, Spain). *Natural Hazards, 30*(3), 297-308, doi:10.1023/B:NHAZ.0000007092.51910.3f.
Fewtrell, T. J., Bates, P. D., Horritt, M., & Hunter, N. M. (2008). Evaluating the effect of scale in flood inundation modelling in urban environments. *Hydrological Processes, 22*(26), 5107-5118, doi:10.1002/hyp.7148.
Fielding, J. (2007). Environmental Injustice or Just the Lie of the Land: an Investigation of the Socio-Economic Class of Those at Risk from Flooding in England and Wales. *Sociological Research Online, 12*(4), 4.
Fisher, P., & Unwin, D. (2005). Re-presenting Geographical Information Systems In P. Fisher, & D. Unwin (Eds.), *Re-presenting GIS*. Chichester: John Wiley and Sons.
FLOODsite (2009). HR Breach Model. http://www.floodsite.net/html/HR_Breach_Model.htm. Accessed 08/05/2014.
Flowerdew, R., & Green, M. (1989). Statistical methods for inference between incompatible zonal systems In M. F. Goodchild, & S. Gopal (Eds.), *Accuracy of Spatial Databases*. London: Taylor and Francis.
Foley, D. L. (1954). Urban Daytime Population: A Field for Demographic-Ecological Analysis. *Social Forces, 32*(4), 323-330.
Frank, A. U. (1998). Different types of "time" in GIS In M. J. Egenhofer, & R. G. Golledge (Eds.), *Spatial and Temporal Reasoning in Geographic Information Systems*. New York: Oxford University Press.

References

Freire, S. (2010). Modeling of spatiotemporal distribution of urban population at high resolution–value for risk assessment and emergency management In *Geographic information and cartography for risk and crisis management* (pp. 53-67): Springer.

Freire, S., Aubrecht, C., & Wegscheider, S. Spatio-temporal population distribution and evacuation modeling for improving tsunami risk assessment in the Lisbon Metropolitan Area. In *Proceedings of Gi4DM 2011, international symposium on geoinformation for disaster management, Antalya, 2011* (pp. 6)

Freire, S., & Aubrecht, C. (2012). Integrating population dynamics into mapping human exposure to seismic hazard. *Nat. Hazards Earth Syst. Sci., 12*(11), 3533-3543, doi:10.5194/nhess-12-3533-2012.

Froehlich, D. (2008). Embankment Dam Breach Parameters and Their Uncertainties. *Journal of Hydraulic Engineering, 134*(12), 1708-1721, doi:10.1061/(ASCE)0733-9429(2008)134:12(1708).

Garb, J. L., Cromley, R. G., & Wait, R. B. (2007). Estimating populations at risk for disaster preparedness and response. *Journal of Homeland Security and Emergency Management, 4*(1).

Gatrell, A. (1994). Density Estimation and the Visualization of Point Patterns In H. M. Hearnshaw, & D. J. Unwin (Eds.), *Visualization in Geographical Information Systems*. Chichester, UK.: John Wiley and Sons Ltd.

Gatrell, A. C., Bailey, T. C., Diggle, P. J., & Rowlingson, B. S. (1996). Spatial Point Pattern Analysis and Its Application in Geographical Epidemiology. *Transactions of the Institute of British Geographers, 21*(1), 256-274.

Gehlke, C. E., & Biehl, K. (1934). Certain Effects of Grouping Upon the Size of the Correlation Coefficient in Census Tract Material. *Journal of the American Statistical Association, 29*(185), 169-170.

Gething, P. W., Atkinson, P. M., Noor, A. M., Gikandi, P. W., Hay, S. I., & Nixon, M. S. (2007). A local space–time kriging approach applied to a national outpatient malaria data set. *Computers & Geosciences, 33*(10-5), 1337-1350, doi:10.1016/j.cageo.2007.05.006.

Gevertz, J. (2014). Aberystwyth University evacuates 600 students following extreme weather reports. The Independent 30/01/2014. http://www.independent.co.uk/student/news/aberystwyth-university-evacuates-600-students-following-extreme-weather-reports-9097255.html. Accessed 17/03/2014.

Giacomelli, L., Perrotta, A., Scandone, R., & Scarpati, C. (2003). The eruption of Vesuvius of 79 AD and its impact on human environment in Pompeii. *Episodes-Newsmagazine of the International Union of Geological Sciences, 26*(3), 235-238.

GLA (2005). London Town Centre Assessment (Stage 1) Convenience Goods Floorspace Need in London. Greater London Authority (GLA). https://www.london.gov.uk/sites/default/files/archives/mayor-planning-docs-convenience_goods_report.rtf. Accessed 15/10/2014.

GMAP (2014). GMAP (Geographical Modelling and Planning) retail centres for Cornwall. Leeds: Callcredit Information Group Ltd.

Goodchild, M. F., & Lam, N. S. (1980). Areal Interpolation: a variant of the traditional spatial problem. *Geo-Processing, 1*, 297-312.

Goodchild, M. F., Anselin, L., & Deichmann, U. (1993). A framework for the areal interpolation of socioeconomic data. *Environment and Planning A, 25*(3), 383-397.

Goodchild, M. F. (2000). GIS and Transportation: Status and Challenges. *GeoInformatica, 4*(2), 127-139, doi:10.1023/a:1009867905167.
Gotway, C. A., & Young, L. J. (2002). Combining Incompatible Spatial Data. *Journal of the American Statistical Association, 97*(458), 632-648.
Grattan, J., & Torrence, R. (2007). Beyond Gloom and Doom: the long term consequences of volcanic disasters In J. Grattan, & R. Torrence (Eds.), *Living under the shadow: cultural impacts of volcanic eruptions*. California: Left Coast Press.
Gregory, D. (2000). Time-geography In R. J. Johnston, D. Gregory, G. Pratt, & M. Watts (Eds.), *The Dictionary of Human Geography*. Oxford: Blackwell.
Gudmundsson, M. T., Pedersen, R., Vogfjörd, K., Thorbjarnardóttir, B., Jakobsdóttir, S., & Roberts, M. J. (2010). Eruptions of Eyjafjallajökull Volcano, Iceland. *Eos Trans. AGU, 91*(21), doi:10.1029/2010eo210002.
Guha-Sapir, D., Rodriguez-Llanes, J. M., & Jakubicka, T. (2011). Using disaster footprints, population databases and GIS to overcome persistent problems for human impact assessment in fl ood events. *Natural Hazards, March 2011 Short Communication*, doi:10.1007/s11069-011-9775-y.
Hägerstrand, T. (1973). The domain of human geography In R. J. Chorley (Ed.), *Directions in Geography*. London: Methuen and Co.
Hägerstrand, T. (1975). Space, time and human conditions In A. Karlqvist, L. Lunqvist, & F. Snickars (Eds.), *Dynamic allocation of urban space*. Farnborough: Saxon House.
Hägerstrand, T. (1976). Geography and the study of interaction between nature and society. *Geoforum, 7*(5-6), 329-334, doi:10.1016/0016-7185(76)90063-4.
Handmer, J., Norton, T., & Dovers, S. (2001). *Ecology, Uncertainty and policy: managing ecosystems for sustainability*. Harlow: Prentice Hall.
Harper, G., & Mayhew, L. (2012). Applications of Population Counts Based on Administrative Data at Local Level. *Applied Spatial Analysis and Policy, 5*(3), 183-209, doi:10.1007/s12061-011-9062-z.
Harriss, R. C., Hohenemser, C., & Kates, R. W. (1978). Our Hazardous Environment. *Environment: Science and Policy for Sustainable Development, 20*(7), 6-41, doi:10.1080/00139157.1978.9928699.
Harvey, D. (1989). *The Condition of Postmodernity*. Oxford: Blackwell.
Harvey, J. T. (2002). Population Estimation Models Based on Individual TM Pixels. *Photogrammetric Engineering & Remote Sensing, 68*(11), 1181-1192.
Hawkin, S. (1995). *A Brief History Of Time: From the Big Bang To Black Holes*. London: Bantam.
Haynes, K., Barclay, J., & Pidgeon, N. (2007). Volcanic hazard communication using maps: an evaluation of their effectiveness. *Bulletin of Volcanology, 70*(2), 123-138, doi:10.1007/s00445-007-0124-7.
He, X. Y., Wang, Z. Y., & Huang, J. C. (2008). Temporal and spatial distribution of dam failure events in China. *International Journal of Sediment Research, 23*(4), 398-405, doi:http://dx.doi.org/10.1016/S1001-6279(09)60010-X.
Heath, S. E., Kass, P. H., Beck, A. M., & Glickman, L. T. (2001). Human and Pet-related Risk Factors for Household Evacuation Failure During a Natural Disaster. *American Journal of Epidemiology, 153*(7), 659-665.

Hein, P. (2013). Expecting the unexpected: a case study on tsunami mitigation in Fujisawa (Japan). *Environmental Hazards*, 1-20, doi:10.1080/17477891.2013.772890.

HESA (2012). All students by HE institution 2011/12. Higher Education Statistics Agency.

HESA (2013). Headline Statistics. http://www.hesa.ac.uk/. Accessed 12/07/2013.

Heuvelink, G. B. M., & Griffith, D. A. (2010). Space–Time Geostatistics for Geography: A Case Study of Radiation Monitoring Across Parts of Germany. *Geographical Analysis, 42*(2), 161-179, doi:10.1111/j.1538-4632.2010.00788.x.

Hewitt, K. (1997). *Regions of risk : a geographical introduction to disasters*. Harlow: Longman.

Hilhorst, D., & Bankoff, G. (2008). Introduction: Mapping Vulnerability In G. Bankoff, G. Frerks, & D. Hilhorst (Eds.), *Mapping vulnerability: disasters,development and people*. London: Earthscan.

Hinks, J. L., & Mason, P. J. (2007). *Environment Agency Ulley Dam: Post-Incident Review*. Swindon: Halcrow Group Limited.

Hinks, J. L., Mason, P. J., & Claydon, J. R. (2008). Ulley Reservoir and high velocity spillway flows In H. Hewlett (Ed.), *Ensuring reservoir safety into the future*. London: Thomas Telford.

HM Government (2010). *A Strong Britain in an Age of Uncertainty: The National Security Strategy*. London: HMSO.

Holling, C. S. (1973). Resilience and Stability of Ecological Systems. *Annual Review of Ecology and Systematics, 4*, 1-23, doi:10.2307/2096802.

Holling, C. S. (1986). The resilience of terrestrial ecosystems: local surprise and global change In W. C. Clark, & R. E. Munn (Eds.), *Sustainable Development of the Biosphere*. Cambridge: Cambridge University Press.

Horner, M. W., & Walsh, P. D. (2000). Easter 1998 Floods. *Water and Environment Journal, 14*(6), 415-418, doi:10.1111/j.1747-6593.2000.tb00287.x.

Horrocks, J. (2010). Rebuilding Ulley Dam. *Hydro Review Worldwide, September 2010*, 22-24.

Horwich, G. (2000). Economic Lessons of the Kobe Earthquake. *Economic Development and Cultural Change, 48*(3), 521-542, doi:10.1086/452609.

Houston, D., Werritty, A., Bassett, D., Geddes, A., Hoolachan, A., & McMillan, M. (2011). *Pluvial (rain-related) flooding in urban areas: the invisible hazard*. York: Joseph Rowntree Foundation.

HSCIC (2010). Hospital Episode Statistics (HES). Health and Social Care Information Centre (HSCIC). http://www.hscic.gov.uk/hes. Accessed 01/08/2014.

HSL (2014). Case Study National Population Database - a tool to support HSE's work. http://www.hsl.gov.uk/resources/case-studies/national-population-database---a-tool-to-support-hse%E2%80%99s-work. Accessed 14/12/2014.

Hudson-Smith, A. (2014). Tracking, Tagging and Scanning the City. *Architectural Design, 84*(1), 40-47, doi:10.1002/ad.1700.

Huppert, H. E., & Sparks, R. S. J. (2006). Extreme natural hazards: population growth, globalization and environmental change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 364*(1845), 1875-1888, doi:10.1098/rsta.2006.1803.

IHS (2011). Global Economic Impact of the Japanese Earthquake, Tsunami, and Nuclear Disaster. http://www.ihs.com/products/global-insight/industry-economic-report.aspx?id=1065929196. Accessed 01/12/2014.
Ipsos-RSL, & ONS (2000). UK Time Use Survey 2000. Colchester, Essex: UK Data Archive.
John Lewis Partnership (2010). *Annual report and accounts 2010*. London: John Lewis Partnership.
Johnston, R. J. (1978). *Multivariate Statistical Analysis in Geography*. London: Longman.
Journel, A. G., & Huijbregts, C. J. (1978). *Mining Geostatistics*. London: Academic Press.
Kaiser, A., Holden, C., Beavan, J., Beetham, D., Benites, R., Celentano, A., Collett, D., Cousins, J., Cubrinovski, M., *et al.* (2012). The Mw 6.2 Christchurch earthquake of February 2011: preliminary report. *New Zealand Journal of Geology and Geophysics, 55*(1), 67-90, doi:10.1080/00288306.2011.641182.
Kates, R. (1976). Experiencing the Environment as Hazard In S. Wapner, S. Cohen, & B. Kaplan (Eds.), *Experiencing the Environment* (pp. 133-156): Springer US.
Kilburn, C. R. J., & McGuire, B. (2001). *Italian Volcanoes*. Edinburgh: Terra.
King, D. (2000). You're on Your Own: Community Vulnerability and the Need for Awareness and Education for Predicatable Natural Disasters. *Journal of Contingencies and Crisis Management, 8*(4), 223-228, doi:10.1111/1468-5973.00143.
King, R., & Ruiz-Gelices, E. (2003). International student migration and the European 'Year Abroad': effects on European identity and subsequent migration behaviour. *International Journal of Population Geography, 9*(3), 229-252, doi:10.1002/ijpg.280.
Klein, R. J. T., Nicholls, R. J., & Thomalla, F. (2003a). Resilience to natural hazards: How useful is this concept? *Global Environmental Change Part B: Environmental Hazards, 5*(1–2), 35-45, doi:10.1016/j.hazards.2004.02.001.
Klein, R. J. T., Nicholls, R. J., & Thomalla, F. (2003b). The Resilience of Coastal Megacities to Weather-Related Hazards In A. Kreimer, M. Arnold, & A. Carlin (Eds.), *Building Safer Cities: The Future of Disaster Risk*. Washington D.C.: The World Bank.
Kyriakidis, P. C. (2004). A geostatistical framework for area-to-point spatial interpolation. *Geographical Analysis, 36*, 259-289.
Lam, N. S. (1983). Spatial Interpolation Methods: A Review. *The American Cartographer, 10*(2), 129-149.
Langford, M., Maguire, D. J., & Unwin, D. J. (1991). The Areal Interpolation Problem: Estimating Population Using Remote Sensing in a GIS Framework In I. Masser, & M. Blakemore (Eds.), *Handling Geographical Information: Methodology and Potential Applications*. New York: Wiley & Sons Ltd.
Langford, M., & Unwin, D. J. (1994). Generating and mapping population density surfaces within a geographical information system. *Cartographic Journal, The, 31*(1), 21-26.
Langford, M. (2007). Rapid facilitation of dasymetric-based population interpolation by means of raster pixel maps. *Computers, Environment and Urban Systems, 31*, 19-32.
Langran, G., & Chrisman, N. R. (1988). A framework for temporal geographic information. *Cartographica, 25*(3), 1-14.

Langran, G. (1992). *Time in Geographic Information Systems*. London: Taylor and Francis.

Langran, G. (1993). Issues of implementing a spatiotemporal system. *International Journal of Geographical Information Science, 7*(4), 305-314.

Lee, D. B. (1973). Requiem for Large-Scale Models. *Journal of the American Institute of Planners, 39*(3), 163-178, doi:10.1080/01944367308977851.

Lenntorp, B. (1999). Time-geography – at the end of its beginning. *GeoJournal, 48*(3), 155-158, doi:10.1023/a:1007067322523.

Lewis, H. (2009). *Boscastle: The flood 16/08/04*. Truro: Cornwall Council.

Lewis, M. (1989, How a Tokyo Earthquake Could Devastate Wall Street and the World Economy. *Manhattan Inc.*

Liu, X. H., Kyriakidis, P. C., & Goodchild, M. F. (2008). Population- density estimation using regression and area to point residual kriging. *International Journal of Geographical Information Science, 22*(4), 431-447, doi:10.1080/13658810701492225.

Lloyd, D., Haklay, M. M., Thurstain-Goodwin, M., & Tobon, C. (2003). Visualising spatial structure in urban data In P. A. Longley, & M. Batty (Eds.), *Advanced Spatial Analysis: The CASA book of GIS*. California: ESRI Press.

Lomas, M. (2007). Catcliffe floods. http://statementimages.co.uk/photo_essays/catcliffe-floods/#.VIV-2vmDl8H. Accessed 08/12/2014.

Longley, P. A., Adnan, M., & Lansley, G. (2015). The geotemporal demographics of Twitter usage. *Environment and Planning A, 47*(2), 465-484.

Maantay, J. A., Maroko, A. R., & Herrmann, C. (2007). Mapping Population Distribution in the Urban Environment: The Cadastral-based Expert Dasymetric System (CEDS). *Cartography and Geographic Information Science, 34*(2), 77-102, doi:10.1559/152304007781002190.

Malcherek, A. (2000). Application of TELEMAC-2D in a narrow estuarine tributary. *Hydrological Processes, 14*(13), 2293-2300, doi:10.1002/1099-1085(200009)14:13<2293::aid-hyp29>3.0.co;2-4.

Malleson, N., Heppenstall, A., & See, L. (2010). Crime reduction through simulation: An agent-based model of burglary. *Computers, Environment and Urban Systems, 34*(3), 236-250, doi:http://dx.doi.org/10.1016/j.compenvurbsys.2009.10.005.

Mark, O., Weesakul, S., Apirumanekul, C., Aroonnet, S. B., & Djordjevic, S. (2004). Potential and limitations of 1D modelling of urban flooding. [Article]. *Journal of Hydrology, 299*(3-4), 284-299, doi:10.1016/j.jhydrol.2001.08.014.

Marsh, T. J., & Dale, M. (2002). The UK Floods of 2000–2001: A Hydrometeorological Appraisal. *Water and Environment Journal, 16*(3), 180-188, doi:10.1111/j.1747-6593.2002.tb00392.x.

Marsh, T. J. (2004). The January 2003 flood on the Thames. *Weather, 59*(3), 59-62, doi:10.1256/wea.212.03.

Marsh, T. J., & Hannaford, J. (2007). *The summer 2007 fl oods in England and Wales – a hydrological appraisal.* Wallingford: Centre for Ecology & Hydrology.

Marsh, T. J. (2008). A hydrological overview of the summer 2007 floods in England and Wales. *Weather, 63*(9), 274-279, doi:10.1002/wea.305.

Martin, D. (1989). Mapping Population Data from Zone Centroid Locations. *Transactions of the Institute of British Geographers, 14*(1), 90-97.

Martin, D., & Bracken, I. (1993). The integration of socioeconomic and physical resource data for applied land management information systems. *Applied Geography, 13*(1), 45-53, doi:http://dx.doi.org/10.1016/0143-6228(93)90079-G.
Martin, D., Tate, N. J., & Langford, M. (2000). Refining Population Surface Models: Experiments with Northern Ireland Census Data. *Transactions in GIS, 4*(4), 343-360, doi:10.1111/1467-9671.00060.
Martin, D. (2009). The role of GIS In A. S. Fotheringham, & P. Rogerson (Eds.), *The SAGE Handbook of Spatial Analysis* (pp. 25-40). London: SAGE.
Martin, D., Leung, S., & Cockings, S. (2009). *Population 24/7 : building time-specific population grid models.* Paper presented at the European Forum for Geostatistics, The Hague, Netherlands, 5-7 October 2009.
Martin, D. (2011). SurfaceBuilder247: User guide. http://www.esrc.ac.uk/my-esrc/grants/RES-062-23-1811/outputs/Read/ece508b5-6438-4e96-99eb-8fd68f1d3b99.
Martin, D., Lloyd, C., & Shuttleworth, I. (2011). Evaluation of gridded population models using 2001 Northern Ireland Census data. *Environment and Planning A, 43*(8), 1965-1980.
Martin, D., Cockings, S., & Leung, S. (forthcoming). Developing a flexible framework for spatiotemporal population modelling. *Annals of the Association of American Geographers*.
Martin, R., Martin, D., & Cockings, S. (2014). *Spatio-temporal population vulnerability to radiological and nuclear hazards.* Paper presented at the Royal Geographical Society (RGS-IBG) Annual International Conference, London, 26-29 August 2014.
Mason, P. J., & Hinks, J. L. (2008). Security of stepped masonry spillways: lessons from Ulley dam. *Dams and Reservoirs, 18*, 5-8.
Mason, P. J., & Hinks, J. L. (2009). Conclusions from the post-incident review for Ulley dam. *Dams and Reservoirs, 19*, 43-44.
Massey, D. (1999). Space-Time, 'Science' and the Relationship between Physical Geography and Human Geography. *Transactions of the Institute of British Geographers, 24*(3), 261-276.
McEwen, L., Hall, T., Hunt, J., Dempsey, M., & Harrison, M. (2002). Flood warning, warning response and planning control issues associated with caravan parks: the April 1998 floods on the lower Avon floodplain, Midlands region, UK. *Applied Geography, 22*(3), 271-305, doi:http://dx.doi.org/10.1016/S0143-6228(02)00008-5.
McGinnigle, J. B. (2002). The 1952 Lynmouth floods revisited. *Weather, 57*(7), 235-242, doi:10.1256/004316502760195894.
McPherson, T. N., & Brown, M. J. (2004). *Estimating daytime and nighttime population distributions in US cities for emergency response activities.* Paper presented at the Symposium on Planning, Nowcasting and Forecasting in the Urban Zone, 84th AMS Annual Meeting, Seattle, 12 January 2004.
McPherson, T. N., Rush, J.F., Khalsa, H., Ivey, A. and Brown, M.J. (2006). *A day-night population exchange model for better exposure and consequence management assessments.* Paper presented at the Sixth Symposium on the Urban Environment American Meteorological Society Atlanta, USA. ,
Melchers, R. E. (2001). On the ALARP approach to risk management. *Reliability Engineering & System Safety, 71*(2), 201-208, doi:http://dx.doi.org/10.1016/S0951-8320(00)00096-X.

Mennis, J. (2003). Generating surface models of population using dasymetric mapping. *Professional Geographer, 55*(1), 31-42.

Mennis, J., & Hultgren, T. (2006). Intelligent Dasymetric Mapping and Its Application to Areal Interpolation. *Cartography and Geographic Information Science, 33*(3), 179-194, doi:10.1559/152304006779077309.

Merriman, P. (2012). Human geography without time-space1. *Transactions of the Institute of British Geographers, 37*(1), 13-27, doi:10.1111/j.1475-5661.2011.00455.x.

Met Office (2009). Cumbria floods November 2009 — a look back. http://www.metoffice.gov.uk/about-us/who/how/case-studies/cumbria-floods. Accessed 29/11/2012.

Met Office (2011). Flooding — Summer 2007. http://www.metoffice.gov.uk/about-us/who/how/case-studies/summer-2007. Accessed 03/04/2013.

Met Office (2014). Winter storms, December 2013 to January 2014. http://www.metoffice.gov.uk/climate/uk/interesting/2013-decwind. Accessed 01/12/2014.

Miller, H. J. (2005). What about People in Geographical Information Science In P. Fisher, & D. Unwin (Eds.), *Re-presenting GIS*. Chichester: John Wiley and Sons.

Mitas, L., & Mitasova, H. (1999). Spatial Interpolation In P. A. Longley, M. F. Goodchild, D. J. Maguire, & D. W. Rhind (Eds.), *Geographical Information Systems: Principles, Techniques, Management and Applications* (2nd (Abridged) ed.). Hoboken, New Jersey: John Wiley & Sons Ltd.

Mitchell, J. K. (1999a). Natural disasters in the context of megacities In J. K. Mitchell (Ed.), *Crucibles of Hazard: Mega-Cities and Disasters in Transition*. Tokyo: United Nations University Press.

Mitchell, J. K. (1999b). *Crucibles of Hazard: Mega-Cities and Disasters in Transition*. Tokyo: United Nations University Press.

MOD (2009). Army footbridge takes shape in Cumbria (Ministry of Defence). http://www.mod.uk/DefenceInternet/DefenceNews/MilitaryOperations/ArmyFootbridgeTakesShapeInCumbria.htm. Accessed 29/11/2012.

Mohamed, M. (2002). *Embankment Breach Formation and Modelling Methods*. PhD Thesis, Open University.

Mohamed, M., Samuels, P. G., Morris, M. W., & Ghataora, G. S. (2002). Improving the accuracy of prediction of breach formation through embankment dams and flood embankments. *River Flow 2002, Proc., Int. Conf. on Fluvial Hydraulics, 1*.

Morris, S. (2012a). Shopping centre tracking system condemned by civil rights campaigners. http://www.guardian.co.uk/business/2012/jan/04/shopping-centre-tracking-system-condemned. Accessed 20/03/2012.

Morris, S. (2012b). Village left to mop up again as new flood defences fail. http://www.guardian.co.uk/uk/2012/nov/25/village-new-flood-defences-fail. Accessed 10/12/2012.

Morrisons (2013). *Annual report and financial statements 2012/13*. Bradford, UK: Wm Morrison Supermarkets PLC.

Mugglin, A. S., & Carlin, B. P. (1998). Hierarchical modeling in geographic information systems: population interpolation over incompatible zones. *Journal of Agricultural, Biological, and Environmental Statistics, 3*(2), 111-130.

Munich Re (2012). Natural catastrophes 2011. Analyses, assessments, positions. *Topics Geo*. München, Germany.
Munich Re (2014). Natural catastrophes 2013: Analyses, assessments, positions. *Topics Geo*. München, Germany.
Muniz, I., Galindo, A., & Garcia, M. A. (2003). Cubic Spline Population Density Functions and Satellite City Delimitation: The Case of Barcelona. *Urban Studies, 40*(7), 1303-1321, doi:10.1080/0042098032000084613.
Murphy, J. M., Sexton, D. M. H., Jenkins, G. J., Boorman, P. M., Booth, B. B. B., Brown, C. C., Clark, R. T., Collins, M., Harris, G. R., *et al.* (2009). *UK Climate Projections science report: Climate change projections* (v3 ed.). Exeter: Met Office Hadley Centre
Murray, S. J., Smith, A. D., & Phillips, J. C. (2012). A Modified Flood Severity Assessment for Enhanced Decision-Support: Application to the Boscastle Flood of 2004. *Weather and Forecasting, 27*(5), 1290-1297, doi:10.1175/waf-d-12-00033.1.
National Rail Timetable (2014). *Table 135: London and Birmingham - Devon and Cornwall*. London: Network Rail.
National Statistics (2004). *Census 2001 definitions*. London: HMSO.
National Statistics (2014). The Taking Part survey: statistical release. https://www.gov.uk/government/statistics/taking-part-201314-quarter-4-statistical-release. Accessed 15/10/2014.
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), 88-95, doi:10.1111/j.1753-318X.2011.01093.x.
Newing, A., Clarke, G., & Clarke, M. (2013a). Visitor expenditure estimation for grocery store location planning: a case study of Cornwall. *The International Review of Retail, Distribution and Consumer Research, 23*(3), 221-244, doi:10.1080/09593969.2012.759612.
Newing, A., Clarke, G., & Clarke, M. (2013b). Identifying seasonal variations in store level visitor grocery demand. *International Journal of Retail & Distribution Management, 41*(6), 477-492, doi:10.1108/09590551311330843.
Newing, A. (2014). Tourist origins and seasonal overnight visitor numbers - Cornwall (Personal communication). School of Geography, University of Leeds, UK.
NFDC (2009). Cabinet 06/07/09 Report A: North Solent Shoreline Management Plan. http://www.newforest.gov.uk/committeedocs/cab/CDR05054.pdf. Accessed 22/03/2013.
Nicholls, R. J. (1995). Coastal Megacities and Climate Change. *GeoJournal, 37*(3), 369-379.
NISRA (2015). Combined Grid Square Product. http://www.nisra.gov.uk/geography/IGC.htm. Accessed 23/03/2015.
Nordbeck, S., & Rystedt, B. (1970). Isarithmic Maps and the Continuity of Reference Interval Functions. *Geografiska Annaler. Series B, Human Geography, 52*(2), 92-123, doi:10.2307/490673.
NRCC (2014). Blue Kenue: Software tool for hydraulic modellers. http://www.nrc-cnrc.gc.ca/eng/solutions/advisory/blue_kenue_index.html. Accessed 08/05/2014.

Oliver, A., & Owen, E. (2007). Scour threat to Ulley Dam. http://www.nce.co.uk/scour-threat-to-ulley-dam/72113.article. Accessed 27/08/2014.
ONS (2004). *Census 2001 definitions*. London: HMSO.
ONS (2006a). Annual Business Inquiry. Office for National Statistics.
ONS (2006b). National Statistics Postcode Directory Office for National Statistics.
ONS (2009). *UK Standard Industrial Classification of Economic Activities 2007 (SIC 2007) Structure and explanatory notes*. Cardiff: Office for National Statistics.
ONS (2011). Mid-2006 Population Estimates: England and Wales (2011). Office for National Statistics.
ONS (2012a). Census 2001: Output Geography - Output Areas. http://www.ons.gov.uk/ons/guide-method/census/census-2001/data-and-products/output-geography/output-areas/index.html. Accessed 25/05/2012.
ONS (2012b). Output geography policy. http://www.ons.gov.uk/ons/guide-method/census/2011/how-our-census-works/how-we-took-the-2011-census/how-we-planned-for-data-delivery/output-geography/output-geography-policy/index.html. Accessed 04/09/2012.
ONS (2012c). *Living Costs and Food Survey: Household expenditure 2001-02 to 2012 COICOP based on current prices (Table 4.3)*. Newport: Office for National Statistics.
ONS (2013a). Information paper: quality and methodology information. Newport: Office for National Statistics.
ONS (2013b). *Information Report: Population Weighted Centroids Guidance*. Newport: Office for National Statistics.
ONS (2014a). Postal Geography. http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/postal/index.html. Accessed 09/12/2014.
ONS (2014b). Out of Term Population of England and Wales: An Alternative 2011 Census Population Base, Office for National Statistics. http://www.ons.gov.uk/ons/dcp171776_377904.pdf.
ONS (2014c). Coastal Communities 2011 Census, Office for National Statistics. http://www.ons.gov.uk/ons/dcp171776_382084.pdf.
Openshaw, S. (1984). *The modifiable areal unit problem*. Norwich: Geo Books.
Orcutt, G. H. (1957). A New Type of Socio-Economic System. *The Review of Economics and Statistics, 39*(2), 116-123.
Ordance Survey (2012). AddressBase Product: User guide. http://www.ordnancesurvey.co.uk/docs/user-guides/addressbase-products-user-guide.pdf. Accessed 16/08/2013.
Path Intelligence (2010). FootPath. http://www.pathintelligence.com/en/products/footpath/about-footpath. Accessed 20/03/2012.
Patterson, L., Urban, M., Myers, A., Bhaduri, B., Bright, E., & Coleman, P. (2007). Assessing spatial and attribute errors in large national datasets for population distribution models: a case study of Philadelphia county schools. *GeoJournal, 69*(1), 93-102, doi:10.1007/s10708-007-9099-3.
Peduzzi, P. (2006). The Disaster Risk Index: Overview of a quantitative approach In J. Birkmann (Ed.), *Measuring Vulnerability to Natural Hazards : Towards Disaster Resilient Societies*. Tokyo: United Nations University Press.

Peduzzi, P., Dao, H., Herold, C., & Mouton, F. (2009). Assessing global exposure and vulnerability towards natural hazards: the Disaster Risk Index. *Nat. Hazards Earth Syst. Sci., 9*(4), 1149-1159, doi:10.5194/nhess-9-1149-2009.

Penning-Rowsell, E., Floyd, P., Ramsbottom, D., & Surendran, S. (2005). Estimating Injury and Loss of Life in Floods: A Deterministic Framework. *Natural Hazards, 36*(1-2), 43-64, doi:10.1007/s11069-004-4538-7.

Per Olof, H. (1991). Review: New Paths for Time-Geography? *Geografiska Annaler. Series B, Human Geography, 73*(3), 199-207.

Perry, R. (2007). What Is a Disaster? In *Handbook of Disaster Research* (pp. 1-15, Handbooks of Sociology and Social Research): Springer New York.

Perry, R. W., & Lindell, M. K. (2003). Preparedness for Emergency Response: Guidelines for the Emergency Planning Process. *Disasters, 27*(4), 336-350, doi:10.1111/j.0361-3666.2003.00237.x.

Pitt, M. (2008). *The Pitt review: Learning lessons from the 2007 floods*. London: Cabinet Office.

Porter, N. N. (2012). Past failures and design lessons In P. B. Sayers (Ed.), *Flood Risk: Planning, design and management of flood defence infrastructure*. London: ICE Publishing.

Pred, A. (1977). The Choreography of Existence: Comments on Hägerstrand's Time-Geography and Its Usefulness. *Economic Geography, 53*(2), 207-221.

Pred, A. (1978). The Impact of Technological and Institutional Innovations on Life Content: Some Time-Geographic Observations. *Geographical Analysis, 10*(4), 345-372, doi:10.1111/j.1538-4632.1978.tb00664.x.

Pred, A. (1981). Social Reproduction and the Time-Geography of Everyday Life. *Geografiska Annaler. Series B, Human Geography, 63*(1), 5-22.

Pred, A. (1984). Place as Historically Contingent Process: Structuration and the Time-Geography of Becoming Places. *Annals of the Association of American Geographers, 74*(2), 279-297, doi:10.1111/j.1467-8306.1984.tb01453.x.

Priest, S., Wilson, T., Tapsell, S., Penning-Rowsell, E., Viavattene, C., & Fernandez-Bilbao, A. (2007). Building a model to estimate Risk to Life for European flood events - Final Report. HR Wallingford.

Puente, S. (1999). Social vulnerability to disasters in Mexico City: An assessment method In J. K. Mitchell (Ed.), *Crucibles of Hazard: Mega-Cities and Disasters in Transition*. Tokyo: United Nations University Press.

Quinn, N. (2014). LISFLOOD dataset for St Austell, Cornwall, UK. (Personal communication). School of Geographical Sciences, University of Bristol.

Raper, J. (2000). *Multidimensional Geographic Information Science*. London: Taylor and Francis.

Raper, J., Miller, H. J., Guhathakurta, S., Muetzelfeldt, R., & Cheng, T. (2005). Time As Well: An Introduction In P. Fisher, & D. Unwin (Eds.), *Re-presenting GIS*. Chichester: John Wiley and Sons.

Rase, W.-D. (2001). Volume-preserving interpolation of a smooth surface from polygon-related data. *Journal of Geographical Systems, 3*(2), 199-213, doi:10.1007/pl00011475.

Remondo, J., González-Díez, A., De Terán, J., & Cendrero, A. (2003). Landslide Susceptibility Models Utilising Spatial Data Analysis Techniques. A Case Study from the Lower Deba Valley, Guipuzcoa

(Spain). *Natural Hazards, 30*(3), 267-279, doi:10.1023/B:NHAZ.0000007202.12543.3a.
RGS (2014). *Small area data: Looking towards a 2021 census*. London: Royal Geographical Society.
Ride, A., & Bretherton, D. (2011). *Community Resilience in Natural Disasters*. New York: Palgrave Macmillan.
RMS (2007a). *U.K. Summer 2007 Floods*. Newark, CA, USA: Risk Management Solutions (RMS).
RMS (2007b). *1947 U.K. River Floods: 60-Year Retrospective*. Newark, CA, USA: Risk Management Solutions (RMS).
Roberts, A., Blunt, I., & Bardsley, M. (2014). Focus On: Distance from home to emergency care. London: The Health Foundation and the Nuffield Trust.
Rouhani, S., & Myers, D. (1990). Problems in space-time kriging of geohydrological data. *Mathematical Geology, 22*(5), 611-623, doi:10.1007/bf00890508.
Rowe, D. (2004). *Boscastle: 16th August 2004 - the day of the flood*. Penryn: Truran.
Ruocco, A., Nicholls, R., Haigh, I., & Wadey, M. (2011). Reconstructing coastal flood occurrence combining sea level and media sources: a case study of the Solent, UK since 1935. *Natural Hazards, 59*(3), 1773-1796, doi:10.1007/s11069-011-9868-7.
Sainsbury's (2014). *Annual Report and Financial Statements 2014*. London: J Sainsbury plc.
Schmitt, R. C. (1956). Estimating Daytime Populations. *Journal of the American Institute of Planners, 22*(2), 83-85.
Sekovski, I., Newton, A., & Dennison, W. C. (2012). Megacities in the coastal zone: Using a driver-pressure-state-impact-response framework to address complex environmental problems. *Estuarine, Coastal and Shelf Science, 96*(0), 48-59, doi:http://dx.doi.org/10.1016/j.ecss.2011.07.011.
Shaw, S.-L., Yu, H., & Bombom, L. S. (2008). A Space-Time GIS Approach to Exploring Large Individual-based Spatiotemporal Datasets. *Transactions in GIS, 12*(4), 425-441, doi:10.1111/j.1467-9671.2008.01114.x.
Shina, C., Sinha, V. d. S., Zinken, J., & Sampaio, W. (2011). When Time is not Space: The social and linguistic construction of time intervals and temporal event relations in an Amazonian culture. *Language and Cognition, 3*(1), 137-169.
Sinton, D. (1978). The inherent structure of information as a constraint to analysis: Mapped thematic data as a case study. *Harvard Papers on GIS, 7*.
Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Cognitive processes and societal risk taking. *Decision Making and Change in Human Affairs, 16*, 7-36.
Slovic, P. (2000). *The Perception of Risk*. London: Earthscan.
Smith, A. D., Martin, D., & Cockings, S. (2014a). Spatio-Temporal Population Modelling for Enhanced Assessment of Urban Exposure to Flood Risk. *Applied Spatial Analysis and Policy*, 1-19, doi:10.1007/s12061-014-9110-6.
Smith, A. D., Goff, C. A., & Panzeri, M. (2014b). Enhancements in reservoir flood risk mapping: example application for Ulley In A. Pepper (Ed.), *Maintaining the Safety of our Dams and Reservoirs* (pp. 295-306). London: ICE Publishing.

Smith, C., & Lawson, N. (2012). Identifying extreme event climate thresholds for greater Manchester, UK: examining the past to prepare for the future. *Meteorological Applications, 19*(1), 26-35, doi:10.1002/met.252.
Smith, G., Arnot, C., Fairburn, J., & Walker, G. (2005). *A National Population Data Base for Major Accident Hazard Modelling*. London: HMSO.
Smith, K., & Petley, D. (2009). *Environmental Hazards: Assessing risk and reducing disaster* (5th ed.). London: Routledge.
Sorokin, P. A., & Merton, R. K. (1937). Social Time: A Methodological and Functional Analysis. *American Journal of Sociology, 42*(5), 615-629.
Statistics Austria (2013). Regional statistical grid units. http://www.statistik.at/web_en/classifications/regional_breakdown/grid/. Accessed 23/03/2015.
Sturcke, J., Weaver, M., & Wainwright, M. (2007). More evacuations as floods threaten to burst dam, The Guardian, 26 June 2007.
Susman, P., O'Keefe, P., & Wisner, B. (1984). Global disasters: a radical interpretation In K. Hewitt (Ed.), *Interpretations of calamitiy*. Boston: Allen and Unwin.
Sustrans (2006). *Shoppers and how they travel*. Bristol: Sustrans.
Tafazolli, R. (2014). *5G Innovation Centre: The UK oppertunities.* Paper presented at the UK Spectrum Policy Forum,
Tait, S. J., Ashley, R. M., Cashman, A., Blanksby, J., & Saul, A. J. (2008). Sewer system operation into the 21st century, study of selected responses from a UK perspective. *Urban Water Journal, 5*(1), 79-88, doi:10.1080/15730620701737470.
Tatem, A. J., Noor, A. M., Von Hagen, C., Di Gregorio, A., & Hay, S. I. (2007). High resolution population maps for low income nations: combining land cover and census in East Africa. *PLoS One, 2*(12), e1298.
Tatem, A. J., Huang, Z., Narib, C., Kumar, U., Kandula, D., Pindolia, D. K., Smith, D. L., Cohen, J. M., Graupe, B.*, et al.* (2014). Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning. *Malaria journal, 13*(1), 52.
Taylor, J., Biddulph, P., Davies, M., & Lai, K. m. (2013). Predicting the microbial exposure risks in urban floods using GIS, building simulation, and microbial models. *Environment International, 51*(0), 182-195, doi:http://dx.doi.org/10.1016/j.envint.2012.10.006.
Tesco (2014). *Annual Report and Financial Statements 2014*. Cheshunt, UK: Tesco plc.
Thrift, N. (1977). *An Introduction to Time Geography* (Concepts and Techniques in Modern Geography). Norwich: Geo Abstracts.
Thurstain-Goodwin, M., & Unwin, D. (2000). Defining and delineating the central areas of towns for statistical monitoring using continuous surface representations. *Transactions in GIS, 4*(4), 305-317.
Thurstain-Goodwin, M. (2003). Data surfaces for a new policy geography In P. A. Longley, & M. Batty (Eds.), *Advanced Spatial Analysis: The CASA book of GIS*. California: ESRI Press.
Tobin, G. A., & Montz, B. E. (1997). *Natural Hazards: Explanation and Integration*. New York: The Guilford Press.
Tobler, W., Deichmann, U., Gottsegen, J., & Maloy, K. (1997). World population in a grid of spherical quadrilaterals. *International Journal of Population Geography, 3*(3), 203-225, doi:10.1002/(sici)1099-1220(199709)3:3<203::aid-ijpg68>3.0.co;2-c.

Tobler, W. R. (1979). Smooth Pycnophylactic Interpolation for Geographical Regions. *Journal of the American Statistical Association, 74*(367), 519-530, doi:10.1080/01621459.1979.10481647.

Torrieri, F., Concilio, G., & Nijkamp, P. (2002). Decision Support Tools for Urban Contingency Policy. A Scenario Approach to Risk Management of the Vesuvio Area in Naples, Italy. *Journal of Contingencies and Crisis Management, 10*(2), 95-112, doi:10.1111/1468-5973.00185.

Trafficmaster (2012). Real time traffic information. http://www.trafficmaster.co.uk/content/1/60/real-time-traffic-information.html. Accessed 15/04/2012.

Twitter (2015). About. https://about.twitter.com/company. Accessed 11/03/2015.

UK Parliament (1994). *Sunday Trading Act*. London: HMSO.

UK Parliament (2012). Flooding. http://www.parliament.uk/topics/flooding.htm. Accessed 16/02/2013.

UKDS (2014). GeoConvert - online geography matching and conversion tool. UK Data Service (UKDS). http://geoconvert.mimas.ac.uk/. Accessed 10/08/2014.

UN (2008). *Urban Agglomerations 2007*. New York: United Nations. Department of Economic and Social Affairs, Population Division.

UNDRO (1979). *Natural Disasters and Vulnerability Analysis Report of Expert Group Meeting (9-12 July 1979)*. Geneva: UNDRO.

UNDRO (1982). *Natural disasters and vulnerability analysis*. Geneva, Switzerland: Office for the United Nations Disaster Relief Co-coordinator.

UNISDR (2009). *The United Nations International Strategy for Disaster Reduction (UNISDR) Terminology on Disaster Risk Reduction*. Geneva, Switzerland: UNISDR.

USDOI (2011). Design Standard No. 13: Embankment Dams Static Stability Analysis, US Department of the Interior Bureau of Reclamation. http://www.usbr.gov/pmts/tech_services/engineering/design/DS13-4.pdf. Accessed 01/09/2014.

van Westen, C. J., Rengers, N., & Soeters, R. (2003). Use of Geomorphological Information in Indirect Landslide Susceptibility Assessment. *Natural Hazards, 30*(3), 399-419, doi:10.1023/B:NHAZ.0000007097.42735.9e.

Visit Cornwall (2011). Value of Tourism 2011 in Cornwall. http://www.visitcornwall.com/sites/default/files/generic_files/Value%20of%20tourism-2011.pdf. Accessed 19/09/2014.

Visit Cornwall (2014). Bodmin Moor and Tarmar Valley. http://www.visitcornwall.com/destinations/bodmin-moor-tamar-valley. Accessed 13/10/2014.

Visit England (2010a). Tourism in English Counties. http://www.visitengland.org/Images/County%20Analysis_tcm30-19530.pdf. Accessed 19/09/2014.

Visit England (2010b). Annual Survey of Visits to Visitor Attractions 2010.

von Glasow, R., Jickells, T., Baklanov, A., Carmichael, G. R., Church, T. M., Gallardo, L., Hughes, C., Kanakidou, M., Liss, P. S., *et al.* (2012). *Megacities in the Coastal Zone.* Paper presented at the EGU General Assembly, Vienna, Austria, 22-27 April 2012.

Wachowicz, M. (1999). *Object-Oriented Design for Temporal GIS*. London: Taylor and Francis.

Wadey, M. P., Nicholls, R. J., & Hutton, C. (2012). Coastal Flooding in the Solent: An Integrated Analysis of Defences and Inundation. *Water, 4*(2), 430-459.
Wahl, T. (2004). Uncertainty of Predictions of Embankment Dam Breach Parameters. *Journal of Hydraulic Engineering, 130*(5), 389-397, doi:doi:10.1061/(ASCE)0733-9429(2004)130:5(389).
Walker, B., Holling, C. S., Carpenter, S. R., & Kinzig, A. (2004). Resilience, Adaptability and Transformability in Social–ecological Systems. *Ecology and Society, 9*(2), 5.
Walker, J. P. (2008). The discontinuance of Boltby reservoir, North Yorkshire, UK. *Dams and Reservoirs, 18*, 17-21.
Wang, D., & Cheng, T. (2001). A spatio-temporal data model for activity-based transport demand modelling. *International Journal of Geographical Information Science, 15*(6), 561-585, doi:10.1080/13658810110046934.
Warren, A. L., & Stewart, E. J. (2008). The Implications of the 2007 Summer Storms for UK Reservoir Safety In H. Hewlett (Ed.), *Ensuring Reservoir Safety into the Future*. London: Thomas Telford.
Welander, P. (1961). Numerical Prediction of Storm Surges In H. E. Landsberg, & J. Mieghem (Eds.), *Advances in Geophysics volume 8* (pp. 316). New York: Academic Press Inc.
WestQuay (2011). Media Pack. http://www.west-quay.co.uk/assets/Documents/Media%20Pack%202011.pdf. Accessed 12/10/2012.
Wilby, R. L., Beven, K. J., & Reynard, N. S. (2008). Climate change and fluvial flood risk in the UK: more of the same? *Hydrological Processes, 22*(14), 2511-2523, doi:10.1002/hyp.6847.
Williams, R., & Glendinning, L. (2007). Three dead and 1,000 evacuated as floods strike. http://www.theguardian.com/uk/2007/jun/26/topstories3.weather. Accessed 05/09/2014.
Wisner, B., Blaikie, P., Cannon, T., & Davis, I. (2004). *At Risk: Natural hazards, people's vulnerability and disasters* (2nd ed.). London: Routledge.
Wood, D., & Fels, J. (1986). Designs on sign/myth and meaning in maps. *Cartographica, 23*(3), 54-103.
Wright, J. K. (1936). A Method of Mapping Densities of Population: With Cape Cod as an Example. *Geographical Review, 26*(1), 103-110.
Wu, S.-s., Qiu, X., & Wang, L. (2005). Population Estimation Methods in GIS and Remote Sensing: A Review. [10.2747/1548-1603.42.1.80]. *GIScience & Remote Sensing, 42*(1), 80-96.
Wu, S.-s., Wang, L., & Qiu, X. (2008). Incorporating GIS Building Data and Census Housing Statistics for Sub-Block-Level Population Estimation. *The Professional Geographer, 60*(1), 121-135, doi:10.1080/00330120701724251.
Yu, H. (2006). Spatio-temporal GIS Design for Exploring Interactions of Human Activities. *Cartography and Geographic Information Science, 33*(1), 3-19, doi:10.1559/152304006777323136.
Yu, H., & Shaw, S. L. (2008). Exploring potential human activities in physical and virtual spaces: a spatio- temporal GIS approach. *International Journal of Geographical Information Science, 22*(4), 409-430, doi:10.1080/13658810701427569.
Yuan, M. (1996). Modelling semantic, temporal and spatial information in geographic information systems. In M. Craglia, & H. Couclelis (Eds.),

Geographic Information Research: Bringing the Atlantic. London: Taylor and Francis.

Zevenbergen, C., Veerbeek, W., Gersonius, B., & Van Herk, S. (2008). Challenges in urban flood management: travelling across spatial and temporal scales. *Journal of Flood Risk Management, 1*(2), 81-88, doi:10.1111/j.1753-318X.2008.00010.x.

Zhong, C., Arisona, S. M., Huang, X., Batty, M., & Schmitt, G. (2014). Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science, 28*(11), 2178-2199, doi:10.1080/13658816.2014.914521.