Quantifying social vulnerability for decision-making on pluvial flood adaptation

Master Thesis

M.W.A. Swinkels BSc (Mark)

Graduation committee

dr. M.S. Krol (Maarten)
E. Bakhshianlamouki MSc (Elham)
D.B. van den Heuvel MSc (Daniël)

UNIVERSITY OF TWENTE.

Witteveen + Bos

Photo on cover from VLN-Nieuws (2022)

Preface

Enschede, 27-08-2023

Dear reader,

with the completion of this master's thesis, my time as a civil engineering student at the University of Twente comes to an end. In the past half a year I dove into the concept of social vulnerability and this journey has led me to all kinds of topics I was (un)happy to encounter during my thesis. Statistics were never my strongest topic and yet I found myself in the library some months ago to borrow the book *Statistics and data analysis in geology*. On the other hand, I got to make some beautiful maps, which was always one of my favourite tasks to do.

My sincerest thanks go out to everyone who has helped me along the way. First of all, all my fellow students with whom I worked together in the past 6 years. Secondly, my housemates. A listening ear to talk to about the long days of studying was never far away in Huize Edelweiss. And then there were my girlfriend and my family. Unconditional support even through the sometimes challenging times in writing this thesis. Thank you all for creating such a nice environment that made me who I am today.

For this thesis in particular, I need to thank Maarten and Elham as supervisors from the university and Daniël and all other colleagues from Witteveen+Bos. Without their constant stream of feedback on every part of the process and the report, the final product would not have been of the same quality as it is.

I hope you enjoy reading this thesis.

Mark Swinkels

Summary

In the context of flood risk modelling, comprehensive studies are carried out on a global scale to gauge the likelihood and consequences of various flooding events. These studies often focus on the translation of risk into monetary values. However, the concept of social vulnerability is rarely taken into account. This research proposes a method to quantify social vulnerability with publicly available data on an as detailed as possible level for the Dutch urban setting.

By tailoring the indicator list as proposed by Cutter et al. (2003) to the Dutch context, two Social Vulnerability Indices were created for the cities of Haarlem and Zwolle. These indices have been analysed and overlayed with maps of bottleneck locations for pluvial floods.

The results showed clusters of high and low socially vulnerable areas in both cities of Haarlem and Zwolle. For both cities residential property appeared as an important factor for social vulnerability. In Haarlem, this was complemented by socioeconomic status and in Zwolle by race and ethnicity. The addition of bottleneck locations to the maps with social vulnerability learned that in Haarlem the less vulnerable part of society is larger but also has more locations at risk in case of a pluvial flood. In Zwolle, this is the other way around. The more vulnerable part of society is larger and has more locations at risk in case of a pluvial flood.

To conclude, the Social Vulnerability Index makes it possible to quantify social vulnerability. This is an important step in applying the concept of social vulnerability in flood risk management. Implementation within the climate adaptation strategies of Haarlem and Zwolle would be possible in the prioritisation of measures to mitigate the effects of extreme precipitation.

Samenvatting

In de context van overstromingsrisicomodellering worden uitgebreide studies op wereldwijde schaal uitgevoerd, om de waarschijnlijkheid en de gevolgen van verschillende overstromingsgebeurtenissen in te schatten. Deze studies richten zich vaak op de vertaling van risico's in monetaire waarden. Er wordt echter zelden rekening gehouden met het concept van sociale kwetsbaarheid. Dit onderzoek stelt een methode voor om sociale kwetsbaarheid te kwantificeren met openbare gegevens op een zo gedetailleerd mogelijk niveau voor de Nederlandse stedelijke omgeving.

Door de indicatorenlijst zoals voorgesteld door Cutter et al. (2003) aan te passen aan de Nederlandse context, zijn twee Sociale Kwetsbaarheidindices gecreëerd voor de steden Haarlem en Zwolle. Deze indices zijn geanalyseerd en gecombineerd met kaarten van knelpuntlocaties voor overstromingen door extreme regenval.

De resultaten lieten clusters van hoog en laag sociaal kwetsbare gebieden zien in zowel Haarlem en Zwolle. Voor beide steden bleek dat woningen een belangrijke factor zijn voor sociale kwetsbaarheid. In Haarlem werd dit aangevuld met sociaaleconomische status en in Zwolle met ras en etniciteit. De toevoeging van knelpuntlocaties aan de kaarten met sociale kwetsbaarheid leerde dat in Haarlem het minder kwetsbare deel van de samenleving groter is, maar ook meer overstromingsrisicolocaties heeft in geval van extreme neerslag. In Zwolle is dit andersom. Het meer kwetsbare deel van de samenleving is groter en heeft meer risicolocaties in het geval van een overstroming.

Tot slot maakt de Sociale Kwetsbaarheidsindex het mogelijk om sociale kwetsbaarheid te kwantificeren. Dit is een belangrijke stap in de toepassing van het concept van sociale kwetsbaarheid in het overstromingsrisicobeheer. Implementatie binnen de klimaatadaptatiestrategieën van Haarlem en Zwolle zou mogelijk zijn bij de prioritering van maatregelen om de effecten van extreme neerslag te beperken.

Table of contents

1 Introduction

1.1 Background

In the context of flood risk modelling, comprehensive studies are carried out on a global scale to gauge the likelihood and consequences of various flooding events. These events involve three flood types: pluvial, fluvial, and coastal floods, each having its own set of challenges and impacts. Firstly, coastal floods occur due to storms forcing water on the shore. Sea level rise, land subsidence, and climate change are increasing the risks of these floods for coastal cities (Shan et al., 2022). Secondly, fluvial floods exist in many forms, ranging from rivers overflowing their banks to failures in levees along waterways (Mohor et al., 2020). Such failures can occur in various ways and degrees, resulting in different types of floods. For instance, a levee breach can lead to a more acute and impactful event compared to the early stages of piping or wave overtopping. Thirdly, pluvial floods are closely linked to extreme rainfall events. During heavy rainfall, the soil, sewage systems, and surface water reservoirs become overwhelmed and incapable of handling the excess stormwater (Haghighatafshar et al., 2020). As a result, water begins to flow into areas where it inflicts damage, particularly in urban regions. These floods are often referred to as urban floods. Over the past few decades, pluvial floods have gained attention as a significant danger, posing considerable risks to numerous cities. (Bulti & Abebe, 2020). The effects have not only been mapped more due to European regulations but simply the fact that high-intensity rainfall events occur more often and more extremely has led to an increase in awareness (Fritsch et al., 2016; Rangari et al., 2018).

The likelihood and consequences of flooding events come together in risk assessments. Risk assessments for flooding in the Netherlands are commonly based on three things. The probability of a flood, the damage per flood in Euros, and the number of casualties per flood (Deltares, 2018). The effects are thus considered to be based on damage costs and the number of casualties. Several studies show that risk in the case of pluvial floods is mainly based on costs. These costs are sometimes separated into physical impact, for example, houses and other buildings, and intangible losses such as traffic delays, loss of recreational value, and inconvenience (Åström et al., 2014; Löwe et al., 2017; Rijkswaterstaat, 2015; Tiepolo et al., 2021; Zhou et al., 2012). This focus on costs and casualties, however, falls short of the comprehensiveness of risk since there are many factors influencing people before they end up being deceased by a flood. Think of social conditions, coping capacity, and the physical environment they live in (Elboshy et al., 2019). These factors combined can be seen as their vulnerability to a flood. Of course, these factors are not only relevant for casualties but also for economic damage. Risk would, therefore, rather be defined as a function of the hazard related to the risk source (e.g. precipitation, high-intensity rain events) and the vulnerability related to the risk object (e.g. buildings, infrastructure, inhabitants) (Hauger et al., 2006).

1.2 Scientific context

1.2.1 What is vulnerability?

In a broad sense, vulnerability to environmental hazards means the potential for loss (Cutter et al., 2003). According to Blaikie (1994), it “involves the characteristics of a person or group in terms of their capacity to anticipate, cope with, resist, and recover”. However, in literature, a multitude of definitions of vulnerability can be found, each with its focus. This highlights that fundamental differences exist in the various visions of what vulnerability entails. Figure 1 shows the focus of the three leading visions on vulnerability with an example of a definition corresponding to the vision. Figure 2 shows the research themes in vulnerability studies and includes some papers indicating the different themes. On top of these different visions and themes, reducing vulnerability to a singular metric or quantifying it easily is a challenging task. (Adger, 2006).

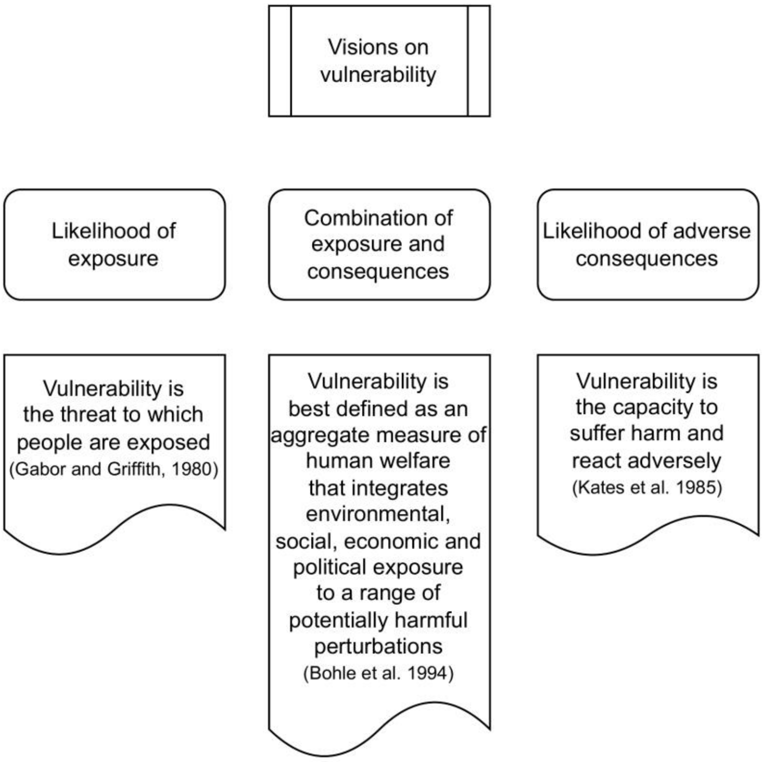

Figure 1: Visions on vulnerability

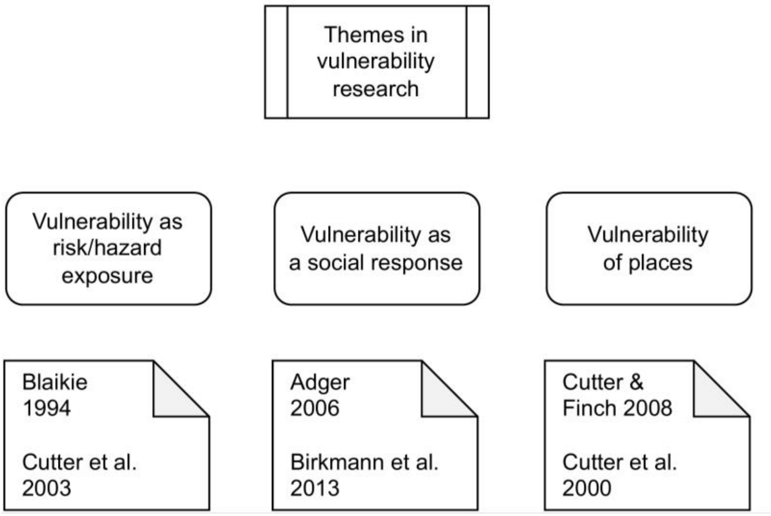

Figure 2: Themes in vulnerability research

1.2.2 What is social vulnerability?

Social vulnerability is often considered to be related to poverty (Philip & Rayhan, 2004), however, social vulnerability is far more complex than just the existence of poverty (Laska & Morrow, 2006). Social vulnerability encompasses all elements closely connected to how hazards intersect with individuals, populations, and communities. This covers the exposure of individuals, socio-demographic and socio-economic characteristics, employment, education, household makeup, demographic composition, and society's ability to manage both hazards and their consequences. (Tascón-González et al., 2020). All those factors show that vulnerability is deeply embedded in social structures, which are often resistant to change (Birkmann et al., 2013). In addition, the social dimensions of vulnerability can help in recognizing and understanding whether certain groups or communities are more sensitive and prone to impacts. (Chakraborty et al., 2020). Based on this knowledge base more targeted solutions and strategies for effective mitigation can be enabled (Tapsell et al., 2010).

1.2.3 Risk

When reading the previous sections on vulnerability and social vulnerability one might think that when replacing the word (social) vulnerability with risk, the text still makes sense. This is indeed somewhat the case. The main differences between vulnerability and risk in the climate and disaster fields are rooted not in conceptual differences but in the usage of different terminologies. (Wolf, 2012). The variation is mainly clarified by the unique origins of these concepts within separate communities. (Janssen & Ostrom, 2006). This thesis does not aim to dive deep into these definitions and the different meanings research communities give to them, nor does it find the similarities or bridge any gap in this context.

1.3 State-of-the-art

The concept of social vulnerability has received attention from researchers since the late twentieth century, mostly in the context of natural disasters (Adger, 2006; Blaikie, 1994; Cutter et al., 2003; Tascón-González et al., 2020). Based on the hypothesis that there is a strong correlation between low socioeconomic status and high vulnerability, tools and methods for measuring and assessing social vulnerability have been developed (Blaikie, 1994; Kuhlicke et al., 2011). This hypothesis is not tested in literature but rather incorporated in visions and views on vulnerability. Socioeconomic status, or social inequality, is seen as a proxy for social vulnerability to some extent. Other factors like race, disability, and age also play a role. This thesis will not dive into this hypothesis, and it adopts the link between socioeconomic status and social vulnerability.

The methods to measure and assess social vulnerability include the use of (1) quantitative indicators such as income, education, and access to healthcare, (2) utility functions for social welfare, as well as more qualitative approaches such as (3) community-based research (Cutter et al., 2003; Kind, 2019; Kind et al., 2020; Kind et al., 2017; Kuhlicke et al., 2011; Pathak et al., 2020; Tate, 2013; Tiepolo et al., 2021). The first method will be used in this thesis and further elaborated on in Chapter 2.2.

In addition to this, social vulnerability has been analysed with the use of geospatial analysis to map and model vulnerability patterns, as well as to identify the drivers and determinants of vulnerability (Forrest et al., 2020; Koks et al., 2015). This geospatial presentation provides the capacity to identify the social vulnerability of places and allows to compare and contrast places (Chakraborty et al., 2020). The interdisciplinary approach has allowed researchers to explore the complex interactions between social, economic, and environmental factors that contribute to vulnerability. Moreover, it allows to develop more effective strategies for reducing vulnerability and building resilience in communities.

1.4 Research gap

Current research that has been done on social vulnerability mainly focuses on the United States of America (USA). This is not surprising since Susan L. Cutter, one of the major scholars in this field, is a professor at the University of South Carolina. The approach to assess social vulnerability with the use of quantitative indicators has originally been made for the context of the United States (Cutter et al., 2003). This is visible in indicators such as the percentage of African Americans, the percentage of Native Americans, and the percentage of houses that are mobile homes, telling something about race and residential property. Using the indicator approach in other contexts than the USA has led to indicators being removed and or added to specify the local context (Fekete, 2009; Forrest et al., 2020; Koks et al., 2015; Tascón-González et al., 2020). These studies do not change the methodology as described by Cutter et al. (2003) much but adapt the approach to their local context. However, the number of studies doing so is limited, especially when comparing scope and scale.

Implementation of social vulnerability is very common in all sorts of policy terrains (e.g., allowance system, welfare benefits, Wmo (social support), legal assistance, scholarships, and energy transition). This is however not the case when looking at climate adaptation. Despite social vulnerability being researched in the context of floods, vulnerability is only considered to be related to exposure in terms of direct damage in this field (Leusink & Swets, 2017; Rozendaal, 2023; Zwolle, 2021). The choice of where and when to implement small-scale measures does not include a social vulnerability component yet.

To overcome this research gap, this research proposes a method to quantify social vulnerability with publicly available data on an as detailed as possible level for the Dutch urban setting. This aspect can then be considered when implementing small-scale measures against pluvial flooding.

1.5 Problem statement

1.5.1 Problem definition

While expressing flood risks in terms of (monetary) costs and casualties has its advantages, it also has clear limitations. For multiple reasons, vulnerable groups often experience the consequences of flooding differently than less vulnerable groups. Examples include (1) less affluent neighbourhoods are oftentimes situated in more flood-prone areas (Blaikie, 1994), (2) vulnerable groups have less financial reserves to repair any flood damage to their properties and (3) vulnerable groups own less (expensive) property and hence suffer fewer damages in absolute terms. As a result, the absolute monetary damages are bound to be less in poorer neighbourhoods, while the relative impact may be similar or higher than in affluent neighbourhoods. It is more or less the current practice in the Netherlands that this perception of vulnerability is not taken into consideration when determining flood adaptation measures in cities.

The relationship between social vulnerability and flood adaptation measures, moreover, remains understudied (El-Zein et al., 2021; Pallathadka, 2021). In the assessment phase, only either physical vulnerability or homogeneous vulnerability of the population is considered (Jongman et al., 2012; Jonkman et al., 2003; Koks et al., 2014). Additionally, in the evaluation phase, social vulnerability is not included either (Koks et al., 2015).

1.5.2 Witteveen+Bos, Deltares, Haarlem, Zwolle

Deltares has asked Witteveen+Bos to join in a TKI (Topconsortium voor Kennis en Innovatie) to create a method to quantify social vulnerability. This can then be used to better incorporate this in projects and decision-making on measures for climate adaptation. Vulnerable groups in society experience the negative effects of floods often differently and traditional methods like a probabilistic risk approach or a cost-benefit analysis often cannot address this well. This is increasingly a point for attention in national and international projects. The goal of Deltares is to standardise this and create a toolbox for organisations worldwide. With this toolbox, they know how and when they can use which tools to map social vulnerabilities to water risks and incorporate this in policy considerations. Witteveen+Bos is interested in this and wants to obtain insights into this topic and state-of-the-art knowledge on incorporating social vulnerability in flood studies and measure programmes. For this, they have come up with a graduation assignment for a master's student to quantify social vulnerability for flood adaptation. The specific department of Witteveen+Bos that is involved has experience with the sewage system in Haarlem and especially for urban pluvial floods the sewage system is an important aspect that can determine the severity. This is the reason Haarlem will be used as a case study.

Furthermore, this thesis is part of Delta Futures Lab Zwolle. This is a thematic working group with the 4TU Centre for Resilience Engineering. Their goal is to explore how spatial developments and investments in the upcoming 10 to 20 years in the Zwolle region interact with a changing water system in the long run (2100). In this project, guidance is available from lecturers of the 4 TU’s, waterboard Drents Overijsselse Delta, and the municipality of Zwolle. Therefore, a case study will also be done on Zwolle.

1.6 Research objective and questions

1.6.1 Research objective

The research objective is to contribute to the consideration of social vulnerabilities in Dutch flood risk assessments by exploring an indicator-based approach and applying this to a case study.

1.6.2 Research questions

Based on the problem statement and the research objective the following research questions have been set up.

1. What is social vulnerability in relation to floods?
2. How can the existing methods to quantify social vulnerability be applied in flood risk management?
3. How can Haarlem and Zwolle incorporate social vulnerability into their climate adaptation strategy?

1.7 Scope

To demarcate this thesis, some limitations are used (see Table 1). The social vulnerability index approach looks at vulnerability as vulnerability of places. Based on data availability, this will be looked at not on an individual level but at a community level. Since the research objective states that the interest is in the Dutch setting, this is a logical limitation as well.

Table 1: Limitations to scope thesis

	Limitation
Vulnerability theme	Vulnerability of places
Social vulnerability	Factors influencing communities, not on an individual level
Hazard vulnerability	Only for implementation in flood risk management
	Only related to pluvial floods
Methods	Indicator approach (Cutter et al., 2003)
Geography	Dutch urban setting, focus on Haarlem and Zwolle

1.8 Structure of this thesis

The remainder of this thesis is structured as follows:

- Chapter two will dive deeper into the indicator approach. Besides, it will explore how the concept of social vulnerability could be considered when deciding on flood prevention measures
- Chapter three describes the methodology and introduces the case studies
- Chapter four presents the results of the case studies
- Chapters five, six and seven are respectively the discussion, conclusion, and recommendations

2 Theoretical background

2.1 Social vulnerability analysis

Social vulnerability analysis (SVA) outlines social characteristics, vulnerability to hazards, and how tangible hazard impacts are distributed (Remo et al., 2016). In this context, social vulnerability refers to the characteristics of an individual or a group and the conditions that impact their ability to anticipate, manage, or recover from the consequences of a hazard (Blaikie, 1994). Strictly speaking, performing a vulnerability analysis entails not simplifying or regarding any aspect of the human-environment system as a mere boundary condition. (Polsky et al., 2007). First, the study area is characterised after which the drivers for vulnerability are identified. The third step is to develop a quantitative vulnerability model and finally, the findings should be communicated to stakeholders (Tate, 2012). The model development starts with the creation of a social vulnerability index (SoVI).

2.2 Social vulnerability index

Progress in conceptual frameworks for social vulnerability and the growing focus on creating quantitative metrics have resulted in a diverse range of methods being used to build indices (Tate, 2012). The underlying hypothesis of these taxonomic approaches to quantify vulnerability is a strong positive connection between lower socio-economic status and increased vulnerability. (Kuhlicke et al., 2011). The basis of most research is a long list of indices by Cutter et al. (2003). They proposed forty-one indicators that contribute to the level of social vulnerability in an area. Other researchers used (part of) this list to determine social vulnerability based on the availability of data or based on the characteristics of the study area (Fekete, 2009; Forrest et al., 2020; Tascón-González et al., 2020). To cluster these indices into one value for the level of social vulnerability there are three common designs. Deductive, hierarchical, and inductive models (Tate, 2012).

2.2.1 Index structures

In a deductive model (Figure 3a), several indicators are normalised and aggregated. Examples are research done by Cutter et al. (2000) and Lein and Abel (2010) in which factors like the number of mobile homes and mean house value or population under 18 years of age and population over 65 years of age respectively are summed up to a total score after being normalised.

A hierarchical model (Figure 3b) has indicators aggregated in groups which are aggregated by themselves in the end. Flanagan et al. (2011), for example, aggregated income, poverty, employment, and education into a domain called socioeconomic status based on a percentile rank. This same method was then used to aggregate four domains (socioeconomic status, household composition, minorities and language, and housing and transportation) to a total SVI score.

The most used model today is the inductive model (Figure 3c). In this model, a large set of indicators is reduced to a set of uncorrelated latent factors using principal component analysis (PCA). The principal component analysis strives to merge a diverse correlation of indicators to include a maximum of information from each original dataset. Specifically, it adeptly identifies data patterns to minimize loss of information while decreasing the dataset's extensive dimensions. (Kim et al., 2021). This model was used in this thesis.

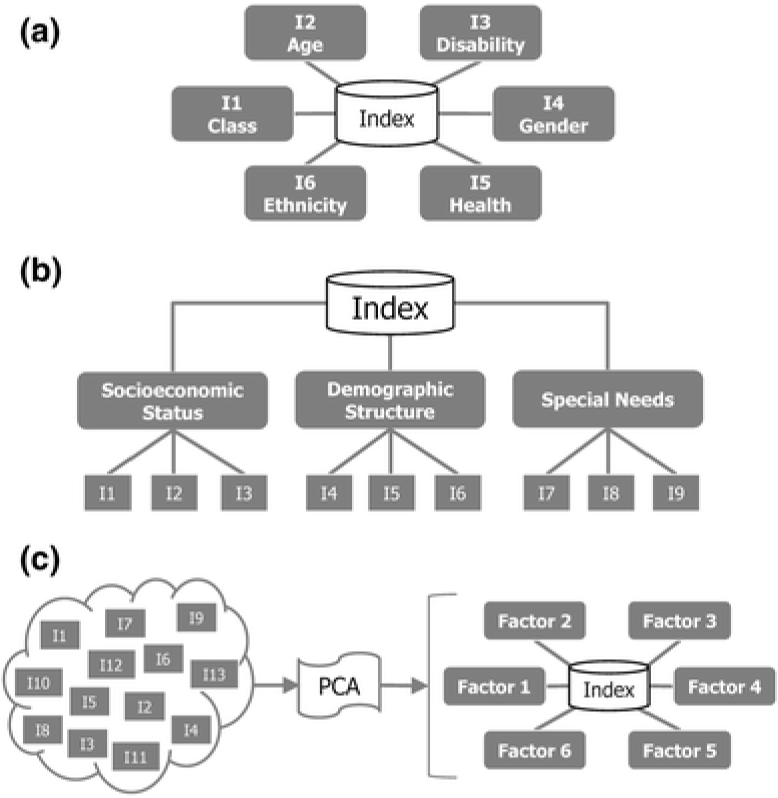

Figure 3: Vulnerability index structural designs (Tate, 2012)

An elaboration of the inductive model has been introduced by Cutter et al. (2003) with a total of 43 possible indicators for social vulnerability. Since then, many researchers have studied their approach and used it for other case studies (Fekete, 2009; Kirby et al., 2019; Koks et al., 2015; Lee, 2014; Schmidtlein et al., 2008; Tate, 2013).

2.2.2 Indicators

Within the social science community, there is a great acknowledgement of population characteristics that influence social vulnerability (Cutter et al., 2001; H. John Heinz Iii Center for Science & the, 2002; Holand & Lujala, 2013) (Table 2). For each population characteristic indicators show if they have a positive or negative impact on the degree of social vulnerability. For age, for example, the elderly and children have a positive impact but for education highly educated has a negative impact. The criteria have been selected based on their contribution to social vulnerability to natural hazards in general. This is not specific to pluvial flood events. A list of that will be presented later in this report.

Table 2: Population Characteristics Influencing Social Vulnerability (Cutter et al., 2001)

Population Characteristic	Description	Increases (+) or Decreases (−) Social Vulnerability
Socioeconomic status (income, political power, prestige)	Status affects the ability to absorb losses and enhance resilience to hazard impacts. Wealth enables communities to absorb and recover from losses more quickly using insurance, social safety nets, and entitlement programs.	High status (+/−) Low income or status (+) (*Sources*: Cutter et al. 2000, Peacock et al. 2000, Puente 1999, Bolin and Stanford 1998, Blaikie et al. 1994, Burton et al. 1993)
Gender	Women often have a more difficult time during recovery than men because of sector-specific employment (e.g., personal services), lower wages, and family care responsibilities.	Gender (+) (*Sources*: Peacock et al. 2000, Enarson and Scanlon 1999, Morrow and Phillips 1999, Enarson and Morrow 1998, Hewitt 1997, Fothergill 1996, Morrow and Enarson 1996)
Race and ethnicity	These factors impose language and cultural barriers and affect access to post-disaster funding and occupation of high-hazard areas.	Non-white (+) Non-Anglo (+) (*Sources*: Pulido 2000, Peacock et al. 2000, Morrow and Phillips 1999, Bolin and Stanford 1998, Bolin 1993)
Age	Extremes of age affect the movement out of harm's way. Parents lose time and money caring for children when day care facilities are affected; the elderly may have mobility constraints or concerns that increase the burden of care and lack of resilience.	Elderly (+) Children (+) (*Sources*: Ngo 2001, Cutter et al. 2000, Hewitt 1997, O'Brien and Mileti 1992)
Commercial and industrial development	The value, quality, and density of commercial and industrial buildings provide indicators of the state of economic health of a community, potential losses in the business community, and longer-term issues with recovery after an event.	High density (+) High value (+/−) (*Sources*: Heinz Center 2000b, Webb et al. 2000)
Employment loss	The potential loss of additional employment following a disaster increases the possible number of unemployed workers in a community. Such losses contribute to a slower recovery from the disaster.	Employment loss (+) (*Source*: Mileti 1999)
Rural/urban	Rural residents may be more vulnerable because of lower incomes and more dependence on a locally based resource economy (e.g., farming or fishing). High-density areas (urban) complicate evacuation out of harm's way.	Rural (+) Urban (+) (*Sources*: Cutter et al. 2000, Cova and Church 1997)
Residential property	The value, quality, and density of residential construction affect potential losses and recovery. Expensive homes on the coast are costly to replace; mobile homes are easily destroyed and less resilient to hazards.	Mobile homes (+) (*Sources*: Cutter et al. 2000, Heinz Center 2000b, Bolin and Stanford 1991)
Infrastructure and lifelines	The loss of sewer, bridges, water, communications, and transportation infrastructure compounds potential disaster losses. The loss of infrastructure may place an insurmountable financial burden on smaller communities that lack the financial resources to rebuild.	Extensive infrastructure (+) (*Sources*: Heinz Center 2000b, Platt 1995)
Renters	People rent because they are transients, do not have the financial resources for home ownership, or do not want the responsibility of home ownership. They often lack access to information about financial aid during recovery. In extreme cases, renters lack sufficient shelter options when lodging becomes uninhabitable or too costly to afford.	Renters (+) (*Sources*: Heinz Center 2000b, Morrow 1999)

(continued)

Population Characteristic	Description	Increases (+) or Decreases (−) Social Vulnerability
Occupation	Some occupations, especially those involving resource extraction, may be severely affected by a hazard event. Self-employed fishermen suffer when their means of production is lost, and they may not have the requisite capital to resume work in a timely fashion; therefore, they may seek alternative employment. Migrant workers engaged in agriculture and low-skilled service jobs (housekeeping, child care, and gardening) may suffer similarly as disposable income fades and the need for services declines. Immigration status also affects occupational recovery.	Professional or managerial (−) Clerical or laborer (+) Service sector (+) (*Sources*: Heinz Center 2000b, Puente 1999, Hewitt 1997)
Family structure	Families with large numbers of dependents and single-parent households often have limited wherewithal to outsource care for dependents and thus must juggle work responsibilities and care for family members. All these factors affect resilience to and recovery from hazards.	High birth rates (+) Large families (+) Single-parent households (+) (*Sources*: Heinz Center 2000b, Morrow 1999, Puente 1999, Morrow 1997, Blaikie et al. 1994)
Education	Education is linked to socioeconomic status in that higher educational attainment affects lifetime earnings, and limited education constrains the ability to understand warning information and access recovery information.	Little education (+) Highly educated (−) (*Source*: Heinz Center 2000b)
Population growth	Counties experiencing rapid growth lack available high-quality housing, and the social services network may not have had time to adjust to increased populations. New migrants may not speak the language and may not be familiar with how to deal with bureaucracies in obtaining relief or recovery information. All these factors increase vulnerability.	Rapid growth (+) (*Sources*: Cutter et al 2000, Heinz Center 2000b, Morrow 1999, Puente 1999)
Health status	The public health literature shows that people with preexisting illnesses may be at risk for death/illness/injury in disaster settings. People with preexisting cardiovascular and respiratory conditions who are exposed to smoke and haze from forest fires may be more at risk for adverse health outcomes; they also may be vulnerable to heart attacks during seismic activity.	Major health problems (+) Minor or no health problems (−) (*Sources*: Parati et al. 2001, Brauer 1999, Brown 1999, Minami et al. 1997)
Medical services	Health care providers, including physicians, nursing homes, and hospitals, are important post-event sources of relief. The lack of proximate medical services lengthens the time needed to obtain short-term relief and achieve longer-term recovery from disasters.	Higher density of medical (−) (*Sources*: Heinz Center 2000b, Morrow 1999, Hewitt 1997)
Social dependence	People who are totally dependent on social services for survival are already economically and socially marginalized and require additional support in the post-disaster period.	High dependence (+) Low dependence (−) (*Sources*: Heinz Center 2000b, Hewitt 2000, Morrow 1999, Drabek 1996)
Special-needs populations	Special-needs populations (infirm, institutionalized, transient, homeless) are difficult to identify, let alone measure and monitor. Yet it is this segment of society that invariably is left out of recovery efforts, largely because of this invisibility in communities.	Large number of special needs (+) Small number of special needs (−) (*Sources*: Morrow 1999, Tobin and Ollenburger 1993)

2.2.3 Aggregation method

As mentioned in section 2.2.1, the most used model for conducting an SVA is the inductive model. Since this thesis utilizes the model mentioned, an elaboration on its aggregation method is needed. This is based on the book Statistics and Data Analysis in Geology by Davis (2002).

Principal component analysis (PCA) is a statistical technique used to analyse and interpret large datasets by reducing their dimensions. A PCA transforms a dataset with multiple variables into a set of new variables called principal components. These components are combinations of the original variables and are generated in a way that maximizes their variance. A PCA works as follows. First, the variance-covariance matrix of the variables is computed. Then, eigenvalues and eigenvectors can be extracted from this. Eigenvalues represent the amount of variance captured by each principal component, while eigenvectors determine the direction or weight of each variable in the principal component. Based on the eigenvalues, the principal components are ranked in order of significance. The first principal component captures the largest amount of variance, followed by the second, third, and so on. By selecting a subset of principal components that explain a significant portion of the total variance, the dimensionality of the dataset can be reduced.

To enhance the interpretability of the factors a PCA can be run with varimax rotation. As Davis (2002) states, “varimax rotation changes the factor loadings so that the original variables have either a high positive or high negative correlation (near ±1) with a factor, or a correlation near zero”. Figure 4 shows this rotation. A common method to select the ultimate components is applying the Kaiser criterion. The Kaiser criterion states that only principal components with eigenvalues greater than 1 should be retained. Retaining components with eigenvalues greater than 1 implies that these components explain more variance than a single original variable would on average.

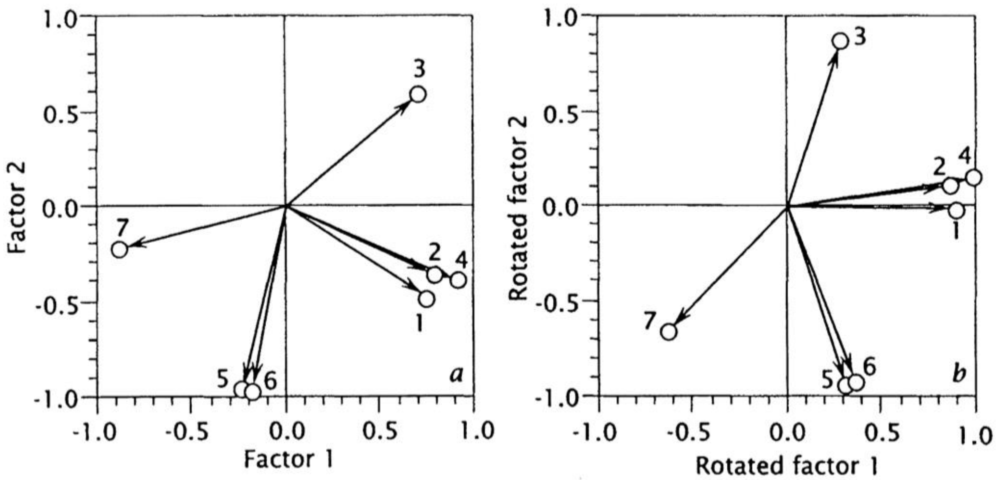

Figure 4: (a) A plot of loadings on two raw factors extracted from measurements on 25 random blocks. (b) A plot of loadings on two factors rotated by the varimax criterion. (Davis, 2002)

3 Methodology

To answer research question three, two case studies have been conducted. This chapter introduces the case studies and explains step by step what data was needed, where it can be found, and how it has been used. This has led to the construction of a social vulnerability index for Haarlem and Zwolle. The resulting index has been compared with places at risk of flooding in case of heavy rainfall.

3.1 Case study approach

3.1.1 Case study selection

This study has executed two case studies to create social vulnerability indices in the municipalities of Haarlem and Zwolle. The case of Haarlem has been chosen based on the interests of the company Witteveen+Bos at which this thesis has been executed. Zwolle has been selected due to the involvement of this thesis with the Delta Futures Lab thematic work group on a climate-resilient Zwolle region. More information on this can be found in section 1.5.2 of this report.

3.1.1.1 Haarlem

Haarlem, a city of around 160,000 inhabitants, is nestled in the province of North Holland. Its strategic location, just 20 kilometres west of Amsterdam, makes it a desirable residential area. The city perfectly balances its historical significance with modern living, boasting boutique shops, cosy cafes, and vibrant events. The annual Bloemencorso flower parade draws crowds. Its proximity to Amsterdam's bustling energy adds to Haarlem's appeal as a residential haven. With its rich cultural scene, artistic heritage, and prime location, Haarlem stands as a captivating example of harmonizing history and contemporary living.

3.1.1.2 Zwolle

Zwolle is a city with a population of around 130,000 residents. Positioned in the Overijssel province, it serves as a central hub with efficient transportation connections, including a significant railway station. The city's distinctive feature is its intricate network of canals that flow through its urban landscape, adding both scenic beauty and practical water management. The IJssel and Vecht rivers further complement this water-centric environment. These aquatic elements not only enhance the city's aesthetics but also contribute to its historical significance, as water has played a pivotal role in Zwolle's development and identity.

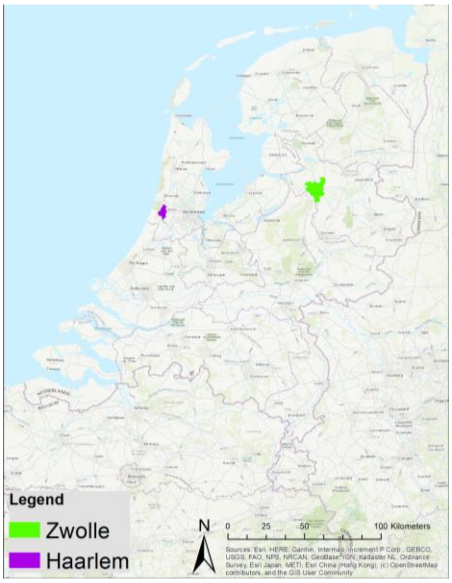

Figure 5: Map of the Netherlands with Haarlem (purple) and Zwolle (green) highlighted

3.1.2 Case study resolution

The resolution of the case study has been partially based on data availability. Nearly all data needed was available at CBS (2022a). Due to privacy, data has been left out in categories to which fewer than five people belong. This is especially a problem when looking at the street scale. Therefore, it has been decided to perform the analysis at the PC5 level, not at PC6. Dutch postcodes have four numbers and two letters and can administratively be divided into three groups. These are PC4, PC5 and PC6, in which PC4 only includes the numbers, PC5 the numbers plus one letter, and PC6 the complete postcode.

3.2 Social vulnerability index construction

Figure 6 shows the steps that have been taken to construct the social vulnerability indices for Haarlem and Zwolle. This flowchart and the methods are based on a paper by Chakraborty et al. (2020). This section of the report will address all steps.

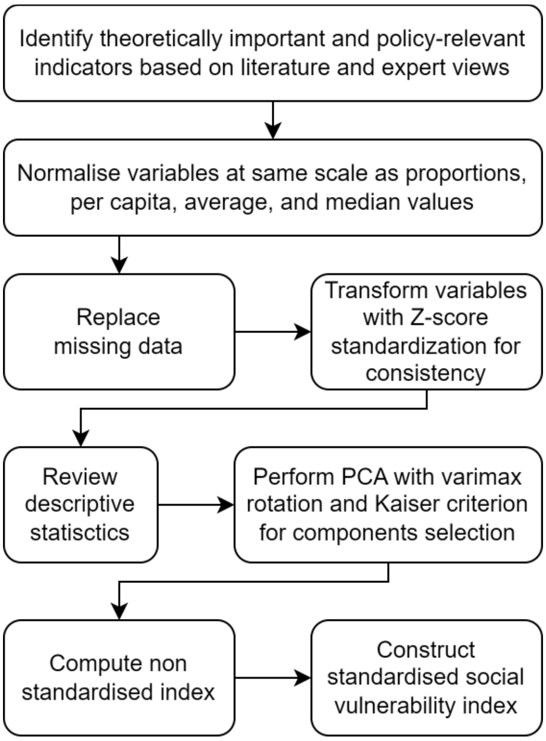

Figure 6: Steps of the Social Vulnerability Index construction

3.2.1 Indicator identification

The indicator approach as developed by Cutter et al. (2003) is, as mentioned earlier in section 1.4, very much focused on the USA. To select and add the right indicators to account for the Western European, and more specific Dutch setting, 5 studies with versions of the indicator list of Cutter et al. (2003) have been put next to each other, see Appendix A – Indicator selection (Fekete, 2009; Forrest et al., 2020; Holand & Lujala, 2013; Koks et al., 2015; Tascón-González et al., 2020). This allows to see the overlap and differences between the indicators. Based on the 5 studies done in Europe, a selection of indicators has been made that fit the Dutch urban setting. This list has been verified by a group of experts from Witteveen+Bos and the University of Twente to indicate not applicable or missing indicators (Bakhshianlamouki, 2023; Klein, 2023; Maas, 2023; Roeleveld, 2023).

The indicator 'Population in need of care' has been suggested by one of the experts to add. People in for example nursing homes or assisted living are almost completely dependent on others, making them very vulnerable. Determining where and how many people live in these conditions in Haarlem and Zwolle was planned to be done using the 'Basisregistratie Adressen en Gebouwen' (BAG) (Kadaster, 2023). Buildings having the combined functions of healthcare and living would fit the description. However, since for example pharmacies, yoga schools, and nail saloons with an apartment on top also have these functions it appeared to be too much work to filter buildings like this. So due to the lack of available data, this indicator has not been taken on for further analysis.

Moreover, all experts spoken to for this thesis have been asked to give the indicators a score from 1 (least important) to 5 (most important). This helps in identifying the parameters that are less important and will improve the outcomes of the research by omitting them. Research shows that differences in indicator selection through a PCA and by expert judgement do not make much difference in places with high or low vulnerability as it does in less extreme areas (Bucherie et al., 2022).

Next, the indicators have been grouped according to population characteristics (also see Table 2). These characteristics are widely acknowledged within the social science community to influence social vulnerability (Cutter et al., 2001; H. John Heinz Iii Center for Science & the, 2002; Holand & Lujala, 2013). By using these groups, it has been more clear what different aspects of vulnerability are taken into account.

3.2.2 Normalise variables

To make the several indicators better comparable, most of them have been transformed into percentages. By using percentages instead of amounts, the number of inhabitants in an area does not weigh in the data of the indicators. Since the number of inhabitants in an area was known, this was an easy transformation. The original units and transformed units can be found in their respective columns in Table 4.

3.2.3 Replace missing data

When data was needed that was not available at the postal code level, data was supplemented from the neighbourhood dataset (CBS, 2022b). PC5 and neighbourhood correspond quite well in terms of scale. With GIS software, the PC5 areas were linked to the neighbourhood to which they have the greatest overlap. Data on the indicator open space was not available through previously mentioned sources. This has been generated based on the (BAG). By subtracting the area covered by buildings from the total area, the amount of open space was approximated (Table 4).

Areas with a population of 0 have been removed from the dataset. Moreover, several indicators have missing data in some areas. This can be because the data is unknown, it is not trustworthy, but in most cases because the data is between 0 and 4. These small values are not allowed to be published by CBS because they could be traced back to individuals. This data has been replaced based on what is suitable per indicator. An overview per indicator can be found in Table 3. When the data was complete, replacing was not applicable.

The birth rate indicator deserves special attention here. The total amount of births in 2020 has been looked up at the municipality (Haarlem, 2023; Zwolle, 2023). The number of births in the dataset has been deducted from this and the remaining have been distributed over areas with missing data based on the population size in that area. However, it is known that missing data is most likely a value between 0 and 4. When many areas have numbers larger than this, the remainder of births would just be distributed equally over all areas with no data available.

Table 3: Missing data handling

Indicators	Missing data replaced with
Median age	N/A
Population under 15 years of age	0
Population over 65 years of age	0
Population with a high level of education	Average surrounding postcodes
Population with a low level of education	Average surrounding postcodes
Birth rate	The remaining amount is distributed over missing areas
Single-parent households	0
Labour force participation	Average surrounding postcodes
Unemployment rate	Average surrounding postcodes
Non-western migrants	Half of (100-%Dutch)
Western migrants	Half of (100-%Dutch)
Renter-occupied houses	0
Houses owned by cooperation	0
Median house value	Median of available data
Construction year of property	N/A
Population density	N/A
Household density	Amount of inhabitants/Average household density/Area
Open space	N/A
Urbanisation	N/A
Social Security recipients	Average surrounding postcodes
Income	N/A
High-income households	Average surrounding postcodes
Percentage living in poverty	Average surrounding postcodes

3.2.4 Z-score transformation

A z-score transformation has been performed on the data because a PCA is sensitive to differences in the units of measurement of variables (Chakraborty et al., 2020). This transformation has been done using Equation 1 and will result in new values that have a mean that is zero and are measured in units of standard deviations (Davis, 2002).

$$z_i = \frac{x_i - \bar{X}}{\sigma} \quad \text{Equation 1}$$

3.2.5 Descriptive statistics

Three methods have been used to test if a PCA is appropriate for the selected variables (Chakraborty et al., 2020). The first one is the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (Kaiser, 1970; Kaiser & Rice, 1974). It determines the degree of common variance among variables, indicating how well the variables are related to each other and whether they can be grouped into underlying factors. The outcome of this test should be above 0.60 but values above 0.8 are considered excellent. A high score indicates that the variables have a high degree of common variance and thus are well-suited for factor analysis.

The second test is Bartlett's Test of Sphericity (Bartlett & Fowler, 1937). Bartlett's test evaluates the null hypothesis that the variables in the dataset are uncorrelated. If the test rejects the null hypothesis, it suggests that there is sufficient correlation among the variables to proceed with factor analysis. The test statistic in Bartlett's test follows a chi-square distribution. A small test statistic indicates that the correlation matrix is close to the identity matrix and is not suitable for factor analysis. On the other hand, a large test statistic indicates that there is enough correlation among the variables for factor analysis. To determine the statistical significance of Bartlett's test, the computed test statistic has been compared with the chi-square distribution critical values at a significance level of 0.01. If the p-value is less than the chosen significance level, the null hypothesis will be rejected, and the factor analysis can be executed.

The third and last test is Cronbach's alpha coefficient (Cronbach, 1951). Cronbach's alpha is a value that quantifies the internal consistency of a set of items designed to measure the same construct. It measures the extent to which these items are interrelated or correlated with each other. Cronbach's alpha ranges from 0 to 1. The higher the value of Cronbach's alpha, the greater the internal consistency among the items. A value close to 1 indicates high internal consistency, suggesting that the items are strongly related to each other, and they are measuring the same underlying construct. An acceptable value for this test is an alpha of above 0.7. Above 0.8 is considered good, and above 0.9 is excellent.

KNO measure and Bartlett's Test of Sphericity have been performed using SPSS software, and Cronbach's alpha coefficient has been determined using STATA software.

3.2.6 PCA

The Principal Component Analysis has been performed in SPSS version 28 using the dimension reduction tool 'factor'. This tool allows to do the KMO and Bartlett's Test. Furthermore, it has varimax rotation and the Kaiser criterion for component selection built-in.

3.2.7 Non-standardised index

To come to a final score for social vulnerability the first step after the PCA was to determine the non-standardised scores per area. The PCA has resulted in several components with eigenvalues greater than one. SPSS can generate the component scores per area. A weighted sum of these factor scores has been used to generate a non-standardised socioeconomic index for area j (NSI_j) as follows:

$$NSI_j = \sum_{i=1}^{n} w_i \cdot PC_i \qquad \textit{Equation 2}$$

where,

$$w_i = \frac{Proportion\ of\ Variance\ for\ Factor}{Total\ Varance\ Explained} \qquad \textit{Equation 3}$$

In this way, the proportion of variance that is explained by a factor towards the total variance explained by the selected components has been the weight of that factor. It is good to mention that determining the weights for a PCA-based composite index analysis lacks a theoretical foundation. (Chakraborty et al., 2020; Cutter & Emrich, 2017).

3.2.8 Social vulnerability index

The last step was to standardize the index to make the scores easily comparable. This has been done using Equation 4.

$$SoVI_{(j)} = \frac{NSI_{(j)} - NSI_{max}}{(NSO_{max} - NSI_{min})} * \ 100 \qquad \textit{Equation 4}$$

3.2.9 PCA post-estimation: goodness-of-fit evaluation

Evaluating the goodness of fit of a factor solution has been done by checking whether the proportion of residuals with scores higher than 0.05 exceeds 50%. Smaller residuals indicate a well-fitting model. The number of residuals with values smaller than -0.05 or greater than 0.05 has been counted in the residual correlation matrix and divided by the total amount of residuals to determine this percentage.

3.3 Locations at risk of flooding

3.3.1 Haarlem

Witteveen+Bos has done model calculations for heavy rainfall in Haarlem. For this, they simulated a rainfall event of 90 mm in one hour, using InfoWorks ICM, according to Leidraad C2100 on hydraulic modelling by Rioned. A rainfall event of 90 mm in one hour corresponds with a frequency of once in about 500 years (Weeren et al., 2018). The calculations include the sewage system, run-off over the street, and any storage at street level. This resulted in an amount of water in the streets after the event, and based on a set sill height it could be determined which buildings will probably have water inside. Based on this information 21 bottleneck locations with clusters of buildings at risk of flood have been identified (Figure 7).

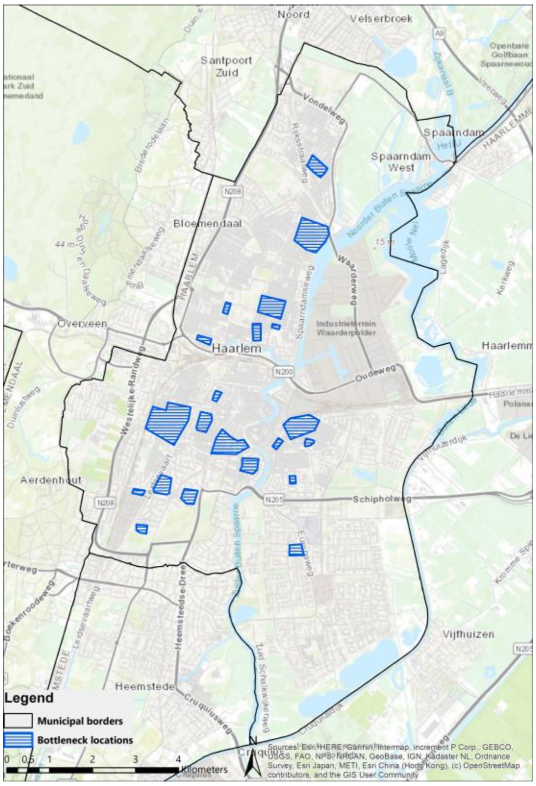

Figure 7: Bottleneck locations in Haarlem

3.3.2 Zwolle

The municipality of Zwolle did the same as Witteveen+Bos did for Haarlem. They have calculated the water level in the streets in case of a rainfall event with 90 mm in one hour. However, they have not done all the calculations for this yet. This means that only for 7 of the 17 neighbourhoods data is available. Since this is too little to do this research, some older data have been used. In 2017, the municipality calculated a rainfall event of 79 mm in one hour. This is still within the statistical margin that it could happen once every 500 years but is more likely to take place every 250 years (Weeren et al., 2018). Also, for Zwolle, based on the amount of water in the streets after the event, and based on the sill height it could be determined which buildings will probably have water inside. Based on this information 21 bottleneck locations with clusters of buildings at risk of flood have been identified (Figure 8). It is important to note that the large rectangle at the right is an industrial park. The relevance of a flood in this area in relation to this thesis is not that large.

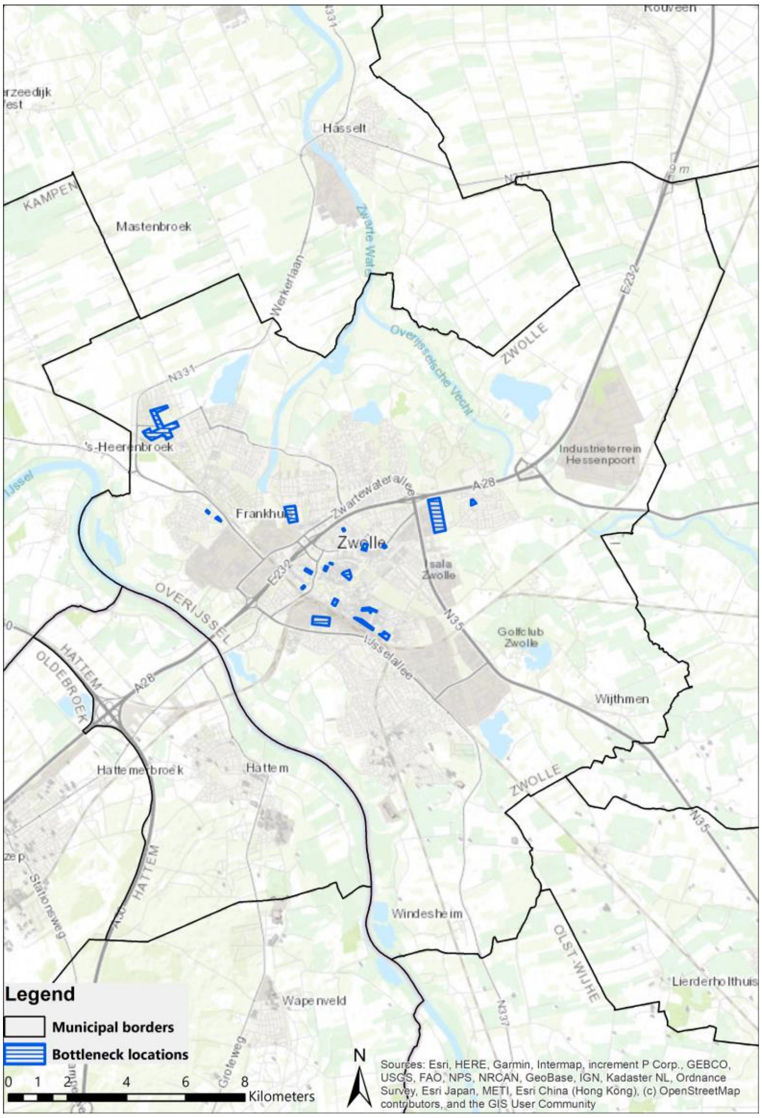

Figure 8: Bottleneck locations in Zwolle

3.4 Cluster analysis

Before combining social vulnerability with flood risk, a cluster analysis has been run over the results of the PCA. A cluster analysis can help in identifying hidden trends and patterns within datasets. It allows to explore the results from another angle and can help interpret the results of this study.

Using ArcGIS software, a Local Moran's I analysis has been run. Local Moran's I calculate spatial autocorrelation statistics for each location in a dataset, generating a map that displays local patterns of clustering. Four types of clusters can be displayed after the analysis:

- HH (High-High): Locations with high attribute values surrounded by locations with high values. This indicates a cluster of high values.
- HL (High-Low): Locations with high attribute values surrounded by locations with low values. This indicates a spatial outlier of high values.
- LH (Low-High): Locations with low attribute values surrounded by locations with high values. This indicates a spatial outlier of low values.
- LL (Low-Low): Locations with low attribute values surrounded by locations with low values. This indicates a cluster of low values.

The Local Moran's I help identify areas with significant local clustering, providing insights into spatial patterns that may not be evident in a global analysis. It is useful for detecting spatial heterogeneity across different regions in the study area.

3.5 Combining social vulnerability with flood risk

To capture the potential social 'risk' of flooding, SVI data has been combined with flood hazard and exposure. This has been done by overlaying the map with social vulnerability scores with the map with bottleneck locations. In this way, a first look can be taken at whether social vulnerability intersects with flood risk and where measures are taken to decrease flood risk.

3.6 Implementation of social vulnerability in flood risk management

After exploring methods to quantify social vulnerability and determining the flood risk in areas within Haarlem and Zwolle, suggestions can be put out on how to incorporate social vulnerability. Since Haarlem and Zwolle have been the only municipalities to which this research looked into their current strategies for flood mitigation, not much can be said specific to the whole Dutch, or even broader context.

4 Results

4.1 Social vulnerability index

4.1.1 PCA input

After going through the indicators listed by (Cutter et al., 2003), the European applications of this list, and the expert views on the selection of indicators, the list in Table 4 has been created to use as input for the Social Vulnerability Index.

Table 4: Indicators, sources, and units

Vulnerability concept	Indicator	Data source	Unit	Transformed unit	Name
Age	Median age	Kerncijfers per postcode	Classes	Median	M_leeftijd
	Population under 15 years of age	Kerncijfers per postcode	Amount	Percentage	P015
	Population over 65 years of age	Kerncijfers per postcode	Amount	Percentage	P65
Education	Population with a high level of education	Wijk en buurtkaart	Amount	Percentage	P_OPL_HG_PC5
	Population with a low level of education	Wijk en buurtkaart	Amount	Percentage	P_OPL_LG_PC5
Family structure	Birth rate	Kerncijfers per postcode	Amount	Amount	A_geboorten
	Single-parent households	Kerncijfers per postcode	Amount	Percentage	P_eenouder
Occupation	Labour force participation	Wijk en buurtkaart	Percentage	Percentage	P_ARB_PP
	Unemployment rate	Wijk en buurtkaart	Amount	Percentage	P_WW
Race and ethnicity	Non western migrants	Kerncijfers per postcode	Percentage	Percentage	P_niet_westerse
	Western migrants	Kerncijfers per postcode	Percentage	Percentage	P_westerse
Renters	Renter-occupied houses	Kerncijfers per postcode	Percentage	Percentage	P_huurwoningen
Residential property	Houses owned by cooperation	Kerncijfers per postcode	Amount	Percentage	P_corporatie woning
	Median house value	Kerncijfers per postcode	Average	Average	G_woz
	Construction year of property	Kerncijfers per postcode	Classes	Median	Median_building_ye ar
Rural/urban	Population density	Kerncijfers per postcode	Amount	Amount/km2	Population_ density
	Household density	Kerncijfers per postcode	Amount	Amount/km2	Household_ density
	Open space	Basisregistratie Adressen en Gebouwen		Percentage	P_Open_space
	Urbanisation	Kerncijfers per postcode	Addresses /km2	Addresses /km2	OAD
Social dependence	Social Security recipients	Kerncijfers per postcode	Amount	Percentage	Pp_met_ uitkering
Socio-economic status	Income	Kerncijfers per postcode	Average	Average	G_inkomen_hh_ 1000
	High-income households	Wijk en buurtkaart	Percentage	Percentage	P_HoogInkP
	Percentage living in poverty	Wijk en buurtkaart	Percentage households	Percentage	P_SocMinH

4.1.2 Descriptive statistics

The descriptive statistics of the dataset before running the PCA were for all three tests within the margins to continue the analysis (Table 5). The KMO measure of sampling adequacy shows the degree of common variance between the indicators. The values are well above the threshold value of 0.6 and nearing 0.8 which would be defined as excellent.

Bartlett's test evaluates the null hypothesis that the variables in the dataset are uncorrelated. To reject this hypothesis, the p-value should be below the threshold. This is the case for both Haarlem and Zwolle, so the hypothesis of uncorrelated variables has been rejected. Correlating is essential to perform a PCA.

Cronbach's alpha coefficient looks at the internal consistency of the datasets. The threshold value of 0.7 has been met in both cases. Values above 0.8 are even considered good.

Table 5: Results of statistical tests before PCA

Tests	Threshold	Haarlem	Zwolle
KMO	>0.60	0.774	0.769
Bartlett's Test of Sphericity	<0.001	0.000	0.000
Cronbach's alpha coefficient	>0.7	0.855	0.867

4.1.3 PCA results

To do the component selection in the PCA a scree plot has been drawn for both cases. This plot shows the eigenvalues of the principal components y-axis and the corresponding factor number on the x-axis. Both plots show six principal components with eigenvalues, explaining the amount of variance per component, of greater than one. Factors with eigenvalues lower than one represent minor variance and thus will not be taken further in the analysis. The six components that are taken on further in the analysis are built up out of several indicators. Dominant population characteristics in these components, as further explained in 2.2.2, are shown in Table 6. When a population characteristic appears twice, the indicators that fall under it have an impact on more than one component. A component with two dominant population characteristics has several indicators of both these population characteristics.

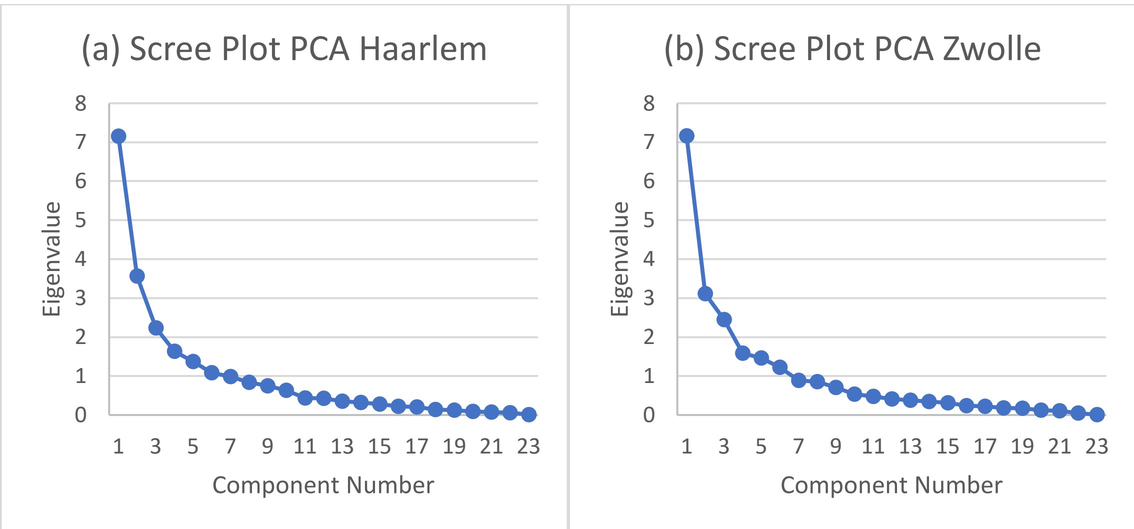

Figure 9: Scree plots of eigenvalues after PCA for (a) Haarlem and (b) Zwolle

Table 6: Dominant population characteristics per component for Haarlem and Zwolle

	Haarlem	Zwolle
Component 1	Residential property Socioeconomic status	Residential property Race and ethnicity
Component 2	Age	Socioeconomic status
Component 3	Rural/urban	Age
Component 4	Rural/urban	Rural/urban
Component 5	Family structure Race and ethnicity	Rural/urban
Component 6	Education	Residential property Family structure

4.1.4 PCA post-estimation: goodness-of-fit evaluation

The proportion of residuals with absolute values greater than 0.05 in the residual correlation matrix showed percentages of 30.83 and 29.35 for respectively Haarlem and Zwolle. This is well within the margin of 50%.

4.2 Expert knowledge

There are some similarities as well as some differences in what experts find important when looking at social vulnerability and where the PCA finds the most variance. Dominant population characteristics have already been shown in Table 6. Table 7 shows the average scores experts gave to the indicators.

Education is considered quite important by experts. Surprisingly, in Haarlem, education lacks significance. For age, this is the other way around. It is underestimated by experts and emerges as an important factor in social vulnerability according to the PCA findings.

Race and ethnicity are not regarded as crucial elements influencing vulnerability by experts, and they score low in the PCA of Haarlem accordingly. In Zwolle however, the PCA shows the race and ethnicity scores in the first component. This highlights regional differences.

Socioeconomic status consistently holds medium to high importance in both expert evaluations and PCA results for Haarlem and Zwolle, other than the rural/urban indicators. While perceived as relatively average by experts, they dominate several PCA components. This indicates the need for further exploration of this factor's influence.

The direction of influence has not been discussed with the experts. It can be assumed that, since they are experts in this field, they for example understand that a high average house value would have a negative impact on social vulnerability.

Table 7: Average of scores given by 4 experts on selected indicators to determine social vulnerability - 1 not very important - 5 very important

Vulnerability concept (Cutter)	Indicator	Average score
Age	Median age	1.3
Race and ethnicity	Western migrants	1.8
Age	Population under 15 years of age	2.3
Renters	Renter-occupied houses	2.3
Rural/urban	Household density	2.3
Age	Population over 65 years of age	2.5
Family structure	Birth rate	2.5
Race and ethnicity	Non-western migrants	2.5
Residential property	Houses owned by cooperation	2.5
Rural/urban	Population density	2.5
Social dependence	Social Security recipients	2.5
Rural/urban	Open space	2.8
Socioeconomic status	High-income households	2.8
Residential property	Median house value	3.0
Occupation	Unemployment rate	3.3
Residential property	Construction year of property	3.3
Rural/urban	Urbanisation	3.3
Socioeconomic status	Median income	3.3
Education	Population with a high level of education	3.5
Education	Population with a low level of education	3.5
Family structure	Single-parent households	3.5
Occupation	Labour force participataion	3.5
Socioeconomic status	Percentage living in poverty	3.5
Medical services	Population in need of care	4.0

NB: The indicator on Population in need or care has later been removed due to lack of data, see 3.2.1.

4.3 Social vulnerability index and bottleneck maps

The figures below show the social vulnerability indices of Haarlem (Figure 10) and Zwolle (Figure 11) overlayed with the bottleneck locations for floods. In Haarlem, the bottlenecks overlap more with the less vulnerable areas (10 to 4, see Table 9).

In Zwolle, the situation is different. The more rural areas in the municipality are not that vulnerable and do not show bottleneck locations. The more vulnerable areas as well as the bottleneck locations are located in the residential areas and thus show an overlap. This is not strange, because bottleneck locations are locations at which multiple houses will flood in case of heavy rainfall. In the more rural areas, there are fewer houses, there is more water storage at street level and the amount of impermeable surface is less. These are important factors for the infiltration capacity. Pluvial floods are thus properly often called urban floods.

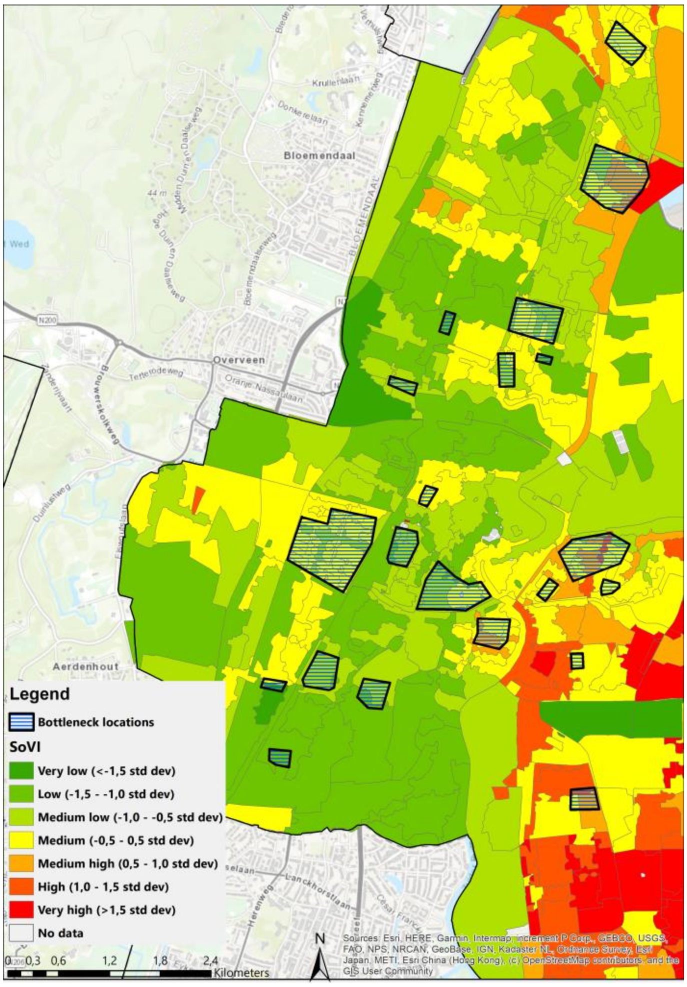

Figure 10: Zoom of the social vulnerability index for Haarlem overlayed with bottleneck locations for floods

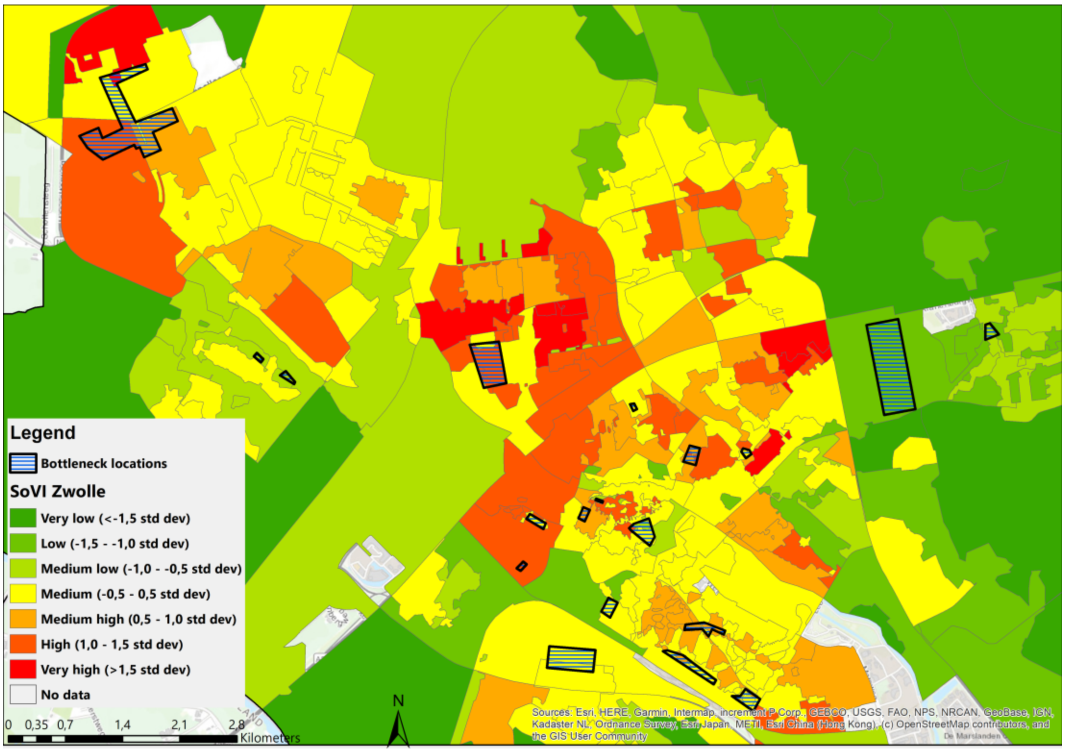

Figure 11: Zoom of the social vulnerability index for Zwolle overlayed with bottleneck locations for floods

4.3.1 Cluster analysis

The outcomes of the Local Moran's I cluster analysis show large clusters of high as well as low vulnerability, both in Haarlem (Figure 12) and Zwolle (Figure 13). The dark blue areas are areas with low vulnerability, surrounded by areas with low vulnerability. Light blue areas represent areas with low vulnerability but are surrounded by areas with high vulnerability. Light red areas are vulnerable but surrounded by areas that are not. Dark red areas are vulnerable areas surrounded by vulnerable areas.

About one-third of the Haarlem map shows a cluster of low vulnerability in the west. On the edges of this cluster, some high-vulnerability areas are located. The southeast shows a big cluster of high vulnerability, surrounded by less vulnerable areas. In Zwolle, the vulnerable areas are clustered in the centre of the map. In this map, it is also clear that the large rural areas are less vulnerable. Interesting to see is that quite some built-up area is not significantly clustered.

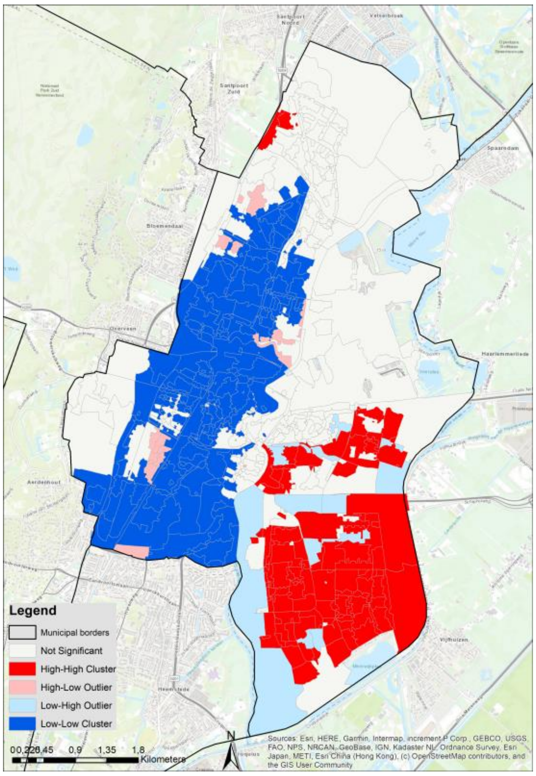

Figure 12: Cluster analysis for social vulnerability in Haarlem. H-H (dark red) - High social vulnerability surrounded by high social vulnerability H-L (light red) - High social vulnerability surrounded by low social vulnerability L-H (light blue) - Low social vulnerability surrounded by high social vulnerability L-L (dark blue) - Low social vulnerability surrounded by low social vulnerability

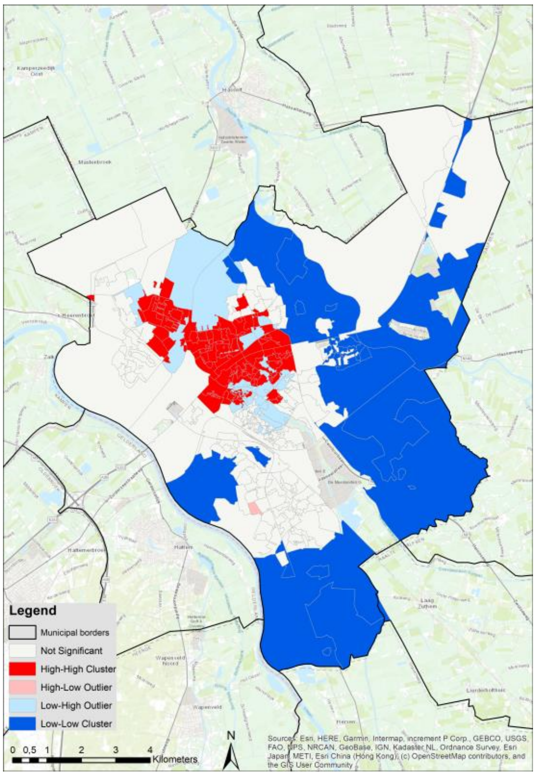

Figure 13: Cluster analysis for social vulnerability in Zwolle. H-H (dark red) - High social vulnerability surrounded by high social vulnerability H-L (light red) - High social vulnerability surrounded by low social vulnerability L-H (light blue) - Low social vulnerability surrounded by high social vulnerability L-L (dark blue) - Low social vulnerability surrounded by low social vulnerability

4.3.2 Indicators Haarlem

Next to the social vulnerability indices, some indicators were mapped separately as well. The first one is the indicator percentage of inhabitants with a high income (Figure 14). This was determined by looking at the share of persons in private households belonging to the national 20% with the highest personal income. This indicator was in the first principal component and is inversely correlated to the social vulnerability index. The map shows, in inverse, some great similarities to the SoVI map but is not entirely the same. Although the extremes overlap, this is not the case for the map showing the areas with inhabitants that are low educated (Figure 15). Low education was in the sixth component and thus not very impactful in the PCA. Areas with more moderate percentages compare less well to the final SoVI map. Some other indicators that have been put on a map can be found in Appendix E – Maps of indicators.

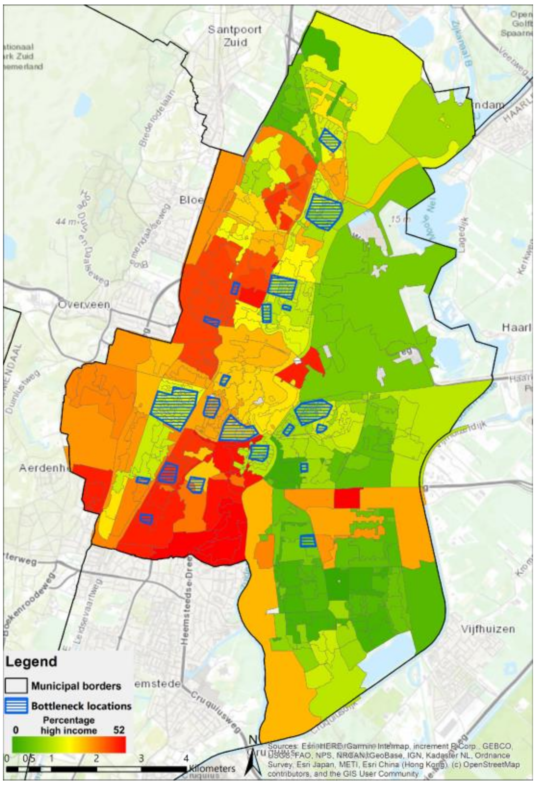

Figure 14: Percentage of people with a high income in Haarlem

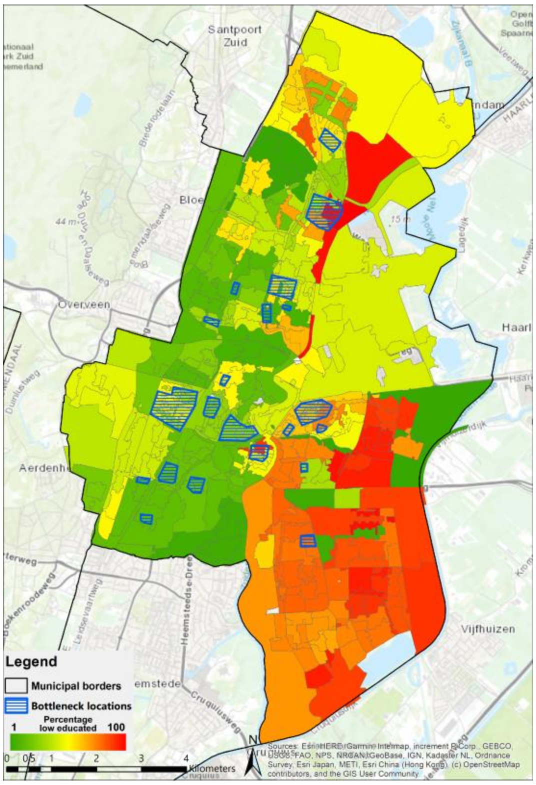

Figure 15: Percentage of people with a low level of education in Haarlem

4.3.3 Indicators Zwolle

Looking at the indicator percentage of rental houses in Zwolle (Figure 16) the big difference between the residential and rural areas becomes visible. This indicator belongs to the first principal component. Nearly all areas with a high percentage of rental houses are relatively socially vulnerable. This is quite different when looking at the birth rate (Figure 17). This indicator scored low in the PCA and the similarities with the social vulnerability index are not apparent. The birth rate is spread very homogeneous over the total municipality of Zwolle, except for the neighbourhood of Stadshagen in the northwest.

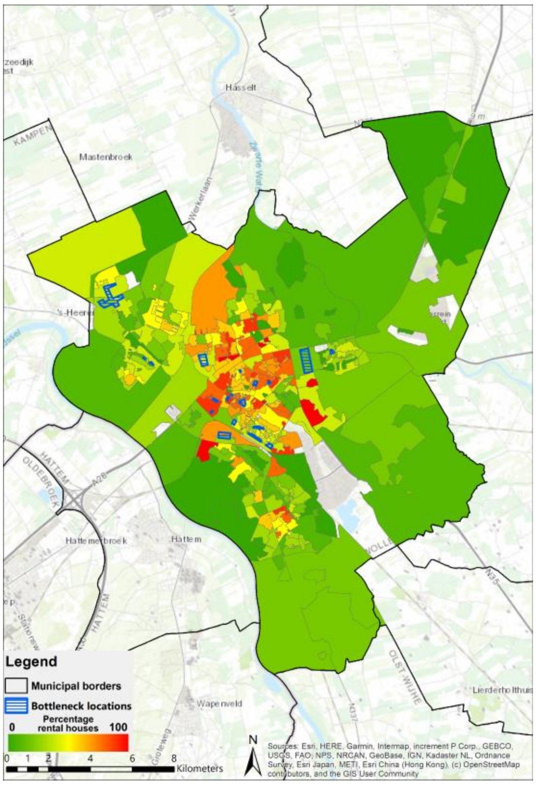

Figure 16: Percentage of rental houses in Zwolle

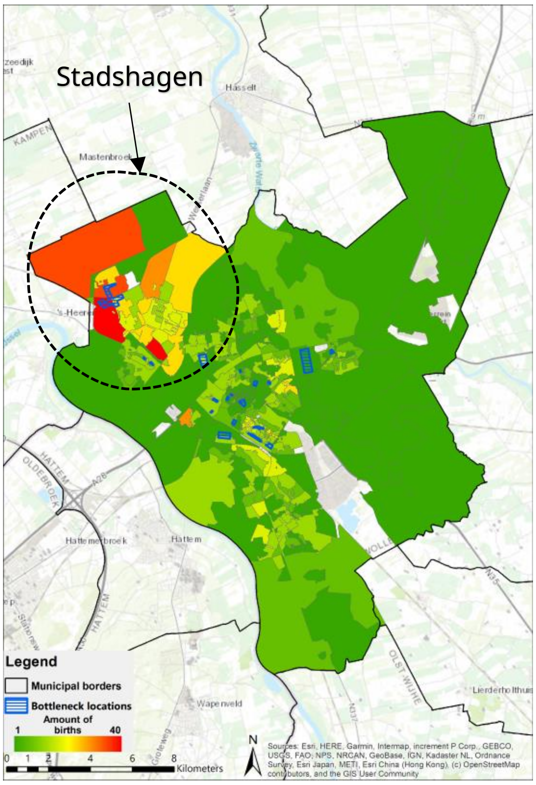

Figure 17: Number of births in Zwolle

4.3.4 Number of inhabitants

The indicators population density, household density, and urbanisation say something about the number of people or households in an area, and those indicators are included in the PCA. However, to look at the specific number of people affected by a pluvial flood is interesting to see those numbers mapped as well (Figure 18 and Figure 19).

In these maps, bottleneck areas appear quite well inhabited and thus a lot of people will experience hindrance from a pluvial flood. The only exception is the industrial area in the east of Zwolle, in which not many people live. When looking at the map of Zwolle it is also apparent that the rural areas have way fewer inhabitants than the residential areas. This is logical but remains relevant when looking at the social vulnerability index.

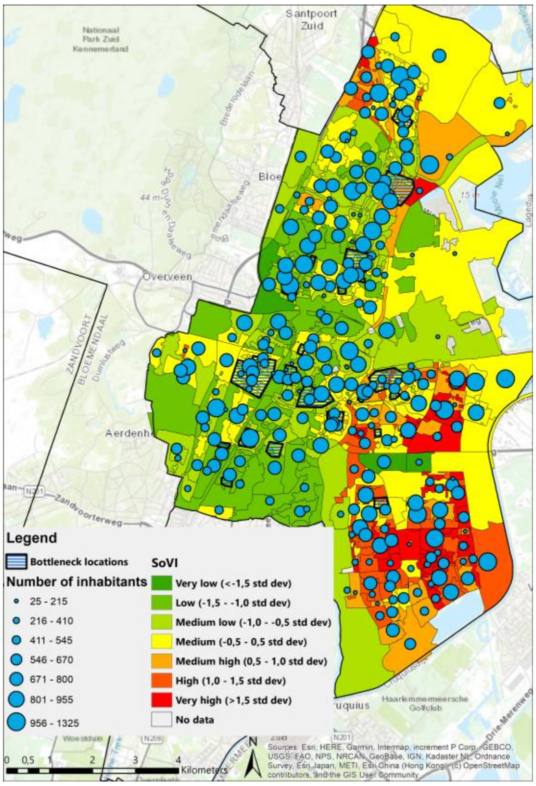

Figure 18: SoVI, bottlenecks and number of inhabitants in Haarlem

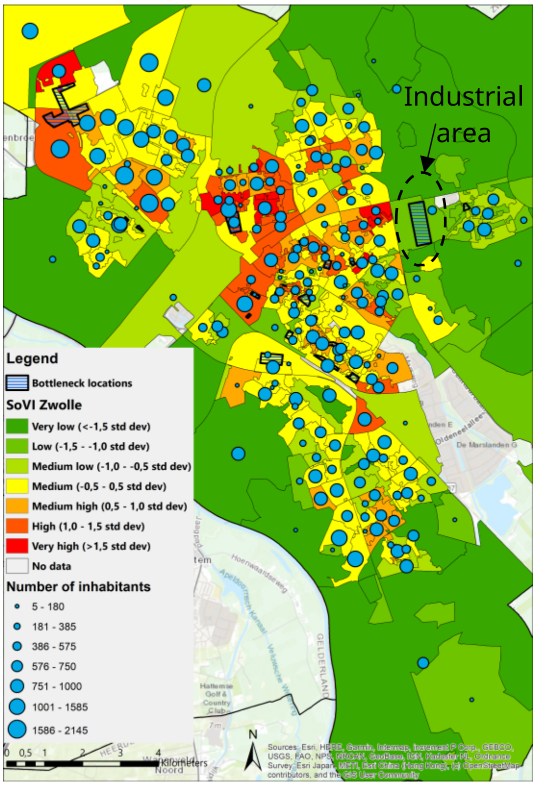

Figure 19: SoVI, bottlenecks and number of inhabitants in Zwolle

Diving a bit deeper into the inhabitant numbers, Table 8 shows the number of inhabitants per social vulnerability class for Haarlem as well as Zwolle. This table shows that in Haarlem the areas with high vulnerability house fewer people (64,170) than the areas with low vulnerability (45,440). In Zwolle, this is the other way around (27,270 versus 44,535).

Table 8: Number of inhabitants per social vulnerability class

Social vulnerability class	Haarlem	Zwolle
Very low (< -1.5 std dev)	2,410	5,050
Low (-1.5 – -1.0 std dev)	22,895	6,080
Medium low (-1.0 – -0.5 std dev)	38,865	16,140
Medium (-0.5 – 0.5 std dev)	53,240	56,875
Medium high (0.5 – 1.0 std dev)	15,395	22,230
High (1.0 – 1.5 std dev)	15,845	17,830
Very high (> 1.5 std dev)	14,200	4,475

Estimating how many people would be affected by a pluvial flood and in which social vulnerability class they fall has been done by assigning the bottleneck locations to the PC5 area they most overlap with. This leads to the results presented in Table 9. Although these are rough estimates, the size of the bottleneck locations does not correspond in any case to the PC5 areas, they are in line with other results.

Table 9: Number of bottlenecks and number of inhabitants in those bottlenecks per social vulnerability class

Social vulnerability class	Haarlem		Zwolle	
	Bottlenecks	Inhabitants	Bottlenecks	Inhabitants
Very low (< -1.5 std dev)	1	700	0	0
Low (-1.5 – -1.0 std dev)	3	2200	1	420
Medium low (-1.0 – -0.5 std dev)	6	4450	2	1190
Medium (-0.5 – 0.5 std dev)	7	5150	9	4825
Medium high (0.5 – 1.0 std dev)	4	2850	3	1625
High (1.0 – 1.5 std dev)	0	0	3	3935
Very high (> 1.5 std dev)	0	0	0	0

In Haarlem, the bottleneck locations overlap more with the less vulnerable than the vulnerable areas. This results in more less vulnerable people being affected by a pluvial flood than more vulnerable people. For Zwolle, it was already visible that the vulnerable people and the bottleneck locations are concentrated in the urban areas. In Table 9 this is reflected in very few bottlenecks that overlap with the low vulnerability classes. For both cities, it is thus the case that the larger group in terms of social vulnerability also has the largest burden in case of a pluvial flood.

5 Discussion

5.1 Interpretation of the results

Social vulnerability is loosely defined, and many visions exist of what it entails. This is one of the reasons that the results of this research are not a number that is good or bad, or a clear solution to a concrete problem. Based on literature, the introductory chapter of this report explained social vulnerability as all factors specifically related to the interactions of hazards with individuals, populations, and communities. This thesis has quantified social vulnerability for Haarlem and Zwolle. The results of this quantification and what can be done with this will be discussed in this chapter.

After the construction of the Social Vulnerability Indices for Haarlem and Zwolle some things become apparent. The results of Haarlem show clear differences between neighbourhoods. The southeast and north seem relatively more vulnerable than other areas in the municipality. The cluster analysis finds the clusters of high vulnerability in the southeast where factors influencing social vulnerability accumulate, and low vulnerability in the middle west. In Zwolle, the built-up areas appear more vulnerable than the rural areas. Clusters of high vulnerability are found in the built-up areas whilst clusters of low vulnerability are found in the rural areas.

These results in Zwolle show that it is hard to compare areas that are not similar to each other in terms of the density of the built environment. This difference lies in indicators like population and household density, urbanisation, open space, rental houses, and house values. Because of the relatively big differences between the urban and rural areas, the nuances in the more vulnerable areas are less visible. Something that is clearer in Haarlem.

The finding that socially vulnerable people tend to live in urban rather than rural areas has been reported before during a social vulnerability analysis. Fekete (2009) found that urban residents are a higher-risk group for flood events. They have to evacuate more often than the population living in rural areas. This underlines the need for location-specific measures, both for flood protection but also to tackle social vulnerability.

Zwolle shows that bottleneck locations overlap with relatively vulnerable areas because they both are located in densely built-up areas. In Haarlem, almost all areas are built up. This provides more opportunities to look at the distribution of bottleneck locations over the several vulnerability classes. 4 of the 21 bottleneck locations overlap with areas with more vulnerable inhabitants while 10 out of 21 overlap with areas with less vulnerable inhabitants. The remainder overlaps with not significantly more or less vulnerable inhabitants.

The starting point to execute this thesis was because Deltares wants to create a worldwide toolbox on flood-related issues and ways to solve them. In this context, social vulnerability came up. After conducting this study, the question arises whether it is possible to standardise SV across regions, or even continents. In Cutter's work (Cutter et al., 2001; Cutter et al., 2003), race and ethnicity are important indicators of social vulnerability. This is merely the case in this study. Although it is present in the first principal component in the analysis for Zwolle, it does not dominate that component solely. For Haarlem, it only appears in the fifth and experts do not judge it as very important as well. In this light, it seems almost impossible to standardise a method to determine relative social vulnerability within areas regional, let alone worldwide. A vast simplification of the indicator list would be an option, but this limits the depth of the study enormously. All this is obviously also subject to data availability.

5.2 Limitations of the study

To execute this study, some assumptions and simplifications have been made. These will be discussed in this section.

To start with, the indicator of assisted living has been left out. This has been done because data on the number of people living in assisted living was not available. According to Schmidtlein et al. (2008), "the SoVI algorithm seems fairly robust to minor changes in variable selection". This would thus mean that not having included this one indicator does not make a huge difference, which seems plausible since by using the PCA the final score is some sort of aggregation of the individual indicators.

Schmidtlein et al. (2008) also concluded that the resolution at which the SoVI algorithm is applied, especially when downscaling, does not change the outcome of the index much. With this, the topic of resolution is brought up. The selected resolution of this study was PC5, mainly due to data availability. To use data that was not available at the PC5 resolution some neighbourhood-level data has been interpolated to the PC5 resolution to make it comparable. For the educational indicators, this meant that the total amount of people with a high or low level of education has been spread over the PC5 areas according to the population per area. For indicators like the percentage of unemployment and the percentage of high-income households, this meant all PC5 areas within that neighbourhood got the same value. When running a PCA and looking for variance this is of course not ideal.

This is immediately the next point of attention. The output of the PCA is dependent on the input. This might seem an open door, but it is important to mention. There is a bias in the results of this study by the limitation of data. Data that is not included is not part of the outcome. As stated above, removing, or adding one indicator would not have changed the results drastically. However, this goes only for minor changes. Removing or adding multiple indicators will have effects on the social vulnerability index.

More on the PCA, the weighting method used to aggregate the six components into one score is debatable. With no substantial information on what is important in assessing social vulnerability, it was chosen to weigh the selected components based on the variance they explain. There were the insights of the experts on what they judge as important and at this moment a dilemma appeared whether to use that information for the weighting. These insights do not take into account possible homogeneity in the data which is included in the variance of the components. Using their importance ranking as a basis for a weighting method may thus lead to a more homogeneous outcome of the Social Vulnerability Index. This has been the main reason to go with the variance of each component as a weight.

How well the PCA performed could have been tested more extensively. Removing one variable and looking at the descriptive statistics and outcomes could have gained some more insights into the influence a certain indicator has on the outcome. The weights could also have been altered to determine whether some aspects are taken into account too much or too little.

On top of this, the usage of data from 2020 has an impact on the results as well (Kuhlicke et al., 2011). All types of indicators can vary quite much over the years. Think of the number of births and the percentage of the population aged 65 and over. 2020 has been chosen because this was the most recent year with a complete dataset. The results, however, might not reflect the current situation.

6 Conclusion

This thesis looked into the theory behind social vulnerability and the Social Vulnerability Index before constructing two indices for Haarlem and Zwolle. These show that differences in social vulnerability exist within both municipalities. For each municipality, two population characteristics are found to be the dominant drivers of social vulnerability, related to the local context: for Haarlem, these are residential property and socioeconomic status. For Zwolle, these are residential property, and race and ethnicity. This chapter will look back at the research questions and provide answers to them.

What is social vulnerability in relation to floods?

Social vulnerability is seen in close relation to a low socioeconomic status. Methods to assess social vulnerability, therefore, use many indicators related to socioeconomic status. However, also indicators like ethnic background, disability, and age are used. The addition of these indicators makes social vulnerability something different than socioeconomic status.

Social vulnerability does not change when looking at different natural hazard types. It is the same for floods as it is for earthquakes or droughts. So, in general, the input in a social vulnerability analysis is relevant to floods, but so it is to other hazards. It is about the capacity of people to deal with a hazard.

How can the existing methods to quantify social vulnerability be applied in flood risk management?

The use of the Social Vulnerability Index is a good way to determine the relative degree of vulnerability in certain areas. It can be applied at different scales and the data needed is publicly available. Another strength of this method is the geospatial presentation which makes the outcomes easily comparable and easy to understand. Moreover, the Social Vulnerability Index is not tailored to floods or any natural hazard specifically. This is a strength of this method because it really looks at the social aspects of vulnerability. In the meantime, taking these social aspects into account is different than the more common hazard-specific vulnerability. Next to that, some weaknesses of this method lie in the limitation of data and the assumptions that need to be taken to get to a result.

The Social Vulnerability Index makes it possible to quantify social vulnerability. This is an important step in applying the concept of social vulnerability in flood risk management. It can help gain insights into social vulnerability levels and identify stakeholders. The geospatial presentation ensures a good combination with flood risk maps.

How can Haarlem and Zwolle incorporate social vulnerability into their climate adaptation strategy?

At this moment pluvial flood adaptation has no priority in municipal climate adaptation policies at either the municipality of Haarlem or Zwolle. When action is taken, this is not solely from the rainwater perspective. According to the municipalities, there always should be multiple reasons to open up the street or to change the layout of an area. Think of a combination of the renewal of sewage pipes, maintenance of gas pipelines, and increasing drainage capacity. This prioritisation method could not be tested on whether or not socially vulnerable people benefit from this, due to a lack of data.

In providing infrastructure services, and for this research flood protection infrastructure, in particular, municipal policies treat everyone equally. Each area within a municipality has to comply with the same rules on water drainage capacities through sewage pipes, and storage in the street. Minima are met everywhere but in Zwolle, new neighbourhoods are given a bit more capacity to adapt to more extreme weather conditions and to create space for excess water coming from surrounding neighbourhoods.

As a recommendation for implementation, it might be a bridge too far to say more socially vulnerable areas should have a different level of protection, but the degree of social vulnerability could be taken into account. The moment climate adaptation does become a reason to start a project, a prioritisation of locations could take into account the degree of social vulnerability in an area. More vulnerable people are less able to prepare for, adapt to, cope with, and recover from hazardous events. It is a choice that has to be made whether certain groups in society should be helped first or more than others, but that is a political one, not one for science.

7 Recommendations

After concluding this thesis some recommendations to improve further research on this topic are presented in this last chapter.

First of all, to improve research in the Dutch setting, the indicator list could be researched much more extensively. This list forms the basis of the analysis and things that are not on the list are not included in the outcome, even if the contribution is minimal.

Secondly, the weighting method used in the PCA might not be appropriate in this study. There is no basis to weigh the selected components based on their explaining variance. In the meantime, there is no scientific base to choose equal weights or to do something else. Further research could be conducted to determine what weighting method is best for a PCA and possibly a PCA for a SoVI in particular.

Thirdly, after completing the Social Vulnerability Index, it should be validated in some way. Exploring which neighbourhoods, the government classifies as vulnerable and comparing them with the outcomes of the Social Vulnerability Index could be a first step in this. Other methods of validation could be researched as well.

Fourthly, to make this research complete, it would have been beneficial to see what the effects are of the current prioritisation methods to work on bottleneck locations. For both Haarlem and Zwolle, this relies on an integral approach that waits for opportunities to combine work. By looking into the locations at which work has been done in the recent past to decrease the effects of heavy rainfall, insights could be gained on the distribution of measures taken over the vulnerability classes.

Fifthly, the results of the Social Vulnerability Index are a bit patchy, especially in Zwolle. When combining a social vulnerability analysis with pluvial flood risk, it would be better to only determine the degree of vulnerability in the areas that are at risk. A further reduction of the study area allows to better see the differences between the areas that are at risk.

Sixthly, social vulnerability can be incorporated into flood risk management in many more terrains than just pluvial flood risk not only in the Netherlands but worldwide. Further research could be done on the implementation of that. From targeted awareness campaigns to complete integration in policy-making (El-Zein et al., 2021).

Finally, there are some other methods to assess social vulnerability. These include intersectionality theory (Boesler, 2022) and utility functions (Kind, 2019). It would be interesting to see the differences or similarities between the outcomes of these methods and the SoVI method. Especially when conducting the same case study.

Bibliography

Adger, W. N. (2006, 2006/08/01/). Vulnerability. *Global Environmental Change, 16*(3), 268-281. https://doi.org/10.1016/j.gloenvcha.2006.02.006

Åström, H., Friis Hansen, P., Garré, L., & Arnbjerg-Nielsen, K. (2014). An influence diagram for urban flood risk assessment through pluvial flood hazards under non-stationary conditions. *Journal of Water and Climate Change, 5*(3), 276-286. https://doi.org/10.2166/wcc.2014.103

Bakhshianlamouki, E. (2023). Personal Communication

Bartlett, M. S., & Fowler, R. H. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences, 160*(901), 268-282. https://doi.org/10.1098/rspa.1937.0109

Birkmann, J., Cardona, O. D., Carreño, M. L., Barbat, A. H., Pelling, M., Schneiderbauer, S., Kienberger, S., Keiler, M., Alexander, D., Zeil, P., & Welle, T. (2013, 2013/06/01). Framing vulnerability, risk and societal responses: the MOVE framework. *Natural Hazards, 67*(2), 193-211. https://doi.org/10.1007/s11069-013-0558-5

Blaikie, P. M. (1994). *At risk : natural hazards, people's vulnerability, and disasters*. Routledge. http://link.library.utoronto.ca/eir/EIRdetail.cfm?Resources__ID=1053430&T=F

Boesler, P. (2022). *Applying Intersectionality Theory to Pluvial Flood Risk - An empirical analysis of social vulnerability to urban pluvial flooding in Europe* Utrecht University]. Utrecht. https://studenttheses.uu.nl/handle/20.500.12932/42779

Bohle, H. G., Downing, T. E., & Watts, M. J. (1994, 1994/03/01/). Climate change and social vulnerability: Toward a sociology and geography of food insecurity. *Global Environmental Change, 4*(1), 37-48. https://doi.org/https://doi.org/10.1016/0959-3780(94)90020-5

Bucherie, A., Hultquist, C., Adamo, S., Neely, C., Ayala, F., Bazo, J., & Kruczkiewicz, A. (2022, 2022/04/15/). A comparison of social vulnerability indices specific to flooding in Ecuador: principal component analysis (PCA) and expert knowledge. *International Journal of Disaster Risk Reduction, 73*, 102897. https://doi.org/https://doi.org/10.1016/j.ijdrr.2022.102897

Bulti, D. T., & Abebe, B. G. (2020, 2020/09/01). A review of flood modeling methods for urban pluvial flood application. *Modeling Earth Systems and Environment, 6*(3), 1293-1302. https://doi.org/10.1007/s40808-020-00803-z

CBS. (2022a). Kerncijfers per postcode. https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/gegevens-per-postcode

CBS. (2022b). *Wijk- en buurtkaart 2022*. https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/wijk-en-buurtkaart-2022

Chakraborty, L., Rus, H., Henstra, D., Thistlethwaite, J., & Scott, D. (2020, 2020/02/01/). A place-based socioeconomic status index: Measuring social vulnerability to flood hazards in the context of environmental justice. *International Journal of Disaster Risk Reduction, 43*, 101394. https://doi.org/10.1016/j.ijdrr.2019.101394

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334. https://doi.org/10.1007/BF02310555

Cutter, S. L. (1996, 1996/12/01). Vulnerability to environmental hazards. *Progress in Human Geography, 20*(4), 529-539. https://doi.org/10.1177/030913259602000407

Cutter, S. L., Boruff, B., & Shirley, W. L. (2001). Indicators of social vulnerability to hazards. *Unpublished paper. Columbia, SC: University of South Carolina, Hazards Research Lab*.

Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2003, 2003/06/01). Social Vulnerability to Environmental Hazards*. *Social Science Quarterly, 84*(2), 242-261. https://doi.org/10.1111/1540-6237.8402002

Cutter, S. L., & Emrich, C. T. (2017). Social Vulnerability Index (SoVI®): Methodology and Limitations. https://www.vulnerabilitymap.org/Portals/0/Files/Social%20Vulnerability%20Index%20SoVI%20-%20Methodology%20and%20Limitations.pdf?ver=Dtf6UG34OrlIsrehg3sjLQ%3D%3D

Cutter, S. L., & Finch, C. (2008, 2008/02/19). Temporal and spatial changes in social vulnerability to natural hazards. *Proceedings of the National Academy of Sciences, 105*(7), 2301-2306. https://doi.org/10.1073/pnas.0710375105

Cutter, S. L., Mitchell, J. T., & Scott, M. S. (2000, 2000/12/01). Revealing the Vulnerability of People and Places: A Case Study of Georgetown County, South Carolina. *Annals of the Association of American Geographers, 90*(4), 713-737. https://doi.org/10.1111/0004-5608.00219

Davis, J. C. (2002). *Statistics and data analysis in geology* (3rd ed. ed.). J. Wiley. http://catdir.loc.gov/catdir/description/wiley041/2004270387.html

Deltares. (2018). *Overstromingsrisico's in Nederland*. https://www.helpdeskwater.nl/publish/pages/132174/overstromingsrisicos_in_nederland_2018_1.pdf

El-Zein, A., Ahmed, T., & Tonmoy, F. (2021, 2021/02/01/). Geophysical and social vulnerability to floods at municipal scale under climate change: The case of an inner-city suburb of Sydney. *Ecological indicators, 121*, 106988. https://doi.org/10.1016/j.ecolind.2020.106988

Elboshy, B., Kanae, S., Gamaleldin, M., Ayad, H., Osaragi, T., & Elbarki, W. (2019, 2019/03/01). A framework for pluvial flood risk assessment in Alexandria considering the coping capacity. *Environment Systems and Decisions, 39*(1), 77-94. https://doi.org/10.1007/s10669-018-9684-7

Fekete, A. (2009). Validation of a social vulnerability index in context to river-floods in Germany. *Nat. Hazards Earth Syst. Sci., 9*(2), 393-403. https://doi.org/10.5194/nhess-9-393-2009

Flanagan, B. E., Gregory, E. W., Hallisey, E. J., Heitgerd, J. L., & Lewis, B. (2011). A Social Vulnerability Index for Disaster Management. *Journal of Homeland Security and Emergency Management, 8*(1). https://doi.org/10.2202/1547-7355.1792 (Journal of Homeland Security and Emergency Management)

Forrest, S. A., Trell, E.-M., & Woltjer, J. (2020, 2020/10/01/). Socio-spatial inequalities in flood resilience: Rainfall flooding in the city of Arnhem. *Cities, 105*, 102843. https://doi.org/10.1016/j.cities.2020.102843

Fritsch, K., Assmann, A., & Tyrna, B. (2016, //). Long-term experiences with pluvial flood risk management [10.1051/e3sconf/20160704017]. *E3S Web Conf., 7*. https://doi.org/ARTN 04017
10.1051/e3sconf/20160704017

Gabor, T., & Griffith, T. K. (1980, 1980/10/02/). The assessment of community vulnerability to acute hazardous materials incidents. *Journal of Hazardous Materials, 3*(4), 323-333. https://doi.org/10.1016/0304-3894(80)80004-5

H. John Heinz Iii Center for Science, E., & the, E. (2002). *Human links to coastal disasters*. H. John Heinz III Center for Science, Economics, and the Environment. https://research.fit.edu/media/site-specific/researchfitedu/coast-climate-adaptation-library/united-states/national/us---heinz-center-amp-nwf-reports/Heinz-Ctr.-2002.-US-Human-Links-to-Coastal-Disasters.pdf

Haarlem, G. (2023). *Haarlm in cijfers*. haarlem.incijfers.nl

Haghighatafshar, S., Becker, P., Moddemeyer, S., Persson, A., Sörensen, J., Aspegren, H., & Jönsson, K. (2020, 2020/10/01/). Paradigm shift in engineering of pluvial floods: From historical recurrence intervals to risk-based design for an uncertain future. *Sustainable Cities and Society, 61*, 102317. https://doi.org/10.1016/j.scs.2020.102317

Hauger, M. B., Mouchel, J. M., & Mikkelsen, P. S. (2006). Indicators of hazard, vulnerability and risk in urban drainage. *Water Science and Technology, 54*(6-7), 441-450. https://doi.org/10.2166/wst.2006.622

Holand, I. S., & Lujala, P. (2013, 2013/05/01). Replicating and Adapting an Index of Social Vulnerability to a New Context: A Comparison Study for Norway. *The Professional Geographer, 65*(2), 312-328. https://doi.org/10.1080/00330124.2012.681509

Janssen, M. A., & Ostrom, E. (2006, 2006/08/01/). Resilience, vulnerability, and adaptation: A cross-cutting theme of the International Human Dimensions Programme on Global Environmental Change. *Global Environmental Change, 16*(3), 237-239. https://doi.org/10.1016/j.gloenvcha.2006.04.003

Jongman, B., Kreibich, H., Apel, H., Barredo, J. I., Bates, P. D., Feyen, L., Gericke, A., Neal, J., Aerts, J. C. J. H., & Ward, P. J. (2012). Comparative flood damage model assessment: towards a European approach. *Nat. Hazards Earth Syst. Sci., 12*(12), 3733-3752. https://doi.org/10.5194/nhess-12-3733-2012

Jonkman, S. N., van Gelder, P. H. A. J. M., & Vrijling, J. K. (2003, 2003/04/04/). An overview of quantitative risk measures for loss of life and economic damage. *Journal of Hazardous Materials, 99*(1), 1-30. https://doi.org/10.1016/S0304-3894(02)00283-2

Kadaster. (2023). *Basisregistratie Adressen en Gebouwen* https://www.kadaster.nl/zakelijk/registraties/basisregistraties/bag/over-bag

Kaiser, H. F. (1970, 1970/12/01). A second generation little jiffy. *Psychometrika, 35*(4), 401-415. https://doi.org/10.1007/BF02291817

Kaiser, H. F., & Rice, J. (1974, 1974/04/01). Little Jiffy, Mark Iv. *Educational and Psychological Measurement, 34*(1), 111-117. https://doi.org/10.1177/001316447403400115

Kates, R. W., Ausubel, J., Berberian, M., & International Council of Scientific Unions. Scientific Committee on Problems of the, E. (1985). *Climate impact assessment : studies of the interaction of climate and society*. Published on behalf of the Scientific Committee on Problems of the Environment (SCOPE) of the International Council of Scientific Unions (ICSU) by Wiley. https://www.cakex.org/sites/default/files/documents/a1985.01.pdf

Kim, J. E., Yu, J., Ryu, J.-H., Lee, J.-H., & Kim, T.-W. (2021, 2021/10/01). Assessment of regional drought vulnerability and risk using principal component analysis and a Gaussian mixture model. *Natural Hazards, 109*(1), 707-724. https://doi.org/10.1007/s11069-021-04854-y

Kind, J. (2019). *Drowning by Numbers - Social Welfare, Cost-Benefit Analysis and Flood Risk Management* https://www.researchgate.net/profile/Jarl-Kind/publication/332671909_Drowning_by_Numbers_-_Social_Welfare_Cost-Benefit_Analysis_and_Flood_Risk_Management/links/61e92486c5e3103375a9bd29/Drowning-by-Numbers-Social-Welfare-Cost-Benefit-Analysis-and-Flood-Risk-Management.pdf

Kind, J., Botzen, W. J. W., & Aerts, J. C. J. H. (2020, Apr). Social vulnerability in cost-benefit analysis for flood risk management. *Environment and Development Economics, 25*(2), 115-134. https://doi.org/10.1017/S1355770X19000275

Kind, J., Wouter Botzen, W. J., & Aerts, J. C. J. H. (2017, 2017/03/01). Accounting for risk aversion, income distribution and social welfare in cost-benefit analysis for flood risk management. *WIREs Climate Change, 8*(2), e446. https://doi.org/10.1002/wcc.446

Kirby, R. H., Reams, M. A., Lam, N. S. N., Zou, L., Dekker, G. G. J., & Fundter, D. Q. P. (2019, 2019/06/01). Assessing Social Vulnerability to Flood Hazards in the Dutch Province of Zeeland. *International Journal of Disaster Risk Science, 10*(2), 233-243. https://doi.org/10.1007/s13753-019-0222-0

Klein, J. (2023). Personal Communication

Koks, E. E., de Moel, H., Aerts, J. C. J. H., & Bouwer, L. M. (2014, 2014/02/01). Effect of spatial adaptation measures on flood risk: study of coastal floods in Belgium. *Regional Environmental Change, 14*(1), 413-425. https://doi.org/10.1007/s10113-013-0514-7

Koks, E. E., Jongman, B., Husby, T. G., & Botzen, W. J. W. (2015, 2015/03/01/). Combining hazard, exposure and social vulnerability to provide lessons for flood risk management. *Environmental Science & Policy, 47*, 42-52. https://doi.org/10.1016/j.envsci.2014.10.013

Kuhlicke, C., Scolobig, A., Tapsell, S., Steinführer, A., & De Marchi, B. (2011, 2011/08/01). Contextualizing social vulnerability: findings from case studies across Europe. *Natural Hazards, 58*(2), 789-810. https://doi.org/10.1007/s11069-011-9751-6

Laska, S., & Morrow, B. H. (2006, //). Social Vulnerabilities and Hurricane Katrina: An Unnatural Disaster in New Orleans. *Marine Technology Society Journal, 40*(4), 16-26. https://doi.org/10.4031/002533206787353123

Lee, Y.-J. (2014, 2014/01/01/). Social vulnerability indicators as a sustainable planning tool. *Environmental Impact Assessment Review, 44*, 31-42. https://doi.org/10.1016/j.eiar.2013.08.002

Lein, J. K., & Abel, L. E. (2010, 2010/01/01). Hazard vulnerability assessment: How well does nature follow our rules? *Environmental Hazards, 9*(2), 147-166. https://doi.org/10.3763/ehaz.2010.0027

Leusink, E., & Swets, A. (2017). *verbreed Gemeentelijk Rioleringsplan Haarlem*. https://gemeentebestuur.haarlem.nl/Vergaderingen/Raad/2017/21-december/19:30/Vaststelling-Verbreed-Gemeentelijk-RioleringsPlan-2018-2023-1/2017361078-2-Bijlage-1-VGRP-2018-2023-1.pdf

Löwe, R., Urich, C., Sto. Domingo, N., Mark, O., Deletic, A., & Arnbjerg-Nielsen, K. (2017, 2017/07/01/). Assessment of urban pluvial flood risk and efficiency of adaptation options through simulations – A new generation of urban planning tools. *Journal of Hydrology, 550*, 355-367. https://doi.org/10.1016/j.jhydrol.2017.05.009

Maas, M. (2023). Personal Communication

Mohor, G. S., Hudson, P., & Thieken, A. H. (2020, 2020/04/01). A Comparison of Factors Driving Flood Losses in Households Affected by Different Flood Types. *Water Resources Research, 56*(4), e2019WR025943. https://doi.org/10.1029/2019WR025943

Pallathadka, A. K. C., Heejun; and Ajibade, Idowu. (2021). The Spatial Patterns of Pluvial Flood Risk, Blue-Green Infrastructure, and Social Vulnerability: A Case Study from Two Alaskan Cities. *International Journal of Geospatial and Environmental Research, 8*(3), 2. https://dc.uwm.edu/ijger/vol8/iss3/2

Pathak, S., Panta, H. K., Bhandari, T., & Paudel, K. P. (2020, 2020/12/01). Flood vulnerability and its influencing factors. *Natural Hazards, 104*(3), 2175-2196. https://doi.org/10.1007/s11069-020-04267-3

Philip, D., & Rayhan, M. I. (2004). Vulnerability and Poverty: What are the causes and how aur they related? [Term paper]. https://www.zef.de/fileadmin/downloads/forum/docprog/Termpapers/2004_3a_Philip_Rayan.pdf

Polsky, C., Neff, R., & Yarnal, B. (2007, 2007/08/01/). Building comparable global change vulnerability assessments: The vulnerability scoping diagram. *Global Environmental Change, 17*(3), 472-485. https://doi.org/10.1016/j.gloenvcha.2007.01.005

Rangari, V. A., Gonugunta, R., Umamahesh, N. V., Patel, A. K., & Bhatt, C. M. (2018). 1D-2D MODELING OF URBAN FLOODS AND RISK MAP GENERATION FOR THE PART OF HYDERABAD CITY. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-5*, 445-450. https://doi.org/10.5194/isprs-archives-XLII-5-445-2018

Remo, J. W. F., Pinter, N., & Mahgoub, M. (2016, 2016/03/01). Assessing Illinois's flood vulnerability using Hazus-MH. *Natural Hazards, 81*(1), 265-287. https://doi.org/10.1007/s11069-015-2077-z

Rijkswaterstaat. (2015). *The National Flood Risk Analysis for the Netherlands*. R. V. P. Office. https://www.helpdeskwater.nl/publish/pages/131663/vnk-rapport-eng-lr.pdf

Roeleveld, P. (2023). Personal Communication

Rozendaal, B. (2023). Personal Communication

Schmidtlein, M. C., Deutsch, R. C., Piegorsch, W. W., & Cutter, S. L. (2008, 2008/08/01). A Sensitivity Analysis of the Social Vulnerability Index. *Risk Analysis, 28*(4), 1099-1114. https://doi.org/10.1111/j.1539-6924.2008.01072.x

Shan, X., Wang, J., Wen, J., Hu, H., Wang, L., Yin, J., & Li, M. (2022, 2022/08/01). Using Multidisciplinary Analysis to Develop Adaptation Options against Extreme Coastal Floods. *International Journal of Disaster Risk Science, 13*(4), 577-591. https://doi.org/10.1007/s13753-022-00421-6

Tapsell, S. M., McCarthy, S., Faulkner, H. P., Alexander, M., Steinführer, A., Kuhlicke, C., Brown, S., Walker, G., Pellizzoni, L., Scolobig, A., De Marchi, B., Bianchizza, C., Supramaniam, M., & Kallis, G. (2010). *Social vulnerability to natural hazards*. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=11c14ea068dc55331c9cd6ae6c56356984bd4dbd

Tascón-González, L., Ferrer-Julià, M., Ruiz, M., & García-Meléndez, E. (2020). Social Vulnerability Assessment for Flood Risk Analysis. *Water, 12*(2). https://doi.org/10.3390/w12020558

Tate, E. (2012, 2012/09/01). Social vulnerability indices: a comparative assessment using uncertainty and sensitivity analysis. *Natural Hazards, 63*(2), 325-347. https://doi.org/10.1007/s11069-012-0152-2

Tate, E. (2013, 2013/05/01). Uncertainty Analysis for a Social Vulnerability Index. *Annals of the Association of American Geographers, 103*(3), 526-543. https://doi.org/10.1080/00045608.2012.700616

Tiepolo, M., Belcore, E., Braccio, S., Issa, S., Massazza, G., Rosso, M., & Tarchiani, V. (2021, 2021/01/01/). Method for fluvial and pluvial flood risk assessment in rural settlements. *MethodsX, 8*, 101463. https://doi.org/10.1016/j.mex.2021.101463

UNDRO. (1980). *Natural disasters and vulnerability analysis :report of Expert Group Meeting, 9-12 July 1979*. https://digitallibrary.un.org/record/95986/files/UNDRO_ExpGrp_1-EN.pdf

VLN-Nieuws. (2022). *NH Nieuws*. https://www.nhnieuws.nl/nieuws/305619/straten-blank-na-korte-heftige-regenbuien-in-haarlem-en-heemstede

Weeren, B.-J. v., Beersma, J., Peerdeman, K., Hakvoort, H., & Talsma, M. (2018). *Overal in Nederland, nu én in de toekomst : een actueel beeld van de kans op extreme neerslag*. STOWA. https://edepot.wur.nl/502598

Wolf, S. (2012, 2012/04/01). Vulnerability and risk: comparing assessment approaches. *Natural Hazards, 61*(3), 1099-1113. https://doi.org/10.1007/s11069-011-9968-4

Zhou, Q., Mikkelsen, P. S., Halsnæs, K., & Arnbjerg-Nielsen, K. (2012, 2012/01/11/). Framework for economic pluvial flood risk assessment considering climate change effects and adaptation benefits. *Journal of Hydrology, 414-415*, 539-549. https://doi.org/10.1016/j.jhydrol.2011.11.031

Zwolle, G. (2021). *Gemeentelijk rioleringsprogramma 2022-2026*. https://zwolle.bestuurlijkeinformatie.nl/Document/View/df0b5480-e63e-4d9a-b65e-50a26f57173d

Zwolle, G. (2023). *Cijfers over Zwolle*. https://cijfersoverzwolle.nl/

Appendices

Appendix A – Indicator selection

Cutter 2003	Fekete 2009	Koks 2015	Forrest 2020	Tascon 2020	Holand 2012	CBS data availability
Average number of people per household, 1990	− Persons per hh					Gemiddelde huishoudgrootte
Birth rate (number of births per 1,000 population), 1990					Birth rate	Geboorten totaal
Earnings (in $1,000) in all industries per square mile, 1990						
General local government debt to revenue ratio, 1992	− Municipality debts per resident				Debt-to-revenu ratio	
Land in farms as a percent of total land, 1992						
Median age, 1990					Median age	Bepaalbaar ahv klassen
Median dollar value of owner-occupied housing, 1990					House value	Gemiddelde WOZ-waarde woningen
Median rent (in dollars) for renter-occupied housing units, 1990						
Net international migration, 1990–1997	− New residents					
Number of commercial establishments per square mile, 1990					Commercial density	Veel soorten voorzieningen in aantal en binnen afstand beschikbaar
Number of housing permits per new						

residential construction per square mile, 1990						
Number of housing units per square mile, 1990	− Population per settlement area	Total number of households			House building density	Woningen
Number of manufacturing establishments per square mile, 1992						
Number of physicians per 100,000 population, 1990	− Residents per doctor				Per capita number of physicians	Aantal huisartesnpraktijken binnen 1,3,5 kolometer
Per capita income (in dollars), 1989	+ GDP per labour force	Average monthly incoome	Individuals on low income (<20%)		Median income	Mediaan huishoudinkomen
Per capita number of community hospitals, 1991	− Medical care centres				Distance to hospital	Aantal ziekenhuizen binnen 5,10,20 kilometer
Per capita residents in nursing homes, 1991					Nursing home residents	
Per capita Social Security recipients, 1990	− Social welfare recipients		Individuals unemployed		Disability pension	Inwoners met WW, bijstand, arbeidsongeschiktheidsuitkering
Percent African American, 1990		Non-European immigrants			Non-Western immigrants	Niet-westerse migratieactergrond
Percent Asian, 1990					--	
Percent employed in primary extractive industries (farming, fishing, mining, and forestry), 1990					Primary industries	
Percent employed in service occupations, 1990					Low-skill service sector	
Percent employed in transportation,						

communications, and other public utilities, 1990						
Percent female-headed households, no spouse present, 1990		One parent households			Single-parent households	Eenouderhuishoudens
Percent females participating in civilian labor force, 1990	+ Female employed					
Percent females, 1990	− Female gender			Population of women		Vrouwen totaal
Percent Hispanic, 1990					--	
Percent living in poverty, 1990					Low income	Mediaan inkomen categorie laag
Percent Native American, 1990					--	
Percent of civilian labor force unemployed, 1991	− Unemployment		Individuals unemployed		Unemployed	
Percent of households earning more than $75,000, 1989			High income		High income	Mediaan inkomen categorie hoog
Percent of housing units that are mobile homes, 1990			Individuals in renting house with low value			Woningen naar bouwjaar
Percent of population 25 years or older with no high school diploma, 1990	− Graduates without basic education			Population without studies	Low level of education	
Percent of population over 65 years, 1990	− Residents age 65 and older	Age 65 years and older		Population over 65 years	Population over 66 years	Inwoners van 65 jaar of ouder
Percent of population under five years old, 1990	− Residents below age 6	Age 0-14 years	Individuals under 4 years old and above 75 years old	Population under 16 years	Population under 6 years	Inwoners tot 15 jaar
Percent of the population participating in the labor force, 1990					Labor force participation	

Percent population change, 1980/1990					Population change	
Percent renter-occupied housing units, 1990					Renters	Huurwoningen totaal
Percent rural farm population, 1990						
Percent urban population, 1990	+ Rural population	Total inhabitants		Population density	Urban population	Stedelijkheid
Value of all property and farm products sold per square mile, 1990						
Vote cast for president, 1992—percent voting for leading party (Democratic)			Turnout in political elections		Voter turnout	
	− Commuters in					
	− Day-care centre					Kinderdagverblijven
	− Elementary Schools per Resident					
	− Foreign females					
	− Foreigners			Foreign population		
	− Handicapped unemployed		Individuals registered as disabled	Population with disabilities		
	− Key funds allocation					
	− Persons in need of care					
	− Population projection age 60+					

	− Rehabilitation centres per Resident					
	− Rent subsidies					
	− Small apartments					
	− Tourist overnight stays			Tourist population		
	+ Building land prices					
	+ Fixed investments					
	+ Foreign employed					
	+ Graduates with High school graduation		Individuals with medium or high levels of education		High level of education	
	+ High qualification employed					
	+ Hospital beds					
	+ Income per hh					
	+ Living space pp					
	+ New apartments					
	+ One and two family homes					
	+ Open space					

	+ Residents from age 30 to 50					Inwoners van 25 tot 45 jaar
	+ University students					
		Average construction year of property			Old houses	Woningen naar bouwjaar
			Indiviuals owning own home			Koopwoningen
				Evacuation time		
					Median assets	
					Employed in health care	
					Outmigration	
					Western immigrants	Westerse migratieachtergrond
					Mortality between 40-49 years	
					Disposable income	
					Local government expences on debt	
					Exit routes	
					Lifelines	
					Municipal roads	

					Age water pipelines	
					Age sewer pipelines	
					High-density urban population	
					Gender equality	
					Social assistance instances	
						Inwoners totaal
						Mannen totaal
						Inwoners van 15 tot 25 jaar
						Inwoners van 45 tot 65 jaar
						Nederlandse achtergrond
						Tweeouderhuishoudens
						Huishoudens totaal
						Eenpersoonshuishoudens
						Meerpersoonshuishoudens zonder kinderen
						Niet bewoonde woningen
						Huurwoningen in eigendom van woningcorporatie
						Meergezinswoningen
						Gemiddeld aardgasverbruik
						Gemiddeld elektriciteitsverbruik
						Afstand tot dichtstbijzijnde huisartsenpraktijk
						Afstand tot dichtstbijzijnde ziekenhuis exclusief buitenpolikliniek

						Afstand tot dichtstbijzijnde ziekenhuis inclusief buitenpolikliniek
						Huisartsenpost
						Apotheek
						Supermakrt
						Overige dagelijkse levensmiddelen
						Warenhuizen
						Cafés
						Cafetaria's
						Restaurants
						Hotels
						BSO
						Basisscholen
						Scholen voortgezetonderwijs
						Scholen VMBO
						Scholen HAVO/VWO
						Oprit hoofdverkeersweg
						Treinstation
						Overstapstation
						Bibliotheek
						Poppodium
						Podiumkunsten
						Attractiepark
						Musea
						Kunstijsbaan
						Bioscoop
						Sauna
						Zonnebank
						Brandweerkazerne

Appendix B – PCA results Haarlem

Table 10: Total Variance Explained PCA Haarlem

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7,152	31,097	31,097	7,152	31,097	31,097	6,767	29,421	29,421
2	3,561	15,483	46,580	3,561	15,483	46,580	2,564	11,147	40,568
3	2,236	9,723	56,303	2,236	9,723	56,303	2,276	9,897	50,465
4	1,638	7,122	63,425	1,638	7,122	63,425	2,133	9,274	59,739
5	1,376	5,981	69,406	1,376	5,981	69,406	1,692	7,356	67,095
6	1,083	4,710	74,116	1,083	4,710	74,116	1,615	7,021	74,116
7	0,986	4,287	78,403						
8	0,840	3,651	82,054						
9	0,752	3,270	85,324						
10	0,636	2,766	88,090						
11	0,437	1,902	89,992						
12	0,424	1,845	91,836						
13	0,356	1,548	93,384						
14	0,322	1,402	94,786						
15	0,277	1,205	95,991						
16	0,222	0,966	96,957						
17	0,209	0,908	97,865						
18	0,140	0,609	98,474						
19	0,120	0,522	98,997						
20	0,088	0,381	99,378						
21	0,075	0,327	99,705						
22	0,056	0,242	99,947						
23	0,012	0,053	100,000						

Table 11: Component Matrix PCA Haarlem

Component Matrix						
	Component					
	1	2	3	4	5	6
P_HoogInkP	-0,885	-0,124	0,008	0,030	-0,088	0,076
P_corporatiewoning	0,845	-0,006	-0,049	-0,106	-0,172	0,256
P_SocMinH	0,845	0,060	0,005	0,058	0,017	-0,009
P_niet_westerse	0,783	0,208	-0,235	0,028	-0,032	-0,068
P_ARB_PP	-0,770	0,317	-0,025	0,000	-0,128	-0,114
G_inkomen_hh_1000	-0,758	-0,290	-0,034	0,131	0,034	0,147
P_huurwoningen	0,751	0,149	0,177	-0,088	-0,417	0,078
G_woz	-0,732	-0,331	-0,094	0,060	0,029	0,284
Pp_met_uitkering	0,716	0,029	0,128	0,002	-0,151	0,108
P_OPL_LG_PC5	0,662	-0,006	-0,078	0,601	0,073	0,079
P_eenouder	0,464	0,187	-0,456	-0,131	-0,064	0,410
M_leeftijd	0,153	-0,744	0,480	-0,023	0,164	0,249
Population_density	-0,016	0,713	0,244	0,032	0,598	0,224
P65	0,249	-0,690	0,465	-0,081	0,092	0,239
Household_density	0,038	0,674	0,416	0,022	0,553	0,138
P_Open_space	0,299	-0,674	-0,220	0,096	0,184	-0,172
OAD	-0,141	0,584	0,393	-0,170	-0,325	0,119
P015	-0,257	0,112	-0,794	-0,030	0,188	0,223
P_westerse	-0,299	0,334	0,442	-0,008	-0,238	-0,288
A_geboorten	-0,014	0,318	-0,404	0,177	0,026	-0,216
P_WW	0,124	0,240	0,100	0,837	-0,179	0,148
P_OPL_HG_PC5	-0,560	-0,060	0,120	0,609	-0,220	0,160
Median_building_year	0,461	-0,274	0,110	0,252	0,317	-0,490

Table 12: Rotated Component Matrix PCA Haarlem

Rotated Component Matrix						
	Component					
	1	2	3	4	5	6
P_HoogInkP	-0,863	-0,008	-0,224	-0,122	-0,046	0,022
P_corporatiewoning	0,836	0,197	-0,034	-0,095	0,271	0,032
P_huurwoningen	0,833	0,133	-0,241	-0,150	-0,080	0,085
P_SocMinH	0,813	0,064	0,201	0,055	0,067	0,087
G_inkomen_hh_1000	-0,810	0,109	-0,031	-0,121	0,079	0,098
G_woz	-0,793	0,170	-0,077	-0,144	0,226	0,060
P_niet_westerse	0,791	-0,210	0,154	0,009	0,155	0,038
Pp_met_uitkering	0,718	0,191	-0,009	-0,027	0,029	0,106
P_ARB_PP	-0,646	-0,368	-0,361	0,046	-0,198	-0,013
M_leeftijd	-0,029	0,898	0,284	-0,086	-0,015	0,009
P65	0,093	0,866	0,227	-0,118	-0,019	-0,027
A_geboorten	0,026	-0,567	0,092	0,029	0,074	0,086
OAD	0,069	-0,076	-0,713	0,228	-0,304	0,017
Median_building_year	0,326	0,065	0,707	0,038	-0,299	0,053
P_Open_space	0,103	0,235	0,677	-0,337	0,155	-0,052
Population_density	0,031	-0,146	-0,166	0,961	0,035	0,036
Household_density	0,091	-0,048	-0,162	0,947	-0,138	0,032
P015	-0,286	-0,480	0,058	-0,022	0,685	-0,094
P_eenouder	0,477	-0,144	-0,146	-0,018	0,614	-0,017
P_westerse	-0,150	-0,110	-0,350	0,098	-0,607	0,026
P_WW	0,122	-0,126	-0,028	0,093	-0,071	0,891
P_OPL_HG_PC5	-0,567	0,015	-0,141	-0,109	-0,104	0,642
P_OPL_LG_PC5	0,560	0,000	0,369	0,067	0,145	0,584

Appendix C – PCA results Zwolle

Table 13: Total Variance Explained PCA Zwolle

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7,159	31,128	31,128	7,159	31,128	31,128	5,202	22,619	22,619
2	3,113	13,537	44,664	3,113	13,537	44,664	2,644	11,497	34,115
3	2,450	10,651	55,315	2,450	10,651	55,315	2,515	10,936	45,051
4	1,582	6,878	62,194	1,582	6,878	62,194	2,428	10,558	55,610
5	1,461	6,354	68,548	1,461	6,354	68,548	2,303	10,013	65,623
6	1,221	5,309	73,857	1,221	5,309	73,857	1,894	8,234	73,857
7	0,889	3,863	77,720						
8	0,853	3,710	81,430						
9	0,711	3,090	84,519						
10	0,538	2,338	86,858						
11	0,480	2,087	88,945						
12	0,409	1,777	90,722						
13	0,381	1,655	92,377						
14	0,347	1,510	93,887						
15	0,312	1,358	95,244						
16	0,236	1,025	96,270						
17	0,220	0,957	97,226						
18	0,180	0,782	98,009						
19	0,170	0,739	98,747						
20	0,126	0,549	99,296						
21	0,106	0,460	99,756						
22	0,048	0,210	99,966						
23	0,008	0,034	100,000						

Table 14: Component Matrix PCA Zwolle

Component Matrix						
	Component					
	1	2	3	4	5	6
P_huurwoningen	0,830	-0,046	-0,040	-0,227	0,112	0,209
G_woz	-0,790	-0,004	0,079	0,127	-0,005	-0,247
P_niet_westerse	0,761	0,061	0,328	-0,149	0,091	0,030
P_corporatiewoning	0,760	-0,074	0,247	-0,019	0,215	0,204
P_SocMinH	0,759	-0,343	0,061	-0,126	-0,218	-0,204
P_HoogInkP	-0,729	0,411	-0,109	-0,039	0,299	0,256
G_inkomen_hh_1000	-0,723	0,084	-0,023	0,121	-0,067	-0,021
Pp_met_uitkering	0,656	-0,130	0,109	-0,184	0,209	0,144
P_WW	0,618	0,198	-0,241	0,313	-0,488	0,075
Household_density	0,615	0,234	-0,216	0,445	0,446	-0,246
Population_density	0,583	0,304	-0,144	0,515	0,420	-0,249
OAD	0,554	0,058	-0,547	-0,220	-0,065	0,093
P_eenouder	0,518	0,324	0,451	0,230	0,136	0,019
P_westerse	0,429	0,230	-0,036	-0,115	-0,083	0,384
M_leeftijd	-0,307	-0,735	-0,128	0,430	0,157	0,257
P65	-0,145	-0,696	-0,195	0,455	0,088	0,271
P_ARB_PP	-0,419	0,678	-0,280	-0,062	0,128	0,151
A_geboorten	0,001	0,614	0,337	0,093	-0,101	0,310
P_Open_space	-0,423	-0,295	0,677	-0,111	-0,031	-0,031
P015	-0,153	0,553	0,598	0,220	-0,098	-0,196
P_OPL_HG_PC5	-0,066	0,425	-0,584	0,171	-0,382	0,120
P_OPL_LG_PC5	0,498	-0,091	0,227	0,396	-0,591	-0,107
Median_building_year	-0,158	0,015	0,401	0,315	-0,047	0,580

Table 15: Rotated Component Matrix PCA Zwolle

Rotated Component Matrix						
	Component					
	1	2	3	4	5	6
P_huurwoningen	0,860	0,131	0,039	0,145	0,126	-0,081
P_corporatiewoning	0,777	0,180	-0,011	-0,113	0,252	0,137
G_woz	-0,772	-0,185	-0,024	-0,251	-0,120	0,001
P_niet_westerse	0,725	0,260	0,254	-0,139	0,198	0,075
Pp_met_uitkering	0,723	0,088	-0,010	-0,083	0,130	-0,054
G_inkomen_hh_1000	-0,664	-0,240	-0,052	-0,042	-0,181	0,122
P_westerse	0,486	-0,002	0,125	0,315	-0,035	0,232
P_OPL_LG_PC5	0,119	0,860	0,024	0,132	0,094	0,212
P_HoogInkP	-0,471	-0,750	0,068	0,060	-0,085	0,261
P_SocMinH	0,580	0,605	0,008	-0,029	0,042	-0,311
P_ARB_PP	-0,335	-0,595	0,331	0,381	0,044	0,184
M_leeftijd	-0,201	0,002	-0,924	-0,174	-0,042	0,030
P65	-0,100	0,093	-0,900	-0,045	0,005	0,023
P015	-0,308	0,126	0,564	-0,260	0,199	0,497
P_OPL_HG_PC5	-0,258	-0,013	0,095	0,799	0,005	0,046
P_Open_space	-0,223	0,066	-0,011	-0,711	-0,364	0,217
P_WW	0,261	0,557	0,070	0,618	0,197	0,131
OAD	0,492	0,024	0,017	0,559	0,054	-0,335
Population_density	0,233	0,088	0,052	0,156	0,931	0,000
Household_density	0,294	0,062	0,007	0,171	0,899	-0,096
P_eenouder	0,377	0,211	0,297	-0,149	0,437	0,401
Median_building_year	0,000	-0,027	-0,235	-0,109	-0,124	0,736
A_geboorten	0,022	-0,095	0,416	0,114	0,033	0,640

Appendix D – Social vulnerability maps

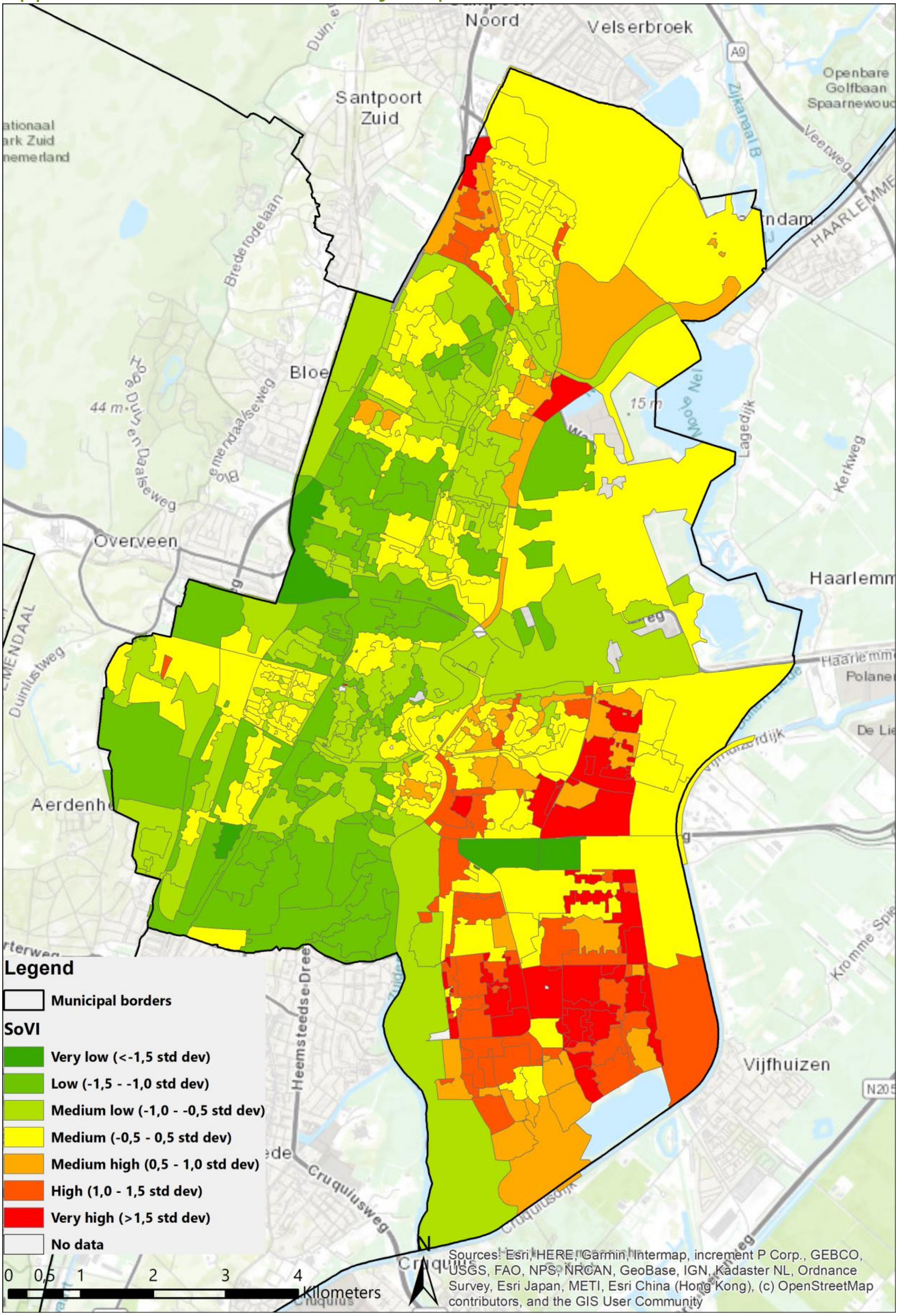

Figure 20: Social Vulnerability Index Haarlem

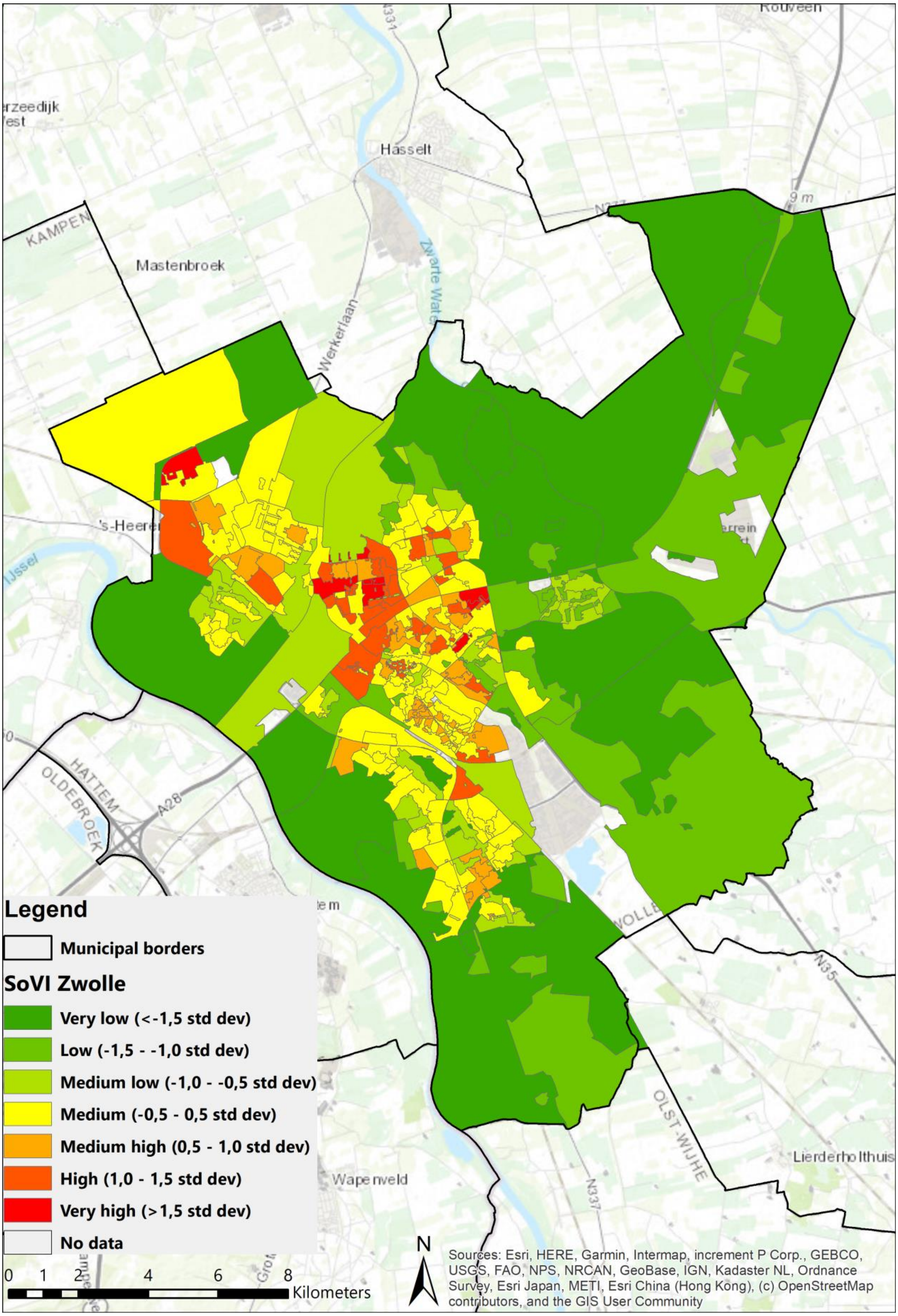

Figure 21: Social Vulnerability index Zwolle

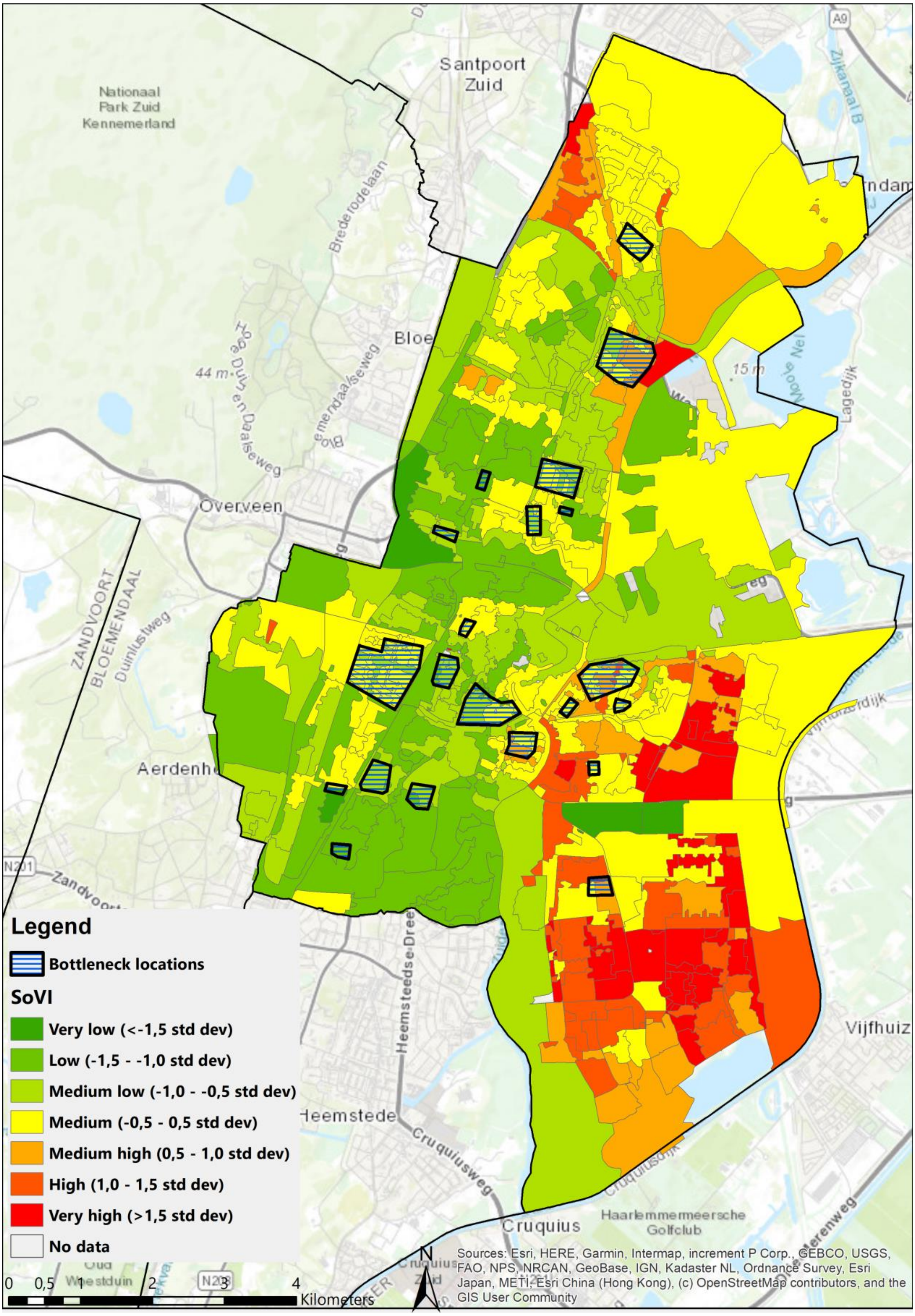

Figure 22: Social vulnerability index for Haarlem overlayed with bottleneck locations for floods

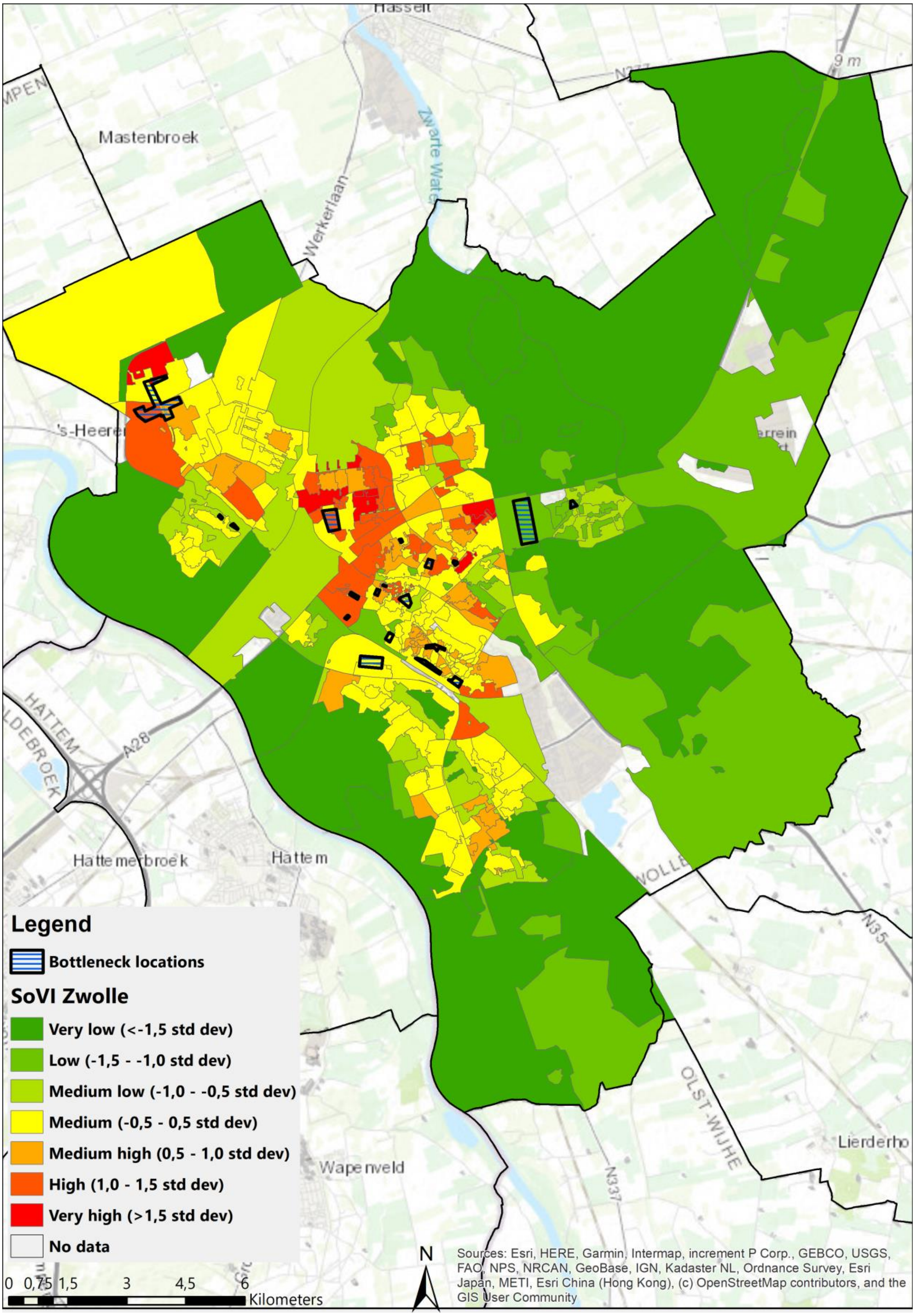

Figure 23: Social vulnerability index for Zwolle overlayed with bottleneck locations for floods

Appendix E – Maps of indicators

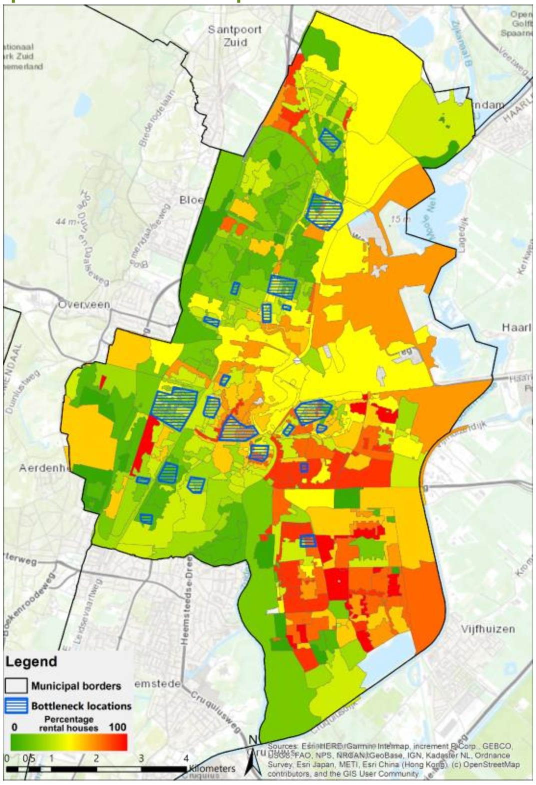

Figure 24: Percentage rental houses per PC5 area Haarlem

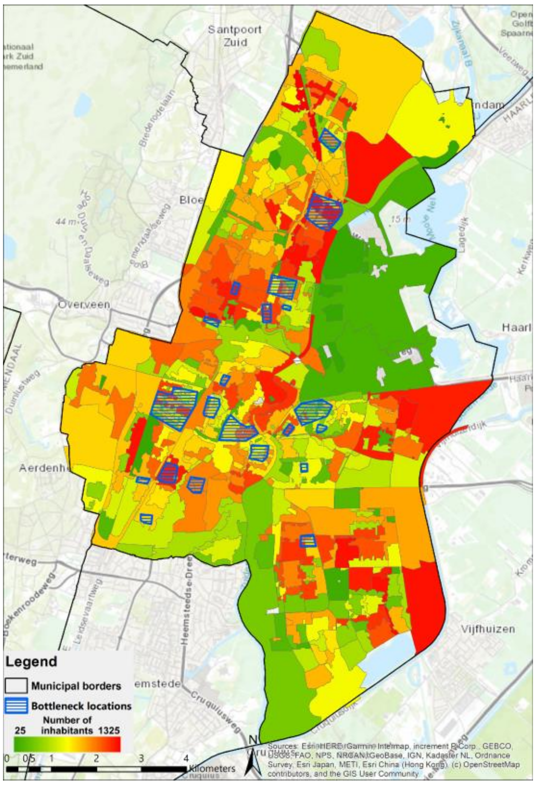

Figure 25: Number of inhabitants per PC5 area Haarlem

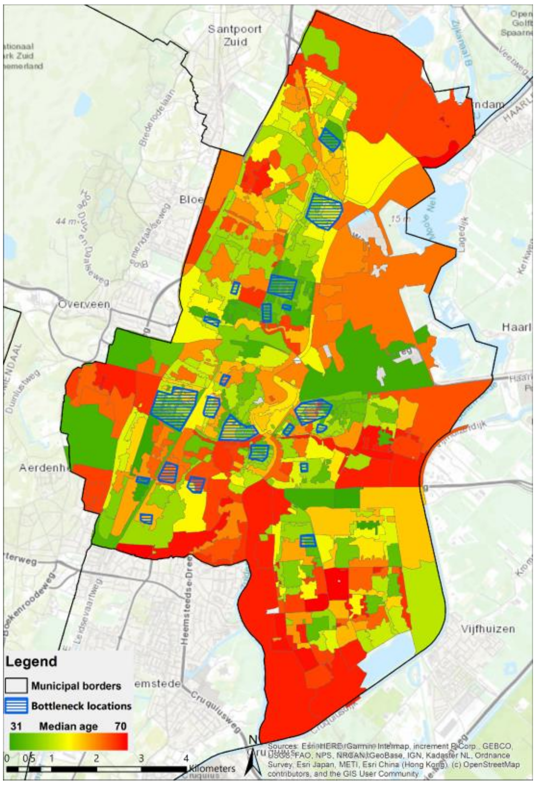

Figure 26: Median age per PC5 area Haarlem

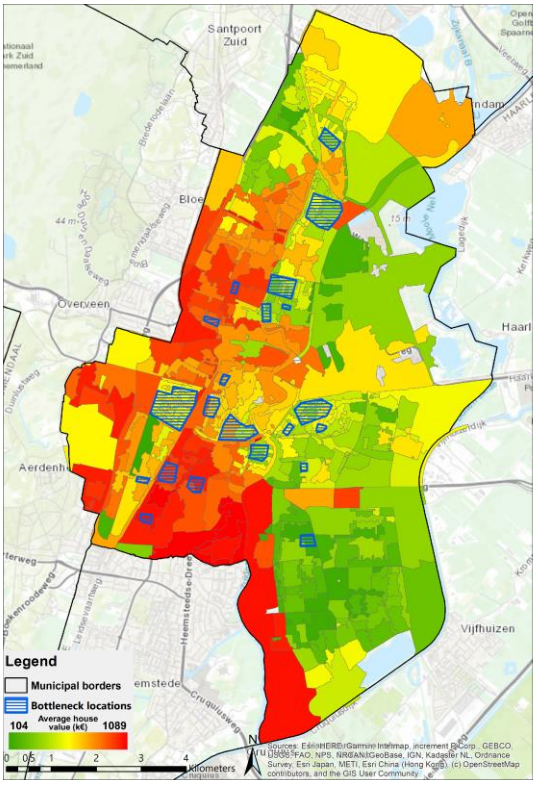

Figure 27: Average house value (k€) per PC5 area Haarlem

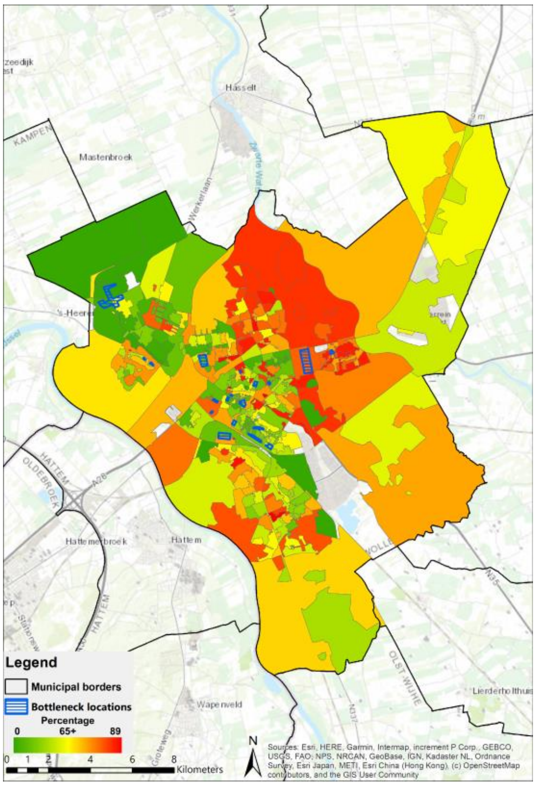

Figure 28: Percentage of the population aged 65+ per PC5 area Zwolle

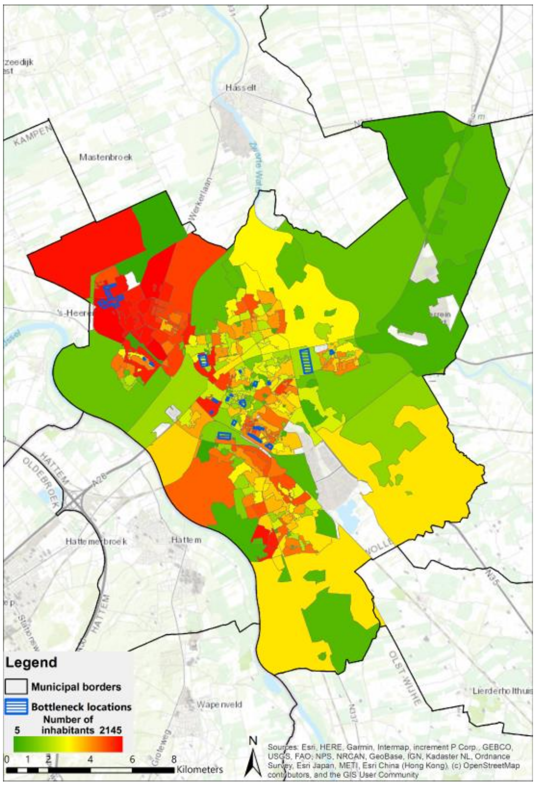

Figure 29: Number of inhabitants per PC5 area Zwolle

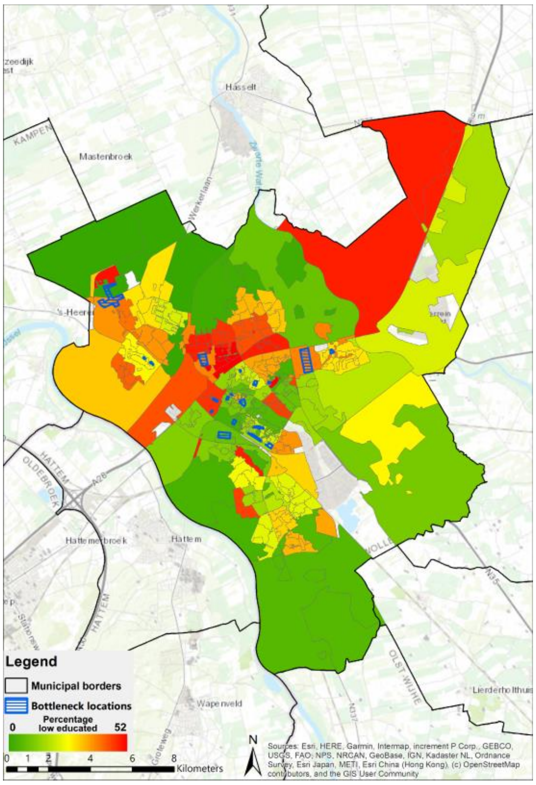

Figure 30: Percentage of the population with a low level of education per PC5 area Zwolle

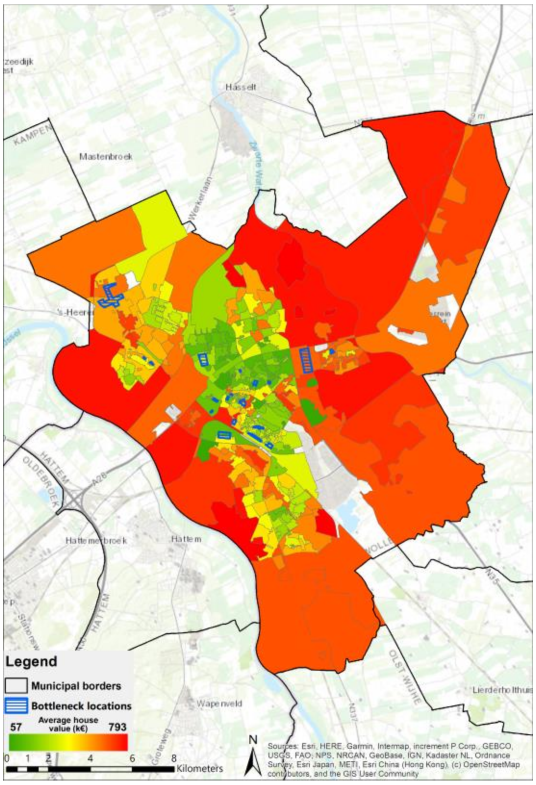

Figure 31: Average house value (k€) per PC5 area Zwolle