Park Prioritisation: Building a Composite Indicator of Green Infrastructure Demand in London

Greg Slater

August 2019

Submitted in part requirement for the Master of Science in Spatial Data Science and Visualisation at the Centre for Advanced Spatial Analysis, Bartlett Faculty of the Built Environment, University College London.

Submission Date: August 30th 2019

Supervisor: Dr Thomas P. Oléron-Evans

Module ID: CASA0012

Word Count: 11,922

Abstract

Policy documents at the national and local scale are increasingly recognising the value of Green Infrastructure in simultaneously delivering a wide range of different benefits. Many city policy documents, and in particular the most recent London Plan, are thus calling for GI interventions to help meet objectives as diverse as mitigating the risk of adverse weather from climate change or helping to improve the physical and mental health of urban populations. This research aims to investigate how such interventions might be optimised, through identifying the highest priority areas and also analysing these patterns of demand in order to unearth opportunities for multifunctional approaches.

Using the theoretical grounding of ecosystem services literature, the services that urban GI delivers are first defined and a wide range of open data sourced which can be used to represent demand for these services. These datasets are then combined in to a single measure of demand for GI. This process uses the methods of composite indicator construction to create a model of cumulative demand across all services. Once complete, uncertainty and sensitivity analysis are conducted to determine the robustness of the model results.

The overall results show that GI demand is focussed in the centre of the city, due to generally higher relative levels of pollution, ambient noise and temperature. But the model is also successful in identifying hotspots further away from the centre where GI interventions might have the greatest impact from being targeted to deliver health-related benefits. Results of cluster analysis using decomposed indicator scores from the composite also highlight a number of focussed priority areas for particular services, for example flood-risk mitigation to the west of the city-centre.

Declaration

I, Greg Slater, hereby declare that this dissertation is all my own original work and that all sources have been acknowledged. It is 11,922 words in length, from the introduction to conclusion inclusive, excluding figures.

Signed:

Date: August 30th 2019

Contents

1 Introduction **10**

1.1 Context and Motivation 10

1.2 Research Questions 11

1.3 Scope 11

1.4 Ethical Considerations 12

1.5 Report Outline 12

2 Literature Review **13**

2.1 Defining Ecosystem Services 13

2.2 Applications to Urban GI 14

2.3 ES Challenges 16

3 Methodology **17**

3.1 Outline 17

3.2 Theoretical Framework 18

3.3 Data Selection 19

3.4 Data Treatment 20

3.4.1 Cultural Indicator Datasets 21

3.4.2 Temperature 21

3.4.3 Air Pollution 23

3.4.4 Ambient Noise 23

3.4.5 Risk of Surface Water Flooding 23

3.5 Multivariate Analysis 23

3.5.1 Shrinkage Estimation 27

3.6 Normalisation 32

3.7 Weighting 33

3.7.1 Principal Component Analysis 34

3.7.2 Health Results 34

3.7.3 Pollution Results 35

3.7.4 Final Weighting Approach 35
3.8 Uncertainty and Sensitivity Analysis 38
3.8.1 Uncertainty Analysis 38
3.8.2 Sensitivity Analysis 42
3.8.3 Discussion 46
3.9 Results Analysis Methods 47

4 Results 48

4.1 GI Demand 48
4.1.1 Decomposition 52
4.1.2 Clustering 55
4.2 Discussion 59
4.3 Limitations and Developments 60

5 Conclusion 62

A Scripts 69

A.1 Data Processing 69
A.1.1 air.r 69
A.1.2 flood.r 70
A.1.3 health.r 71
A.1.4 noise.r 74
A.1.5 temperature.r 75
A.2 Index Methods 77
A.2.1 dataConsolidate.r 77
A.2.2 dataShrinkage.r 78
A.2.3 dataPcaHealth.py 79
A.2.4 dataPcaAir.py 80
A.2.5 dataAggregation.r 80
A.3 Index Testing 81
A.3.1 functions.r 81
A.3.2 testingUA.r 82
A.3.3 testingSA.r 86

List of Figures

3.1 Conceptual data structure of a composite indicator 20
3.2 Descriptive statistics for final datasets 24
3.3 Distributions of final datasets 25
3.4 Correlation plot for final datasets 26
3.5 Health indicator rates and area populations 28
3.6 Example neighbourhood definitions for Westminster LSOA 29
3.7 Distribution of neighbours per LSOA for different neighbourhood definitions 29
3.8 Indicator distributions after shrinkage estimation using different neighbourhood definitions 30
3.9 Dementia indicator scores in Westminster after shrinkage estimation using different neighbourhood definitions 31
3.10 Effect of exponential transformation on dementia indicator 33
3.11 Process flow of all model steps 37
3.12 LAD ranking for all uncertainty tests (1 = lowest GI demand, 33 = highest GI demand) 39
3.13 LAD ranking for all uncertainty tests - coloured by uncertainty factor $X3$ 41
3.14 LAD ranking for all uncertainty tests - coloured by uncertainty factor $X2$ 41
3.15 Distribution of the standard deviation in mean LSOA ranking across factor values . 42
3.16 Ternary diagram of generated weight samples using different α values for Q_3 . . . 43
3.17 Mean variable weights used in each aggregation stage sensitivity test 44
3.18 Regulating sensitivity - Mean $\bar{R}_s$ by alpha value for each indicator test 45
3.19 Cultural sensitivity - Mean $\bar{R}_s$ by alpha value for each indicator test 46
3.20 GI sensitivity - Mean $\bar{R}_s$ by alpha value for each indicator test 46

4.1 Deciles of final LSOA demand ranking (1 = lowest demand, 10 = highest demand) 49
4.2 London boroughs - for reference 50
4.3 Demand deciles for Westminster 51
4.4 Mean GI demand index score by LAD, with decomposed indicator scores highlighted 52
4.5 Decomposed indicator scores - Westminster 53
4.6 Decomposed indicator scores - Four highest and four lowest scoring LADs 53

4.7 Indicator effect scores for London – proportional contribution to area index score per indicator 54
4.8 Indicator effect scores for Westminster – proportional contribution to area index score per indicator 55
4.9 k-means test results – Silhouette score and SHS for $k = [1, 20]$ 56
4.10 Cluster descriptions - clusters 1 to 5 57
4.11 Cluster descriptions - clusters 6 to 10 58
4.12 Cluster map for Westminster 59

List of Tables

2.1 Ecosystem Service categories from the CICES 14

3.1 List of GI related quotes from relevant policy documents 18

3.2 Urban GI services and benefits 19

3.3 Urban GI services, benefits and demand indicators 20

3.4 Key details of final datasets 22

3.5 PCA results for health data - after varimax rotation 34

3.6 PCA results for air pollution data - before varimax rotation 35

3.7 Input factors used in uncertainty analysis 38

3.8 Number of sample weights used in sensitivity tests for each aggregation stage . . 45

Acronyms

CI	Composite Indicator - also referred to as index
CICES	Common International Classification of Ecosystem Services
EB	Empirical Bayes
ES	Ecosystem Services
GI	Green Infrastructure
IMD	Index of Multiple Deprivation
LAD	Local Authority District
LAEI	London Atmospheric Emissions Inventory
LSOA	Lower Super Output Area
MA	Millennium Ecosystem Assessment
NEA	National Ecosystem Assessment
QOF	Quality Outcomes Framework
SA	Sensitivity Analysis
UA	Uncertainty Analysis
UN	United Nations

Acknowledgements

I would first of all like to thank my supervisor, Dr Thomas P. Oléron-Evans, for his support throughout this research. Thomas's feedback and suggestions have been a great help. Thanks too, go to Kirsten Zeller and Jo Gay at Westminster City Council. Kirsten for her organisational effort and many generous introductions to colleagues. Jo for her time and feedback on the aims of the research and insight into the environmental policy process.

I owe an enormous debt of gratitude to the open source community for the many freely available tools used in this work, and the generosity of the developers who provide such excellent advice on their use in communities like stack overflow. The same goes for those at organisations who have made the effort to publish incredibly rich data sets in standard formats.

Finally, thank-you to the friends and family who have helped maintain my motivation throughout.

Chapter 1

Introduction

1.1 Context and Motivation

Green Infrastructure (GI) can be defined as "a network of green spaces - and features such as street trees and green roofs - that is planned, designed and managed to provide a range of benefits" (GLA, 2012a). The subject of GI has become more prominent in urban policy documents as recognition has grown that the benefits it provides can help tackle the many environmental and social challenges cities face; from improving resilience to higher frequency climate extremes predicted by the IPCC, to reducing the stress levels of rapidly growing urban populations (Elmqvist et al., 2016). The draft London Plan 2019 makes an explicit call for strategic GI interventions to address such challenges, demanding that interventions are planned in an integrated way to "deliver multiple benefits" (GLA, 2019b: p216). Both the London Plan and the Westminster City Plan (City of Westminster, 2019a) list a broad range of policy objectives which GI can help support, from promoting mental and physical health, supporting recreation and encouraging active travel, to improving air and water quality and enhancing biodiversity and ecological resistance. This sentiment is echoed in the UK's National Planning Policy Framework, which calls out GI's potential to simultaneously deliver benefits at a range of scales for individuals, society, the economy and the environment (Ministry of Housing, Communities Local Government, 2016).

What is clear from these policy documents is that GI's multifunctional nature represents a great opportunity for urban policy makers. What seems to be missing, however, is guidance on the targeting of such interventions. It is well known that exposure to various environmental and health problems across different populations is far from equal (Clark et al., 2014; World Health Organisation, 2010). If the problems that GI can help alleviate are not evenly spread across the city it seems it would be worthwhile considering how their distribution – which might be considered a type of GI 'demand' - could inform planning for GI interventions. The planning literature contains some examples of work in this direction, where authors have integrated different

GI benefit criteria into spatial planning models that can be used to identify high-priority areas (Meerow and Newell, 2017) or reveal ‘multifunctionality hotspots’ (Hansen and Pauleit, 2014). The aim of this dissertation is to combine research into such approaches with methods developed for building Composite Indices in order to create a single measure of GI demand. The hope is that doing so might provide a different perspective on the spatial priorities for GI in London, provide some insight into how multifunctionality objectives might be met, and ultimately aid more informed policy discussions on GI interventions.

1.2 Research Questions

These aims outlined above can be summarised in the following two questions:

- Can publicly available data be used to create a comprehensive measure of GI demand across London, which accounts for the many benefits it can deliver?
- Can such a measure provide insight into localised patterns of GI demand in order to highlight opportunities for multifunctional GI approaches?

1.3 Scope

The way to approach answering these questions can be made clearer by defining some boundaries for the research. There is no aim to cover the causal links between GI and the benefits listed in policy documents; this is a vast field of literature in itself spanning many different disciplines. While evidence exists for many of the benefits - see for example van den Berg et al. (2015) for a systematic review of epidemiological studies into health effects of GI – it is accepted that evidence may not be equal for all benefits and the research here must take a necessarily simplified approach to the complex and overlapping issues GI can address. The approach will, however, be grounded on a solid conceptual framework, to be outlined more clearly in the literature review. It should be made clear too that any representation of demand will be purely indicative rather than an empirical measure based on exposure to or use of GI. Rather than any sort of economic model which might seek to assign an actual monetary value to different GI benefits the approach here will be to build a relative measure of different indicators which represent demand for those benefits. In this sense, the demand measure should not be thought of as a definitive solution, rather - as Nardo et al. (2005) suggest in their guide to creating composite indices – it is best considered as a valuable starting point for discussion or a tool to use in a decision-making process.

1.4 Ethical Considerations

There are no active participants involved in this research as all data is secondary and freely available in the public domain. All research outputs will be shared with Westminster City Council in order to inform their GI strategy. While the aim of the work is to improve the outcomes of decisions on GI siting and functionality by considering demand, there is a risk that the model doesn't accurately reflect this. Any risks of negative outcomes as a result are considered to be extremely minimal by the fact that this research will never be the only factor considered in a decision. Further, any changes to GI will follow due process within WCC, be informed by policy officers with detailed knowledge of the areas, and subject to relevant planning procedures.

1.5 Report Outline

The full report follows the structure described below:

- **Chapter 2** – Literature Review – introduces the concept of Environmental Services, briefly explaining its evolution and challenges before considering how it can be used as a theoretical framework to break down and analyse the functions, services and benefits of urban GI in a structured manner.
- **Chapter 3** – Methodology – begins with a brief summary of the composite indicator approach before outlining each of the steps taken in creating the index of GI demand. The aim of this section is to explain both the theory behind each of these various methods and the results of their specific application in this research.
- **Chapter 4** – Results – breaks down and visualises the various outputs of the final composite indicator, drawing out drivers of performance and trends across London. Results of some further analysis using these outputs are also presented before finally considering whether the two research questions have been effectively answered, as well as some of the limitations of the research.
- **Chapter 5** – Conclusion – summarises the overall findings.

Chapter 2

Literature Review

2.1 Defining Ecosystem Services

The concept of Ecosystem Services (ES) is an attempt to understand how ecosystems provide benefits to people, through characterising the ecological structures and processes that support them, in a way that allows them to be analysed (Potschin and Haines-Young, 2016). It is a fundamentally anthropocentric view, as it focusses on the contribution of ecosystems to human welfare, thus making it a fitting lens through which to inform policy-making (DEFRA, 2011). This idea of ecosystems providing 'services' to humans first appeared in a 1981 environmentalist work Extinction, by Paul Elrich (Potschin et al., 2016). But it is in more recent years that the idea has become more mainstream, driven by the influence of the 2005 Millennium Ecosystem Assessment (MA) (ibid.). This project, initiated by UN Secretary General Kofi Annan in 2001, examined the consequences of ecosystem change for human well-being, and outlined the advantages of such a framework:

> 'By examining the environment through the framework of ecosystem services, it becomes much easier to identify how changes in ecosystems influence human well-being and to provide information in a form that decision-makers can weigh alongside other social and economic information'
>
> (Millennium Ecosystem Assessment, 2005)

Since the MA was published the field has expanded rapidly, spurring governments into completing their own assessments, including the UK's National Ecosystem Assessment (Watson and Albon, 2011).

Though there are certainly many differing approaches within the field, it's widely accepted that there is some kind of pathway though which ES are delivered. This pathway is described in the cascade model set out by Potschin and Haines-Young (2016). The 'cascade' begins with a function carried out by a biophysical structure or process in the environment. This function can be characterised as a service that provides a benefit to the social and economic system, and can thus be ascribed some value (ibid.). If we take a wetlands as an example, slowing the passage of water is one function carried out in that ecosystem. We can consider this a service and call it flood protection, which creates human benefit through reduced likelihood of property damage (Hansen and Pauleit, 2014). The MA built a comprehensive list of all such services - classified into different categories - and has become the foundation for other slightly varying frameworks (Elmqvist et al., 2016). Most agree on three top-level categories of services: provisioning, regulating and cultural, and some also include a final category of habitat. Table 2.1 gives the definitions used in the Common International Classification of Ecosystem Services (Haines-Young and Potschin, 2018).

Category	*Description*
Provisioning	All nutritional, non-nutritional material and energetic outputs from living systems as well as abiotic outputs (including water).
Regulation & Maintenance	All the ways in which living organisms can mediate or moderate the ambient environment that affects human health, safety or comfort, together with abiotic equivalents.
Cultural	All the non-material and non-consumptive, outputs of ecosystems that affect physical and mental states of people.

Table 2.1: Ecosystem Service categories from the CICES

This ES framework supports a clear structure in analyses of ecosystems by making it easier to break down and classify their different aspects. Most importantly though, it is the explicit linking of functions all the way through to the benefits humans derive from them that allows the worth of ecosystems to be expressed in terms that are more intelligible to policy-makers and various institutions (Chan et al., 2012).

2.2 Applications to Urban GI

There are many different ways that the Ecosystem Services framework can be applied. Work such as the MA or the UK NEA might be considered as very broad and comprehensive assessments of natural systems. These aim to provide a “coherent body of evidence about the state of our natural environment and the services it provides”, thus enabling the full value of these services to be taken account in decision making (Watson and Albon, 2011: p4). But the ES framework can also be used to support work with a more targeted focus, such as demonstrating the value of a particular type of ecosystem or natural asset like Green Infrastructure. The Natural Capital Accounts (Vivid Economics, 2017) is a project which aimed to attribute a monetary value to the benefits that public parks and green spaces provide to Londoners. It covers a range of services and combines calculations such as the avoidance cost to mental health services and the value added to property prices into an overall value figure of £91 billion. Doing so informs an important debate on the funding for such spaces, which is happening in the context of constraints on public funding for provision and maintenance (ibid.). The London i-Tree Eco Project has similar aims but focusses specifically on London's urban forest, using a breakdown of services from the MA categorisation which can be delivered specifically by this type of ecosystem - though it should be noted that only a subset of these are then measured in economic terms (England Forestry Commission, 2015). The work aims to support more informed decisions about the management and maintenance of this natural asset as well as encourage further investment to improve it.

Decision-support is a theme that appears repeatedly in work using ES, which is a result of its broad and multifunctional approach. Some work makes this decision-making aspect the explicit focus; for example Meerow and Newell (2017) integrate six benefits of urban GI into a spatial planning model. This model allows the incorporation of varying stakeholder priorities, enabling them to "visualise the implications of their preferences and identify trade-offs in policy goals" (ibid: p66). Work in a similar direction by Lovell and Taylor (2013: p1458) uses a framework of ES relevant to urban GI to develop decision-making tools with a participatory focus, the goal being to "identify and represent the interests of all stakeholders", thus driving "more equitable and sustainable outcomes with broad-based public support." These examples demonstrate the value from using ES to integrate knowledge from many diverse fields concerned with green spaces in cities (Elmqvist et al., 2016), allowing users of the resulting analysis to develop a comprehensive understanding of the system in question. This is a clear improvement over previous approaches where urban GI was "researched and implemented from the perspective of a single benefit" (Meerow and Newell, 2017: p63). Hansen and Pauleit (2014) consider three key advantages conferred from this multifunctional approach: multifunctionality hotspots can be unearthed, meaning GI providing multiple services can be prioritised and benefit maximised; it's possible to gain a better understanding of synergies and trade-offs between services, for example through understanding that intensive recreation is unlikely to support the protection of sensitive species in the same space; and finally, the supply and demand of services can be considered to identify shortfalls and improvement areas . This sort of understanding makes the ES framework a useful tool to approach the challenges raised in documents like the London Plan 2019, i.e. that urban GI interventions should address multiple policy objectives in an integrated manner.

There is one final way the ES framework can be applied that is worth considering, and that is demand. It is a critical part of the equation as by their definition of providing benefits to humans, an ecosystem service cannot exist without demand for that benefit (ibid.). The UK NEA considers demand in its comprehensive assessment of UK ES in order to investigate whether our use of ecosystems is sustainable and forecast future scenarios (Watson and Albon, 2011). Burkhard et al. (2012) are interested in the spatial patterns of demand; they use land use types to estimate demand for every ES in the MA classification and analyse how demand has changed in a region of Germany over time. This broad approach considers services from all ecosystems rather than GI specifically, and though comprehensive, the model uses a rather arbitrary scale of demand from 0-5; so a particular type of urban land cover is scored 1 (low demand) for water purification services but 5 (high demand) for air quality regulation services. This highlights the challenge of quantifying demand across so many different domains, and is particularly difficult for non-material (i.e. cultural) ES where demand for recreation or aesthetics is hard to quantify.

2.3 ES Challenges

The ES framework is not totally devoid of challenges. As the field has expanded quickly, it covers many "disciplinary starting points" and as such achieving a common, accepted conceptual framework is difficult (Postchin et al., 2016: p11). The same problem means that terminology is often used interchangeably or terms are conflated (Chan et al., 2012). The broadness of ES means comprehensive analysis involves combining data from many different fields, which presents a number of challenges, perhaps most obviously that data used to measure different services must be comparable (Hansen and Pauleit, 2014). There is often not clear empirical evidence of the cited benefits for a service (Elmqvist et al.,2016), which makes measuring the value of those benefits unreliable. And furthermore, there are many different dimensions across which value could be measured, such as economic, socio-cultural or ecological (Chan et al., 2012). The challenge of quantifying the benefits and value of services under the cultural category has meant that much work arguably has a relatively limited economic focus (ibid.). This is clear from the London-based examples of Natural Capital and iTree project; both reports clearly demonstrate a real value for an important natural resource, but as Chan et al. highlight, it is just one dimension of their value. Reflecting the true value of cultural ES such as 'recreation and relaxation' or 'social cohesion' is problematic as these often intangible values are far less subject to the rules of trade-off, protection, loss or gain which economic based analyses often involve.

Chapter 3

Methodology

3.1 Outline

It is well understood that many socio-economic phenomena are made up of different dimensions; the aim of a Composite Indicator (CI) - also referred to as an index - is to combine these different dimensions into a single measure which might be considered a proxy of the phenomenon (Mazziotta and Pareto, 2013). This practice has become increasingly common as the demand for methods to interpret and consolidate ever-greater amounts of data increases, and the popularity of CIs has been boosted by their adoption by many global institutions (Greco et al., 2019). Despite this popularity, the construction of a CI is no simple task; it is a process of many steps and assumptions which should each be assessed carefully (Saltelli, 2007). One of the most widely cited resources on the processes involved is the OECD's guide, which breaks its technical guidelines into 10 clearly laid out steps which aim to "provide builders of composite indicators with a set of recommendations on how to design, develop and disseminate a composite indicator" (Nardo et al., 2008: p15). The remainder of this section uses the OECD 10-step outline to describe the construction of a CI which aims to measure demand for Green Infrastructure in London, using Environmental Services as the theoretical framework.

It is important to make clear here that this process is not strictly linear; mostly it has been carried out iteratively with adjustments and changes in approach made throughout. Furthermore, the OECD list is by no means a one size fits all approach to building a CI, rather it breaks down the key aims of each step and the many possible ways they can be achieved. It does, however, provide a useful framework to discuss the methods used as well as the reasoning behind their selection and particular application.

3.2 Theoretical Framework

The conceptual framework is a foundational part of a CI as it provides the basis for the selection and combination of variables into a meaningful measure (Nardo et al., 2008). The aim here is to use the ES framework to first define a clear list of the services provided specifically by urban GI. This can then be used to define indicators of demand for each of these services, which will then form the Composite Indicator of overall demand. Here, three relevant policy documents are used to build a list of potential ES delivered by urban GI: the National Planning Policy Framework (Ministry of Housing, Communities and Local Government, 2016), the London Plan 2019 (GLA, 2019b), and the Westminster Open Spaces & Biodiversity Strategy (City of Westminster, 2019b). Table 3.1 shows all quotes which reference a benefit of or aim for GI. Immediately clear is the very broad range of impacts covered and also the lack of distinction between what might be considered a service – e.g. 'play, sport or recreation' - or benefit – e.g. 'improve public health'.

Category	Policy Reference (quoted)	Page no.	Policy Document
Cultural	improve public health	para030	NPPF
	delivering mental and physical health benefits	para030	NPPF
	providing opportunities for recreation and exercise	para030	NPPF
	promoting mental and physical health and well-being	p340	London Plan
	encouraging walking and cycling	p340	London Plan
	supporting landscape and heritage conservation	p340	London Plan
	learning about the environment	p340	London Plan
	play, sport and recreation	p340	London Plan
	providing opportunities for relaxation and interaction, which can impact on other concerns such as loneliness and social isolation	p8	Westminster Open Spaces & Biodiversity Strategy
	promote more active lifestyles ... by providing space for physical activity	p8	Westminster Open Spaces & Biodiversity Strategy
	reduce stress and improve mental health	p10	Westminster Open Spaces & Biodiversity Strategy
	help foster a sense of community	p10	Westminster Open Spaces & Biodiversity Strategy
	allow people to experience a closer relationship with nature	p11	Westminster Open Spaces & Biodiversity Strategy
Regulating	Green infrastructure also helps reduce air pollution, noise and the impacts of extreme heat and extreme rainfall events	para030	NPFF
	adapting to the impacts of climate change and the urban heat-island effect	p213	London Plan
	improving air and water quality	p213	London Plan
	help to cool the city, lowering both surface and air temperatures	p13	Westminster Open Spaces & Biodiversity Strategy
	help to reduce flood risk [from] intense rainfall	p13	Westminster Open Spaces & Biodiversity Strategy
	help improve air quality by absorbing pollutants	p13	Westminster Open Spaces & Biodiversity Strategy
Habitat	supporting food growing and conserving and enhancing biodiversity and ecological resilience	p213	London Plan
	provide habitats for wildlife and support pollinators	p10	Westminster Open Spaces & Biodiversity Strategy

Table 3.1: List of GI related quotes from relevant policy documents

Table 3.2 represents an attempt to consolidate these quotes into a clear breakdown of services and their human benefits, using as a reference a comprehensive list of ES provided by urban GI created by Elmqvist et al. (2016) from both the CICES and MA frameworks.

Category	Ecosystem Service	Benefit
Cultural	Recreation & relaxation	Improved physical and mental health
	Sense of place & social capital	Improved networks and social trust
	Environmental Learning	Improved resilience and adaptive capacity in urban systems
Regulating	Air filtration	Reduced health impacts from air pollution
	Temperature regulation	More amenable living environment
	Surface water capture	Reduced likelihood of flooding and property damage
	Noise Regulation	More amenable living environment
Habitat	Biodiversity conservation	Improved resilience of green spaces

Table 3.2: Urban GI services and benefits

From this list it is now possible to decide how best to represent demand for each of these services. This involves considering what human and environmental data points the services and their benefits impact. For example, considering the air filtration service, demand could be represented as high where air pollution is high and low where pollution is low. Or for recreation and relaxation and the benefit of improved physical health, demand could be represented as high where some measure of physical health is high and low where the same measure is lowest. In doing so the aim is not create a measure with a specific unit, e.g. minutes of GI exposure per day, but rather allow a relative comparison of demand to aid decisions for prioritisation of GI overall as well as particular services. The construction of a CI from this sort of representation is a similar approach to that of the Index of Multiple Deprivation, in the sense that it is an area based model which characterises demand in an area relative to other areas (Smith et al., 2015).

3.3 Data Selection

Upon searching for data to use as demand indicators, one of the key challenges of index building is immediately apparent: availability of data (Mazziotta and Pareto, 2013). This challenge is doubled for cultural services, where representing or measuring the value of benefits is already a challenge. In order to ensure robustness of this index it is necessary to use a somewhat restricted framework (table 3.3), limited to services where it is possible to make a clear logical connection between the service, benefit and demand indicator. Services in the cultural category will be focussed on two of Westminster's key GI policy objectives, physical and mental health (City of Westminster, 2019b). In order to produce results at a granular spatial level, data is further restricted to that available at the level of Lower Super Output Area (LSOA), of which there are 4,835 in London. Figure 3.1

shows the conceptual flow of CI construction, from individual data sets to the final CI. Here, the aim is for each domain to represent an ES category using indicators which represent different services.

Category	Ecosystem Service	Benefit	Demand Indicator
Cultural	Recreation & relaxation	Improved physical health	Obesity prevalence
			Cardiovascular disease prevalence
		Improved mental health	Depression prevalence
			Dementia prevalence
	Sense of place & social capital	Improved networks and social trust	
	Environmental Learning	Improved resilience and adaptive capacity in urban systems	
Regulating	Air filtration	Reduced health impacts from air pollution	Air pollution levels
	Temperature regulation	More amenable living environment	Ambient temperature
	Surface water capture	Reduced likelihood of flooding and property damage	Surface water flood risk
	Noise Regulation	More amenable living environment	Ambient noise
Habitat	Biodiversity conservation	Improved resilience of green spaces	

Table 3.3: Urban GI services, benefits and demand indicators

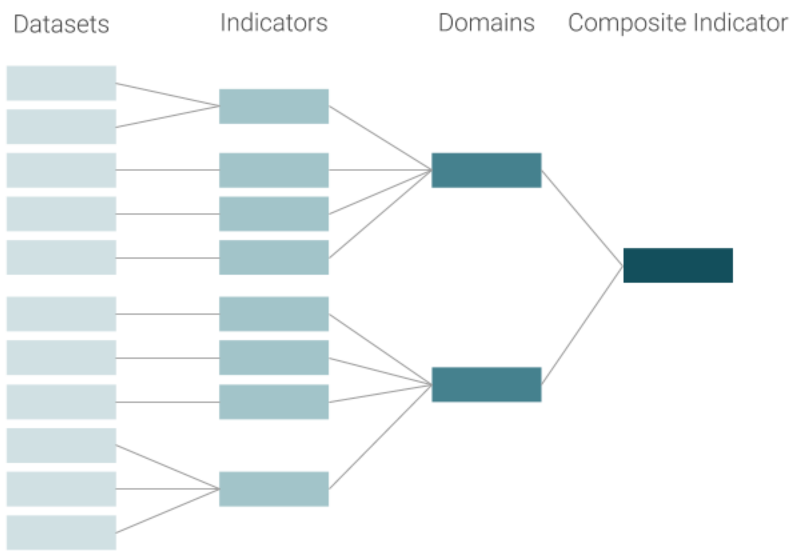

Figure 3.1: Conceptual data structure of a composite indicator

3.4 Data Treatment

Table 3.4 describes key details for the datasets and the following section explains key processing steps for each. The cultural indicator data and each of the regulating indicator datasets were processed differently due to their different sources, formats and spatial scales. For reasons explained further in section 3.5, what is important here is not ensuring that each of the final measures is directly comparable – e.g. all describing the proportion of LSOA area above a certain safe threshold – but that each is a measure of best to worst and allows accurate comparisons between LSOAs.

3.4.1 Cultural Indicator Datasets

The NHS Quality and Outcomes Framework (QOF) has run since 2004, and records data for a number of health indicators. Data is publicly available as numbers of patients for each indicator and GP surgery, and in 2018 had coverage of 94.8% of all GP surgeries (NHS Digital, 2018). By combining this data with GP register-to-area reference data it is possible to transform it to areal prevalence counts per LSOA. Four indicators were chosen for the framework here: obesity and cardiovascular disease because high BMI and high blood pressure have been identified as in the top 10 risks to health globally and are known to be linked to physical activity (GBD 2015 Risk Factor Collaborators et al., 2016), and depression and dementia to align with Westminster's aims to improve mental health. All indicators apart from dementia are measured only for specific age groups.

For each GP surgery in London the distribution of its patient register across LSOAs was calculated and an estimation made for the number of patients per LSOA per indicator. The prevalence rate was then calculated by summing all counts for each LSOA and then dividing by the appropriate age-range population. It should be noted there are two sources of uncertainty here: the surgery to LSOA mapping may not be fully accurate due to changes in addresses; and patient distribution across LSOAs may not be uniform across age-bands as is assumed here. Measures to account for estimation errors are explained in section 3.5.

3.4.2 Temperature

The temperature data used is from 'UrbClim' - a model designed to study the Urban Heat Island effect – run to calculate the mean temperature at midnight for 2011 during an average summer (GLA, 2017). The relative coarseness of the temperature data means it is only possible to generate a measure of mean temperature per LSOA. To process, the raster data was converted to point, and the mean calculated for all points within each LSOA. On average each LSOA contains 5.2 point records, though there are 84 which do not contain any – for these, the mean temperature of all LSOAs sharing a boundary was taken.

Category	Demand Indicator	Data Source	Type	Resolution / Scale	Measure
Cultural	Depression prevalence	NHS Quality and Outcomes Framework Prevalence data 2017 – 2018 NHS GP registration data 2018 ONS LSOA population estimates 2017	count per GP practice	Number of affected patients on GP practice register	Patients per LSOA with new diagnosis of depression / population of indicator age group per LSOA
	Dementia prevalence				Patients per LSOA with new diagnosis of dementia / population of indicator age group per LSOA
	Obesity prevalence				Patients per LSOA with BMI ≥ 30 / population of indicator age group per LSOA
	Cardiovascular disease prevalence				Patients per LSOA with new diagnosis of hypertension / population of indicator age group per LSOA
Regulating	Air pollution levels	London Atmospheric Emissions Inventory modelled 2020 concentrations of NO2	point	20m grid	Mean concentration of NO2 per LSOA
		London Atmospheric Emissions Inventory modelled 2020 concentrations of Nox	point	20m grid	Mean concentration of Nox per LSOA
		London Atmospheric Emissions Inventory modelled 2020 concentrations of PM2.5	point	20m grid	Mean concentration of PM2.5 per LSOA
		London Atmospheric Emissions Inventory modelled 2020 concentrations of PM10	point	20m grid	Mean concentration of PM10 per LSOA
	Ambient temperature	VITO 'UrbClim' simulation of mean temperature at midnight during summer 2011 at 250m resolution	raster	250m grid	Mean ambient temperature per LSOA
	Risk of Surface Water Flooding	DEFRA Risk of Flooding from Surface Water Hazard: 3.3 percent annual chance	polygon	smallest polygon = $4m^2$	Proportion of LSOA area with any risk of surface water flooding
	Ambient noise	DEFRA road noise night time annual average noise level results in dB (night defined as 2300 – 0700)	polygon	smallest polygon = $0.3m^2$	Proportion of LSOA area with noise level above 50dB
		DEFRA rail noise night time annual average noise level results in dB (night defined as 2300 – 0700)	polygon	smallest polygon = $0.2m^2$	Proportion of LSOA area with noise level above 50dB

Table 3.4: Key details of final datasets

3.4.3 Air Pollution

This data is from the 2016 London Atmospheric Emissions Inventory (GLA, 2016), which provides estimates of concentrations for four pollutants at 20m grid resolution. The higher resolution means better coverage, with each LSOA containing 814 points on average. Files for each of the four pollutants were processed in turn by calculating the mean concentration of all points within each LSOA.

3.4.4 Ambient Noise

The data used is night time annual average noise levels in dB for both road and rail noise sources in 2012 (GLA, 2012b). WHO guidelines recommend ambient levels at night should be below 45dB and 44dB for road and rail respectively due to the adverse health effects associated with long-term exposure (World Health Organisation, 2018). Rather than covering the whole of London, the polygon dataset shows areas predicted to be 50dB and above in 5dB bands, which is all above the WHO safe threshold. So for each LSOA the proportion of its area covered by any noise level above 50dB was calculated.

3.4.5 Risk of Surface Water Flooding

DEFRA publish data from a Risk of Surface Water Flooding (RoSWF) model which shows the hazard mapping for “flooding from surface water that could result from a flood with a 3.3% chance of happening in any given year” (DEFRA, 2013). It is at a very high level of detail, though like the noise data it only shows risk areas. For this reason a similar measure was created by calculating the proportion of each LSOA covered by any area of flood risk.

3.5 Multivariate Analysis

Before creating a CI it is important to assess the datasets. Analysis here can provide a structural overview of the data being used, which can inform any necessary error correction as well as the next steps of the methodology (Saltelli, 2007). It is also useful to understand correlation between datasets as ideally, selected indicators should be less correlated in order to minimise redundancy (Mazziotta and Pareto, 2013).

Immediately clear from the summary statistics in table 3.2 is the wide range of scales. This emphasises the importance of standardising each dataset to ensure they can be combined in a meaningful way. The number of missing values in some of the regulating data also stands out, particularly rail noise. This unfortunately is an artefact of the calculation method as the model doesn’t cover the whole study area (rail lines clearly only cover a small proportion of London).

Category	*Indicator*	*Data*	*Data set name*	*Unit*	*Count*	*Min*	*Max*	*Mean*	*Std*	*0 Values*
cultural	mental health	Depression Prevalence	DEP	% of indicator appropriate LSOA population	4,835	0.02055	0.21816	0.08088	0.02082	0
cultural	mental health	Dementia Prevalence	DEM	% of indicator appropriate LSOA population	4,835	0.00044	0.01956	0.00563	0.00217	0
cultural	physical health	Obesity Prevalence	OB	% of indicator appropriate LSOA population	4,835	0.01416	0.27185	0.08806	0.03060	0
cultural	physical health	Cardiovascular Disease Prevalence	CVDPP	% of indicator appropriate LSOA population	4,835	0.00102	0.03602	0.01177	0.00369	0
regulating	air pollution	Mean NO2	no2	microgramme per cubic metre	4,835	18.326	46.614	29.020	3.953	0
regulating	air pollution	Mean NOx	nox	microgramme per cubic metre	4,835	21.428	88.453	42.163	8.237	0
regulating	air pollution	Mean PM10	pm10	microgramme per cubic metre	4,835	20.741	28.607	23.318	1.145	0
regulating	air pollution	Mean PM25	pm25	microgramme per cubic metre	4,835	13.018	17.159	14.339	0.573	0
regulating	ambient noise	Road Noise Risk Area	noise_road	% of area at high risk	4,835	0.000	1.000	0.144	0.172	1,134
regulating	ambient noise	Rail Noise Risk Area	noise_rail	% of area at high risk	4,835	0.000	0.666	0.050	0.096	2,992
regulating	ambient temperature	Ambient Temperature	temp	degrees celcius	4,835	14.262	17.181	16.614	0.379	0
regulating	surface water flood risk	Flood Risk Area	flood	% of area at high risk	4,835	0.000	0.368	0.030	0.035	99

Figure 3.2: Descriptive statistics for final datasets

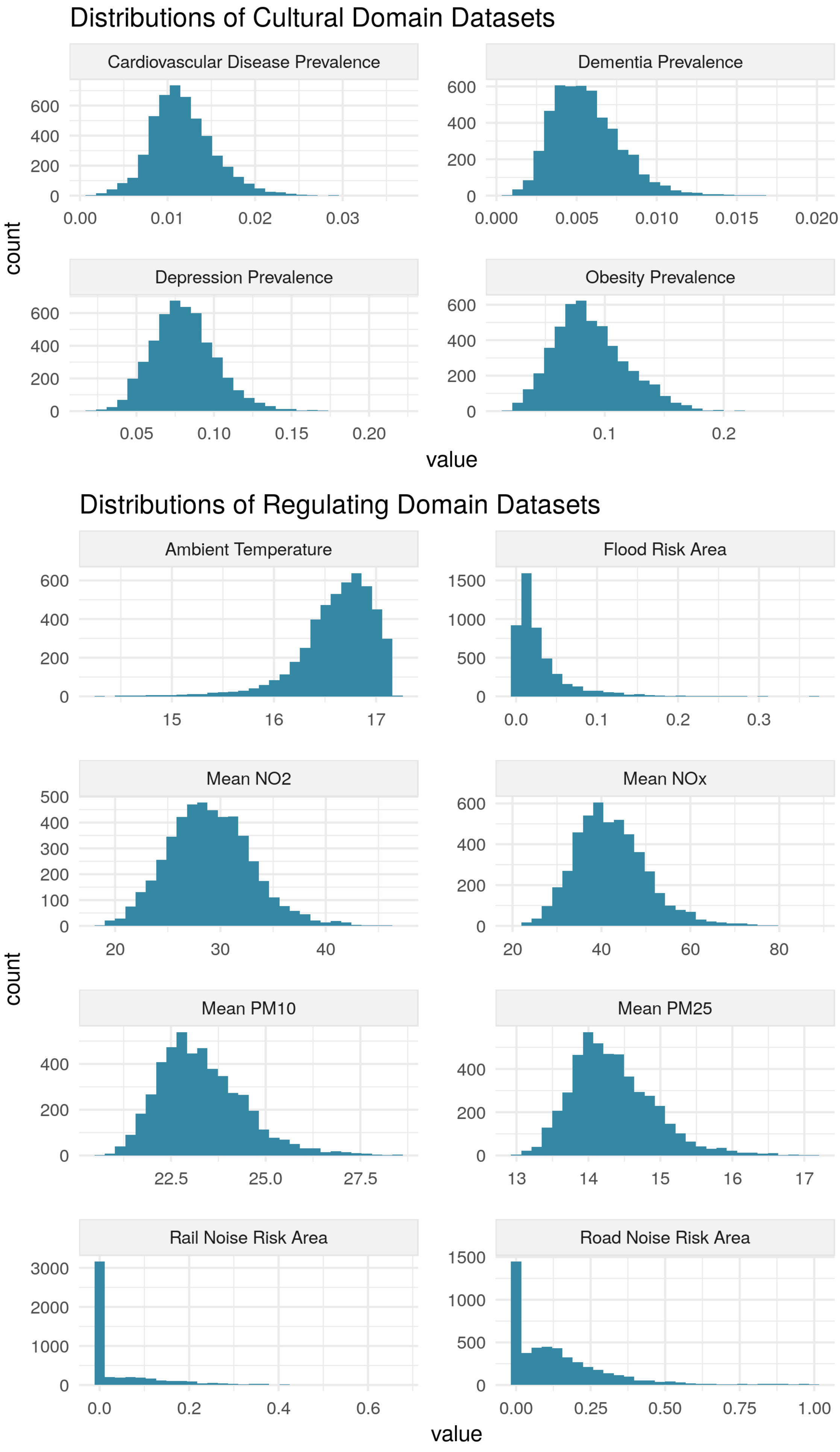

Figure 3.3: Distributions of final datasets

While it is certainly possible to choose a standardisation method in which the effect of 0 values is minimised there is a point beyond which some knock-on impact is inescapable. Over 60% 0 values for the rail noise data was judged too high and it was removed, leaving just road noise as an indication of ambient noise. The proportion of 0 values in the other datasets (23% in road noise and 2% in flood risk) is more manageable.

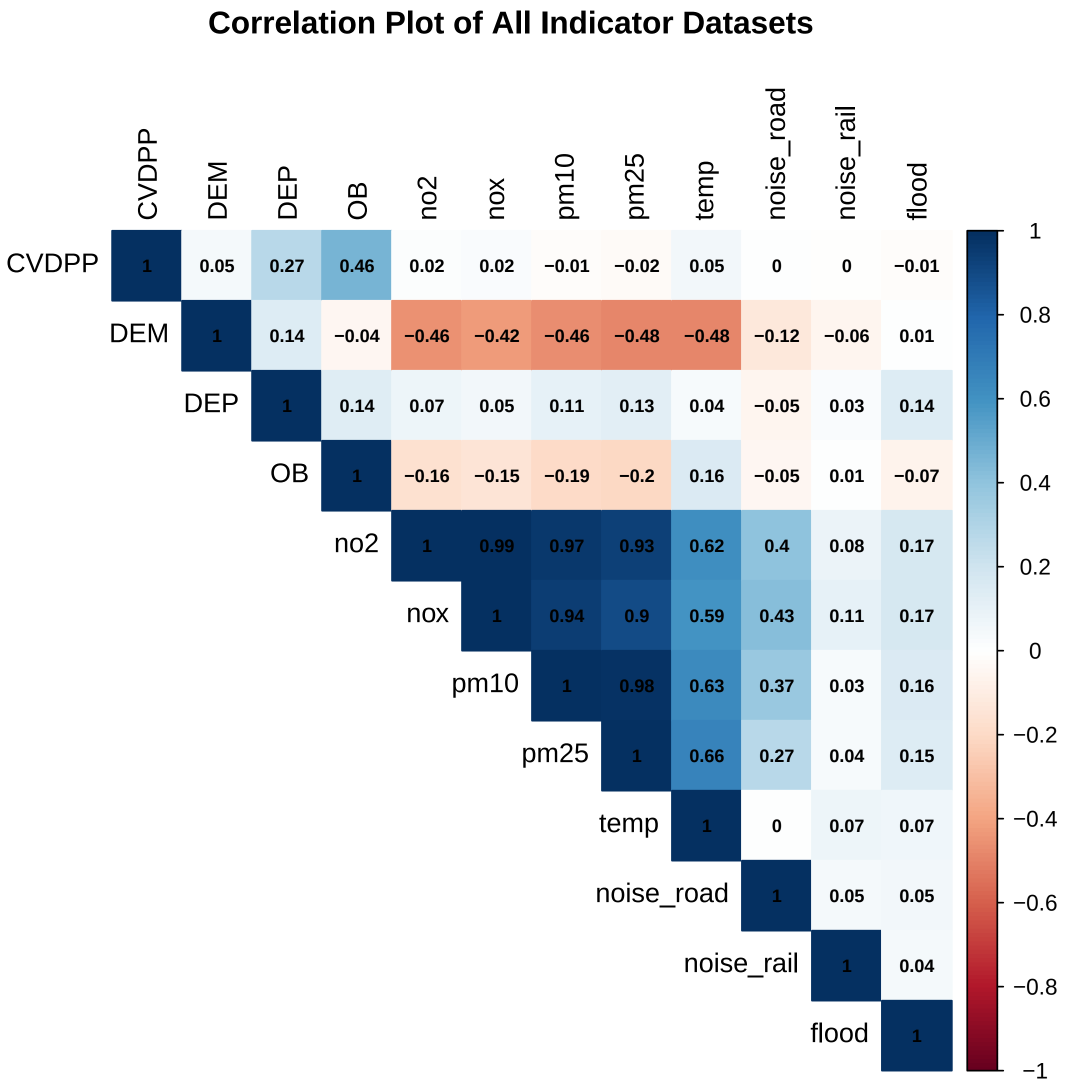

Figure 3.4: Correlation plot for final datasets

The correlation plot in figure 3.4 can help suggest if datasets are representing a wide range of demand, as is the aim. The most obvious pattern is the high correlation between all air pollution data – this is to be expected and suggests these datasets together will make a good indicator for overall air pollution. Interestingly, both the temperature and road noise datasets also show some positive correlation with the air pollution data. This is possibly due to the road network and land use data which is an important factor in each of these modelled datasets (GLA, 2012b; 2016; 2017), and should be considered when deciding the methods to aggregate these indicators into a regulating domain score. Correlation between the cultural indicators is generally low. The exception is between cardiovascular disease and obesity prevalence, possibly suggesting their similar link to general physical health. The negative correlation between dementia and the air pollution and temperature data is also striking and suggests that the results may indeed show some interesting patterns of demand.

Considering the histograms in figure 3.3, the health and air pollution datasets are the most clearly normal distributions, while almost all others are positively skewed – the negative skew of ambient noise being the only exception. Different shapes of distribution should not be a challenge here as they will be manipulated in the normalisation section; as mentioned earlier, what is important is that each dataset give a reliable way to compare scores between LSOAs for each indicator. So what is more important is that the risk of errors in each dataset is considered. When working with data at a granular spatial level one of the most straightforward ways to reduce the impact of measurement error is to aggregate the data to a coarser level (Hampton et al., 2011). For the regulating datasets this is essentially already the case as each have been aggregated from more granular datasets. While this does mean that due to the Modifiable Areal Unit Problem some variance in the original data is likely to be lost (Wong, 2009), it reduces the risk that this variance is from error. The health data, however, presents the opposite challenge – it has in effect been disaggregated to a more granular spatial level by transforming measurements at GP surgery level to LSOA. The next section outlines steps taken to reduce risk with this data.

3.5.1 Shrinkage Estimation

Working with data at small area level can present measurement problems even in very large samples (Assuncao et al., 2005). When analysing the occurrence of an event across space the measure of interest is the rate, i.e. occurrences / population. In small areas the population denominator is usually small, meaning an increase of just one or two occurrences can dramatically influence the rate (Devine et al., 1994). This challenge is known as the ‘small number problem’, and the danger is that the underlying spatial pattern of interest is obscured (Hampton et al., 2011). Figure 3.5 shows that some of the highest observed rates in the health data are in in fact from LSOAs with the smallest populations, showing the data needs to be handled with caution - this is where shrinkage estimation can help.

Two sources of variability can be considered in the distribution of a rate across space: the true - and unobservable - underlying rate, and the variation of the observed rate around this true rate caused by randomness or measurement error (Devine et al., 1994). There are many methods which aim to provide a more accurate estimation of this first rate, from simple spatial smoothing to Bayesian Monte Carlo Markov Chain estimators; of these, a local Empirical Bayes (EB) approach can be a fair balance between reliability, interpretability and computational ease (Yasaitis et al., 2015). An EB estimator estimates the variability of each area relative to a larger sample mean,

Figure 3.5: Health indicator rates and area populations

which then allows for a weighted combination of the observed rate with the mean of the larger sample, with the weight determined by the ratio between variances (Marshall, 1991). The result is that the weight given to an area with a small population is low, and in this way the estimator 'borrows strength' from larger regions to create a more stable estimate (Devine et al., 1994 ; Hampton et al., 2011).

A key question in using a local EB approach is how to define an appropriate larger region, or neighbourhood, for each of the small areas. With the small area as LSOA, it is tempting to simply use the level above in the UK's administrative geography. This is the approach in the Index of Multiple Deprivation (IMD), where the Local Authority District (LAD) is used for shrinkage estimation (Smith et al., 2015). While this may be a sensible approach for a country-wide study, a single LAD might contain enormous variation in a diverse and dense urban region such as London. For this reason, as well as a neighbourhood definition using the LAD, two other approaches were tested: one using queen contiguous areas and another using a k-nearest neighbours approach with k = 10. Figure 3.6 shows examples of these definitions for an LSOA in Westminster. In R the package spdep's EBlocal method - which is an implementation of Marshall's (1991) local Empirical Bayes method – was used to carry out shrinkage estimation using these three different neighbourhood definitions (Bivand et al., 2019).

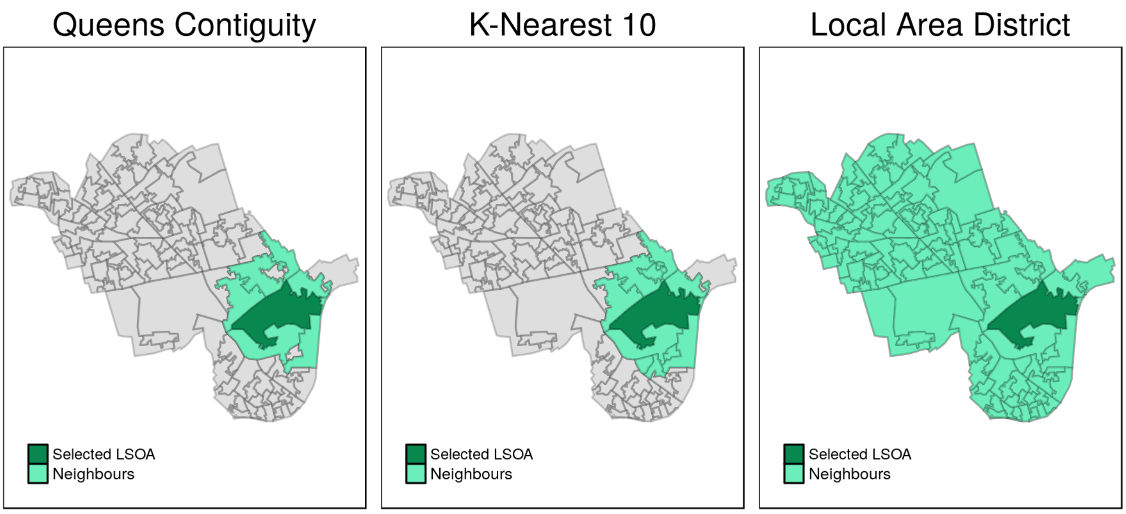

Figure 3.6: Example neighbourhood definitions for Westminster LSOA

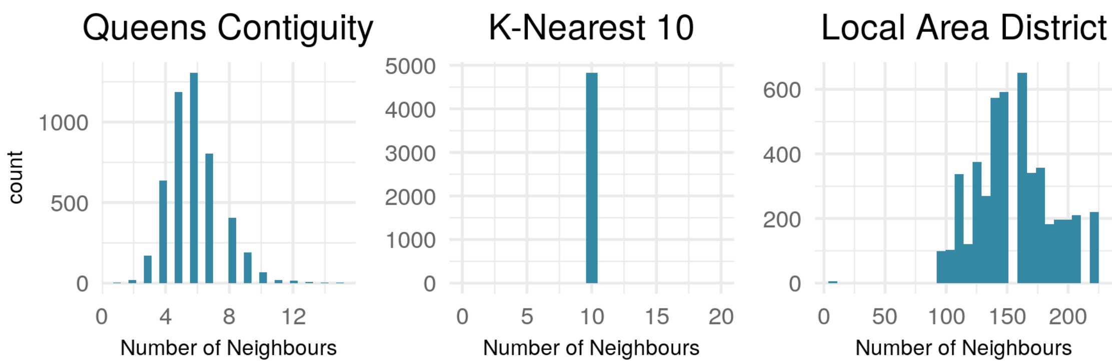

Figure 3.7: Distribution of neighbours per LSOA for different neighbourhood definitions

Figure 3.7 shows that the Queen's and LAD approach can result in some LSOAs being treated very differently in terms of their neighbours, City of London being an extreme example in the LAD approach. It is clear to see from figure 3.8 the different effects of each approach on the post-processing distributions; the LAD approach possibly results in 'overshrinkage' where the variance is dramatically reduced (Devine et al., 1994). This is also clear in the processing results for the dementia indicator in Westminster (figure 3.9), where it appears the LAD definition has reduced the rate to almost a single value. After testing different values for k, the KN-10 approach was selected as the best configuration as it pulls in the most extreme outliers while not shrinking the distribution too dramatically.

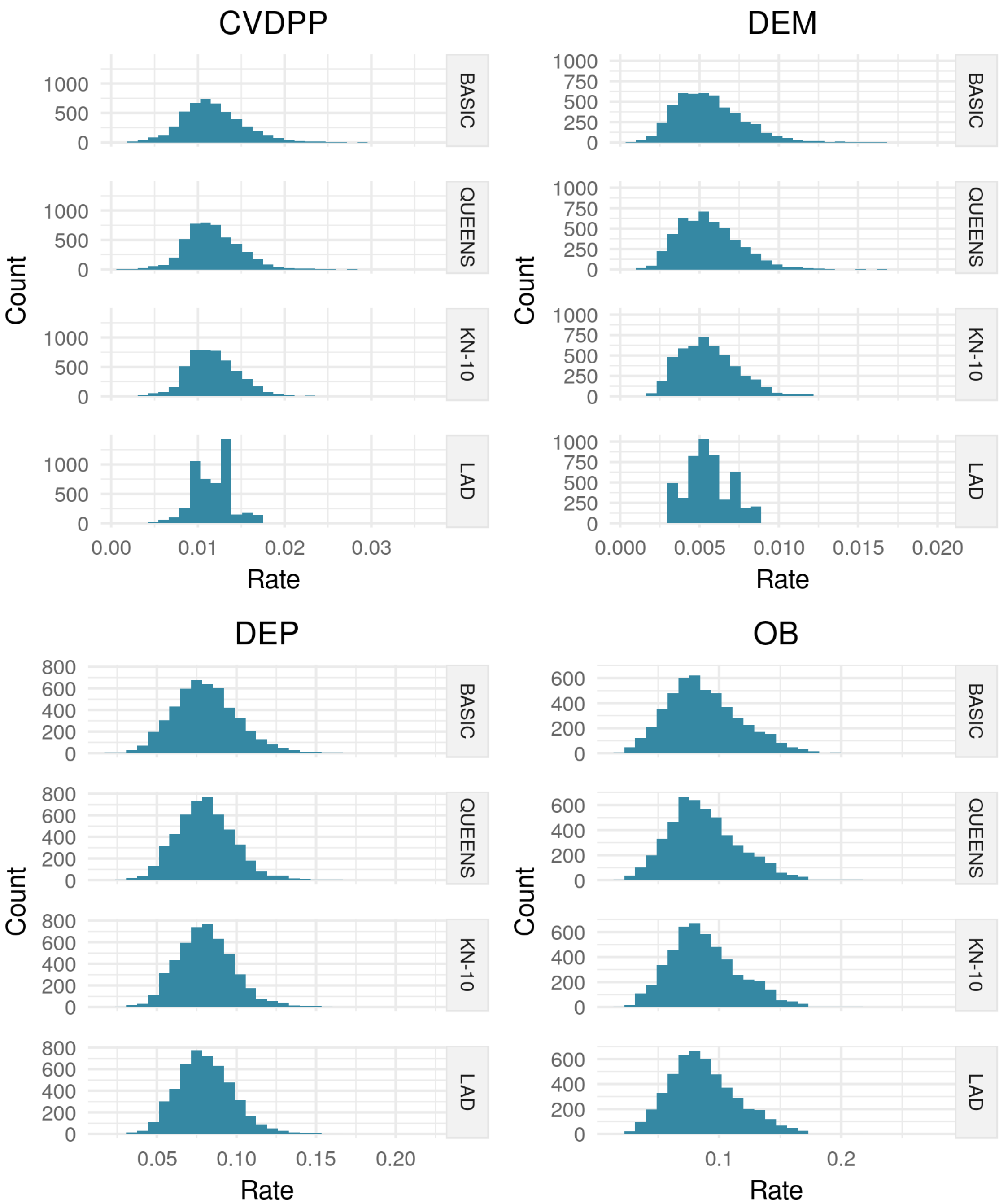

Figure 3.8: Indicator distributions after shrinkage estimation using different neighbourhood definitions

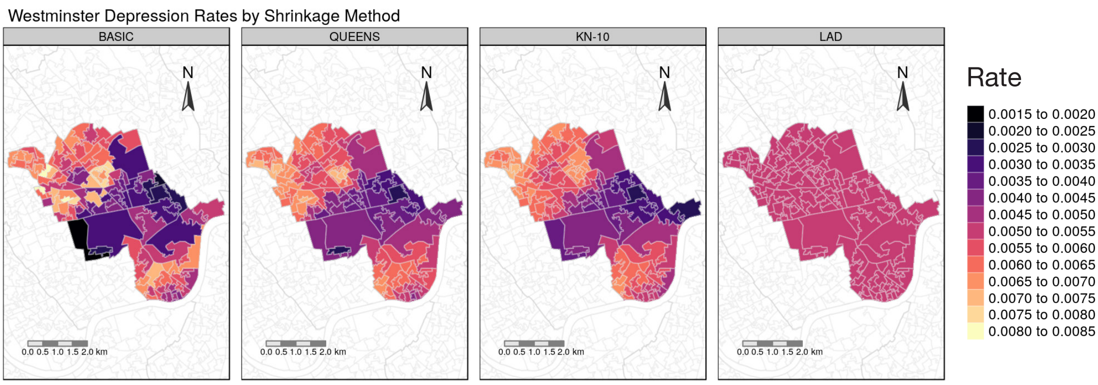

Figure 3.9: Dementia indicator scores in Westminster after shrinkage estimation using different neighbourhood definitions

3.6 Normalisation

Normalisation is a critical step in building a CI as it renders the data for different indicators comparable (EU Science Hub, 2015). The methods used here to normalise also integrate aspects of compensability, so it makes sense to discuss that issue here rather than section 3.7, which it falls under in the 10-step guide.

The issue of compensability forces the developer of a CI to consider whether the indicators being used are substitutable or non-substitutable, i.e. “if a deficit in one component may be compensated by a surplus in another” (Mazziotta and Pareto, 2013: p71). Depending on the decision made, a compensatory or non-compensatory aggregation method is chosen. This is a key conceptual point in the model, and determines whether the resulting index will balance out high and low demand across multiple indicators, or whether high demand in any indicator is not compensated by low in another. As the aim of this model is to highlight areas of demand and investigate opportunities for multifunctionality the latter approach is considered more appropriate and a method to achieve this borrowed from the IMD. The designers of the IMD decide upon a weighted cumulative model which greatly minimises ‘cancellation effects’, i.e. it is non-compensatory (Smith et al., 2015). The method involves ranking areas for each indicator score and then transforming the ranks to an exponential distribution:

For any LSOA, find its rank on the domain R, scaled to the range $[0, 1]$. $R = 1/N$ for the lowest indicator score and $R = N/N$ (i.e. $R = 1$) for the highest indicator score, where N = the number of LSOAs in Greater London. The transformed indicator score X is given by:

$$X = -23\ln(1 - R(1 - \exp^{-100/23}))$$

The result is a distribution of the range $[0, 100]$ with the highest ranked decile spread between 50 and 100 – thus focussing the distribution on the highest scoring LSOAs. The formulation allows precise control over cancellation: a scaling constant of 23 is used as it produces a cancellation effect of roughly 10% (ibid.). For example, in an extreme case if an LSOA is ranked lowest for temperature and highest for pollution, should these two indicators be combined with equal weighting it would end up being ranked in the 90th percentile in the combined score, compared to the 50th percentile had the untransformed ranks or a normal distribution been used (ibid.). Figure 3.10 shows the change in distribution for the dementia indicator before and after transformation, with the highest ranked decile highlighted in each chart.

The transformation has the simultaneous effect of standardising all data to the same range, making the indicator data suitable to aggregate. It is important to consider at which points in the model the exponential transformation should be executed. As the aim is for the model to be cumulative across indicators and domains it will be carried out once a final score is derived for each

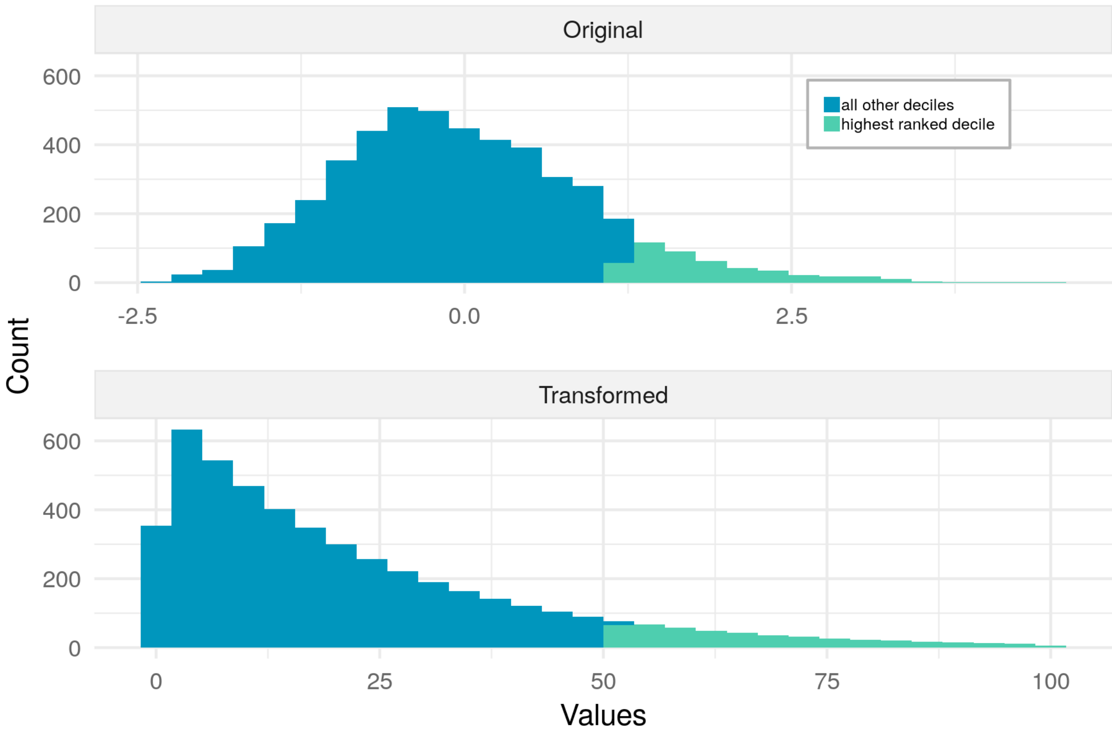

Figure 3.10: Effect of exponential transformation on dementia indicator

of these levels. Data at the sub-indicator level will not be transformed before being aggregated to an indicator score. By shifting values to the low end of the distribution, this transformation will have the added effect of minimising the impact of the 0 values in the noise and flood data.

3.7 Weighting

A large part of the challenge with aggregation has been dealt with by the exponential transformation: a simple arithmetic mean of the transformed score will combine indicator data in a satisfactory way (as described above). However, using the arithmetic mean assumes equal weights between indicators, but this is an entirely subjective decision as the weight essentially dictates the 'explicit importance' given to each indicator (Greco et al., 2019). Weighting can therefore significantly affect the results of a CI and as a result the question of most appropriate method is a contentious one in the literature. The one point of agreement is that no weighting system is beyond criticism (ibid.). Factor analysis methods can help investigate appropriate weights to use as well as how to combine data sources to ensure that each indicator is measuring a clearly separate aspect of the objective measure (Smith et al., 2015). Here PCA is used in cases where there is more than one dataset for a ecosystem service, i.e. all health data in the cultural category and the pollution data in the regulating category.

3.7.1 Principal Component Analysis

A factor can be understood as “a dimension or construct which is a condensed statement of the relationship between a set of variables” (Kline, 1993: p5). Principal Component Analysis (PCA) is a commonly used method which transforms variables using the standardised form of the co-variance matrix: the correlation matrix. What results is an uncorrelated set of new variables (components), in the order of the amount of variance they account for in the original data (Nardo et al., 2005). In PCA the components are real factors – that is, actual combinations of variables – and the factor loadings show the correlation of these components with the variables (Kline, 1993). Once PCA has been conducted a key final step is factor rotation. There is essentially an infinite combination of factors and loadings which all describe the same thing; rotation is the term used to describe a process which rotates the axes of the Euclidean representation of the factors and their loadings while maximising a particular objective (idbid.). The varimax process aims to achieve a so-called ‘simple-structure’ to final factors by minimising the number of original variables which have high loadings on the same factor (Nardo et al., 2008). It is these final weights which are a useful tool in the construction of a CI, as they can help decide if and how the variables should be grouped, and with what weights.

3.7.2 Health Results

Variable	*Factor Loadings*			*Squared Factor Loadings*		
	Factor 1	*Factor 2*	*Factor 3*	*Factor 1*	*Factor 2*	*Factor 3*
DEMENTIA	-0.085	**0.301**	0.123	0.007	0.091	0.015
DEPRESSION	0.098	0.142	**0.301**	0.010	0.020	0.090
CARDIOVASCULAR	**0.651**	-0.017	0.209	0.424	0.000	0.044
OBESITY	**0.651**	-0.194	0.033	0.424	0.037	0.001
Eigenvalue	0.864	0.149	0.15			
Variance proportion	74.3%	12.8%	12.9%			

Table 3.5: PCA results for health data - after varimax rotation

The eigenvalues – calculated by summing the squares of the factor loadings for each factor - describe the total amount of variance that can be explained by each factor (Kline, 1993). When selecting factors to keep for dimension reduction it is common practice to keep those larger than 1 and which cumulatively represent over 80% of explained variance (Nicoletti et al., 2000). The relatively low scores here suggest that there may not be a strong underlying factor of these variables. This is perhaps not a surprise given the low levels of correlation between the variables observed in the multivariate analysis, as correlation between the variables is generally a pre-requisite to successful PCA (ibid.). As a result of this it may be best to keep some of these variables as separate indicators. Here, the factor loadings can be used as a guide: the Cardiovascular and Obesity variables remaining grouped on the same factor both before and after rotation suggests they could be combined into a physical health indicator. Combining these two variables with equal weights – due to the equal loading on the factor – would result in three final indicators to each represent a different area of demand in the cultural category.

3.7.3 Pollution Results

Variable	*Factor Loadings*			*Squared Factor Loadings*		
	Factor 1	*Factor 2*	*Factor 3*	*Factor 1*	*Factor 2*	*Factor 3*
NO2	**0.990**	-0.120	0.028	0.980	0.001	0.000
NOX	**0.975**	-0.211	-0.010	0.950	0.000	0.000
PM10	**0.990**	0.112	-0.046	0.980	0.002	0.003
PM25	**0.969**	0.221	0.028	0.939	0.001	0.000
Eigenvalue	3.849	0.121	0.004			
Variance proportion	96.9%	3.0%	0.1%			

Table 3.6: PCA results for air pollution data - before varimax rotation

The PCA results for air pollution data in table 3.6 clearly show a strong relationship between all variables. The rotation step is useful if variables load across multiple factors and the aim is to make these distributions as clear as possible. However, the results here are clearly very different to the health data results, with the first factor having a very high eigenvalue, high loadings across every variable and representative of 96.9% of the total variation. With such strong loading on one factor it seems sensible to instead use this as the final pollution indicator rather than use the factor loadings to manually aggregate them. Using PCA for this technique means confidence in precisely how much of the variance in the original data is retained while the dimensions are reduced (Nardo et al., 2008).

3.7.4 Final Weighting Approach

The final weights decided above can be used to bring all data to indicator level, which can then be exponentially transformed. This gives the final set of indicator scores, which then need to be aggregated to domain-level scores, and those in turn exponentially transformed and aggregated to the final CI. Because each of the indicators have been selected to represent specific parts of the urban GI ES framework outlined in section 2, and because a key aim of the final index is to highlight opportunities for multi-functional GI interventions, the final approach will treat all indicators with equal weights. This will ensure the final index score is a holistic view of opportunities for GI intervention, rather than skewed towards a particular ES or service category. Because of the importance of weighting and the impact it can have, indicator weights are the particular focus of the sensitivity analysis conducted on the final model. Figure 3.11 outlines all the steps in this final model.

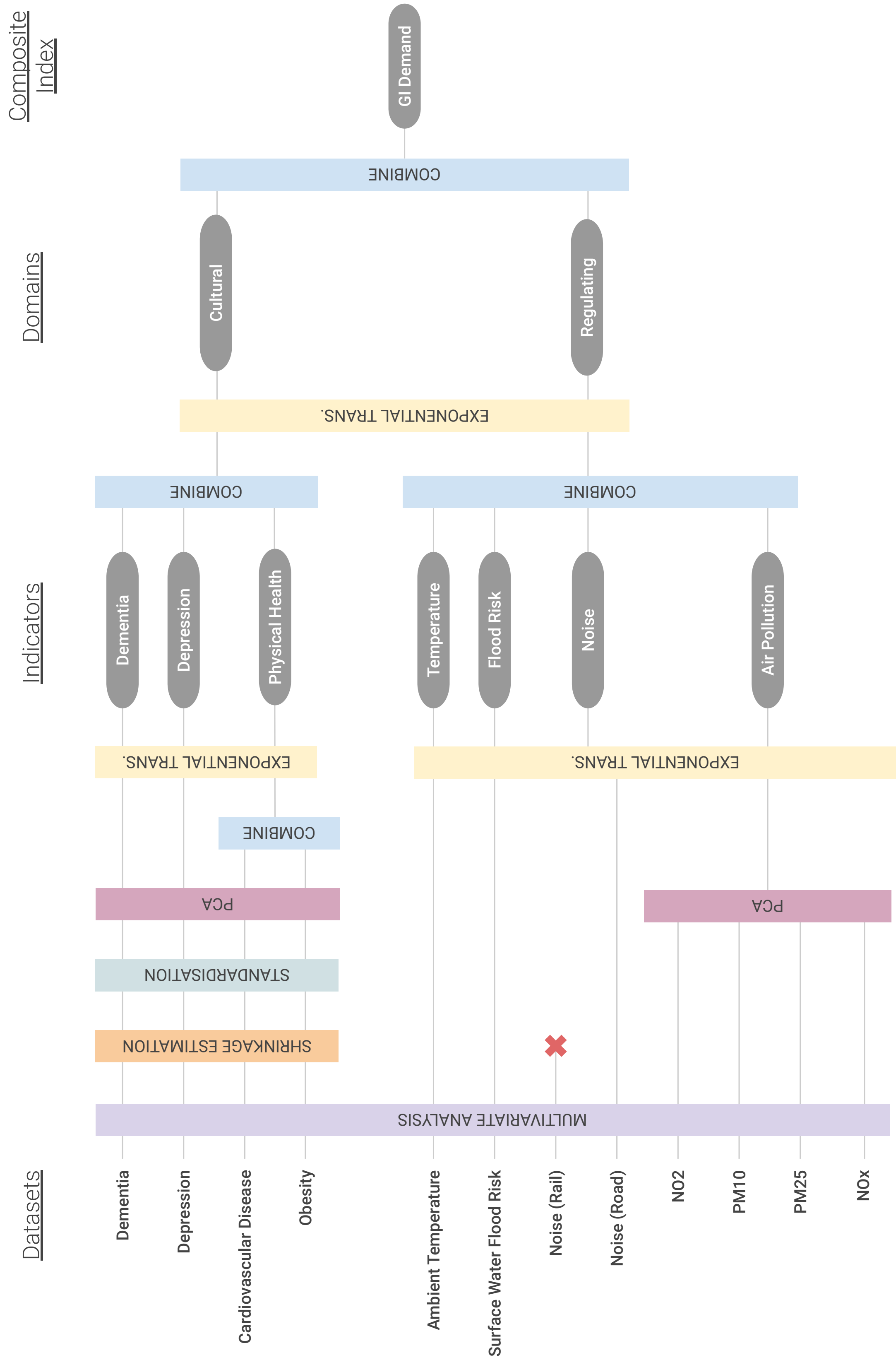

Figure 3.11: Process flow of all model steps

3.8 Uncertainty and Sensitivity Analysis

From each of the steps involved up to this stage it is clear there are many distinct points which might be considered a source of uncertainty in the final output. It is thus critical to test an index for robustness and gain some understanding of how sensitive it is to changes in the construction process (Greco et al., 2019). Typical methods to do so are Uncertainty Analysis (UA) and Sensitivity Analysis (SA). Nardo et al. (2008: p34) describe the two processes as follows:

> Uncertainty Analysis: "focuses on how uncertainty in the input factors propagates through the structure of the composite indicator and affects the composite indicator values"
>
> Sensitivity Analysis: "assesses the contribution of the individual source of uncertainty to the output variance"

3.8.1 Uncertainty Analysis

Approach

If each source of uncertainty in the model is understood as an input factor, the UA procedure is essentially a series of simulations carried out to understand the changes in a final measurement (Saisana et al., 2005). In their study, Saisana et al. run a Monte Carlo simulation of a model by assigning a Probability Distribution Function to each input factor Xi, and sampling from these factors for many model runs. A simplified approach is taken here, by selecting input factors considered important to the model's output and evaluating the resulting LSOA index ranking for every possible combination of each input factor $Xi, i = 1, 2, \ldots k$. Table 3.7 describes the factors chosen.

Input Factor	*Name*	*Description*	*Range*	*Values*
X1	Air pollution data	Varies the calculation of the air pollution indicator.	[1,3]	1 = Only NO2 used 2 = All data sources equally combined 3 = Principal component used
X2	Shrinkage neighbourhood	Varies the neighbourhood definition used in shrinkage estimation	[1,4]	1 = No shrinkage estimation 2 = Queen's contiguity 3 = K-Nearest 10 neighbours 4 = Local Area District
X3	Cultural indicators	Tests exclusion of each indicator in the Cultural domain	[1,4]	1 = All cultural domain indicators included 2 = Physical health excluded 3 = Dementia excluded 4 = Depression excluded
X4	Regulating Indicators	Tests exclusion of each indicator in the Regulating domain	[1,5]	1 = All regulating domain indicators included 2 = Flood risk excluded 3 = Ambient temperature excluded 4 = Noise excluded 5 = Air pollution excluded

Table 3.7: Input factors used in uncertainty analysis

The model is run for N combinations of input factors, $Xl, l = 1, 2, \ldots N$, where $X = X_1^l, X_2^l, ..., X_k^l$. In this case, by testing each possible factor combination, N = 240. The final index score and rank is calculated per LSOA for the N input factor combinations. As well as considering the distribution and variance of these rankings, Nardo et al. (2008) suggest measuring the final ranking from each test against a baseline version of the model and calculating a statistic which represents the overall variance from that test. Let CI be the index value for LSOA $g, g = 1, 2, \ldots, M$. For any run of the model we can record the rank of the resulting CI score for each LSOA, i.e. $Rank(CIg)$, and then explore the mean shift in LSOA rankings. This statistic is the average of the absolute differences in LSOAs' ranks in relation to a reference ranking over the M LSOAs:

$$\bar{R}_s = \frac{1}{M} \sum_{g=1}^{M} \mid Rank_{ref}(CI_g) - Rank(CI_g) \mid .$$

The model used as a reference will be the final configuration shown in figure 3.11. For clarity, in relation to the input factors, in the reference model $X1$ = 3, $X2$ = 3, $X3$ = 1, $X4$ = 1 and indicators and domains are combined with equal weighting.

Results

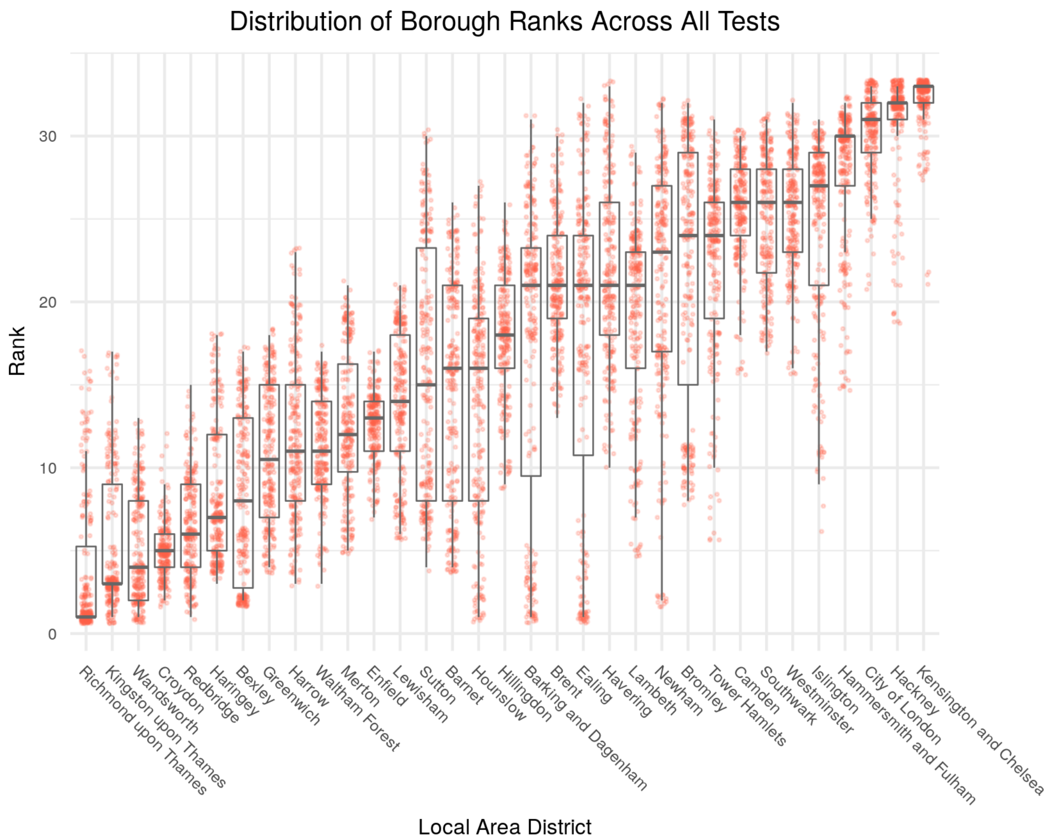

Figure 3.12: LAD ranking for all uncertainty tests (1 = lowest GI demand, 33 = highest GI demand)

To simplify interpretation of charts, results are aggregated by ranking LADs based on their mean LSOA score for each test. In figure 3.12, for each LAD every point represents their final ranking from one of the 240 tests. It is clear that the general variance in rank is fairly high across all LADs. This is possibly because of the relatively small number of simulations carried out here, and that at least two of the input factors – $X3$ and $X4$ – are likely to have a large impact on the model by removing one indicator entirely from the final seven used. As might be expected, variance is generally higher amongst the LADs in the middle range of rankings, suggesting it is harder to make clear distinctions in this range compared to the more extreme ends of the scale where there is less variance. This is a possible effect of the exponential transformation used: low-medium indicator scores are essentially restricted to a small range, so rankings between areas with scores in that range might be more likely to change between tests. There also appears to be an almost clustering-like effect in rank values for some LADs, possibly caused by a certain factor having a particularly strong effect on end rankings. This is possible to visualise by taking Nardo et al.'s suggested scatter plots a step further and colouring points based on the values of one of the input factors.

For each LAD, every point represents a test with a different configuration of all four input factors. Using Ealing as an example in Figure 3.12: the median rank across all tests is just over 20, but there are a group of test results where the ranks are all below 10. Figure 3.13 shows for most of these results factor X3 had a value of 2, which (referring back to table 3.7) means that the physical health indicator was excluded. However, when the points are coloured by factor X2 values (as in figure 3.14), each shrinkage method is fairly evenly distributed throughout the ranking results. This suggests that the value of some factors have a far stronger influence on the final ranking result than others. While the CI literature contains some sophisticated measures of sensitivity, there seem to be relatively few straightforward measures to compare levels of uncertainty. Here a method of comparison is devised using the standard deviation of ranks between factor values, calculated using the following steps:

- Beginning with the first input factor, $X1$, for every possible value the mean rank for each LSOA across every test result is calculated. So considering all 240 test configurations, there are 80 different test results for each value of factor $X1$, for each LSOA their mean rank in those 80 tests is calculated.
- For each LSOA the standard deviation in their mean rank across each input factor value is calculated.
- This is repeated for each remaining input factor.

The aim of this measure is to show, for every factor, the range of variance in mean rank as the value of that factor is changed in the testing.

Figure 3.15 clearly shows the variance caused by changes to the cultural input factor is far higher than others. This is possibly caused by the fact it consists of fewer variables, so removing one may have a larger impact. The variance from the shrinkage and air factors is much more

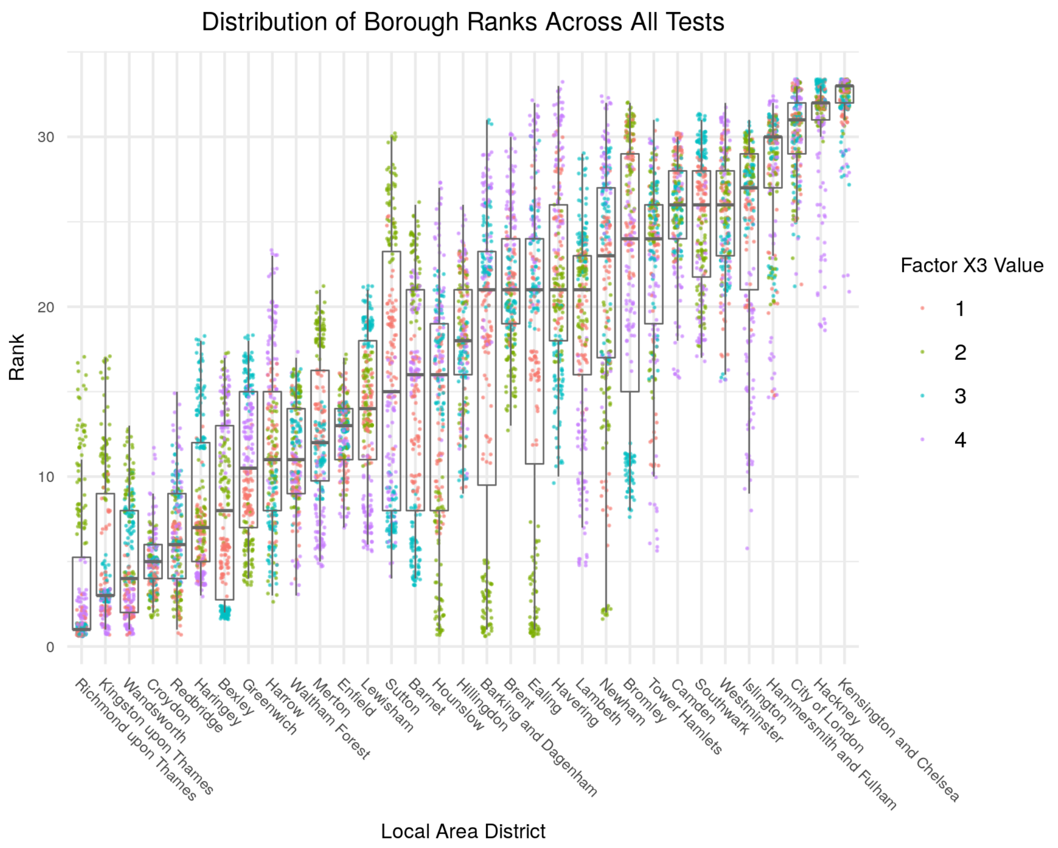

Figure 3.13: LAD ranking for all uncertainty tests - coloured by uncertainty factor $X3$

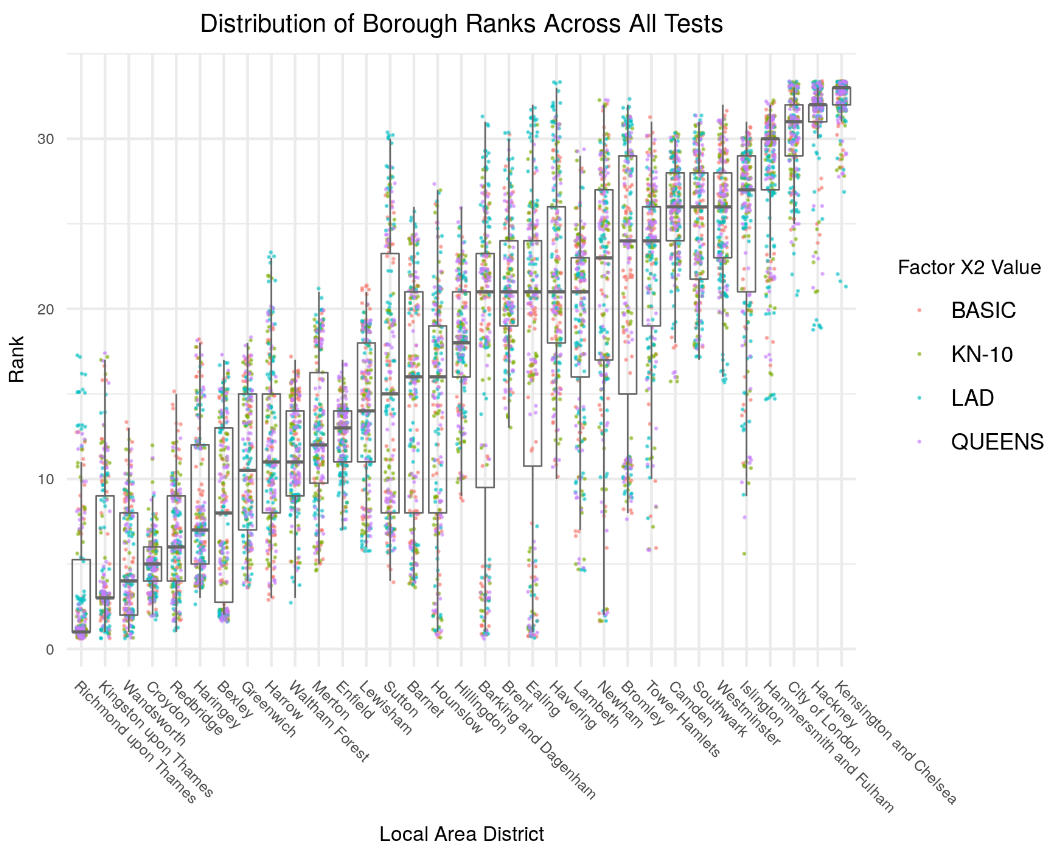

Figure 3.14: LAD ranking for all uncertainty tests - coloured by uncertainty factor $X2$

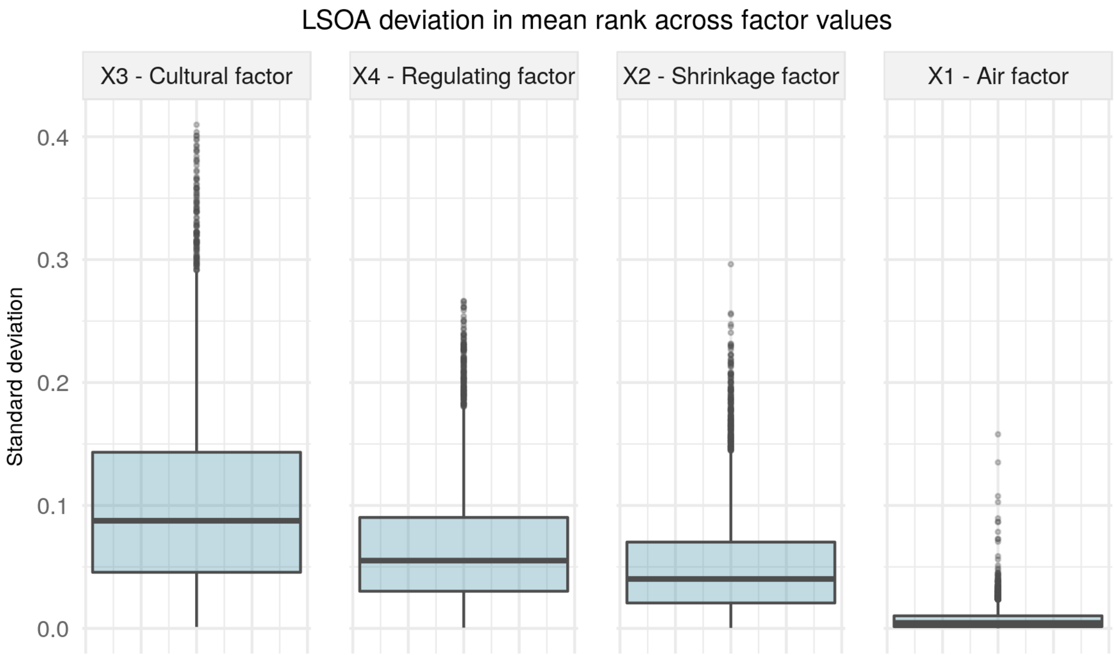

Figure 3.15: Distribution of the standard deviation in mean LSOA ranking across factor values

minimal, suggesting the model is more robust to changes at these earlier stages.

3.8.2 Sensitivity Analysis

Approach

The simulation approach used in the UA can be focussed in order to to judge the sensitivity of the model to changes in one very specific aspect. Here a method of sampling indicator weights using the Dirichlet distribution is devised in order to measure sensitivity to the weighting of each indicator in turn. The Dirichlet distribution can generate $Q = [Q_1, Q_2, ..., Q_k]$ a random probability mass function, that is $Q_i \geq 0$ for $i = 1, 2, ..., k$ and $\sum_{i=1}^{k} Q_i = 1$ (Frigyik et al., 2010). The values of Q can thus be used as weights for a set of variables V, where $V = [V_1, V_2, ..., V_k]$, in any aggregation step. The final score, M of this aggregated set of variables can be calculated as follows:

$$M = \sum_{i=1}^{k} Q_i V_i$$

An input vector of k number of α values to the dirichlet sampling process affects the distribution of values in Q in a predictable way, so by adjusting the values of α it is possible to produce a range of weights which gradually increase the explicit importance of a particular variable in any aggregation. A ternary diagram can help visualise a set of sample weights taken for three variables: each plot in figure 3.16 shows 250 samples generated using a different α value for Q_3, while the alpha for Q_1 and Q_2 is kept at 1. It is clear how as α for Q_3 increases, the distribution of samples shifts from random between all weights to concentrated at the high end of Q_3.

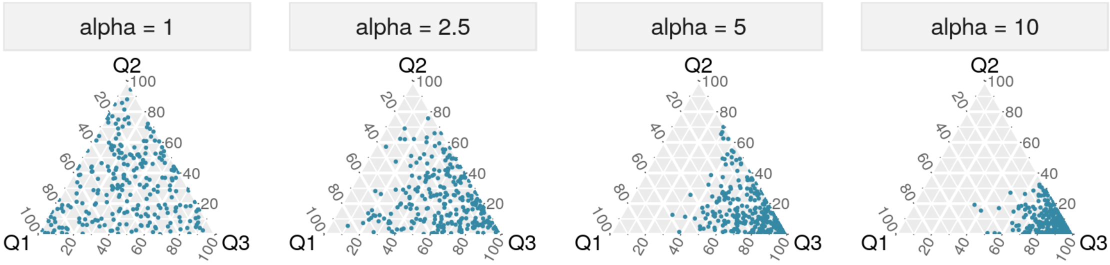

Figure 3.16: Ternary diagram of generated weight samples using different α values for Q_3

Using this process to produce sample weights, the aim here is to test each stage in the model where data is aggregated:

- 3 cultural indicators into the cultural domain score
- 4 regulating indicators into the regulating domain score
- cultural and regulating domain scores into the overall CI score

The following process is followed for each of these stages:

- Create 250 sample weights for the k input variables where the α for each variable is set to 1.
- Increment the α value for the first variable through a range $[1, 10]$ creating 250 samples for each increment and keeping the α values for all other variables at 1.
- Reset all α values to 1 and repeat the step above for each remaining variable.
- Run the model for every sample weight in turn.

Figure 3.17 shows for each aggregation stage, the mean of the 250 sample weights for each variable as the alpha is incremented in every individual variable test. By calculating the $\bar{R}_s$ statistic for model runs using every sample weight it is possible to analyse how each indicator in turn affects model uncertainty.

Results

Figure 3.18 shows there is quite a difference between some of the regulating indicators, with the flood indicator causing a steep rise in mean $\bar{R}_s$ as its weight is increased, while the air indicator stays almost flat in comparison. The noise and temperature indicators in contrast follow a very similar trajectory. The strong effect of the flood indicator is particularly interesting as this dataset

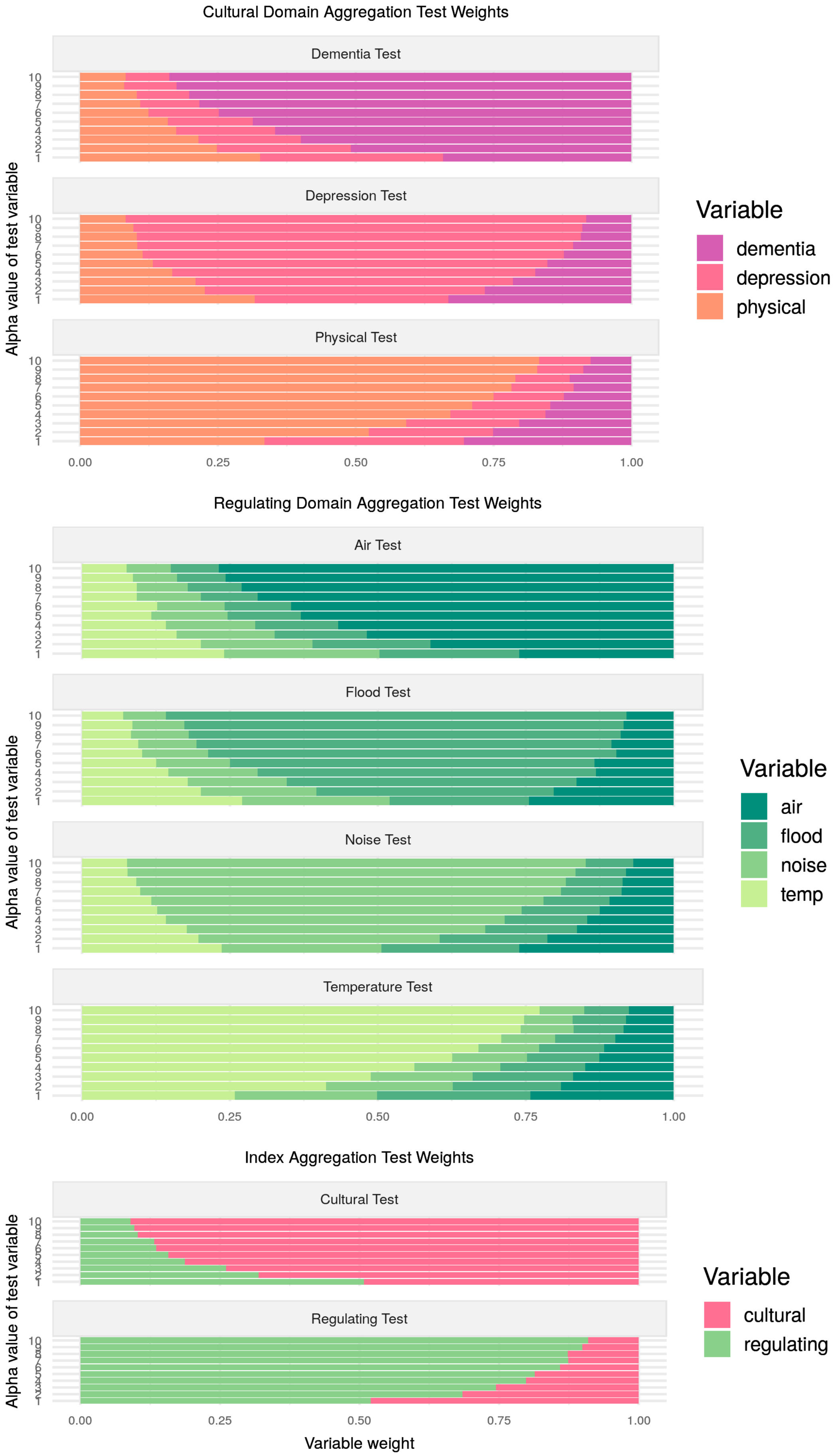

Figure 3.17: Mean variable weights used in each aggregation stage sensitivity test

Aggregation Stage	*No. Variables Aggregated*	*Total Sample Weights*
Cultural domain	3	7,500
Regulating domain	4	10,000
GI Demand	2	5,000

Table 3.8: Number of sample weights used in sensitivity tests for each aggregation stage

had a large number of 0 values, which – despite the exponential transformation – could possibly explain the greater variance from the indicator. In the cultural domain (figure 3.19) it is the depression indicator which has the least effect on mean $\bar{R}_s$, while the other two follow a similar and slightly steeper climb. While the highest $\bar{R}_s$ score for any indicator between each domain is not dissimilar - between 0.14-0.15 – it is interesting to note that when α = 1 the mean $\bar{R}_s$ for the cultural domain indicators is substantially higher at 0.9 compared to 0.7 for the regulating domain. This echoes the trend seen in the uncertainty analysis where deviation in average rank was greater for cultural indicator removal tests than the same test for regulating indicators. Here again is evidence that changes to the cultural domain generally have a greater effect on the model, which is possibly caused by the fact that it is constructed from fewer indicators.

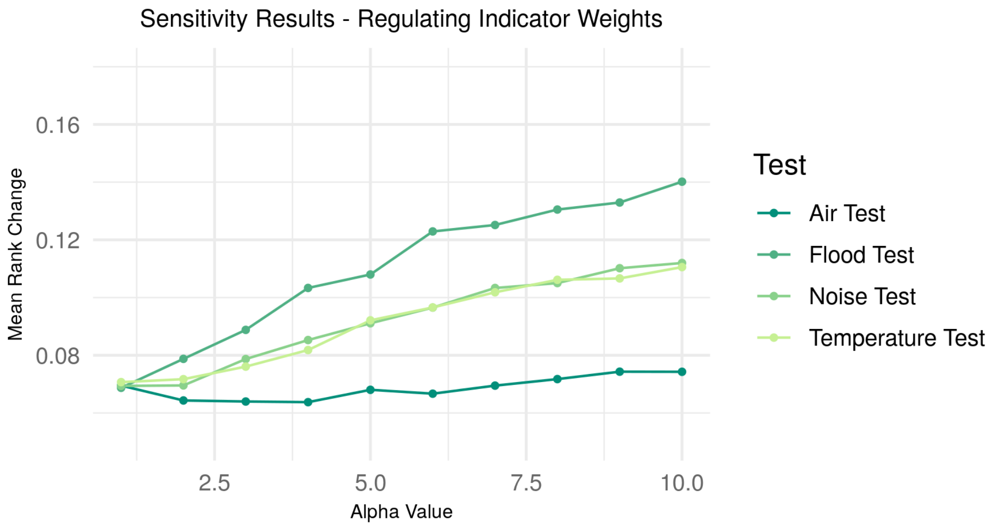

Figure 3.18: Regulating sensitivity - Mean $\bar{R}_s$ by alpha value for each indicator test

This trend can also be seen in the weight tests for the final index (figure 3.20), where the mean $\bar{R}_s$ is generally slightly higher for the cultural domain test. Furthermore, the scale of $\bar{R}_s$ is higher than that the cultural and regulating indicator tests, reinforcing the fact that changes at a later level of the model have a greater impact.

It is important to remember for each of these charts that the mean $\bar{R}_s$ is not a purely objective measure of model variance, rather it is simply a measure of the magnitude of change from the baseline model – that is, a model in which domain indicators and index domains were combined

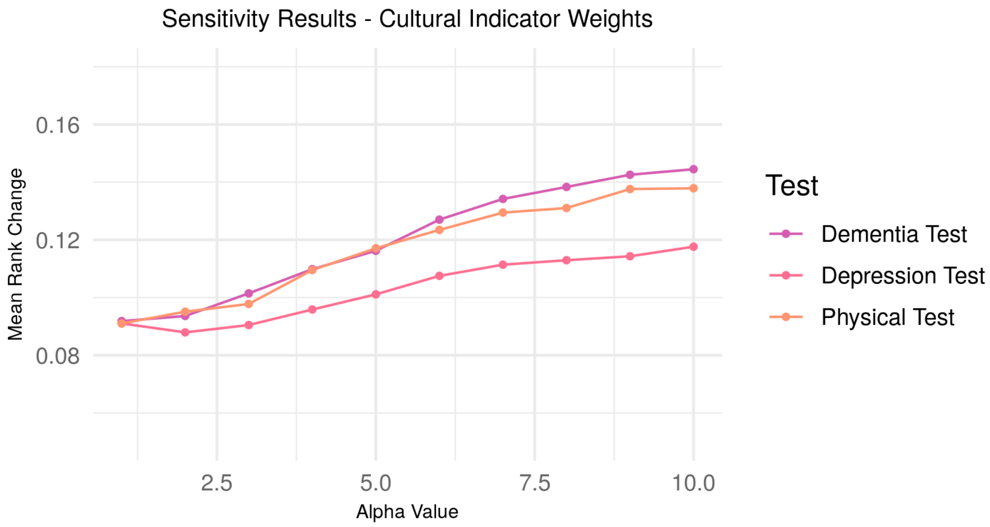

Figure 3.19: Cultural sensitivity - Mean $\bar{R}_s$ by alpha value for each indicator test

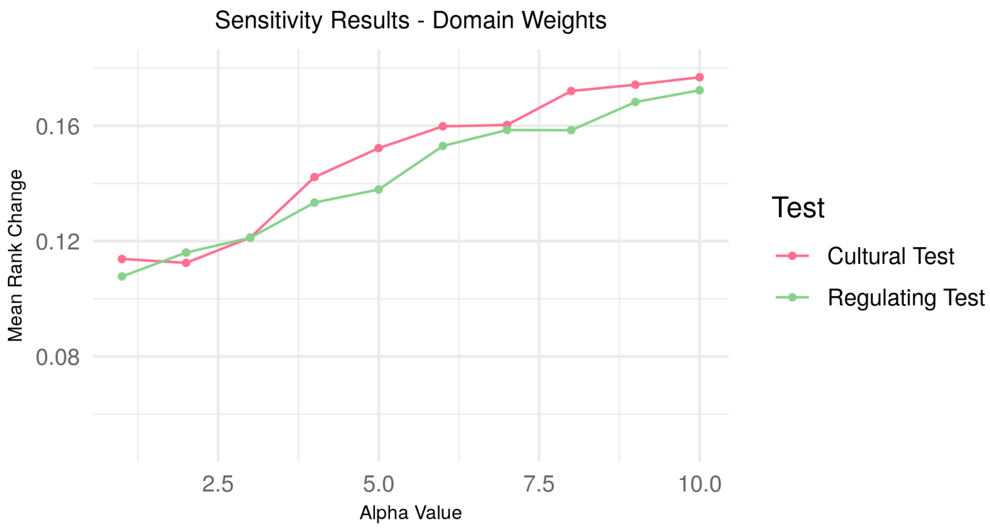

Figure 3.20: GI sensitivity - Mean $\bar{R}_s$ by alpha value for each indicator test

with equal weighting. Nonetheless it is still a useful measure to try and start to unpick each indicator's contribution to the model.

3.8.3 Discussion

The uncertainty analysis showed a generally high level of variance in the model, but this effect appeared to be largely caused by input factors $X3$ and $X4$, which tested the removal of individual indicators from the model. This impact is understandably quite dramatic as that indicator can in effect constitute between 12 and 17% of the model, depending on its source domain. The model shows far greater robustness from other important decisions like the shrinkage estimation method used for the health indicators and the construction of the air quality indicator from multiple datasets. Here it is possible to be more confident that the underlying trend is less impacted by small methodological changes.

The sensitivity analysis digs deeper into the challenge of weighting indicators and shows that not all indicators are equal: the flood risk indicator in particular contributed an over-proportional level of variance in test results. It is harder to judge from these results exactly what causes this, and also to judge what overall level of uncertainty is acceptable in the model. What does seem clear is that the number of variables used in a domain, and the number of domains used in the index makes an important contribution to model stability. Here the relatively low number of seven indicators is likely a key factor in the fairly high uncertainty scores in some tests, where rankings can change up to an average of 15%. This should be an important consideration in interpreting results and considering how appropriate the model might be to re-run in the future with updated data. This level does seem far from suggesting the model is totally unusable however; there is clearly an underlying pattern of demand which is largely being retained.

3.9 Results Analysis Methods

The following chapter consolidates the final three steps of the OECD guide. Best practice should involve validation against a comparable index. There is one similar model for London – the GLA's GI Focus Map (GLA, 2019a), released during the course of this research – but comparison is unfortunately not possible as the results data are not accessible. Instead, the results section focuses on developing a number of visualisation methods suggested in the CI literature - to aid interpretability charts are aggregated to LAD level.

Analysis is extended beyond the final CI score by decomposition, which involves calculating the contribution of each indicator to the final score for each area (Nardo et al., 2008). Additionally, cluster analysis using K-means is used to further explore this decomposed data for patterns of demand, using SSE and the silhouette score to judge the results (Tan et al., 2006).

Chapter 4

Results

4.1 GI Demand

Most clear from the overall GI demand map in figure 4.1 is the focus of high demand in the centre of London, though there is some variation and pockets of lower demand within. Beyond this the picture is much more varied, with areas of high demand scattered and some particularly focussed at the edges of the region. Some of the drivers of these patterns can be observed in the domain score maps, particularly the large area of high cultural demand in the south-east and the very clear central focus of the regulating domain scores. The very centre of London which has the highest regulating scores is also where some of the lowest cultural scores are focussed, though these low scores increase immediately to the north and in the south towards Lambeth and Southwark. Focussing on Westminster specifically (figure 4.3) it's clear that the borough is largely made up of high scoring areas, though there is still a scattering of lower scoring LSOAs. The domain maps show a substantial difference between cultural and regulating scores: the regulating indicators are generally high scoring and cover much of the borough, whereas the cultural indicator scores are lower with just two high scoring areas at the north-western and south-eastern boundaries.

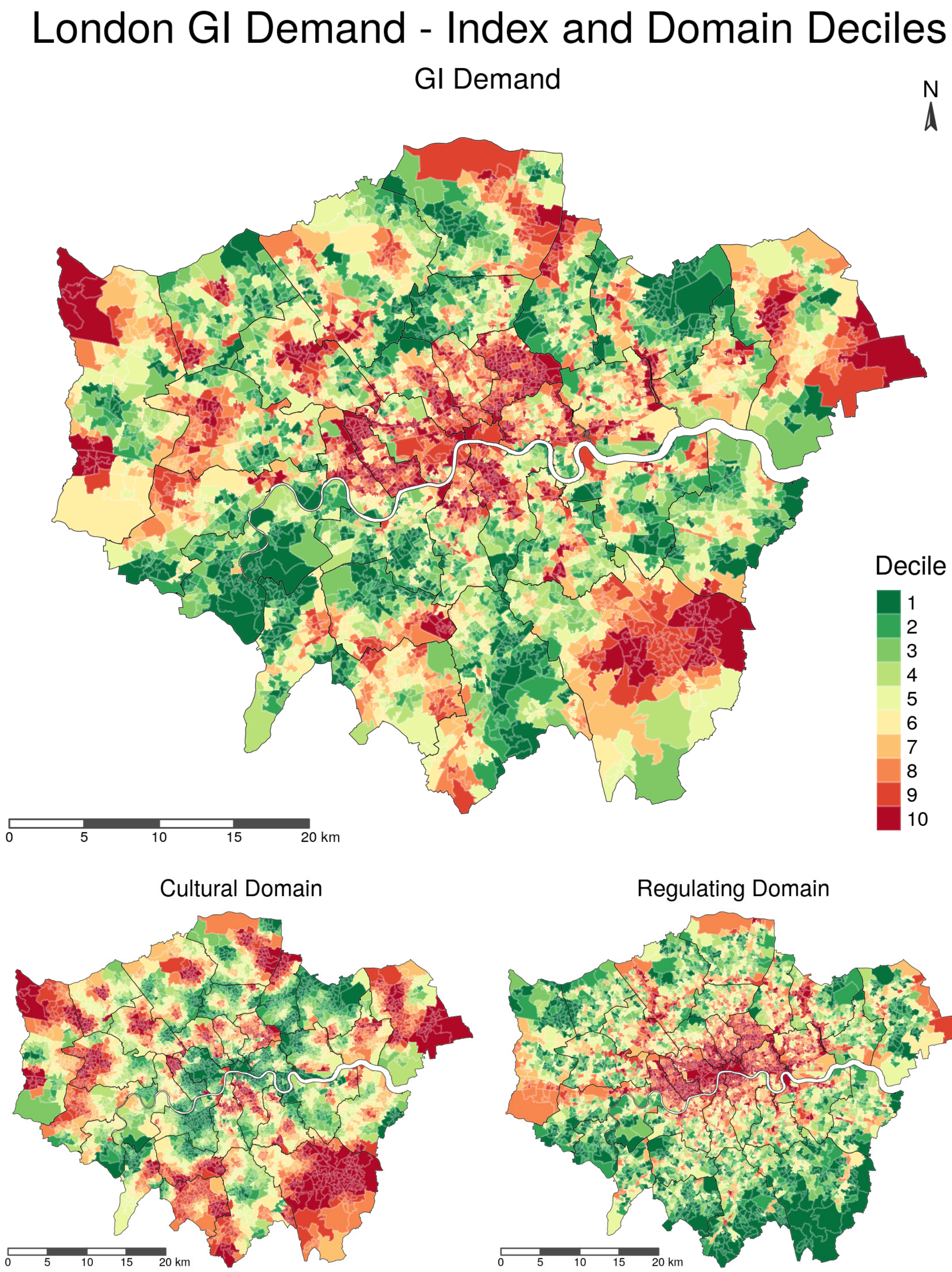

Figure 4.1: Deciles of final LSOA demand ranking (1 = lowest demand, 10 = highest demand)

Figure 4.2: London boroughs - for reference

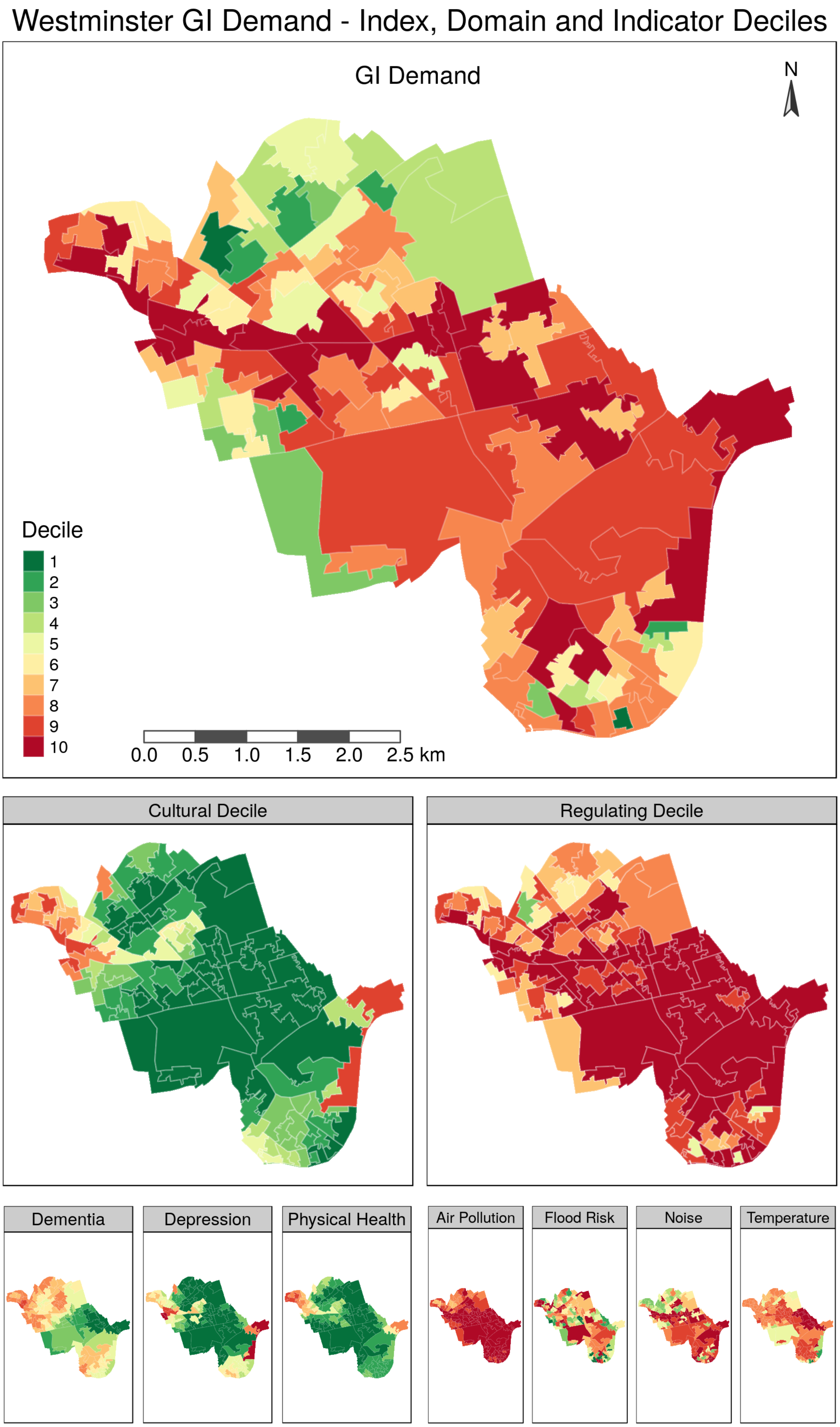

Figure 4.3: Demand deciles for Westminster

4.1.1 Decomposition

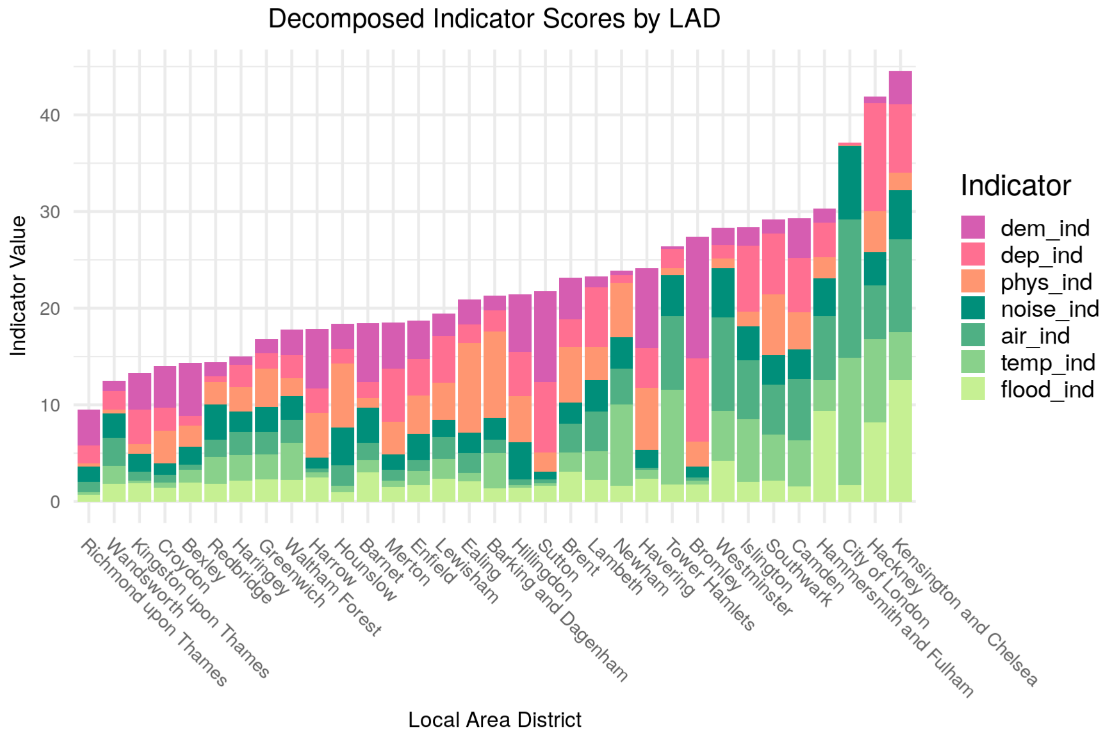

Figure 4.4: Mean GI demand index score by LAD, with decomposed indicator scores highlighted

The eight highest ranked LADs are all inner London boroughs. An interesting pattern here is the number of LADs with high scores in just one domain, particularly in the top part of the ranking: Bromley and Havering are dominated by the cultural domain indicators, while regulating indicators contribute significantly to scores for the rest, with Westminster and City of London as two extreme examples. It can be easier to see this breakdown for one area at a time – figure 4.5 shows Westminster's scores for each indicator as a proportion of the maximum score of any LAD. Here it's clear that all the cultural indicators have similarly low scores, while the regulating indicator scores dominate, especially air and noise. Figure 4.6 shows the same chart for the four highest and lowest overall scoring boroughs. Here the strong influence of the regulating domain on high scores is even clearer – all of the high scoring LADs have significant scores in this domain while generally cultural indicator scores stay lower, though Hackney and Kensington & Chelsea do have high depression scores.

This pattern of indicator influence can be unpacked in greater detail by visualising the decomposed indicator scores for any area as a proportion of that area's final index score (figure 4.7). This, rather than showing where demand is highest, shows for each indicator where it has its greatest effect on the final GI demand score. While the air indicator doesn't have the highest effect scores it does show a clear spatial pattern, with its highest effect focussed in the centre of London and diminishing outwards. The temperature indicator follows a similar pattern but with its focus just to the north-east of the centre in Tower Hamlets and Newham. The other indicators

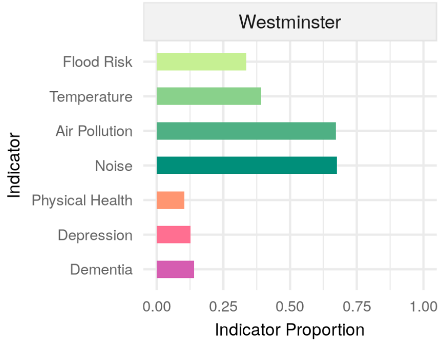

Figure 4.5: Decomposed indicator scores - Westminster

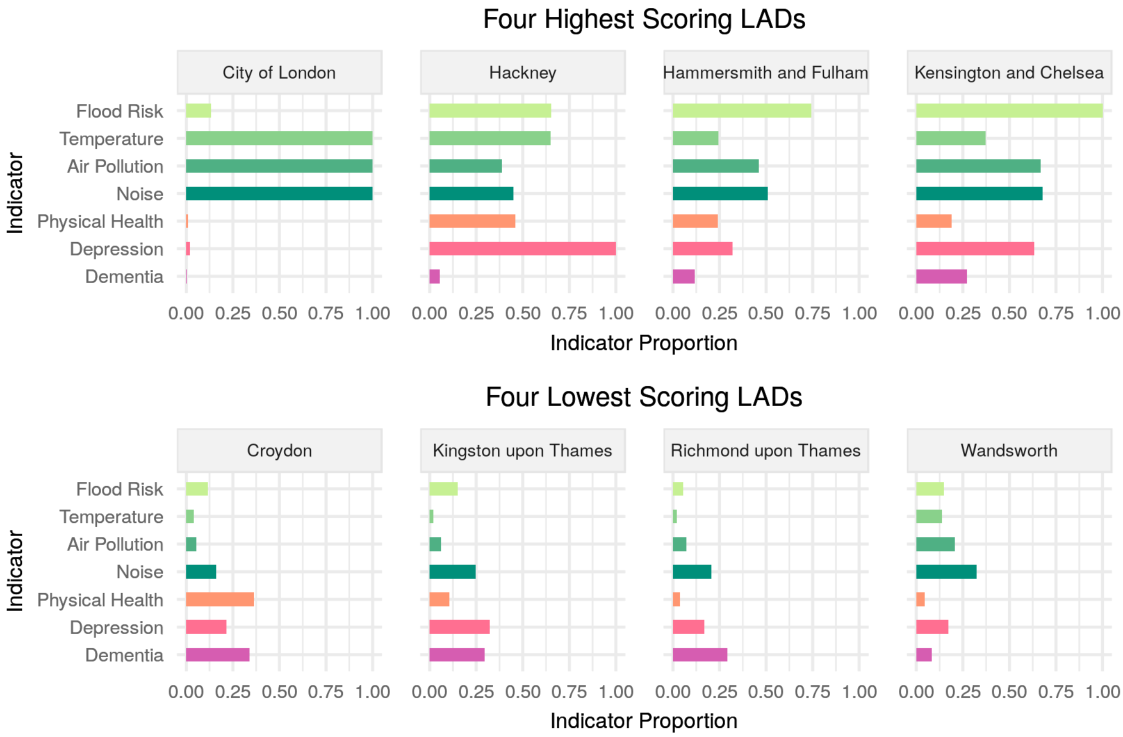

Figure 4.6: Decomposed indicator scores - Four highest and four lowest scoring LADs

show much more varied patterns, with small hotspots scattered around the city. The dementia indicator, however, stands out with by far the highest effect scores: these are largely focussed on the outer edges of London, particularly in parts of Richmond, Sutton and Croydon in the south where it almost completely dominates the score for some LSOAs. The same data displayed for just Westminster in figure 4.8 shows how the very high regulating indicator scores dominate the index results, and that proportionally the cultural indicators are contributing far less to the final score for most LSOAs.

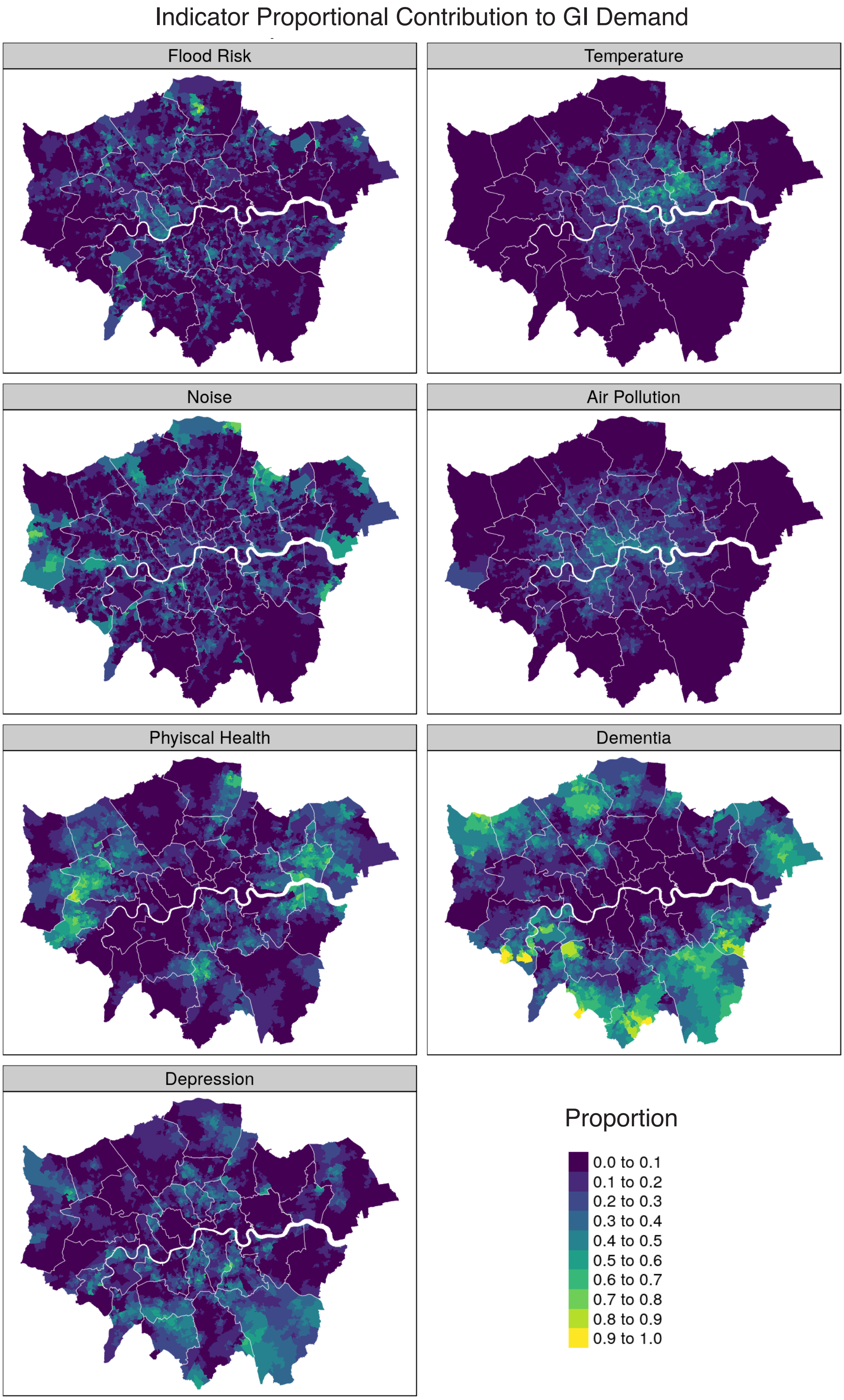

Figure 4.7: Indicator effect scores for London – proportional contribution to area index score per indicator

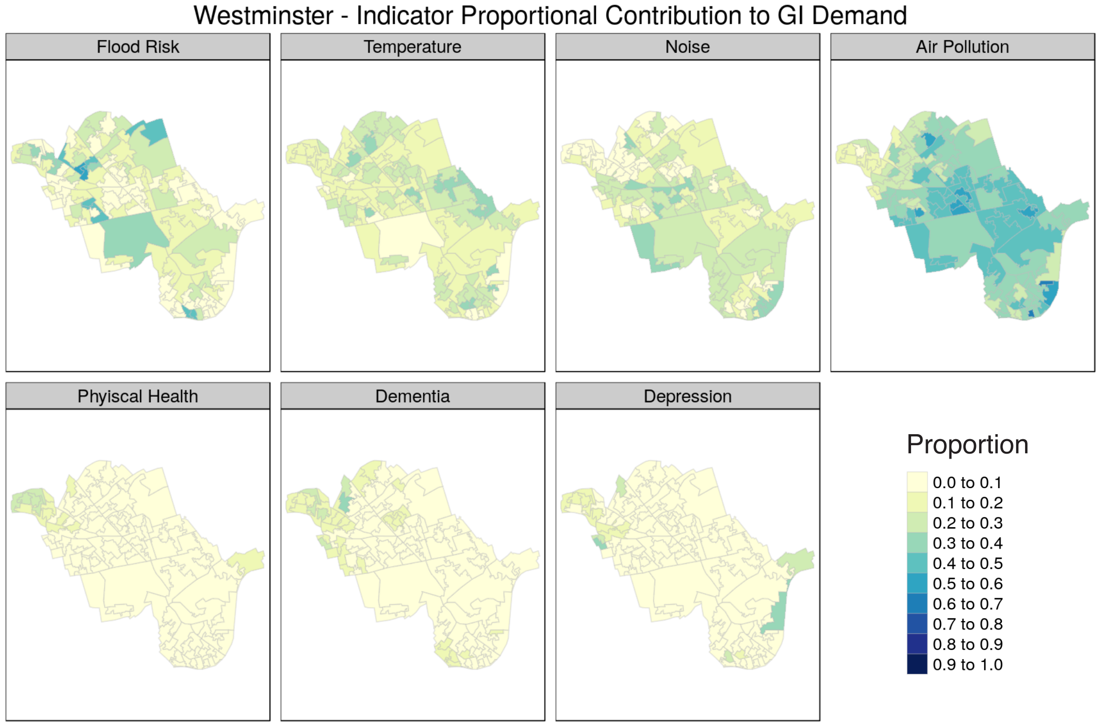

Figure 4.8: Indicator effect scores for Westminster – proportional contribution to area index score per indicator

4.1.2 Clustering

The initial results of k-means for the decomposed indicator scores showed that there aren't extremely clear clusters in the data. Slightly improved results were obtained by using an extra categorical variable which classes each LSOA as inner or outer London. This was included due to the different patterns in results for central and outer areas of London observed in the section above – by grouping the data this way the algorithm may be able to better distinguish clusters within inner and outer London separately. Figure 4.9 shows the silhouette score is relatively low throughout the range of k values tested, though beyond the initial peak at k=2 there is a second hump at k=10. Though it isn't matched by the desired clear elbow in the SSE measure the results are still worth investigating as while the algorithm may not be identifying clear clusters it is still partitioning the LSOAs into groups which are as different to one another as possible (Tan et al., 2006). The same visualisation used in figure 4.6 to compare relative indicator scores between boroughs can be used here to examine the different patterns of demand between the 10 clusters. Figures 4.10 and 4.11 combine these charts with a further chart showing the proportion of LSOAs in Inner and Outer London for each cluster, as well as mapping the cluster distribution across London.

Immediately clear from the cluster maps is the large areas of London taken up by clusters 2 and 5, which both represent LSOAs that have relatively low scores across all indicators. Inner London is largely made up of 3 and 8 – 3 describing extremely high scores in air, noise and

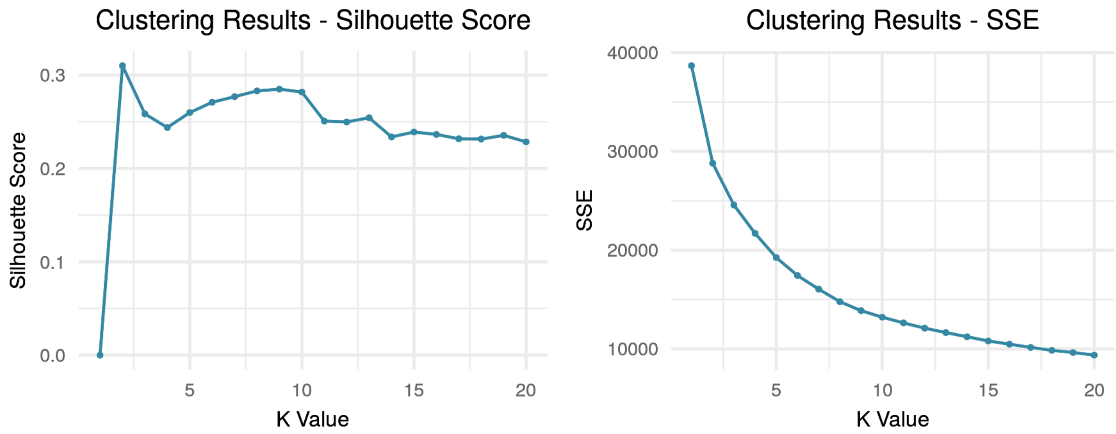

Figure 4.9: k-means test results – Silhouette score and SHS for $k = [1, 20]$

temperature and with a clear concentration around the very centre of London, and 8 slightly lower scores in those same indicators but with an emphasis on temperature while all others stay low. Cluster 4 appears to highlight the only areas in inner London which have high cultural domain scores, particularly in depression and physical health indicators. While many of the clusters cover a range of areas, 1 and 7 , along with 3, show perhaps the clearest spatial focus: 7 representing high flood scores to the west of Westminster, and 1 high noise scores in Newham and Waltham Forest. Interestingly, both of these areas correspond with where these two indicators have their greatest effect on final GI demand as shown in figure 4.7, showing that those trends are strong enough for those areas to be grouped in clusters. Apart from cluster 1 and 5, the rest of Outer London is grouped into pockets of 6, 9, or 10, representing higher scores in either the noise, physical or dementia indicators. Of these, 10 covers the largest area, with a particularly large area in the south-east covering Bromley – this is the same area which had very high cultural domain scores in figure 4.1 and high dementia effect in figure 4.7.

Aside from the spatial patterns, what is most clear is that few areas appear to have high demand in both the cultural and regulating domains: for all clusters apart from one, when demand is high in one domain it is low in the other. The only real exception is cluster 4, with scores in the depression, physical and temperature indicators. Furthermore, the cluster results reinforce the central focus of high regulating scores suggested in earlier results, with cluster 4 the only one of five predominantly inner London clusters with high cultural scores. Meanwhile, clusters 1 and 6 are the only predominantly outer London clusters with any high regulating scores.

All of the clusters in Westminster are focussed around the regulating domain. There are some areas with high flood scores (cluster 7), but the borough is predominantly cluster 3 – high air, temperature and noise scores – and cluster 8 – high air and temperature scores – with the remainder made up of 2, which represents low demand. This again shows how the index in inner London areas is dominated by the regulating domain, whose indicators generally have their highest scores in these areas.

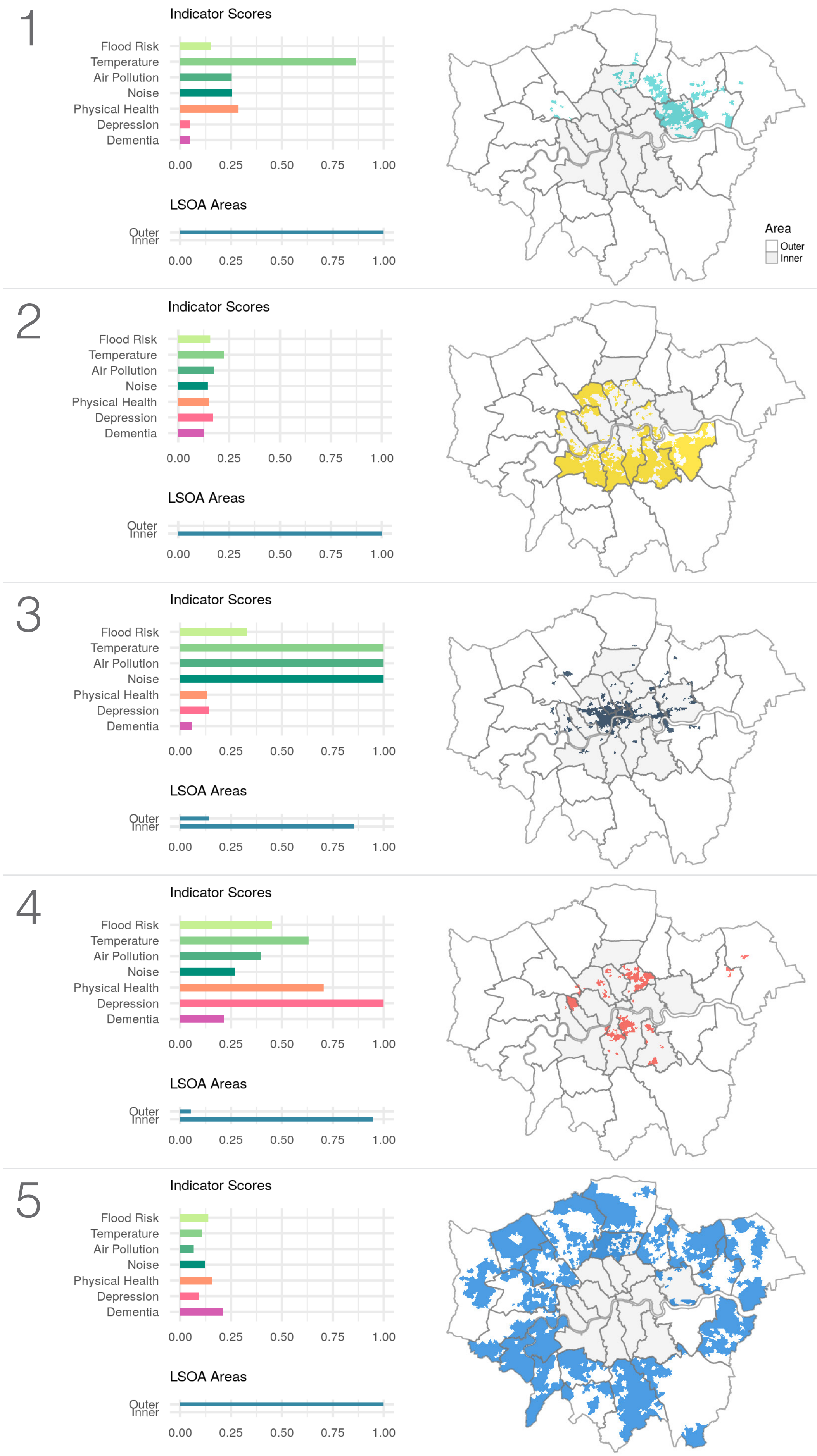

Figure 4.10: Cluster descriptions - clusters 1 to 5

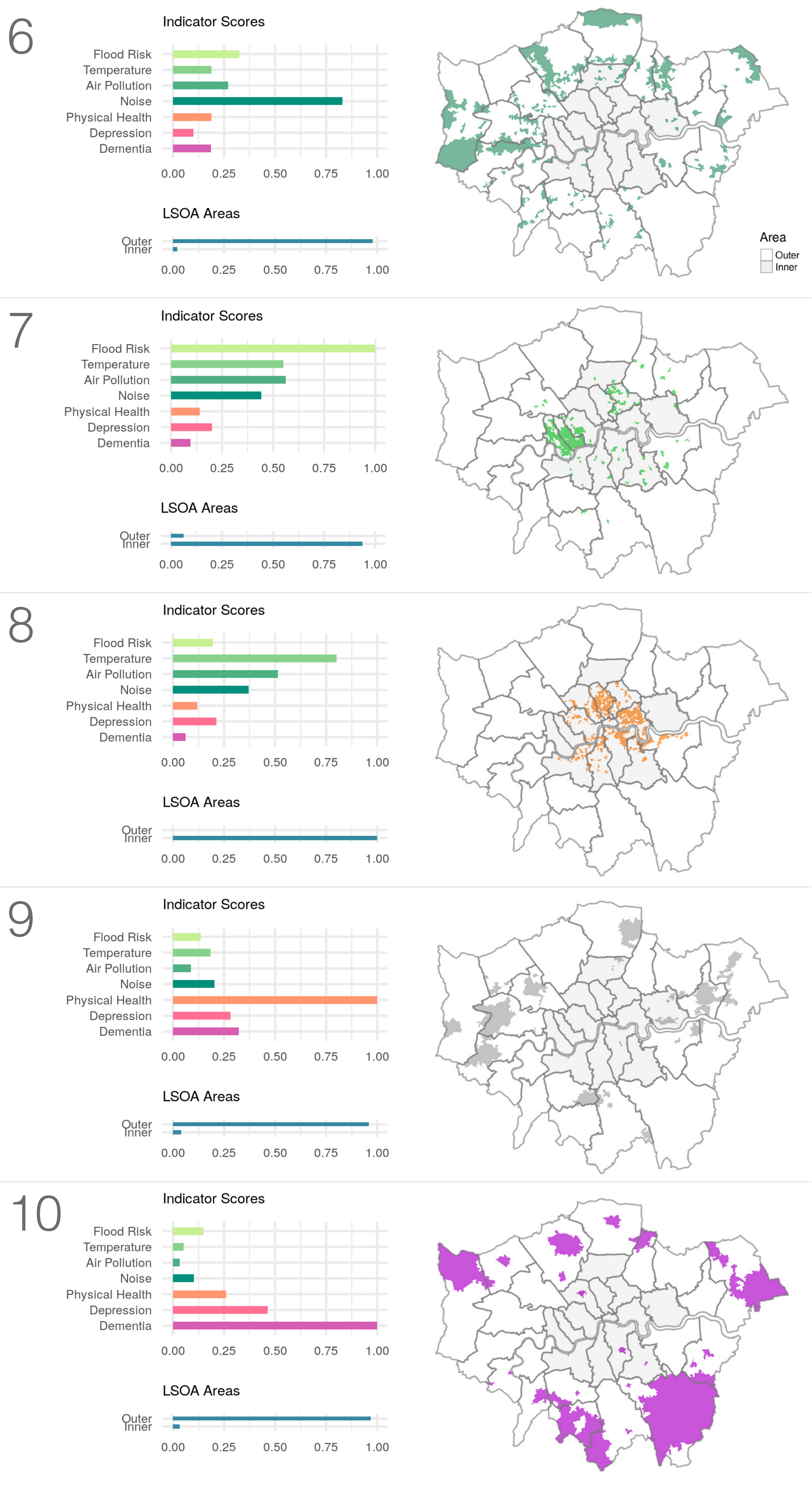

Figure 4.11: Cluster descriptions - clusters 6 to 10

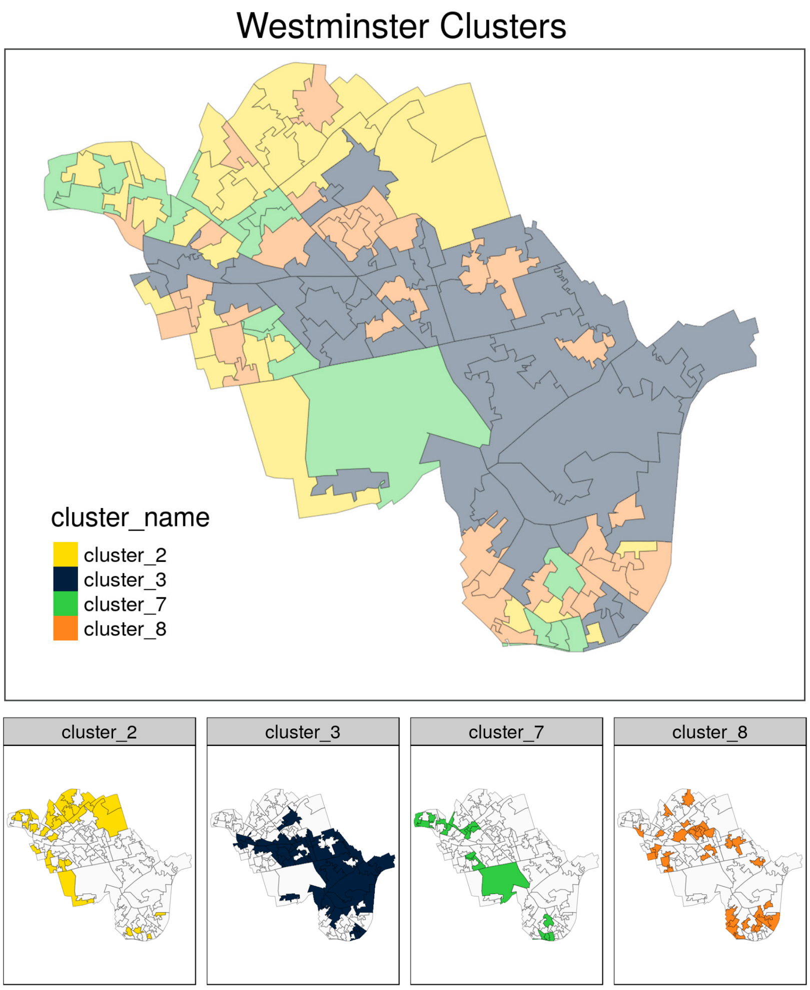

Figure 4.12: Cluster map for Westminster

4.2 Discussion

This research set out to answer the following questions:

- Can publicly available data be used to create a comprehensive measure of GI demand across London, which accounts for the many benefits it can deliver?
- Can such a measure provide insight into localised patterns of GI demand in order to highlight opportunities for multifunctional GI approaches?

The results show that these datasets can be successfully combined to create a measure of GI demand. In this case the final measure may not completely integrate all the benefits urban GI can provide, but still provides a valuable picture of the highest priority areas for London. Perhaps unsurprisingly, much of the high demand is focussed in the centre as a result of the very central focus of some of the regulating indicators. But it is by no means a blanket coverage of high demand, and the granularity of the LSOA level results mean that there is still valuable insight into relative priorities at the borough level, as shown in Westminster. Furthermore, the index highlights some areas in outer London which should be considered equally high priority in terms of GI demand as Central London. Some of these areas, such as Bromley in the south-east, are driven by very high scores in the cultural domain - showing the holistic picture gained by integrating different data.

Decomposing the final index scores allowed the local patterns of demand between indicators to be investigated. These results showed that most of the LADs with highest overall demand had high scores in the regulating domain – though within these high scores the balance between regulating indicators shifts, with some LADs in the top five having far higher flood indicator scores than others for instance. A strong underlying factor in the local balance of demand appears to be contrasting inner / outer focus of the regulating and cultural domains – the air and dementia indicators were highlighted as two key drivers of this in the effect maps in figure 4.7. Clustering this decomposed data proved effective at highlighting areas with clear priorities for different indicators, some of which are very clearly focussed in particular areas. This could be a very valuable input into London-wide strategies for GI service priorities. However, the classification of large areas of London as low-scoring across all indicators shows the challenge of understanding both relative demand and multifunctionality from the index results. The comparative and cumulative elements of the model - designed to highlight high demand - mean variance across low-ranking scores for each indicator is reduced, while high scores are emphasised. The clusters which describe high scores in just one indicator show the model's value in emphasising the highest priority areas for particular services, but slightly less success at highlighting opportunities for multifunctional approaches to GI services.

4.3 Limitations and Developments

A key limitation of this work is the relatively narrow focus of the data used, which in the cultural domain really only captures health aspects. This is largely a result of the lack of reliable data which describes demand for cultural ES that produce more intangible benefits like 'social cohesion'. Being able to integrate some aspect of these ES could greatly increase the diversity of results and improve the robustness of the model. Furthermore, the handling of the existing datasets could likely be improved: each is a very rich source of data and in some cases the transformation from original spatial scale to LSOA could be made more sophisticated in order to capture a greater and more reliable level of detail. This transformation also raises the challenge of the scale part of the Modifiable Areal Unit Problem, in which variance from source data is lost as it is aggregated to a higher spatial scale (Wong, 2009). In this case the aggregation is relatively small scale, but investigating the use of different spatial scales would certainly be worthwhile, in particular comparing the value of the LSOA approach versus the hex-grid approach taken in the GLA's GI Focus Map.

An interesting challenge also arises around the interpretation of the results, and the inherently subjective nature of a CI. While the demand model here successfully measures relative GI demand across London, some of the trends observed – such as the central focus of high regulating indicator demand – suggest that different definitions of relativity might be useful to different users. To use Westminster as an example, it may be less important to WCC that some areas of the borough have extremely low cultural demand relative to the rest of London, and more important to consider cultural demand scores in that area relative to the rest of the borough, or even possibly just relative to other indicator scores in a specific area. This again shows the tension highlighted from the clustering results between using the same index to understand London-wide GI demand and opportunities for multifunctionality at a smaller local scale. As the framework and running of the model is established, it would be possible to investigate the value of running the model at different spatial scales to obtain different understandings of local relative demand.

Possibly the most valuable improvement of this work would be moving beyond considering demand in isolation: there is only so much value to understanding demand when it is viewed in the absence of supply. A direct comparison of the two would deliver even greater insight into where interventions should be prioritised. Creating even a simple measure of supply at LSOA level presents some challenges however, not least in finding a reliable and well-classified source of GI site data. For that reason supply moved beyond the scope of this research, but should certainly be a priority for any future work in this area. Furthermore, the analysis of results here has largely steered clear of suggestions of causality – this is because the final index does not represent a naturally observed phenomenon but rather an artificial composite of many datasets, each of which are influenced by many social and environmental patterns. That said, it is likely investigating the relationship between the index and other data-points not used in its construction could prove informative.

Chapter 5

Conclusion

The aim of this research was to use the methods of Composite Indicator construction, supported by an Environmental Services framework, to build a single measure of demand for Green Infrastructure in London and investigate opportunities for multifunctional GI interventions.

The ES approach proved extremely valuable, lending structure to the selection of services from policy documents and logic to their connection to indicators of demand. The one comparable measure – the GLA's GI focus map – might be open to more criticism for the equal combination of 13 datasets with fairly little conceptual grounding. The ES categories aid too in interpretation of the results, allowing clear breakdowns and insight in to the interplay between different domain indicators. The challenge of the approach comes from data selection, and primarily finding data which might be used as an indication of demand for ES with more intangible value - particularly those in the cultural domain. While the model proved relatively robust to some methodological changes, the relatively limited range of indicators here presented challenges with sensitivity in that the complete removal or re-weighting of individual indicators produced substantial effects on the final results. Thus there would be a double benefit to integrating more data, through widening the scope of the model as well as likely increasing its robustness to small changes in source data or adjustments in weighting. This latter point is important to solve in order to produce a stable and re-runnable model which could be used to track change in demand over time.

Despite these challenges, there has certainly been value in combining these diverse datasets, and it was possible produce results at a level of granularity that could be valuable for a range of users. Particular data-handling methods developed – such as the transformation of QOF health data from GP surgery to LSOA and subsequent shrinkage estimation - could be re-used in many different areas. The overall results delivered an insightful breakdown of demand priorities across London, showing both the highest priority areas as well as how priorities for the different urban GI services shifts across these areas. Clustering proved a useful way to delve deeper into these results, successfully highlighting priority areas for services or combinations of them across London. There were some limitations with these results in their application to the second research question, however. While they certainly highlighted some localised patterns of demand, finding patterns of multifunctionality was slightly less successful. This highlights the key nature of a Composite Indicator: there are so many methodological decisions throughout the process that affect the end outcome, each must be made with a very specific application in mind. In this case the application is relative, London-wide demand for Urban GI services, and this focus doesn't lend itself as well to investigating multifunctionality. Further work could adjust the model approach to find a balance between these two aims. Adjusting the level of compensability between indicators could reduce the emphasis on high-demand scores, or running the model for particular areas could transform the results of relative demand comparisons - essentially creating a more local-scale measure of demand. Indeed, different scales might in fact be useful to different users, and would be worth exploring.

Bibliography

Assuncao, Renato M. et al. (2005). “Empirical Bayes Estimation of Demographic Schedules for Small Areas”. In: *Demography* 42.3, pp. 537–558. ISSN: 1533-7790. DOI: 10.1353/dem.2005.0022. URL: http://link.springer.com/10.1353/dem.2005.0022 (visited on June 27, 2019).

Berg, Magdalena van den et al. (2015). “Health benefits of green spaces in the living environment: A systematic review of epidemiological studies”. In: *Urban Forestry & Urban Greening* 14.4, pp. 806–816. ISSN: 1618-8667. DOI: 10.1016/j.ufug.2015.07.008. URL: http://www.sciencedirect.com/science/article/pii/S1618866715001016 (visited on May 5, 2019).

Bivand, Roger et al. (2019). *spdep: Spatial Dependence: Weighting Schemes, Statistics and Models*. Version 1.1-2. URL: https://CRAN.R-project.org/package=spdep (visited on July 25, 2019).

Burkhard, Benjamin et al. (2012). “Mapping ecosystem service supply, demand and budgets”. In: *Ecological Indicators*. Challenges of sustaining natural capital and ecosystem services 21, pp. 17–29. ISSN: 1470-160X. DOI: 10.1016/j.ecolind.2011.06.019. URL: http://www.sciencedirect.com/science/article/pii/S1470160X11001907 (visited on May 9, 2019).

Chan, Kai M. A., Terre Satterfield, and Joshua Goldstein (2012). “Rethinking ecosystem services to better address and navigate cultural values”. In: *Ecological Economics* 74, pp. 8–18. ISSN: 0921-8009. DOI: 10.1016/j.ecolecon.2011.11.011. URL: http://www.sciencedirect.com/science/article/pii/S0921800911004927 (visited on May 23, 2019).

City of Westminster (2019a). *Westminster City Plan 2019-2040*. URL: https://www.westminster.gov.uk/cityplan2040 (visited on Jan. 4, 2019).

— (2019b). *Westminster Open Spaces and Biodiversity Strategy*. URL: https://www.westminster.gov.uk/openspacesbiodiversity (visited on Apr. 1, 2019).

Clark, Lara P., Dylan B. Millet, and Julian D. Marshall (2014). “National Patterns in Environmental Injustice and Inequality: Outdoor NO2 Air Pollution in the United States”. In: *PLOS ONE* 9.4, e94431. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0094431. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0094431 (visited on Dec. 17, 2018).

DEFRA (2011). *An introductory guide to valuing ecosystem services*. URL: https://www.gov.uk/government/publications/an-introductory-guide-to-valuing-ecosystem-services (visited on May 6, 2019).

DEFRA (2013). *Risk of Flooding from Surface Water Hazard: 3.3 percent annual chance*. URL: https://environment.data.gov.uk/dataset/924d4380-d465-11e4-bf2a-f0def148f590 (visited on July 19, 2019).

Devine, Owen J., Thomas A. Louis, and M. Elizabeth Halloran (1994). “Empirical Bayes Methods for Stabilizing Incidence Rates before Mapping:” in: *Epidemiology* 5.6, pp. 622–630. ISSN: 1044-3983. DOI: 10.1097/00001648-199411000-00010. URL: https://insights.ovid.com/crossref?an=00001648-199411000-00010 (visited on June 27, 2019).

Elmqvist, Thomas, Erik Gómez-Baggethun, and Johannes Langemeyer (2016). *Ecosystem Services Provided by Urban Green Infrastructure*. Routledge Handbooks Online. ISBN: 978-1-138-02508-0 978-1-315-77530-2. DOI: 10.4324/9781315775302-39. URL: http://www.routledgehandbooks.com/doi/10.4324/9781315775302-39 (visited on May 2, 2019).

England Forestry Commission (2015). *London iTree eco project*. Treeconomics London. URL: https://www.forestry.gov.uk/london-itree (visited on Jan. 31, 2019).

EU Science Hub (2015). *10-Step guide for the construction of a composite indicator*. URL: https://ec.europa.eu/jrc/en/coin/10-step-guide/overview (visited on May 3, 2019).

Frigyik, Bela A., Amol Kapila, and Maya R. Gupta (2010). “Introduction to the Dirichlet distribution and related processes”. In: *Department of Electrical Engineering, University of Washignton, UWEETR-2010-0006* 6, pp. 1–27. URL: https://pdfs.semanticscholar.org/775e/5727f5df0cb9bf834af2ea2548a696c27a38.pdf (visited on July 4, 2019).

GBD 2015 Risk Factors Collaborators (2016). “Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015”. In: *Lancet (London, England)* 388.10053, pp. 1659–1724. ISSN: 1474-547X. DOI: 10.1016/S0140-6736(16)31679-8.

GLA (2012a). *All London Green Grid*. London City Hall. URL: https://www.london.gov.uk//what-we-do/environment/parks-green-spaces-and-biodiversity/all-london-green-grid (visited on Aug. 13, 2019).

— (2012b). *Noise Pollution in London - London Datastore*. URL: https://data.london.gov.uk/dataset/noise-pollution-in-london (visited on July 19, 2019).

— (2016). *London Atmospheric Emissions (LAEI) 2016 - London Datastore*. URL: https://data.london.gov.uk/dataset/london-atmospheric-emissions-inventory--laei--2016 (visited on July 19, 2019).

— (2017). *London's Urban Heat Island - Average Summer - London Datastore*. URL: https://data.london.gov.uk/dataset/london-s-urban-heat-island---average-summer (visited on July 19, 2019).

— (2019a). *Green Infrastructure Focus Map - London Datastore*. URL: https://data.london.gov.uk/dataset/green-infrastructure-focus-map (visited on Aug. 16, 2019).

GLA (2019b). *The London Plan: The spatial development strategy for greater London - draft showing minor suggested changes*. Greater London Authority. URL: https://www.london.gov.uk/sites/default/files/draft_london_plan_-showing_minor_suggested_changes_july_2018.pdf (visited on Feb. 13, 2019).

Greco, Salvatore et al. (2019). "On the Methodological Framework of Composite Indices: A Review of the Issues of Weighting, Aggregation, and Robustness". In: *Social Indicators Research* 141.1, pp. 61–94. ISSN: 1573-0921. DOI: 10.1007/s11205-017-1832-9. URL: https://doi.org/10.1007/s11205-017-1832-9 (visited on May 3, 2019).

Haines-Young, Roy and Marion Potschin (2018). *Common International Classification of Ecosystem Services (CICES) V5.1 and Guidance on the Application of the Revised StructureGuidance on the Application of the Revised Structure*. URL: www.cices.eu (visited on May 30, 2019).

Hampton, Kristen H et al. (2011). "Adjusting for sampling variability in sparse data: geostatistical approaches to disease mapping". In: *International Journal of Health Geographics* 10.1, p. 54. ISSN: 1476-072X. DOI: 10.1186/1476-072X-10-54. URL: http://ij-healthgeographics.biomedcentral.com/articles/10.1186/1476-072X-10-54 (visited on June 27, 2019).

Hansen, Rieke and Stephan Pauleit (2014). "From Multifunctionality to Multiple Ecosystem Services? A Conceptual Framework for Multifunctionality in Green Infrastructure Planning for Urban Areas". In: *AMBIO* 43.4, pp. 516–529. ISSN: 1654-7209. DOI: 10.1007/s13280-014-0510-2. URL: https://doi.org/10.1007/s13280-014-0510-2 (visited on May 9, 2019).

Kline, Paul (1993). *An Easy Guide to Factor Analysis*. London, UNITED KINGDOM: Routledge. ISBN: 978-1-317-72560-2. URL: http://ebookcentral.proquest.com/lib/ucl/detail.action?docID=1639423 (visited on July 4, 2019).

Lovell, Sarah Taylor and John R. Taylor (2013). "Supplying urban ecosystem services through multifunctional green infrastructure in the United States". In: *Landscape Ecology* 28.8, pp. 1447–1463. ISSN: 1572-9761. DOI: 10.1007/s10980-013-9912-y. URL: https://doi.org/10.1007/s10980-013-9912-y (visited on May 9, 2019).

Marshall, Roger J. (1991). "Mapping Disease and Mortality Rates Using Empirical Bayes Estimators". In: *Applied Statistics* 40.2, p. 283. ISSN: 00359254. DOI: 10.2307/2347593. URL: https://www.jstor.org/stable/10.2307/2347593?origin=crossref (visited on June 27, 2019).

Mazziotta, Matteo and Adriano Pareto (2013). "Methods for Constructing Composite Indices: One for All or All for One?" In: *Rivista Italiana di Economia Demografia e Statistica* 2, p. 14.

Meerow, Sara and Joshua P. Newell (2017). "Spatial planning for multifunctional green infrastructure: Growing resilience in Detroit". In: *Landscape and Urban Planning* 159, pp. 62–75. ISSN: 0169-2046. DOI: 10.1016/j.landurbplan.2016.10.005. URL: http://www.sciencedirect.com/science/article/pii/S0169204616302250 (visited on Jan. 24, 2019).

Millennium Ecosystem Assessment (2005). URL: https://www.millenniumassessment.org/en/About.html (visited on June 1, 2019).

Ministry of Housing, Communities and Local Government (2016). *Planning practice guidance: The National Planning Policy Framework and relevant planning practice guidance*. URL: https://www.gov.uk/government/collections/planning-practice-guidance (visited on May 6, 2019).

Nardo, Michela et al. (2005). *Tools for Composite Indicators Building*. EUR - Scientific and Technical Research Reports JRC31473. Joint Research Centre - European Commission.

— (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. OECD. ISBN: 978-92-64-04345-9 978-92-64-04346-6. DOI: 10.1787/9789264043466-en. URL: https://www.oecd-ilibrary.org/economics/handbook-on-constructing-composite-indicators-methodology-and-user-guide_9789264043466-en (visited on May 3, 2019).

NHS Digital (2018). *The Quality and Outcomes Framework – Prevalence, Achievements and Exceptions report – Technical Annex*. QOF Online Resources. URL: https://digital.nhs.uk/data-and-information/publications/statistical/quality-and-outcomes-framework-achievement-prevalence-and-exceptions-data/2017-18.

Nicoletti, Giuseppe, Stefano Scarpetta, and Olivier Boylaud (2000). *Summary Indicators of Product Market Regulation with an Extension to Employment Protection Legislation*. Working Paper 226. Paris: Organization for Economic Co-operation and Development. URL: https://www.oecd-ilibrary.org/economics/summary-indicators-of-product-market-regulation-with-an-extension-to-employment-protection-legislation_215182844604 (visited on July 4, 2019).

Potschin, Marion and Roy Haines-Young (2016). *Defining and Measuring Ecosystem Services*. Routledge Handbooks Online. ISBN: 978-1-138-02508-0 978-1-315-77530-2. DOI: 10.4324/9781315775302-4. URL: http://www.routledgehandbooks.com/doi/10.4324/9781315775302-4 (visited on May 3, 2019).

Potschin, Marion, Roy Haines-Young, et al. (2016). *Ecosystem Services in the Twenty-First Century*. Routledge Handbooks Online. ISBN: 978-1-138-02508-0 978-1-315-77530-2. DOI: 10.4324/9781315775302-1. URL: http://www.routledgehandbooks.com/doi/10.4324/9781315775302-1 (visited on May 2, 2019).

Saisana, M., A. Saltelli, and S. Tarantola (2005). “Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators”. In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 168.2, pp. 307–323. ISSN: 1467-985X. DOI: 10.1111/j.1467-985X.2005.00350.x. URL: http://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-985X.2005.00350.x (visited on July 9, 2019).

Saltelli, Andrea (2007). "Composite Indicators between Analysis and Advocacy". In: *Social Indicators Research* 81.1, pp. 65–77. ISSN: 1573-0921. DOI: 10.1007/s11205-006-0024-9. URL: https://doi.org/10.1007/s11205-006-0024-9 (visited on May 3, 2019).

Smith, Thomas et al. (2015). *The English Indices of Deprivation 2015: Technical report.* OCLC: 1003863484. Ministry of Housing, Communities & Local Government. ISBN: 978-1-4098-4689-5. URL: https://www.gov.uk/government/publications/english-indices-of-deprivation-2015-technical-report (visited on Feb. 5, 2019).

Tan, Pang-Ning, Michael Steinbach, and Vipin Kumar (2006). *Introduction to data mining*. Pearson Education India.

Vivid Economics (2017). *Natural Capital Accounts for Public Green Space in London*. URL: https://www.london.gov.uk/sites/default/files/11015viv_natural_capital_account_for_london_v7_full_vis.pdf (visited on Feb. 1, 2019).

Watson, Robert and Steve Albon (2011). *UK National Ecosystem Assessment: Synthesis of the Key Findings*. Reports from the UK National Ecosystem Assessment, pp. 177–218. URL: http://link.springer.com/10.1007/s10640-010-9418-x (visited on Aug. 24, 2019).

Wong, David (2009). "The modifiable area unit problem (MAUP)". In: *The SAGE handbook of spatial analysis*. Ed. by A. Stewart Fotheringham and Peter Rogerson. Los Angeles [i.e., Thousand Oaks, Calif.]: SAGE. ISBN: 978-1-4129-1082-8.

World Health Organisation (2010). *Environment and health risks: a review of the influence and effects of social inequalities*. background document to the Fifth Ministerial Conference on Environment and Health. Parma. DOI: http://www.euro.who.int/en/health-topics/health-determinants/gender/publications/2010/environment-and-health-risks-a-review-of-the-influence-and-effects-of-social-inequalities. URL: http://www.euro.who.int/en/health-topics/health-determinants/gender/publications/2010/environment-and-health-risks-a-review-of-the-influence-and-effects-of-social-inequalities (visited on Dec. 21, 2018).

— (2018). *Environmental Noise Guidelines for the European Region*. The Regional Office for Europe of the World Health Organizatio. URL: http://www.euro.who.int/en/publications/abstracts/environmental-noise-guidelines-for-the-european-region-2018 (visited on June 12, 2019).

Yasaitis, Laura C., Mariana C. Arcaya, and S.V. Subramanian (2015). "Comparison of Estimation Methods for Creating Small Area Rates of Acute Myocardial Infarction Among Medicare Beneficiaries in California". In: *Health & place* 35, pp. 95–104. ISSN: 1353-8292. DOI: 10.1016/j.healthplace.2015.08.003. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5072888/ (visited on June 28, 2019).

Appendix A

Scripts

A.1 Data Processing

A.1.1 air.r

```

library(sf)
library(tmap)
library(rgdal)
library(tidyverse)

# DATA IN ------------------------------------------------------------------>

PM10d <- read_csv('inputs/PostLAEI2013_2020_PM10d.csv')
PM10d_sf <- st_as_sf(PM10d, coords=c('x', 'y'), crs=st_crs(27700)$proj4string)

oa_ref <- read_csv('inputs/Output_Area_to_Lower_Layer_Super_Output_Area_to_Middle_Layer_Super_Output_Area_to_Local_Authority_District_December_2017_Lookup_in_Great_Britain__Classification_Version_2.csv')
lsoa_lon <- filter(oa_ref, RGN11NM == 'London') %>% select(LSOA11CD, LAD17CD, LAD17NM) %>% distinct()

lsoa <- st_read('inputs/LSOA/Lower_Layer_Super_Output_Areas_December_2011_Generalised_Clipped__Boundaries_in_England_and_Wales.shp')
lsoa <- inner_join(lsoa, lsoa_lon, by=c('lsoa11cd'='LSOA11CD'))
lsoa <- st_transform(lsoa, 27700)


# PROCESSING ------------------------------------------------------------------>

# LOOP INPUTS
LADs <- distinct(filter(oa_ref, RGN11NM == 'London'), LAD17NM) # distinct list of all LADs in London, to loop through
final_output_df <- data.frame() # empty frame to load data into

# loop through LADs
for (lad in seq(1, nrow(LADs))) {

  print(paste('PROCESSING DATA FOR: ', LADs$LAD17NM[lad]))

  # filter LSOA sf to just current LAD
  working_lsoa_sf <- (filter(lsoa, LAD17NM == LADs$LAD17NM[lad]))

  print('- creating subset of air data')
  # create temporary sf of just NO2 points in current LAD, for faster processing
  working_NO2_index <- st_contains(st_union(working_lsoa_sf), PM10d_sf, sparse = T)
  working_NO2_sf <-  PM10d_sf[working_NO2_index[[1]],]

  # create matrix to store output for each LAD
  out_matrix <- matrix(ncol=6, nrow=nrow(working_lsoa_sf))
  LAD17NM <- working_lsoa_sf$LAD17NM[[1]]

  # loop through lsoas in the current LAD and work out NO2 scores
  for (i in seq(1, nrow(working_lsoa_sf))){  #nrow(working_lsoa_sf)

    index <- st_contains(working_lsoa_sf[i,1], working_NO2_sf, sparse=T)
    points <- working_NO2_sf[index[[1]],] %>% st_set_geometry(NULL)
```

```

    lsoa11cd <- working_lsoa_sf$lsoa11cd[[i]]
    count <- nrow(points)
    mean <- mean(points[['conct']])
    max  <- max(points[['conct']])
    min  <- min(points[['conct']])

    try(
      out_matrix[i,] <- c(LAD17NM, lsoa11cd, count, mean, max, min)
    )
    if (i %% 50 == 0 ){
    print(paste('-- loaded result no.', i, 'out of', nrow(working_lsoa_sf), ' - ', round(i/nrow(working_lsoa_sf)*100), '% complete'))
    }
  }

  print(paste('- calculations complete for: ', LADs$LAD17NM[lad]))
  # load LAD output into a dataframe and name columns
  output_df <- data.frame(out_matrix)
  names(output_df) <- c('LAD17NM', 'lsoa11cd', 'count', 'mean', 'max', 'min')

  # if it's the first LAD turn into to final_output, else join to existing final_output
  if (lad == 1){
    final_output_df <- output_df
  } else {
    final_output_df <- rbind(final_output_df, output_df)
  }
  # write final_output table to csv
  write_csv(final_output_df, 'outputs/LAD_all_output_PM10d.csv')
  print('- all data loaded to output table')
}


# read all pollution data and join to one table

no2 <- read_csv('outputs/LAD_all_output_NO2.csv') %>% select(lsoa11cd, no2_mean = NO2_mean)
nox <- read_csv('outputs/LAD_all_output_NOx.csv') %>% select(lsoa11cd, nox_mean = mean)
pm10 <- read_csv('outputs/LAD_all_output_PM10.csv') %>% select(lsoa11cd, pm10_mean = mean)
pm25 <- read_csv('outputs/LAD_all_output_PM25.csv') %>% select(lsoa11cd, pm25_mean = mean)

all_air <-  no2 %>% left_join(nox, by = 'lsoa11cd') %>%
  left_join(pm10, by = 'lsoa11cd') %>%
  left_join(pm25, by = 'lsoa11cd')

write_csv(all_air, 'lon_lsoa_pollution_all.csv')
```

A.1.2 flood.r

```

library(sf)
library(tidyverse)
library(tmap)
library(ggplot2)
library(RColorBrewer)

# DATA IN  -------------------------------------------------------------------->

# read in London LSOA sf
lsoa_lon_sf <- st_read('../initial/data_georef/london/london.shp')
lsoa_lon <- lsoa_lon_sf %>% st_set_geometry(NULL)

# flood data
flood <- st_read('flood/RoFSW_london_LSOA/RoFSW_london_LSOA.shp') %>%
  st_transform(27700) %>%
  select(hazard, shape_area, tile_id, LSOA11CD, LAD11NM)


# PROCESSING  -------------------------------------------------------------------->

# empty lists to store results
lsoa_areas <- NULL
```

```
lsoa_flood_areas <- NULL
errors <- NULL

# loop through all London LSOAS
for (row in 1:nrow(lsoa_lon)){

  # string of current LSOA
  lsoa_name <- (paste(lsoa_lon[row, "LSOA11CD"]))
  # print progress every 50 recs
  if (row %% 50 == 0){
    print(paste('working on record', row, '--', lsoa_name))
  }

  # filter flood and LSOA main files down to LSOA in question
  flood_polys <- filter(flood, LSOA11CD == lsoa_name)
  lsoa_poly <- filter(lsoa_lon_sf, LSOA11CD == lsoa_name)

  # ERROR HANDLING
  # create polys of only overlapping areas
  flood_overlaps <- try(st_intersection(flood_polys, lsoa_poly))
  # if intersection errors log lsoa to errors list and append 0s to others
  if(class(flood_overlaps) == 'try-error'){

    print(paste('ERROR at record', row, '--', lsoa_name))
    errors <- append(errors, lsoa_name)
    lsoa_areas <- append(lsoa_areas, 0)
    lsoa_flood_areas <- append(lsoa_flood_areas, 0)

  } else {
  # calculate areas
  lsoa_area <- units::drop_units(st_area(lsoa_poly))
  lsoa_flood_area <- units::drop_units(sum(st_area(flood_overlaps)))

  # append areas to list
  lsoa_areas <- append(lsoa_areas, lsoa_area)
  lsoa_flood_areas <- append(lsoa_flood_areas, lsoa_flood_area)
  }
}

# load results to lon_lsoa table
lsoa_lon$lsoa_area <- lsoa_areas
lsoa_lon$lsoa_flood_area <- lsoa_flood_areas
lsoa_lon$flood_percent <- lsoa_lon$lsoa_flood_area / lsoa_lon$lsoa_area

# check distinct
nrow(lsoa_lon)
nrow(distinct(lsoa_lon))

write_csv(lsoa_lon, 'flood/lon_lsoa_flood_area_percentage.csv')
```

A.1.3 health.r

```

library(sf)
library(tidyverse)
library(tmap)
library(ggplot2)

# DATA IN ------------------------------------------------------------------->

# all prevelance data - 149,100 rows
prev_all <- read_csv('data_health/qof-1718-csv/PREVALENCE.csv')
# all GP surgery data - 806,289 rows
surg_all <- read_csv('data_health/gp_LSOA/2018/gp-reg-pat-prac-lsoa-all.csv')
# organisation reference lookup, filter to just practices in London, check distinct
org_ref_lon <- read_csv('data_health/qof-1718-csv/ORGANISATION_REFERENCE.csv') %>%
  filter(REGION_NAME == 'LONDON') %>%
  select(PRACTICE_CODE, REGION_NAME)

# London LSOA sf
lsoa_lon_sf <- st_read('../initial/data_georef/london/london.shp')
lsoa_lon <- lsoa_lon_sf %>% st_set_geometry(NULL)

# read in population data
pop <- read_csv('data_health/population_age_lsoa/population_age_lsoa.csv')

```

```

# FILTER  ------------------------------------------------------------------>

# filter prevelance data to depression and dementia - 14,200 rows
prev <- prev_all %>% filter(INDICATOR_GROUP_CODE %in% c('DEP', 'DEM', 'CVDPP', 'OB'
    ))
nrow(prev)

# join prevelance to org lookup to restrict to just London practices - 1,286 unique
     GP surgeries
# some surgeries have NA in register count, so replace with 0
prev_lon <- inner_join(prev, org_ref_lon, by = "PRACTICE_CODE") %>%
  mutate(REGISTER = replace_na(REGISTER, 0))

head(prev_lon, 10)
distinct(prev_lon, INDICATOR_GROUP_CODE, PATIENT_LIST_TYPE)
nrow(prev_lon)
nrow(distinct(prev_lon, PRACTICE_CODE))

# Create table of total patient list size per London GP surgery - 1,286 surgeries
    in London
lon_tot_gp_pat_count <- prev_lon %>% filter(PATIENT_LIST_TYPE == 'TOTAL') %>%
  select(PRACTICE_CODE, PATIENT_LIST_SIZE) %>%
  filter()

nrow(lon_tot_gp_pat_count)


# restrict surgery > LSOA lookup to just London GP surgeries by inner joining to my
     London surgeries,
# then create field proportion of patients in each LSOA
surg_lon <- inner_join(surg_all, lon_tot_gp_pat_count, by = 'PRACTICE_CODE') %>%
  select(-c(PUBLICATION, EXTRACT_DATE, PRACTICE_NAME, SEX), LSOA_PATIENT_COUNT = '
    Number of Patients') %>%
  mutate(LSOA_PATIENT_PROP = LSOA_PATIENT_COUNT / PATIENT_LIST_SIZE)

head(surg_lon, 10)
nrow(surg_lon)

# check that the number of patients across LSOA from the distribution table adds up
     to the patient size list from QOF data
count_comparison <- surg_lon %>%
  group_by(PRACTICE_CODE) %>%
  summarise(lsoa_pat = sum(LSOA_PATIENT_COUNT), qof_pat = max(PATIENT_LIST_SIZE),
    prop_sum = sum(LSOA_PATIENT_PROP)) %>%
  mutate(dif = lsoa_pat - qof_pat)

head(count_comparison, 10)

# correct 4,835 LSOAs present, with 1,281 surgeries in total (we seem to have lost
    5 from the pure prevelance data)
print(paste('number of LSOAs in London surgery data:', nrow(distinct(surg_lon, LSOA
    _CODE))))
print(paste('number of GP surgerys in London:', nrow(distinct(surg_lon, PRACTICE_
    CODE))))

# still too many LSOAs in the data, so need to join to just London LSOAs
surg_lon_fin <- surg_lon %>% inner_join(lsoa_lon, by = c('LSOA_CODE' = 'LSOA11CD'))

# check counts again:
print(paste('number of LSOAs in London surgery data:', nrow(distinct(surg_lon_fin,
    LSOA_CODE))))
print(paste('number of GP surgerys in London:', nrow(distinct(surg_lon_fin,
    PRACTICE_CODE))))

# check how many LSOAs there are for each GP surgery, and the max no. of patients
    in any LSOA for each surgery
surg_lsoa_count <- surg_lon_fin %>% group_by(PRACTICE_CODE) %>%
  summarise(n_lsoa = n(), max_pat = max(LSOA_PATIENT_COUNT))

head(surg_lsoa_count, 10)

# check how many surgeries and patients per LSOA
lsoa_surg_count <- surg_lon_fin %>% group_by(LSOA_CODE) %>%
  summarise(practices = n(), patients = sum(LSOA_PATIENT_COUNT))

```

```
# Joining all info together, and create field for proportion of each practice's
    prevalence counts represented in each LSOA
prev_lon_lsoa <- inner_join(prev_lon, surg_lon_fin, by = 'PRACTICE_CODE') %>%
  mutate(INDICATOR_LSOA_COUNT = REGISTER * LSOA_PATIENT_PROP)

# check counts again:
print(paste('number of LSOAs in London surgery data:', nrow(distinct(prev_lon_lsoa,
     LSOA_CODE))))
print(paste('number of GP surgerys in London:', nrow(distinct(prev_lon, PRACTICE_
    CODE))))

head(prev_lon_lsoa, 10)


# FINAL COUNTS PER LSOA, group by LSOA and indicator and sum the count for each
    LSOA
lsoa_prev <- prev_lon_lsoa %>% group_by(LSOA_CODE, INDICATOR_GROUP_CODE) %>%
  summarise(INDICATOR_COUNT = sum(INDICATOR_LSOA_COUNT))

head(lsoa_prev, 10)
print(paste('number of LSOAs in final prev counts:', nrow(distinct(lsoa_prev, LSOA_
    CODE))))
print(paste('number of NAs in final prev indicator count:', nrow(filter(lsoa_prev,
    is.na(INDICATOR_COUNT)))))


# POPULATION COUNTS
    --------------------------------------------------------------------->

# In order to standardise the number of patients in each LSOA, we need to know the
    population of the age band they were meaured from
# DEM - TOTAL
# DEP - 18OV
# CVDPP - 30_74
# OB - 18OV

# population for London LSOAS
pop_lon <- pop %>% rename(LSOA11CD = 'Area Codes', LSOA_POP = 'All Ages') %>%
  inner_join(lsoa_lon, by = 'LSOA11CD')

nrow(pop_lon)

ggplot(data = pop_lon, aes(LSOA_POP)) + geom_histogram()

# total pop
pop_lon_dem <- pop_lon %>% select(LSOA11CD, LSOA_POP) %>%
  mutate(INDICATOR_GROUP_CODE = 'DEM', AGE_BAND = 'TOTAL') %>%
  select(LSOA11CD, INDICATOR_GROUP_CODE, AGE_BAND, LSOA_POP)

head(pop_lon_dem, 10)

# age pops for dep ages
pop_lon_dep <- pop_lon %>% select(LSOA11CD, '18':'90+') %>%
  mutate(INDICATOR_GROUP_CODE = 'DEP', AGE_BAND = '18OV') %>%
  gather('18':'90+', key = 'AGE', value = 'POP') %>%
  group_by(LSOA11CD, INDICATOR_GROUP_CODE, AGE_BAND) %>%
  summarise(LSOA_POP = sum(POP))

head(pop_lon_dep, 10)

# age pops for cvt ages
pop_lon_cvt <- pop_lon %>% select(LSOA11CD, '30':'74') %>%
  mutate(INDICATOR_GROUP_CODE = 'CVDPP', AGE_BAND = '30_74') %>%
  gather('30':'74', key = 'AGE', value = 'POP') %>%
  group_by(LSOA11CD, INDICATOR_GROUP_CODE, AGE_BAND) %>%
  summarise(LSOA_POP = sum(POP))

head(pop_lon_cvt, 10)

# age pops for OB ages
pop_lon_ob <- pop_lon %>% select(LSOA11CD, '18':'90+') %>%
  mutate(INDICATOR_GROUP_CODE = 'OB', AGE_BAND = '18OV') %>%
  gather('18':'90+', key = 'AGE', value = 'POP') %>%
  group_by(LSOA11CD, INDICATOR_GROUP_CODE, AGE_BAND) %>%
  summarise(LSOA_POP = sum(POP))
```

```r

head(pop_lon_ob, 10)

# combine into one narrow table of pop counts per indicator list per LSOA
pop_lon_inds <- plyr::rbind.fill(pop_lon_dep, pop_lon_dem, pop_lon_cvt, pop_lon_ob)

head(pop_lon_inds, 10)
nrow(pop_lon_inds)
nrow(distinct(pop_lon_inds, LSOA11CD))

# check table looks correct for each indicator
pop_lon_inds %>% group_by(INDICATOR_GROUP_CODE, AGE_BAND) %>%
  summarise(count_lsoa = n(), mean_pop = mean(LSOA_POP), max_pop = max(LSOA_POP), min_pop = min(LSOA_POP))


# % POPULATION CALCS ----------------------------------------------------------------->

# Now we have lsoa pop for each list type age band we can turn our prevalence counts per lsoa to % of pop per LSOA

# left join prevelance data to population by LSOA and the indicator
lsoa_prev_pop <- lsoa_prev %>% left_join(pop_lon_inds, by = c('LSOA_CODE' = 'LSOA11CD', 'INDICATOR_GROUP_CODE' = 'INDICATOR_GROUP_CODE'))

head(lsoa_prev_pop, 10)
sum(is.na(lsoa_prev_pop))

lsoa_prev_pop$INDICATOR_PERC <- lsoa_prev_pop$INDICATOR_COUNT / lsoa_prev_pop$LSOA_POP

# write_csv(lsoa_prev_pop, 'data_health/lon_lsoa_prev_and_pop_narrow2_wOB.csv')
lsoa_prev_pop <- read_csv('data_health/lon_lsoa_prev_and_pop_narrow2_wOB.csv')

# create wide version using spread
lsoa_prev_wide <- lsoa_prev_pop %>%
  select(LSOA_CODE, INDICATOR_GROUP_CODE, INDICATOR_PERC) %>%
  spread(key = INDICATOR_GROUP_CODE, value = INDICATOR_PERC)

head(lsoa_prev_wide, 10)
sum(is.na(lsoa_prev_wide))
nrow(lsoa_prev_wide)
nrow(distinct(lsoa_prev_wide))

write_csv(lsoa_prev_wide, 'data_health/lon_lsoa_ind_prevalence_percentage_of_pop2_wOB.csv')
```

A.1.4 noise.r

```r

library(sf)
library(tidyverse)
library(tmap)

# DATA IN ------------------------------------------------------------->
# read in London LSOA sf
lsoa_lon_sf <- st_read('../initial/data_georef/london/london.shp')
lsoa_lon <- lsoa_lon_sf %>% st_set_geometry(NULL)

# NOISE DATA
road_lon <- st_read('data_noise/Road_Lnight_London/Road_Lnight_London.shp') %>%
  st_transform(27700)

rail_lon <- st_read('data_noise/Rail_Lnight_London/Rail_Lnight_London.shp') %>%
  st_transform(27700)


# take random sample to check category counts
road_lon_samp <- road_lon[sample(1:nrow(road_lon), 1000), ]
# check classes
road_lon_samp %>% st_set_geometry(NULL) %>%
  count(NoiseClass)

rail_lon_samp <- rail_lon[sample(1:nrow(rail_lon), 1000), ]
# check classes
```

```
rail_lon_samp %>% st_set_geometry(NULL) %>%
  count(NoiseClass)

# LSOA LEVELS ------------------------------------------------------------>

# list of boroughs
boroughs <- lsoa_lon_sf %>% st_set_geometry(NULL) %>% distinct(LAD17NM)
# output table
final_output <- tibble()
# log for error LSOAs
errors <- NULL

for (b in seq(1, nrow(boroughs))){

  lsoas <- filter(lsoa, LAD17NM %in% boroughs[b, ])
  b_boundary <- st_union(lsoas)

  print(paste('calculating noise overlap with', boroughs[b, ]))

  borough_noise <- st_intersection(b_boundary, rail_lon) %>%
    st_union() %>%
    st_buffer(0)

  # empty lists to store results from below
  lsoa_area <- NULL
  noise_area <- NULL

  print(paste('running through', nrow(lsoas), 'LSOAs in', boroughs[b, ]))

  for (l in seq(1, nrow(lsoas))){

    print(l)
    lsoa_overlap <- try(st_intersection(lsoas[l, ], borough_noise))

    if(class(lsoa_overlap) == 'try-error'){

      err <- lsoas[l, ] %>% st_set_geometry(NULL) %>% select(LSOA11CD)
      errors <- append(errors, err[1,])

      lsoa_area <- append(lsoa_area, units::set_units(0, m^2))
      noise_area <- append(noise_area, units::set_units(0, m^2))

    } else {
      print('success')

      # work out areas for LSOA and noise intersection
      l_area <- st_area(lsoas[l, ])
      n_area <- st_area(lsoa_overlap)
      # append l_area to list
      lsoa_area <- append(lsoa_area, l_area)
      # only append n_area if it has a value, else append 0
      noise_area <- append(noise_area, ifelse(length(n_area) > 0, n_area, units::
    set_units(0, m^2)))
    }
  }

  lsoa_vals <- lsoas %>% st_set_geometry(NULL)
  borough_output <- tibble(lsoa_vals$LSOA11CD, lsoa_area, noise_area)
  final_output <- rbind(final_output, borough_output)
  write_csv(final_output, 'noise/Rail_Lnight_LSOA_area_covered.csv')
}

nrow(final_output)

# lsoa_errors <- tibble(errors = errors)
# write_csv(lsoa_errors, 'noise/Road_Lnight_LSOA_area_errors.csv')
```

A.1.5 temperature.r

```

library(raster)
library(sf)
library(tidyverse)
library(tmap)

# DATA IN ------------------------------------------------------------>
```

```
# LONDON SHAPEFILE >> SF @ 27700CRF
lon <- st_read('london_boundaries/LSOA_2011_London_gen_MHW.shp') %>%
  st_transform(27700) %>%
  select(LSOA11CD, LAD11NM)

# LONDON TEMPERATURE DATA
temp <- raster('temperature/London_Tmin_midnight_2011.tif')

# TEMP DATA >> SF @ 27700CRF
temp_point <- rasterToPoints(temp, spatial = TRUE) %>%
  st_as_sf() %>%
  rename(temp = London_Tmin_midnight_2011) %>%
  st_transform(27700)


# TRANSFORM POINTS TO POLY ------------------------------------------->
# join points to the london LSOA polys, group and summarise
lon_temp <- st_join(lon, temp_point, join = st_contains) %>%
  group_by(LSOA11CD) %>%
  summarise(count = sum(!is.na(temp)), mean = mean(temp))


# DEAL WITH MISSING VALUES ------------------------------------------->
lon_temp_missings <- filter(lon_temp, count == 0)
print(paste("missing values n =", nrow(lon_temp_missings)))

# empty lists
LSOA11CD <- NULL
temp <- NULL

for (m in seq(1, nrow(lon_temp_missings))) {
# for (m in seq(1, 3)) {

  # LSOA we're missing data for
  missing <- lon_temp_missings[m,]
  # all surrounding LSOAs, making sure no NAs included
  # (there are 11 of the 84 LSOAs which have NAs in surrounding LSOAs)
  touching <- lon_temp[missing, op = st_overlaps] %>%
    filter(!is.na(mean))
  # remove geometry and take just character value of LSOA11CD
  st_geometry(missing) <- NULL
  missing <- as.character(droplevels(missing$LSOA11CD))
  # caluclate mean of all touching LSOAs
  missing_temp <- mean(touching$mean)
  # append values to list
  LSOA11CD <- append(LSOA11CD, missing)
  temp <- append(temp, missing_temp)
}

# load into table
temp_missings <- tibble(LSOA11CD, temp)
# check for NAs
filter(temp_missings, is.na(temp))
# append inferred values to missing temps SF
lon_temp_missings <- left_join(lon_temp_missings, temp_missings, by = 'LSOA11CD')


# JOIN BACK WITH VALID DATA  ------------------------------------------->

# filter all temps SF to no missing values, and create temp column
lon_temp_valid <- filter(lon_temp, count > 0)
lon_temp_valid$temp <- lon_temp_valid$mean
# double check missing + valid is same length as starting table
print(paste("all LSOAs =", nrow(lon_temp)))
print(paste("missing + valid LSOAs =", nrow(lon_temp_missings) + nrow(lon_temp_valid)))
# table of just london boroughs
lon_regs <- lon %>% st_set_geometry(NULL)

# bind valid and missings, select fields, and add back on boroughs
lon_temp_fin <- rbind(lon_temp_valid, lon_temp_missings) %>%
  select(LSOA11CD, count, temp) %>%
  left_join(lon_regs, by = 'LSOA11CD')

# check final table is distinct
```

```
print(paste("rows n =", nrow(lon_temp_fin)))
print(paste("distinct rows n =", nrow(distinct(lon_temp_fin))))


# SAVE ------------------------------------------------------------------->

st_write(lon_temp_fin, 'data_temperature/lon_tmin_LSOA_final/lon_lsoa_tempmin.shp')
lon_temp_fin %>% st_set_geometry(NULL) %>% write_csv('temperature/lon_lsoa_tempmin.csv')
```

A.2 Index Methods

A.2.1 dataConsolidate.r

```

library(tidyverse)


# script to consolidate all processed data so far, and produce summary
# stats for all

# DATA IN ------------------------------------------------------------------->

# all indicator data
health     <- read_csv('../initial/data_health/lon_lsoa_ind_prevalence_percentage_of_pop2_wOB.csv') %>% rename(LSOA11CD = LSOA_CODE)
pollution  <- read_csv('../initial/data_pollution/lon_lsoa_pollution_all.csv') %>% select(LSOA11CD = lsoa11cd, no2 = no2_mean, nox = nox_mean, pm10 = pm10_mean, pm25 = pm25_mean)
temp       <- read_csv('../initial/data_temperature/lon_lsoa_tempmin.csv') %>% select(LSOA11CD, temp)
noise_road <- read_csv('../initial/data_noise/lon_lsoa_Road_Lnight_area_covered_perc.csv') %>% select(LSOA11CD, noise_road = noise_perc)
noise_rail <- read_csv('../initial/data_noise/lon_lsoa_Rail_Lnight_area_covered_perc.csv') %>% select(LSOA11CD, noise_rail = noise_perc)
flood      <- read_csv('../initial/data_flood/lon_lsoa_flood_area_percentage.csv') %>% select(LSOA11CD, flood = flood_percent)

ind_lookup <- read_csv('dataset_ind_cat_lookup.csv')

# London LSOA data - 4,835 rows
lsoa_lon <- read_csv('../initial/data_georef/Output_Area_to_Lower_Layer_Super_Output_Area_to_Middle_Layer_Super_Output_Area_to_Local_Authority_District_December_2017_Lookup_in_Great_Britain__Classification_Version_2.csv') %>%
  filter(RGN11NM == 'London') %>%
  select(LSOA11CD, LAD17CD, LAD17NM) %>%
  distinct()

lon_all <- lsoa_lon %>% left_join(health, by = 'LSOA11CD') %>%
  left_join(pollution, by = 'LSOA11CD') %>%
  left_join(temp, by = 'LSOA11CD') %>%
  left_join(noise_road, by = 'LSOA11CD') %>%
  left_join(noise_rail, by = 'LSOA11CD') %>%
  left_join(flood, by = 'LSOA11CD')

nrow(lon_all)
nrow(distinct(lon_all))

head(lon_all, 10)
head(ind_lookup)
# write_csv(lon_all, 'lon_all_data.csv')
# lon_all <- read_csv('lon_all_data.csv')

# SUMMARY TABLE ------------------------------------------------------------------->

tibble(names(lon_all))

lon_all_nar <- lon_all %>% gather('CVDPP':'flood', key = 'data', value = 'value') %>%
  inner_join(ind_lookup, by = 'data')


# summary_stats <-
```

```r
summary <- lon_all_nar %>% group_by(category, cat_order, indicator, ind_order, data_name) %>%
  summarise(count = n(), min = min(value), max = max(value), mean = mean(value), std = sd(value))

write_csv(summary, 'summary_stats.csv')
```

A.2.2 dataShrinkage.r

```r

library(tidyverse)
library(ggplot2)
library(sf)
library(sp)
library(spdep)
library(tmap)
library(gridExtra)

# script to carry out shrinkage estimation using local empirical bayes
# with three differentneighbourhood definitions for each lsoa. All
# results stored, which can then be used in uncertainty and sensitivity tests

# DATA IN ------------------------------------------------------------------>

# London LSOA data - 4,835 rows
lsoa_lon_sf <- st_read('../initial/data_georef/london/london.shp')
lsoa_lon <- lsoa_lon_sf %>% st_set_geometry(NULL)


lad_sf <- st_read('../london_boundaries/London_Borough_Excluding_MHW.shp') %>%
  st_transform(27700)

prev_all <- read_csv('../initial/data_health/lon_lsoa_prev_and_pop_narrow2_wOB.csv')

# join all prev data to london sf, this new sf will contain 4 records per LSOA
lon_all_sf <- lsoa_lon_sf %>% left_join(prev_all, by = c('LSOA11CD' = 'LSOA_CODE'))



# SPATIAL WEIGHTS ---------------------------------------------------------------->

lsoa_lon_sp <- as(lsoa_lon_sf, 'Spatial')

# Create QUEENS weights matrix
nb_queens <- poly2nb(lsoa_lon_sp, queen=TRUE)

# plot distribution of number of neighbours for each LSOA
nb_queen_tally <- tibble(n_neighbours = card(nb_queens))

# Create K-N weights matrix
# input fields
coords <- coordinates(lsoa_lon_sp)
ids <- row.names(as(lsoa_lon_sp, 'data.frame'))

# calculate k nearest neighbours
nb_kn10 <- knn2nb(knearneigh(coords, k = 10), row.names = ids)


# tally and plot for k10
nb_kn10_tally <- tibble(n_neighbours = card(nb_kn10)) %>% group_by(n_neighbours) %>% summarise(count = n())


# calculate BLOCK weight matrix based on LAD membership
nb_lad <- read.gal('lon_lad_blocks.gal', override.id = TRUE)
nb_lad_tally <- tibble(n_neighbours = card(nb_lad))


# EMPIRICAL BAYES ---------------------------------------------------------------->

# list of indicators for loop
indicators <- c('CVDPP', 'DEM', 'DEP', 'OB')
# empty table to load results into
```

```r
eb_results <- tibble()

for (i in seq(1, 4)){

  print(paste('Running Empirical Bayes Smoothing for indicator', indicators[i]))

  # filter sf to just current indicator
  lon_sf <- filter(lon_all_sf, INDICATOR_GROUP_CODE == indicators[i])
  # create sp from the london sf
  lon_sp <- as(lon_sf, 'Spatial')
  # run local empirical bayes for each weight matrix
  eb_queens <- EBlocal(lon_sp$INDICATOR_COUNT, lon_sp$LSOA_POP, nb_queens)
  eb_kn10 <- EBlocal(lon_sp$INDICATOR_COUNT, lon_sp$LSOA_POP, nb_kn10)
  eb_lad <- EBlocal(lon_sp$INDICATOR_COUNT, lon_sp$LSOA_POP, nb_lad)
  # load outputs from each to table, also just calculating basic rate
  q <- tibble(LSOA11CD = lon_sf$LSOA11CD, INDICATOR_GROUP_CODE = indicators[i], METHOD = 'QUEENS', rate = eb_queens$est)
  k <- tibble(LSOA11CD = lon_sf$LSOA11CD, INDICATOR_GROUP_CODE = indicators[i], METHOD = 'KN-10', rate = eb_kn10$est)
  l <- tibble(LSOA11CD = lon_sf$LSOA11CD, INDICATOR_GROUP_CODE = indicators[i], METHOD = 'LAD', rate = eb_lad$est)
  # bind all into one table
  out <- plyr::rbind.fill(q, k, l)
  # bind to final results table
  eb_results <- rbind(eb_results, out)
}

# check all results loaded, and distinct
count(eb_results, INDICATOR_GROUP_CODE, METHOD)
nrow(eb_results)
nrow(distinct(eb_results))
head(eb_results)

# calculate basic rate from prev table and load into results too
basic_rate <- tibble(LSOA11CD = prev_all$LSOA_CODE, INDICATOR_GROUP_CODE = prev_all$INDICATOR_GROUP_CODE,
                     METHOD = 'BASIC', rate = prev_all$INDICATOR_COUNT / prev_all$LSOA_POP)

eb_results <- rbind(eb_results, basic_rate)

# check distribution of results across ind and method
results_check <- eb_results %>% group_by(INDICATOR_GROUP_CODE, METHOD_F) %>%
  summarise(count = n(), min = min(rate), max = max(rate), mean = mean(rate), sd = sd(rate))

results_check

write_csv(eb_results, 'lsoa_health_shrunk_scores_all_final.csv')
```

A.2.3 dataPcaHealth.py

```python

from sklearn.preprocessing import StandardScaler
import numpy as np
from factor_analyzer import FactorAnalyzer
import pandas as pd
import factor_analyzer

health_raw = pd.read_csv('../shrinkage/lsoa_health_shrunk_scores_all_final.csv',
                         usecols = ['LSOA11CD', 'INDICATOR_GROUP_CODE', 'METHOD', 'rate'])

# take kn-10 shrinkage results health data
health_raw = health_raw[(health_raw['METHOD'] == 'KN-10')]
health_raw = health_raw[['LSOA11CD', 'INDICATOR_GROUP_CODE', 'rate']]

# pivot to wide table
health = health_raw.pivot(index = 'LSOA11CD', columns = 'INDICATOR_GROUP_CODE', values = 'rate')

health.columns.name = None
health = health.reset_index()

# check correct no. of LSOAs
print(len(health))
```

```
health.head()

# take just indicators and standardise
cols = ['DEM', 'DEP', 'CVDPP', 'OB']

vals = health[cols].values
ss = StandardScaler()
health_s = pd.DataFrame(ss.fit_transform(vals), columns = cols)
health_s.head()


# run factor analysis

fa = FactorAnalyzer()
fa.analyze(health_s, 4, method = 'principal', rotation = 'varimax')

fa.loadings.to_csv('pca_results_health.csv')
```

A.2.4 dataPcaAir.py

```

from sklearn.preprocessing import StandardScaler
from factor_analyzer import FactorAnalyzer
import pandas as pd
import factor_analyzer

air = pd.read_csv('../initial/data_pollution/lon_lsoa_pollution_all.csv')
air.head()

# take just indicators and standardise
vals = air.iloc[:,1:].values

ss = StandardScaler()
air_s = pd.DataFrame(ss.fit_transform(vals), columns = air.columns[1:])

air_s.describe()

# run factor analysis

fa = FactorAnalyzer()
fa.analyze(air_s, 4, method = 'principal', rotation = None)

fa.loadings.to_csv('pca_results_air_.csv')
```

A.2.5 dataAggregation.r

```

library(tidyverse)

# script to aggregate all final CI data using equal weights in
# all aggregations

# DATA IN ------------------------------------------------------------------->

lon_all <- read_csv('../descriptive_stats/lon_all_data.csv')
health_shrunk <- read_csv('../shrinkage/lsoa_health_shrunk_scores_all_final.csv')
air <- read_csv('../pca/lon_lsoa_pollution_factors.csv')

# VARIABLES & FUNCTIONS
    ---------------------------------------------------------------------->

# define shrinkage method to use for health data
shrinkage_method <- 'KN-10'

# read in functions
source('functions.r')

# PROCESSING ---------------------------------------------------------------->

# HEALTH
# select health data for appropriate shrinkage method and spread to wide table
health_cur <- filter(health_shrunk, METHOD == shrinkage_method) %>%
  select(-c('METHOD', 'METHOD_F')) %>%
  spread(key = INDICATOR_GROUP_CODE, value = rate)

head(health_cur)
```

```
nrow(health_cur)

# standardise the indicator fields
std <- health_cur %>% select(-(LSOA11CD)) %>% mutate_all(list(~scale(.) %>% as.vector))
std['LSOA11CD'] <- health_cur['LSOA11CD']


# domain weighting
health_w1 <- std %>% mutate(PHYS = (CVDPP * 0.5) + (OB * 0.5))
health_w1[, 'phys_exp'] <- rank_exp(health_w1, PHYS)
health_w1[, 'dem_exp'] <- rank_exp(health_w1, DEM)
health_w1[, 'dep_exp'] <- rank_exp(health_w1, DEP)

health_w1 <- health_w1 %>%
  mutate(health_score = (phys_exp / 3) + (dem_exp / 3) + (dep_exp) /3)

health_w1[, 'health_exp'] <- rank_exp(health_w1, health_score)
health_fin <- select(health_w1, LSOA11CD, phys_exp, dem_exp, dep_exp, health_score, health_exp)

# NOISE
noise_w1 <- select(lon_all, LSOA11CD, noise_road)
noise_w1[, 'noise_exp'] <- rank_exp(noise_w1, noise_road)
noise_fin <- select(noise_w1, LSOA11CD, noise_exp)


# AIR POLLUTION
air_temp <- select(air, LSOA11CD = lsoa11cd, value = factor_1)

air_w1 <- air_temp
air_w1['air_exp'] <- rank_exp(air_w1, value)
air_fin <- select(air_w1, LSOA11CD, air_exp)

# REMAINING REG INDICATORS
reg_all <- select(lon_all, LSOA11CD)
reg_all['flood_exp'] <- rank_exp(lon_all, flood)
reg_all['temp_exp'] <- rank_exp(lon_all, temp)

reg_all <- reg_all %>% left_join(noise_fin, by = 'LSOA11CD') %>%
  left_join(air_fin, by = 'LSOA11CD')

# REGULATING DOMAIN
reg_w1 <- reg_all %>%
  mutate(reg_score = (flood_exp / 4) + (temp_exp / 4) + (noise_exp / 4) + (air_exp/4))

reg_w1['reg_exp'] <- rank_exp(reg_w1, reg_score)
reg_fin <- select(reg_w1, LSOA11CD, flood_exp, temp_exp, noise_exp, air_exp, reg_score, reg_exp)


# FINAL INDEX
gi_temp <- reg_fin %>%
  left_join(health_fin, by = 'LSOA11CD')

gi_w1 <- gi_temp %>%
  mutate(gi_score = (reg_exp / 2) + (health_exp / 2))

gi_w1['gi_exp'] <- rank_exp(gi_w1, gi_score)

gi_w1 <- gi_w1 %>% mutate(gi_rank = min_rank(gi_score),
                          gi_decile = ntile(gi_score, 10))


write_csv(gi_w1, 'outputs/gi_basic_scores.csv')
```

A.3 Index Testing

A.3.1 functions.r

```

# EXPONENTIAL TRANSFORMATION AND RANKING

# exponential transformation function
```

```
exp_trans <- function(var){
  -23 * log(1 - (var * (1 - (exp(-100/23)))))
}

# function to rank and exponentially transform a variable in a given table
rank_exp <- function(df, var){

  var <- enquo(var)
  # select relevant data from df
  data <- select(df, LSOA11CD, !! var)
  # create a ranking, rescale to 0-1, then apply exponential transformation
  data %>%
    mutate(rank = min_rank(!! var),
           rank_s = scales::rescale(rank, to = c(0, 1)),
           exp = exp_trans(rank_s)) %>%
    select(exp)
}


# DIRICHLET WEIGHT SAMPLES

# move last item from a list back n positions
lCycle <- function(x, n = 1) {
  if (n == 0) x else c(tail(x, -n), head(x, n))
}

# this function creates a table of nsamples from the dirichlet distribution for
    nweights (no. of indicators)
# each of the nweights will be treated with the alpha values in turn, while the
    others remain at 1
dirSampleSpread <- function(alpha_vals, nsamples, nweights){

  out <- data.frame()
  # loop through alpha values
  for (a in seq(1, length(alpha_vals))){
    # create a list of 1 for each weight, with the current alpha value at the end
    weights <- c(rep(1, (nweights - 1)), alpha_vals[a])

    for (w in seq(0, nweights - 1)){

      # cycle the weight list so that each weight gets sampled with the current
    alpha treatment
      w_list <- lCycle(weights, w)
      # create dirichlet samples and turn to data frame
      weight_samples_m <- rdirichlet(nsamples, w_list)
      weight_samples <- as.data.frame(weight_samples_m)
      # rename weight variables and then store alpha value
      names(weight_samples) <- paste('w', seq(1, nweights), sep = '')
      # weight_samples$alpha <- paste('alpha = ', alpha_vals[a])
      weight_samples$alpha <-alpha_vals[a]
      weight_samples$w_focus <- nweights - w

      # # bind to output table
      out <- rbind(out, weight_samples)
    }
  }
  out
}
```

A.3.2 testingUA.r

```

library(tidyverse)
library(ggplot2)
library(gridExtra)
library(gtools)
library(tmap)
library(tmaptools)
library(sf)

# processing for uncertainty analysis - essentially puts the index
# aggregation script in a really big nested loop (...) to test all
# combinations of input factors. Stores final CI score for every LSOA
# in one big table, as well as summary stats for each test variation

# DATA IN  ------------------------------------------------------------------>
```

```

# read in London LSOA sf
lsoa_lon_sf <- st_read('../initial/data_georef/london/london.shp')
# indicator data
lon_all <- read_csv('../descriptive_stats/lon_all_data.csv')
# health data after shrinkage
health_shrunk <- read_csv('../shrinkage/lsoa_health_shrunk_scores_all_final.csv')
# all air pollution indicator data
air <- read_csv('../initial/data_pollution/lon_lsoa_pollution_all.csv')
# factorised air pollution data
air_factor <- read_csv('../pca/lon_lsoa_pollution_factors.csv')


# FUNCTIONS ------------------------------------------------------------------->

# read in functions
source('functions.r')

# define lists to use in looping
shrinkage_tests <- c('BASIC', 'QUEENS', 'KN-10', 'LAD')
noise_tests <- c('basic')
air_tests <- c('no2_only', 'all_equal', 'factor')

regulating_weights <- list('1' = c(0.25, 0.25, 0.25, 0.25),
                           '2' = c( 0, 0.25, 0.25, 0.25),
                           '3' = c(0.25, 0, 0.25, 0.25),
                           '4' = c(0.25, 0.25, 0, 0.25),
                           '5' = c(0.25, 0.25, 0.25, 0))

cultural_weights <- list('1' = c(1/3, 1/3, 1/3),
                         '2' = c(0, 0.5, 0.5),
                         '3' = c(0.5, 0, 0.5),
                         '4' = c(0.5, 0.5, 0))

# PROCESSING ------------------------------------------------------------------->

ua_output <- tibble()

# run loop - will create 240 versions of index and store results for each
for (air_t in seq(1, length(air_tests))){

  for (reg_w in seq(1, length(regulating_weights))){

    # NOISE
    noise_w1 <- select(lon_all, LSOA11CD, noise_road)
    noise_w1[, 'noise_exp'] <- rank_exp(noise_w1, noise_road)
    noise_fin <- select(noise_w1, LSOA11CD, noise_exp)

    # AIR POLLUTION
    # no2 only
    air_temp <- select(lon_all, LSOA11CD, no2)
    air_no2 <- air_temp
    air_no2['air_exp'] <- rank_exp(air_no2, no2)
    # all equal
    air_all <- select(air, LSOA11CD = lsoa11cd, no2_mean, nox_mean, pm10_mean, pm25
    _mean) %>%
      mutate(all = (no2_mean / 4) + (nox_mean / 4) + (pm10_mean / 4) + (pm25_mean /
     4))

    air_all['air_exp'] <- rank_exp(air_all, all)

    # factor
    air_f <- select(air_factor, LSOA11CD = lsoa11cd, factor_1)
    air_f['air_exp'] <- rank_exp(air_f, factor_1)

    # select final based on air test var
    if (air_t == 1){
      air_fin <- select(air_no2, LSOA11CD, air_exp)
    }
    if (air_t == 2){
      air_fin <- select(air_all, LSOA11CD, air_exp)
    }
    if (air_t == 3){
      air_fin <- select(air_f, LSOA11CD, air_exp)
    }

```

```
# REMAINING REG INDICATORS
reg_all <- select(lon_all, LSOA11CD)
# exponential transform for remaining two indicators
reg_all['flood_exp'] <- rank_exp(lon_all, flood)
reg_all['temp_exp'] <- rank_exp(lon_all, temp)
# join to final noise and air tables
reg_all <- reg_all %>% left_join(noise_fin, by = 'LSOA11CD') %>%
  left_join(air_fin, by = 'LSOA11CD')

# REGULATING DOMAIN - final weighting
reg_weight <- reg_all %>%
  mutate(reg_score = (flood_exp * regulating_weights[[reg_w]][1]) + (temp_exp * regulating_weights[[reg_w]][2]) +
           (noise_exp * regulating_weights[[reg_w]][3]) + (air_exp * regulating_weights[[reg_w]][4]))

reg_weight['reg_exp'] <- rank_exp(reg_weight, reg_score)
reg_fin <- select(reg_weight, LSOA11CD, reg_exp)


for (shr_t in seq(1, length(shrinkage_tests))){

  # HEALTH
  # select health data for appropriate shrinkage method and spread to wide table
  health_cur <- filter(health_shrunk, METHOD == shrinkage_tests[shr_t]) %>%
    select(-c('METHOD', 'METHOD_F')) %>%
    spread(key = INDICATOR_GROUP_CODE, value = rate)

  # standardise the indicator fields
  std <- health_cur %>% select(-(LSOA11CD)) %>% mutate_all(list(~scale(.) %>% as.vector))
  std['LSOA11CD'] <- health_cur['LSOA11CD']

  for (cul_w in seq(1, length(cultural_weights))){

    # DOMAIN WEIGHTING
    # combine obesity and cvd into physical indicator
    health_w1 <- std %>% mutate(PHYS = (CVDPP * 0.5) + (OB * 0.5))
    health_w1['phys_exp'] <- rank_exp(health_w1, PHYS)
    health_w1['dem_exp'] <- rank_exp(health_w1, DEM)
    health_w1['dep_exp'] <- rank_exp(health_w1, DEP)

    # use weights from list to calculate score
    health_w1 <- health_w1 %>%
      mutate(health_score = (phys_exp * cultural_weights[[cul_w]][1]) + (dem_exp * cultural_weights[[cul_w]][2]) +
               (dep_exp * cultural_weights[[cul_w]][3]))

    # transform to exponential and create final table
    health_w1['health_exp'] <- rank_exp(health_w1, health_score)
    health_fin <- select(health_w1, LSOA11CD, health_exp)

    # FINAL INDEX
    gi_temp <- reg_fin %>%
      left_join(health_fin, by = 'LSOA11CD')

    gi_w1 <- gi_temp %>%
      mutate(gi_score = (reg_exp / 2) + (health_exp / 2))

    gi_w1['gi_exp'] <- rank_exp(gi_w1, gi_score)

    gi_w1 <- gi_w1 %>% mutate(gi_rank = min_rank(gi_score),
                              gi_decile = ntile(gi_score, 10))

    # print current iteration
    print(paste('shr: ', shrinkage_tests[shr_t], '   cul_w: ', cul_w, '   air_t: ', air_tests[air_t], '   reg_w: ', reg_w))

    # create final output table
    gi_fin <- select(gi_w1, LSOA11CD, gi_score, gi_exp, gi_rank, gi_decile)

    # add variable fields
    gi_fin$shrinkage_test <- shrinkage_tests[shr_t]
    gi_fin$cultural_weight <- cul_w
    gi_fin$air_test <- air_tests[air_t]
```

```
        gi_fin$noise_test <- 'basic'
        gi_fin$regulating_weight <- reg_w
        gi_fin$gi_weight <- 'basic'

        ua_output <- rbind(ua_output, gi_fin)
      }
    }
  }
}

write_csv(ua_output, 'uncertainty_analysis_raw_output.csv')

# filter to just the test we're calling the baseline
baseline <- filter(ua_output, shrinkage_test == 'KN-10', cultural_weight == 1, air_test == 'factor', regulating_weight == 1) %>%
  select(LSOA11CD, gi_rank_b = gi_rank)

# save baseine to csv for use in sensitivity analysis
write_csv(baseline, 'outputs/gi_baseline_scores.csv')

# left join baseline to full table, so we can work out each iterations score compared to the baseline
ua_final <- ua_output %>% left_join(baseline, by = 'LSOA11CD') %>%
  mutate(gi_rank_change = abs(gi_rank - gi_rank_b), rank_change_perc = gi_rank_change / 4835)

# write / read as necessary
# write_csv(sens_final, 'outputs/UA_all_tests_rank_change.csv')
# ua_final <- read_csv('outputs/UA_all_tests_rank_change.csv')

# summary stats for each variation of the model run
ua_vars <- ua_final %>% group_by(shrinkage_test, cultural_weight, air_test, noise_test, regulating_weight, gi_weight) %>%
  summarise(mean_rank_change = mean(rank_change_perc), min_rank_change = min(rank_change_perc),
            max_rank_change = max(rank_change_perc),sd_rank_change = sd(rank_change_perc), count = n())


# check we're grouping all used indicators, result below should be 0
filter(ua_vars, count != 4835)

write_csv(ua_vars, 'sensitivity_analysis_variants.csv')


# LSOA RANK VARIANCE PER TEST ---------------------------------------------------------------------->

lsoa_lon <- lsoa_lon_sf %>% st_set_geometry(NULL)

# join to get LAD field and create unique test identifier to group by
results <- ua_final %>% left_join(lsoa_lon, by = 'LSOA11CD') %>%
  unite(test, c('shrinkage_test', 'air_test', 'cultural_weight', 'regulating_weight'), remove = FALSE)

nrow(distinct(results, test))

# for each test factor calculate mean rank for each LSOA then deviation in mean rank across each test
dev_shrinkage <- results %>% group_by(LSOA11CD, shrinkage_test) %>%
  summarise(mean_rank = mean(gi_rank)) %>%
  group_by(LSOA11CD) %>%
  summarise(rank_dev = sd(mean_rank) / 4835, test = 'shrinkage')

dev_air <- results %>% group_by(LSOA11CD, air_test) %>%
  summarise(mean_rank = mean(gi_rank)) %>%
  group_by(LSOA11CD) %>%
  summarise(rank_dev = sd(mean_rank) / 4835, test = 'air')

dev_cultural <- results %>% group_by(LSOA11CD, cultural_weight) %>%
  summarise(mean_rank = mean(gi_rank)) %>%
  group_by(LSOA11CD) %>%
  summarise(rank_dev = sd(mean_rank) / 4835, test = 'cultural_weights')

dev_regulating <- results %>% group_by(LSOA11CD, regulating_weight) %>%
  summarise(mean_rank = mean(gi_rank)) %>%
```

```
  group_by(LSOA11CD) %>%
  summarise(rank_dev = sd(mean_rank) / 4835, test = 'regulating_weights')

# bind all tables together
dev_all <- plyr::rbind.fill(dev_shrinkage, dev_air, dev_cultural, dev_regulating)
head(dev_all)
# check we have 4835 records for each test
dev_all %>% group_by(test) %>% summarise(count = n())
```

A.3.3 testingSA.r

```

library(tidyverse)
library(ggplot2)
library(gridExtra)
library(gtools)

# processing for sensitivity analysis - similar set up to the UA processing,
# but here the loop is focussed just on one aggregation step and runs through
# a big table of different weight samples created using the dirichlet
# function at the beginning. Here only Rs for each test is recorded.

# DATA IN  ------------------------------------------------------------------>

lon_all <- read_csv('../descriptive_stats/lon_all_data.csv')
health_shrunk <- read_csv('../shrinkage/lsoa_health_shrunk_scores_all_final.csv')
air <- read_csv('../pca/lon_lsoa_pollution_factors.csv')

baseline <- read_csv('../sensitivity/outputs/gi_baseline_scores.csv') %>%
  select(LSOA11CD, gi_rank_b = gi_rank)

# FUNCTIONS AND VARIABLES ------------------------------------------------------------------>

# read in functions
source('functions.r')

# create table of sample weights - 3 variables, 250 samples for each iteration, alpha ranging from 1-10
weights <- dirSampleSpread(alpha_vals = seq(1, 10, 1), nsamples = 250, nweights = 3)

# sense check output
weights <- arrange(weights, w_focus, alpha)
nrow(weights)
head(weights)
head(filter(weights, w_focus == 1))
head(filter(weights, w_focus == 2))
head(filter(weights, w_focus == 3))
weights %>% group_by(w_focus, alpha) %>% summarise(n = n())


# PROCESSING  ------------------------------------------------------------------>

# below runs through standard index aggregation stages, but looping a particular stage using sample weights from above

# NOISE
noise_w1 <- select(lon_all, LSOA11CD, noise_road)
noise_w1[, 'noise_exp'] <- rank_exp(noise_w1, noise_road)
noise_fin <- select(noise_w1, LSOA11CD, noise_exp)

# AIR POLLUTION
# factor
air_f <- select(air, LSOA11CD = lsoa11cd, factor_1)
air_f['air_exp'] <- rank_exp(air_f, factor_1)
air_fin <- select(air_f, LSOA11CD, air_exp)

# REMAINING REG INDICATORS
reg_all <- select(lon_all, LSOA11CD)
# exponential transform for remaining two indicators
reg_all['flood_exp'] <- rank_exp(lon_all, flood)
reg_all['temp_exp'] <- rank_exp(lon_all, temp)

# join to final noise and air tables
reg_all <- reg_all %>% left_join(noise_fin, by = 'LSOA11CD') %>%
```

```
  left_join(air_fin, by = 'LSOA11CD')

# REGULATING DOMAIN - final weighting
reg_weight <- reg_all %>%
  mutate(reg_score = (flood_exp / 4) + (temp_exp / 4) +
           (noise_exp / 4) + (air_exp / 4))

reg_weight['reg_exp'] <- rank_exp(reg_weight, reg_score)
reg_fin <- select(reg_weight, LSOA11CD, reg_exp)


# CULTURAL
# select health data for appropriate shrinkage method and spread to wide table
health_cur <- filter(health_shrunk, METHOD == 'KN-10') %>%
  select(-c('METHOD', 'METHOD_F')) %>%
  spread(key = INDICATOR_GROUP_CODE, value = rate)

# standardise the indicator fields
std <- health_cur %>% select(-(LSOA11CD)) %>% mutate_all(list(~scale(.) %>% as.
    vector))
std['LSOA11CD'] <- health_cur['LSOA11CD']


# DOMAIN WEIGHTING
# combine obesity and cvd into physical indicator
health_w1 <- std %>% mutate(PHYS = (CVDPP * 0.5) + (OB * 0.5))
health_w1['phys_exp'] <- rank_exp(health_w1, PHYS)
health_w1['dem_exp'] <- rank_exp(health_w1, DEM)
health_w1['dep_exp'] <- rank_exp(health_w1, DEP)


# LOOP FOR SENSITIVITY
score_list <- NULL

for (w in seq(1, nrow(weights))){

  # use weights from list to calculate score
  health_w1 <- health_w1 %>%
    mutate(health_score = (phys_exp * weights[w, 'w1']) + (dem_exp * weights[w, 'w2
    ']) +
             (dep_exp * weights[w, 'w3']))

  # transform to exponential and create final table
  health_w1['health_exp'] <- rank_exp(health_w1, health_score)
  health_fin <- select(health_w1, LSOA11CD, health_exp)

  # FINAL INDEX
  gi_temp <- reg_fin %>%
    left_join(health_fin, by = 'LSOA11CD')

  gi_w1 <- gi_temp %>%
    mutate(gi_score = (reg_exp / 2) + (health_exp / 2))

  gi_w1 <- gi_w1 %>% mutate(gi_rank = min_rank(gi_score))

  # create final output table - joining to baseline to calculate rank change per
    LSOA
  gi_fin <- select(gi_w1, LSOA11CD, gi_rank) %>%
    left_join(baseline, by = 'LSOA11CD') %>%
    mutate(rank_change = abs(gi_rank - gi_rank_b))

  mean_r_change <- mean(gi_fin$rank_change)
  score_list <- append(score_list, mean_r_change)
}

# append score list (of mean rank change per test) back to weights table
weights$rank_change <- score_list
weights$rank_change_percent <- score_list / 4835

write_csv(weights, 'outputs/sensitivity_dirichlet_cultural_results.csv')
```