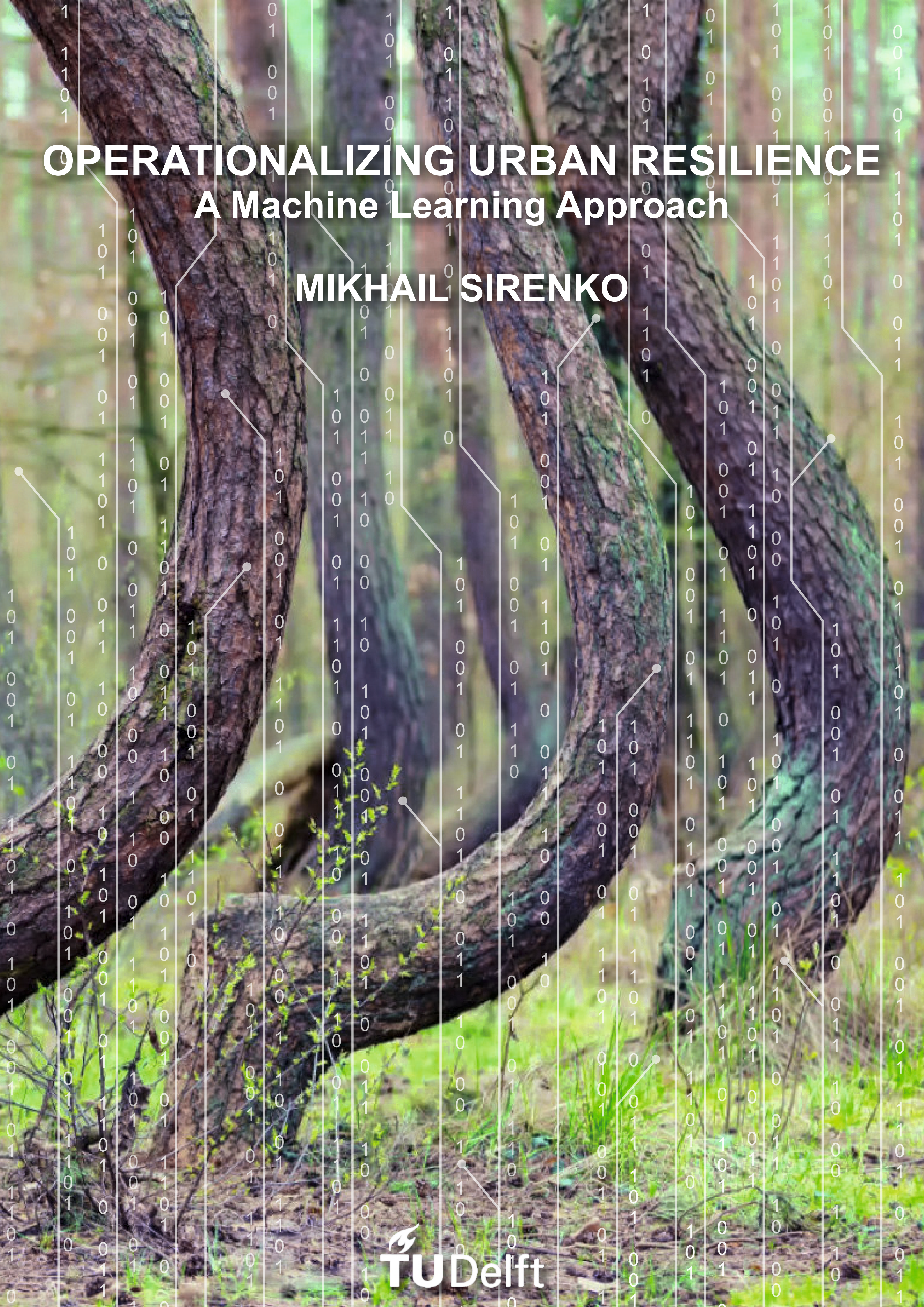

OPERATIONALIZING URBAN RESILIENCE
A Machine Learning Approach

MIKHAIL SIRENKO

TU Delft

Operationalizing Urban Resilience

A Machine Learning Approach

by

Mikhail Sirenko

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday August 30, 2019 at 09:00 AM.

Student number:	4737881	
Project duration:	April 30, 2019 – August 30, 2019	
Thesis committee:	Prof. dr. ir. A. Verbraeck,	TU Delft, chair
	Dr. S. Cunningham,	TU Delft, first supervisor
	Dr. Y. Huang,	TU Delft, second supervisor

An electronic version of this thesis is available at http://repository.tudelft.nl/.

Associated code and models are available at
http://github.com/mikhailsirenko/urban-resilience-ml.

TU Delft

Executive summary

The rapid development of urban areas poses numerous problems for both urban planners and citizens. The uncertainties associated with climate change are making the situation even more complicated. It is unknown when an extreme climatic event will strike, what will be its duration and power. The concept of urban resilience perceived to be promising in that regards. Planning resilient “fail-to-safe” instead of conventional “fail-safe” cities provide us with evidence in a positive future for urban areas. However, the practice of urban resilience is holding up due to a variety of discrepancies: about its meaning, how it can be assessed, and what are the potential solutions for it. During the last decade, cities have been accumulating a lot of data that is available now for public use. Advances in hardware and software engineering boosted machine learning field and provided it with a possibility efficiently process and analyze complex data sets. To bridge the gap between the theory and practice of urban resilience, it is proposed to use open data and machine learning algorithms for its assessment.

This study has three major outcomes. First, a new conceptual framework is created. This framework combines vulnerability concept with socio-ecological resilience. Resilience, in this case, is not general, but shock-specific: a city should be resilient to extreme heat. To operationalize the components of socio-ecological resilience, three theory-based infrastructure solutions are proposed: critical, green, and social. Critical infrastructure is represented by ambulance services. Green infrastructure stands for vegetation in the form of trees, shrubs, and grass. Social infrastructure is an umbrella term for places where people can gather, build up, or improve their relationships. Important to mention that the proposed framework is relatively flexible. It can be reconfigured for other shocks such as flooding or an earthquake.

Secondly, an advanced computational framework is developed. The goal of this framework is to convert proposed concepts to measurable indicators. To do so, it uses a variety of open data sources and machine learning algorithms. The data sets are covering: ambulance calls and temperature, socioeconomic, socio-demographic attributes and health indicators, amenity locations, street networks, and places popularity. The concept of vulnerability is operationalized with time series forecasting (to study the impact of the heat wave on the population), dimensionality reduction and clustering (map vulnerable people in the city), and regression (identify relations between the heat wave calls and population characteristics). Socio-ecological resilience and corresponding infrastructures are analyzed with network analysis (the lengths of the routes) and geospatial analysis (clustering the objects into groups).

Thirdly, created computational framework is applied to the case study: The Hague. The results are divided into three parts: extensive analysis of the recent heat wave, assessment of population vulnerability, and resilience capacities. It was found that the recent heat wave of July 2019 had a significant impact not only The Hague’s population but also on the citizens of Rotterdam and Amsterdam. This fact is demonstrated by the number of ambulance calls that has doubled and even tripled during the heat wave. Besides, the study confirmed the fact the vulnerability is unequally distributed in time and space. There are certain periods where the number of calls is higher, and there are specific areas in the city where from more calls were made. Prediction of ambulance calls is possible but highly challenging if an extreme even is taken into account. The analysis revealed that the number of ambulance calls made from the urban heat islands in the city center was higher than on average.

Vulnerability assessment strengthens the conclusions of the previous part. Vulnerable population is unequally distributed across the city. Based on identified profiles (a unique combination of citizens attributes), the areas for specific attention were found. Importantly, the granularity of the data allows us to go beyond the discussion of districts or neighborhoods.

500 by 500 m^2 resolution provides us with a possibility for more precise interventions. Regression analysis revealed two important insights. The number of citizens above 65 years old cannot fully explain the number of ambulance call made during the heat wave. It is quite controversial. The simplest vulnerability model imposes that the age group is a single major factor. Second, the improved vulnerability model based on the profiles did not demonstrate significant relations with ambulance calls. This finding stresses the importance of concepts of green and social infrastructure for combating extreme heat threat.

Resilience capacities assessment is divided by the type of infrastructure: assessment of critical, green, and social. It was found that all parts of the city are relatively well-connected to the hospitals. The lengths of the shortest driving route from any hospital to any polygon within the city do not exceed 4.2 kilometers. It turned out that the number of trees in the city center is relatively small in comparison to the southwest. Importantly, within this area, there is a high concentration of vulnerable individuals. The study revealed that for inhabitants of certain parts of the city, southwest, for instance, the libraries (the chosen type of social infrastructure) are hardly reachable. Again, this part of the city has a high concentration of vulnerable individuals and communities.

Most publications on urban resilience use a single case study or innovative solutions as a source for analysis and conclusions. Opposite to it, this research uses open data and machine learning algorithms to operationalize urban resilience. Such an approach promotes reusability of the results and back up existent empirical evidence with numbers. For instance, vulnerability, as well as resilience capacities, are unequally distributed across the city. Besides, the mismatch between the number of ambulance calls and demography pointed out the need for bottom-up resilient solutions. The data sets used in this study can be used for establishing better data collection practices. This was quite limited by the data available on critical, green, and social infrastructures. With more extensive data sets defined in line with the created conceptual framework, a much better picture of how resilient is the city can be obtained.

Based on these findings, a policy advice is formulated. There is a need for investments in The Hague's green and social infrastructures. Such solutions installed in specific parts of the city will enhance absorptive and adaptive capacities and make The Hague more resilient. In simple terms, vegetation will decrease the temperature within the areas of its placement and combats urban heat island effect. Libraries promote establishing better relationships that are extremely useful when extreme event happening.

Acknowledgements

I would like to thank first my committee: Prof.dr.ir. Alexander Verbraeck for providing me with great advises and bringing clarity to my research, Dr. Scott Cunningham showing me out-of-box way of looking on algorithms and support, Dr. Yilin Huang for patience and highly valuable feedback. I would also like to thank Olivier Hendrick, Senior Strategic Advisor Resilience at the Municipality of The Hague, for interesting discussions and bringing the research back to the ground. Last but not least I would also like to acknowledge the support given by my newfound friends: Ammar Faiq, Shajeeshan Lingeswaran, Yubin Cho, Merih Karabulut, Connor McMullen and many other.

Mikhail Sirenko
Delft, August 2019

Contents

1 Introduction

This chapter discusses some of the major problems of modern urban areas such as urbanization, climate change, and extreme heat. The complexity of the challenges requires an alternative to conventional planning approach: urban resilience. However, it is unclear how the assessment of such an abstract concept can be made. The proposed method is a combination of open data and machine learning algorithms. This discussion is summarized in the research gap and research questions.

1.1. Problem

1.1.1. Grand challenges: urbanization, climate change and extreme heat

During the last decade, the evidence for the ubiquitous growth of cities and their future importance has become highly recognizable. According to the latest projections, the urban population will increase to 68 percent by 2050 (United Nations, 2017). Today more than 80 percent of the world GDP is generated by cities, and therefore they can be considered as the main contributors or barriers to sustainable development. As a response, governmental bodies and international organizations have raised the priority of urbanization problems in their agenda.

The need for inclusive, safe, resilient, and sustainable cities was recognized by the United Nations and transferred in Sustainable Development Goal number eleven (United Nations, 2018b). This goal poses a set of seven targets that should be achieved by 2030. The set covers multiple aspects of urban life, for instance, poverty, climate change, and healthcare. The key players in the political arena have built their agendas around this goal. For example, the World Bank aimed at facilitating the building of sustainable cities and communities. To achieve this, it invests on average $6 billion in urban development and resilience projects every year (World Bank, 2019).

Alongside policy-makers, academia also stresses the challenges of such rapid urbanization. The consequences vary: from massive degradation of ecological systems (Alberti, 2005; Pickett et al., 2011), to plausible disruptive effects caused by flooding and urban heat islands (McGranahan et al., 2007; Rizwan et al., 2008). As much as is known about specific issues and cases, there is a lack of comprehensive understanding of cities (Batty, 2013). Usually, researchers are driven by a specific issue of interest and tend to narrow the study. Bettencourt et al. argued that there is a need for predictive and quantitative theory to ensure sustainable development. They proposed a set of empirically found relations, so-called "scaling laws" (Batty, 2013), between the total population of a city and a variety of indicators: serious crimes, HIV cases, and GDP. As the population grows, the corresponding indicators also grow exercising urban planning in place.

Apart from the issues mentioned above, significant attention is given to climate change. An agency of the United Nations, UN-Habitat, responsible for human settlements and sustainable urban development, emphasized the disastrous impacts of climate change on the future of cities (UN-Habitat, 2011). According to UN-Habitat, the rise in temperature will result in more frequent storms and cyclones as well as in extreme heat and heatwaves. Problematically, the population groups that are the most vulnerable to climate change related problems are usually urban poor and elderly. In the latest report by Intergovernmental Panel on Climate Change (IPCC) - a part of UN responsible for accessing climate change, it was stated that climate change is a complex issue that is connected to multiple areas such as water scarcity, drought and health (Intergovernmental Panel on Climate Change, 2018).

Various researchers highlighted the devastating effects of climate change on urban areas. One of the main problems here is extreme temperatures. The extreme heat in the form of heatwaves spreads unequally across the city, thus creating microclimates: urban areas where the temperature is significantly higher than in other city areas called *urban heat islands*. Bowler et al. (2010) stressed the connection between heat waves and heat strokes, hyperthermia, and increased death rates. The authors presented empirical evidence in favor of greening to reduce health problems caused by climate change. Rizwan et al. (2008) stated that rapid urbanization leads to accumulation of heat in cities and as a result, increase energy consumption and, again, can serve as a cause of increased mortality rates. Importantly, some population groups are more vulnerable to extreme heat: old, urban poor, people with existent health problems of certain types, and so on (Klinenberg, 2003; Swalheim and Dodman, 2008; Zaidi and Pelling, 2015).

Thus, rapid urbanization and climate change amplify already existent problems and challenges and require an alternative approach (Feliciotti et al., 2017).

The following Figure 1.1 summarizes the relations between the problems mentioned above and demonstrates the scope of this study (depicted with dashed line):

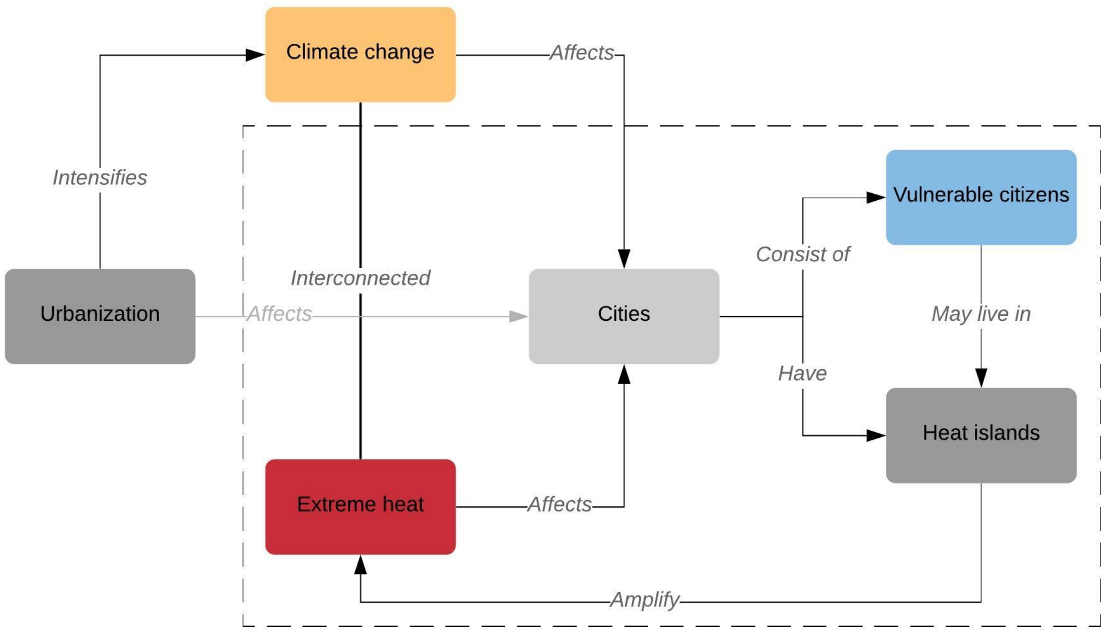

Figure 1.1: Simplified relations between the upcoming challenges of cities and the scope of this research

1.1.2. Opportunities: urban resilience, open data and machine learning

Urban resilience

One promising attempt to combat sustainability challenges is the concept of *resilience* (Meerow et al., 2016). When applied to urban areas, regardless of the system of interest (transportation, cyber, economic, etcetera), the concept usually called *urban resilience*. Remarkably, there is no agreement on the definition of urban resilience and its major concepts yet (Bourdic et al., 2012; Chelleri et al., 2015; Leichenko, 2011). Usually, a scholar interprets the concept and the ideas based on his or her scientific domain and system of interest. Consequently, a new definition and a conceptual framework often proposed. In the study made by Meerow et al. (2016), the authors named a variety of subject areas researching urban resilience: agriculture and biological sciences, engineering, social sciences, business management, and accounting, etcetera. As a result, there are many different types of resilience.

This fact counts in favor of interdisciplinary nature of resilience and partly explain the amount of attention given to it during the last decade. Not surprisingly, the notion of resilience was quickly taken onboard by city scientists who tend to think of cities as complex entities with a variety of interconnected subsystems (Batty, 2008; Godschalk, 2003). Despite the debates and discrepancies, most of the researchers pointed to the work of Holling (1973) made in system ecology field in the early 70s as the foundation for this concept. Holling (1973) proposed the following definition of resilience:

Definition 1.1.2.1 **Holling (1973)**

Resilience is the measure of the persistence of systems and of their ability to absorb change and disturbance and still maintain the same relationships between populations or state variables.

As can be seen from this definition above, it stresses the importance of *absorption capacity* and the ability to maintain stability. Other research on resilience usually has these concepts as a core. However, as mentioned previously, scholars tend to emphasize a variety of city layers and see resilience from different angles. For instance, from the engineering perspective of Lamond and Proverbs (2009) urban resilience defined as follows:

Definition 1.1.2.2 **Lamond and Proverbs (2009)**

Urban resilience encompasses the idea that towns and cities should be able to recover quickly from major and minor disasters.

Engineering resilience is interlinked with a concept of *robustness*, or in simple terms, how much damage an engineering system within a city, electric power system, for example, can handle before a collapse. Resilience, in this case, viewed as the ability to ”bounce back” fast to the original state. The definition taken from the social sciences by Desouza and Flanery (2013) is not so straightforward as the previous one:

Definition 1.1.2.3 **Desouza and Flanery (2013)**

Urban resilience is the ability to absorb, adapt and respond to changes in urban systems.

This definition introduces two more concepts: *adaptability* and *transformation*. A city as a social system not necessary should bounce back to its original state. Instead, it can adapt or transform and reach a new, yet unknown, but beneficial state. For example, during a heatwave, certain individuals may take care of their more vulnerable friends. That is, when faced a threat, they realized the importance of it, adapted and took a series of actions.

After the immediate danger of the threat is gone, the system can transform. Continuing with the previous example, vulnerable individuals may want to undertake a radical change; for instance, move to a place where there is no threat of heatwaves. In this example, adaptation is a short- or mid-term response, whereas transformation is a long-term. As can be seen, these concepts are not easy to operationalize.

Thus, it is clear that resilience is a complex theory that requires more elaborate research to make sense of it. However, the question that is left, what are the reasons for adopting it? First of all, resilience is tightly bound with the notion of *uncertainty*, what makes it highly relevant while dealing with such complex systems as cities (Bourdic et al., 2012; Chelleri et al., 2015). It is uncertain how a city as a system will respond to a shock or stress, which part of it will collapse first and how this collapse will affect other parts. Consider a failure of an IT system. Due to unknown reasons, some IT system of the large telecommunication company collapsed. Importantly, this company provides services to a national emergency network. As a result, all emergency numbers are unavailable for a couple of hours. Coupled with a high temperature in a city, and as a result, increased need for ambulances, this accident can have serious consequences. This situation happened in the Netherlands on June 24, 2019 (Nederlandse Omroep Stichting, 2019).

Previously, the intention of city planners was focused on making urban areas as robust as possible. The stable state once achieved, should be in there for generations. Ahern (2011) formulated this paradigm as **"fail-safe."** However, practice demonstrated the need in the opposite. Modern cities can be better described as dynamically changing nonlinear systems (Rodin, 2014; Tyler and Moench, 2012) that can "fail." Instead of continuous empowerment of infrastructure in place, there is a call for **"safe-to-fail"** resilient cities.

Apart from the uncertainty and complexity of contemporary urban areas (internal), there are uncertainties brought by the shocks and stresses (external). Climate change is a complex phenomenon that often causes unpredictable before disruptive events. As a result, to prepare for the climate change, policy-makers cannot count on typical urban planning practices with a core in known future and planning horizon in 10 years anymore (Cretney, 2014; Henstra, 2012; Ruth and Coelho, 2007). Ernstson et al. (2010) taking this argument even further by stating that the traditional paradigm of planning for predictable future is not only insufficient, but it may, in some ways, also be destructive. However, opposite the risk management, it promotes mid- and long-term planning instead of short- and mid-term (Mehmood, 2016; Yamagata and Sharifi, 2018).

Open data

Over the past years, the amount of data collected by cities has become more than significant. It happened due to a variety of reasons. Partly because of the development of technologies: data storage is cheaper now, there are new ways of collecting the data, for example, with sensors or satellites. Partly due to the intention of city managers to brand themselves as "smart" or "digital" city. Without a doubt, there are more rationales for this data gathering trend. However, in spite of this, the bottom line is that more extensive usage of data can contribute to better decision-making.

The concept of *open data* goes alongside with this idea. In simple terms, open data aimed at promoting public use of collected data sets. It can be made just by publishing the data online and giving an open access to it with easy to use API. From there, the data sets can be downloaded and used by citizens enthusiasts or scholars to conduct a study and possibly contribute to some positive changes within an urban area.

Of course, there are some drawbacks of extensive data collection and usage — for instance, privacy issues. Van Zoonen (2016) and many others extensively discussing privacy concerns of smart cities. The main questions are: what data, on what resolution and for what purpose? In many cases, these questions are addressed with a specific law. As a result, published data sets only include aggregated averages that do not violate the rules.

However, data does not speak for itself. Therefore, there is a need for algorithms that can be used to analyze it and derive meaningful insight.

Machine learning

Machine learning (ML) defined as automated detection of meaningful patterns in data (Shalev-Shwartz and Ben-David, 2014). With increased computational power, development of open-source programming languages such as Python and R, and newly available data sources, ML provides a variety of opportunities for urban analytics. On the one hand, it helps to analyze large data sets and come with a better understanding of urban areas. On the other side, it can help to test theories or strength empirical findings.

Taken together, urban resilience, open data and ML provide us with means and ends for the problem of extreme heat.

1.1.3. Case Study: The Netherlands and the city of The Hague

The history of the heatwaves in the Netherlands count three exceptional cases: the heatwaves of 2006, 2010 and 2018. The reports posted by the Central Agency for Statistics (CBS) highlight the key facts about these three cases (Centraal Bureau voor de Statistiek, 2018a; Harmsen and Garssen, 2006, 2010).

2006's heat waves were lasting from 26 June to 6 July the first one and from 15 July to 30 July the second one respectively. July of 2006 was the hottest month ever recorded by KNMI. The temperature was 6.6 °C higher than the average. Consequently, there was a significant increase in deaths - 1000 deaths more than on average, mostly among elderly of two age categories 65-79 and over 80 years old.

The upcoming heatwave of 2010 also resulted in an increase in around 500 deaths. Lasting from 23 June to 12 July it claimed 500 extra lives. The increase in temperature was again around 6 °C more than an average in 28 °C.

The recent heatwaves of 2018 continued this trend. However, according to the report from CBS, it causes less damage than the previous two. The three vertical color bars on Figure 1.5 represent the periods of extreme heat. The blue one depicts the heatwave of 2006 and the two pink ones, the heatwaves of 2018 overlapped with the second heat wave of 2006. As can be seen from the graph, there is an increase in deaths for both 2006 and 2018 line charts.

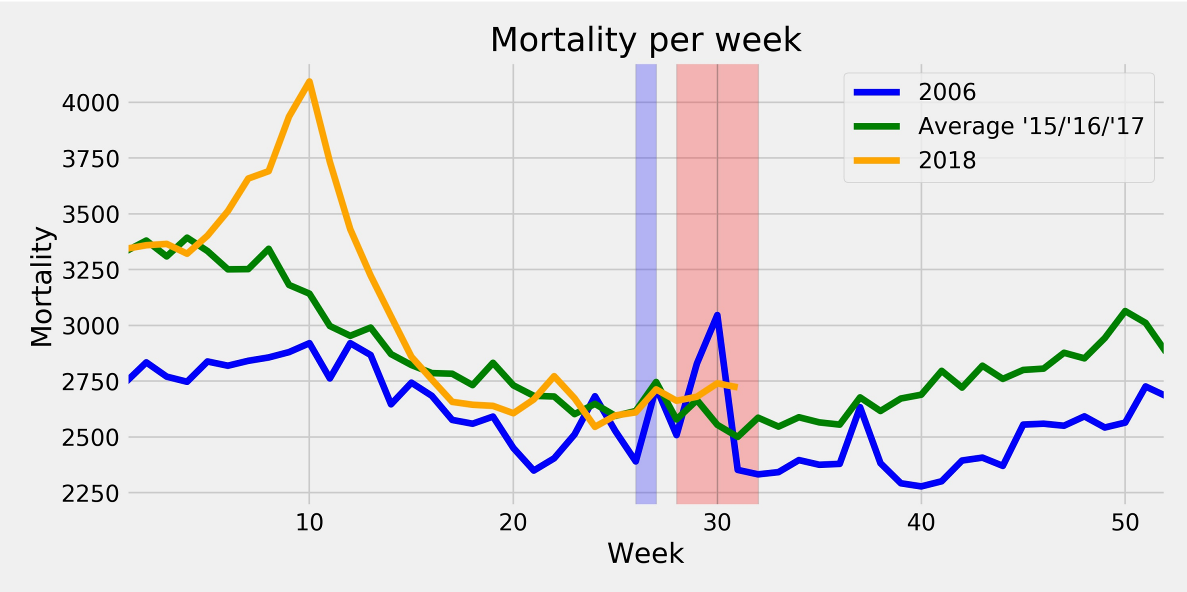

Figure 1.2: Differences in mortality per week: during the heatwaves and the average values for 2015, 2016, and 2017. Data of the graph available at Centraal Bureau voor de Statistiek (2018a)

The next Figure 1.3 demonstrates what age groups were affected the most. According to the data, the most affected are citizens of age from 65 to 79 and over 80 years old.

A couple of valuable insights can be derived from these graphs. First, talking chronologically, there is a significant difference between the average values and 2006. The heat waves of 2006 were one of the most devastating in the Netherlands. The peak of mortality in 2006 is 3046, whereas on average it is 2554. The difference is striking, and it is almost 500. Second, it is hard to compare 2018 values with others. At first glance, it seems that the increase in mortality was not significant as it was for 2006. However, a remarkable event preceded the 2018 heatwave. An epidemic of flu significantly affected the vulnerable population during the winter period (Centraal Bureau voor de Statistiek, 2018b). An increase can be seen from week 3 to a peak in week 10. Extra graphs emphasizing these points can be found in Appendix A.1.

At the moment of finalizing this study, another heat wave reached the Netherlands. The effect of it was so heavy that many Dutch cities renewed the historical maximums in temperature. For instance, a new record in 39.2 °C was recorded at the Gilze-Rijen weather

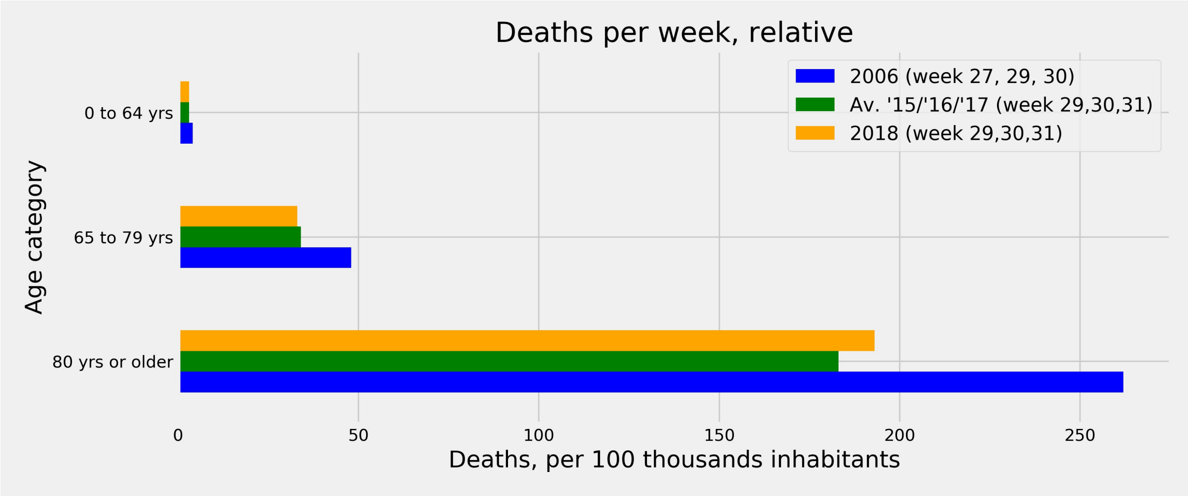

Figure 1.3: Comparison of deaths by age group. Data of the graph available at Centraal Bureau voor de Statistiek (2018a)

station near Breda. It lasted for only six days: from June 22 to June 27, but, again, caused an increase in mortality. On August 9, preliminary numbers on mortality were published by Centraal Bureau voor de Statistiek (2019a). Figure 1.4 demonstrates morality values of three heatwaves: 2006, 2018 and 2019. As can be seen from the graph, the numbers of 2019 almost reached 2006 at its peak.

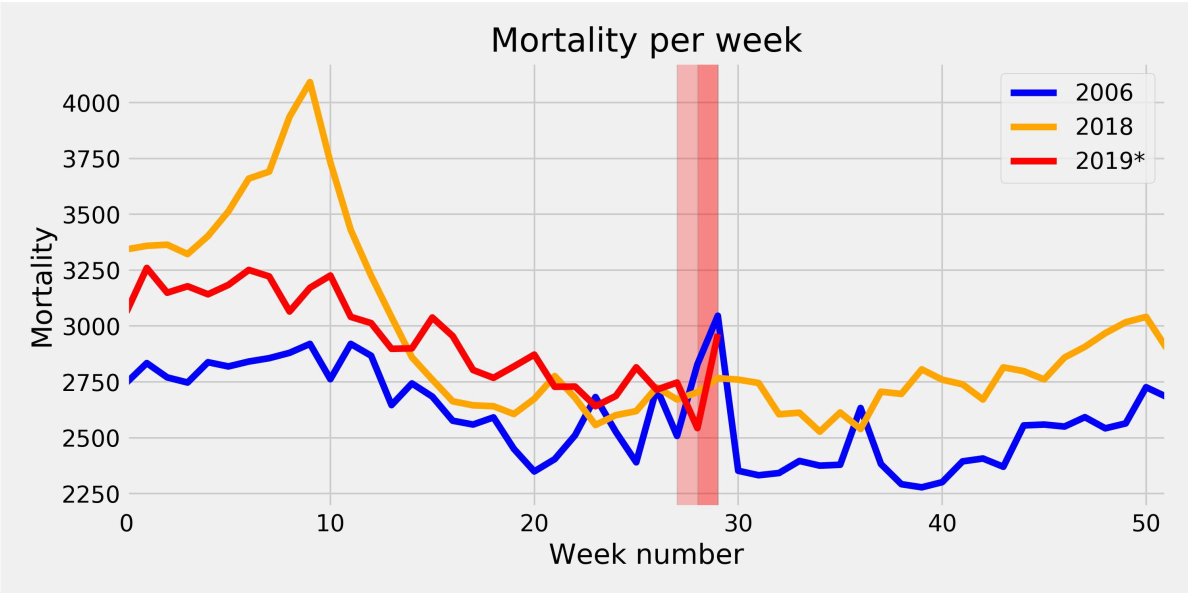

Figure 1.4: Preliminary numbers of 2019 heat wave mortality in comparison with 2006 and 2018 values. Data of the graph available at Centraal Bureau voor de Statistiek (2019a)

According to the data, the age group affected the most is above 80 years old. There is a minor increase for the age group of from 65 to 79 years old as well, but not as significant as for the first one. Also, CBS reported that there is a difference in mortality between the regions of the Netherlands. Due to some reasons, it is higher in the east (corresponding graphs can be found in Appendix A.1).

Important to mention that mortality is not the only way to measure of the damage made by heat waves. Besides the fatal outcomes, extreme heat significantly worsens person health, but keep him or her alive. Thus, these numbers clearly indicate the importance of the problem, but not cover the whole broad range of consequences.

The Netherlands recognizes the importance of the extreme heat threat. In 2007, the Netherlands National Institute for Public Health and the Environment (RIVM), an independent agency within the Ministry of Health, Welfare and Sport, created the first version of a National Heat Plan (Rijksinstituut voor Volksgezondheid en Milieu, 2011). According to the RIVM, long-lasting periods of hot weather can cause serious health problems from severe respiratory distress to a heart attack. The research shows that certain population groups as more vulnerable to such issues: people who have chronic illnesses or obesity, and children. The National Heat Plan defines the following:

- Conditions under which there is a threat;
- Responsible actors to whom the threat should be communicated;
- Actions that should be taken by the citizens to minimize the consequences.

The current version of the plan 27 °C daily temperature for four days threshold. If the forecast of the Royal Netherlands Meteorological Institute (KNMI) indicates so, the plan must be activated, and all key actors are informed. From the perspective of the actions, the RVIM has a series of straightforward recommendations such as: drink more water, be in the shade, ventilate living places. Besides that, the RIVM also recognizes the importance of taking care of each other by neighbors as an essential action. The last action in the recommendation list emphasizes it with a statement to pay extra attention to people in need.

The climate puts significant pressure on the population of the Netherlands. The fact that the population of the Netherlands is experiencing ageing is amplifying the importance of this problem.

As was stated earlier, metropolitan areas are experiencing numerous shocks and stresses. According to the report prepared by 100 Resilient Cities network one of the largest cities of Netherlands - The Hague, is expected to suffer from the following continuous stress - climate change and corresponding shock - extreme heat (Resilient The Hague, 2018). The Municipality of The Hague recognizes that climate change will put significant pressure on vulnerable individuals communities. According to the studies of Klinenberg (2003), Swalheim and Dodman (2008), Zaidi and Pelling (2015), some specific socio-demographic attributes of the vulnerable population makes it hard for them to adapt during extreme events. The part of the policies of the Municipality of The Hague is built around its vulnerable citizens (der Sangen, 2018) what makes us think that it is especially relevant to get a better insight into the distribution of such vulnerable individuals and communities across the city.

According to Resilient The Hague (2019) The Hague is a segregated city, and it can amplify these problems. The population structure of different districts and neighborhoods varies widely in terms of income, employment, educational level, and health. Low rates of social integration can prevent acts of community response during crises. From the technical side, to improve the policy-making process, the city of The Hague together with CBS launched an Urban Data Center Den Haag (Centraal Bureau voor de Statistiek, 2017). The data collected by the center is open for public use. Thus, researchers and scholars can use it to develop data-informed solutions for combating extreme heat.

1.2. Research gap

There are a variety of conceptual tensions regarding the concept of urban resilience starting from its definition (Meerow et al. (2016) counted at least 25 definitions that can be met in major scientific publications for the last 15 years) to the unified version of its assessment (Bourdic et al., 2012) and the type of institutions that should be built to ensure its implementation (Zaidi and Pelling, 2015). Such discrepancies, partly caused by multidisciplinary nature and novelty of the concept, considered by some as benefits since they allow stakeholders to come up with their own understating of resilience (Engle, 2011). However, many other argue that it creates obstacles for practitioners and make them doubt the ideas behind urban resilience (Chelleri et al., 2015; Jabareen, 2013; Leichenko, 2011; Meerow et al., 2016). The questions like ”resilience where” and ”resilience of what” (demarcation of urban system), ”resilience to what” (identification of external factors), ”resilience for what purpose” (defining goals and objectives) are often left without unambiguous answers. Thus, there is a need for a clear, but flexible specification of urban resilience concept in regards to a particular shock: extreme heat, in a specific city: The Hague. The word flexible here refers to the possibility of using it for either similar nature-based shocks (flooding, earth earthquake, etcetera) or other cities (in the Netherlands or any other country).

Apart from it, most of the frameworks available operate on a highly abstract level, allows multiple interpretations and lack of practical utility. Resilience criteria built by (Sharifi and Yamagata, 2016) to measure urban resilience seems to cover all possible aspects of urban life and do not provide insight to resilience itself and rather just resembles general impact assessment models. We argue for a conceptual framework that should be, on the one hand, built on the solid theoretical foundations (to preserve major beneficial ideas of resilience), but on the other side be operationalizable.

Another aspect of urban resilience that requires research is the solutions realize it. Typically resilience scholars briefly mention possible measures that might be helpful in one or another case Chelleri et al. (2015); Jabareen (2013). It is especially the case while dealing with non-engineering, but social systems: vulnerable individuals and communities. Here we are seeking for explicit theory-based or empirically proven solutions that can become parts of urban resilience framework.

There are different ways of how the assessment of conceptual framework can be done. One is to try innovative ideas in a particular city like it was done by Crowe et al. (2016); other is to identify proxies and formulate variables that sufficiently reflect on the main components of the framework Bourdic et al. (2012). Joined with open data and open-source ML packages, the second provides opportunities for future reusability of results. Moreover, empirical findings regarding the usefulness of certain solutions should be carefully assessed with the data before actual implementation. However, currently, urban analytics is still in its early stages (Batty, 2019). Thus, there is a lack of knowledge about what algorithms should be used for comprehensive analysis of city problems.

1.3. Research questions

The research gap discussed above pushes us to formulate two equally important research questions. The first one oriented towards bridging the gap between **theoretical** foundations of urban resilience and its implementation, whereas the second one relates to the **methodology** and methods.

Due to a variety of discrepancies, the actual practice of it is held up. An extensive literature review should be conducted to: identify and fix the main components of it; formulate operationalizable frameworks; and, finally find potential solutions. Since the focus of the study is on a specific shock: extreme heat, the notion of it should be included in the formulation. Thus, the corresponding research question can be formulated as follows:

> Research question 1
>
> What concepts of urban resilience are useful for planning in case of extreme heat?

The second research question reflects on the lack of knowledge about using machine learning algorithms for city problems (apart from being almost absent in urban resilience literature):

> Research question 2
>
> Using what machine learning algorithms these concepts of urban resilience can be operationalized?

The search is limited by two factors: first, the availability of data, and second, the nature of the task. The next step after the concepts were fixed, and corresponding proxies formulated, is to identify the data that is sufficiently covering them. Based on this, the algorithms can be selected. The prior hypothesis is to use three types of algorithms: unsupervised learning, regression analysis and deep learning for time series forecasting.

The final question covers the **case study**: the city of The Hague. Now we know the concepts and how they can be operationalized. More specifically, the concept of urban resilience is tightly connected to the concept of vulnerability. Besides, resilience capacities depend on different types of infrastructure solutions (emergency, green, and social). The following two questions reflect on it.

> Research question 3
>
> How vulnerable individuals and communities are distributed across The Hague?

> Research question 4
>
> How various areas of The Hague differ in terms of resilience capacities?

The final research question summarizes the findings of the previous ones. It formulated rather broad and ambiguous, but on purpose. Conventionally, when dealing with such complex concepts as urban resilience, an unified metric is proposed. For example, a single livability index based on 96 variables. Such an approach makes it easy to operate with since a single number is assigned to each neighborhood. However, the weights, the contribution of each factor are highly arguable and can cause a lot of debates and wrong decisions. Opposite to it, I argue for a more transparent approach. It should be clear what are the variables that make the city resilient. Each of the interventions and their contribution should be discussed separately. Only then a more comprehensive assessment and resulting from it policy change can be proposed.

> Research question 5
>
> Is The Hague resilient to extreme heat?

1.4. Research flow

The thesis consists of 5 chapters. The first chapter consistently introduces the problem of the study: from urbanization and climate change to extreme heat and its effect on vulnerable population. Then, it discusses a promising approach to tackle it: urban resilience in combination with open data and machine learning. This discussion is followed by the description of the case study: the Netherlands and the city of The Hague. After that, the chapter describes the research gap that needs to be filled: there is a lack of clarity and practical utility in urban resilience literature. Besides, it is not clear what machine learning algorithms can be used for assessment of urban resilience. The research gap is followed by the research questions. The chapter ends with the research flow of the study and a brief summary of the points made.

The second chapter is dedicated to an extensive literature review. It covers urban resilience concept, essential aspects of extreme heat threat and proposes theory-based and empirical solutions to tackle it. The chapter ends with the formulation of two conceptual frameworks that should be further operationalized. Thus, it answers the first research question.

Chapter three illustrates how a transfer from a conceptual to a computational framework can be made. To do so, it suggests a set of variables, lists available data sets and proposes four types of algorithms that can be used in the analysis. These findings summarized in a computational framework that can be used for assessment of resilience to extreme heat. That is, this chapter seeks to answer the second research question.

The fourth chapter presents the results of the case study analysis. The algorithms are sequentially applied to, first, map vulnerability, and, second, analyze a set of capacities. Therefore, it provides answers to research question 3 and 4.

The final fifth chapter concludes the study with explicit answers to the posed research questions. In addition, it discusses the benefits and limitations of operationalizing urban resilience with machine learning. After that, it draws the avenues of future work. It also reflects on the case study findings and presents a policy advice for the city of The Hague.

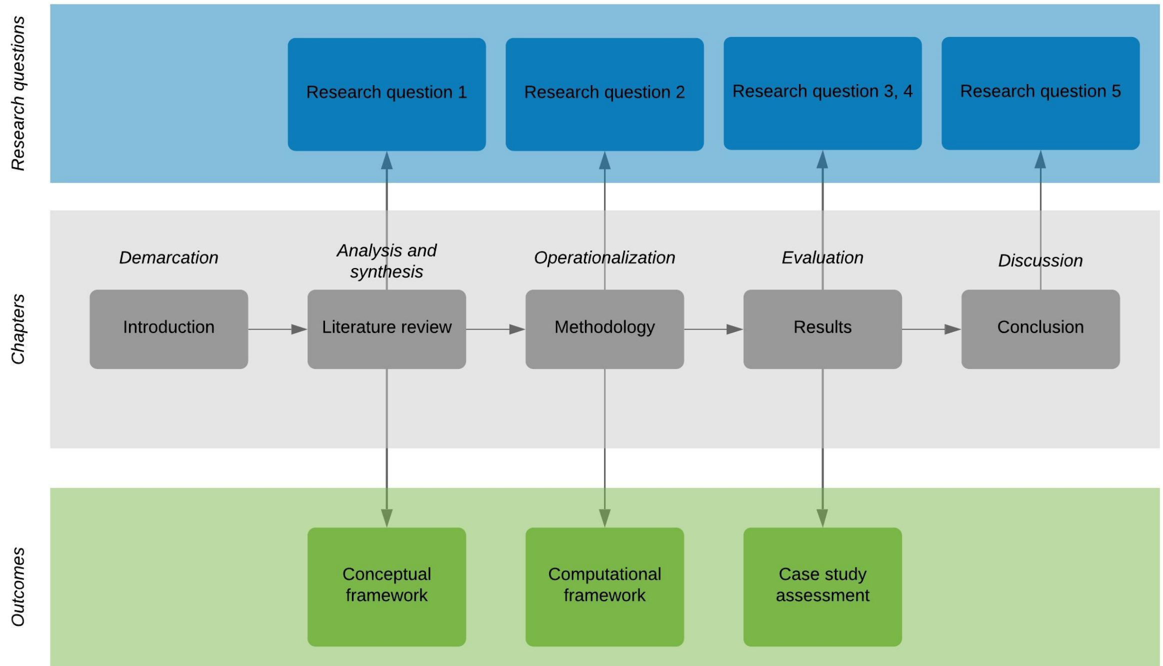

Figure 1.5: Simplified representation of the research flow

1.5. Conclusion

Rapid urbanization is a complex process that puts a lot of stress on modern society. Climate change and extreme heat put vulnerable population of cities at risk. A phenomenon known as urban heat island effect intensifies it by keeping higher temperatures in certain areas. Take together, they require a better approach rather than conventional urban planning.

The concept of urban resilience seen by scholars and policy-makers as an opportunity to create safer cities that can sustain under a wide array of shocks and stresses. To operationalize it and promote further reusability, we propose to use open data with open-source machine learning packages. The city of The Hague, along with other cities of the Netherlands, just recently in July 2019 experienced disruptive effects of the heatwaves. Now, the Municipality of The Hague is seeking for a way to make the city more resilient to extreme heat. However, due to a variety of tensions around the resilience concept, implementation rather tricky. Besides that, there is a lack of examples of how its complex and often ambiguous frameworks can be operationalized. This thesis aimed at bringing more clarity and practical utility to the field of urban resilience. It proposes a computational framework that can gain insight on how to protect vulnerable individuals and communities from the danger of heat waves.

2

Literature Review

This chapter aims to provide a background and lay down the foundations for this study. First, it discusses recent findings in urban resilience literature: what are the main concepts of it and how they can be applied. Second, it briefly introduces essential aspects of an extreme heat threat: the heat wave phenomenon, the urban heat island concept, and defines people who are vulnerable to extreme heat. Finally, this chapter describes different types of infrastructure solutions that can be used to combat extreme heat threat. Thus, this chapter seeks to answer the first research question.

2.1. Urban resilience

During the last decade, the concept of urban resilience has received attention from international organizations, city planners, and academia. The United Nations connected it with goal number eleven and a specific unit of it - UN-Habitat is actively involved in its implementation (United Nations, 2018b). The World Bank is also facilitating the process of making cities more resilient (World Bank, 2019).

Many cities have started to build up their strategies around it. Now, they are aimed at becoming more resilient against a variety of threats: earthquakes, climate change, terrorist attacks, etcetera (100 Resilient Cities). One of the initiatives - 100 Resilience Cities network, united more than 30 cities from all over the world. The member cities identified the challenges that they have experienced in the past or likely to face in the future. The promise is that these problems can be successfully addressed by urban resilience.

However, from the research side, the number of approaches to it and corresponding frameworks make it hard to put urban resilience into practice (Jabareen, 2013; Leichenko, 2011; Meerow et al., 2016). Moreover, there is an opinion that the theory is still relatively vague and understudied and, therefore, practitioners cannot really benefit from it (Chelleri et al., 2015).

There is a need for a clear but flexible specification of the urban resilience concept. The word flexible here refers to the possibility of using it for either similar shocks and stresses or being applied by different cities around the world.

In order to bridge this gap, I propose to conduct an extensive analysis of the concept. The analysis is divided into 8 steps, represented in Figure 2.1. The first three steps are aimed to decompose the concept into the central parts: what are the types of resilience, related concepts and components, resilience definitions. The fourth one dedicated to the discussion of conceptual frameworks: how the ideas of previously identified elements can be put into practice? The fifth step provides an overview of assessment models. On contrast with the conceptual frameworks, the assessment models operate with specific proxies and indicators. They propose ways to ”measure” or ”quantify” urban resilience. The sixth step gives a short overview of the practice of applying urban resilience. In step seven, I discuss the critique of urban resilience: is it a positive concept, and if not, what are the drawbacks of it? Conventionally analysis ends with the summary of findings.

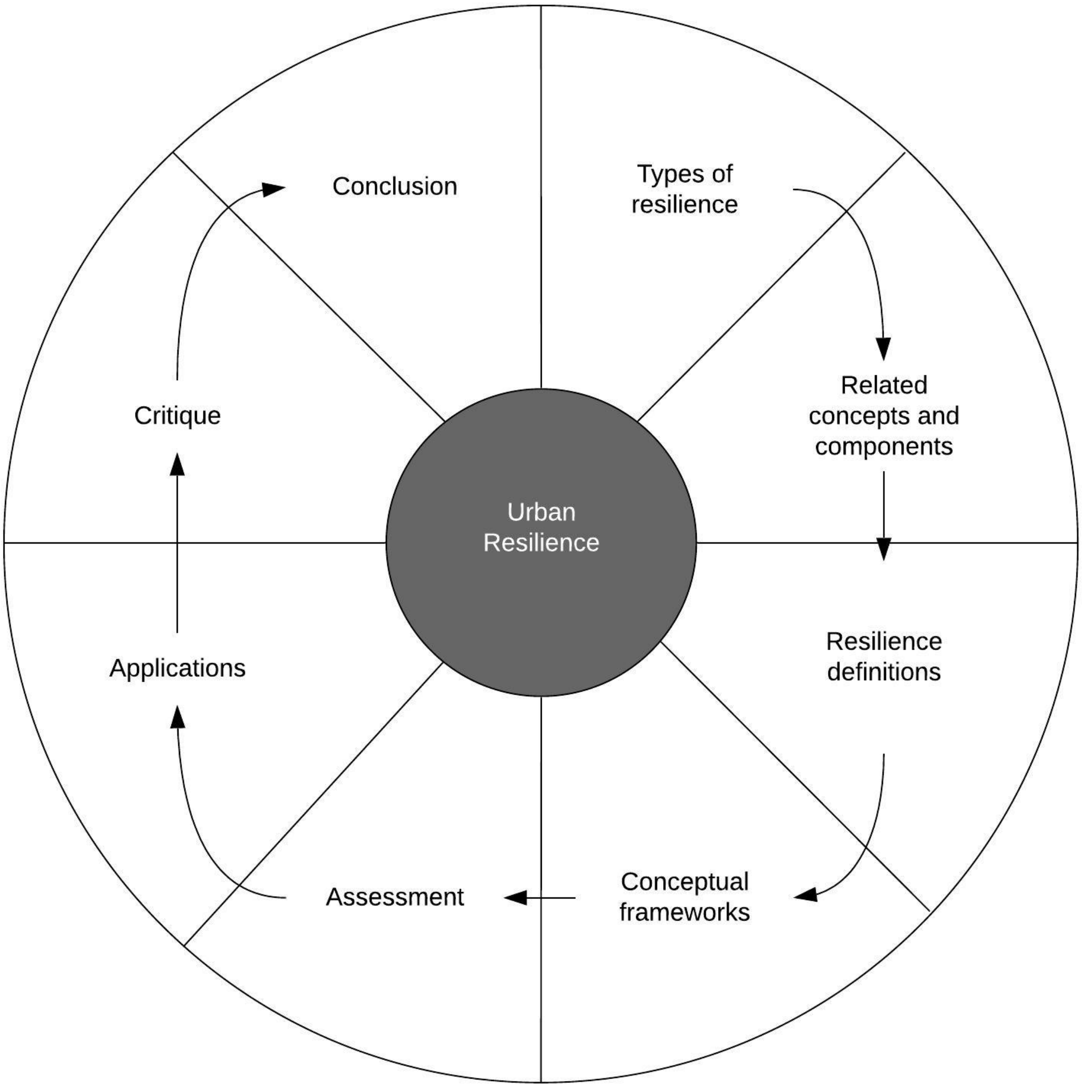

Figure 2.1: Schematic overview of the content of the urban resilience section

2.1.1. Origins and current stage

Let us start by introducing the definition of resilience that can be found in one of the most popular dictionaries - the Oxford Dictionary. Originated from a Latin word *resilio* - ”to bounce back,” nowadays, resilience defined by the Oxford Dictionary of English (Stevenson, 2010) in the two ways:

Definition 2.1.1.1 **Stevenson (2010)**

Resilience is the capacity to recover quickly from difficulties; toughness.

Definition 2.1.1.2 **Stevenson (2010)**

Resilience is the ability of a substance or object to spring back into shape; elasticity.

These definitions highlight the essential parts of resilience that also can be met in the scientific literature. The first one introduces a set of 4 interconnected concepts: *capacity*, *recovery*, and *time* and *difficulties*. Capacity here is a way to measure resilience, recovery - a type of response to external shock, or to difficulty, the time represented by a word *quickly*: if something is resilient it must recover quickly. However, the synonym that is given: toughness, resembles *robustness* more than resilience (see Figure 2.2). **Robustness perceived by academia differently from resilience** (Scholz et al., 2012). Something robust, not necessary resilient, and vice versa.

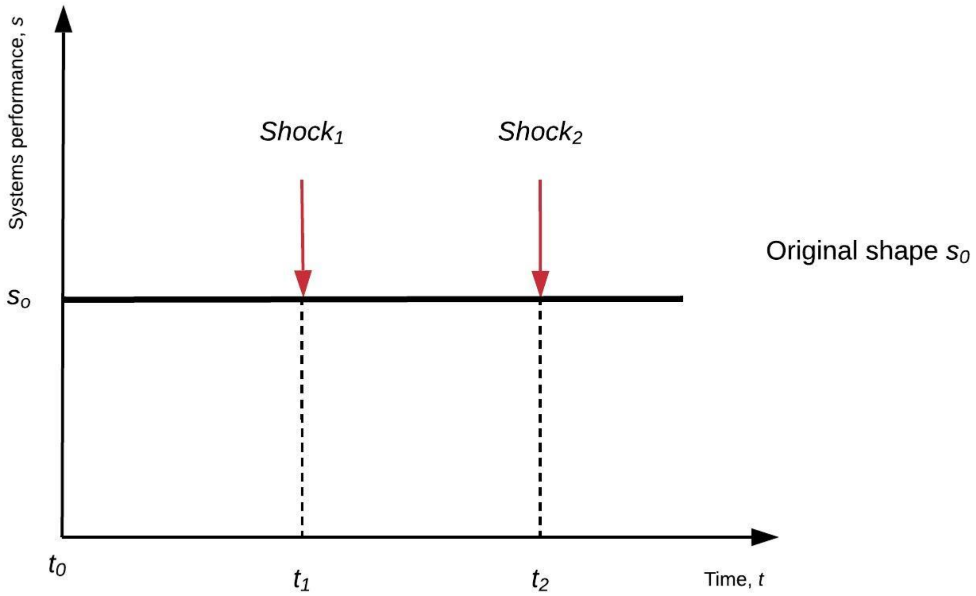

Figure 2.2: Performance of robust system under a set of shocks

> Consider an example of the track on a railway or railroad. It is relatively robust. Daily, it experiences a lot of stress made by trains and maintains the same shape. A single piece of it may preserve the same shape even under a TNT explosion. However, if you bend it (of course, with the usage of special equipment), it breaks. Thus, it has high robustness, but a low resilience. An opposite example is a bamboo tree. Bend it and release and it will go to the original shape. However, try to hit multiple times with a hammer, and it will break.

The second one operates with two essential concepts: *spring back* and *shape*. Spring back refers to a type of actions that returns an object to its original shape. In scientific literature, this shape called an equilibrium and resilience that operates with one possible shape - one-equilibrium resilience. An example of such a resilient system is a rubber. It can be easily stretched and then go back to the original shape.

The visual representation of these definitions is given by Figure 2.3. Let us explain it in details. The system of interest was in its original state or equilibrium s_o at time t_0. It faced a shock at time t_s, and after that, the performance of the system has started to deplete until it reached the minimum s_b at time t_b. That is a starting point for a system to bounce back to the original shape. In total it took $t_r - t_s$ to recover.

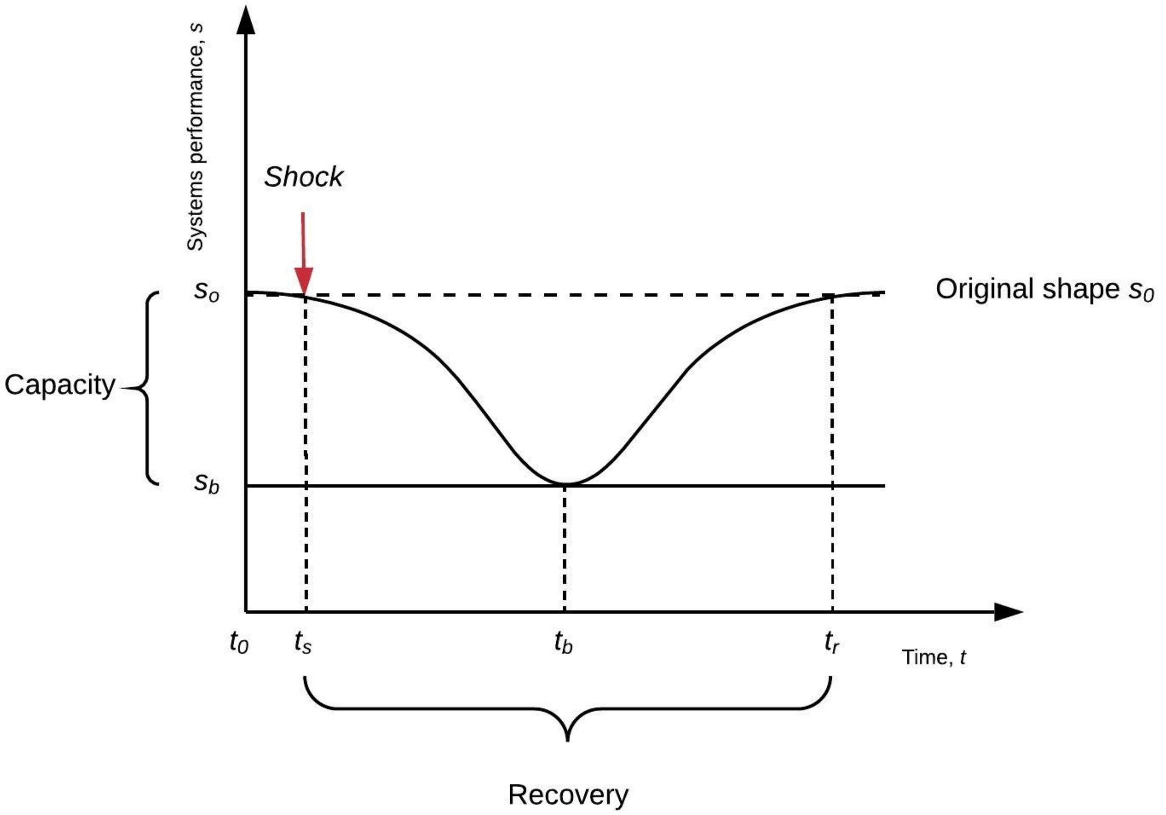

Figure 2.3: Performance of resilient system under a shock

The above-described definitions provide us with a general understanding of resilience. However, when used in scientific publications, the concept has a much broader meaning.

The opinions on the roots of resilience vary. One named psychology and physics (Barata-Salgueiro and Erkip, 2014; Maruyama, 2016). Psychology treats resilience as an ability to recover from a shock or a trauma (Fletcher and Sarkar, 2013), and this partly resembles the Definition 2.1.1.1.

> For example, a person was bitten by a dog in childhood. Such an event can be treated as a shock or trauma. If a person has psychological resilience, he or she can go back to the original state when there was no fear of dogs.

Physics or more specifically - materials science, consider resilience as an ability of an object to go back to its original state after being stretched (Treloar, 1975). Such an approach is similar to Definition 2.1.1.2 and hardly requires further elaboration to cover the behavior of physical systems.

> Consider a physical object - elastic band. This object has been stretched or deformed but below its elastic limit. Therefore, it will not be torn. If this elastic band is of good quality, it will get to its original shape without any signs of distortion.

However, most of the scholars refer to Holling (1973) with the theory of *ecological systems* as the primary contributor to resilience in its modern understanding (Ahern, 2011; Leichenko, 2011; Meerow et al., 2016). In this early work, Holling described an important **difference between resilience and stability** in ecological systems. For example, in a population of fish, owls, or mice. These systems experience a variety of external stresses and shocks that can lead to the extinction of the species. From his point of view, resilience is an attribute of a system that determines the capacity to absorb changes and persist, whereas stability represents the speed of returning to the original state. The faster the system goes back to the original, the more stable it is.

Definition 2.1.1.3 **Holling (1973)**

Resilience is the ability of these systems to absorb changes of state variables driving variables and parameters and still persist.

However, such an approach considers only **one equilibrium** state, for example, when the number of predators in the system balances the number of preys. As a result, it criticized for being static and does not suit for more complex socio-ecological systems (Barata-Salgueiro and Erkip, 2014).

In the next work, the author took a broader perspective (Holling, 2001). He suggested to consider human-nature interactions as complex adaptive system (Levin and Lelvin, 1999). This implies having a large system that consists of multiple subsystems of different nature: natural, human, combined human-nature, and socio-ecological ones. These subsystems interact through a *panarchy* - a combination of adaptive cycles of growth, accumulation, restructuring and renewal (Holling, 2001). Similar to the early works, to sustain the larger system should persist and be resilient. **Resilience now is a measure of vulnerability**: how vulnerable is the system to unexpected events and changes? However, the notion of **adaptation** puts a new perspective on the way how it can be achieved. The system must continuously adapt to sustain.

The following publication of Folke et al. (2002) emphasized the importance of **learning and development** as a path to become more resilient (see Definition 2.1.1.4). The work proposes to cover the concept of resilience from multiple perspectives. First, resilience relates to the magnitude of shock that the system can absorb. Second, it is a degree of self-organization. Third, it characterizes the capacity to learn and develop. Thus, starting from rather a static concept with one equilibrium state that must be preserved, resilience became a conceptual framework. Now it prescribes what the desired system attributes are and how they can be achieved (Folke et al., 2002).

Definition 2.1.1.4 **Folke et al. (2002)**

Resilience is the capacity to buffer change, learn and develop.

However, what always in place is **importance of uncertainty**. Whatever the type of the system is: natural, ecological, or socio-ecological, it faces multiple challenges that have not been expected.

The ”branch” of resilience discussed above known as *socio-ecological resilience* (Barata-Salgueiro and Erkip, 2014; Meerow et al., 2016). The researchers of this branch formed Resilience Alliance (Walker and Salt, 2012) - an interdisciplinary research network that addresses the problems of sustainability via the panarchy model and resilience thinking.

Apart from physics, psychology, and theory of ecological systems, there are other opinions on the roots of resilience that worth to mention. Ernstson et al. (2010) highlighted complex adaptive systems theory with its important concept of self-organization, whereas Simmie and Martin (2010) emphasized the contribution of complexity theory and the corresponding idea of evolution.

Given these promising foundations, the idea of resilience has started to spread across a variety of disciplines. Now applications of resilience can be found in engineering (Fiksel, 2006), energy systems(McLellan et al., 2012), economics (Rose, 2004; Simmie and Martin, 2010), political sciences and institutions (Zaidi and Pelling, 2015), security studies (Coaffee, 2016), etcetera. But most importantly, city scientists have started to use the concept of resilience and corresponding ideas. To avoid ambiguity, let us note that when applied to cities and urban areas, the concept named *urban resilience*. Without a doubt, urban resilience as an attribute for a city seems to be promising and desired. Using definitions mentioned above, a resilient city can sustain, adapt and develop under a variety of unexpected shocks and stresses. The city is conventionally perceived as a combination of different subsystems. As a result, different scholars proposed to consider resilience towards either a specific threat or resilience of a specific subsystem. Leichenko (2011) proposed to categorize urban resilience literature into four main groups: urban-ecological resilience, urban hazards and disaster risk management, urban and regional economies, urban governance and institutions. In addition to this, Meerow et al. (2016) highlighted studies on social resilience and vulnerability (Cutter et al., 2003). Remarkably, the vast majority of publications on urban resilience is dedicated to climate change and cities (Leichenko, 2011; Wardekker et al., 2010).

As can be seen from the discussion above, the concept of resilience is relatively new and complex. In various fields of studies, it has been developing differently. Physical systems are still can be relatively well described by Definition 2.1.1.2. In socio-ecological systems, the concept slowly progressed from a static ”persist to sustain” with one equilibrium to more comprehensive, but more ambiguous ”absorb, adapt and develop.” The complexity that has arisen can be explained by a **mismatch between the system and the model** (Ackoff and Gharajedaghi, 2003).

For instance, try to explain an engineering system - city sewerage, with an ecological model - predator-prey relations. Of course, some similarities between these two can be found. The resilience of the sewerage system can be explained via ecological resilience - it must be able to handle the designed amount of rainfall and sanitary. However, when adding humans to it, it stops being just engineering systems and requires more a comprehensive model.

Besides, applied to different systems of interest, the corresponding concepts of absorption, adaptation, and development will be inevitably explained differently. On top of it, the researchers have their specific domain, and even by stating that the city requires a holistic approach, there is no consensus on such a model.

Thus, there a lot of discrepancies about the key concepts, their explanations, and definitions. Some argue that it is favourable. Resilience is a flexible concept that can be taken onboard by scholars with different backgrounds and scientific domains and provide flexibility (Brand and Jax, 2007; Vale, 2014). But most of the researchers see it as an obstacle (Chelleri et al., 2015; Jabareen, 2013; Meerow et al., 2016).

2.1.2. Types, concepts and definitions

To find common ground, it is essential to, first, decompose the concept of urban resilience into parts, second, highlight similarities and differences, and, finally select the ones that are useful while dealing with extreme heat, and, moreover can be operationalized. To do so, let us start the first name the sources of discrepancies or conceptual tensions:

1. What is the scientific domain behind the research and what model of a city it propagates?
2. What type of resilience it implies?
3. What is the system of interest?
4. What are the key concepts and components?
5. What is the definition of urban resilience?

Similar approach has been taken by Meerow et al. (2016) and Maruyama (2016). Meerow et al. (2016) proposed to differentiate the tensions regarding the definition, characterization of ”urban,” (what is the city), the notions of resilience (key concepts), whether resilience a positive concept, pathways to urban resilience, and, finally understanding of adaptation. Maruyama (2016) introduced a taxonomy of urban resilience. The taxonomy consists of the type of shock or stress, characteristic of the target system, and, the type of recovery. Figure 2.4 is a schematic representation of how the concepts of urban resilience interconnected. The directed arrow denotes ”A defines B” type of relationships.

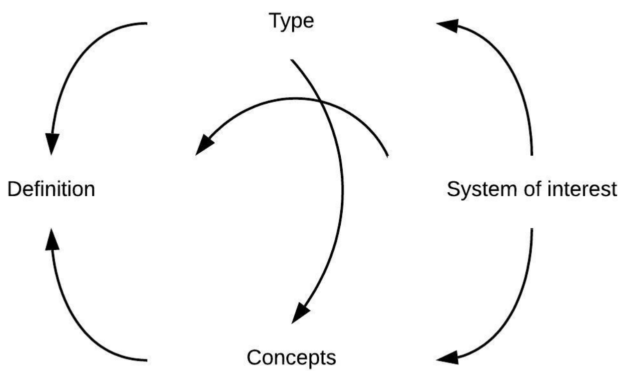

Figure 2.4: The interrelations between the components of urban resilience concept

Scientific domain and model of city

Now, let us discuss the proposed classification in details. The first discrepancy occurred due to various scientific domains and the implied model of a city. A scientific domain is a combination of theories that are similar in some sense. They can have the same object of studies, a problem of interest, or phenomenon under consideration. Explicitly or not, when discussing resilience, the domain uses a particular model of the city. It may emphasize only certain parts of it. For example, the problems of urban poor need to be addressed through a set of subsidies: a city as a political system. Or the traffic congestion is the most import and need to be tackled by traffic engineers: a city as a transportation system.

Opposite to a conventional standpoint, many scholars argued in favour of using a holistic model of a city. For example, urban ecology percieves a city as complex adaptive system (Wu, 2014). The elements of the system combined into networks of nodes. The material exchange between these nodes is made via flows. Hazards and disaster management tends to see a city as a system of failing infrastructures (Vale and Campanella, 2005). If an extreme event occurs, it will lead to, first, a collapse of one infrastructure: telecommunication system, for example. From there, other infrastructure objects will be damaged: hospitals, fire stations, etcetera), thus, creating a cascade effect. Urban and regional economies while considering urban resilience prefer to treat a city also as a complex adaptive system. The city can evolve under shocks and stresses and reach new, unknown before, states. Governance and institutions literature stresses the importance of how the political system of cities is organized. Institutional arrangements, social contracts are the main instruments to build up resilience against climate change, for instance (O'Brien et al., 2009; Ostrom, 2010). Climate change science usually operates with a socio-ecological model of the city where the urban population alongside businesses pollute the environment (Wardekker et al., 2010). The final domain: urban theory, examines the city as a social system. The remarkable work by Jacobs (1961) lay down the foundations of social resilience. She emphasized the importance of diversity in all its manifestations: buildings types, businesses and people located and living nearby. There are other domains the are using the concept of urban resilience, for example, landscape ecology (Botkin, 1990). However, most of the work has been done in the domains that were mentioned earlier.

Thus, the classification proposed by Leichenko (2011) can be enlarged from 4 to 6 categories and summarized with the Table 2.1.

№	Scientific domain	Model of city
1	Urban ecology	Complex adaptive system
2	Hazards and disaster risk reduction	Combination of critical infrastructures
3	Urban and regional economies	Evolutionary market
4	Governance and institutions	Political system
5	Climate change and adaptation	Socio-ecological system
6	Urban theory	Social system

Table 2.1: Classification of resilience by scientific domain and corresponding model of a city

Type of resilience

The second discrepancy follows from the first one. Resilience is a flexible theory that can be taken onboard by scholars with different backgrounds and scientific domains (Brand and Jax, 2007; Vale, 2014). Consequently, by looking at it from a different angle, the type of resilience somehow is introduced. Ideally, together, the theory and the model should define the type of resilience. For instance, if resilience is an engineering one, then, the authors will stress the importance of infrastructure systems such as electric power system or a sewerage network.

Engineering resilience is often associated with a concept of robustness and treated as an opposite term to vulnerability (Engle, 2011). The system should have a certain capacity to absorb and then recover to its original state.

For example, after an earthquake (shock) a set of power stations is off (see Figure 2.5). However, a city has a set of emergency generators for such a case. In resilience literature, this feature of the city system called *redundancy*. There must be a certain number of redundant elements that should be in place to ensure systems performance around a desired threshold. However, these generators cannot provide the full capacity, and performance of the city goes down (s_b). Some homes or even the whole neighborhoods do not have access to electricity. Further, a specific agency within the city need to care of repairing and after a while (recovery time) initial electricity supply will be restored (s_o).

Socio-ecological resilience is a more complex concept and it focuses around the interactions between humans and nature, but oppositely to engineering resilience, it goes beyond of being a ”flipside” of static vulnerability concept (Chelleri et al., 2015). To combat a shock socio-ecological system can *adapt* and reach a new state that was not anticipated before (s_a) (Folke et al., 2002). The shock here plays the role of a trigger that launches the unpredictable process of evolution. Usually, the new state of the system treated as a desirable one.

Consider the following example. The population of a certain area used to flooding. City managers took the care of this threat by heightening up the dykes. However, this time, the waves are much higher, and the barriers are not that useful anymore. The water reaches a critical point, and the situation quickly becomes life-threatening. This is the moment when individuals and communities are forced to adapt to survive: climb on the roofs, try to escape from the area on boats.

As can be seen from this example, it is harder to map out the path towards a new state, since the fact the water will cross the barriers was not anticipated.

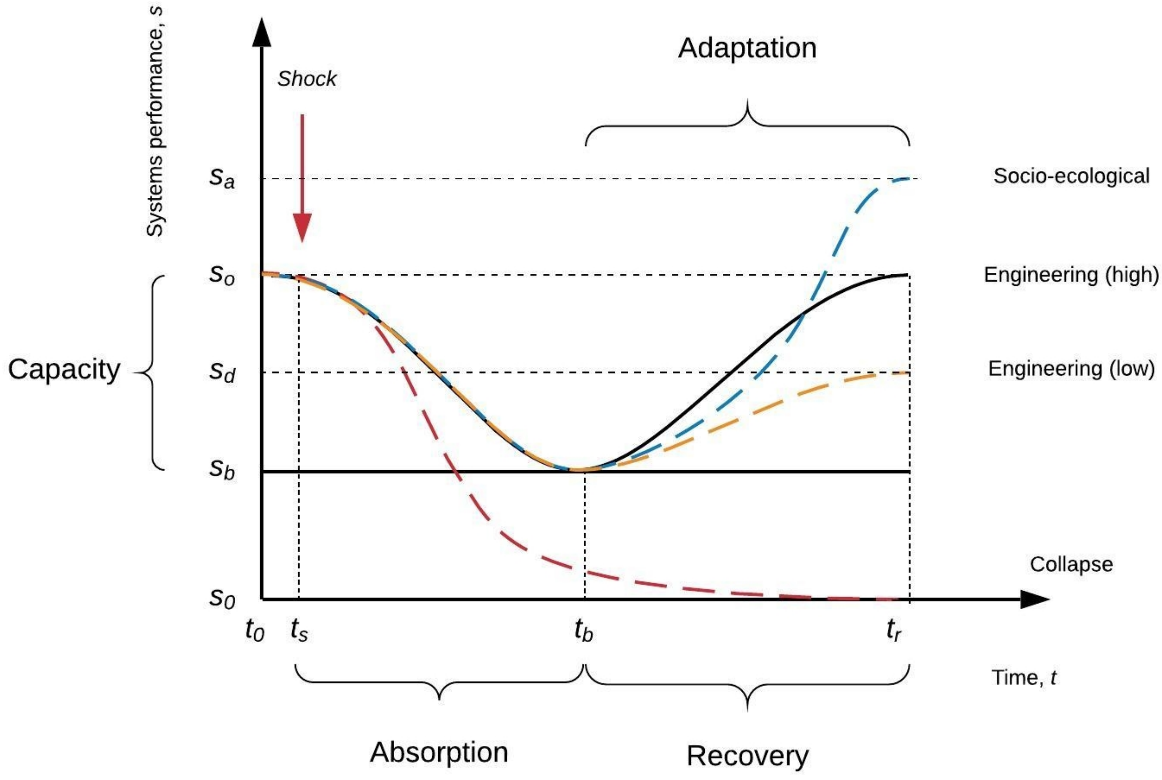

Figure 2.5: The performance of a system under a shock depending on the type of resilience

Social resilience stresses the importance of social relations and power (Friend and Moench, 2013). Since the city can be thought just as large human settlement, the questions of equity and diversity can be central in defining social resilience. When facing a threat, what population groups will be affected the most? How to ensure that vulnerable citizens such as urban poor, sick, and elderly will receive help?

Economic and governmental resilience also operate on the top of the socio-ecological concepts (Barata-Salgueiro and Erkip, 2014; Duit et al., 2010). However, the perspective of Duit et al. (2010) is more pragmatic. The authors proposed to consider resilience as a problem of social framing. Different actors have various preferences that should be satisfied and the trade-offs between them addressed.

Besides that, there is an important distinction between *general* and *specific* resilience. Having general resilience implies that the urban system should sustain no matter what challenge will appear: flooding, disease outbreak, or cyberattack (Carpenter et al., 2001; Maruyama, 2016). Opposite to it, specific resilience prescribes city to be to handle only a particular type of problem - natural disasters, for example (Carpenter et al., 2001; Chelleri et al., 2015).

One more distinction can be made between *resilience of a city* and *resilience in city* (Ernstson et al., 2010; Leichenko, 2011). In a way, these two can be treated as ”scales.” The first is the lower one, and, the second is the higher one. Resilience in cities refers to the interactions between urban form, land-use, and, population, on the one hand, and ecological processes that arise from these interactions, on the other hand. Usually, the biggest attention is given to these elements and intersections placed in a specific city. Therefore, the analysis implies that cities are closed systems. On the contrary, the resilience of a city suggests considering a bigger picture: ”systems of cities” (Batty, 2013).

To conclude the discussion about the types of resilience, let us note that the majority of publications are dedicated to two types of resilience: socio-ecological and engineering ones

(Duit et al., 2010; Meerow et al., 2016).

№	Scientific domain	Type of resilience
1	Urban ecology	Socio-ecological
2	Hazards and disaster risk reduction	Engineering
3	Urban and regional economies	Economic
4	Governance and institutions	Social and socio-ecological
5	Climate change and adaptation	Socio-ecological
6	Urban theory	Social

Table 2.2: Classification of resilience by scientific domain and type

System of interest

Differences in the systems of interest create the third source of discrepancies. Again, even the system of interest is usually defined as the city as the whole, often, researchers tend to emphasize a specific part of it: resilience of the water supply system - in case of the drought, it must be able to handle the demand, maybe with a short delay (recovery time), but to bounce back to the original state; resilient communities - one of the recent observation backed up with the data is that local communities are often can have much more influence in saving peoples lives, bigger than international organizations (partly it follows from the fact that these people are geographically closer to the epicenter of the disaster); etcetera. (Maruyama, 2016) proposed to name this division as differences in the target system.

№	Type of resilience	Predominant system of interest
1	Socio-ecological	Environment, people and physical infrastructure
2	Engineering	Infrastructure systems
3	Economic	City economy, businesses and organizations
4	Social	Individuals and communities

Table 2.3: Classification of resilience by type and system of interest

Related concepts

Therefore, by defining: the type of resilience, the system of interest, related concepts, the definition often also predefined. Partly the related concepts were already named above, but let us describe them sequentially. There is no strict division between the concepts that are used in case of different type of resilience. There is often an overlap. However, usually, researchers emphasize some concepts over other ones.

The literature on socio-ecological and economic resilience mostly puts the accent on the *adaptation* and *transformation* (Crowe et al., 2016; Engle, 2011; Ernstson et al., 2010; Mehmood, 2016). Sometimes transitioning, learning, and development (Folke et al., 2002). These can be treated as desired city attributes when transferred to measurable capacities (adaptability or adaptive capacity), or as the ”pathways” to resilience (via adaptation). For example, to what extent the city can adapt? Or, to survive the city must adapt. The exact meanings of the adaption and adaptability concepts are ambiguous, and there are a lot of debates around it (Meerow et al., 2016). Barnett (2001) stated that adaptation is hard to grasp since it requires system-wide understanding and analysis. Besides that, there are a variety of possible ways how a city can adapt which are defined by the system of interest (a part of the city that should do so). He defined the adaptation in the general terms (see Definition 2.1.2.1). This definition operates with ecological and social systems and proposes to modify them. Now the difference between engineering and socio-ecological resilience becomes more clear. **Instead of maintaining status-quo and often improving robustness or speed of bouncing**

back, socio-ecological resilience proposes a more complex task of "modifying" the city.

Definition 2.1.2.1 **Barnett (2001)**

Adaptation is the task of modifying the ecological and social system to accommodate climate change impacts such as accelerated sea-level rise so that systems can persist over time.

The Committee on Climate Change also considered adaptation as an action that goes beyond simple recovery to the original state (Jabareen, 2013) (see Definition 2.1.2.2)

Definition 2.1.2.2 **Jabareen (2013)**

Adaptation is adjustment of behavior to limit harm, or exploit beneficial opportunities, arising from climate change.

Meerow et al. (2016) highlighted the differences between *general adaptability*, so the city has a set of actions that must be taken in case of any king of shock or stress (similar to having general and specific resilience), and *specific adaptations*, for example, a course of measures in case of an earthquake.

By considering adaptability as an attribute of the city instead of as a set of action, we loosely equate it to the concept of *adaptive capacity* (Walker et al., 2004). Engle (2011) explained adaptive capacity in simple terms as the ability of a system to adapt. The more adaptive capacity city has, the higher the ability of it to adapt under uncertain and unpredictable challenges.

Definition 2.1.2.3 **Berkes et al. (2000)**

Adaptive capacity is the capacity to adapt to and shape change.

One of the most often overlapping concepts with resilience (apart from the robustness) is the concept of *vulnerability*. In broad terms, vulnerability can be defined as susceptibility to harm. According to the IPCC, vulnerability consists of three components: *exposure*, *sensitivity*, and *adaptive capacity*. Exposure defined as the extent to which the system is physically in harm's way. For example, inhabitants who live close to the sea are more exposed to a tsunami. Sensitivity determines to what extent the system can be affected by stress or a shock. Given a lot of protective measures coastal population might be well-protected from the waves and therefore have low sensitivity to it. The final component is adaptive capacity. The notion of adaptive capacity is similar to the one used in urban resilience literature. It defines the system's ability to prepare for and adjust to stress or a shock, mainly to lessen the negative impacts and take advantage of the opportunities (Adger et al., 2007; Smit and Wandel, 2006). Therefore, adaptive capacity is one of the factors that can decrease exposure and sensitivity.

The concepts of transformation and corresponding transformative capacity seem to be understudied. Transformation in resilience literature implies a fundamental change, radical shifts, and, thus, is different from adaptation (Moore et al., 2014; Walker and Salt, 2012).

Definition 2.1.2.4 **Moore et al. (2014)**

Transformation involves fundamental change, which in the context of sustainability, requires radical, systemic shifts in values and beliefs, patterns of social behavior, and multilevel governance and management regimes.

Definition 2.1.2.5 **Walker and Salt (2012)**

Transformative capacity is the capacity to create a fundamentally new system (including new state variables, excluding one or more existing state variables, and usually operating at different scales) when ecological, economic, and/or social conditions make the existing system untenable.

Without a doubt, such transformative changes in systems structure are often needed. However, many modern cities locked-in with previously accepted development plans, or, in terms of complex adaptive systems, path-dependant. The past defines the future, and the amount of resources needed to transform the city is usually much higher than to persist with its current structure. The time required for transformation is higher than for adaptation and absorption (Chelleri et al., 2015).

Definition 2.1.2.6 **Walker and Salt (2012)**

Transformative capacity is the capacity to create a fundamentally new system (including new state variables, excluding one or more existing state variables, and usually operating at different scales) when ecological, economic, and/or social conditions make the existing system untenable.

The biggest challenge when dealing with the concepts of socio-ecological resilience is to come up with a way to transfer these concepts into quantifiable constructs (Engle, 2011).

Addressing the problems of climate change researchers often introduce the concept of *mitigation* as an import part of resilience (Mehmood, 2016; O'Malley et al., 2014). In simple terms, mitigation refers to the changes aimed at reducing the sources that contribute to climate change (Jabareen, 2013). However, let us note that mitigation is out of the scope of this research and will not be considered further.

The scholars of general and engineering resilience stressed the importance of certain systems attributes that should be enhanced: *robustness, multifunctionality, modularity, redundancy, connectivity, efficiency* (Feliciotti et al., 2017; Friend and Moench, 2013).

Social resilience literature emphasizes *diversity, equity, integration, distribution of power* (Jabareen, 2013).

№	Type of resilience	Important concepts
1	Socio-ecological	Absorption, adaptation, transformation, learning, development, mitigation
2	Engineering	Absorption, equilibrium, status-quo, robustness, recovery time
3	Social	Diversity, equity, integration, power

Table 2.4: Classification of resilience by type and corresponding key concepts

Urban resilience definitions

Finally, everything is summarized by the definition of urban resilience. Some authors proposed definitions of general resilience. For example, Yamagata and Sharifi (2018) have a broad perspective on resilience: a city must survive during any shock either a natural one - earthquake, or human-made - economic crisis. They perceive shocks as an opportunity for a system to become stronger. Maruyama (2016) has a similar position. According to him, resilience is *"an ability of a system to withstand shocks and stresses and recover from failure."* Despite the fact that he has taken a perspective of a biological system, the definition does not operate with the notions of adaptation and transformation. In his definition, he contrasted the ability of organisms to survive, and countries and commercial companies combat unexpected changes. From his point of view, the concept of resilience is similar to biological robustness and can be also framed as a definition useful from the engineering resilience perspective. As was discussed previously, modern socio-ecological literature proposes to go beyond a static concept of vulnerability and stresses the importance of adaptation, learning, and development (Folke et al., 2002). Lamond and Proverbs (2009) provides a great example of urban resilience from the engineering perspective. If a disaster occurs a city should be able to go back to its original ”pre-disaster” state: electricity, water, and food supplies are restored, roads and houses are fixed, etcetera. Remarkably that not all definitions given by social science domains directly address the concepts of equity and diversity.

№	Type of resilience	Definition example
1	General	The capacity to withstand shocks and stresses while retaining system properties
2	Socio-ecological	Resilience is the capacity to buffer change, learn and develop
3	Engineering	Urban resilience encompasses the idea that towns and cities should be able to recover quickly from major and minor disasters
4	Social	Urban resilience is ability to absorb, adapt and respond to changes in urban systems

Table 2.5: Classification of resilience by type and definition

Apart from the scholars, international organizations involved in building resilient cities developed their definitions (see Table 2.6). Notably, these definitions are more explicit regarding who should benefit from urban resilience: inhabitants, individuals, or communities. For these definitions, the ones as mentioned above are generic. The similarity is in the measures that should be taken in case of crisis: adaptation, transformation, and growth.

№	Definition	Organization
1	The measurable ability of any urban system, with its inhabitants, to maintain continuity through all shocks and stresses, while positively adapting and transforming towards sustainability	UN-Habitat
2	The capacity of individuals, communities, institutions, businesses, and systems within a city to survive, adapt, and grow no matter what kind of chronic stresses and acute shocks they experience	100 Resilient Cities Network

Table 2.6: Definitions of urban resilience proposed by international organizations

Ideally, there should be a consistency between all 5 aspects of urban resilience: domain, type, a system of interest, related concepts, and definition. However, often literature demonstrates the opposite. Coaffee (2016) proposed the following definition of urban resilience: *”...the capacity to withstand and rebound from disruptive challenges.”* It seems that this definition taken from social science domains do not fully cover the whole complexity of social systems. Nevertheless, the fact that there is a mismatch between the system and the model does not mean that there are no benefits of using the concept. It merely means that one should be careful when trying to put it into practice.

Without a doubt, the review proposed here does not fully cover all aspects and advances in urban resilience literature. The concept has spread across a variety of disciplines and now has different definitions and key ideas. It is still somewhat challenging to obtain a comprehensive understanding and immediately put it into practice and operationalize. However, the review clarified the crucial differences between major types of urban resilience: engineering and socio-ecological ones, discussed what capacities of the city should be enhanced, what are the possible pathways to resilience: absorption, adaptation, and transformation, and finally examined example definitions that are common in the field. The next step is to examine all this can be combined in a conceptual framework.

2.1.3. Conceptual frameworks and assessment

According to Jabareen (2009) conceptual framework is *"a network, or 'a plane,' of interlinked concepts that together provide a comprehensive understanding of a phenomenon or phenomena."* Conceptual frameworks are often built to push forward the implementation and play the role of guidelines. However, the process of building conceptual frameworks for urban resilience is complicated by the tensions mentioned above. The vast majority of publications is still focused around defining urban resilience: what is "the right" definition, what are measures that can ensure transitioning to resilience and so on. It results in the lack of the frameworks, and, slows down the practice of urban resilience.

The number of frameworks present in the literature is relatively small. Among all of them, two seems to cover the theory of urban resilience from the right angle: by Jabareen (2013) and Chelleri et al. (2015). So far, the applicability of the concepts has not been discussed. By the right angle, we mean that it should be **useful** (provide relevant insights) and be **operationalizable** (quantifiable to a certain extent) for extreme heat threat and protecting vulnerable individuals and communities. In the terminology defined above, it should have a focus around socio and socio-ecological resilience to a specific shock - heat waves.

The framework proposed by Jabareen (2013) defines four main concepts that are enabling urban resilience. Namely: vulnerability analysis matrix (VAM), prevention, urban governance, uncertainty-oriented planning. The first one is essential for this study. In short, it can be used for **socio-demographic mapping of risks and vulnerabilities.** Even the debates about relations between vulnerability and resilience are still ongoing, these concepts are rather complementary than contradictory and can strengthen each other Engle (2011).

VAM consists of four main components: *uncertainty, informality, demography, and spatiality.* Uncertainty refers to unpredictable environmental challenges posed to modern urban areas. Conventional risk assessment often ignores uncertainties and operate based on probabilistic predictions. According to the author, such an approach showed its inability to deal with climate change impacts. The second component seeks to answer the question on the demography of vulnerability: who is vulnerable, what are the main attributes of these individuals and communities, etcetera. The third component reflects the fact that urban areas often develop not according to a plan, but rather chaotically and unorganized. As a result, informal settlements appear. These settlements are often not registered and underrepresented in the processes of planning. They lack essential infrastructures and usually vulnerable due to specific socio-demographic attributes. The final component - spatial distribution of vulnerability, stresses the importance of the fact that the effects of the threat spread unevenly across the city. For example, coastal are more subject to a tsunami; residents who live in an urban heat island experience higher temperatures. What is essential here is that fact that more vulnerable areas usually populated with vulnerable individuals. Sometimes, it happens because of the land price differences. The more the place is subject to damage, the cheaper it is.

It seems beneficial to start the analysis from VAM, and only after that shift the focus on other concepts of the framework: developing preventive measures, addressing specifics of the governance process and putting planning measures in practice. The aspects of VAM is summarized in Figure 2.6.

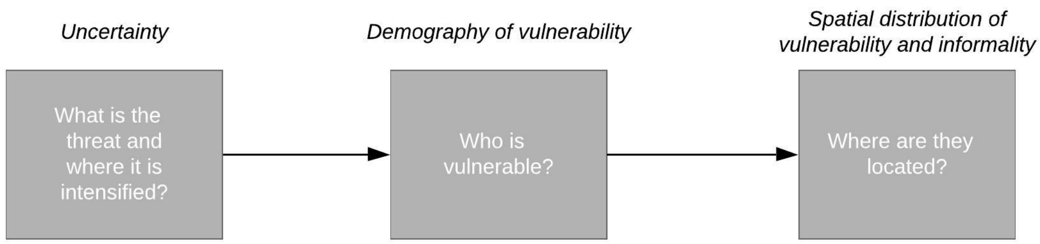

Figure 2.6: The components of Vulnerability Analysis Matrix concept connected. Adapted from Jabareen (2013)

The research of Chelleri et al. (2015) and Meerow et al. (2016) defined possible approaches of transition to urban resilience. The way how vulnerable individuals and communities can tackle a shock may resemble a set of sequential steps. First, **absorb and recover if possible**. Second, **adapt if needed**. And finally, **perform a radical transformation to become fully ”safe-to-fail”** (Ahern, 2011).

For example, when heat wave strikes, ”a part” of it will be absorbed by vegetation around a specific place (O'Malley et al., 2014). Emergency services will respond and help the population to recover. However, some people do not call an ambulance. Ehey may not be aware of the danger or underestimate it. In this case, other individuals (friends, neighbours or relatives) could take care of those who are more prone to heat and vulnerable (Klinenberg, 2003). Finally, the lessons learned can lead to radical transformations in the future: rebuilding a city in a way that it will not have an urban heat island effect, or buying an air-conditioner (not a sustainable solution though).

More formally, Chelleri et al. (2015) explained recovery as an immediate response to the system shocks and aimed at returning the system in its initial state (engineering resilience). Adaptation is defined as *”the process of adjustment to actual or expected changes and its consequences, disregarding system boundaries by moving the thresholds to make persist within the same regime.”* The third measure proposes long-term structural transformations are a response for shocks that cannot be tackled by the previous two. Figure 2.7 describes the framework.

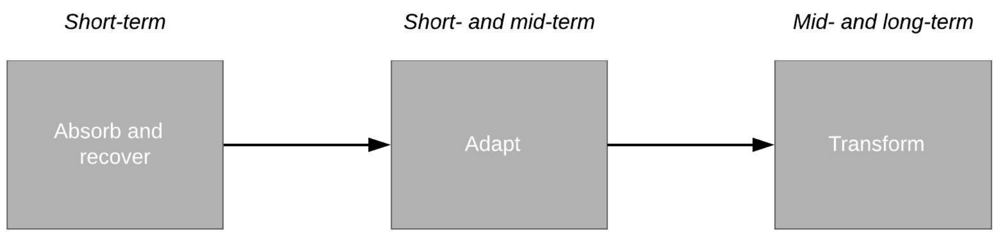

Figure 2.7: Approach to urban resilience: absorb and recover, adapt and transform. Modified from Chelleri et al. (2015)

Importantly, many authors argued in favour of the **combination of top-down and bottom-up measures to ensure resilience** (Chelleri et al., 2015; Swalheim and Dodman, 2008; Zaidi and Pelling, 2015). The example of a top-down action is, again, emergency services that are run by the city, well-informed and ready to help people in need. The bottom-up approach implies an initiative coming from individuals or communities. Neighbour is taking care of the neighborhood. To promote such actions, *social infrastructure* can be used (Klinenberg, 2018).

In a nutshell, social infrastructure is an umbrella term that refers to various places where people gather and create or empower social ties, for instance, libraries. Empirical research showed that libraries play an essential role in bringing people with various background and social status. It is especially relevant for individuals who can be considered as vulnerable to extreme heat: elderly, low-income, lonely.

A few words can be said about the other way to operationalize urban resilience through assessment models. Usually, such models are build-up on the top of conceptual frameworks. Thus, along with the complications related to building conceptual frameworks, a new problem arises: how to translate sometimes rather vague and abstract ideas into numbers (Jabareen, 2013). For example, the work of Sharifi and Yamagata (2016) and Bourdic et al. (2012) is far from being specific and does not differ from general assessment models. These models put attention on the variety of broad indicators such as a stable economy, strong demographics, modern transportation system, and so on. Somehow it seems that they neglect the complex interplay between the parts of the urban system.

Two frameworks described above provide fertile soil for the defined problem: how to protect the vulnerable population from extreme heat. Combine VAM and the measures of socio-ecological resilience to:

1. map the threat, vulnerable individuals and communities;

2. identify the possible measures that should be taken to ensure resilience.

2.1.4. Applications

Scientific literature does not have many examples of putting the urban resilience into practice. There are several reasons for that: the novelty of the concept, and, therefore lack of knowledge; scientific debates around it, and thus, inability to precisely define it; vagueness or ambiguity of related ideas, for example, how to quantify adaptive capacity?

Crowe et al. (2016) argued in favor of innovations to built up resilience. The authors stressed the importance of imagination, creative approach, and design thinking. They presented five experiments that were aimed to promote urban resilience in Dublin. In short: a website `www.reusingdublin.ie` to map underused places in the city; a smartphone application ”Geo-timeline” to capture the current state of a place and keep track of the development; another web platform ”Community gains” to share the resources that are needed to realize a certain project; etcetera. Ahern (2011) supported such an approach by emphasizing the importance of ”learning-by-doing” and ”experimental probes.”

Oppositely, O’Malley et al. (2014) proposed to incorporate a conventional solution into urban resilience framework - vegetation to fight urban heat island effect. The simulation model that was built by authors for the city of London, showed a positive impact of *green infrastructure* on reducing the average temperature.

The largest challenge in urban resilience practice is to propose solutions that are consistent with the theory. So far, they either framed as experiments (Crowe et al., 2016) or aimed at only one aspect of urban resilience O’Malley et al. (2014).

2.1.5. Critique

Urban resilience is a relatively new concept, and as a result, it is still understudied and has a lot of debates around it. The biggest obstacle is its vagueness, ambiguity and complexity (Engle, 2011; Friend and Moench, 2013). Having a variety of definitions makes it hard to put it into practice. Moreover, there is no comprehensive theory about the object of urban resilience: cities (Batty, 2013). Cities can be studied from a variety of angles that are all desire equal attention. What model should be used to ensure that every essential aspect and relationship of it covered ”good enough?” And what is good enough?

Another critique relates to whether or not it urban resilience a positive concept (Chelleri et al., 2015; Meerow et al., 2016). A growing body of literature seeks to answer the questions of equity, transparency, and inclusiveness (Leichenko, 2011). Resilience has its cost, and since the distribution of vulnerability is usually unequal, how to make sure that everyone benefits from it, especially the urban poor, minorities and vulnerable?

2.1.6. Concluding remarks

The goals of this section were: to provide an overview of urban resilience concept, then decompose and explain main parts of it, and lastly, identify those are relevant for extreme heat and can be operationalized.

It was found that urban resilience is a beneficial and complex concept. Urban systems should be resilient in order to survive during the times of hitherto unseen problems. But there are a lot of components to unfold quite carefully. First, what is the predominant theory that is used to explain the performance of the city? What type of resilience is desired? What is the system of interest? What are the key attributes of an urban area that should be enhanced? What measures can be used to ensure the transitioning to urban resilience? On the one hand, these questions make it hard to operationalize urban resilience, but, on the other side, provide a researcher with flexibility.

Given the fact that the focus of this study is on a specific threat (extreme heat) and system of interest (vulnerable individuals and communities), it will be fair to assume that urban ecology theory along with socio-ecological resilience will be a beneficial standpoint. Consequently, the main focus will be first on mapping vulnerabilities using Vulnerability Analysis Matrix by Jabareen (2013), and, second on the three essential processes and corresponding capacities described in Chelleri et al. (2015): absorption and recovery, adaptation, and transformation.

However, we will not limit ourselves by considering urban population. A significant contribution is made by different types of infrastructures: critical, green and social. These will be discussed in the following chapter.

Unfortunately, none of the definitions present in the literature does not fully cover all of the assumptions made above. Therefore, we will use a definition that is as close as possible. The one given by (Desouza and Flanery, 2013) seems to summarize the essence quite precisely (see Definition 2.1.6.1).

Definition 2.1.6.1 **Desouza and Flanery (2013)**

Urban resilience is an ability to absorb, adapt and respond to changes in urban systems.

2.2. Extreme heat

This section aims to cover the essential aspects of the extreme heat threat. First, it provides a short note on climate change and explains the heat wave phenomenon. Second, it describes urban heat island effect and how it amplifies the effect of heat waves. Finally, it defines vulnerable to extreme heat population.

More frequent higher temperatures are one of the future challenges for all world countries (Straub, 2018; United Nations, 2018a). Even for those that have not had a great history in combating such climatic problems and the Netherlands is not an exemption. The scientific evidence is summarized in the latest report made by The Intergovernmental Panel on Climate Change (IPCC) (Pachauri and Mayer, 2015). The visualization made by Ed Hawkins, a climate scientist at the University of Reading in the UK and a lead author of the IPCC's Sixth Assessment Report, provides great insight into these changes (Hawkins, 2019). It demonstrates the annual average temperatures for the Netherlands from 1901-2018 using temperature data from Berkeley Earth. The average temperature in 1971-2000 is set as the boundary between blue and red colors, and the color scale varies from +/- 2.6 standard deviations of the annual average temperatures between 1901-2000. Figure 2.8 demonstrates that the average yearly temperature, along with the frequency of hot extremes, is rising.

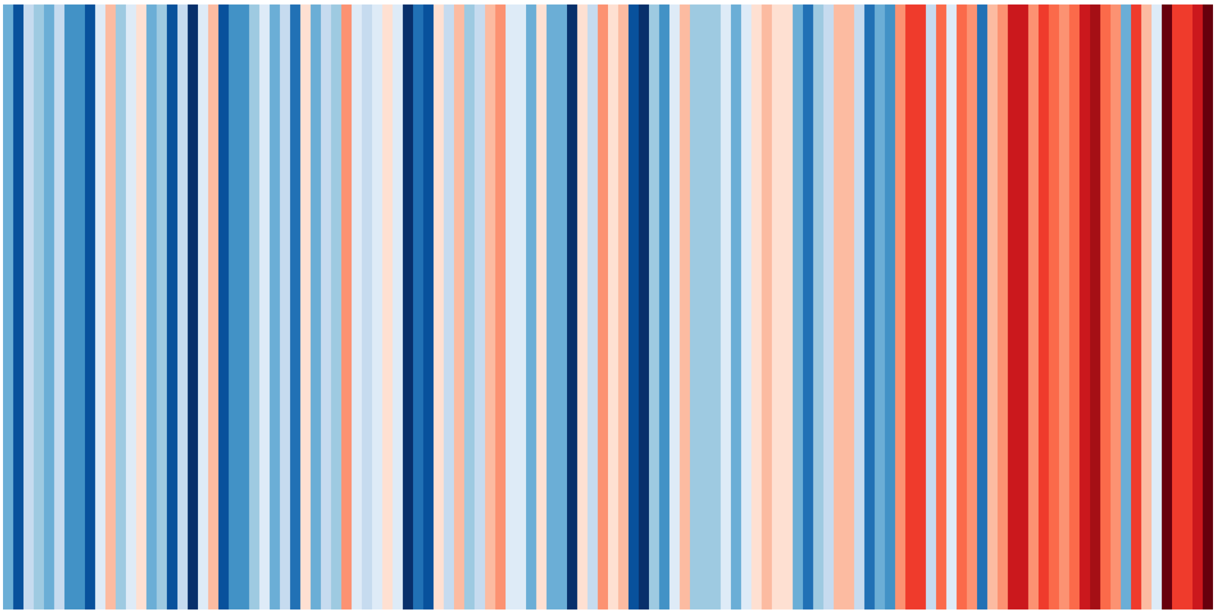

Figure 2.8: Annual average temperatures for Netherlands from 1901-2018 using data from Berkeley Earth. Visualization made by Hawkins (2019).

2.2.1. Heat wave

Heat waves are a relatively well-studied phenomenon (Wu et al., 2013). Broadly speaking, the primary cause of it is climate change. The damage made by the recent heat waves is tremendous. For instance, the Chicago heat wave of 1995 was a cause for around 700 deaths, the European heat wave of 2003 claimed the lives of 70000 people (Åström et al., 2011). Given the fact that average temperatures are rising, the problem requires extra attention.

The situation is complicated by the uncertainties about: the location (where another heat wave will take place), and the frequency (how often heat waves can be expected to appear). Urban resilience literature stressed the fact that cities should be ready for all sort of uncertainties to survive (Ahern, 2011; Leichenko, 2011). Thus, even the past history of a specific city has no examples of heat waves, it should be ready to face it. Due to uncertainties related to climate change, for instance, the amount of greenhouse gas emission, it is hard to predict the number of upcoming heat waves. As a result, the conventional infrastructure measures sometimes are not ready for the shock.

Another important factor is the aging population. Typically, people who are under the risk of heat wave are elderly. According to population projections made for the Netherlands in 2010, the percentage of people aged over 65 expected to increase to 26% in 2035 (Smits et al., 2014). Centraal Bureau voor de Statistiek reported that *"half of the Dutch adults will be over 50 in 2019"* (Centraal Bureau voor de Statistiek, 2014).

The definitions of heat wave usually have an accent on two critical attributes: *temperature threshold* (Definition 2.2.1.1) and *duration* (Definition 2.2.1.2)

Definition 2.2.1.1 **Stefanon et al. (2012)**

A heat wave is an extreme event that is defined when the temperature exceeds a given threshold.

Depending on the country this threshold may vary. For example, the threshold set by the Netherlands National Institute for Public Health and the Environment (RIVM) is 27 °C (Rijksinstituut voor Volksgezondheid en Milieu, 2011).

Definition 2.2.1.2 **Zaidi and Pelling (2015)**

A heat wave is a period of hot weather lasting for more than five consecutive days.

The duration is also a point of discussion. RIVM recently have lowered down the duration from 5 to 4 consecutive days.

Other important factors can be considered: night temperature, humidity, etcetera. However, the crucial role is played by the temperature and the number of days that it lasts.

2.2.2. Urban heat island

One of the factors that can amplify the effect of the heat wave is urban heat island (UHI). Formally, it can be defined as:

Definition 2.2.2.1 **O'Malley et al. (2014)**

Urban heat island is a phenomenon where the significant temperature difference between inner micro-climates of a city and their neighborhoods micro-climates can be perceived.

Thus, inhabitants who live in UHI experience higher temperatures and, therefore, are more vulnerable.

Some argue that the urban heat island is relatively easy to find. It is just an area within the city with the highest population density. The more extensive approach revealed the following important contributors to creating of UHI: land coverage, emission, and energy consumption (O'Malley et al., 2014).

2.2.3. Extreme heat vulnerability

The higher temperatures challenge the cities in many ways: failing infrastructure, disruptions in food and water supply, but most importantly, it causes deaths. According to the studies, certain population groups are more *vulnerable* to the effects of the heat waves than others. The **vulnerability here is susceptibility to harm that can made by a heat wave**. The factors that characterize these groups can be split into the following categories: age group, health problems, other socio-demographic or socioeconomic characteristics.

Conventionally, researchers highlight the importance of the age factor (Knowlton et al., 2008; Naughton et al., 2002; Zaidi and Pelling, 2015). Others stressed the fact that heat wave can exacerbate existent deceases, for instance, cardiovascular, cerebrovascular, and respiratory problems (O'Neill et al., 2009; Schwartz, 2005). The third branch of the research emphasizes the importance of income, education, lifestyle, etcetera. Swalheim and Dodman (2008) named the urban poor as the main category that is vulnerable to climate change and extreme heat. Klinenberg (2003) in his famous study additionally pointed out leaving alone individuals, people with mobility restrictions, and those who do not have access to transportation.

These factors mentioned above dictate the choices of future data. To summarize, we are looking at the following variables or their proxies:

- age;
- sex;
- ethnicity;
- household composition;
- income;
- education

- employment;
- health issues.

2.3. Solutions for resilience capacities

This section aims to explore possible solutions for urban resilience. Given the focus on a specific threat: heat waves, and a system of interest: vulnerable individuals and communities, the question is: what measures can be useful?

There are different ways to become more resilient. The approaches and related complications were discussed previously (see Subsection 2.1.4). The solutions can be split into the two groups: innovative, framed as experiments, with some possible positive contribution, and, conventional, that is used to combat a specific threat and has been proved via simulation models. Overall, there is a lack of practical examples of how urban resilience should be operationalized and what are the possible, proven with data, ways to do so.

One way to proceed further is to assess urban resilience based on existent theories. The framework of socio-ecological resilience defined in Subsection 2.1.3) requires us to find some solutions that will support each of the components: absorption, adaptation, and transformation. We proposed to use different types of *infrastructures* for it (see Figure 2.9). These infrastructures should increase the corresponding capacities, and, promote the actions (denoted by the bold the lines). Importantly, sometimes there is an overlap between the type of the infrastructure and the capacity that it should enhance (denoted with the colored lines). For example, green infrastructure, when presented by a park, can increase adaptive capacity as well as absorptive. In the following subsections, we will discuss these infrastructures in more details.

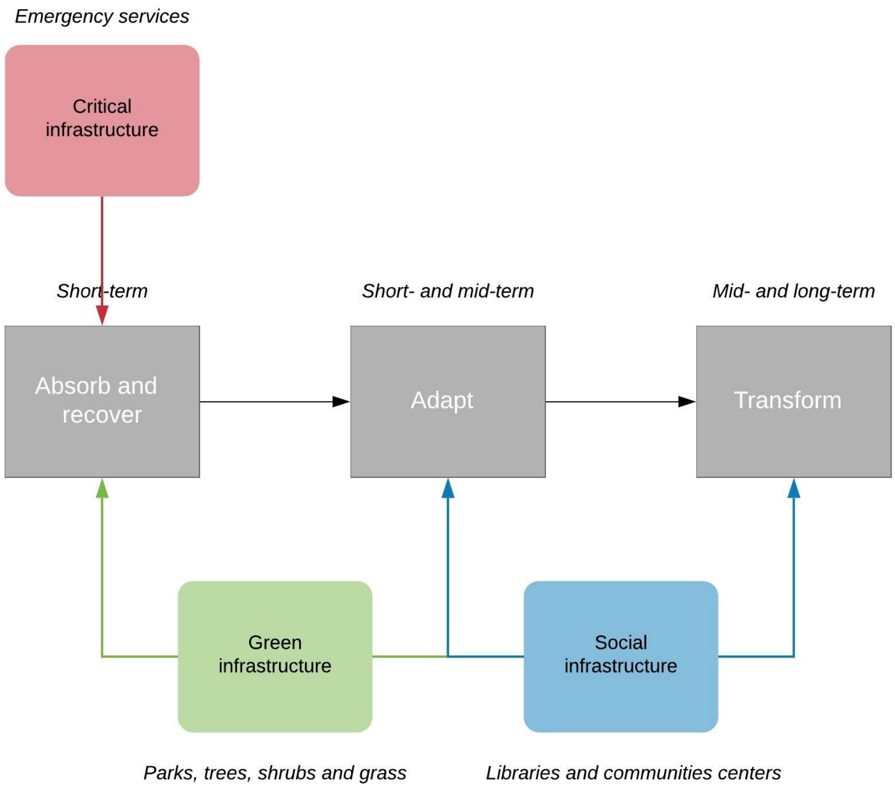

Figure 2.9: Conceptual framework of socio-ecological resilience supported by the different types of infrastructure solutions

2.3.1. Critical infrastructure

The importance of *critical infrastructure* cannot be underestimated. When a shock occurs, the immediate response is required. In case of extreme heat threat, critical infrastructure is represented by ambulance services. Heat wave is a continuous phenomenon that can last for days (during the daytime and nighttime as well). This fact pushes ambulance services beyond the borders of its capacities. This type of infrastructure contributes to absorptive capacity and helps vulnerable individuals and communities to recover.

Evaluation of critical infrastructure can be done differently. For example, by demand and supply ratio. How many cars and doctors are available in case of an extreme situation? Or by measuring the lengths of the routes from vulnerable individuals and communities to the closest hospital. Some parts of the city are not easy to reach due or common traffic congestion, or it is just physically far. The third way is to analyze ”preparation time.” How long will it take to get personnel ready? Since the heat wave may strike unpredictability, for example, during the autumn, some staff can be on vacation when the help is needed. Without a doubt, more comprehensive metrics for assessment of critical infrastructure exist. However, for the sake of this research, we will limit ourselves only with the first and the second. Ideally, critical infrastructure should be able to tackle any increase in demand caused by an unexpected heat wave. Besides, the hospitals should be well connected to the places where the most vulnerable people live.

The following Figure 2.10 is a simplified representation of critical infrastructure concept. It depicts an abstract city with one hospital highlighted.

Figure 2.10: Sketched representation of critical infrastructure concept. Adapted from Envato (2019)

2.3.2. Green infrastructure

The concept of *green infrastructure* (sometimes called *blue and green*) is an integral part of sustainability and resilience concepts. In short, green infrastructure is a term used to denote vegetation in the form of trees, shrubs, and grass.

The support that green infrastructure provides varies from the immediate response (absorption) in case of standalone trees to mid-term (adaptation) in case of parks and places where people can gather and create or improve social ties O’Malley et al. (2014); Peters et al. (2010).

According to multiple studies, such infrastructure plays an essential role in decreasing urban heat island effect and decreasing the temperature. In the simulation study conducted by O'Malley et al. (2014) 3 strategies to tackle urban heat island effect were selected: vegetation, water, use of materials with albedo rating. The vegetation strategy performed the best. Places with a high amount of vegetation showed lower temperatures.

Jabareen (2013) also highlighted the role of *greening.* Apart from the positive contribution to urban climate, it has also improved biodiversity and citizens health (MacKillop, 2012; Swanwick et al., 2003). Ahern (2011) also argued in favour of green infrastructure. From his point of view, there is a big overlap between resilience and sustainability. *greenification* is usually a part of sustainable development strategies. Modern cities require massive redevelopment. As a result of this process, future cities need to be much greener.

In Sharifi and Yamagata (2016), the authors proposed a set of indicators that can be used to measure resilience. Some of them are directly dedicated to green infrastructure: the number of parks and forest conservation. However, neither the connection of green infrastructure to extreme heat problem nor discussed its implementation were discussed. For example, how many parks should be in a city?

In simple terms, there should be more green spaces in cities. Both trees and parks. Especially where the vulnerable individuals and communities are concentrated.

Figure 2.11: Sketched representation of green infrastructure concept. Adapted from Envato (2019)

2.3.3. Social infrastructure

Social infrastructure is an umbrella concept recently proposed by Klinenberg (2018) for the places where people can gather, build up or improve their relationships. The concept refers to libraries, community centers, parks. Klinenberg argued that social infrastructure not equal to "social capital" - the concept that is commonly used to measure people's relationships, since it has a place in space.

There is empirical evidence on the usefulness of informal activities organized by libraries and other community centers to build up relationships and prepare people for extreme heat Klinenberg (2015). Importantly, these activities may not be directed to the extreme event itself, but rather aimed at socializing and fun. Having healthy relationships usually implies taking care of each other. People with different statuses, backgrounds and ethnicity, vulnerable and not, who met at the library, become aware of each other status. As a result, during a heat wave they may offer help to each other. Social infrastructure is connected to mid-term adaptation and long-term transformation. It is a bottom-up solution since it promotes proactive behavior of citizens when a shock is in place. In addition, since it is a physical space, it can be used as a place to gather in case of an extreme event, and therefore, contributing to absorptive capacity and recovery phases.

In his work, Klinenberg named certain factors that affect the process of building up social ties even if the social infrastructure is in place. Examples are high crime rates and constant migration (Klinenberg, 2018). However, he showed in the cases how the presence of social infrastructure could gradually transform even quite problematic neighborhoods. He argues that the development of it contributes to extreme heat resilience and much more important than investing in conventional infrastructure measures.

Figure 2.12 emphasizes one type of social infrastructure that will be studied later: public libraries.

Figure 2.12: Sketched representation of social infrastructure concept. Adapted from Envato (2019)

2.4. Conclusion

This chapter had three main goals that are connected to the first research question. First, to explore the current stage of urban resilience and identify its key components and operationalizable frameworks for extreme heat threat. Second, identify the main aspects of extreme heat shock, its amplifier: urban heat island, and understand who is the most vulnerable in this case. The final third goal was to explore what are the possible infrastructure strategies that can tackle the extreme heat threat.

Resilience is a complex concept. Originated from systems ecology it has spread across a variety of disciplines: from economics to social sciences. Recently it has adopted by city scientists and became urban resilience. During the literature review, it was found that there are three main types of urban resilience: general, engineering and socio-ecological. The last two are connected to specific study domains and have some model of the city to operate with: a city as a combination of engineering systems and a system as complex adaptive system. Besides that, they have certain key components: robustness, vulnerability, absorption and recovery, adaptation, transformation. Due to the nature of the problem and focus on vulnerable individuals and communities it seems useful to adopt urban ecology theory and socio-ecological resilience. It comes with the notions of absorption and recovery, adaptation, transformation. To support the computations two conceptual frameworks were chosen and merged into one from Jabareen (2013) and Chelleri et al. (2015).

The main attributes of the heat wave threat highlighted in the scientific literature are temperature threshold and duration period. Based on those, a specialized agency of the country plans the emergency actions. The opinions of the researchers regarding heat wave vulnerability vary. Some tend to emphasize age as the key factor, when others stress the importance of other attributes: education, for example. In this study, I will merge those opinions. The vulnerable individuals are people above a certain age threshold (typically, 65), living alone, with low income and set of health issues. The urban heat island plays a role of an amplifier. If a vulnerable individual lives in a heat island, it is a point of concerns.

From the solution perspective, three types of measures can be successfully incorporated into the framework of socio-ecological resilience — first, green infrastructure or vegetation in the form of trees, shrubs and grass. Vegetation absorbs the heat and extremely helpful to combat urban heat island effect. Second, critical infrastructure or ambulance services. When the heat strikes, certain people will ask for help via an emergency network. Ideally, all vulnerable individuals should be easily accessible by ambulance cars. Besides, the personnel of hospitals must be ready for unexpected peaks in demand caused by heat waves. The third component: social infrastructure is aimed at adaptation and transformation. This concept denotes physical spaces where people can gather and build up healthy social ties, for instance, libraries. Opposite to the previous two solutions, social infrastructure promotes proactive behaviour of certain citizen groups, and therefore, it is a bottom-up solution. A resilient city should is a city that has well-developed social infrastructure, especially in places, where a lot of vulnerable individuals live.

The next chapter describes how a transfer from identified conceptual framework to a single computational model can be made.

The following Figure 2.13 concludes the theoretical discussion and summarizes the findings of this chapter.

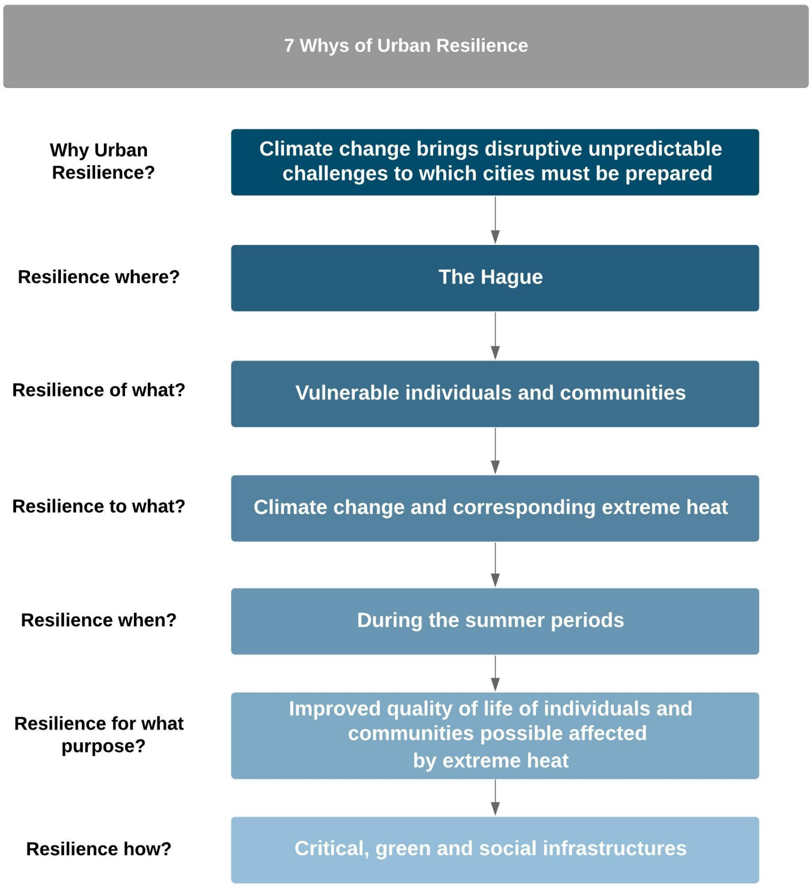

Figure 2.13: *The Seven Whys of Urban Resilience* modified from Meerow et al. (2016) and Ospina (2018)

3

Methodology

This chapter aims to match the concepts described previously with data and methods that are used to operationalize them. First, it discusses the essential details of how the chosen conceptual frameworks are operationalized. In line with this, it presents an overview of the open data sources supporting this process. After that, it covers the important aspects of machine learning algorithms. Finally, it proposes a computational framework. Thus, this chapter provides the answer to the second research question.

3.1. Introduction

The components of urban resilience concept are hard to operationalize. They are sophisticated and allow multiple interpretations. For this study, we need to propose a way to operationalize, first, Vulnerability Analysis Matrix (VAM) by Jabareen (2013), and, second, the framework of socio-ecological resilience of Chelleri et al. (2015) Meerow et al. (2016) combined with infrastructure solutions (SER). Besides, we also need to present evidence in the extreme heat threat and discuss urban heat island effect. To promote future reusability of results, I will start by explaining what kind of data required generally, and only after that, introduce case-specific data sets.

3.2. Data

This study heavily relies on open data. On the one hand, this fact promotes future reusability of the research, but on the other side, the data available predefines the way how certain concepts can be operationalized. Ideally, operationalization of VAM requires to have microdata about the city of interest: The Hague. *Microdata* in statistics usually denotes the information collected on the level of individuals: income, education, occupation, and other. However, due to rising data privacy concerns, such an approach is undesirable. Therefore, the study forced to use aggregated data sets. The same argument on data availability holds for the resilience capacities assessment. Hence, this study also can be used to formulate future data needs. Especially in regards to social infrastructure. The examples are: numbers of library members, their living addresses, how often do they participate in the events organized by the library.

Now, let us dive in to the details of the VAM. Let us recall that it seeks to answer the following questions:

1. What is the threat and where it is intensified?
2. Who is vulnerable?
3. Where are they located?

Therefore, first, there is a need for historical data about heat waves and corresponding damage made to a specific city (The Hague in our case), or at least to a particular country (the Netherlands). The damage can be measured differently. The conventional metrics are all variations of **mortality**. Plus, we need to have two types of **temperature data**: average temperature across the city (to fix the presence of a heat wave) and local temperatures (to map urban heat islands within the a city). Thirdly we need the sets of **socio-demographic and socioeconomic indicators** of citizens to analyse vulnerability. Importantly, these indicators should be allocated geographically. In addition, I propose to use an extra variable to analyze vulnerability: the number of **ambulance calls**. Ideally, we would like to know the reasons for these calls: does a person experience a discomfort because of heat, there is an accident, or protracted illness caused some complications? If this information is unavailable, then we will need make an assumption about relations between the temperature and the number of ambulance calls. The hotter it is, the more calls hospitals will receive.

The SER, together with the solutions, can be operationalized in the following way. The first component of it: absorption and recovery, is covered by two types of infrastructures: green and critical. In simple terms, the aim for green infrastructure is to analyze to what extent a specific unit, for example, a district, is covered by **vegetation**. The more vegetation a district has, the more absorptive capacity is there. The contribution of critical infrastructure can be measured through street network analysis: how close are the hospitals to vulnerable individuals and communities? The closer they are, the faster the recovery can happen. Basically, the variable is **proximity**. Besides, the ambulance calls data to provide an interesting insight into the demand for emergency services. The hospitals around a specific area should have enough capacity (simply personnel and cars) to tackle an increased number of requests.

The adaptive capacity is the most complex to measure. There are many different types of social infrastructure: libraries, community centers, churches, and so on. The critical assumption is if the infrastructure is in place and easily reachable, then relationships will be built. Again, the variable is **proximity**. Besides, I propose to use the data source called *Popular times*. Recently introduced by Google, it has information about **popularity** of a certain place: ratings, distribution of costumer other the day, average visiting times. Certain factors can degrade social infrastructure. These factors are much easier to measure. It is constant migration and high crime rates.

3.2.1. Heat wave and urban heat island

The data sources that are used to demonstrate the importance of the heat wave threat and urban heat island effect showed in Figure 3.1. These data sets do not require further elaboration and just need to be carefully analyzed.

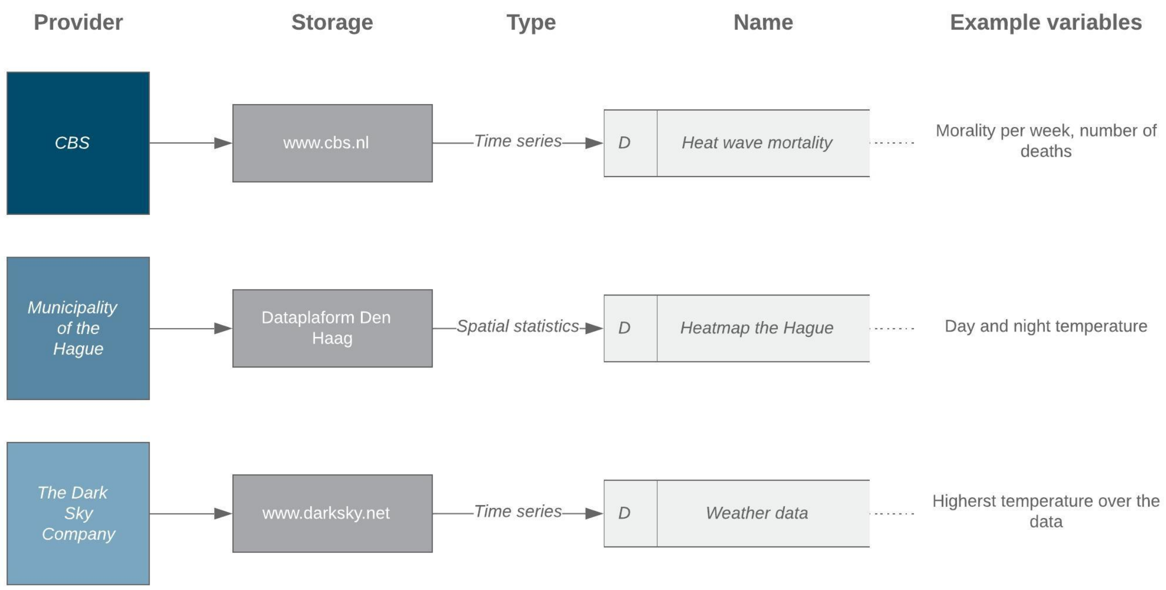

Figure 3.1: Data sources used to map the extreme heat threat

Let us explain the data set from Figure 3.1 in details. The first one has the data about increases in mortality during the periods of extreme heat in 2006, 2010, 2018 and 2019 (Centraal Bureau voor de Statistiek, 2018a, 2019a). The graphs showing these statistics can be found in Section 1.1.3 and Appendix A.1.

The second one has results of the study conducted by Hoeven (2018). In 2017 the authors analyzed the daytime and nighttime temperatures using remote sensing and crowd-sourced weather data from around 200 stations. The observations were aggregated and resulted measurements reported on the scale 100 by 100 square meters. The data set of the study is open-access and can be found online at Hoeven (2017). The result of the analysis can be found in Chapter 4.

The final third data set contains information about the temperature in The Hague. The data is provided by The Dark Sky Company that is specialized in weather forecasting and visualization. To access the data, I used a Python package *darkskylib* and *The Dark Sky API* (Lukáš, 2019; The Dark Sky Company, 2019). It is free for a 1000 requests and allows to collect temperature data in real-time.

3.2.2. Extreme heat vulnerability

The main data sources that are used to operationalize the other two components of VAM are demonstrated in Figure 3.2:

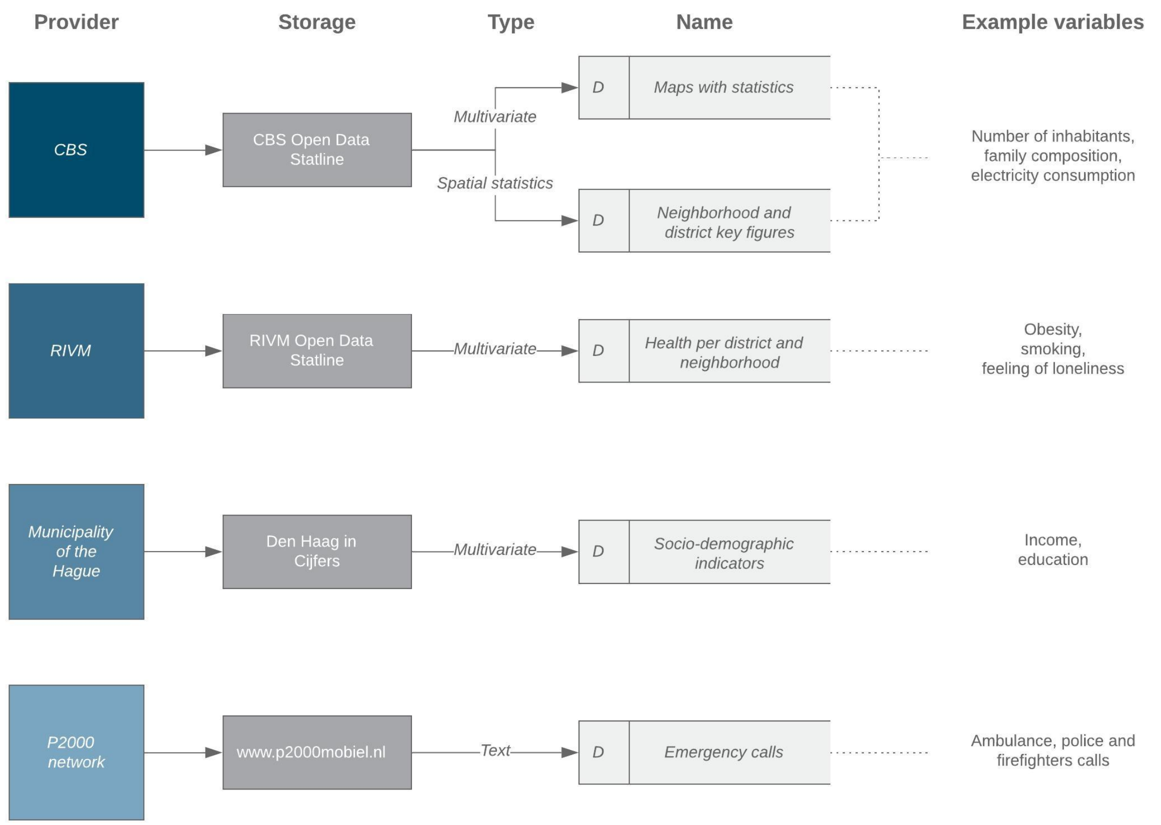

Figure 3.2: Data sources used to map vulnerability

Socio-demographic and socioeconomic variables

The first three data sources reflect on the discussion of vulnerability from Section 2.2.3. People with specific socio-demographic and socioeconomic attributes are more prone to the damage of heat waves; therefore, more vulnerable. The data provided by CBS and the municipality of The Hague is census data and reported yearly. Important to remember that the actual changes may happen monthly or even daily: moving, marriage, change of workplace.

Besides that, the variables of these data sets are allocated according to the address of registration. Of course, an individual spends some of his time at this location. However, his or her daily behavior patterns are mostly hidden: travel to work, walking around or going on vacation, etcetera. Nevertheless, we can still benefit from using such data. As it was found during the literature review, vulnerable individuals are those who above a certain age (above 50 or 65 in some cases), have limitations in mobility and health issues. It is fair to assume that such individuals will spend a significant amount of time in their living places. Therefore, a map of the city, which represents such vulnerability is highly relevant and can provide us with insight.

As can be seen, the necessary data about the inhabitants of The Hague is available at different sources: the data provided directly by Centraal Bureau voor de Statistiek (CBS), the data collected by Rijksinstituut voor Volksgezondheid en Milieu (RIVM) and provided by the infrastructure of CBS, and the data collected and published by the Municipality of The Hague.

Quite recently, CBS has started to publish *Kaart met statistieken* (*Maps with statistics* in English). These data sets are maps with assigned attributes. They have different *resolution*: 100 by 100 m^2, 500 by 500 m^2, 6-Digit Postal Codes, 4-Digit Postal Codes. The resolution here is the size of the area covered by the data. That is, the data with 100 by 100 m^2 resolution reports only about the citizens and built environment that registered within this area. Important to mention that the higher the resolution, the less data is available. For instance, is income data is not available for 100 by 100 m^2 resolution. Since manipulation with postal codes is somewhat complicated, it was decided to use only 100 by 100 and 500 by 500 m^2 data sets.

The name of the other CBS data set is *Kerncijfers wijken en buurten* (*Key figures of districts and neighborhoods* in English). The variables of this data set available on three levels: city, district and neighborhood. The total number of variables here is higher than for Maps with statistics. It has such attributes as family composition, income and crime rates.

In 2016 RIVM collected a variety of health indicators for all Dutch cities (Rijksinstituut voor Volksgezondheid en Milieu, 2016). For example, obesity, limited mobility, and so on. This data is available on district and neighborhood levels. The resulting data set called *Gezondheid per wijk en buurt 2016* (*Health per district and neighborhood 2016* in English).

Apart from CBS, cities themselves also collecting the data. Municipality of The Hague has a more broad range of variables that characterize its population. For instance, there is data about migration and occupation. This data is also available with district and neighborhood resolution.

Thus, we have three data sources and four data sets that complement each other and should cover the concept of extreme heat vulnerability. However, two major problems should be addressed. First, the availability of the data. The Maps with statistics are not available for 2018 and 2019 years. Same holds for RIVM study that was conducted in 2016. To overcome this limitation and preserve consistency, I will use the data for 2016 for further analysis. Second, the resolutions of the data do not match. For instance, the variables about family composition are available only on the neighborhood level, whereas desired data resolution is 100 by 100 m^2. I propose the following solutions: create a *multi-resolution model*. In simple terms, we want to assign the variables available on the neighborhood level to the higher resolution of 100 by 100 m^2. Such assignment is pretty complicated if the difference between the areas is large (a neighborhood can consist of around 50 squares of 100 by 100 m^2. Therefore, to minimize inconsistencies, let us keep working with the data with 500 by 500 m^2 resolution. The assignment will be done as follows. Take a single neighborhood and a variable of interest. From a geometric point of view, this neighborhood is covered by n squares of the lower space. Let us uniformly distribute the variable of interest between these squares. Visual representation of this procedure is depicted with Figure 3.3.

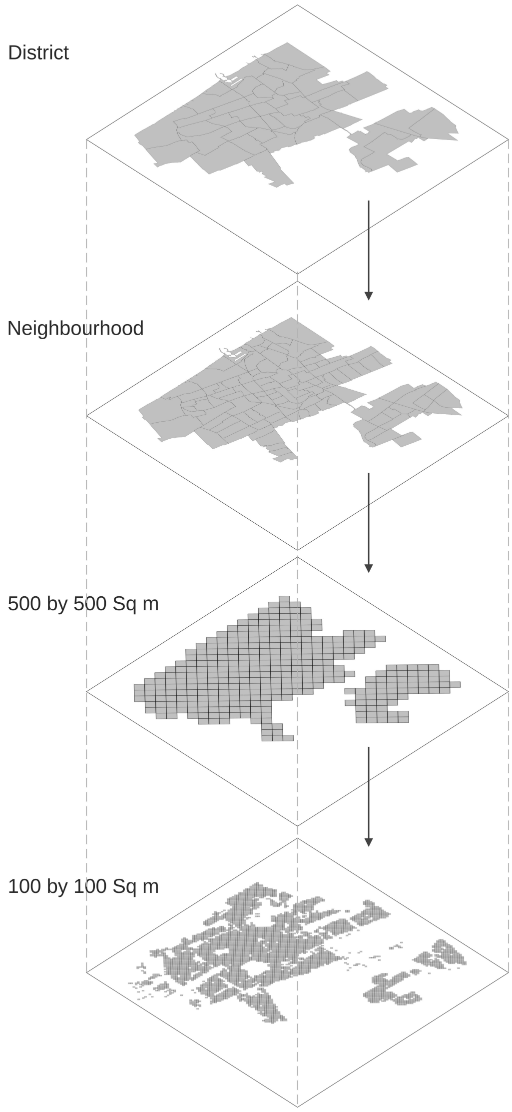

Figure 3.3: Multi-resolution assignment: from district level to 100 by 100 m^2

Emergency network

Emergency calls are an excellent data source for analysis of vulnerability. For this study, we will use the calls collected from *P2000 network*: a one-way communications network used for emergency services in the Netherlands.

There are three hypotheses to test with this data set. The first one relates to the distribution of the calls. It is highly likely that it is unequal over a day, week, and month. For example, there can be more calls during the morning or late afternoon. Second, the calls are made from different geographical locations. The number of calls made from certain parts of the city can be higher, thus, creating certain "hotspots" of high vulnerability. The final hypothesis examines the connection between the number of ambulance calls and the temperature in the city. Whether the assumption: the higher the temperature, the more calls are made, holds.

The data sets discussed previously were preprocessed and published for public use by governmental institutions. Therefore, they can be downloaded, unpacked, and almost immediately used (some preprocessing is still required). Oppositely, the data on emergency calls need to be collected first, second, carefully preprocessed, and only after that used to get an insight. Consequently, it is necessary to describe the data collection process.

There are a variety of websites that are translating calls from P2000 network. `www.`

p2000mobiel.nl was chosen because of its relatively simple interface (the way how the information is presented). The web site does not state any limitations of the usage or collection of its data. Therefore, it was assumed that the data could be collected via
textitweb scraping. Web scraping is a direct collection of information from a web site utilizing a script in any programming language. This script first collects and then transfers the information published on the web site to a server where the script was launched. Table 3.1 represent a set of records scraped from the web site.

№	Call type	Original message	Date	Preprocessed message
1	Ambulance	[[1520031] Ambulance Haaglanden]	2019-06-25 12:16:22	A2 Pisuissestraat SGRAVH : 15131
2	Ambulance	[[1520014] GGD Den Haag-Waldorpstraat (Ambula...	2019-06-25 11:56:28	A1 Huisartsenpraktijk Loosduinse Hoofdstraat S...
3	Ambulance	[[1520021] Witte Kruis Den haag (Ambulance 15...	2019-06-25 11:41:29	A2 Pisuissestraat SGRAVH : 15121

Table 3.1: Sample records of P2000 network

Important to mention that by default, a record does not have coordinates. Instead, it has only a postal code written in the message. In case of The Hague or Haaglanden region, it is 4-digits Postal Code (6-digits Postal Code available for Rotterdam and Amsterdam). Thus, the assignment should be made manually, for example, by finding a middle point of a certain postal code. As a result, there can be a certain shift or bias. To match the postal codes with coordinates, I used one extra data set that has not been discussed yet. *OpenAddress* is a free and open global address collection. Stored online at www.openaddresses.io street names, postal codes, and corresponding coordinates can be easily downloaded in CSV format.

3.2.3. Solutions for resilience capacities

The data sources that are used for the analysis of resilience capacities are depicted in Figure 3.4:

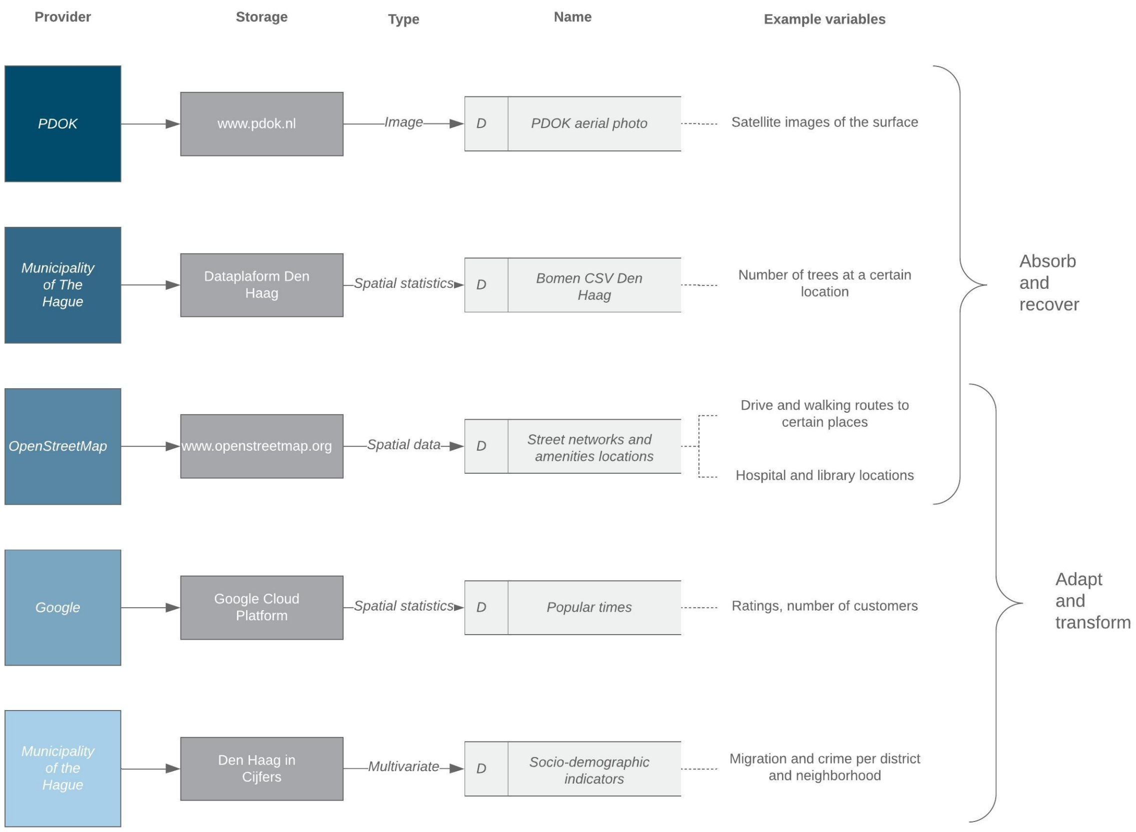

Figure 3.4: Data sources used to analyze resilience capacities

Critical infrastructure

The data provided by OpenStreetMap (OSM) is an excellent source for urban analytics. It has extensive and up-to-date information about street networks and amenities around a specific location.

To analyze the critical infrastructure of the city, at first, we need to collect hospital locations. After that, we can analyze the driving routes of ambulance cars using street network data. The question is: what is the length of the shortest route from a hospital to a vulnerable individual or community? Since we are operating with 500 by 500 m^2 polygons, the distances will be measured from a hospital to the centroid of a polygon. There are multiple hospitals, and numerous driving routes should be analyzed. The result of this analysis is an isochrone map representing the lengths of the shortest paths from any hospital to any polygon in the city.

Let us provide a short note on the data collection procedure. The data can be derived with either an API provided by OSM or via a package written in Python programming language. For this study, a Python package called *OSMNx* is chosen (Boeing, 2017). It is relatively stable and provides a variety of built-in functions to analyze and visualize street networks.

Green infrastructure

The analysis of green infrastructure can be done in two ways: by using satellite images or with statistical information given by the municipality of The Hague. In both cases the aim is to understand to what extent a certain part of the city is covered with vegetation.

The satellite images of the Netherlands are provided by *Publieke Dienstverlening Op de Kaart* (*Public Services On the Map* in English) Publieke Dienstverlening Op de Kaart. These images are stored online in WMS databases at `https://www.pdok.nl/datasets`. To retrieve these images in JPEG format, a Python package *owslib* is used. For the sake of convenience, the whole image of The Hague area is split into n parts.

The second data set created by the municipality of The Hague represents a permanent revision of the trees in the city. It has a variety of attributes such as tree type, the age of planting, and so on. The attributes that are important for this study: the number of trees and their coordinates.

Social infrastructure

The analysis of social infrastructure and corresponding adaptive and transformative capacities is based on three data sources: OSM, Google, and the municipality of The Hague.

OSM data used in the similar to critical infrastructure way. We aimed at collecting library locations and analyze the shortest routes to them. But for now, these routes are walking ones.

The data set from Google, generally called *Popular times data*, provides a set of variables about "popularity" of a certain place: live visit information, wait times, and typical visit duration. Google states that this is aggregated information collected from Google Location History. Thus, if a person has a smartphone with installed Google Maps application on it and a feature to share information about the location on (by default it is on), then the fact that he or she was present at the particular place will be included in the overall statistics. The goal is to come up with an additional measurement of how much a specific library is used. Without a doubt, certain places located in the city center might be used not only by the citizens of the adjacent territories but by tourists, people working closeby. However, taking together with proximity data from OSM will provide us with a better insight about to what extent public libraries are used. Popular times data can be accessed via an API prodvided by Google. Again, similarly to OSM I will adopt a Python package to download necessary variables. The package called *populartimes* and available at `https://github.com/m-wrzr/populartimes`.

As was discussed in Section 2.3, particular factors might degrade social infrastructure and prevent establishing strong social ties. The data sets provided by the municipality of The Hague have such indicators: migration, crime. Besides, since transformative capacity defined as a radical change, for example, change of living place to less heat-sensitive (there is some evidence that recently built housing is more heat-proof), the income and housing prices can provide us with some insight.

Concluding remarks

The data sources discussed above can be classified in the following way:

1. Data about the threat: heatwave, and its amplifier: UHI;
2. Data about citizens: a set of variables representing to what extent they are vulnerable to extreme heat;
3. Data about infrastructure: vegetation, emergency services, and libraries.

All these three data sources sufficiently cover previously defined conceptual frameworks: Vulnerability Analysis Matrix and socio-ecological resilience combined with infrastructure solution. Without a doubt, there are more data sets that are useful in that regards. However, since the practice of operationalizing of urban resilience is lacking, such an approach seems to be an acceptable first step.

Let us conclude by mentioning that a more extensive description of all data sets can be found in Appendix B.

3.3. Methods

The aims of this section are the following. First, to explain the choice of the methods, and, second, to cover the necessary fundamentals of the chosen algorithms.

Different approaches can be proposed to analyze urban resilience. Some of them are more qualitative; some are more quantitative. Quite recently, cities became a part of the data collection race. Some of the data collected for the last couple of years is published online as open-access. Besides the cities, private and non-profit companies such as the OpenStreetMap Foundation, for example, also provide access to their extensive data set collection. Taken together, these facts promote a shift towards using computational methods to deal with urban problems. One the other hand, the last ten years characterized by the rapid development of field know as *machine learning*. Defined as automated detection of meaningful patterns in data (Shalev-Shwartz and Ben-David, 2014), machine learning can be used for better understanding of cities and improve decision-making process. However, often, there is a lack of understanding and practical examples regarding what methods should be used for such novel data sets. Multiplied by the complex nature of cities, the task of using machine learning for urban problems desires extra attention.

The impact of heat waves can be analyzed in the following way. The data from the P2000 network has two important dimensions: time and geographical location. Thus, it can be presented in the format of time series with assigned coordinates. The type of algorithms that are used for the analysis of such data is **time series forecasting**. In simple terms, the tasks of time series forecasting are to a). Extract daily, weekly, or any other time pattern b). Based on this, predict the future values. Therefore, we can adopt time series forecasting to understand how the demand for ambulance services is distributed over time and what is the geography of ambulance calls.

Applied to the VAM machine learning can help to reveal interesting patterns. The concept of vulnerability has a sophisticated structure. As was found in Subsection 2.2.3, there are a variety of factors: demographic, health, behavioral, and so on, define whether a not a person is vulnerable to extreme heat. To handle such a complicated task, we propose to use two types of unsupervised learning techniques: **clustering and dimensionality reduction**. Unsupervised learning is one of the fundamental learning types of modern machine learning. In unsupervised learning, the algorithm is not provided with a set of ”true” labels of classes. The goal for using these algorithms is understand how vulnerable individuals and communities are grouped and distributed across the city.

The three components of SER can be studied as follows. First, the data sets defined previously for green infrastructure allow us to use three types of algorithms: **clustering** and classification for the satellite data, and **geospatial analysis** for the trees data set. Since the computational resources required for classification are quite high, it was decided to use only clustering on satellite images. Without a doubt, such an approach is far from being perfect. It provides us with quite a raw numbers regarding how green is a certain part of the city. However, such ”proof of concept” might be quite useful in case of the lack of computational resources. Second, to analyze critical infrastructure I propose to use network analysis with calculation of short paths lengths from hospitals to vulnerable individuals and communities. The same approach can be used for social infrastructure. In addition, for Google’s Popular times data set I suggest to use conventional algorithms of geospatial analysis.

Finally, to match extreme heat vulnerability (represented by the number of ambulance calls made during the heat wave) and distribution of vulnerable individuals and communities (results of dimensionality reduction algorithms) I propose to use the methods of **regression analysis**.

№	Type of technique	Package name	Reference
1	Time series forecasting	fbprophet, scipy	Taylor and Letham (2018), Jones et al. (2001)
2	Dimensionality reduction, clustering and regression analysis	scikit-learn, yellowbrick, hdbscan	Pedregosa et al. (2011), Bengfort et al. (2018), McInnes et al. (2017)
3	Non-negative matrix factorization	nimfa	Zupan et al. (2012)
3	Network analysis	OSMNx	Boeing (2017)
4	Geospatial analysis	pysal	Rey and Anselin (2007)

Table 3.2: Software implementation of the chosen algorithms

The following subsections explain the necessary fundamentals of proposed techniques and corresponding algorithms.

3.3.1. Time series forecasting

Time series forecasting can be defined as predicting future values of a particular phenomenon based on the previous observations (Azoff, 1994). Opposite to regression techniques, usually TSF does not require a set of predictors; instead, it operates only with the information extracted from the past data.

Time series forecasting has a set of limitations. First, it is very sensitive to the amount of data that was provided. The less data is provided to the algorithm, the less accurate predictions. Second, the predictions can be complicated by outliers. Some algorithms are robust to outliers. That is, they can extract the trend even if the time series curve has a lot of abnormalities and noise.

For this study I propose to use *The Prophet* (fbprophet) time series forecasting model proposed by Taylor and Letham (2018). It is based on decomposable time series model from Harvey and Peters (1990), allows to extract non-linear trends, robust to missing data and outliers.

As was discussed previously, we are going to use time series forecasting for analysis of ambulance calls. Let us finalize the discussion of these type of techniques by providing an example. Consider the following time series: the number of ambulance calls made in The Hague from 23 of May to 1 of June 2019 (see Figure 3.5)

Figure 3.5: Number of ambulance calls made in The Hague from 23-05-2019 to 01-06-2019

The goal is to understand the daily patterns and come up with a prediction of future demand. Let us make a prediction for next four days or 96 hours. As can be seen from Figure 4.9 the algorithm is able to catch the overall trend (depicted with the blue line). However, since the number of data points is relatively low, the actual values (black dots) are often lay above the uncertainty interval (blurred light blue curve).

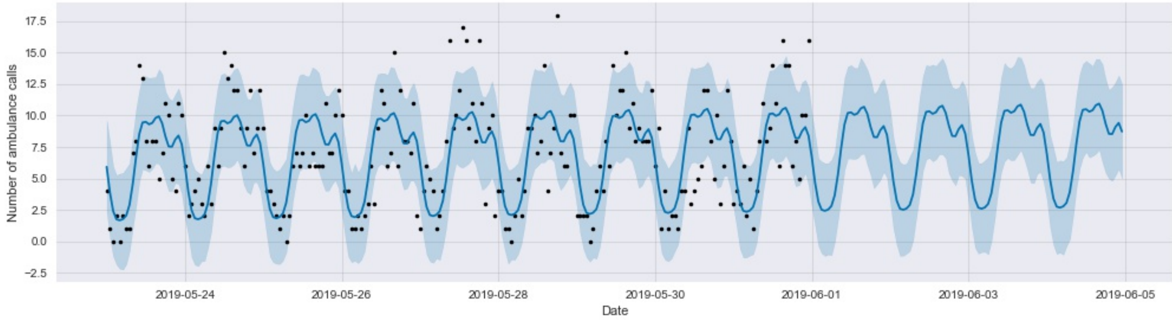

Figure 3.6: Prediction of ambulance calls by fbprophet

The final figure demonstrates the decomposition of the time series into the trend and daily pattern. It seems that there are hills and caveats. Not surprisingly, there are fewer calls during the night. However, what is interesting is that the curve has a certain number of peaks. Such a result provides us with an insight into how vulnerability distributed over time.

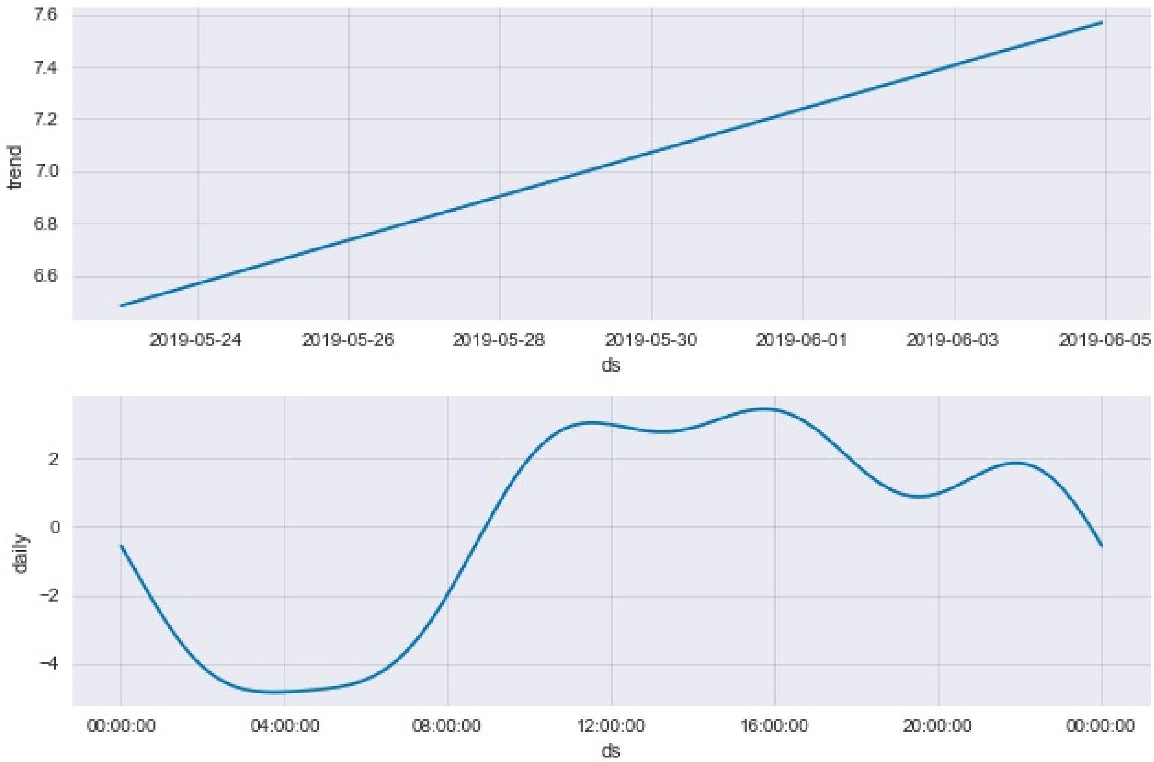

Figure 3.7: Extracting trend and daily pattern with fbprophet

3.3.2. Dimensionality reduction

Dimensionality reduction (DR) is a common technique of unsupervised machine learning. The task of DR is to **reduce the number of variables to consider**. For example, if the original sample of two observations has three attributes, they can be mapped into three-dimensional space. For the sake of simplicity, let us assume that these observations have converted into cylinders (see Figure 3.8). By reducing the number of dimensions n from 3 to 2, they transformed into circles, and after setting up n to 1, to segments.

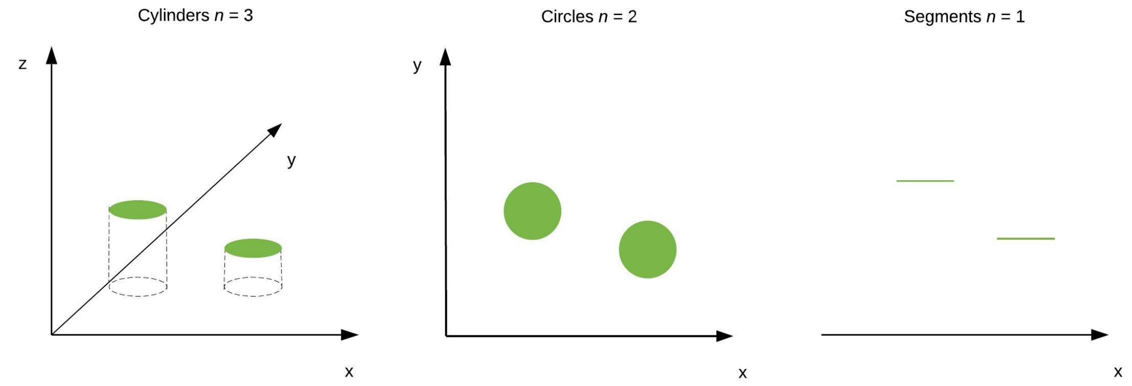

Figure 3.8: The task of dimensionality reduction

Typically, machine learning problems involve many variables, from hundreds to thousands. In some cases, the information represented by specific variables can be treated as unnecessary or redundant. For example, one of the main applications of DR is to compress images. By reducing the initial number of variables, the weight of the image file will be reduced as well. However, the quality of the output can still be satisfactory. The Figure 3.10 demonstrates how *Non-negative matrix factorization* algorithm works on image data. From left to right, the initial number of dimensions was reduced to 5.

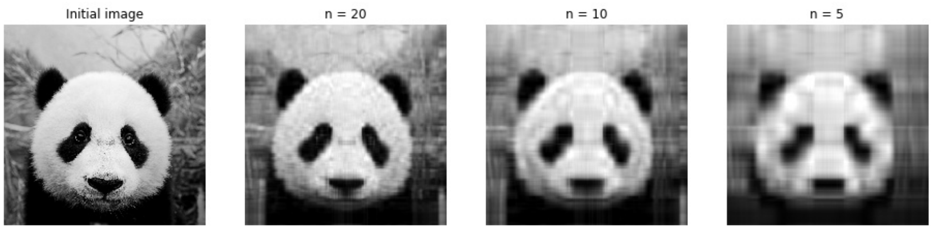

Figure 3.9: Example of non-negative matrix factorization on image data

Some variables in the initial data set can be correlated. Therefore, by merging those, there is no loss of information, to a certain extent. Another point of using DR is to clear up data set and reduce the noise and thereby improve the performance of the next algorithm in the pipeline (usually, supervised learning algorithm: regression or classification).

There are two main approaches top dimensionality reduction: *projection* and *manifold learning*. The key assumption of the projection approach is that all initial observations lie within a much lower-dimensional subspace of the high-dimensional space. Manifold learning relies on the *manifold assumption* which holds that most real-world high-dimensional data sets lie just *close* to a much lower-dimensional manifold.

The input of any DR algorithm is the original data set of n observations with m attributes. The main parameter of all algorithms independent of the approach is the number of dimensions d to which m should be reduced. The output is composed data set of n observations with d attributes. The evaluation of performance depends on the algorithm. Often, there is a need for expert validation, or at least the resulted variables should be checked by the researcher.

DR algorithms can be used for the VAM as follows. Since the number of individual, household and environmental attributes is relatively high (at least it is hard to operate with the initial m attributes), there is a need for reduced set of k variables. Besides, many of these attributes are correlated what makes it hard to perform regression analysis on it. The merged variables will contribute to a **profile** of the individuals living at a certain place. By a profile, I mean a specific combination of predominant characteristic. For example, middle-age, Turkish, high-income, with kids. The assumption is that the city consists of a certain number of unique profiles.

Necessary Principal Component Analysis

Principal component analysis (PCA) is one of the most popular dimensionality reduction algorithms (Jolliffe and Cadima, 2016). The earliest publications about the PCA dated to the early twentieth century by Pearson (1901) and Hotelling (1933). PCA is a projection method. Therefore, it first seeks for the hyperplane that lies closet to the data, and the projects the data onto it.

PCA creates a set of new orthogonal, and, therefore, uncorrelated, variables called *principal components*. These variables can be interpreted as directions of maximum variance.

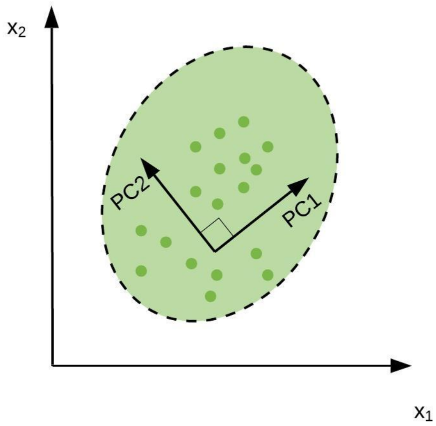

Figure 3.10: Contrasting original axes x_1 and x_2 with principal components PC1 and PC2

In short, in order to perform PCA we need to construct a $n \times m$-dimensional transformation matrix W. This matrix allow to map a sample vector x onto a new d-dimensional feature subspace:

$$x = [x_1, x_2, \dots, x_m], \quad x \in \mathbb{R}^m$$

$$\downarrow xW, \quad W \in \mathbb{R}^{m \times d}$$

$$z = [z_1, z_2, \dots, z_d], \quad z \in \mathbb{R}^d$$

The resulted first principal component will have the largest possible *explained variance*. Explained variance stands for the amount of variance accumulated in the component. The more the variance, the more import is the component for recreating the overall picture. Therefore, the second and all subsequent components will have less variance than the first one.

The first step of the algorithm is to construct covariance matrix. For example, covariance between two features x_j and x_k is equal to:

$$\sigma_{jk} = \frac{1}{n} \sum_{i=1}^{n} \left(x_j^{(i)} - \mu_j\right)\left(x_k^{(i)} - \mu_k\right)$$

where μ_j and μ_k are the sample means. In this terms covariance matrix for three features equals:

$$\sum = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}$$

The eigenvectors of this matrix are principal components and the corresponding eigenvalues define their ”magnitude.” After eigenvectors and eigenvalues were defined, the final step is to obtain eigenpairs. An eigenvector v satisfies the following condition:

$$\sum v = \lambda v$$

where λ is a scalar and eigenvalue. From here we can use λ_j to analyze the performance of PCA algorithm. The following equation defines variance explained ratios:

$$\frac{\lambda_j}{\sum_{j=1}^{d} \lambda_j}$$

The ”right” number of dimensions is usually defined with *elbow method*. The method points on how many dimensions should be chosen to cover 95% of the variance. Often, the graph kind of breaks in this point and create an elbow.

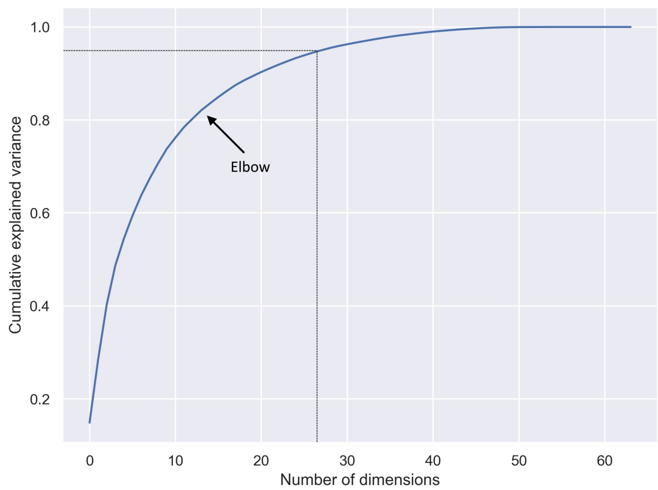

Figure 3.11: Demonstration of the elbow method for PCA

Like any other method, PCA has its limitations. One of the crucial ones it relies on the assumption of linearity. Secondly, it relies on orthogonal transformations. Currently, there are many modifications of the algorithm that can overcome those (see Jolliffe and Cadima (2016)).

Non-negative Matrix Factorization

Non-negative matrix factorization is a relatively novel algorithm for dimensionality reduction (Wang and Zhang, 2013). In some sense, NMF works similarly to previously discussed PCA: each data point is represented as ”weighted” sum of some higher-dimensional components. The difference is that these components must be non-negative, whereas for PCA just orthogonal. Remarkably, since NMF operates with only non-negative components it is easier to interpret the results.

In simple terms, NMF factorizes (breaks down) the feature matrix into matrices representing the latent relationships between observations and their features.

Let X be the original matrix $n \times m$ composed out of observations and corresponding attributes. The task of NMF is to find a solution for $V \approx WH$ where W $n \times r$ and H $r \times m$. Thus, each column of V, v_i, can be found through:

$$v_i = Wh_i$$

where h_i is i column vector of matrix H.

The process of finding good approximation is optimization problem with respect to W and H. Formally we have

$$\min_{W\in\mathbb{R}^{n\times r}, H\in\mathbb{R}^{r\times m}} \|X - WH\|, \quad \text{such that} \quad W \geq 0 \text{ and } H \geq 0$$

This problem can be solved by using in a couple of ways. For instance, *alternating least squares* or *multiplicative update rules* methods (Lee and Seung). In simple terms, the first one is using initial guesses, random number, for W given H is correct and other way around.

Ding et al. (2005) noted that NMF is equivalent to spectral and K-Means clustering.

One of the complications of NMF is evaluation of its performance. Opposite to PCA it does not have a metrics similar to explained variance.

Other algorithms and metrics

Apart from PCA and NMF, in this study I used the following dimensionality reduction algorithms: *Factor Analysis*, *Singular Value Decomposition*, *Independent Component Analysis* and a variety of implementation of NMF. More information about these algorithms can be found in Pedregosa et al. (2011), Zupan et al. (2012). The metrics that are used to evaluate the performance of these algorithms are also extensively discussed in the sources mentioned above.

3.3.3. Clustering

Clustering is one the techniques of unsupervised learning machine learning. In simple terms, clustering is **grouping of similar object into a certain number of sets** or clusters. The grouping is done based on some rule, for example, based on the density: if there are a lot of data points within a certain area - it is a cluster, and the rest treated as a noise. The number of clusters can predefined by a user or can be defined by the clustering algorithm itself.

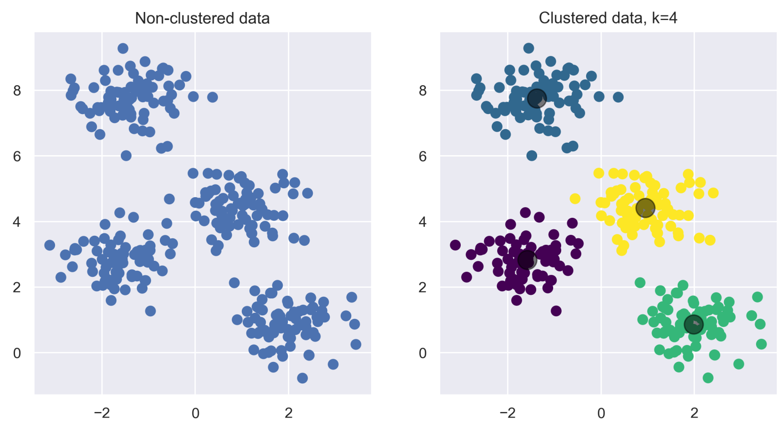

Figure 3.12: The task of clustering

Further, I will use clustering to operationalize VAM in the following way. As was discussed in the previous section, the result of any dimensionality reduction algorithm is a composed set of individual attributes or profiles. Each of these profiles is distributed over the city. In some places, one profile is predominant, in other areas another one. When we apply clustering on these profiles, we will see **how specific profiles are grouped together**. The hypothesis is that certain parts of the city will be highly homogeneous clusters. This is undesirable if a homogeneous cluster is represented by a vulnerable profile and takes up a lot of space.

Necessary K-Means

The K-Means algorithm is one the most popular and simple ones (Jain, 2010). First published in 1982 now it is a point of departure of any cluster analysis. K-means refers to partitioning algorithms, These algorithms are based on specifying an initial number of clusters, and then iteratively reallocating objects among clusters to convergence.

Other algorithms and metrics

Apart from K-Means, in this study I used the following clustering algorithms: *Mini Batch K-Means*, *Affinity Propagation*, *Mean Shift*, *DBSCAN*, *HDBSCAN*, *Spectral Clustering*, *Agglomerative Clustering*, *OPTICS*, *Birch* and *GaussianMixture*. More information about these algorithms can be found in Pedregosa et al. (2011), McInnes et al. (2017). The metrics that are used to evaluate the performance of these algorithms are also extensively discussed in the sources mentioned above.

3.3.4. Regression

Regression is a type of statistical analysis used to measure relationships between a set of variables. In this study, I will use regression algorithms to identify relations between the number of ambulance calls and the attributes of the citizens within a specific area. For example, if the amount of highly vulnerable people within a certain area is high, then the number of ambulance calls made during the heat wave should also be high. Besides, there is an essential factor affecting this: the presence of urban heat island.

In this study I used the following regression algorithms and techniques: *Multiple-linear regression*, *Ridge*, *Elastic Net*, *Support Vector Regression*, *Huber Regressor*. More information about these algorithms can be found in Pedregosa et al. (2011). The metrics that are used to evaluate the performance of these algorithms are also extensively discussed in the source mentioned above.

3.4. Computational framework

Conceptual framework, data, and methods joined together create a *computational framework*. A computational framework should provide us with a clear and understandable way how the concept of urban resilience can be operationalized utilizing machine learning algorithms. In some sense, a computational framework should be a representation of a data-driven assessment model and help to evaluate the current state of the urban system and estimate its future performance. Formally we have two conceptual frameworks: one is aimed at vulnerability and another one at resilience capacities, there should be two computational frameworks (see Figure 3.13 and Figure 3.14).

3.5. Reusable machine learning study

Reusability of computational models is complicated due to a couple of reasons, first, because of the high prices. To get statistical software or access to the full potential of a programming or simulation language, you need to buy a license that is usually pretty expensive. Second, due to the way the project is organized, and the code is written. Without a clear template, extensive documentation or well-organized API, it is highly likely that potential users will be scared away. Finally, the data used in the study should be appropriately stored and easily accessible.

To overcome these problems, it was decided to use the following approach. First, as was briefly discussed above, the software implementation of computational models will be made with Python programming language. It is high-level, general-purpose language, and it is freely distributed and accessible for a single installation from `www.python.org` or with the Anaconda distribution at `www.anaconda.com`. Currently, Python is one of the most popular programming languages in the areas of data science, machine learning, and software engineering. Thus, it is likely that on the one hand, the next users will get familiar with the way how the study was made, but on the other side, they will get the access to it for free.

There are different approaches to code organization. In short, the fields of data science and machine learning is aimed at creating certain flexible standards. IBM, Google, Amazon propose their templates and guidelines. The solution that was chosen for this study is called *Cookiecutter* (Driven Data Labs, 2019). Cookiecutter is an open-source solution for creating project structures. The structure generated by Cookiecutter by default has the following folders: data, docs, models, notebooks, etcetera. Such simple at first glance folder organization has a long story of improvements, modification, and critique. Besides the overall structure of the project, the machine learning project itself needs to be organized a set of consecutive steps. Again, there are various ways to conduct a study. The most common approaches are: *CRISP-DM* and *OSEMN*. Géron (2017) proposed one more great template in the form step-wise approach. The common ground between these three was found in the following steps: problem understanding, data gathering and preprocessing, future preparation, modeling, and interpretation of results. That is, I split assessments into separate folders. Each folder has a set of notebooks that have names of corresponding steps. Each notebook is supervised with necessary comments on how and why certain actions were made.

All data sets that are used in this study are open-access; therefore, they can be collected by any party. However, some may find the data collection procedure complicated. Besides, it is essential to provide two types of data: raw and processed. Since sometimes, processing can be done differently and significantly affect the results, any researcher can have a possibility to do in his or her way. To satisfy these needs, I stored the main data sets in two folders: *raw* and *processed*.

The following sections of this chapter will present the main results of the analysis. For the sake of saving space, specific preprocessing will not be discussed. However, in case of doubts regarding a particular procedure, the answers can be found on `www.github.com/mikhailsirenko/urban-resilience-ml`.

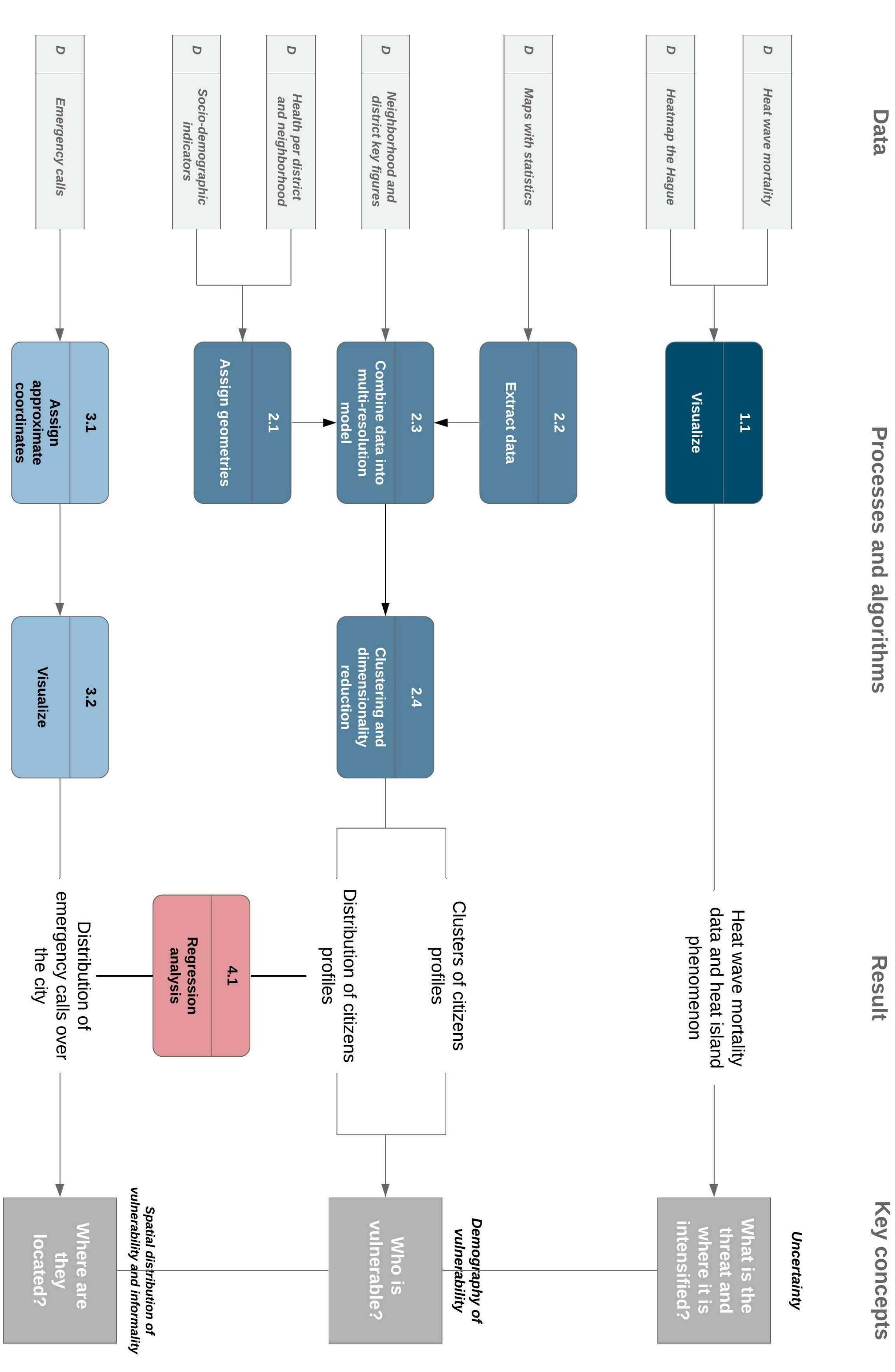

Figure 3.13: Computational framework for Vulnerability Analysis Matrix

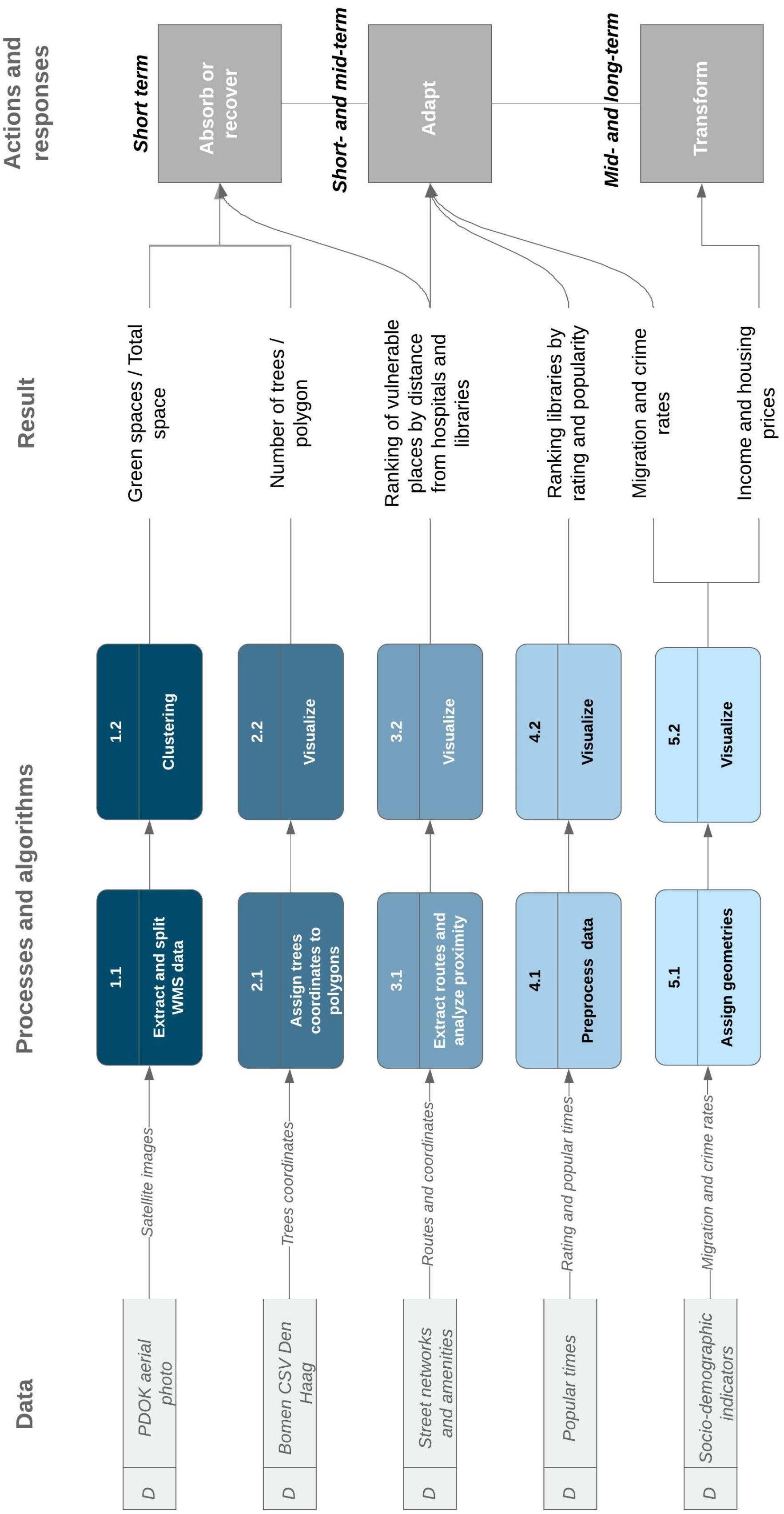

Figure 3.14: Computational framework for resilience capacities assessment

3.6. Conclusion

This chapter had three goals, namely:

1. Propose a way to transfer identified conceptual frameworks into a set of measurable indicators;

2. Discuss open data sources that are sufficiently cover these indicators;

3. Introduce the methods that can be used for the analysis of these data sets.

The discussion of these points should result in a computational framework, and, therefore, the answer to the second research question.

Operationalization of Vulnerability Analysis Matrix implies, first, understanding of heat waves and urban heat islands, and, second, analysis of socio-demographic and socioeconomic variables associated with inhabitants of the city and their locations. The following variables can cover the first component: mortality, the average temperature across the city, and the temperature measured locally. The variables of the second one: age, ethnicity, family and household composition, income, education, occupation, health. Besides, the data about ambulance calls can be used to analyze vulnerability. It especially insightful to compare the distribution of ambulance calls during a heat wave with an average day.

Operationalization of the framework of socio-ecological resilience in combinations with solutions (SER) is more complicated. Critical infrastructure can be assessed via proximity metrics or how close are vulnerable individuals and communities to hospitals. Besides, the number of ambulance calls made during a heat wave provides us with a good indication ща potential demand during crisis events. The variables for green infrastructure are the number of trees planted within the city and vegetation depicted at satellite images. The final component, social infrastructure, assessed, again, with proximity and with Google's Popular times data. The last data set includes ratings, the average number of clients, etcetera.

To analyze these data sets, four types of techniques are proposed: time series forecasting, dimensionality reduction, and regression. Besides these major types, some methods geospatial and network analysis are used. Time series forecasting aims to extract trend, daily and weekly patterns. Also, try to predict the number of ambulance calls based on past data. By means of geospatial analysis, I will try to match the ambulance calls made during the heat wave with the locations of urban heat islands. The goal of dimensionality reduction algorithms is to decompose numerous indicators into uncorrelated sets - profiles. These profiles will demonstrate the distribution of vulnerable and non-vulnerable individuals over the city. Clustering algorithms provides us with an opportunity to group these profiles to understand the complex demography of the city better. To what extent vulnerable citizens are mixed with non-vulnerable? The final type of algorithms, regression analysis, is used to match the profiles with the number of ambulance calls. To what extent extreme heat vulnerability (presented by socio-demographic and socioeconomic attributes) results in the ambulance calls during the heat wave. The study of resilience capacities is based on street network analysis along with geospatial methods.

The data and the methods joined together create a computational framework and provide the answer to the second research question.

4
Results

This chapter aims to present the results of the case study analysis. That is, in this chapter, I apply the proposed computational framework to understand: how resilient is the Hague to extreme heat. It starts with a discussion of recent heat waves, reports time series forecasting and outputs of geospatial analysis on ambulance calls. In addition, it explores the contribution of urban heat island effect into the number of ambulance calls. The second section reports the results of the vulnerability assessment. Here dimensionality reduction and clustering algorithms are applied to identify vulnerable individuals and communities and their distribution across the city. The discussion of vulnerability is followed by resilience capacities assessment. Geospatial and network analysis are utilized to understand how various parts of the city differ in terms of vital infrastructure. Thus, the chapter seeks to answer the third and fourth research questions.

4.1. Analysis of heat waves and urban heat islands

Background information

In Chapter 1 of this manuscript we have already extensively discussed the history of heat waves in the Netherlands. Before 2019, there were three major cases: in 2006, 2010, and 2018. At the moment of finalizing this thesis a new 6-days heat wave covered Dutch cities: from 22 to 27 July 2019. This time the heat was so strong that some cities renewed their temperature records. The Royal Netherlands Meteorological Institute (KNMI) issued code orange and The Netherlands National Institute for Public Health and the Environment (RIVM) activated National Heat Plan (Rijksinstituut voor Volksgezondheid en Milieu, 2019). However, according to the preliminary numbers reported by CBS the heat wave has devastating consequences. The heat wave caused more than 400 extra deaths from all over the Netherlands. Hence, the 2019 heat wave caught up with the heat wave of 2006 in the damage made to society. Figure 4.1 depicts the maximum daily temperature during the heat wave:

Ambulance calls

The concept of vulnerability from Vulnerability Analysis Matrix (VAM) states that **susceptibility to harm distributed unequally over time and space**. That is, vulnerable individuals and communities can be affected differently: during mornings or nighttime; at home or workplace. The data on ambulance calls can provide us with an opportunity to reflect on both. First, the calls are made at a particular time. The calls have time stamps. Second, a person asks for help from a specific location. For this, we info have postal codes. 4-digits or 6-digit Postal Codes are written in the message. The goal of this analysis is to understand time and space differences. I will start the from general pattern and after that proceed to the study of the calls made during the heat wave.

The data on ambulance calls was collected for the period from 23 May to 1 August, 2019. Apart from The Hague, the data was collected for Rotterdam, Amsterdam and Eindhoven.

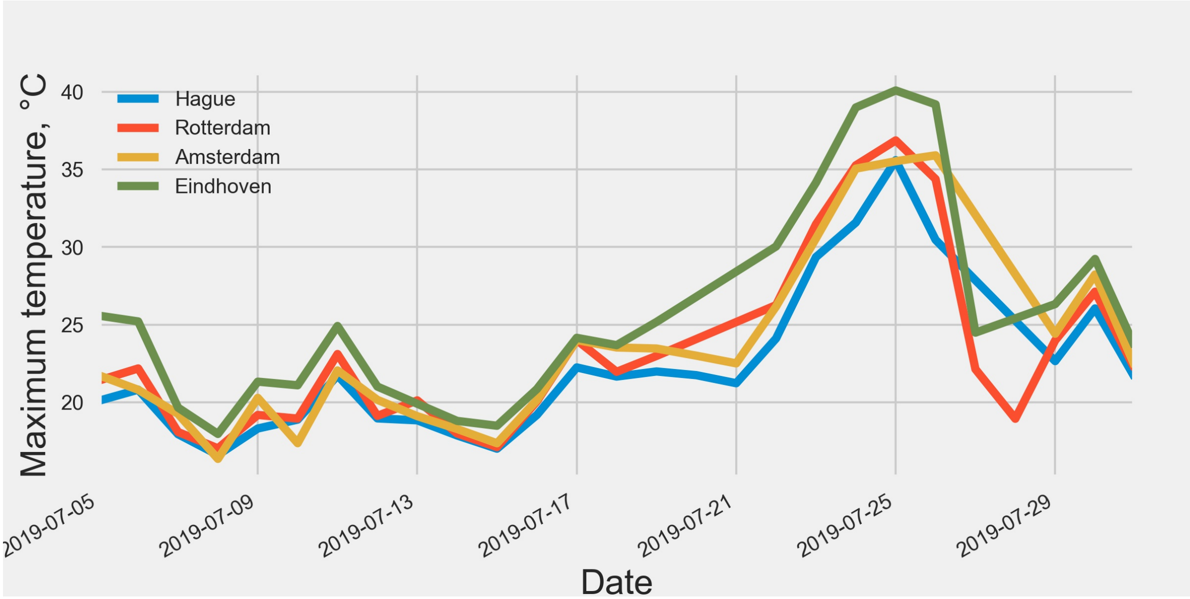

Figure 4.1: The changes in temperature during the heat wave between different cities. The data is collected with Dark Sky API (The Dark Sky Company, 2019)

Temporal patterns

Let us start the analysis from the time factor: *how are the ambulance calls distributed over day, week and month?* Figure 4.2 shows the distribution of calls over a day averaged over a month. The number of data points collected for each of the months is different. Therefore, there are some "discrepancies" on the x-axes (it holds only for May and August). Even though this figure provides us with an exciting insight, we can observe certain regularities. It seems that there are at least two flat peaks (for May and June): the first is around 14:00, and the second close to 22:00. Somehow, these peaks indicate at what time vulnerability is the highest. Consequently, during these times, the demand for ambulance services is also the highest. The data of July is slightly different, and it pushes us to explore the second type of pattern: the monthly one.

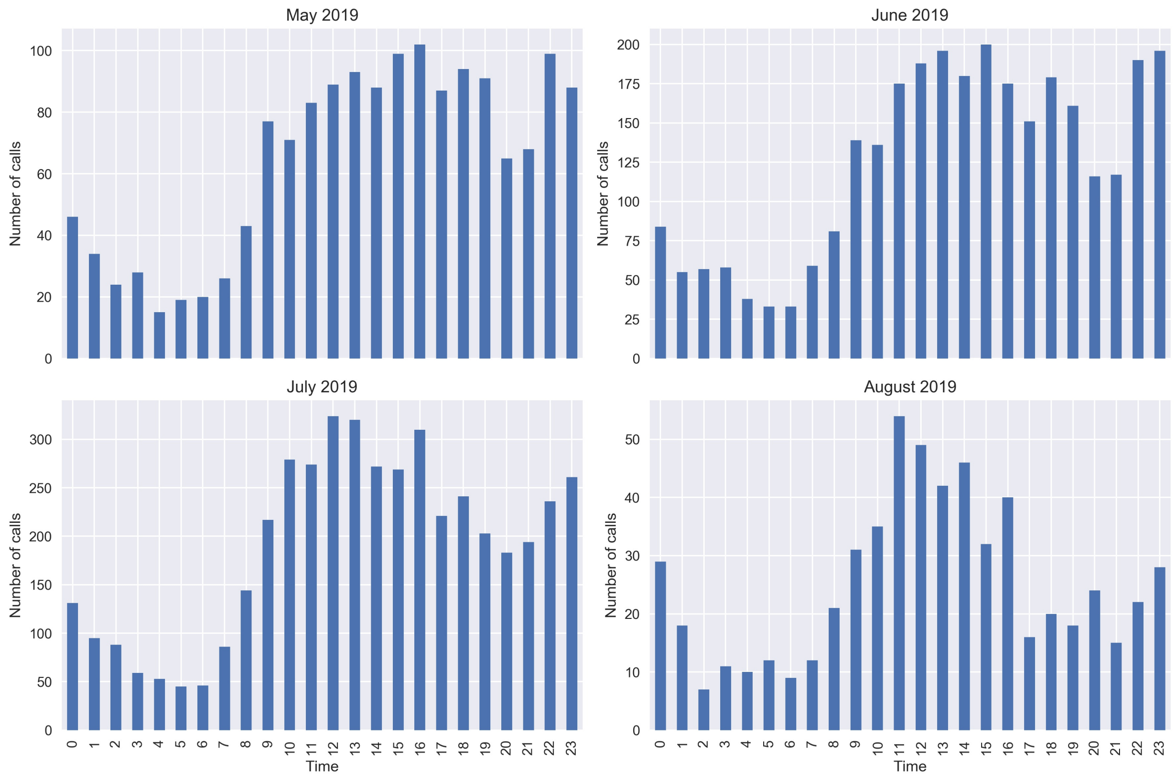

Figure 4.2: Average number of ambulance calls made during the day for May. June. July and August 2019 respectively

Figure 4.3 how many ambulance calls were made in the Hague during May and June 2019. It is interesting that there a certain average value that keeps for many days. This value equals to 180. However, there is a single peak on June 25. As we will see later, the temperature on this day was higher than on average. This is one of the possible explanations for such an increase.

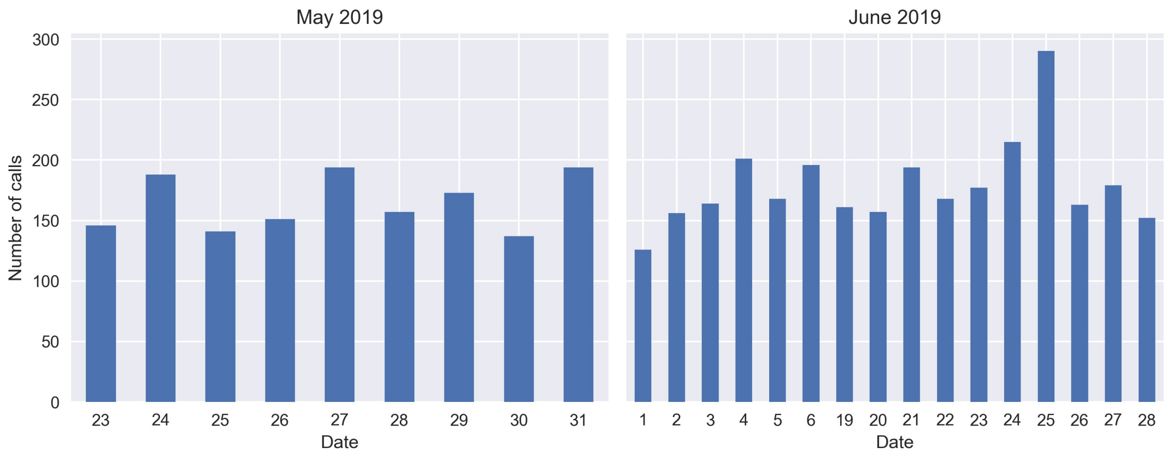

Figure 4.3: Distribution of ambulance calls made in the Hague in May and June 2019

The observations for July 2019 when plotted monthly has a striking difference (see Figure 4.4). The number of ambulance calls made since 23 of July has almost doubled! That is a clear indication that an extreme event happened. Let us note that the data for 26 of July is incomplete for The Hague due to the failure of data collection script. Based on the data for Rotterdam and Amsterdam I expect it to be not less than 350 (similar value of 24 of July).

As we know, this is the heat wave hit the Netherlands at that time. We will elaborate on it in the following subsection.

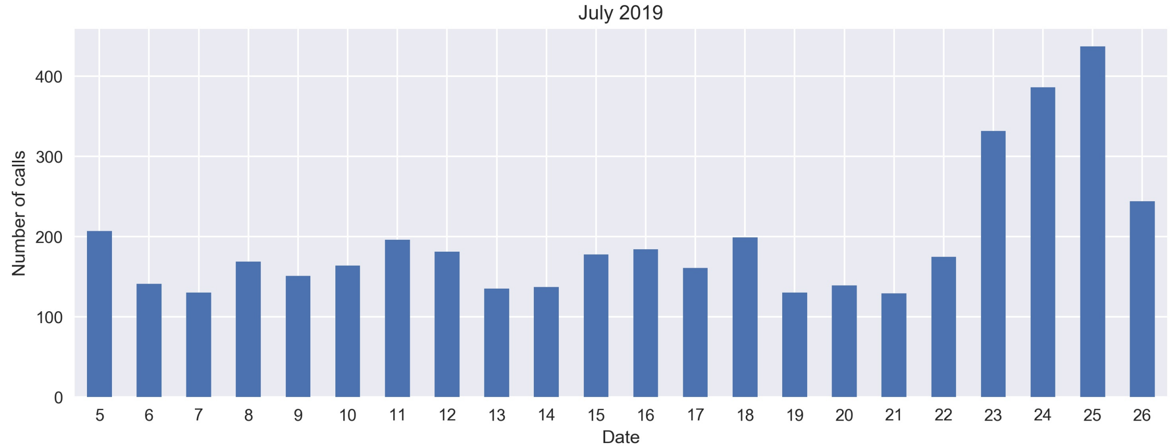

Figure 4.4: Distribution of ambulance calls made in the Hague in July 2019

Before proceeding to time series forecasting, let us explore the heat wave data in more details. Figure 4.5 provide us with a better understanding of how the beginning of July differs from the heat wave period.

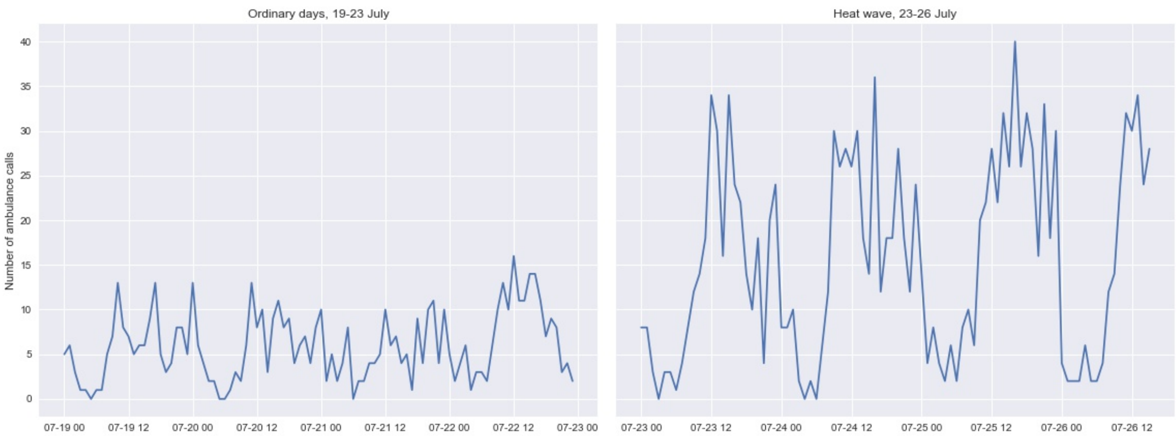

Figure 4.5: Comparison of a relatively ordinary section with the heat wave period

It is interesting to see how the number of ambulance calls made during the heat wave days varies from the average values. The hypothesis here is that the need for help will increase during the night periods. The first figure of the second row demonstrates such a pattern. As the temperature has risen, the more calls have been registered. The heat has accumulated, and some people have started to experience problems. In addition, during the heat wave period, one more peak has arisen: around 22:00.

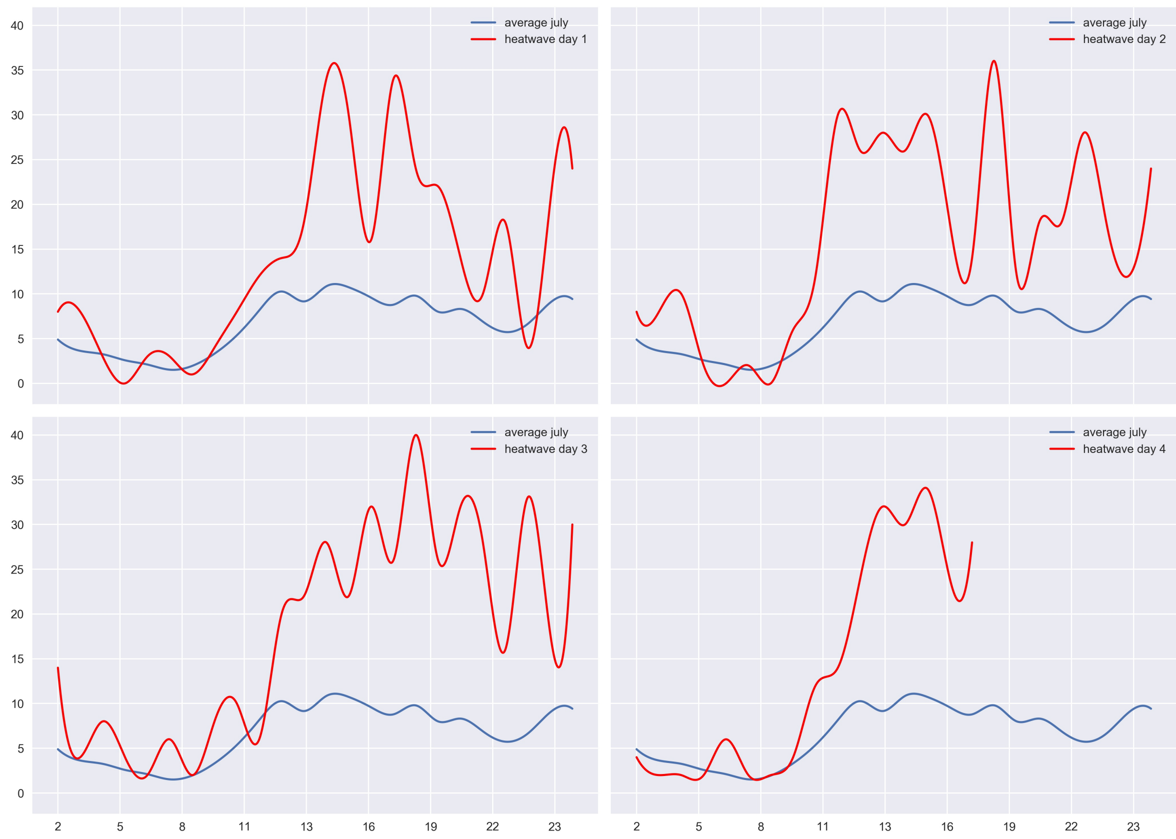

Figure 4.6: Comparison of the calls made during the heat wave wit the average values of July 2019

Time series forecasting

Now let us utilize methods of time series forecasting to get more method-based evidence in the existence of patterns. The analysis of time-series usually starts from calculating *autcorellations* or generating *autocorellation plots*. In simple terms, autocorrelation represents the similarity between data points as a function of the time lag between them. Autocorrelation used a preliminary metric to make sure that time-series is not purely random. Figure 4.7 shows autocorrelation plots for three consecutive months. These plots can be read in the following way. If the observations are purely random, no pattern can be found. In the case of May 2019, the plot creates a clear pattern. However, since June has some missing values and July has a set of peaks caused by the heat wave, the plots do not demonstrate such a distinct similarity as for May. Thus, our hypothesis about a presence of a pattern derived from Figure 4.2, now strengthen by a statistical measure.

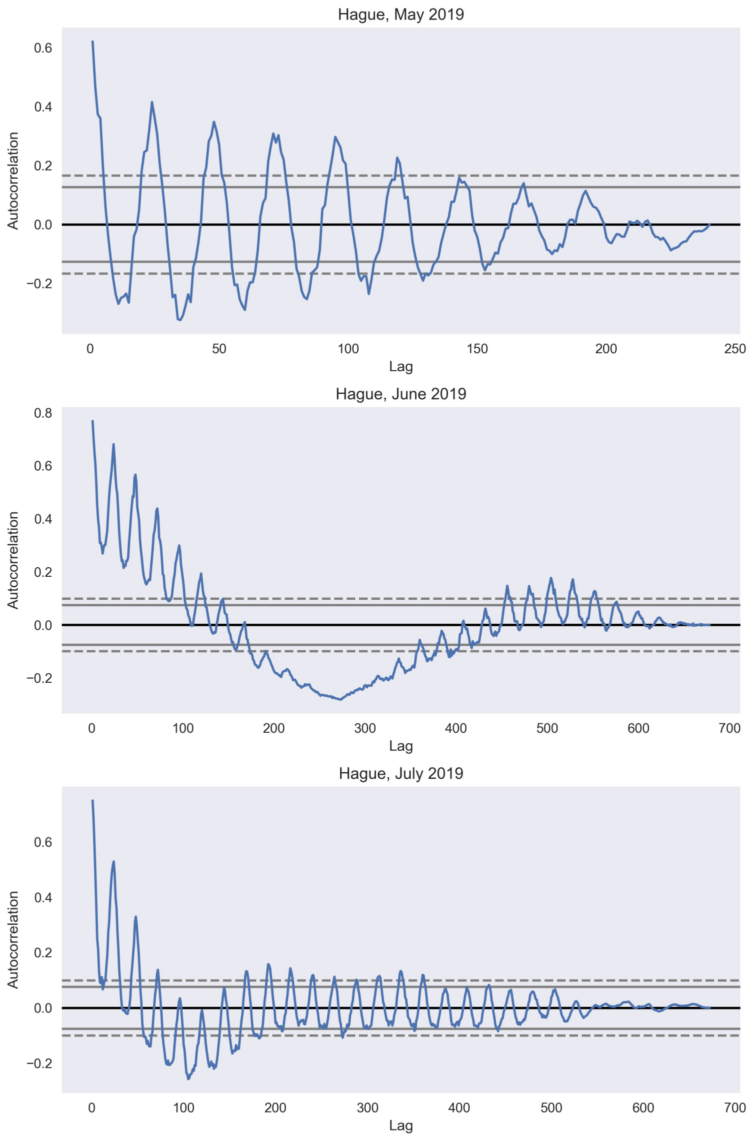

Figure 4.7: Autocorrelation plots based on the monthly ambulance calls data for the Hague

The next step is to apply the algorithms of time series forecasting provided by *fbprophet* infrastructure. The goals remain the same: a). Extract the trend and the patterns; b) try to predict the future number of ambulance calls based on past data. The input into the algorithm is time-series data for either a single month or a set of consecutive months. The specification and limitations of the algorithm are extensively discussed in Taylor and Letham (2018).

Let us take 14 days as an example: from 23 of May to 7 of June 2019. Conventionally, we split this data set into two parts. The first serves to train the algorithm, and the second has the real values to compare with algorithm outcomes. The data looks as follows:

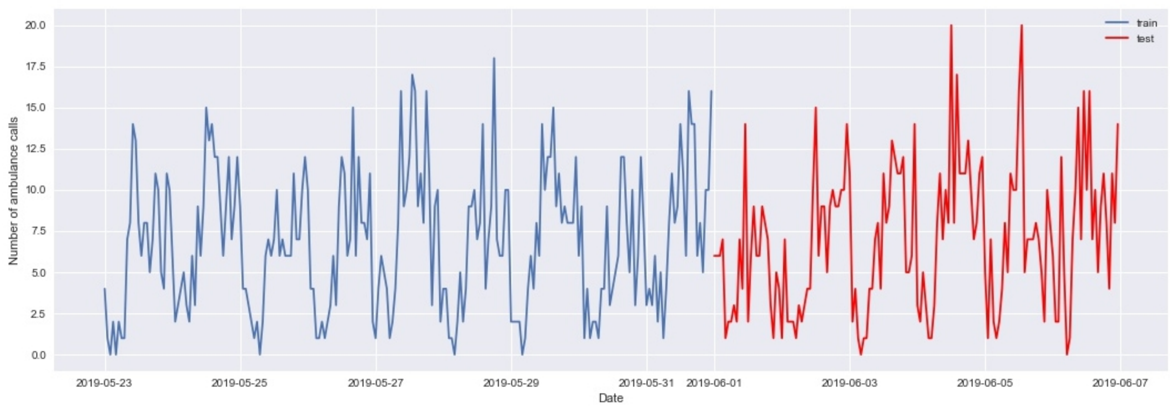

Figure 4.8: Time-series data used for forecasting from 23 May to June 7

As can be seen, this segment is relatively stable and do not demonstrate some contrasting peaks. The prediction will be made for 6 days. Besides, we need to extract the trend, daily and weekly patterns. The results summarized with the following figure:

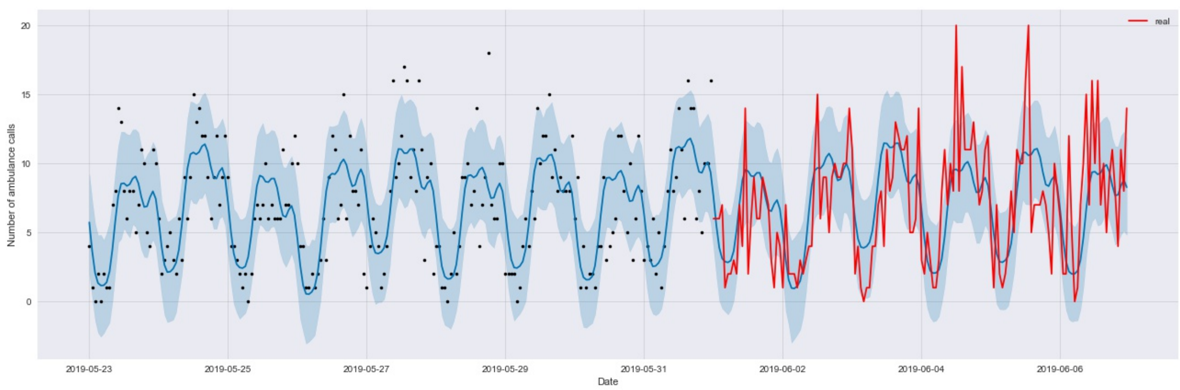

Figure 4.9: Prediction of 6 days made by fbprophet

This graph requires explanations. The left part of it represents the training data. The black dots are the real values, the blue curve is a prediction, and the blurred blue line is uncertainty ranges (a metric demonstrating the ranges of possible deviations). The red curve on the right part of it show real values superimposed on predictions. Based on such a limited data set, the algorithm demonstrated highly satisfactory results.

The next set of graphs show time-series decomposition. As expected, there is a stable daily pattern: the highest peak during the noon with two lower bumps around 16:00 and 22:00. Extraction of the weekly pattern is complicated, and of course, the curve of the trend cannot be used for the reasoning.

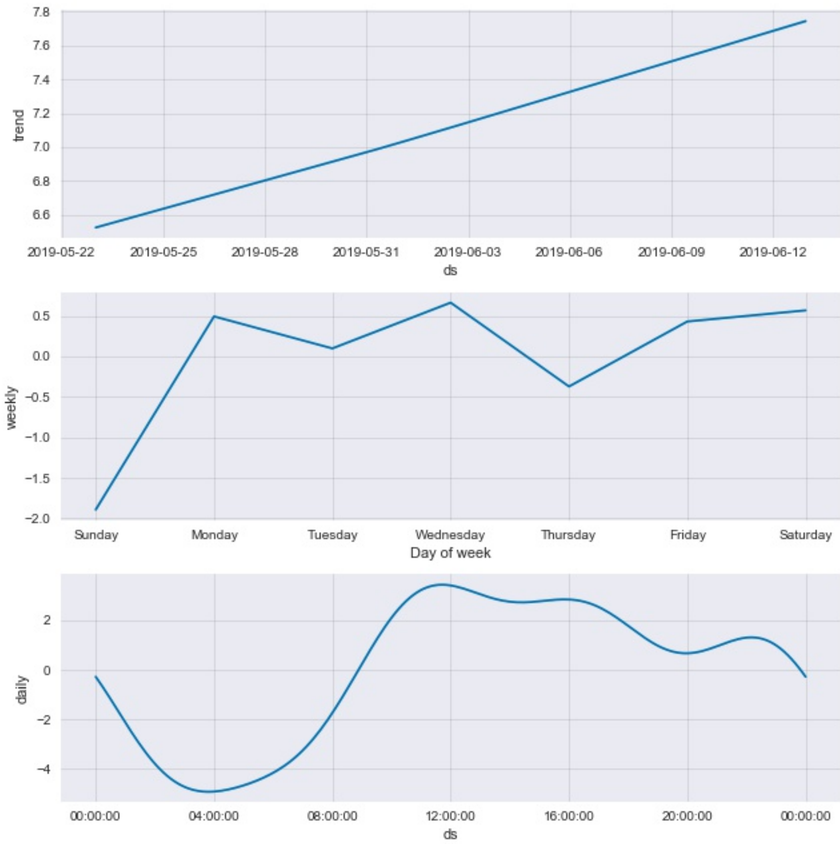

Figure 4.10: Time-series decomposition

The next task is to test the performance of the algorithm on the data collected during the heat wave. Figure 4.11 represents the initial time-series data.

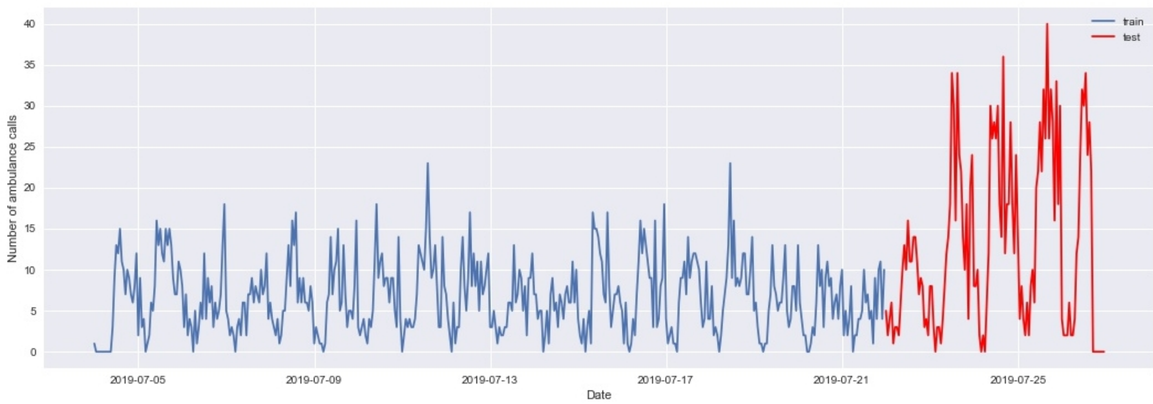

Figure 4.11: Time-series data used for forecasting from 5 July to July 26

As can be seen from the graph, there is a couple of striking peaks from July 22 to July 26. Now, the predicted values are very far from being accurate.

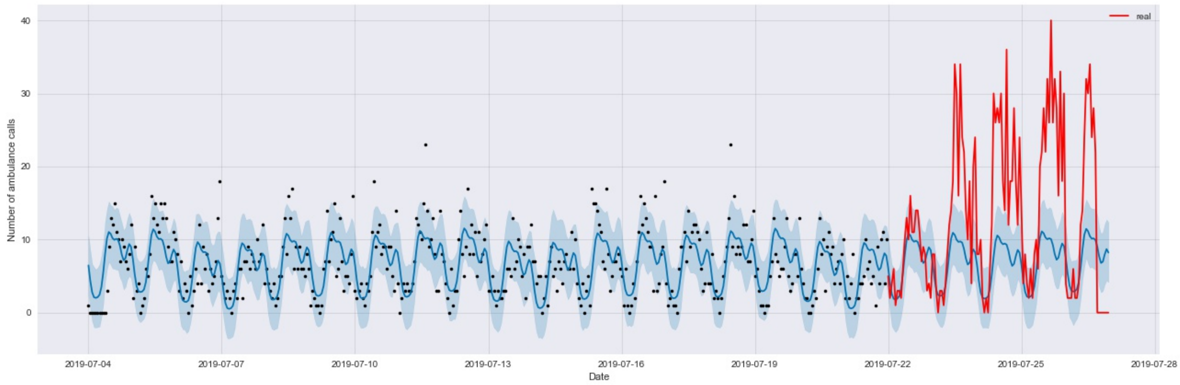

Figure 4.12: Prediction of 3 days (heat wave) with fbprophet

The conducted analysis provided us with three important takeaways. First, the number of ambulance calls made during the recent heat wave is strikingly different from the average days. It is twice as much for The Hague. Such a difference can serve as an indicator of need in an increase of hospital capacities. Second, there are certain times of day when the demand for ambulance services is higher: 12:00, 16:00, and 22:00. Again, it can be used for optimizing the work of critical infrastructures. Finally, quite advanced time series forecasting algorithms should be used quite carefully. Stand-alone learning on the past in case of such sudden external shocks seems to be unreliable.

The performance metrics can be found in the online repository.

Geographical patterns

Let us start from examining a general pattern, or *how ambulance calls distributed across the city on average* and then contrast it with *how the distribution of ambulance calls has changed during the recent heat wave*.

According to the vulnerability concept, the calls should be distributed unequally; certain parts of the city will have more requests than others. Since the total number of calls collected for almost three months is almost 20000, it is hard to infer some insight immediately.

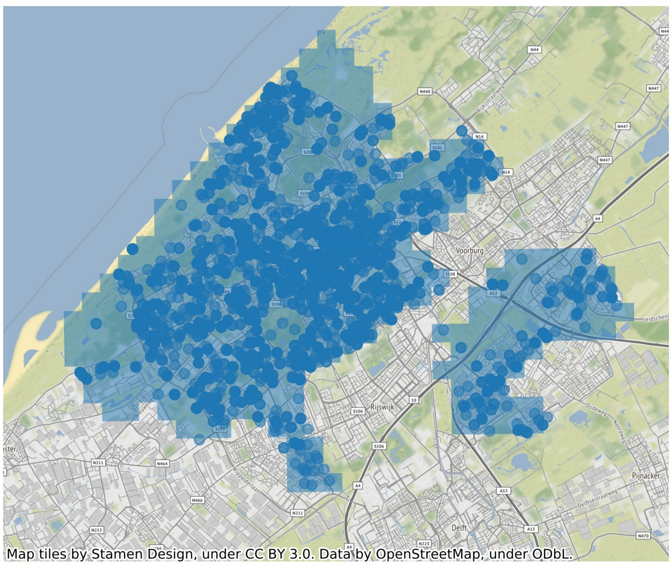

Figure 4.13: Distribution of ambulance calls over The Hague

To overcome this difficulty, let us assign the calls to 500 by 500 m^2 polygons and classify them by the number of calls made. It was found that around one-fourth of the calls were made from the central medical institutions of The Hague. The following figure shows how this may shift the overview. The second figure, the choropleth map, demonstrates the calls combined into eight groups and assigned to the areas of 500 by 500 m^2. The red circles on the third map represent the main hospitals in the city. Contrasting these two, we can see that three main were generated nearby the hospital. The next step is to remove these calls and see how the picture will change.

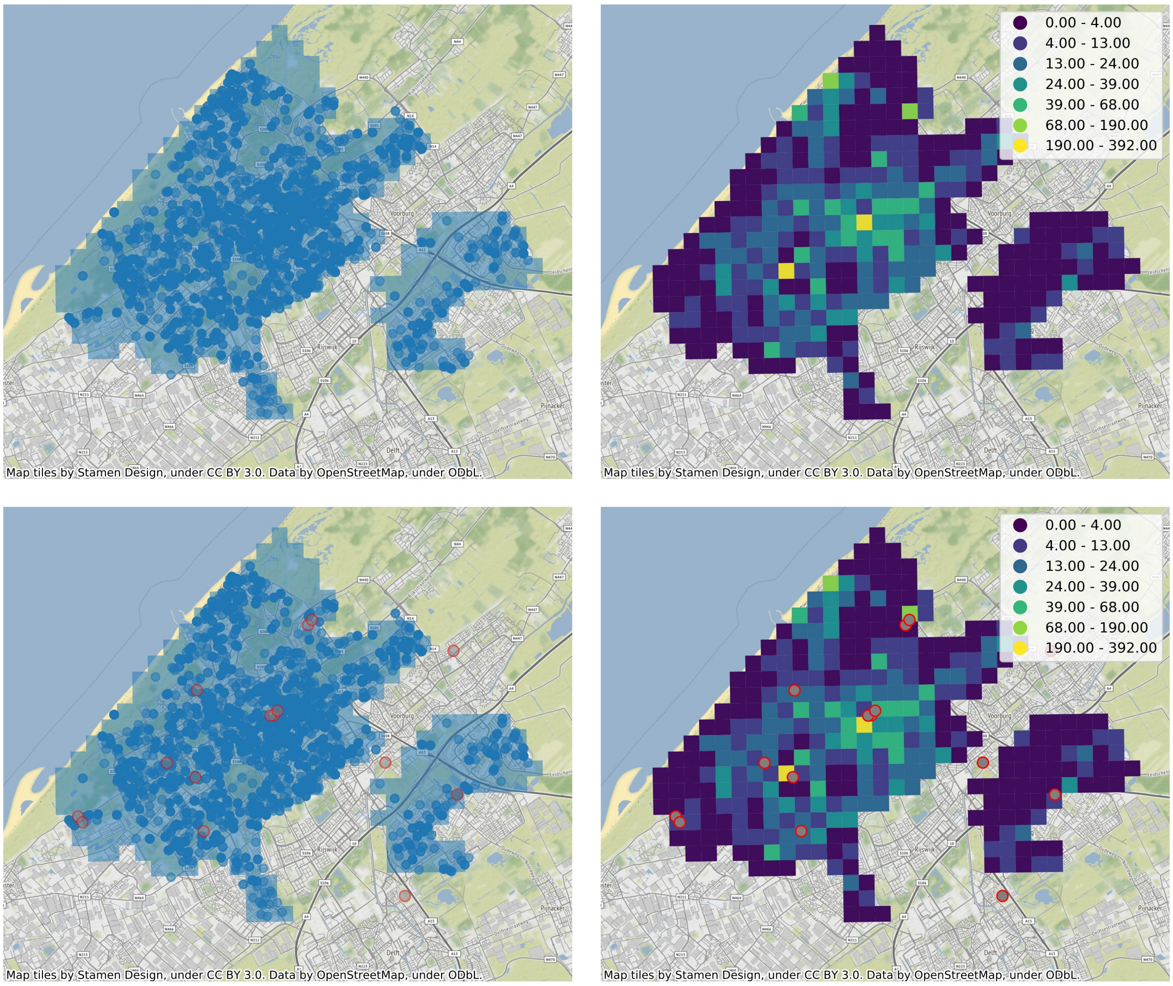

Figure 4.14: Calls generated by hospitals highlighted

After the calls made from the hospital locations were excluded, the new points for attention appeared. The first ones are ones that are close to the beach area. Given the fact that the records were collected during the summer, the appearance of such a pattern is understandable. Besides that, we start to see how unequally the calls are distributed over the city. The first thing that catches your eye is that the area nearby to the city center (occasionally, it is in the center of the map). It has a higher number of ambulance calls comparing to the periphery (see Figure 4.15).

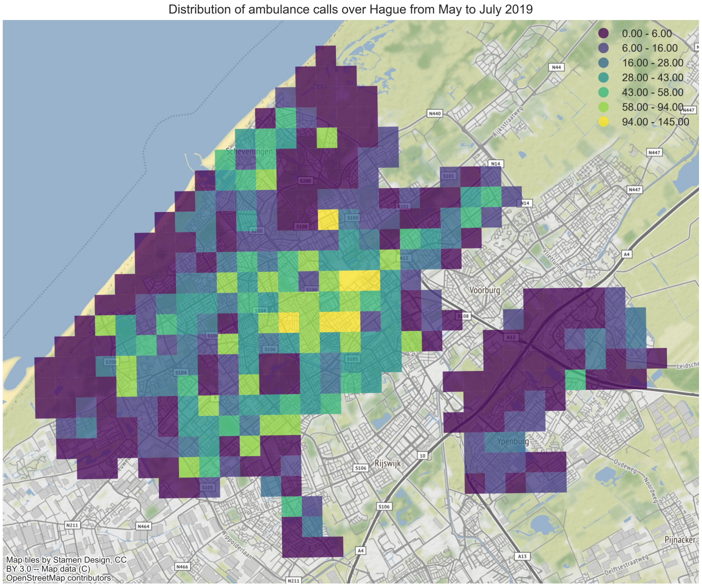

Figure 4.15: Ambulance calls cleaned

The next step is to contrast the average number of calls made from May to July 2019 with the calls made during the heatwave. The resulting Figure 4.16 is highly insightful. The average number of calls reflects on ”general” vulnerability of the city. All types of accidents are included. It might be a car crash, scooter accident, or any other extreme case. Of course, the same holds and for the heat wave data set. This is one of the reasons why we see an overlap represented by cluster 2 colored with light green. However, since the number of data points is relatively low, it is more likely that the left figure points us to a different conclusion. In the highlighted areas, people were asking for help significantly higher number of times. **The hypothesis is that the population of these areas is highly vulnerable to extreme heat**. The discussion of this idea will be continued in Section 4.2 on vulnerability assessment.

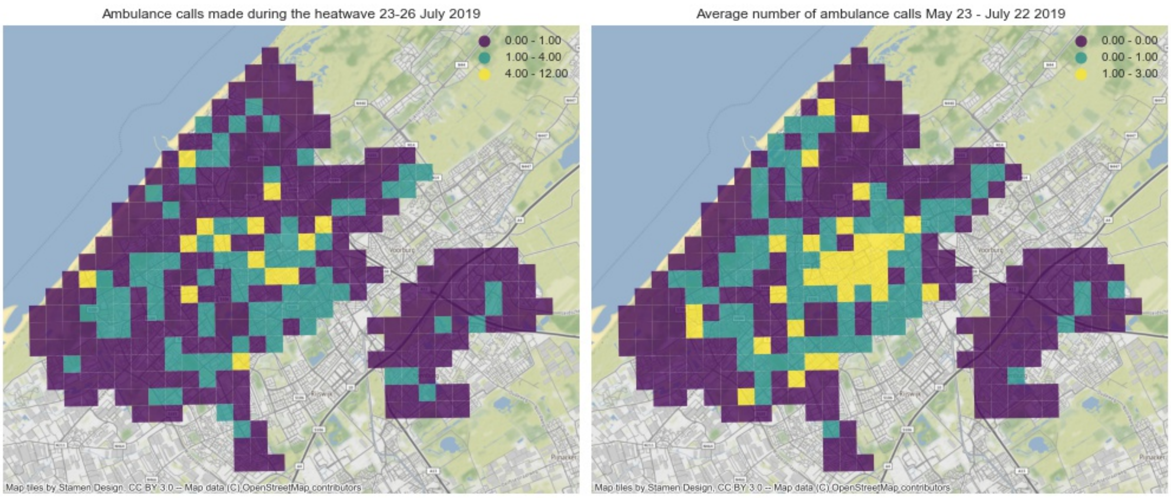

Figure 4.16: Heat wave calls contrasted with the average values

Ambulance calls and urban heat island

The study conducted by Hoeven (2018) allows us to explore relationships between The Hague's urban heat islands (UHI) and ambulance calls. The hypothesis is that *there will be more calls made from UHI since the temperature there is higher*.

Figures 4.17, 4.18 demonstrate one of the main results of the study: daytime and night-time temperatures on May 27 and May 28 respectively. The surface of the districts with the highest temperatures is painted over. It was concluded that there are some areas where the heat island effect is significant. Namely, the following districts: Scheveningen, Centrum, Schildersbuurt, Transvaalkwartier, and Binckhorst.

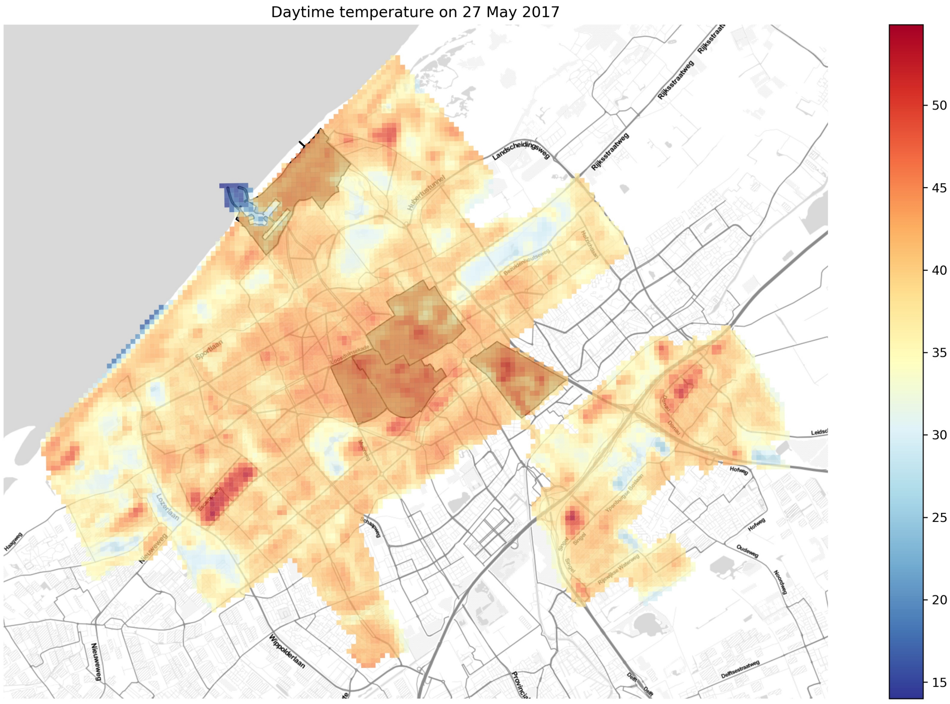

Figure 4.17: Daytime temperature in The Hague on 27 May 2017

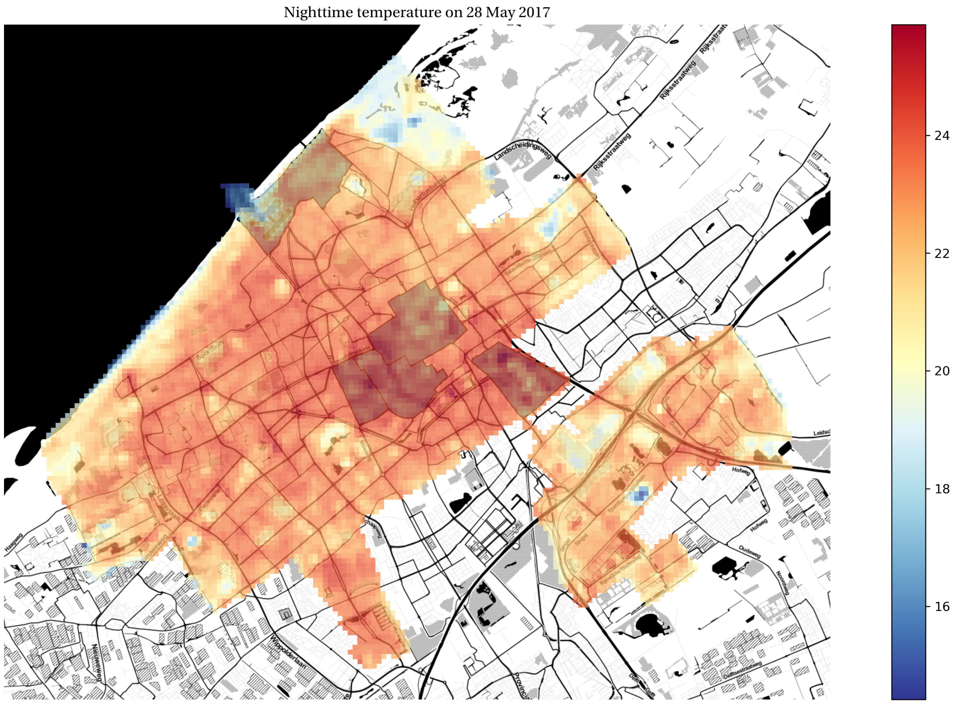

Figure 4.18: Nighttime temperature in The Hague on 28 May 2017

Now let us try to match the ambulance calls made during the heat wave with the UHI. Figure 4.19 consists of two maps: ambulance calls on the left and the maximum temperatures on the right. The orange polygons on the left one are the districts with the highest UHI effect. Indeed, **these districts have peaks of ambulance calls**.

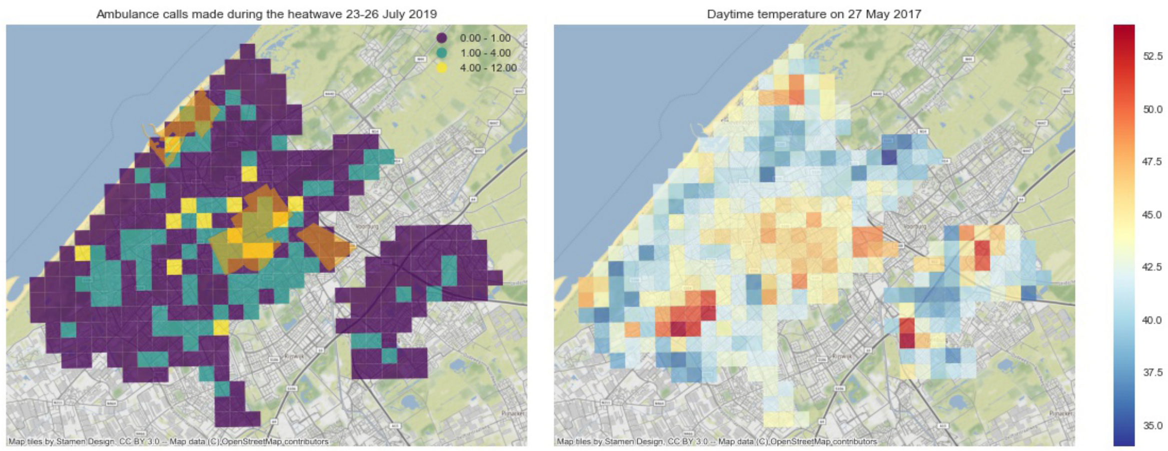

Figure 4.19: Ambulance calls overlapped with heat island effect

Ambulance calls and temperature

So far, we have become quite familiar with two facts. First, in July 2019 there was a heat wave in the Netherlands, and The Hague, consequently. Second, the number of ambulance calls has almost doubled during that time. Let us elaborate on relationships the temperature and the number of the ambulance a bit more.

Figure 4.20 demonstrate an immediate increase in the number of ambulance calls with a strike of the heat wave:

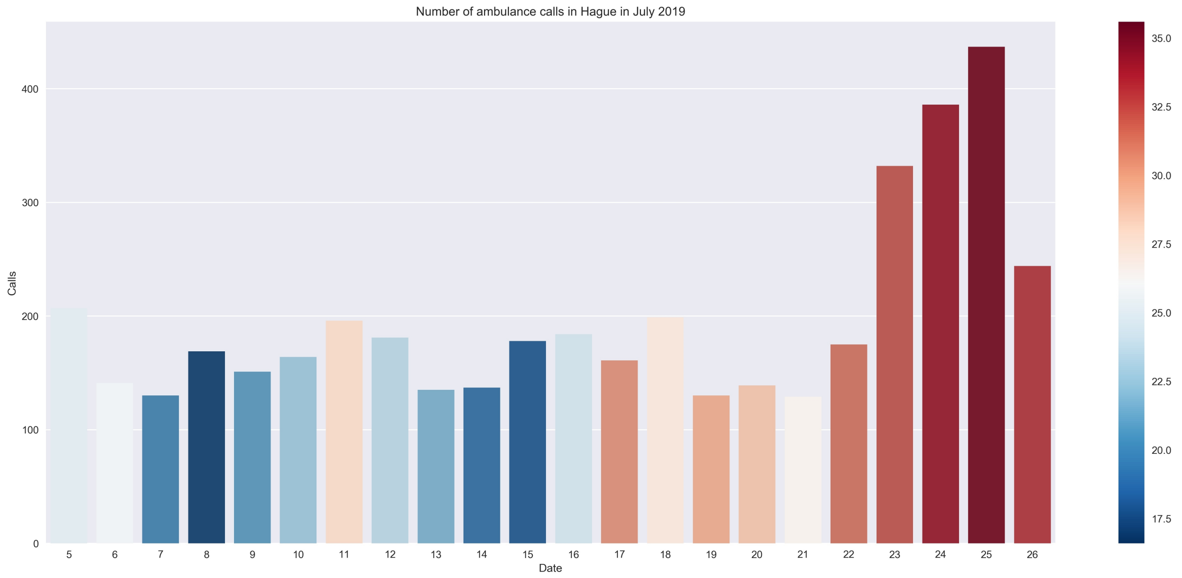

Figure 4.20: Dependency between ambulance calls and temperature for The Hague

Importantly, the same type of relationships hold for other two largest Dutch cities: Rotterdam and Amsterdam (see Figure 4.21 and 4.22. For Rotterdam, the difference between the average values and the heat wave numbers is almost three times!

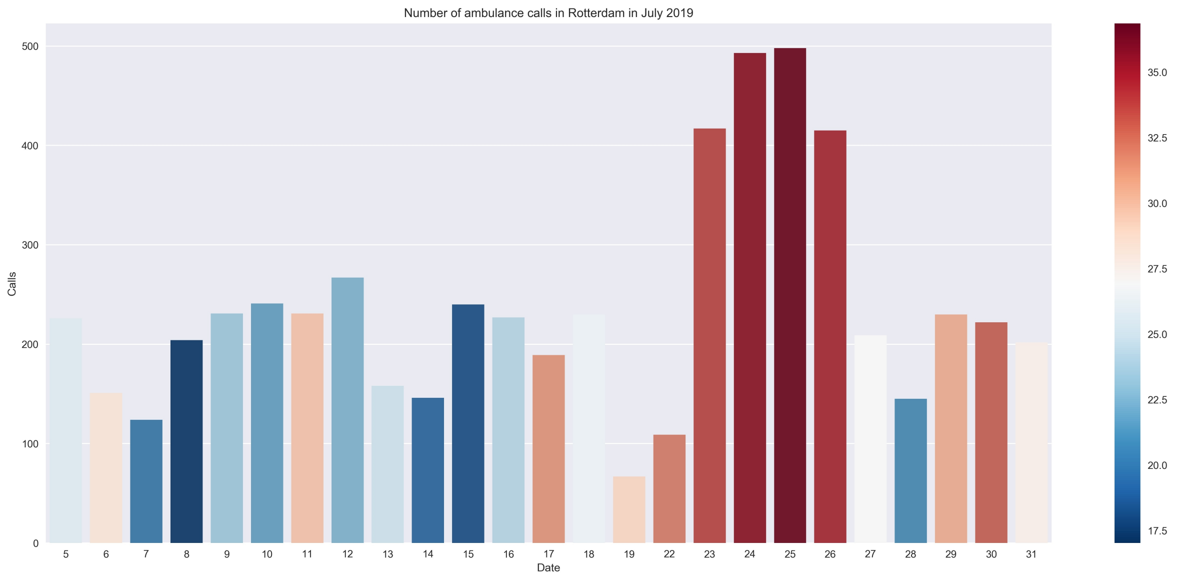

Figure 4.21: Dependency between ambulance calls and temperature for Rotterdam

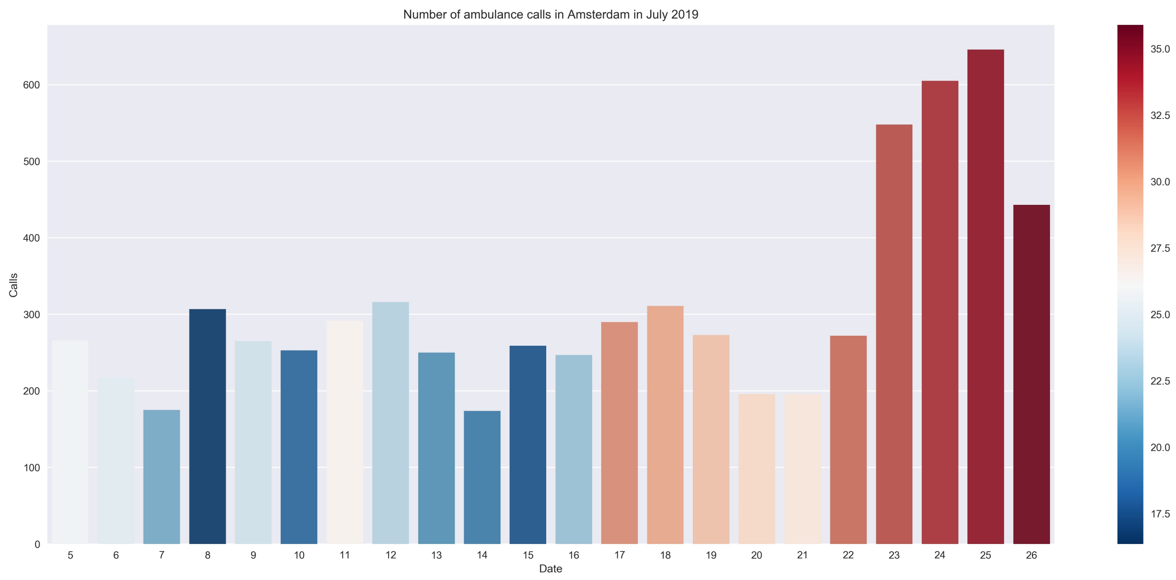

Figure 4.22: Dependency between ambulance calls and temperature for Amsterdam

These graphs communicate two critical ideas. First, they are clearly showing that **extreme case demand is bigger in two or even three times than the average**. Such information can be used for planning future hospital capacities. Second, based on the data of three cities, **there are strong relationships between the temperature and vulnerability represented by the number of ambulance calls**. The population of Dutch cites is actively seeking for help against the extreme heat threat.

Concluding remarks

The ambulance calls data along with temperature and UHI provided us with a great insight into vulnerability and importance of the extreme heat for the population of The Hague. First, we saw that the distribution of the calls is unequal across time and space. Generally, there are two peaks in demand: around 12:00 and 22:00. During the heat wave, this pattern slightly moved, and new peaks appeared. Time series forecasting with fbprophet is a useful tool for planning even with a relatively small data set. However, it should be used quite carefully while dealing with such problems as heat wave. The number of calls made during the heat wave indicates a need to reconsider hospital capacities. Finally, in The Hague case, we observed that the more calls were made from the places with high UHI effect.

4.2. Vulnerability assessment

The methods that are proposed to analyze vulnerability: dimensionality reduction, clustering, and regression analysis. Vulnerability assessment with dimensionality reduction should identify where vulnerable to extreme heat individuals are concentrated. Clustering can give an insight into how more vulnerable and less vulnerable individuals are coexisting. Methods of regression analysis utilized to analyze relationships between citizens attributes and the number of ambulance calls.

4.2.1. Dimensionality reduction

With dimensionality reduction algorithms, we seek to understand how vulnerable individuals are distributed across the city. Opposite to clustering, when a ”hard” decision is made by assigning a single cluster, dimensionality reduction shows relative numbers, ”proportions” of vulnerable individuals living in a particular area.

The input data is defined by the collection of the data sources, namely from CBS, RIVM, and the municipality of the Hague. These data sets have information about individuals and households. The main input variables are age group, ethnicity, household composition, whether or not an individual receiving benefits (a proxy chosen to represent income) and two health factors (mobility issues and having a serious disease). The resolution of the data is 500 by 500 m^2. The total number of observations is 339. Table 4.1 shows original variable names and their meaning:

№	Variable name	Explanation	Type
1	INW_014	Residents from 0 to 14 years old	Real number
2	INW_1524	Residents from 15 to 25 years old	Real number
3	INW_2544	Residents from 25 to 44 years old	Real number
4	INW_4564	Residents from 45 to 64 years old	Real number
5	INW_65PL	Residents over 65 years old	Real number
6	P_AUTOCHT	Native Dutch	Percentage
7	P_WALLOCH	Western foreigners	Percentage
8	P_NWALLOCH	Non-western foreigners	Percentage
9	UITKMINAOW	Benefits	Real number
10	TOTHH_EENP	One-resident household	Real number
11	TOTHH_MPZK	Multi-person household without children	Real number
12	HH_ENOUD	One parent with children	Real number
13	HH_TWEEOUD	Two parents with children	Real number
14	ErnstigBep	Seriously limited because of health	Percentage
15	Mobiliteit	Mobility impairment	Percentage

Table 4.1: The list of the variables used

Let us start the study from analyzing results of Principal Component Analysis (PCA) algorithm. The decision on the right number of components can be based on the amount of variance (see Figure 4.23). As in the case of clustering, it is relatively hard to decide on the desired number of dimensions. Let us stop with four components since the amount of explained variance added by the fifth and other components is relatively low.

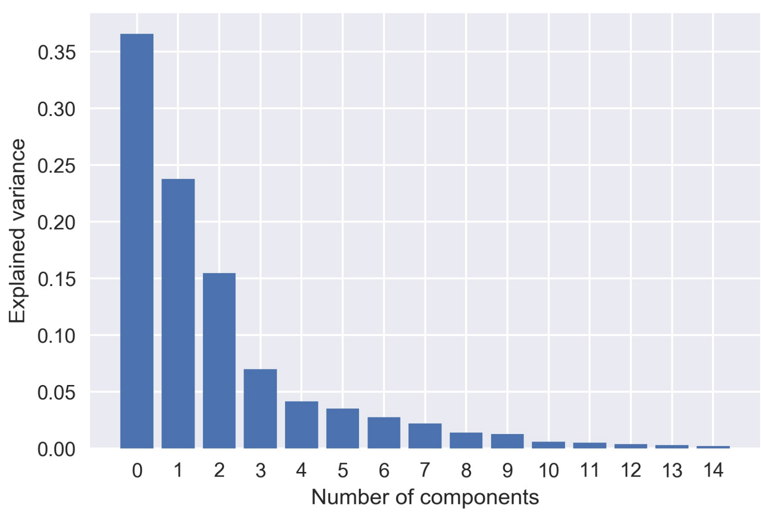

Figure 4.23: Variance of PCA components

To understand the components of PCA, we can use a ”heat map” of its values. As can be seen from Figure 4.24, it is relatively hard to infer the results immediately. However, three remarkable patterns are present. Let us name a component of dimensionality reduction algorithm as a **profile**. Each profile is a unique combination of attributes. Some of those attributes are predominant, and some share similar values. Profile represented by component 1 has the highest number of non-western foreigners within a specific age group who receive benefits. Profile 2 consists mostly of Dutch inhabitants with kids. Profile 3 characterizes older people with health issues. The main profiles of interest for us is 1 and 3.

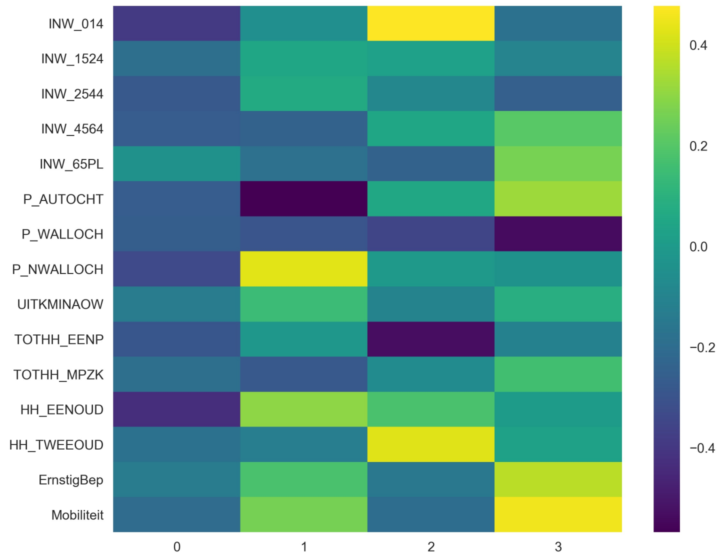

Figure 4.24: Components of PCA

Let us explore the distribution of these profiles on the map. In contrast to the clustering, profiles created by a dimensionality reduction algorithm reflect on the complex nature of urban demography. The presence of each profile can be found elsewhere in the city. For instance, we can see that profile number 1 dominates in the city center, the highest concentration of profile number 3 in on the southwest of the city.

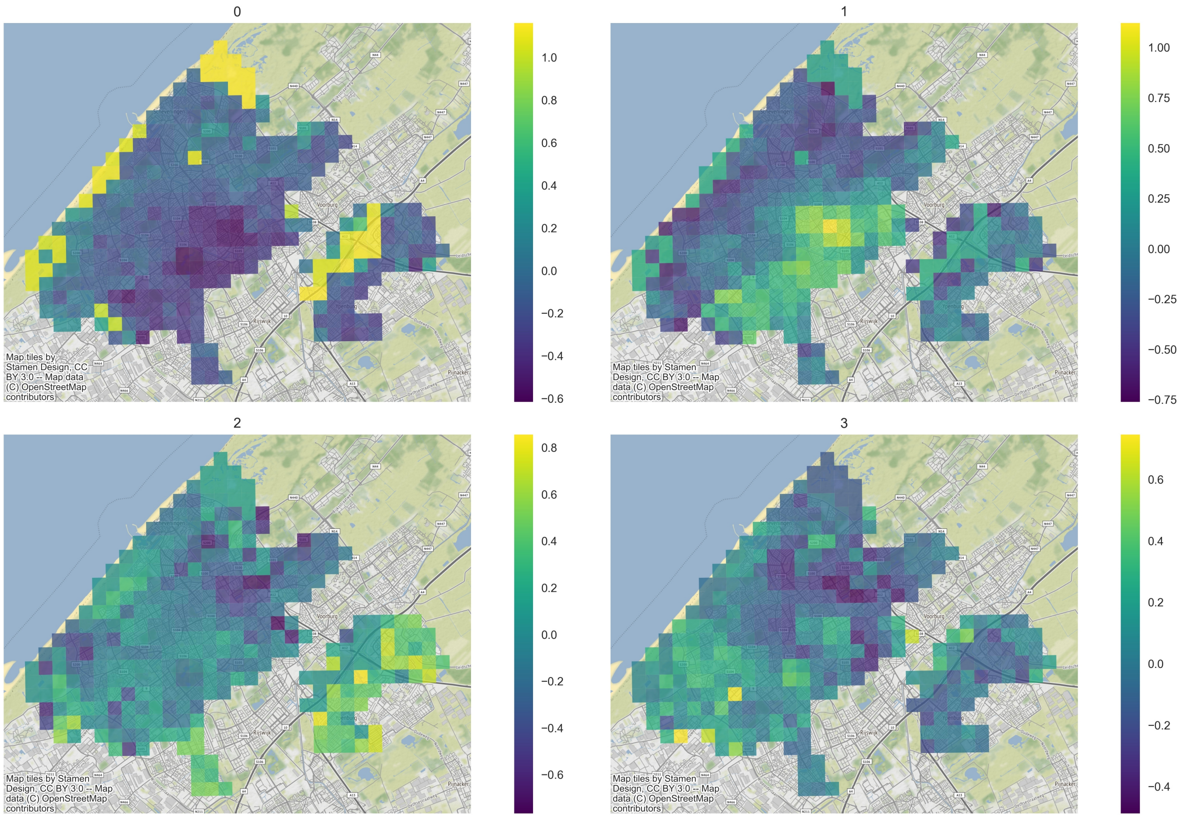

Figure 4.25: Distribution of PCA components on map

On the one hand, having the knowledge of The Hague demographics it is possible to validate such results. On the other hand, to ensure the robustness of the conclusions, let us compare the performance of PCA with other dimensionality reduction algorithms. It is also necessary due to the complexity of the task. The original data set has some non-linearities that PCA might not be able to handle. As was discussed in Chapter 3, the other algorithms are: Non-negative Matrix Factorization (NMF), Factor Analysis (FA), Singular Value Decomposition (SVD) and Independent Component Analysis (ICA). The metric that is used to compare the performance of the algorithms: explained variance. The metric is calculated on the train data (80%) and test data sets (20%).

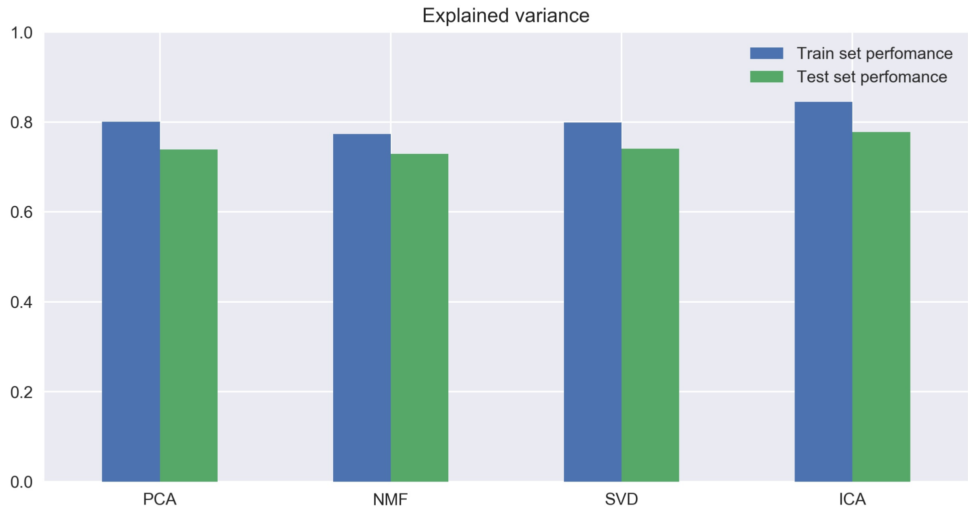

Figure 4.26: Comparison of different dimensionality reduction algorithms

As can be seen from Figure 4.26, all algorithms demonstrate relatively the same performance. ICA is doing slightly better than the rest, whereas NMF has the lowest score. The next step is to compare the profiles identified by the algorithms. The choice of the technique should not be based only on the metric, but also on the fact that the results are easily interpretable.

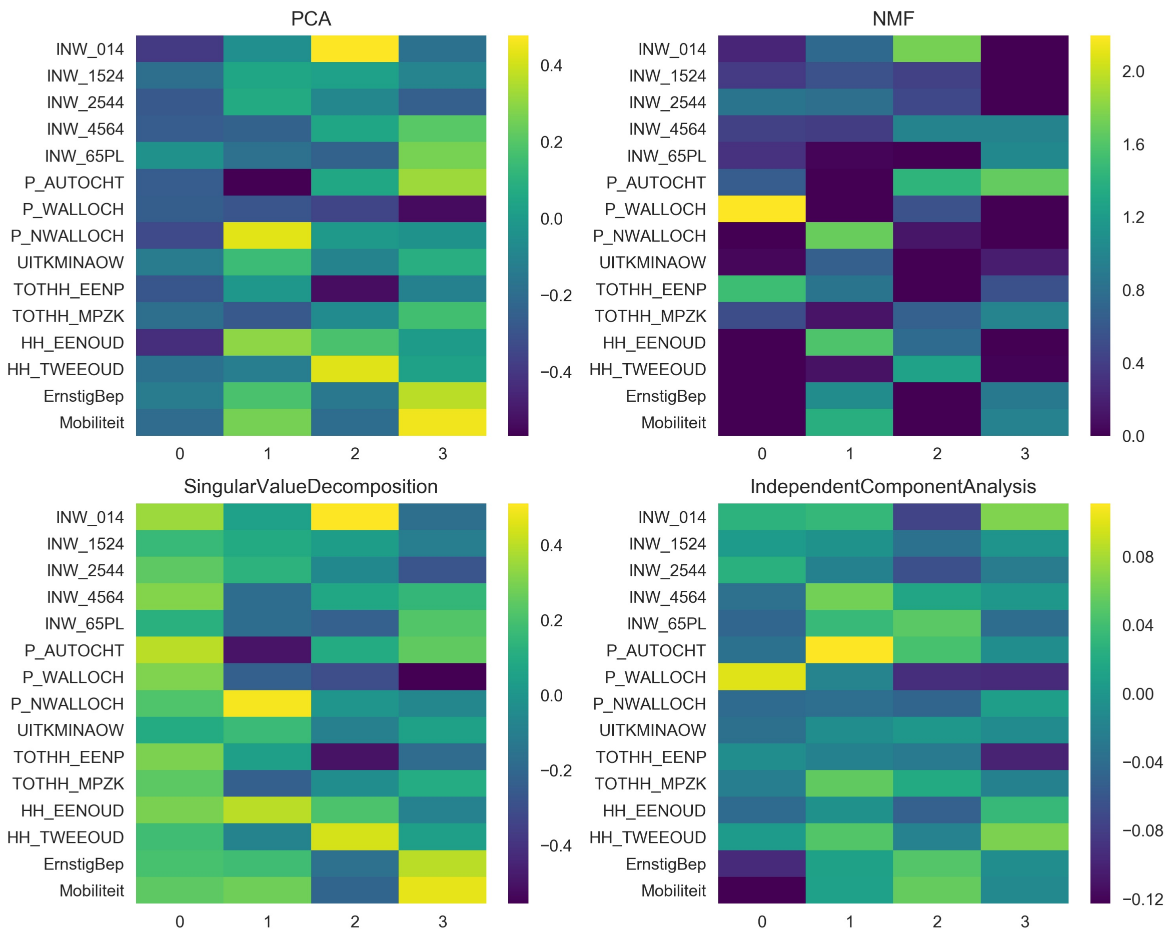

Figure 4.27: Comparison of components of different dimensionality reduction algorithm

However, such an approach is hard to apply here. Due to the nature of the data set, the algorithms identified slightly different patterns. In simple terms, it means that some of them failed to decompose the initial variables in independent components successfully. Note that extra metrics and other useful visual instruments can be found in the online repository. Importantly, one algorithm - NMF, it is the easiest to interpret and operate with, even though its explained variance was the lowest (4.28). It highlighted the most remarkable attributes of the data set. Besides its components are non-negative, which makes it useful when working with population data. The components of interest here are number 1 and 3. The first one stands for low-income diverse age non-western foreigners with health issues, and the second one represents mostly Dutch elderly with health issues.

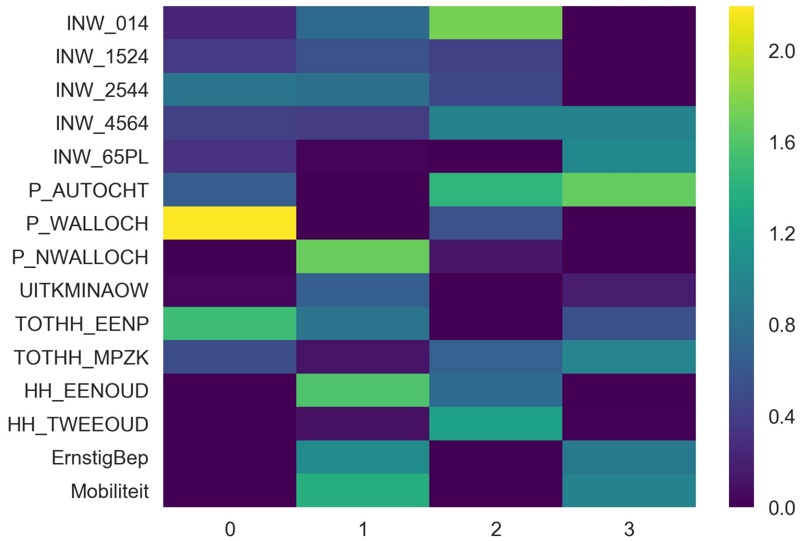

Figure 4.28: Components of NMF

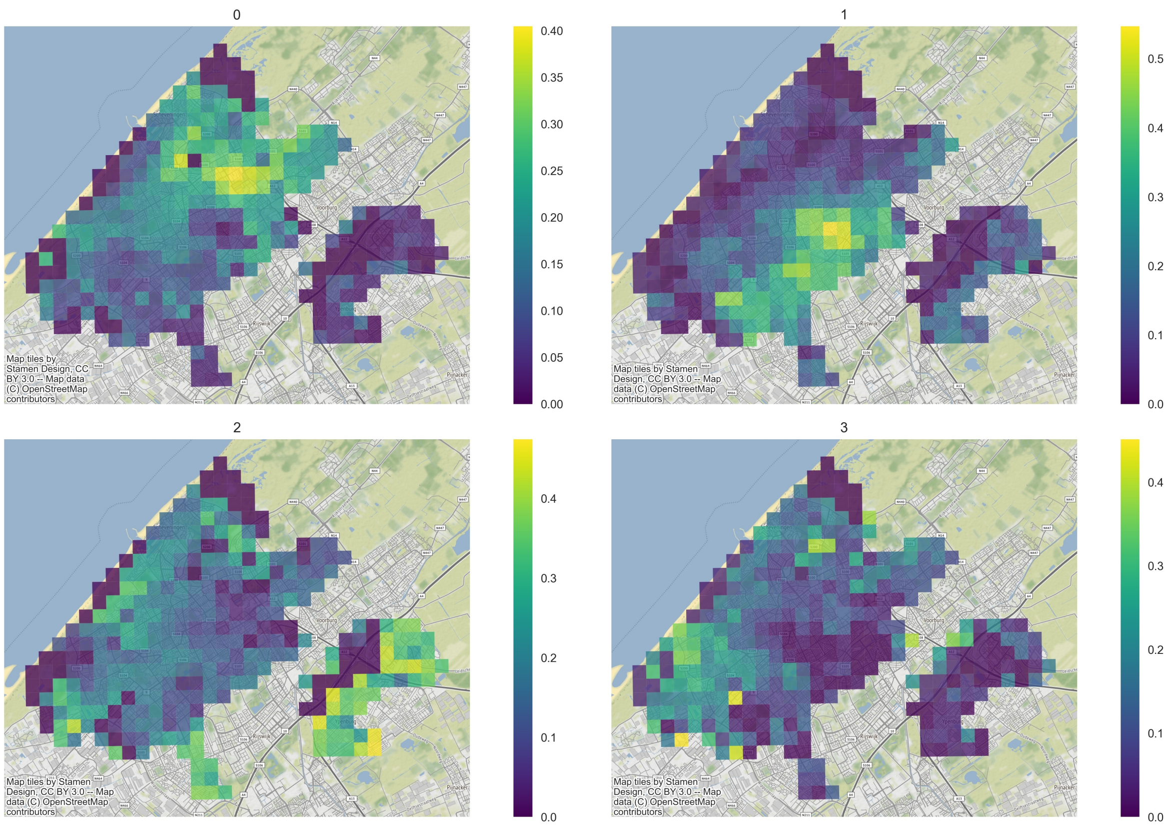

Figure 4.29: Distribution of NMF components on map

Important to mention that the number of attributes of interest can be enlarged further. NMF proved to be robust to outliers and provided easy to interpret results. Besides, the number of components can also be increased to get a more precise picture of such a diverse city like The Hague.

Concluding remarks

Dimensionality reduction algorithms and primarily NMF provided us with the distribution of vulnerable individuals over the city. This insight is highly valuable since it can help to plan and allocate solutions for resilience capacities.

4.2.2. Clustering

By applying clustering, we seek to understand the complex demography of the city. More specifically, we aimed to understand how previously defined profiles are grouped, or in simple terms, whether vulnerable citizens living close by to less vulnerable.

The task of clustering differs from the task of dimensionality reduction. The second one provides us with a distribution of profiles over the city, whereas in the case of the first on a hard decision on a cluster should be made. Initially, clustering is highly challenging because of the nature of the data sets. Multiple features are correlated, and typically, the algorithm will fail to come up with interpretable results. However, the profiles generated by NMF are not correlated any more. Thus, the task is easier and we can expect more straightforward results.

As was discussed previously, the analysis will be started from the K-Means algorithm. K-Means requires the researcher to define the number of clusters. The *elbow method* can provide support in choosing the right number of clusters. However, when dealing with complex data, the elbow method may not work. As can be seen from Figure 4.30 there is no specific place where inertia is suddenly stabilizing. Instead, it is relatively smooth while the number of clusters is increasing. The best guess is to select the number of clusters equal to 5, since after that there is something that slightly resembles an elbow. Thus, we still need to take a look at the composition of the clusters and use expert knowledge about the case.

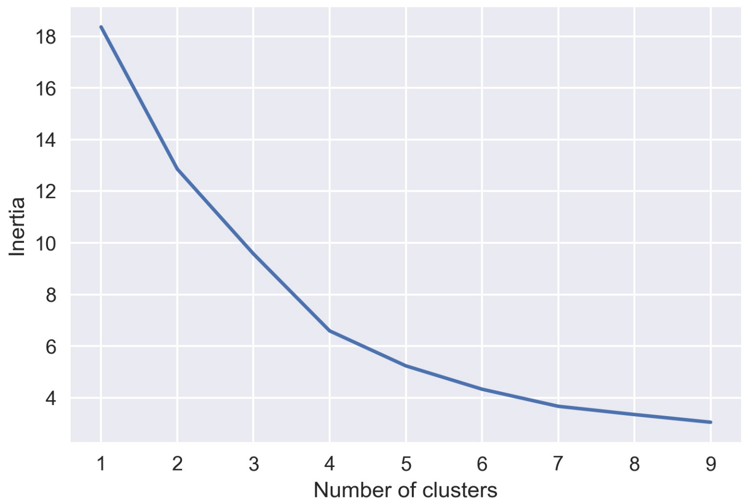

Figure 4.30: Elbow method for optimal number of clusters

There are different ways to evaluate the structure of the complex multi-dimensional data. The one that is highly useful for clustering is *parallel coordinates plot* (Wegman, 1990). In short, a multi-dimensional observation is represented as a polyline with the vertices on the parallel axes. These vertices are values of the attributes. Let us explore the results of the K-Means algorithm using parallel coordinates plot. As can be seen from the graph below, clusters still often share the same values of the same attributes. It is not surprising since we are working with the demographic data, what complicates the analysis. What is remarkable is that the cluster number 4 significantly relies on a certain profile: NMF_1 (which stands for the vulnerable population living in the city center). Cluster number 2 on NMF_3 that consist of the elderly with health problems.

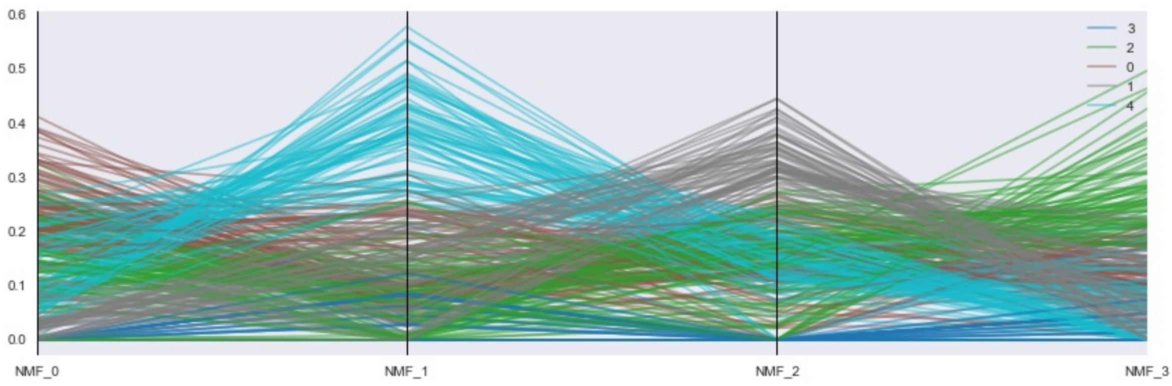

Figure 4.31: Representation of clusters with parallel coordinates plot

Now let us evaluate clustering results from a geographical perspective. From Figure 4.32, we see that certain individual clusters dominate in specific areas. For instance, in the south of the Hague, the cluster number 4 prevails. The north of the city is separated between clusters 2, 1 and 0. Cluster 3 mostly covers the areas with low population and parks.

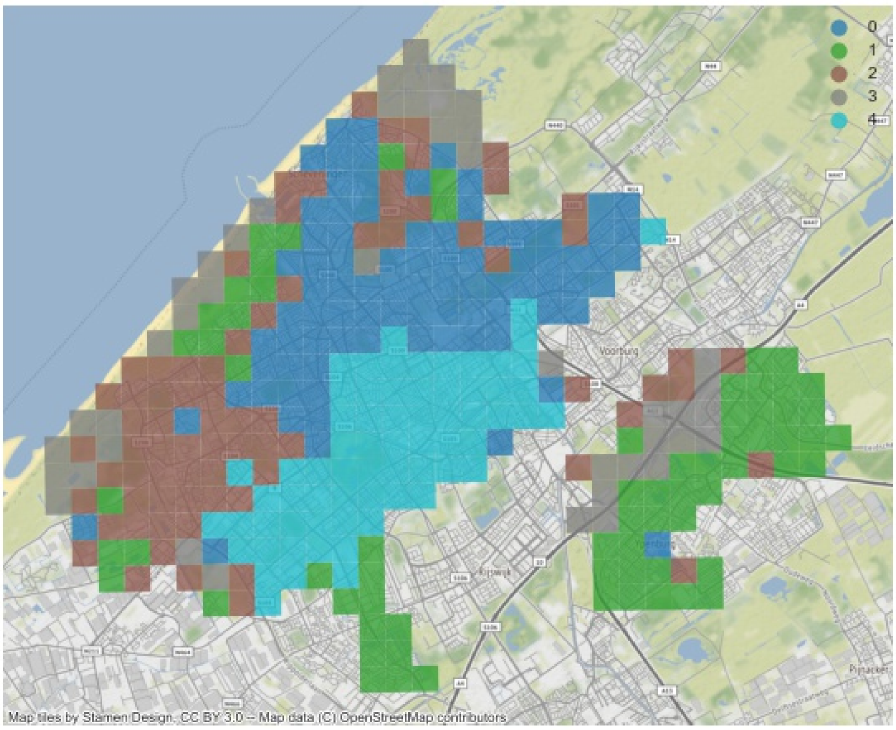

Figure 4.32: K-means clusters on map

Again, to ensure the robustness of the conclusions, we need to analyze the performance metrics and consider the outcomes of other clustering algorithms besides the K-Means. Two standard metrics were chosen: Silhouette Score and Davies-Bouldin Index. These metrics are used in the real labels for the clusters are unknown. The extensive explanations of the parameters can be found in Pedregosa et al. (2011). In short, we are looking for an algorithm with the highest Silhouette Score and the lowest Davies-Bouldin Index. There is often a trade-off between these two. The algorithms that have similar scores are K-Means, Minibatch K-Means, Spectral Clustering, and Ward. Since the differences are relatively low, we will continue to operate with the K-Means. The outcomes of the other algorithms can be found in the online repository.

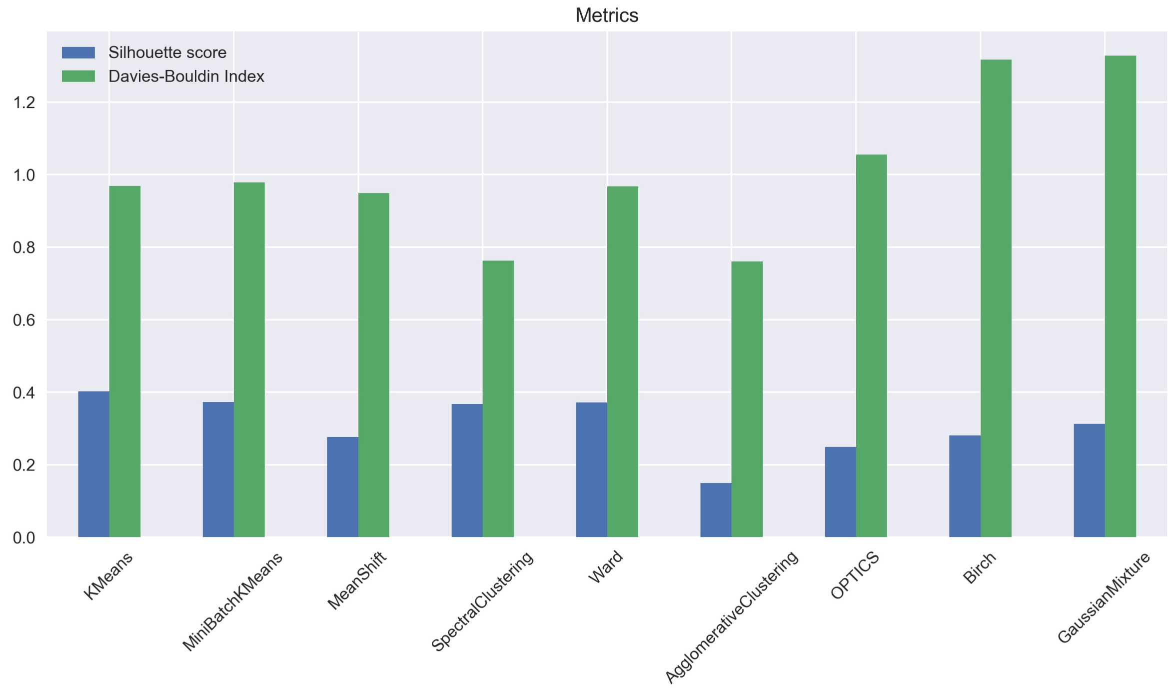

Figure 4.33: Comparison of performance metrics for clustering algorithms

Concluding remarks

The performed clustering algorithms pointed us to the fact that vulnerable individuals are often grouped in vulnerable communities, as in the case of cluster 4 and cluster 2. Such a finding highlights the importance of social infrastructure for building better social ties and establish healthy relationships among people with different backgrounds.

4.2.3. Regression analysis

Regression analysis is utilized to identify relationships between the number of ambulance calls and different characteristics of residents of each of The Hague's 500 by 500 m^2 polygon. As was discussed in Chapter 2, there are a variety of attributes that define extreme heat vulnerability: age, household composition, health, and so on. Let us denote these as predictors X_i. The fact that a person calls an ambulance somehow reflects on his or her vulnerability. The hypothesis is: if a person is vulnerable, it is likely that he or she will call an ambulance during a heat wave. Let us denote the number of ambulance calls as a predictive variable y_i.

Let us first ensure the presence of relationships between the total number of ambulance calls made from May to August 2019 and population. It is fair to assume that certain places in the city have a higher number of ambulance calls because of their "function," so to say. For example, city center may have a higher number of ambulance calls because of the daily visitors. People come there, experience specific problems, and call an ambulance. However, the general pattern should be still in place. The local population should make a certain percentage of the calls. That is, the more people leave within the area, the more calls will be made. Indeed as Figure 4.34 shows, there are some relationships between the number of calls and population. The middle line represents the regression curve, whereas the blurred blue line shows regression confidence intervals drawn with translucent band lines.

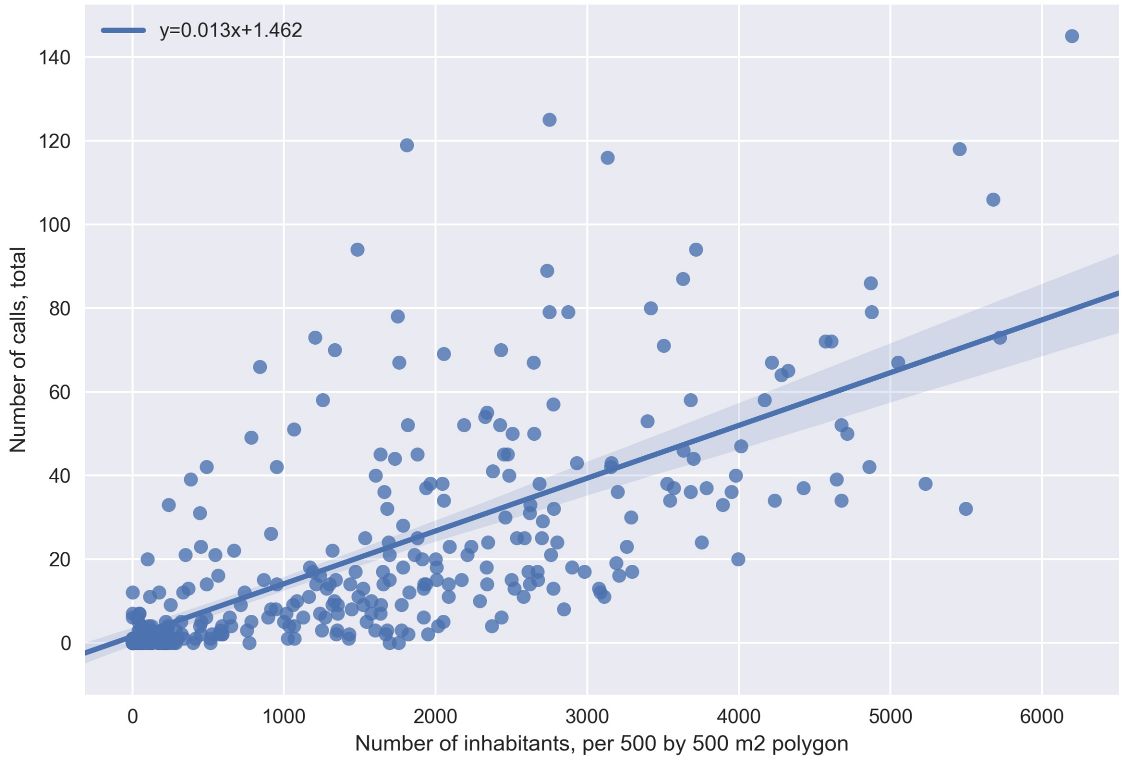

Figure 4.34: Regression plot on population and number of ambulance calls

The R^2 metric is equal to **0.482** which stands for low predictive accuracy. Informally, the relationships are present, but the chosen predictor cannot fully explain the phenomenon. The same impression can be derived from the graph. Importantly, the points above the regression line are not part of a single district, city center, for instance. Instead, they are spread all over the map (visual representation of the calls can be found in 4.15.

Let us continue the analysis with the simplest vulnerability model that can be found in the literature. This model is based only on one predictor X - the age group. If a person is older than 65 (sometimes 80) years, he or she is highly vulnerable to extreme heat. This time linear regression shows even lower accuracy. R^2 is equal to **0.266**.

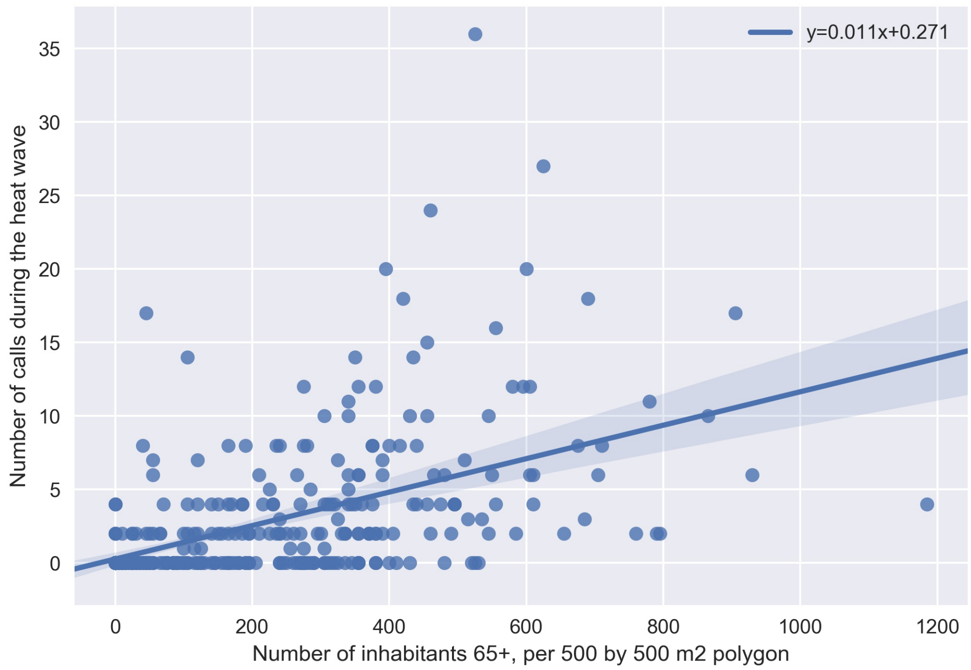

Figure 4.35: Regression plot on population over 65 years old and number of ambulance calls made during the heat wave

The final attempt is to connect the profiles with the number of calls. The profiles or the components of W matrix of NMF can serve as predictors. That is, a unique combination of different citizens attributes may serve as a better predictor for the ambulance calls made during the heat wave. More formally, the predictors are the components or columns of W matrix (see Figure 4.28):

$$y_i = W_{0,i} + W_{1,i} + W_{2,i} + W_{3,i},$$

where i the number of polygon = $\overline{1,339}$. Note that the relationships between y and W_j are not necessary linear. To test out this hypothesis we used a set of linear and non-linear regression algorithms. It was found that that the predictive accuracy remains low. Figure 4.36 contrast the real values with predictions. As can be seen almost all of the algorithms failed to identify the pattern.

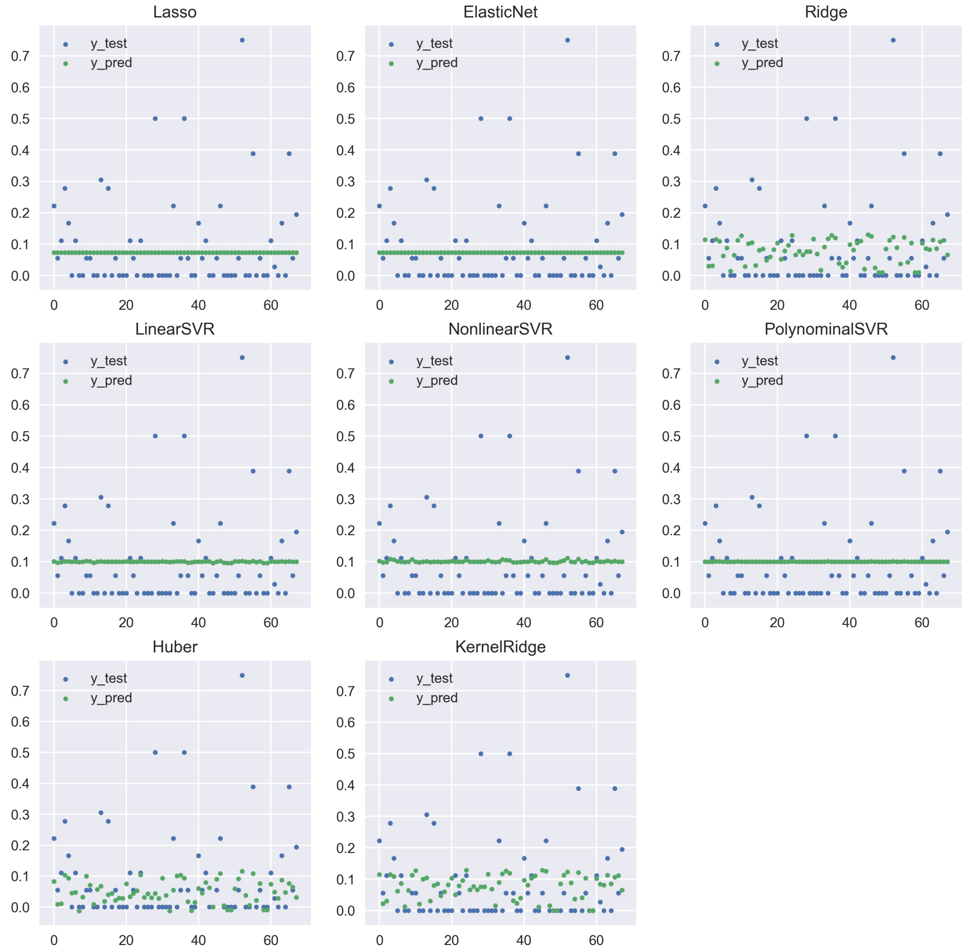

Figure 4.36: Contrasting real values with prediction

Figure 4.37 represent performance metrics of these regression algorithms.

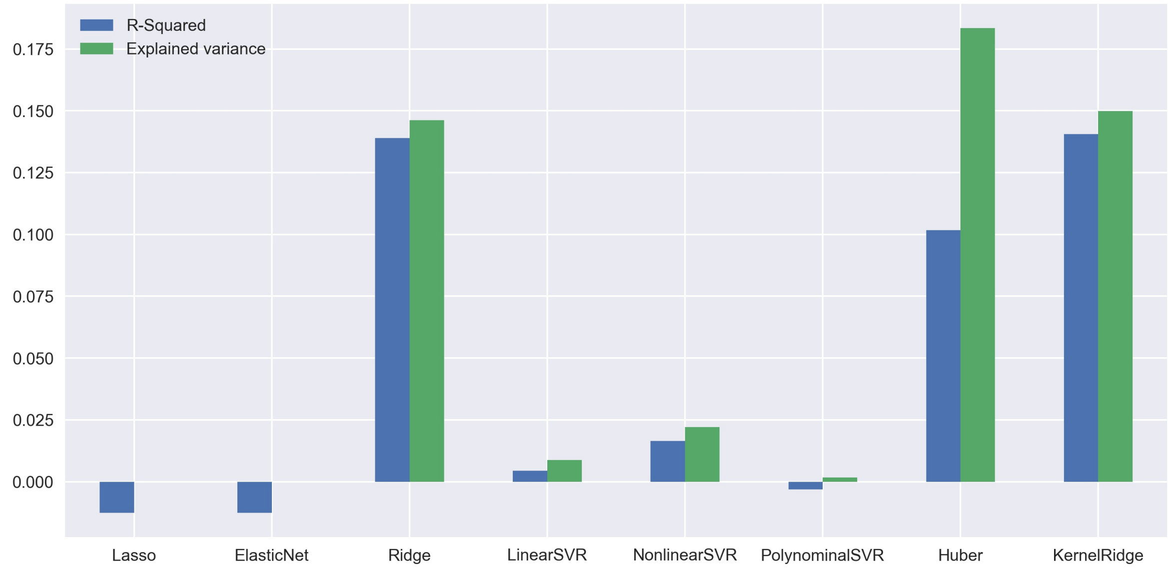

Figure 4.37: Comparison of performance metrics for regression algorithms

Concluding remarks

The goal of regression analysis was to analyze relationships between the demographics and the number of ambulance calls. It was found that such relationships are present. The more the population within a certain area, the more ambulance calls are made. Further, when explaining the number of heat wave ambulance calls with previously defined vulnerability models, no significant relationships were found. There are two ways to explain such an outcome. The first one is to assume that vulnerable people feel good. They do not require help and therefore, do not call an ambulance. However, such a conclusion is contradictory to the official statistics. As we saw from preliminary data provided by CBS, mortality increased significantly. The heat wave influenced vulnerable people. But some of them did not call the ambulance. Thus, there is a need for other types of solutions apart from ambulance services.

4.3. Resilience capacities assessment

This section aims to assess resilience capacities. As was discussed in Chapter 2, they are tightly connected to three types of infrastructures: critical, green, and social. The first two responsible for absorption capacity: vegetation absorbs heat, whereas ambulance services provide immediate help to people in need. Social infrastructure is attached to adaptation and transformation capacities. Having a place close by, for instance, a public library, allows people to gather and build better social ties. Such relationships can play a crucial role in the case of extreme situations. Instead of relying on conventional measures, people can adapt and behave proactively by taking care of their more vulnerable friends and acquaintances.

The methods that are utilized for analysis resilience capacities: geospatial and network analysis.

4.3.1. Absorptive capacity

Critical infrastructure

In this study, critical infrastructure is represented by hospitals. The assessment of it can be done it two ways: by assessment of capacities (number of cars and staff) and by street network analysis (proximity). The raw estimates for the first one were presented in the first section of this Chapter. There, we compared the average demand with the demand during the heat wave. The second one aims to evaluate the reachability vulnerable individuals and communities for ambulance services.

According to OpenStreetMap (OSM) data gathered on July 31 2019, there are 15 hospitals located nearby to The Hague administrative border.

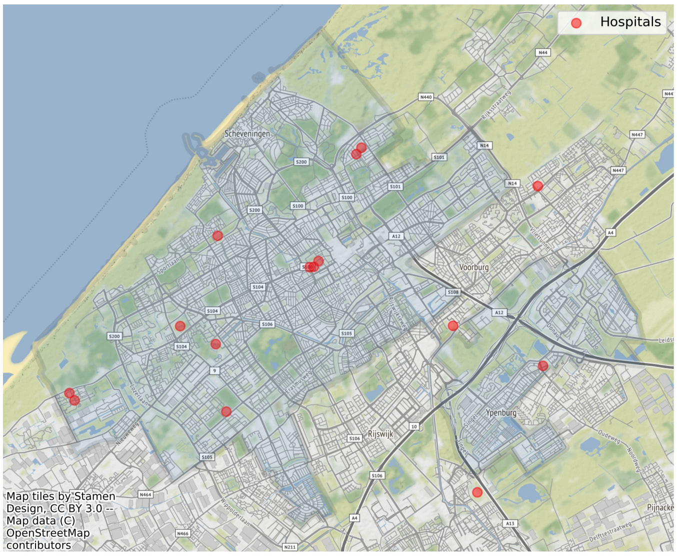

Figure 4.38: Locations of the hospitals nearby The Hague

Now we are going to analyze the driving routes from the hospitals to each of The Hague's polygons. The areas of particular interest are with a higher concentration of profiles number 1 and 3 (vulnerable individuals and communities). Figure 4.39 contrasts the hospitals with the distribution of profiles. As can be seen, each of the profiles is relatively well-placed with respect to the hospitals. There are two hospitals nearby the area of the highest concentration of profile 1. The points of profile number 3 are rounded with the hospitals.

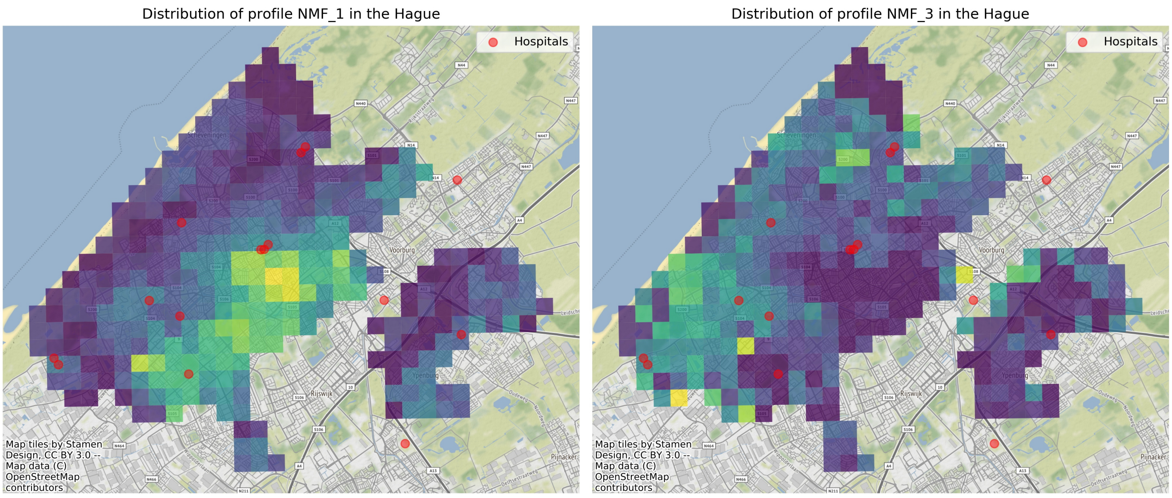

Figure 4.39: Distribution of profiles contrasted with the hospitals

To have a complete picture, let us utilize the methods of network analysis. Using the graph of The Hague's streets, we are aimed to calculate the distances of the shortest routes between hospitals and polygons. Example of such an approach for a single polygon represented in Figure 4.40. The routes are depicted with colored lines. The closer you are, the redder the line, The purple dot is a polygon with a high concentration of profile number 1.

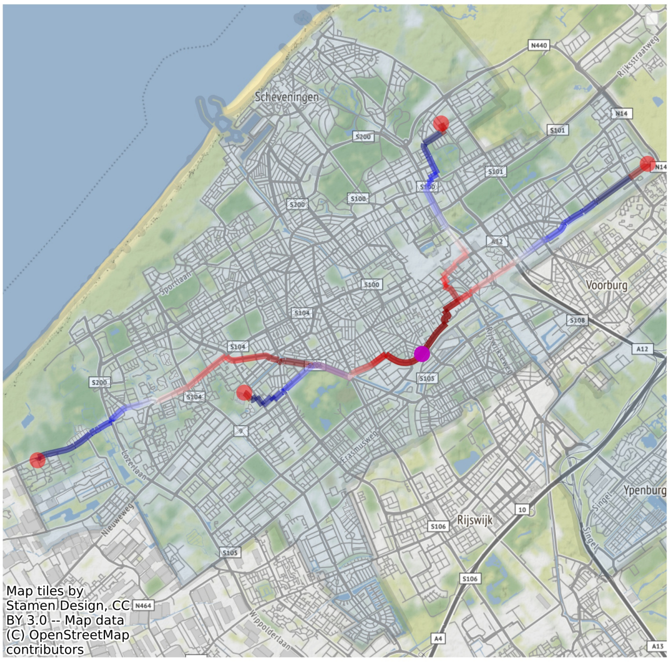

Figure 4.40: Shortest driving routes from 4 hospitals to vulnerable population

After iterating over all polygons we obtain the following proximity map. The polygons are grouped into four clusters based on the length of the shortest path. The color here stands for the distance. As can be seen from the figure, the length of any path does not exceed a relatively small limit in 4.2 kilometers. The places with a high concentration of profile 1 and 3 have even lower distances. Thus, the vulnerable population is relatively well-located.

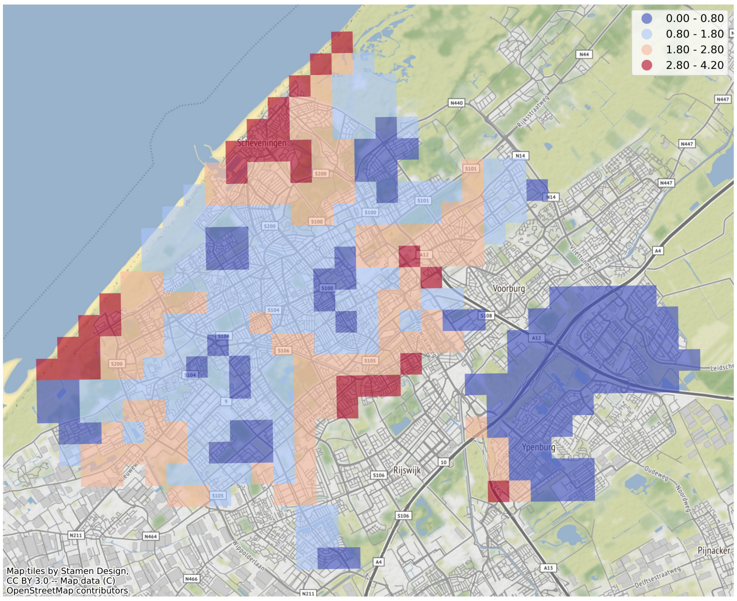

Figure 4.41: The distance of a shortest route to closet hospital in The Hague, km

Green infrastructure

Green infrastructure in the form of trees, grass, and shrubs plays an essential role in combating extreme heat. Vegetation absorbs heat and cool down the temperature around its area. The analysis of it can be performed differently. Apart from the analysis of satellite images with clustering or classification, there are some conventional methods, for instance, by analyzing the number of trees per area. Such a metric will not reflect on drastic changes that are happening during the year, but still can be a relatively good proxy.

The aim is to have more vegetation in the areas with the urban heat island effect. Besides, green infrastructure is vital in the areas with a high concentration of vulnerable individuals and communities. As was discussed previously, green and social infrastructure is essential when ambulance services are not used. Vulnerable people do not realize his or her vulnerability and do not call for help. Their function is performed continuously independent on a person willing to use it.

Figure 4.42 demonstrates that the trees distributed highly unequal across the city. Speaking broadly, the city center has a lack of green. In addition, it has higher concentration of vulnerable citizens of profile number 1. On the contrary, the southwest of the city has a significant amount of vegetation. That the part of the city where citizens with profile number 3 mostly located. Thus, there is need for more vegetation in the areas close to the city center.

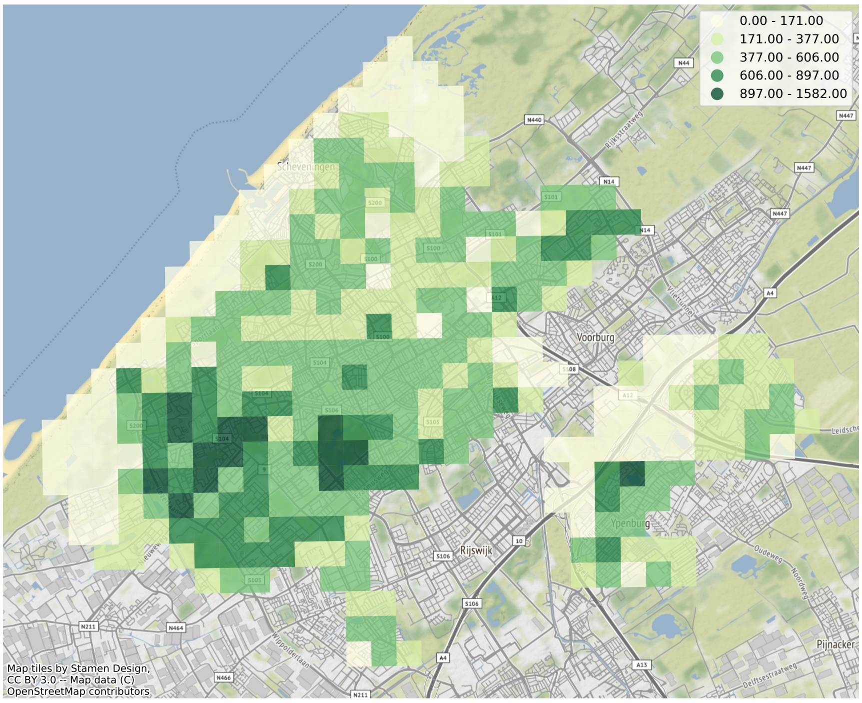

Figure 4.42: The number of trees per 500 by 500 m^2 polygon

4.3.2. Adaptive and transformative capacities

The social infrastructure concept contributes to adaptive and transformative capacities. It promotes establishing better or new social ties that are essential to have in case of extreme heat. The analysis of social infrastructure can be performed in similar to critical infrastructure way. Now, the shortest walking paths to the nearest libraries will be analyzed. Besides the proximities, we will study the popularity of the libraries based on the data provided by Google.

Again, we are using the data collected from OSM on 31 of July 2019. There are 25 libraries located nearby to The Hague administrative border:

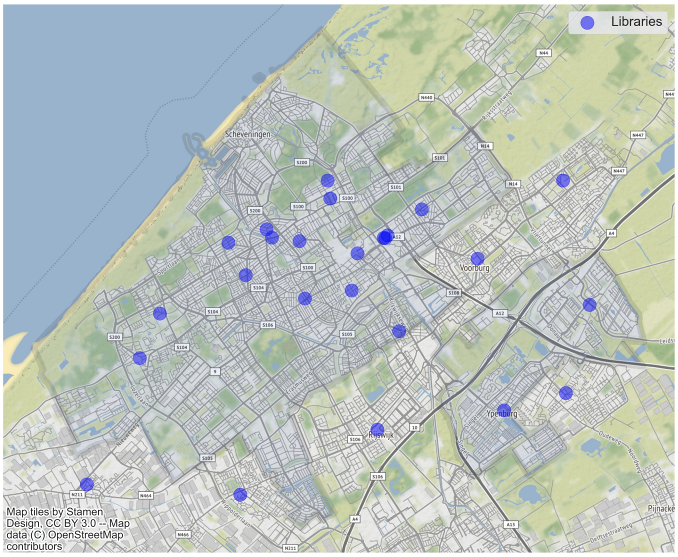

Figure 4.43: Locations of the libraries nearby The Hague

The goal is to analyze walking routes to the libraries from each of The Hague's polygons. The areas of interest remain the same: with a higher concentration of profiles number 1 and 3. Figure 4.44 contrasts distribution of profiles with the libraries. As can be seen, that is a certain number of libraries located closely to the highest concentration of the profile number 1. However, the other areas with the less number of citizens of these profiles do not have easily accessible libraries nearby. There are also a couple of libraries in the southwest of the city that might to reachable to profile number 3.

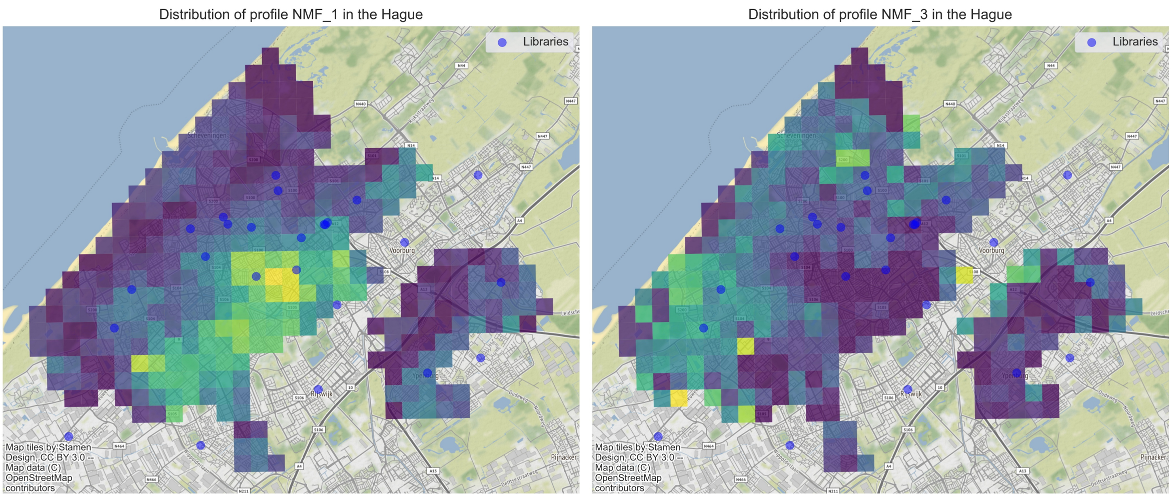

Figure 4.44: Distribution of profiles contrasted with libraries

The network analysis is based on the shortest walking routes from polygons to the libraries. After iterating over all polygons, we obtain a proximity map similar to the one for critical infrastructure. As can be seen from the figure, the distance to the closest library for many parts of the city is more than 1 kilometer (especially for the regions on the north of the city). This fact seems to be quite critical since we are operating with vulnerable individuals that have severe illnesses and mobility issues. For such people, such infrastructure might be hard to reach.

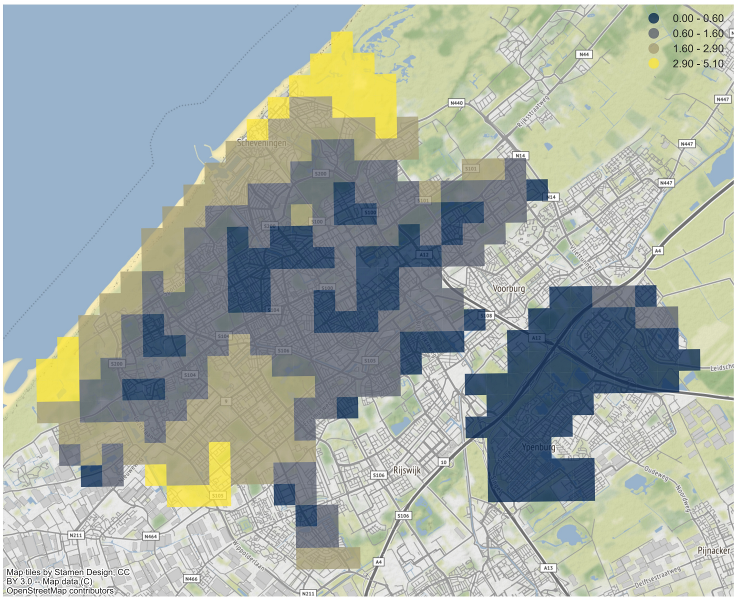

Figure 4.45: The distance of a shortest route to the closest library in The Hague, km

The final part of the analysis is based on the recently introduced Google's Popular times data. The variables that are provided in this data set: the rating of a place and the number of visitors over the day. The assumption here is that these two variables reflect on how popular is the library among the citizens. Figure 4.46 demonstrates the rating multiplied by the number of reviews. The large point in the middle represents The Hague's central library. It is not surprising since the place is highly popular not only among the citizens but also tourists. On the contrary, other city libraries seem to be underrated.

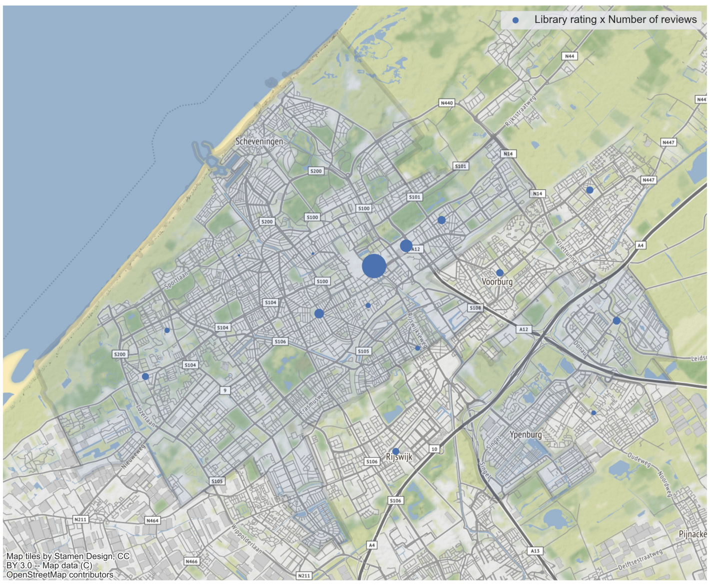

Figure 4.46: Rating of The Hague's libraries

4.4. Conclusion

The goals of this chapter were: analyze the ambulance calls made during the recent heat wave; identify how vulnerable individuals and communities are distributed across the city; assess resilience capacities represented by different types of infrastructure.

It was found that the number of ambulance calls made during the recent heat wave is significantly higher than the average. Besides The Hague, the same holds for Rotterdam and Amsterdam. Such a finding indicates a high sensitivity of the urban population to extreme heat and can be used for planning of hospital capacities for the next heat wave. Interestingly, the ambulance calls are distributed unequally in time and space. On average there are two peaks in demand around 12:00 and 22:00. When the heat wave came, this pattern has changed slightly: a new peak around 23:00 appeared. In addition, the number of ambulance calls made during the night has increased. There are certain places in the city where the number of ambulance calls is higher. Importantly, during the heat wave there was an increase (in comparison to the average) of ambulance calls made from the urban heat islands.

The outputs of dimensionality reduction algorithms strengthen the fact the vulnerability distributed unequally across the city. Two profiles generated by Non-negative matrix factorization have specific areas of concentration. With clustering, it was found that these profiles are poorly mixed. That is, a single profile is mostly dominating with an area. Or in simple terms, vulnerable people living close by to vulnerable people.

Regression analysis revealed a set of significant findings. First, the number of ambulance calls can be partly explained by the population of an area where the calls were made. Such a conclusion allowed us to test the second type of regression: the number of ambulance calls made during the heat wave by the number of residents over 65 years old. Interestingly, the performance metrics of this regression are lower in comparison to the previous ones. Finally, the set of regression algorithms was applied to match the profiles with the number of heat wave ambulance calls. Quite controversially, no significant relations were found. There are two explanations for that. Either vulnerable individuals are not exposed to extreme heat and therefore, do not call an ambulance. Or even they are exposed, but they do not call an ambulance. This explanation is partly in line with the preliminary mortality data provided by CBS and highly dangerous. A person that is exposed and vulnerable will face some severe issues without some help supplied from outside. Such a finding indicates the importance of other types of infrastructure besides critical represented by ambulance services.

The final section of the chapter was dedicated to resilience capacities assessment. Critical infrastructure represented by hospitals was analyzed with street network analysis. It was found that the length of the shortest driving route for any part of the city does not exceed 4.2 kilometers. Such a distance can be considered as relatively low. The green infrastructure was studied with the means of geospatial analysis. It turned out that the number of trees in the city center is relatively small in comparison to the southwest. Importantly, within this area, there is a high concentration of vulnerable individuals of profile 1. Assessment of adaptive and transformative capacities was simplified to the analysis of social infrastructure represented by libraries. The study revealed that for inhabitants of certain parts of the city, southwest, for instance, the libraries are hardly reachable. Notably, this part of the city has a high concentration of profile number 1. In addition, according to Google's data, some libraries are underrated and perhaps not very popular among the citizens.

5 Conclusion

This chapter aims to summarize the results of the research. After reiterating the accomplishments of this research, answers to all of the formulated research questions are provided. After that, a short discussion of findings and policy advice are following. The chapter ends with limitations and prospects for future work.

5.1. Research summary

The rapid development of urban areas poses numerous problems for both urban planners and citizens. The uncertainties associated with climate change are making the situation even more complicated (Jabareen, 2013). It is unknown when an extreme climatic event will strike, what will be its duration and power. The concept of urban resilience perceived promising in that regards. Planning resilient “fail-to-safe” instead of conventional “fail-safe” cities provide us with evidence in a positive future for urban areas (Ahern, 2011). However, the practice of urban resilience is holding up due to a variety of discrepancies: about its meaning, how it can be assessed, and what are the potential solutions for it. During the last decade, cities have been accumulating a lot of data that is available now for public use. Advances in hardware and software engineering boosted the field of machine learning and provided it with the ability to efficiently process and analyze complex data sets. To bridge the gap between the theory and practice of urban resilience, I propose to use open data and machine learning algorithms for its assessment.

This study resulted in three major outcomes. First, I proposed an innovative conceptual framework. This framework combines the concept of vulnerability with socio-ecological resilience (Chelleri et al., 2015; Jabareen, 2013). Resilience, in this case, is not general, but shock-specific. A city should be resilient to extreme heat in the form of heat waves. To operationalize the components of socio-ecological resilience, I identified three theory-based infrastructure solutions: critical, green, and social (Klinenberg, 2018; Zaidi and Pelling, 2015). Critical infrastructure in this study is represented by ambulance services. Green infrastructure stands for vegetation in the form of trees, shrubs, and grass. Social infrastructure is an umbrella term for places where people can gather, build up, or improve their relationships. It is important to mention that the proposed framework is relatively flexible. It can be reconfigured for other shocks such as flooding or an earthquake.

Second, I developed an advanced computational framework on this base. The goal of this framework is to convert proposed concepts to measurable indicators. To do so, the computational framework uses a variety of open data sources and machine learning algorithms. The data sets include: ambulance calls and temperature, socioeconomic, socio-demographic attributes and health indicators, amenities locations, street networks, and places popularity. The concept of vulnerability is operationalized with time series forecasting (to study the impact of the heat wave on the population), dimensionality reduction and clustering (map vulnerable people in the city), and regression (identify relations between the heat wave and population characteristics) (Pedregosa et al., 2011; Taylor and Letham, 2018). Socio-ecological resilience and corresponding infrastructures are analyzed with network analysis (the lengths of the routes) and geospatial analysis (clustering the objects into groups) (Boeing, 2017).

Third, I demonstrated how the created computational framework could be used in the case study: The Hague. The results of this demonstration are divided into three parts: extensive analysis of the recent heat wave, assessment of population vulnerability, and resilience capacities. It was found that the heat wave had a significant impact not only The Hague's population but also on the citizens of Rotterdam and Amsterdam. This fact is demonstrated by the number of ambulance calls that doubled and even tripled during the heat wave. Additionally, the study confirmed that vulnerability is unequally distributed in time and space. There are certain periods where the number of calls is higher, and there are specific areas in the city where from more calls were made. Prediction of ambulance calls is possible but highly challenging if an extreme even is taken into account. The analysis revealed that the number of ambulance calls made from the urban heat islands in the city center was higher than on average.

Vulnerability assessment strengthens the conclusions of the previous part. A vulnerable population is unequally distributed across the city. Based on identified profiles (a unique combination of citizens' attributes), the areas for specific attention were found. Importantly, the granularity of the data allows us to go beyond the discussion of districts or neighborhoods. 500 by 500 m^2 resolution provides us with a possibility for more precise interventions. Regression analysis revealed two important insights. The number of citizens above 65 years old cannot fully explain the number of ambulance call made during the July 2019 heat wave. It is quite controversial. The simplest vulnerability model imposes that the age group is a single major factor. Second, the improved vulnerability model based on the profiles did not demonstrate significant relations with ambulance calls. This finding stresses the importance of concepts of green and social infrastructure for combating extreme heat threat.

Resilience capacities assessment was divided by the type of infrastructure: assessment of critical, green, and social. It was found that all parts of the city are relatively well-connected to the hospitals. The lengths of the shortest driving route from any hospital to any polygon within the city do not exceed 4.2 kilometers. It turned out that the number of trees in the city center is relatively small in comparison to the southwest. Importantly, within this area, there is a high concentration of vulnerable individuals. The study revealed that for inhabitants of certain parts of the city, southwest, for instance, the libraries (the chosen type of social infrastructure) are hardly reachable. Again, this part of the city has a high concentration of vulnerable individuals and communities.

Currently, most publications use a single case study or innovative solutions as a source for analysis and conclusions. Opposite to it, I used open data and machine learning algorithms to operationalize urban resilience. Such an approach promotes reusability of the results and back up existent empirical evidence with numbers. For instance, we saw that vulnerability, as well as resilience capacities, are unequally distributed across the city. Besides, the mismatch between the number of ambulance calls and demography pointed out the need for bottom-up resilient solutions. The data sets used in this study can be used for establishing better data collection practices. In this research, I was quite limited by the data available on critical, green, and social infrastructures. With more extensive data sets defined in line with the created conceptual framework, we can get a much better picture of how resilient is the city.

To summarize, this study is one of the first advances towards a more computational approach to urban resilience. It identified what, showed how, and demonstrated the outcomes.

5.2. Research questions and conclusions

This study had 5 research questions. The first one is theoretical, the second one is methodological, third and fourth are case study related. The final fifth research questions summarizes the findings. Let us start answering them one by one. The first research question was formulated as follows:

Research question 1

What concepts of urban resilience are useful for planning in case of extreme heat?

The second chapter of the manuscript was aimed to answer this question. During the extensive literature review, it was found that three streams of thoughts dominate in urban resilience literature: general, engineering, and socio-ecological. Since the problem of interest is the effect of extreme heat on vulnerable individuals and communities, it was decided to select socio-ecological resilience as the point of departure. The related concepts are **vulnerability, absorption, recovery, adaptation, and transformation**. These concepts were adapted and linked via two conceptual frameworks from Jabareen (2013) and Chelleri et al. (2015). Each of these concepts is associated with capacity, for example, absorptive capacity. This capacity is a relative measure representing to what extent this attribute is developed. The more capacity the system has, the more resilient it is. To support these resilience capacities, three types of solutions were identified: **green, critical, and social infrastructures** (Klinenberg, 2018; Zaidi and Pelling, 2015). Critical infrastructure represented by ambulances services is a conventional top-down solution that is highly required in case of extreme heat. Green infrastructure or vegetation can absorb the heat created by a heat wave and minimize the effect of its amplifier - urban heat island. The concept of social infrastructure is an excellent example of a bottom-up solution. It denotes places where people gather and build up relationships. People with stronger social ties can take care of each other, and, therefore provide vital help in case of extreme situation. The conceptual framework build on these concepts differs from the ones that are present in the literature in two ways. First, it explicitly states the types and example of solutions that positively contribute to resilience. Importantly, these solutions provide a certain level of flexibility and can be reused for other shocks such as flooding, for example. Second, the framework can be operationalized. That is, it allows a researcher to define specific indicators and measure corresponding capacities.

The second research question was stated as:

Research question 2

Using what machine learning algorithms these concepts of urban resilience can be operationalized?

The third chapter was dedicated to this question. It presented the data sources and the algorithms that should be used for operationalization of created conceptual framework. Ideally, more extensive data sets are needed. Such data sets should include micro-data, mortality rates, and death causes. However, even with the limited number of socio-demographic and socioeconomic attributes, a compelling insight can be generated. Importantly, only open data was proposed for this research.

We suggested using four types of algorithms: **time series forecasting, dimensionality reduction, clustering, and regression**. Besides that, **street network analysis** and **geospatial analysis** played an important role in resilience capacities assessment. Time series forecasting was utilized to analyze the ambulance calls (an indicator of extreme heat vulnerability). Such a method allowed us to a). Identify how the calls are distributed over the day, week, and month; b). Predict the number of calls based on past data. Dimensionality reduction and clustering algorithms were used to map vulnerable individuals and communities. That is, identify who is vulnerable and where they are located. Regression techniques were utilized to understand the connection between the need for help, represented by the ambulance calls, and specific population attributes. Street network analysis was used to evaluate how reachable are specific parts of the city to ambulance services. The aim is to ensure that the most vulnerable people will receive help as fast as possible. The same approach was applied to analyze the reachability of libraries (the chosen type of social infrastructure).

During the analysis of the case study, it was found that proposed algorithms generate valuable insights. **Decomposable time series model** of *fbprophet* allowed us to identify the distribution of calls over the day and week Taylor and Letham (2018). However, one should be extremely careful when relying on such an approach for predicting the number of ambulance calls. Since it operates only with the past data as an input, it cannot predict peaks caused by unexpected extreme events.

Additionally, the study showed that specific algorithms perform better than others. We argue in favor of relatively novel method in machine learning - **Non-negative matrix factorization** (NMF) (Wang and Zhang, 2013). The results of NMF on socio-demographic and socioeconomic data are relatively robust and easy to interpret. The algorithm always highlights the most remarkable features of the data set. Such a feature is extremely useful when dealing with large amounts of variables. Plus, as it was recently reported, NMF can be treated as a clustering algorithm. The main drawback relates to the evaluation of results. The researcher should be supported with information about the case; otherwise, it will be hard to identify the right number of components. **K-Means** and **Mimi Batch K-Means** clustering algorithms showed the best performance (Jain, 2010). This result is due to the nature of the data. After being composed by NMF, the features became grouped in a way that is easily tractable by these algorithms. Regression analysis is was highly challenging and does not reveal the best algorithm in terms of performance.

Important to mention, that the same algorithms can be used for analysis of any other city or other shocks such as flooding, for example. In case of the other city, given the data of the same type, a researcher will be provided with the same output. For another shock only new vulnerability model should be defined.

The third and fourth questions relate to the analysis of the case study:

Research question 3

How vulnerable individuals and communities are distributed across The Hague?

Research question 4

How various areas of The Hague differ in terms of resilience capacities?

The fourth chapter presents the results of the case study analysis and answers third and fourth research questions. It was found that **vulnerable individuals and communities are highly unequally distributed over The Hague**. Analysis of ambulance calls showed that certain parts of the city have more calls than others. These findings hold for both the calls made during the heat wave and the average number of calls. During vulnerability assessment, it was found that some areas of The Hague have higher concentrations of vulnerable individuals and communities. For instance, old citizens that have mobility issues and serious sicknesses leave on the northwest of the city. The areas close to the city center have a high number of low-income single-household inhabitants with the same health problems. According to the literature, both these groups represent people that are vulnerable to extreme heat. Since these places have such demographics, they require more attention. Plus, the neighborhoods close to the city center were identified as urban heat islands. The wast majority of the aged population live on the north of the city, and therefore, they are not subject to urban heat island effect. However, the second group of people is.

Resilience capacities assessment was based on three types of infrastructure: critical, green, and social. Similar to vulnerability, **resilience capacities are also distributed unequally over The Hague**. It was found that the length of the shortest driving route for any part of the city does not exceed 4.2 kilometers. Such a distance can be considered as relatively low. It turned out that the number of trees in the city center is relatively small in comparison to the southwest. Importantly, within this area, there is a high concentration of vulnerable individuals. The study revealed that for inhabitants of certain parts of the city, southwest, for instance, the libraries are hardly reachable. And again this part of the city has a high concentration of vulnerable citizens.

5.3. Discussion

The discussion section will be driven by the final fifth research question:

> **Research question 5**
>
> Is The Hague resilient to extreme heat?

The conducted study showed that extreme threat is a real problem for The Hague. Besides The Hague, other Dutch cities proved to be highly sensitive to it: Amsterdam and Rotterdam. The rise in ambulance calls and preliminary mortality data seem to be a call for urgent measures. The findings of the vulnerability assessment provided us with points of attention. Now we know how vulnerable individuals and communities are distributed across the city. Some parts of The Hague have a higher concentration of the elderly with mobility problems, whereas others have a higher number of low-income single-household residents with prolonged health issues. Resilience capacities assessment revealed points for potential intervention. Different parts of The Hague have a different amount of vegetation, accessibility to libraries. The Hague has complex demographics. There a variety of reasons why specific population groups are concentrated within certain areas: historical, economic, etcetera. Since the equal distribution of vulnerability seems to be hard to reach if not impossible, the improvements in resilience capacities are vital. The ultimate goal here is to have high resilience in places with high vulnerability. Currently, it is not the case. Achilles' heel of this study is the proposed metrics for absorptive, adaptive, and transformative capacities. They are simplifications. However, I believe that this research will lay down the foundations for more elaborate data-based studies of urban resilience. A more in-depth understanding of how each of these concepts can be operationalized will allow us to have a straightforward and hopefully positive answer to this question.

5.4. Policy advice

Climate change continuously exercising modern urban areas. Some with hurricanes and earthquakes, other with extreme heat. Uncertainties involved in climatic processes do not allow us neither fix the date of the next heat wave coming nor estimate its capacity. Urban resilience seems to be the right if not only one planning possible approach in that regards. No matter when the heat wave will strike, and no matter how strong it will be, The Hague must absorb, adapt, and transform. More specifically, the damage made to vulnerable individuals and communities show be minimized. Vulnerable to extreme heat people conventionally defined as older than 65. The fact that the Netherlands experiences population ageing amplify the importance of measures aimed at this population group. Other vulnerability models operate with a more complex combination of individual attributes: low-income single-household individuals with health issues. The policy measures need to be concentrated around both of those. To tackle extreme heat threat, we argue for investments in critical, green, and social infrastructures. Critical infrastructure represented by the hospitals, and it is the first instance for people in need. We need to ensure that The Hague's hospitals are ready for sudden increases in demand. When such a conventional measure fails, green and social infrastructure start to play an essential role. There is a lack of vegetation in the city center, and it has already resulted in an urban heat island effect. Importantly, this part of the city is populated by vulnerable individuals. Besides, there is a need for the development of various types of social infrastructure: libraries, community centers, and so on. Such places promote establishing new social ties and relationships among citizens. It was empirically proven having such connections is crucial when the extreme event is happening.

5.5. Limitations and further work

This study is one of the first attempts of quantitative analysis of such a complex concept as urban resilience. More specifically, its operationalization with the methods of machine learning. Quantification always brings certain challenges to the researcher. It implies simplifications and simplifications, in turn, limit the potential of the concept. We stated that critical, green, and social infrastructures would positively contribute to urban resilience. But the metrics proposed did not allow us to make an absolute decision in favor of being resilient or not.

For instance, we have not talked about the actual amount of vegetation required to decrease the temperature below a certain threshold. Instead, we were driven by am assumption: the more, the better. The assessment of critical infrastructure was also narrow. Two facts did not allow us to come up with a straightforward conclusion. First, the actual capacity of each hospital is unknown. The demand for ambulance services skyrocketed during the heat wave, and perhaps some extra facilities were used. If so, we need to have more extensive data sets to create a more comprehensive assessment model. Second, we do not know how the calls are distributed across the hospitals. Possible, almost all requests are handled by a single hospital, whereas we assume that all of them have the same ”priority.” Therefore, the proposed metric (proximity) is rather a call for a more elaborate analysis of critical infrastructure than the absolute decision in favor of high absorptive capacity and as a result, urban resilience. A similar argument holds for adaptive and transformative capacities assessment. When simplifying it to only reachability of libraries, the actual outcomes were not measured.

Thus, it is hard to measure resilience capacities, and, moreover, account for, with a specific degree of certainty, that proposed measures will lead to the most resilient city. One can argue that these measures are not the best ones when others find better ways to quantify them. To overcome this, we argue for more extensive theoretical, what is more important, and, empirical research.

Modern cities generate more and more data. It pushes researchers to find ways to bridge the theory with the data sets that recently became available. With more data, for instance, mortality rates and causes, better results can be obtained. However, even with the most extensive data set, data analysis can misinterpret such a complex system as a city. To overcome this limitation, we propose to conduct a simulation study based to test the assumptions of this study. Given the fact that there is only empirical evidence and a lack of data that can demonstrate the full potential of social infrastructure, a simulation study also can be useful.

Bibliography

100 Resilient Cities. What is urban resilience. URL http://www.100resilientcities.org/resources/. Last accessed 22 June 2019.

Russell L Ackoff and Jamshid Gharajedaghi. On the mismatch between systems and their models. *URL www. acasa. upenn. edu/System_MismatchesA. pdf*, 2003.

W Neil Adger, Shardul Agrawala, M Monirul Qader Mirza, Cecilia Conde, Karen o'Brien, Juan Pulhin, Roger Pulwarty, Barry Smit, and Kiyoshi Takahashi. Assessment of adaptation practices, options, constraints and capacity. *Climate change*, 200, 2007.

Jack Ahern. From fail-safe to safe-to-fail: Sustainability and resilience in the new urban world. *Landscape and urban Planning*, 100(4):341–343, 2011.

Marina Alberti. The effects of urban patterns on ecosystem function. *International regional science review*, 28(2):168–192, 2005.

Daniel Oudin Åström, Forsberg Bertil, and Rocklöv Joacim. Heat wave impact on morbidity and mortality in the elderly population: a review of recent studies. *Maturitas*, 69(2):99–105, 2011.

E Michael Azoff. *Neural network time series forecasting of financial markets*. John Wiley & Sons, Inc., 1994.

Teresa Barata-Salgueiro and Feyzan Erkip. Retail planning and urban resilience – an introduction to the special issue. *Cities*, 36:107–111, Feb 2014. ISSN 02642751. doi: 10.1016/j.cities.2013.01.007.

Jon Barnett. Adapting to climate change in pacific island countries: the problem of uncertainty. *World development*, 29(6):977–993, 2001.

Michael Batty. *The size, scale, and shape of cities*. American Association for the Advancement of Science, 2008.

Michael Batty. *The new science of cities*. MIT press, 2013.

Michael Batty. Urban analytics defined. *Environment and Planning B: Urban Analytics and City Science*, 46(3):403–405, Mar 2019. ISSN 2399-8083. doi: 10.1177/2399808319839494.

Benjamin Bengfort, Rebecca Bilbro, and Tony Ojeda. *Applied Text Analysis with Python: Enabling Language-aware Data Products with Machine Learning.* " O'Reilly Media, Inc.", 2018.

Fikret Berkes, Carl Folke, and Johan Colding. *Linking social and ecological systems: management practices and social mechanisms for building resilience.* Cambridge University Press, 2000.

L. M. A. Bettencourt, J. Lobo, D. Helbing, C. Kuhnert, and G. B. West. Growth, innovation, scaling, and the pace of life in cities. 104(17):7301–7306. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0610172104. URL http://www.pnas.org/cgi/doi/10.1073/pnas.0610172104.

Geoff Boeing. Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65:126–139, 2017.

Daniel B Botkin. Discordant harmonies: a new ecology for the twenty-first century. 1990.

Loeiz Bourdic, Serge Salat, and Caroline Nowacki. Assessing cities: a new system of cross-scale spatial indicators. *Building Research & Information*, 40(5):592–605, 2012.

Diana E Bowler, Lisette Buyung-Ali, Teri M Knight, and Andrew S Pullin. Urban greening to cool towns and cities: A systematic review of the empirical evidence. *Landscape and urban planning*, 97(3):147–155, 2010.

Fridolin Brand and Kurt Jax. Focusing the meaning (s) of resilience: resilience as a descriptive concept and a boundary object. *Ecology and society*, 12(1), 2007.

Steve Carpenter, Brian Walker, J Marty Anderies, and Nick Abel. From metaphor to measurement: resilience of what to what? *Ecosystems*, 4(8):765–781, 2001.

Centraal Bureau voor de Statistiek. Half of dutch adults will be over 50 in 2019, 2014. URL https://www.cbs.nl/en-gb/news/2014/38/half-of-dutch-adults-will-be-over-50-in-2019. Last accessed 01 July 2019.

Centraal Bureau voor de Statistiek. Cbs urban data centre the hague, 2017. URL https://www.cbs.nl/en-gb/our-services/urban-data-centres/locaties/cbs-urban-data-centre-the-hague. Last accessed 27 May 2019.

Centraal Bureau voor de Statistiek. Heatwaves caused minor rise in mortality, 2018a. URL https://www.cbs.nl/en-gb/news/2018/33/heatwaves-caused-minor-rise-in-mortality#15ce4860-4d04-49e1-b88a-242dbbfaf925. Last accessed 27 May 2019.

Centraal Bureau voor de Statistiek. More deaths in winter, 2018b. URL https://www.cbs.nl/en-gb/news/2018/07/more-deaths-in-winter. Last accessed 14 June 2019.

Centraal Bureau voor de Statistiek. Hogere sterfte tijdens recente hittegolf, 2019a. URL https://www.cbs.nl/nl-nl/nieuws/2019/32/hogere-sterfte-tijdens-recente-hittegolf. Last accessed 12 August 2019.

Centraal Bureau voor de Statistiek. Kaart van 100 meter bij 100 meter met statistieken, 2019b. URL https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische%20data/kaart-van-100-meter-bij-100-meter-met-statistieken.

Lorenzo Chelleri, James J Waters, Marta Olazabal, and Guido Minucci. Resilience trade-offs: addressing multiple scales and temporal aspects of urban resilience. *Environment and Urbanization*, 27(1):181–198, 2015.

Jon Coaffee. *Terrorism, risk and the global city: Towards urban resilience*. Routledge, 2016.

Raven Cretney. Resilience for whom? emerging critical geographies of socio-ecological resilience. *Geography Compass*, 8(9):627–640, 2014.

Philip R Crowe, Karen Foley, and Marcus J Collier. Operationalizing urban resilience through a framework for adaptive co-management and design: Five experiments in urban planning practice and policy. *Environmental science & policy*, 62:112–119, 2016.

Susan L Cutter, Bryan J Boruff, and W Lynn Shirley. Social vulnerability to environmental hazards. *Social science quarterly*, 84(2):242–261, 2003.

Miriam Van der Sangen. Vulnerable groups at the core of the hague local policy, 2018. URL https://www.cbs.nl/en-gb/corporate/2018/03/vulnerable-groups-at-the-core-of-the-hague-local-policy. Last accessed 29 May 2019.

Kevin C Desouza and Trevor H Flanery. Designing, planning, and managing resilient cities: A conceptual framework. *Cities*, 35:89–99, 2013.

Chris Ding, Xiaofeng He, and Horst D. Simon. On the equivalence of nonnegative matrix factorization and spectral clustering. In *Proceedings of the 2005 SIAM International Conference on Data Mining*, page 606–610. Society for Industrial and Applied Mathematics, Apr 2005. ISBN 978-0-89871-593-4. doi: 10.1137/1.9781611972757.70. URL https://epubs.siam.org/doi/10.1137/1.9781611972757.70.

Driven Data Labs. Cookiecutter data science, 2019. URL https://drivendata.github.io/cookiecutter-data-science/#contributing. Last accessed 19 Aug 2019.

Andreas Duit, Victor Galaz, Katarina Eckerberg, and Jonas Ebbesson. Governance, complexity, and resilience, 2010.

Nathan L Engle. Adaptive capacity and its assessment. *Global environmental change*, 21(2): 647–656, 2011.

Envato. Portfolio, 2019. URL https://graphicriver.net/user/alexzel/portfolio.

Henrik Ernstson, Sander E Van der Leeuw, Charles L Redman, Douglas J Meffert, George Davis, Christine Alfsen, and Thomas Elmqvist. Urban transitions: on urban resilience and human-dominated ecosystems. *Ambio*, 39(8):531–545, 2010.

Alessandra Feliciotti, Ombretta Romice, and Sergio Porta. Urban regeneration, masterplans and resilience: the case of the gorbals in glasgow. *Urban Morphology*, 21(1):61–79, 2017.

Joseph Fiksel. Sustainability and resilience: toward a systems approach. *Sustainability: Science, Practice and Policy*, 2(2):14–21, 2006.

David Fletcher and Mustafa Sarkar. Psychological resilience: A review and critique of definitions, concepts, and theory. *European psychologist*, 18(1):12, 2013.

Carl Folke, Steve Carpenter, Thomas Elmqvist, Lance Gunderson, Crawford S Holling, and Brian Walker. Resilience and sustainable development: building adaptive capacity in a world of transformations. *AMBIO: A journal of the human environment*, 31(5):437–441, 2002.

Richard Friend and Marcus Moench. What is the purpose of urban climate resilience? implications for addressing poverty and vulnerability. *Urban Climate*, 6:98–113, 2013.

Aurélien Géron. *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems.* " O'Reilly Media, Inc.", 2017.

David R Godschalk. Urban hazard mitigation: creating resilient cities. *Natural hazards review*, 4(3):136–143, 2003.

Carel Harmsen and Joop Garssen. July heat causes one thousand extra deaths, 2006. URL https://www.cbs.nl/en-gb/news/2006/35/july-heat-causes-one-thousand-extra-deaths. Last accessed 27 May 2019.

Carel Harmsen and Joop Garssen. Recent heat claims 500 extra lives, 2010. URL https://www.cbs.nl/en-gb/news/2010/29/recent-heat-claims-500-extra-lives. Last accessed 27 May 2019.

Andrew C Harvey and Simon Peters. Estimation procedures for structural time series models. *Journal of Forecasting*, 9(2):89–108, 1990.

Ed Hawkins. Warming stripes for netherlands from 1901-2018, 2019. URL https://showyourstripes.info/. Last accessed 01 July 2019.

Daniel Henstra. Toward the climate-resilient city: extreme weather and urban climate adaptation policies in two canadian provinces. *Journal of Comparative Policy Analysis: Research and Practice*, 14(2):175–194, 2012.

Wandl A. Hoeven, FD Van der. Hittekaart den haag, 2017. Data was retrieved from Den Haag Dataplatform, https://denhaag.dataplatform.nl/#/data/83b26801-e622-48b8-80bc-ee15fdeb7edf.

Wandl A. Hoeven, FD Van der. Haagse hitte: Het haagse warmte-eiland in kaart gebracht. Technical report, 2018. URL https://books.bk.tudelft.nl/index.php/press/catalog/book/isbn.9789463660037. Last accessed 03 June 2019.

Crawford S Holling. Resilience and stability of ecological systems. *Annual review of ecology and systematics*, 4(1):1–23, 1973.

Crawford S Holling. Understanding the complexity of economic, ecological, and social systems. *Ecosystems*, 4(5):390–405, 2001.

Harold Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, 24(6):417, 1933.

Intergovernmental Panel on Climate Change. Ar5 climate change 2014: Impacts, adaptation, and vulnerability: Urban areas. Technical report, 2018. URL https://www.ipcc.ch/site/assets/uploads/2018/02/WGIIAR5-Chap8_FINAL.pdf. Last accessed 23 April 2019.

Yosef Jabareen. Building a conceptual framework: philosophy, definitions, and procedure. *International journal of qualitative methods*, 8(4):49–62, 2009.

Yosef Jabareen. Planning the resilient city: Concepts and strategies for coping with climate change and environmental risk. *Cities*, 31:220–229, 2013.

Jane Jacobs. The death and life of great american cities. 1961.

Anil K. Jain. Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 31(8): 651–666, Jun 2010. ISSN 01678655. doi: 10.1016/j.patrec.2009.09.011.

Ian T. Jolliffe and Jorge Cadima. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202, Apr 2016. ISSN 1364-503X, 1471-2962. doi: 10.1098/rsta.2015.0202.

Eric Jones, Travis Oliphant, Pearu Peterson, et al. Scipy: Open source scientific tools for python. 2001.

Eric Klinenberg. Review of heat wave: social autopsy of disaster in chicago. *New England Journal of Medicine*, 348(7):666–667, 2003.

Eric Klinenberg. *Heat wave: A social autopsy of disaster in Chicago*. University of Chicago Press, 2015.

Eric Klinenberg. *Palaces for the people: How social infrastructure can help fight inequality, polarization, and the decline of civic life*. Crown, 2018.

Kim Knowlton, Miriam Rotkin-Ellman, Galatea King, Helene G Margolis, Daniel Smith, Gina Solomon, Roger Trent, and Paul English. The 2006 california heat wave: impacts on hospitalizations and emergency department visits. *Environmental health perspectives*, 117(1): 61–67, 2008.

Jessica E Lamond and David G Proverbs. Resilience to flooding: lessons from international comparison. *Proceedings of the Institution of Civil Engineers-Urban design and planning*, 162(2):63–70, 2009.

Daniel D Lee and H Sebastian Seung. Algorithms for non-negative matrix factorization. page 7.

Robin Leichenko. Climate change and urban resilience. *Current opinion in environmental sustainability*, 3(3):164–168, 2011.

Simon A Levin and Simon Lelvin. *Fragile dominion: complexity and the commons*. Number 574.5248 L4. Perseus Books Reading, MA, 1999.

Kubiš Lukáš. darkskylib 0.3.91, 2019. URL `https://pypi.org/project/darkskylib/`.

Fionn MacKillop. Climatic city: Two centuries of urban planning and climate science in manchester (uk) and its region. *Cities*, 29(4):244–251, 2012.

Hiroshi Maruyama. Taxonomy and general strategies for resilience. In *Urban Resilience*, pages 3–21. Springer, 2016.

Gordon McGranahan, Deborah Balk, and Bridget Anderson. The rising tide: assessing the risks of climate change and human settlements in low elevation coastal zones. *Environment and urbanization*, 19(1):17–37, 2007.

Leland McInnes, John Healy, and Steve Astels. hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software*, 2(11), mar 2017. doi: 10.21105/joss.00205. URL `https://doi.org/10.21105%2Fjoss.00205`.

Benjamin McLellan, Qi Zhang, Hooman Farzaneh, N Agya Utama, and Keiichi N Ishihara. Resilience, sustainability and risk management: A focus on energy. *Challenges*, 3(2):153–182, 2012.

Sara Meerow, Joshua P Newell, and Melissa Stults. Defining urban resilience: A review. *Landscape and urban planning*, 147:38–49, 2016.

Abid Mehmood. Of resilient places: planning for urban resilience. *European Planning Studies*, 24(2):407–419, 2016.

Michele-Lee Moore, Ola Tjornbo, Elin Enfors, Corrie Knapp, Jennifer Hodbod, Jacopo A Baggio, Albert Norström, Per Olsson, and Duan Biggs. Studying the complexity of change: toward an analytical framework for understanding deliberate social-ecological transformations. *Ecology and Society*, 19(4):54, 2014.

Mary P Naughton, Alden Henderson, Maria C Mirabelli, Reinhard Kaiser, John L Wilhelm, Stephanie M Kieszak, Carol H Rubin, and Michael A McGeehin. Heat-related mortality during a 1999 heat wave in chicago. *American journal of preventive medicine*, 22(4):221–227, 2002.

Nederlandse Omroep Stichting. Storing kpn: 112 niet bereikbaar, landelijk noodnummer. *NOS Journaal*, 2019. URL `https://nos.nl/artikel/2290486-storing-kpn-112-niet-bereikbaar-landelijk-noodnummer.html`.

Karen O'Brien, Bronwyn Hayward, and Fikret Berkes. Rethinking social contracts: building resilience in a changing climate. *Ecology and Society*, 14(2), 2009.

Angélica V. Ospina. Big data for resilience: Storybook. Technical report, 2018. URL `https://www.iisd.org/sites/default/files/publications/big-data-resilience-storybook.pdf`. Last accessed 06 May 2019.

Elinor Ostrom. Polycentric systems for coping with collective action and global environmental change. *Global environmental change*, 20(4):550–557, 2010.

Christopher O'Malley, Poorang AE Piroozfarb, Eric RP Farr, and Jonathan Gates. An investigation into minimizing urban heat island (uhi) effects: A uk perspective. *Energy Procedia*, 62:72–80, 2014.

Marie S O'Neill, Rebecca Carter, Jonathan K Kish, Carina J Gronlund, Jalonne L White-Newsome, Xico Manarolla, Antonella Zanobetti, and Joel D Schwartz. Preventing heat-related morbidity and mortality: new approaches in a changing climate. *Maturitas*, 64(2): 98–103, 2009.

R. K. Pachauri and Leo Mayer, editors. *Climate change 2014: synthesis report*. Intergovernmental Panel on Climate Change, Geneva, Switzerland, 2015. ISBN 978-92-9169-143-2. OCLC: 914851124.

Karl Pearson. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, 1901.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Karin Peters, Birgit Elands, and Arjen Buijs. Social interactions in urban parks: stimulating social cohesion? *Urban forestry & Urban greening*, 9(2):93–100, 2010.

Steward TA Pickett, Mary L Cadenasso, J Morgan Grove, Christopher G Boone, Peter M Groffman, Elena Irwin, Sujay S Kaushal, Victoria Marshall, Brian P McGrath, Charles H Nilon, et al. Urban ecological systems: Scientific foundations and a decade of progress. *Journal of Environmental Management*, 92(3):331–362, 2011.

Publieke Dienstverlening Op de Kaart. Hét platform voor hoogwaardige geodata. URL `https://www.pdok.nl/`. Last accessed 08 July 2019.

Resilient The Hague. The hague 100 resilient cities: Preliminary resilience assessment. Technical report, 2018. URL `http://100resilientcities.org/wp-content/uploads/2018/03/The-Hague-PRA-English.pdf`. Last accessed 07 May 2019.

Resilient The Hague. The hague resilience strategy. Technical report, 2019. URL `http://100resilientcities.org/wp-content/uploads/2019/05/Resilience-Strategy-The-Hague-English.pdf`. Last accessed 13 Aug 2019.

Sergio J. Rey and Luc Anselin. PySAL: A Python Library of Spatial Analytical Methods. *The Review of Regional Studies*, 37(1):5–27, 2007.

Rijksinstituut voor Volksgezondheid en Milieu. Nationaal hitteplan, 2011. URL `https://www.rivm.nl/hitte/wat-doet-rivm`. Last accessed 28 May 2019.

Rijksinstituut voor Volksgezondheid en Milieu. Gezondheid per wijk en buurt 2016, 2016. URL `https://statline.rivm.nl/portal.html?_la=nl&_catalog=RIVM&tableId=50052NED&_theme=71`.

Rijksinstituut voor Volksgezondheid en Milieu. Higher mortality during heat-wave in july in the netherlands, 2019. URL `https://www.rivm.nl/en/news/higher-mortality-during-heatwave-in-july-in-netherlands`.

Ahmed Memon Rizwan, Leung YC Dennis, and LIU Chunho. A review on the generation, determination and mitigation of urban heat island. *Journal of Environmental Sciences*, 20 (1):120–128, 2008.

Judith Rodin. *The resilience dividend: being strong in a world where things go wrong*. Public Affairs, 2014.

Adam Rose. Defining and measuring economic resilience to disasters. *Disaster Prevention and Management: An International Journal*, 13(4):307–314, 2004.

Matthias Ruth and Dana Coelho. Understanding and managing the complexity of urban systems under climate change. *Climate Policy*, 7(4):317–336, 2007.

Roland W Scholz, Yann B Blumer, and Fridolin S Brand. Risk, vulnerability, robustness, and resilience from a decision-theoretic perspective. *Journal of Risk Research*, 15(3):313–330, 2012.

Joel Schwartz. Who is sensitive to extremes of temperature?: A case-only analysis. *Epidemiology*, 16(1):67–72, 2005.

Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, May 2014. ISBN 978-1-139-95274-3. Google-Books-ID: Hf6QAwAAQBAJ.

Ayyoob Sharifi and Yoshiki Yamagata. Urban resilience assessment: Multiple dimensions, criteria, and indicators. In *Urban Resilience*, pages 259–276. Springer, 2016.

James Simmie and Ron Martin. The economic resilience of regions: towards an evolutionary approach. *Cambridge journal of regions, economy and society*, 3(1):27–43, 2010.

Barry Smit and Johanna Wandel. Adaptation, adaptive capacity and vulnerability. *Global environmental change*, 16(3):282–292, 2006.

Carolien HM Smits, Hugo K van den Beld, Marja J Aartsen, and Johannes JF Schroots. Aging in the netherlands: State of the art and science. *The Gerontologist*, 54(3):335–343, 2014.

Marc Stefanon, Fabio D'Andrea, and Philippe Drobinski. Heatwave classification over europe and the mediterranean region. *Environmental Research Letters*, 7(1):014023, 2012.

Angus Stevenson. *Oxford dictionary of English*. Oxford University Press, USA, 2010.

Stefan Straub. Natural catastrophe review: Series of hurricanes makes 2017 year of highest insured losses ever, 2018.

Sarah Swalheim and David Dodman. Building resilience: how the urban poor can drive climate adaptation. *IIED, London*, 2008.

Carys Swanwick, Nigel Dunnett, and Helen Woolley. Nature, role and value of green space in towns and cities: An overview. *Built Environment (1978-)*, pages 94–106, 2003.

Sean J Taylor and Benjamin Letham. Forecasting at scale. *The American Statistician*, 72(1): 37–45, 2018.

The Dark Sky Company. Dark sky api, 2019. URL `https://darksky.net/dev`.

Leslie Ronald George Treloar. *The physics of rubber elasticity*. Oxford University Press, USA, 1975.

Stephen Tyler and Marcus Moench. A framework for urban climate resilience. *Climate and development*, 4(4):311–326, 2012.

UN-Habitat. Cities and climate change: Global report on human settlements 2011. Technical report, 2011. URL `https://unhabitat.org/books/cities-and-climate-change-global-report-on-human-settlements-2011/#`. Last accessed 23 April 2019.

United Nations. World population prospects: Key findings & advances tables. Technical report, 2017. URL `https://esa.un.org/unpd/wpp/publications/Files/WPP2017_KeyFindings.pdf`. Last accessed 23 April 2019.

United Nations. World 'Nearing Critical Point of No Return' on Climate Change, Delegate Warns, as Second Committee Debates Sustainable Development, 2018a. URL `https://www.un.org/press/en/2018/gaef3500.doc.htm`.

United Nations. The sustainable development goals report 2018. Technical report, 2018b. URL `https://unstats.un.org/sdgs/files/report/2018/TheSustainableDevelopmentGoalsReport2018-EN.pdf`. Last accessed 23 April 2019.

Lawrence J Vale. The politics of resilient cities: whose resilience and whose city? *Building Research & Information*, 42(2):191–201, 2014.

Lawrence J Vale and Thomas J Campanella. *The resilient city: How modern cities recover from disaster*. Oxford University Press, 2005.

Liesbet Van Zoonen. Privacy concerns in smart cities. *Government Information Quarterly*, 33 (3):472–480, 2016.

Brian Walker and David Salt. *Resilience thinking: sustaining ecosystems and people in a changing world*. Island press, 2012.

Brian Walker, Crawford S Holling, Stephen Carpenter, and Ann Kinzig. Resilience, adaptability and transformability in social–ecological systems. *Ecology and society*, 9(2), 2004.

Y. Wang and Y. Zhang. Nonnegative matrix factorization: A comprehensive review. *IEEE Transactions on Knowledge and Data Engineering*, 25(6):1336–1353, Jun 2013. ISSN 1041-4347. doi: 10.1109/TKDE.2012.51.

J Arjan Wardekker, Arie de Jong, Joost M Knoop, and Jeroen P van der Sluijs. Operationalising a resilience approach to adapting an urban delta to uncertain climate changes. *Technological Forecasting and Social Change*, 77(6):987–998, 2010.

Edward J Wegman. Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association*, 85(411):664–675, 1990.

World Bank. Urban development overview, 2019. URL `https://www.worldbank.org/en/topic/urbandevelopment/overview`. Last accessed 23 April 2019.

Jianguo Wu. Urban ecology and sustainability: The state-of-the-science and future directions. *Landscape and Urban Planning*, 125:209–221, May 2014. ISSN 01692046. doi: 10.1016/j.landurbplan.2014.01.018.

Jianyong Wu, Ying Zhou, Yang Gao, Joshua S Fu, Brent A Johnson, Cheng Huang, Young-Min Kim, and Yang Liu. Estimation and uncertainty analysis of impacts of future heat waves on mortality in the eastern united states. *Environmental health perspectives*, 122(1): 10–16, 2013.

Yoshiki Yamagata and Ayyoob Sharifi. *Resilience-Oriented Urban Planning: Theoretical and Empirical Insights*, volume 65. Springer, 2018.

R Zehra Zaidi and Mark Pelling. Institutionally configured risk: Assessing urban resilience and disaster risk reduction to heat wave risk in london. *Urban Studies*, 52(7):1218–1233, 2015.

BlaĹž Zupan et al. Nimfa: A python library for nonnegative matrix factorization. *Journal of Machine Learning Research*, 13(Mar):849–853, 2012.

List of Tables

List of Figures

A

Supplementary information

This appendix chapter aims to provide the reader with some extra information about the findings of this study. To simplify the navigation, it was decided to make the structure of it identical to the structure of the thesis. Thus, the sections of it have the same names as the chapters of the original manuscript. The sections consist of additional graphs, tables and explanations.

A.1. Introduction

In the first chapter, while discussing the heat wave mortality data from Centraal Bureau voor de Statistiek (2018a), a couple of points were emphasized. Let us discuss them in more details with some new graphics. First, the heat wave of 2006 was the most devastating comparing to 2010 and 2018. The peak of mortality in 2006 depicted with a red dot on the figure below. The corresponding number is 3046, whereas the average value of the blue line is 2554. The difference is striking: almost 500 extra deaths.

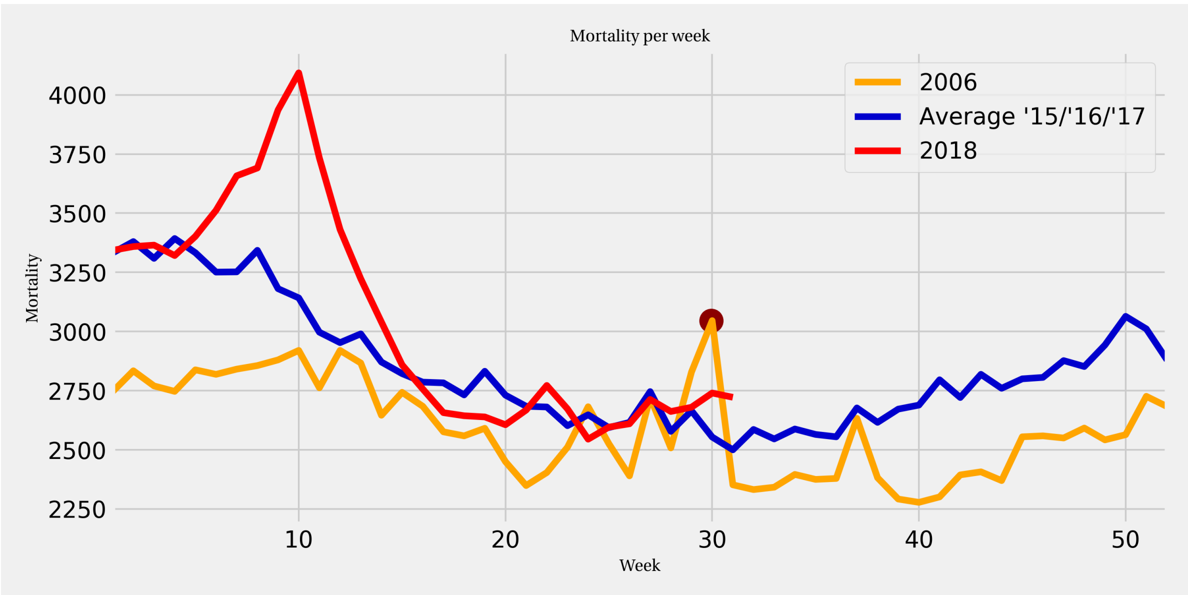

Figure A.1: Contrasting the maximum mortality of the 2006 heatwave with the average value. Data of the graph available at Centraal Bureau voor de Statistiek (2018a).

Second, the actual comparison of 2018 with 2006 hardly possible. At first glance, it seems that the increase in mortality in 2018 was not so significant as it was for 2006. However, a remarkable event preceded the 2018 heatwave. An epidemic of flu significantly affected the vulnerable population during the winter period (Centraal Bureau voor de Statistiek, 2018b). An increase can be seen from week 3 to a peak in week 10 (thick red line on the graph below). Thus, by using only these numbers it is difficult to evaluate the efficiency of responsible parties such as ambulance services, informal caregivers, etcetera.

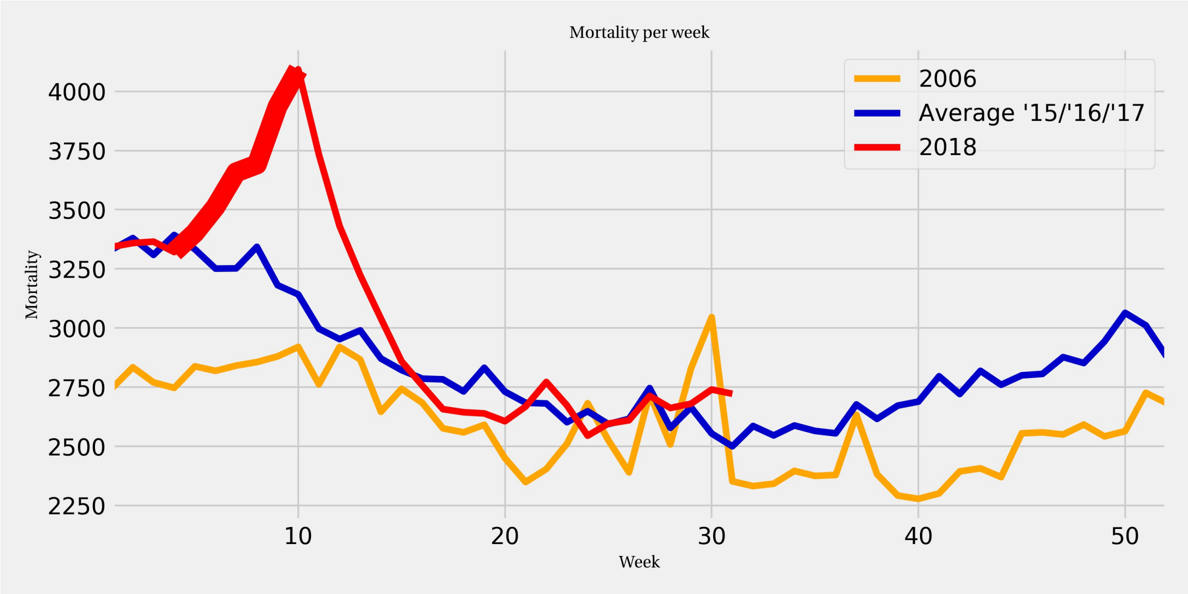

Figure A.2: The possible effect of the flu on mortality in 2018 highlighted. Data of the graph available at Centraal Bureau voor de Statistiek (2018a).

According to the preliminary data, 2019 heat wave mortality almost reached the values of 2006 with the value of 2964. The difference is only 82. And again, this value is much higher than the average depicted with a blue line.

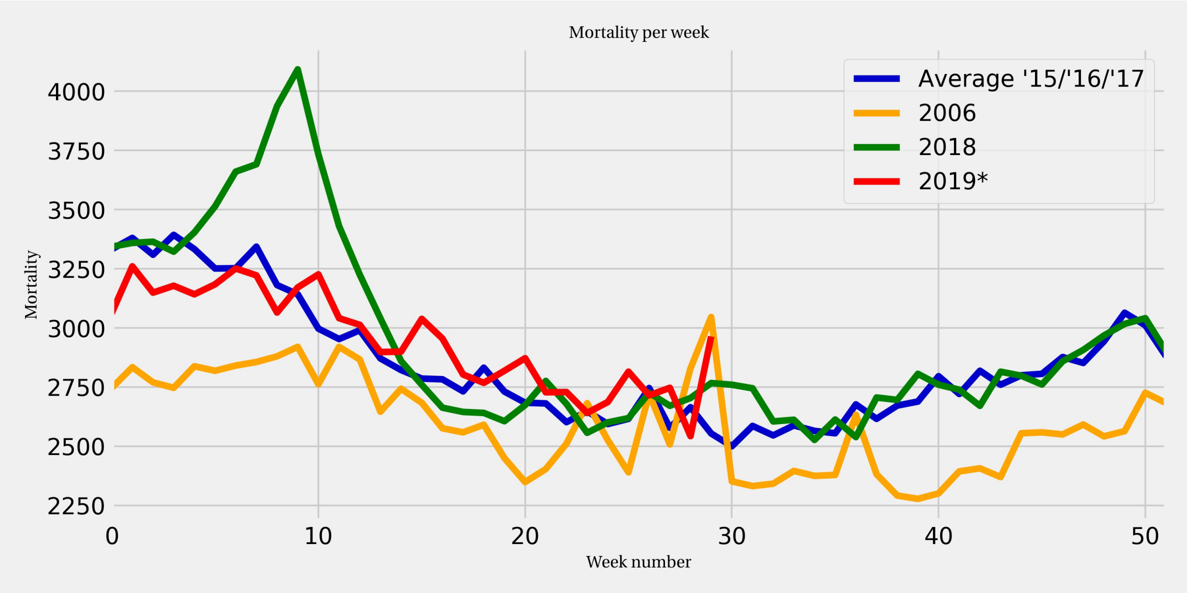

Figure A.3: Comparison of mortality during the major heat waves with averages from 2015, 2016, and 2017. Data of the graph available at Centraal Bureau voor de Statistiek (2019a).

The age group affected the most is above 80 years old, then age group of from 65 to 79, and finally the effect is the least for others (below 65 years old). The figure A.4 shows this difference.

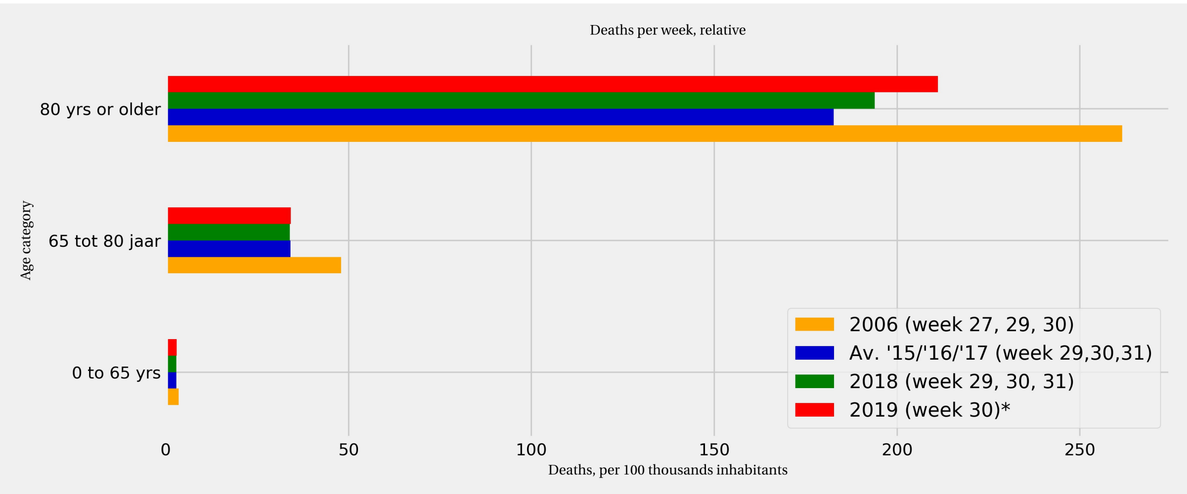

Figure A.4: Differences in deaths in various age groups. Data of the graph available at Centraal Bureau voor de Statistiek (2019a).

The final remark relates to the differences in numbers between the regions of the Netherlands. The east of the Netherlands has the highest mortality numbers for above 80 years old age group. The west and south regions share the second position. The north of the country has a decrease in numbers compared to 2018 and a slight increase in comparison with the average.

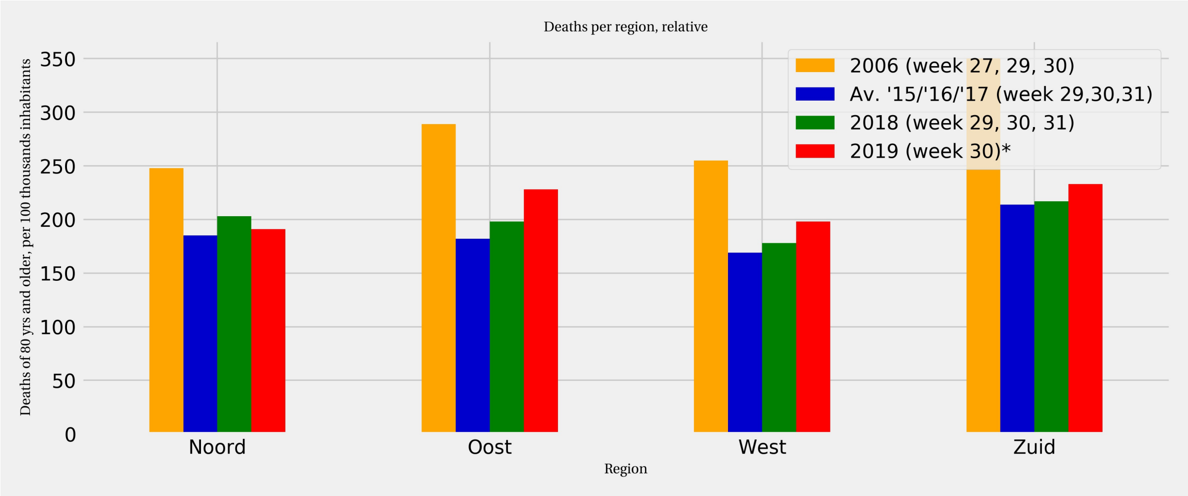

Figure A.5: Comparison of deaths for above 80 years old citizens in different regions. Data of the graph available at Centraal Bureau voor de Statistiek (2019a).

A.2. Results

A.2.1. Analysis of heat waves and urban heat islands

The following figures demonstrates the distribution of ambulance calls for Rotterdam and Amsterdam.

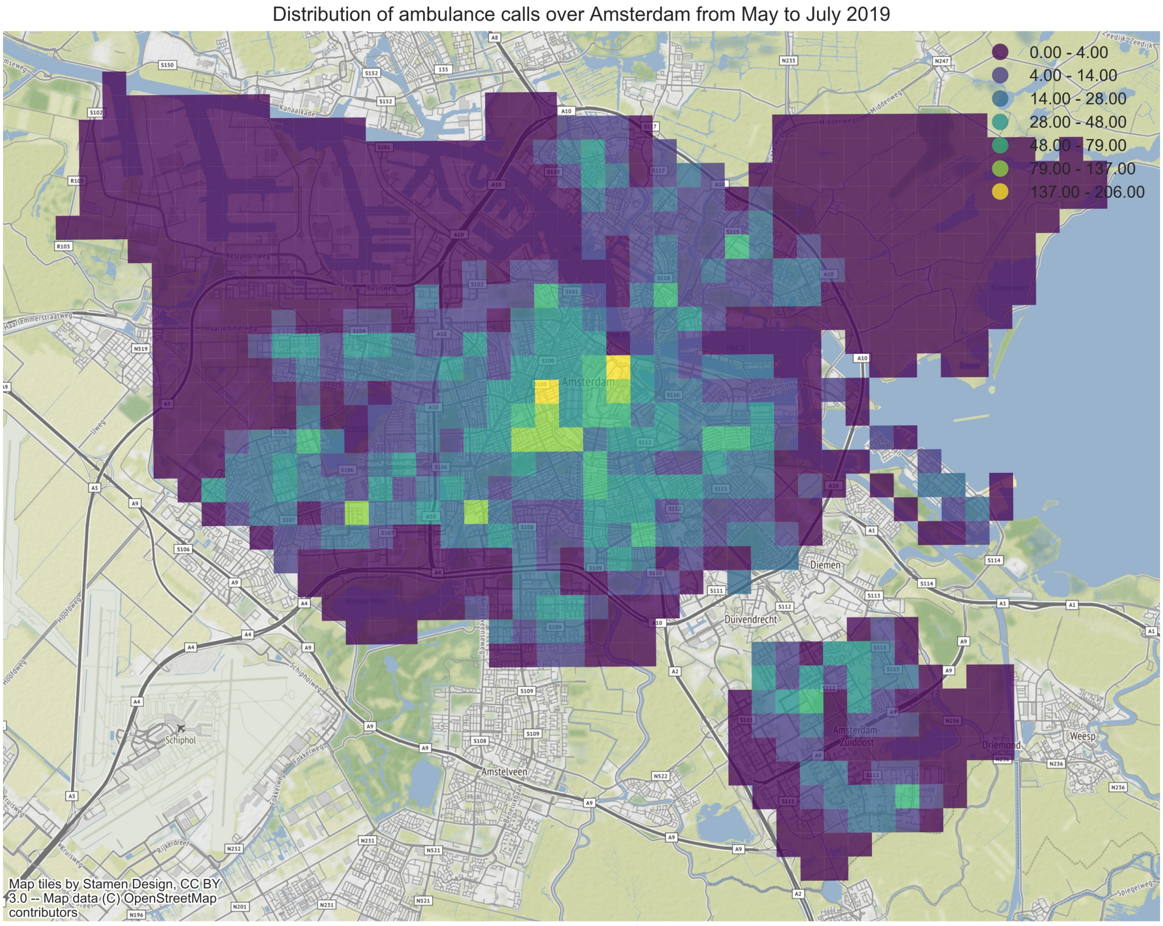

Figure A.6: Distribution of ambulance calls in Amsterdam

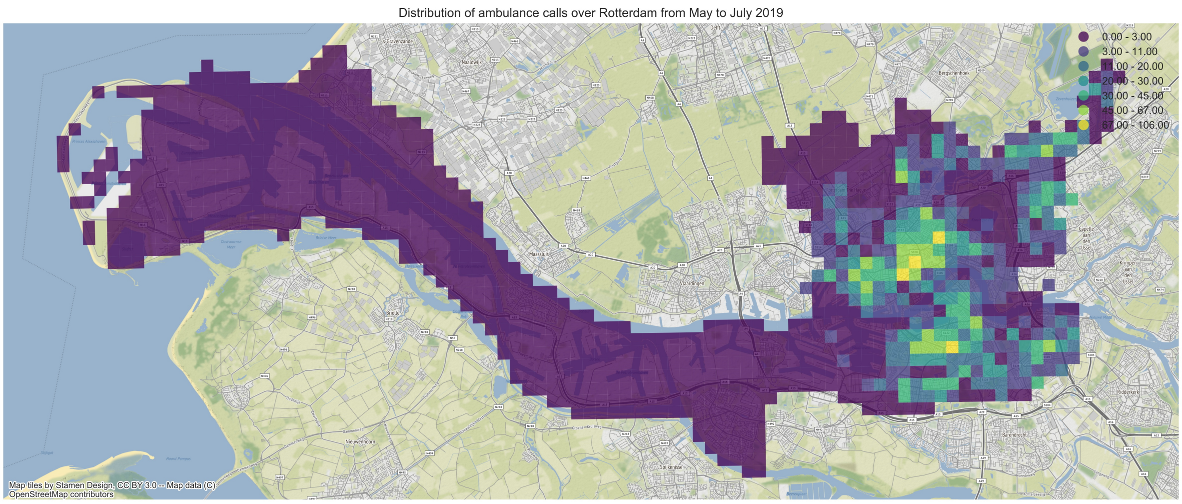

Figure A.7: Distribution of ambulance calls in Rotterdam

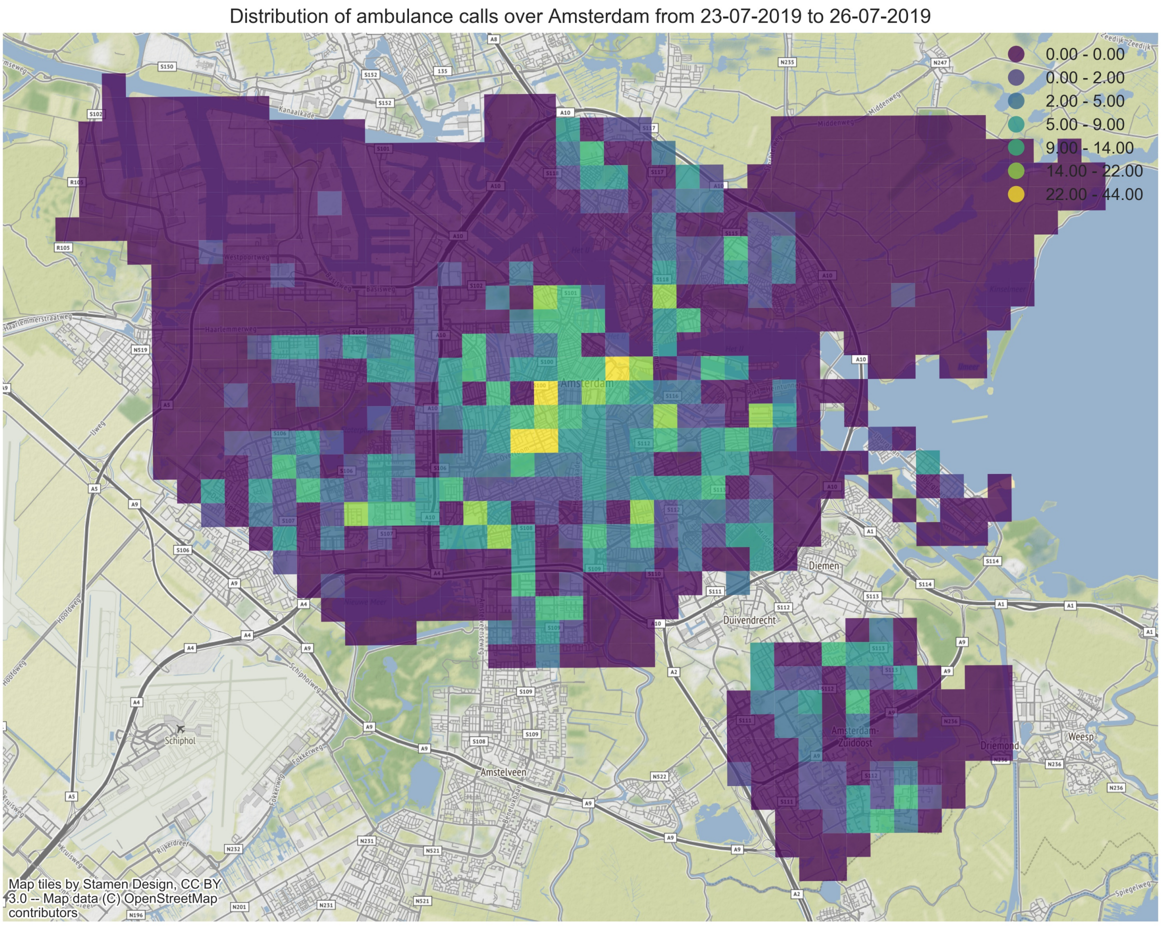

Figure A.8: Distribution of ambulance calls in Amsterdam during the July 2019 heat wave

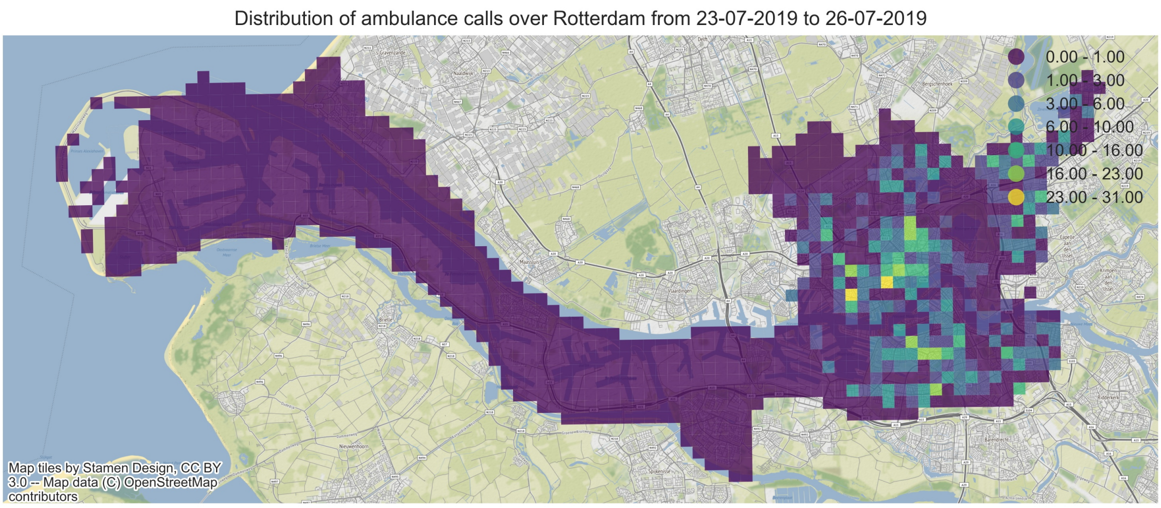

Figure A.9: Distribution of ambulance calls in Rotterdam during the July 2019 heat wave

B

Data description

This research highly relies on the data. Therefore it seems meaningful to describe additional details of the data sets used. The are several different data sources used in this study: the key data sets and corresponding variables described in Chapter 3. Here we present a more elaborate description of each data set. Note that all of the data set are open access and can be easily downloaded without any restriction. Besides, the raw, as well as processed data, is stored in the online repository. The most comprehensive data description can be found in the attached to each data set PDF files.

B.1. Maps with statistics

The two data sets *Kaart van 100 meter bij 100 meter met statistieken* and *Kaart van 500 meter bij 500 meter met statistieken* are similar in terms of variable names and their ranges. Therefore, let us provide the description only for one of them.

B.2. Den Haag in Cijfers

The online platform Den Haag in Cijers provides a lot data regarding population, businesses and infrastructure of the city of the Hague.

B.3. Other data sets

№	Translated name	Type	Values	Note	Original name
1	Number of residents	Real number			INWONER
2	Male	Real number			MAN
3	Female	Real number			VROUW
3	Individuals 0-14 years old	Real number			INW_014
4	Individuals 15-24 years old	Real number			INW_1524
5	Individuals 25-44 years old	Real number			INW_2544
6	Individuals 45-64 years old	Real number			INW_4564
7	Individuals over 65 years old	Real number			INW_65
8	Births	Real number			GEBOORTE
9	Native Dutch	Percent			P_AUTOCHT
10	Western foreigners	Percent		Europe, North America, Oceania and Japan	P_WALLOCH
11	Non-western foreigners	Percent		Africa, Latin America, Asia, Turkey	P_NWALLOCH

Table B.1: Kaart van 100 meter bij 100 meter met statistieken data set description

Category	Subcategory	Subsubcategory	Variable	Observation years
Population	Population characteristics		Age	1993-2019
			Gender	
			Ethnic group (CBS)	
			Marital status	
			Family composition	
	Population developments	Births	Number of births	1996-2017
		Deaths	Number of deaths	
		Total migration	Total arrivals	1995-2017
			Total departures	
		Migration within the Hague	Migration within municipality	2010-2017
		Migration outside the Hague	Total arrivals from outside the Hague	2010-2017
Economics	Income		Average personal income of individuals 52 weeks	2005-2016
			Average disposable income of private households	1994-2016
			Average standardized disposable income of private households	2000-2016
		Sources of income	4 classes	2011-2016
		Distribution of income	Low-average-high income	1994-2016
	Employment	Number of persons employed in (SBI 9 classes)	9 classes	2000-2018
		Labour force	Potential labour force	1993-2019
	Unemployment	Job seekers until 2014	Number of unemployed persons, total, Number of unemployed job-seekers	1995-2017
			Education, total, 5 classes	2006-2014
			Age, age, total, 9 classes	1996-2014
		Job seekers from 2015	Age, total 3 classes	2014-2017

Table B.2: Description of the data sets retrieved from Den Haag in Cijfers

Observations	9066
Total variables	19
Variables of interest	2
Source	Hoeven (2017)
Data Collection Period	2017
Data Reporting Period	2018
Date of Data Acquisition	25-05-2019

Table B.3: Short description of the *Heatmap The Hague* data set created by Hoeven (2017)

Observations	162
Total variables	10
Variables of interest	2
Source	The Dark Sky Company (2019)
Data Collection Period	2017
Data Reporting Period	2019
Date of Data Acquisition	01-04-2019

Table B.4: Short description of the The Dark Sky *Weather data*. The data collected from The Dark Sky Company (2019)

Observations	377600
Total variables	33
Variables of interest	21
Source	Centraal Bureau voor de Statistiek (2019b)
Date of Data Acquisition	25-05-2019 - 01-08-2019

Table B.5: Short description of *Map 100 by 100 square meters with statistics*. The data retrieved from Centraal Bureau voor de Statistiek (2019b)

Observations	159
Total variables	32
Variables of interest	7
Source	
Data Collection Period	2016
Data Reporting Period	2016
Last Modification	16-11-2018
Date of Data Acquisition	21-06-2019

Table B.6: Short description of the *"Health per district and neighborhood 2016"* data set

Observations	25040
Source	`https://p2000mobile.nl`
Data Collection Period	From 23-05-2019 to 31-07-2019

Table B.7: Description of the *"P2000 network"* data set

C

Reproducible research

The source code is available at `https://github.com/mikhailsirenko/urban-resilience-ml`. Moreover, almost all of the data sources can be downloaded via notebooks that are provided. Thus, a researcher can operate with raw data, run prepossessing scripts, and verify the results. For the data sources that were retrieved manually, a comprehensive description of the process is provided.