Vulnerability of cities to soil moisture and groundwater droughts

Ilias Machairas

TU Delft

Vulnerability of cities to soil moisture and groundwater droughts

by

Ilias Machairas

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on Wednesday October 28, 2020 at 10:00 AM.

Student number:	4801040	
Project duration:	January 1, 2020 – October 28, 2020	
Thesis committee:	Frans van de Ven	TU Delft, Deltares (supervisor)
	Markus Hrachowitz	TU Delft
	Nazli Yonca Aydin	TU Delft
	Marco Hoogvliet	Deltares

This thesis is confidential and cannot be made public until October 28, 2020.

Picture cover page: earth.org (2020)

TU Delft

Abstract

Due to climate change, extreme phenomena like droughts are going to be intensified. Even though droughts in agriculture have been studied, regarding urban environment their consequences are rather unexplored. Cities are susceptible to droughts and the estimation of their vulnerability is the first step for their protection. The objective of this thesis is to determine vulnerability of cities to groundwater and soil moisture droughts.

Since drought is a complex phenomenon which is difficult to define, its analysis is not straightforward especially in cities. For that reason, an urban drought categorization framework is created. In the current research, two of the four drought categories were studied (groundwater and soil moisture droughts). A case study for the vulnerability calculation was chosen for Leiden, the Netherlands. Before determining its vulnerability, different techniques to identify drought exposure characteristics were investigated. More specifically, regarding deficit and duration, the following methods were used: (i) fixed, (ii) variable, (iii) moving window method, and (iv) median groundwater level as threshold for the case of groundwater droughts. Spatial analysis was performed to estimate the areal extent of droughts whereas frequency distribution analysis is assessed of the minima of monthly blocks. A similar approach was applied for soil moisture droughts. Soil moisture is modeled via the lumped Urban Water Balance Model.

Vulnerability was estimated as the aggregation of exposure and sensitivity (physical and social). For both components of vulnerability, their indicators were normalized. Out of the four drought characteristics (deficit, duration, spatial extent, and frequency), only the two first were included as exposure indicators for the vulnerability estimation. The indicators' weights were computed using the Analytical Hierarchy Process (AHP). Vulnerability estimation was applied for both groundwater and soil moisture droughts separately.

One of the main results is that variable threshold performance is higher than fixed threshold and moving window for deficit and duration estimation regarding groundwater droughts. However, a combination of fixed and variable threshold can provide a profound insight into drought exposure. That applies to soil moisture droughts too. To increase performance of deficit and duration identification, pooling can be applied. An inter-event time of around 10 days for the case of variable threshold and 30^{th} percentile is suggested regarding groundwater droughts. The analysis on soil moisture droughts was not sufficient to draw conclusions regarding inter-event time.

Vulnerability follows sensitivity patterns - rather than exposure pattern - for both groundwater and soil moisture droughts at both fine and coarse space resolution. Besides, out of all indicators used, those which contribute to vulnerability variation the most were determined for both studied types of droughts. Those indicators are: 'land use' and 'percentage of households belonging to the lowest 40% income nationwide' for groundwater droughts whereas for soil moisture ones, it is 'green areas'. Another result is that the differences in vulnerability vary marginally using different techniques to identify drought events for both studied types of droughts. Therefore, even applying a sophisticated technique to identify drought events, will not lead to significant improvement regarding vulnerability estimation. All aforementioned conclusions regarding vulnerability are highly uncertain since drought experts assigned different indicators' weights; the convergence of opinions was low.

Based on the proposed methodology, water managers would be able to determine vulnerability of cities to droughts and policy makers would be able to protect the regions which are highly vulnerable. Consequently, the adverse impacts of droughts on cities could be mitigated, reducing residents' hardship.

Preface

This thesis is the last part of my master in Civil Engineering, Water Management at the TU Delft. Studying at TU Delft was a unique experience and helped me to open my horizons developing new skills.

I would like to thank all members of my thesis committee. More specifically, Frans van de Ven for his guidance and support throughout my thesis especially when I faced hardships. I would also like to express my gratitude to Marco Hoogvliet for sharing his insights and experience for water projects in the Netherlands. Our meetings were helpful and his contribution was valuable. Thanks to Markus Hrachowitz for his critical comments about hydrological modeling, and to Nazli Yonca Aydin for her feedback helping me not to ignore the social aspect of my thesis. I would also like to thank Toine Vergroesen, creator of the Urban Water Balance Model, for his technical support.

Finally, I would like to thank my family and friends for their emotional support throughout my studies.

Ilias Machairas
Delft, October, 2020

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter highlights how and to what degree droughts and specifically urban droughts influence societies' well-being. Sections 1.1 and 1.2 describe the impacts of droughts in general and on water quality respectively. Section 1.3 provides examples of the consequences of droughts on cities whereas Section 1.4 highlights the need for an urban drought framework based on which urban droughts can be classified. The main objective of the thesis is presented in Section 1.5.

1.1 Impacts of droughts

Climate change leads to extreme weather conditions. Droughts are going to be intensified for both scenarios of global temperature increase of 1.5 °C and 2°C (the intensification is higher for the scenario of 2°C increase) (Intergovermental Panel on Climate Change, 2018). Droughts may occur in humid and dry climate conditions and are sometimes confused with aridity which is a characteristic of a region and is based on the difference of precipitation and potential evaporation. They occur when water input is lower than the normal one regardless of the aridity.

The impacts of droughts do not rise abruptly; thus, not much attention is given by stakeholders until the point at which the consequences are severe (Wilhite and Vanyarkho, 2000). As a result, the estimation of the damage is difficult (structural losses are not common either) and may be lower than the actual one (Wilhite, 1994). In the USA, droughts are the second most damaging events after tropical cyclones and the total damage for the period of 1980-2019 amounts to $249.1 billion having been adjusted based on the consumer price index (NOAA, National Centers Environmental Information (NCEI), 2019). In the European Union, the total damage cost due to droughts for the period from 1977 to 2007 amounted to 100 billion € and it is estimated that the annual cost is 6.2 billion € (DG Environment European Commission and others, 2007). In developing countries, the consequences are severe as well since the crop production decreases considerably influencing millions of people. For instance, in eastern India, considering the damage and the mitigation measures for farmers, the total annual cost amounts to $500 million (Pandey et al., 2000). All of the aforementioned estimations contain assumptions; therefore, the cost may have been underestimated. Moreover, most of these cost estimates are focusing on agricultural production losses and do not include damage in the urban environment.

The consequences of droughts vary and can be categorized into economic, environmental, and social impacts (Wilhite and Vanyarkho, 2000). Some examples of economic impact are: (i) the reduction of crop yield which leads to less sales, (ii) the loss of livestock production, and (iii) the increased cost for transport since the depth of streams is lower,; therefore the ships cannot travel with a lot of load. During droughts, there is not only shortage of drinking water but also other critical resources are limited, such as electricity. For instance, in Rio de Janeiro, in 2001, due to a severe drought the

hydroelectric power stations' performance reduced considerably and restrictions on residents' energy consumption by 20-25% were applied (De Sherbinin et al., 2007).

Regarding environmental impact, species of fauna and flora may face problems and get extinct, so the biodiversity may get reduced. Additionally, the increased salinity intrusion and the higher possibility of fires are two important consequences (Wilhite and Vanyarkho, 2000). For instance, the 2010-2011 drought in Texas led to fires in three counties resulting to the destruction of 21,000 acres (almost 85 km^2) and burning of 66 million trees (Schmidt and Garland, 2012). As for the changes in flora, they are dramatic for many ecosystems, such as wetlands, rivers and lakes. In wetlands where not many changes in the water level occur during droughts, the variety of macrophytes (aquatic plants) increases but in case the frequency of drought periods is high, only a minority of species is able to survive and most species become extinct (Greening and Gerritsen, 1987). Reduction of macrophytes has been found in river Kennet, England during the dry period 1996/1997 when the local ecosystem has been disturbed due to lower discharge (Wright et al., 2002). In shallow lakes, the winter droughts created a shift in the dominant macrophyte species and the reason was the higher NH_4-N based on the study of McGowan et al. (2005) in Saskatchewan, Canada.

As for social aspect, conflicts among water users tend to occur since the impact of drought may be unequal. Social riots, political changes and even civil war and armed conflicts between countries can take place due to droughts such as occurred in Syria in 2011 (Gleick, 2014) and in Morocco in 1980s (Swearingen, 1992). A further consequence of droughts is the migration of large populations from rural to urban area due to droughts; the problem is relocated to a city nearby which is not often able to handle extra inhabitants. Therefore, the pressure for resources increases. For instance, in the South-East part of Brazil, large migration has occurred in many periods in the past during droughts, since the economy of this region was based on rain-fed agriculture (Adamo, 2010) and the same happened in Africa (Henderson et al., 2017).

1.2 Influence of droughts on water quality

In addition, droughts (mostly the hydrological ones) have an influence on the water bodies' quality changing a host of parameters. The reduced number of rainfalls in Otago, New Zealand during the drought which occurred in 1998-1999 reduced the concentration of nutrients which derived from agricultural fields since the run-off was less (Caruso, 2002). On the other hand, in Meuse river, higher nitrite (NO_2^-) concentration has been observed and the possible explanation has been the reduced dilution (Van Vliet and Zwolsman, 2008). On South Platte river, in Colorado nitrate (NO_3^-) and NO_2^- concentration were different for each land use where the sample was taken (Sprague, 2005); thus, general conclusions and patterns cannot be drawn.

Temperature is another important parameter of water quality and there is not absolute consistency about its change during droughts. Van Vliet and Zwolsman (2008) found a mediocre increase of 2°C on Meuse river whereas Ha et al. (1999) spotted an increase of around 7°C on Nakdong River, South Korea. On the other hand Wilbers et al. (2009) and Mosley et al. (2012) noticed that the water temperature remained almost constant on the Dommel River, the Netherlands and on river Murray, Australia accordingly since the majority of Dommel river stems from groundwater sources (less influence from hydrological droughts), and the fact that the air temperature near river Murray did not increase respectively.

During a drought, salinity increases whereas dissolved oxygen decreases. As for salinity, on Murray river and lakes in Australia, it increased when the flows became too low since salt intrusion was more dominant (Mosley et al., 2012). The outflow from lakes was reduced and at the same time evaporation was higher during droughts leading to higher salt concentration. Dissolved oxygen is a key parameter for water quality and has been reduced in many studies, such as on Pearl river, China (Dai et al., 2006), and Rhine (Zwolsman and Van Bokhoven, 2007); the reasons have been less dilution and higher

temperature which has been negative correlated to dissolved oxygen.

Droughts also influence the ecology of the water body and its surrounding area. Regarding algae, there is contradictory research since García-Prieto et al. (2012) and Van Vliet and Zwolsman (2008) found an increase of them during droughts in the province of Salamanca, Spain and on Meuse river respectively but according to Boar et al. (1995) there was no increase of algae production at the 1989-1992 drought on river Nar in England. Some examples of droughts' impact to fauna are: the population of invertebrates decreased sharply in the Thames catchment during the drought period of 1989-1992 (Attrill and Power, 2000), and the total fish population became extinct in 1953-1954 on Smiths branch, Illinois (Larimore et al., 1959). Droughts may create ideal conditions for eutriphication which can be devastating for fish predators due to lack of oxygen early in the morning. Based on the aforementioned research, it is not clear how droughts influence water bodies' quality but the majority of studies revealed adverse consequences on many water parameters indicating that droughts' effects ought to be mitigated.

1.3 Droughts and urban environment

The consequences of droughts are devastating not only for rural environment (e.g. wilting of crops) but also for urban one considering that about 66% of the global population is predicted to live in cities by 2050 (United Nations, 2014). This will lead to more pressure in all kind of resources and to increasing the sensitivity of the cities. The vulnerability of cities to droughts will be higher because of urbanization even without considering the climate change effects (Güneralp et al., 2015). The impact of droughts on cities has not be studied intensively (Zhang et al., 2019) despite the fact that many vulnerable assets are clustered there. It is common that only mega cities are taken into account regarding vulnerability; cities of medium size whose population is ever increasing are ignored (Birkmann et al., 2016).

Many problems can be created by droughts in urban environment and the most important ones are mentioned below. During drought period, the amount of working hours which one employee can work decreases considerably and the unemployment rate increases as well (Desbureaux and Rodella, 2019). For instance, in Hong Kong during the 1963 drought, dyeing industries faced severe problems (the production was reduced by 35-40%) and as a result the laboring class suffered from salary reduction and unemployment (Chen and Chan, 2007). Furthermore, the spread of contagious diseases is larger during droughts especially in cities in developing countries where a proper sewage system is absent or malfunctions and as a result the productivity is reduced (drop of 6.5% in cities in Latin America) (Kovats et al., 2003). Furthermore, droughts can decrease the groundwater level (GWL) and subsidence may occur in lowland areas causing damages not only on buildings on piles but also other buildings (Ven et al., 2011). A major problem however is the deterioration of wooden pile foundations.

Wilting of the trees in urban parks due to low soil moisture is another adverse impact of drought in cities but it is strongly dependent on the species of trees (Ven et al., 2011); the reduced evaporation leads to higher temperature and an enhanced urban heat island effect. Another consequence of droughts is that fires which occurred due to droughts could cause reduction of the pressure in the drinking water network, assuming that some amount of water is extracted from the network in order to extinguish the fire; as a result, back-flow could take place which is quite dangerous regarding the spread of illnesses (Knutson et al., 1998). Water quality degradation in the canals of a city is also possible due to droughts if the city has any. The adverse impacts on water characteristics which were mentioned in Section 1.2 hold in urban environment too. The last example is about Cape Town which was the first large city globally which faced up with a serious water crisis due to a drought in 2015-2017 (which was considered as one of the most severe since 1993)(Wolski, 2018); the timely taken actions reduced drastically the consumption of water and taps were not switched off.

For the reasons mentioned above, cities should be protected from droughts. Countermeasures ought

to be taken making sure that the water supply is enough (or sufficient volumes of water are stored in all parts of the urban water cycle, such as unsaturated zone) given that the demand is going to increase in the future.

1.4 Need for an Urban Drought Framework

A universal definition of droughts is impossible since (i) there are other components besides precipitation which play an important role to a water system, and (ii) different stakeholders are influenced under different circumstances which are not possible to be included in one all-inclusive definition (Lloyd-Hughes, 2014; Heim Jr, 2002). The simplest definition has been given by Sheffield and Wood (2012)[p.11] which is "a deficit of water relative to normal conditions". The inability of providing a clear definition for droughts leads to ineffective research about this phenomenon (Yevjevich, 1967).

There is confusion regarding urban drought definition and urban drought groups since (i) different parts of the water cycle are influenced by urban droughts, and (ii) the fact that urban droughts vary over the world. Based on Wilhite and Glantz (1985), droughts can be categorized into meteorological, agricultural, hydrological, and socio-economic ones; each category focuses on a different component of the water cycle. Urban drought can be considered as a type of socio-economic drought where the demand is higher than the supply for a period of time (not permanently) (Zhang et al., 2019). However, Rossi et al. (1992) considered it not as a sub-category of the socio-economic drought but as a distinct category.

It is also important to highlight the differences between urban drought and water shortage since there is confusion about them. The former one is a temporary water stress whereas the latter one is a permanent status of water resources lack; urban droughts can cause a water scarcity and not the other way around (Zhang et al., 2019). The focus of urban droughts is on change whereas for water scarcity is on status. In order to distinguish drought from water scarcity, estimation of the naturalized situation is required using hydrological modeling (van Loon, 2013) but this is not part of the current study.

Regarding urban drought categorization, it is a difficult task but can be very beneficial so that policy makers can take relevant measures in time and protect valuable assets in a city effectively. Zhang et al. (2019) classified urban droughts based on the causes into the following categories: precipitation-induced, run-off induced, pollution-induced, and demand-induced but there are no steps which policy makers can follow to test which type of drought their city faces. Thus, there is a need for a comprehensive framework for urban drought categorization to overcome this problem.

1.5 Research Objective and Research Questions

The main goals of this thesis report is: (i) to define urban droughts and to create a framework based on which policy makers of city can classify the drought which their city faces, (ii) to analyze the vulnerability to urban droughts specifically regarding soil moisture and groundwater, and (iii) to analyze the features of the aforementioned droughts. The main research questions are:

- **How to define an urban drought?**
  An urban drought framework is required so that policy makers can identify if an urban drought affects their city and what its category is. What is the process which needs to be followed for the categorization?
- **How can an urban drought be quantified? How can the inherent characteristics of groundwater and soil moisture urban drought be analyzed?**
  Those characteristics are deficit, duration, areal extent, and frequency. There is no clear methodology and pattern which one researcher can follow in order to estimate those for an urban drought.

- **How can the vulnerability of cities to soil moisture and groundwater droughts be determined?**
  This question is about cities in in developed countries. There is a host of vulnerability frameworks available which are helpful but which of them fits well for the current study? Which indicators reveal that a specific part of a city is vulnerable to soil moisture and groundwater droughts?

The reminder of this thesis is as follows: Chapter 2 describes relevant previous studies showing the approaches and types of indicators which have been used for vulnerability assessment. In that way, the scarcity of drought vulnerability analyses on cities is revealed. Chapter 3 describes the methodology used highlighting the framework for urban drought categorization. The results of groundwater and soil moisture droughts' analysis are covered in Chapter 4. The former were based on measurements whereas the latter on modeling given the lack of data regarding soil moisture. In Chapter 5, the main points of interest, and limitations of the applied method are discussed. The last chapter (6) of this thesis is the Conclusion and Recommendations where: (i) the summary of this thesis is presented highlighting what this thesis offered to current knowledge, and (ii) recommendations for further research are suggested.

Chapter 2

Literature Review

Section 2.1 describes six vulnerability frameworks analyzing their main components. Subsequently, Section 2.2 emphasizes on droughts vulnerability specifically presenting methods which have already been used. Drought Risk is described shortly in Section 2.3 whereas the characteristics of a drought (deficit, duration, frequency, and spatial extent) are analyzed in Section 2.4. Finally, previous research for droughts regrading soil moisture, and groundwater are mentioned in Sections 2.5, and 2.6 respectively. It is worth noting that a summary is provided at the end of each section.

2.1 Vulnerability Frameworks

According to Intergovermental Panel on Climate Change (2007)[p.27] "Vulnerability is the degree to which a system is susceptible to, and unable to cope with, adverse effects of climate change, including climate variability and extremes". This is one out of many vulnerability definitions and it is mainly focused on climate change. In order to assess vulnerability, there is a host of available vulnerability frameworks which emphasize on different aspects of a hazard and consist of different components. There are advantages and and disadvantages for each vulnerability framework; therefore, policy makers need to to consider them to select the most suitable framework. Six vulnerability frameworks are shortly presented below.

2.1.1 First Framework

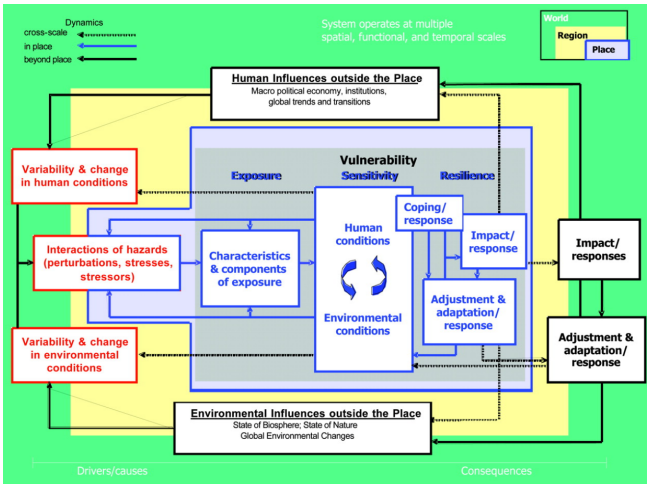


Figure 2.1: First Vulnerability Framework with its components and their interrelationships, Source: (Turner et al., 2003)

Turner et al. (2003) created a vulnerability framework for sustainability which was based on (i) variability in human and environmental conditions, and (ii) the link between humans and environment (see Figure 2.1). They employed three spatial scales: globe, region, and place but most interactions occur between region and place. The hazards can be derived from inside the human-environment system or even outside of it.

There are three components: exposure, sensitivity and resilience; the first two elements are related to human-environment conditions whereas the third one includes all kinds of response of the coupled human-environment system. Adaptive capacity shows the flexibility of a system to maintain its normal conditions with some acceptable deviation given that slow changes over long time occur such as climate change, and urbanization (Turner et al., 2003). This term stresses how easily the system can be adapted to the new situation and the new requirements of society.

2.1.2 Second Framework

Another framework for vulnerability was suggested by Bohle et al. (1994) and its main components are: risk exposure, coping capacity, and recovery potential (potentiality) as can be seen in Figure 2.2. There are many terms in this framework which are not straightforward but are analyzed below to avoid confusion. Human ecology represents the interaction between nature and society; expanded entitlements show the entitlements which a group of people need to own for welfare; and political economy is straightforward. Regarding the inner part of the figure, endowments refer to the groups or elements which may be unable to cope with the climate variability due to physical reasons e.g. geographic location. Class relations and empowerments represent the social relationships which can withstand the change and the formation of rights respectively. Political ecology shows how the resources are managed and contribute to potentiality (the degree to which assets can recover from the variability). This framework was used for social vulnerability in developing countries and is not that suitable for other types of research.

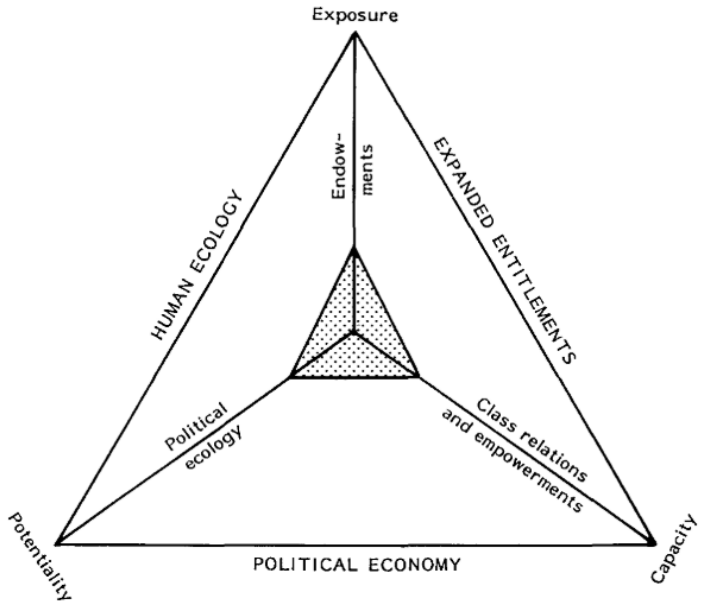


Figure 2.2: Second Vulnerability Framework with its three main components, Source: (Bohle et al., 1994)

2.1.3 Third Framework

Another framework is presented by Fritzsche et al. (2014) highlighting all the steps from planning to implementation for a vulnerability assessment. Exposure shows the climate variability. Weather

variables are used to quantify this change. Exposure underlines how often a damaging event occurs and what its severity is. Sensitivity is another component which shows how the system changes due to exposure (when specific events happen) based on the inherent features of the system. Human intervention may also have led to changes of the physical characteristics of the system. For instance, the location of critical infrastructure (e.g. hospitals, administration and electricity buildings) or the materials used in buildings can be sensitivity factors of a city to a natural hazard. The higher the sensitivity, the larger the possibility for the system to change due to increased exposure. The combination of exposure and sensitivity leads to potential impact.

Adaptive capacity is also part of this vulnerability framework and shows the flexibility of a system to adapt to new conditions and to mitigate damage or even take advantage of the conditions if it is possible. It is not only about withstanding the impact of the event but it also includes the ability of the system to recover after the event. The right governance in conjunction with the right policy and institutional environment play a key role for adaptation. Not only is the effective decision-making required but also the implementation of the decisions. Exposure and sensitivity have a positive relationship with vulnerability whereas adaptive capacity has a negative one. In this framework, adaptive capacity can be categorized into knowledge (awareness and education), technology (means which can be used to reduce the impact), institutions, and economy. Figure 2.3 illustrates the basic components of this vulnerability framework. Another term used in that framework is resilience which has a negative relationship with vulnerability but it is considered as not just the opposite of vulnerability. To quantify these components, suitable indicators (variables) are used.

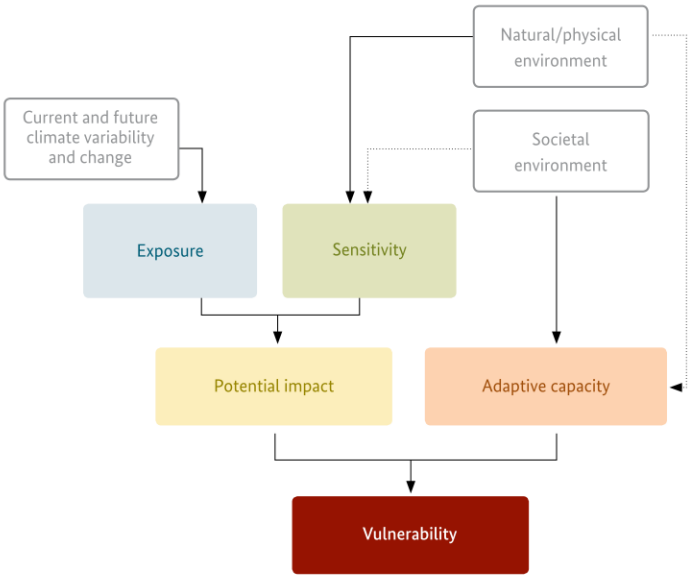


Figure 2.3: Third Vulnerability Framework with its main components, Source: (Fritzsche et al., 2014)

Space and time are two components which need to be decided at an early stage of vulnerability analysis and need to be in line with the project main objectives. The spatial scale of vulnerability determines the stakeholders and institutions from which one researcher needs to extract data. The geographical scope can vary from a local community to the national level; it is based on the goal of vulnerability analysis. As for temporal dimension, vulnerability assessment can be conducted for the current situation based on past events (a duration of 30 years ago is considered satisfactory) but also for the prospective status (assessment for a long future period needs to be avoided due to high uncertainties) (Fritzsche et al., 2014).

In that framework, the steps for indicators' determination in order to describe exposure, sensitivity,

and adaptive capacity are:

- selection of spatial and temporal scale
- determining the indicators and the their sources
- checking if these indicators are available to the suitable format
- finding alternative sources of the indicators in case the previous step was not successful
- making adjustments on the spatial and temporal scale in case the two previous steps failed

This process is iterative if suitable data for an indicator cannot be identified. Indicators can be derived from: (i) measurements, censuses and surveys (mainly for sensitivity and adaptive capacity components), (ii) modeling (mainly for exposure component), and (iii) expert judgment.

2.1.4 MOVE Framework

As for the MOVE framework (Birkmann et al., 2013), the four components of vulnerability are: exposure, susceptibility, societal response capacity or lack of resilience, and hazard. The main dimensions of vulnerability are: physical, social, ecological, economic, cultural, and institutional (see Figure 2.4). Exposure in this framework includes the systems (infrastructure and humans) which can be influenced due to a hazard event. Susceptibility shows the predisposition of the system to be disrupted and is not influenced by exposure; the term fragility is considered identical with susceptibility. Lack of resilience or societal response capacity refers to the inefficient resources and provision thereof which could be used to tackle the hazard. Lack of resilience can be classified into three phases (before, during, and after the hazard event).

Adaptation includes learning from damaging events which leads to (operational and institutional) changes so that exposed assets can increase their protections for future hazard events. When successful adaptation is implemented, resilience is activated and less coping is required. An important difference between coping capacity and adaptation is that adaptation does not include direct response to the hazard. Hazard is also used in the framework and shows the possibility of damaging events (derived from nature, or society, or humans) to influence exposed elements.

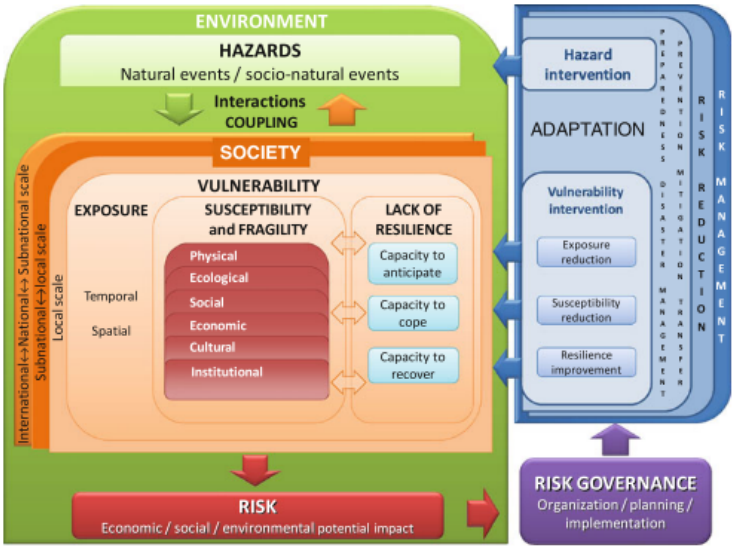


Figure 2.4: MOVE Framework with its main components, Source: (Birkmann et al., 2013)

This framework provides the basic concepts of vulnerability but no methods to quantify the concepts described. It stresses: (i) the coupling between environment and society, and (ii) the fact that spatial and temporal scale influence the vulnerability analysis since vulnerability indicators tend to be

applicable only to specific scales and cannot be used at different ones (Birkmann et al., 2013). One downside is that factors to describe cultural and institutional dimensions of vulnerability are difficult to be found since these dimensions are not that concrete compared to the others used in this framework (Birkmann et al., 2013).

2.1.5 Eight step approach

The main objective of this framework, developed by Schröter et al. (2005), is to provide critical information to stakeholders regarding adaptation measures. It is focused on climate change and comprises eight steps (the first three ones do not include modeling but they do the last five ones). These steps are:

- Study Area Definition
  Both researchers and stakeholders need to select the study area after a thorough discussion regarding financial and time limits. The study scale should be similar as that where policy makers can make decisions but smaller or larger scales may be necessary to complement the vulnerability analysis.

- Familiarizing with study area
  The main goal of this step is for the researcher to identify vulnerability drivers. Not straightforward information (such as how the system works, what the values of the community are) are not easily recognized but they can add value to the analysis.

- Hypothesizing vulnerability
  Preliminary hypothesis can be formulated after discussion with stakeholders. Not having a clear focus is a common mistake in that phase trying to increase the spectrum of the analysis. Time and budget also need to be considered in that step since they influence the focus scale.

- Developing a causal model
  Identifying the causes of vulnerability is the goal of this step and the model can link causes and consequences. Diagrams can help in better representation of the model.

- Finding relevant indicators
  The indicators are region-specific and before their selection it needs to be sure that they fit the context and are understood by stakeholders.

- Model operationalization
  The causal model uses the indicators as input data and produces vulnerability in conjunction with assigned weights. The model needs to be able to handle time series data, tackle systems with high complexity, and apply computer-based methods.

- Projection of future vulnerability
  Projections which consider the variability of the driving forces in full range need to be examined. Even though validation of the model cannot often be applied since there are no observations, the consequences of past events can be employed to show how reasonable the model is.

- Communicate vulnerability
  Stakeholders may not be able to conceive the uncertainties of the model; therefore, cooperation with researchers is required. Active participation of stakeholders can ease the process. This step needs to be applied to all the previous steps and not only at the end of the analysis.

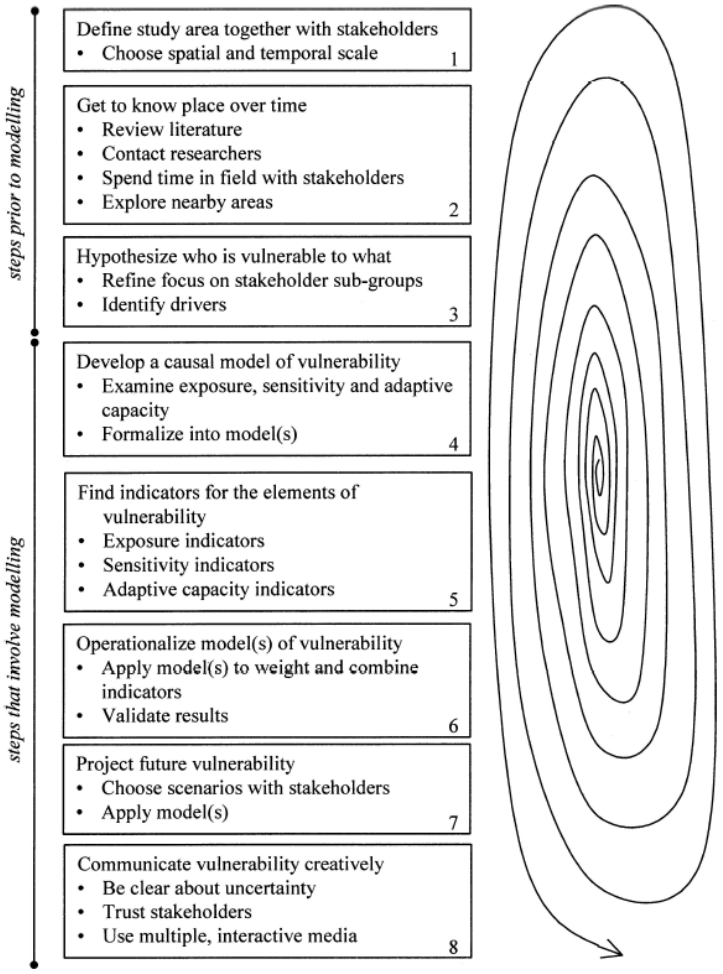


Figure 2.5: Eight Step vulnerbility assessment, Source: (Schröter et al., 2005)

2.1.6 Vulnerability Scoping Diagram

A vulnerability framework able to be used under any different context could be helpful but more research is required in that scope. Lack of resources and time limitations may hinder the implementation of vulnerability analysis. In areas where vulnerability assessment is not possible to be performed, conclusions from vulnerability assessment in other case studies can be useful; this knowledge can lead to adaptation actions which work to some degree compared to inaction (Polsky et al., 2007). In order to cover this gap, Polsky et al. (2007) introduced the Vulnerability Scoping Diagram (VSD); a generalized framework for vulnerability of different threats with different indicators so that comparison among different hazards and areas can be conducted. Firstly, hazard and exposure units need to be selected which are: (i) the driver that disrupts the human-environment system, (ii) the exposed groups and assets respectively. Subsequently, VSD can be created which consists of three rings: (i) dimensions of vulnerability (exposure, sensitivity, and adaptive capacity), (ii) components of dimensions (features of the human-environment system which may be affected), (iii) and measures of components. Measures of components are (qualitative and quantitative) indicators. Groups of them create the components of dimensions.

Summary regarding vulnerability frameworks

- In the first framework, the focus in on the interaction between human and nature. There are three spatial scales (globe, region, and place).
- The second framework is mostly applicable for social vulnerability; its components are quite

abstract.

- The third framework consists of three components: exposure, sensitivity, and adaptive capacity. A detailed guidance for its implementation is provided.
- Based on MOVE Framework, vulnerability consists of four components: exposure, susceptibility, lack of resilience, and hazard. No methods for their quantification are provided. It is worth noting that exposure and adaptive capacity are defined differently than the previous framework.
- The eight-step approach describes thoroughly how vulnerability can be estimated. A causal model which can handle complex systems is required. This framework also highlights the cooperation between modelers and stakeholders. It has been neglected in the other frameworks.
- Vulnerability Scoping Diagram is an approach to determine vulnerability under any different context. It seems a promising method but more research is required.

2.2 Drought Vulnerability

The aforementioned frameworks are useful for a variety of issues including droughts but are quite generic. On the contrary, this section is more specific reviewing previous studies specifically on drought vulnerability.

The vulnerability of city to droughts has not been thoroughly studied (Zhang et al., 2019). Due to this lack, the vulnerability to droughts in general and in rural environment has been studied in this Section since it can provide valuable information about the methods used even though urban and rural areas are different. The main difference between drought vulnerability in the rural and urban environment is that other indicators are used. Some indicators used in agricultural droughts may be useful for cities in developing countries but they do not probably provide any information for cities in developed countries because none of the elements (e.g. residents, and households) meet the criteria, such as illiteracy rate and percentage of households without electricity.

Only a few studies have focused on droughts vulnerability in cities and are presented below. Chang et al. (2015); Wang et al. (2019) have researched the vulnerability of cities in China employing the entropy evaluation approach (Iyengar and Sudarshan, 1982) for the estimation of weights, taking advantage of a variety of variables regarding economy, society, ecology, and social resources. Vulnerability of 65 cities have also been studied in China by Yuan et al. (2015) using variables based on exposure, sensitivity, and adaptive capacity. The authors estimated the weights using Analytical Hierarchy Process (AHP) which is based on expert knowledge. AHP has been used in a variety of scientific fields even for seismic (Panahi et al., 2014) and landslide (Ahmed, 2015) vulnerability. Another worthwhile study has been performed by Corti et al. (2009, 2011) who have investigated the damage of buildings due to subsidence based on insurance inventories having created vulnerability curves which converted the soil moisture deficit to monetary damage.

From this paragraph and below, vulnerability to droughts in general and in rural environment are analyzed. Hoque et al. (2020) have employed an integrated method taking advantage of various criteria calculating the vulnerability for each one out of the four types of droughts (meteorological, hydrological, agricultural, and socio-economic). The weights have been estimated using AHP and the combination with the four previous vulnerability maps has led to the final output which has been assessed using validation with the Relative Operating Characteristic (ROC) method.

Dumitraşcu et al. (2018) determined the drought vulnerability in southern Romania splitting it into three indices: the first one regarding demographics and society, the second one for economy, and the last one for water-related infrastructures. For each of these categories, a host of variables were available but only some of them were selected based on expert knowledge; these variables were normalized before the aggregation for the estimation of the three indices (Dumitraşcu et al., 2018). The final index of vulnerability was estimated using weighted average but it was basically just the arithmetic mean since

the selected weights were equal. The study was focused on agriculture which means that most of the variables used are unsuitable for an analysis in cities but there are some exceptions which are applicable to both cases. These are: from the first index: population over 65 years old or younger than 5 years old, and population of a minority; from the second one: unemployment rate as a proxy of economic status; and from the third one: the total amount of delivered water.

Thomas et al. (2016) developed two separated maps for vulnerability to droughts in central India (the first one based on variables which were constant time-wise (in the short term) such as land use and soil type, and the second one with variables which varied temporally such as groundwater and soil moisture availability). The final vulnerability map was created combining the maps mentioned above and in that way it provided an integrated assessment of how vulnerable the area was. However, they did not use a specific method for assigning the weights and the latter ones were based on their critical judgment which means that there was some subjectivity.

Another interesting study was performed by Gurwin (2014) who investigated groundwater drought vulnerability for a catchment in Poland using Geographical Information Systems (GIS). More specifically, weighted overlay based on groundwater monitoring wells and Kriging interpolation were applied. The number of monitoring wells at which the GWL was below a specific threshold was one of the used parameters for drought vulnerability determination. Data regarding land use and (recommended and actual) groundwater exploitation were used for the weighted GIS overlay.

Social aspect of vulnerability is often neglected and all the attention is focused on the physical components (e.g. geology) as Gleick (2014) did in Iran. Some useful variables for the determination of the social vulnerability index to all natural hazards for the counties of the United States have been the following ones: personal wealth, race and ethnicity, the occupation of residents, social status, education, medical services density, and transport infrastructure (Cutter et al., 2003). These variables can be used for drought vulnerability pinpointing areas where residents are not able to handle the impacts of droughts and recover from them. However, some variables may not contribute that much compared to other ones and are redundant; thus, multicollinearity tests are needed to reduce the number of variables.

As mentioned in Section 2.1, there is not one universal vulnerability framework since in each discipline researchers analyzed vulnerability from a different perspective focusing on different aspects. As a result, comparison among studies for which different frameworks have been used are unfeasible. In order to overcome it, Alcamo et al. (2008) introduced an inference modeling approach and applied to three case studies which faced droughts crises. The analysis was qualitative at the beginning recognizing the relationships between vulnerability factors and their impact, but at the following phase it became quantitative collecting data for indicators from high levels to low ones and the other way around. The authors also validated their results with media articles and interviews but these sources are not trustworthy even though they used them as binary variables.

For the studies mentioned above, mostly vector layers (polygons) have been employed and analyzed with the help of GIS. Only a few studies have used raster images. The reason is that vector layers are capable of analyzing vulnerability over spatial units such as counties and provinces. Most data of indicators have been collected in that level by the responsible authorities in different countries; therefore, it has been more convenient for researchers to perform their analysis in that level and not at a finer spatial scale.

2.3 Drought Risk

Besides drought vulnerability, drought risk was calculated in many studies (Kim et al., 2015; Nasrollahi et al., 2018; Rahman and Lateh, 2016; Shahid and Behrawan, 2008; Zhang et al., 2015). However, in their analysis, vulnerability was mostly related to socioeconomic factors of the study area. In addition,

the term hazard was used which includes exposure and considers the occurrence frequency of drought events for different classes of intensity. It is worth noting that the approach described in this Section is not in line with the vulnerability frameworks in Sections 2.1.1 and 2.1.3 since what is described in this approach as vulnerability for them is sensitivity.

Drought risk was estimated based on equation (2.1), where DRI: is the Drought Risk Index, DHI: is the Drought Hazard Index, and DVI: is the Drought Vulnerability Index (Wilhite, 2000).

$$DRI = DHI * DVI \tag{2.1}$$

If the DHI is zero, then the DRI is zero regardless of the vulnerability value. DHI is based on drought characteristics (deficit and occurrence frequency). In order to estimate it, Rahman and Lateh (2016) combined the three categories of drought deficit (moderate, severe, and extreme) based on Standardized Precipitation Index (SPI) as can be seen in formula (2.2), where DHI_i: is the Drought Hazard Index for i station, O_m, O_s, O_e: are the percentages of occurrence for moderate, severe, and extreme droughts respectively whereas W_m, W_s, W_e: are their weights accordingly based on AHP. The occurrence frequency of a specific drought category was the ratio of the total drought duration of this drought category to the total time series for a precipitation station. Kim et al. (2015); Nasrollahi et al. (2018); Shahid and Behrawan (2008); Zhang et al. (2015) calculated the DRI based on drought deficit using SPI as well but employed other drought categories. For the study of Shahid and Behrawan (2008), these categories were: moderate, severe, and very severe droughts. Additionally, instead of percentages of occurrence they assigned ratings/rankings based on them (the percentages were categorized according to the natural break method before the ratings' assignment). Formula (2.2) may be more extensive based on the number of used drought categories.

$$DHI_i = (O_m \times W_m) + (O_s \times W_s) + (O_e \times W_e) \tag{2.2}$$

On the other hand, Murthy et al. (2017) determined DHI with a similar formula but employing other parameters having normalized them before such as number of rainy days, rainfall mean, and its variance. In addition, the weights of DHI were not calculated with AHP but using entropy evaluation method.

In all aforementioned studies, maps of DHI, DVI, and DRI were the main outputs of their analysis. Kriging and IDW interpolation from rain gauges are two popular and sophisticated techniques for the creation of DHI maps. These techniques can be applied two different ways: (i) calculate first interpolate later (CI), and (ii) interpolate first and calculate later (IC). Based on Bechini et al. (2000), the error is higher for IC if the formula for the calculation of the dependent variable contains many independent variables. The increased number of required interpolations compared to CI lead to higher mean squared errors. Most studies used the CI method.

DVI maps were produced by the overlay of different thematic maps which represented vulnerability indicators. The maps of DHI and DRI were produced for each time scale/step (i.e. 3-month, 6-month, etc) but the DVI map was constant (there were no time scales).

Summary for drought vulnerability and risk

- A limited number of studies have focused on vulnerability of cities to droughts. Most of the current literature is related to agricultural droughts (especially in developing countries).
- Entropy evaluation and AHP have been used to determine the weights of vulnerability's indicators.
- Most indicators which have been used to determine vulnerability to agricultural droughts in developing countries are not applicable to cities.
- Vector analysis has mostly been applied in the current literature for vulnerability estimation than raster analysis.

- Terminology differs over different studies. Table 2.1 shows the correspondence between different terms. Not strictly defined, there are two main groups. The following terms: (i) vulnerability, (ii) sensitivity, and (iii) exposure of the first group correspond to (i) risk, (ii) vulnerability, and (iii) hazard respectively to the second group of studies.

Table 2.1: Different terminology among studies

Terminology	
1st Group	2nd Group
vulnerability	risk
sensitivity	vulnerability
exposure	hazard

- Standardized indices (such as SPI) have mostly been used to quantify exposure (or hazard if the terminology of the second group is followed).

2.4 Drought Characteristics

As mentioned above urban drought is not clearly defined and it has not be studied individually. Therefore, analyzing its characteristics tends to be similar as analyzing those from other types of droughts. This section refers to four drought characteristics (deficit, duration, frequency and spatial extent) for droughts in general; they can be applicable to urban droughts too. Copulas are mentioned as well based on which more than one drought characteristics can be analyzed in combination.

2.4.1 Deficit

Indices can be employed to determine drought deficit (severity). They can also be useful for the estimation of other characteristics such as duration, frequency, and areal extent. Further calculations are required, though. Indices can be separated into two large groups: standardized ones and threshold-level method taking into account loose criteria (Van Loon, 2015). Most standardized indices belong to one out of the four groups based on the four types of droughts (meteorological, hydrological, agricultural, and socio-economic). The main disadvantage of those is that drought is expressed in a relative way and not in an absolute one; thus they do not provide volumes of deficit which are useful for water management analysis (Van Loon, 2015). As for threshold level approach, the threshold can be fixed (there is subjectivity in its selection, though) or variable. The main advantage of the variable threshold is that the deficit can be estimated not only during the dry periods but also during the wet ones. Another major advantage of threshold method is that it can be used for the determination of drought characteristics (not only deficit but also duration) in different parts of the hydrological cycle. This cannot be done by indices which are applicable only for a specific domain of the hydrological cycle such as SPI for meteorological droughts only (Heudorfer and Stahl, 2017). There are different variable threshold approaches which can be employed (e.g. daily, and monthly quantiles) and the same applies for the smoothing to threshold function which is required subsequently. Beyene et al. (2014) showed that using a moving window quantile of 30 days as threshold outperformed the monthly and daily quantiles for five European catchments regarding streamflow droughts.

2.4.2 Duration

Applying standardized indices determines not only drought deficit but also duration. It is the period at which the index remains negative consecutively without any positive value to interrupt. When the index becomes positive, this is the end of the drought event.

Regarding threshold method, duration is the time when the hydrological variable (e.g. soil moisture) remains below the selected threshold. The main drawback is that large drought events may be split

into more than one since the value of the studied hydrological variable exceeded the threshold for just a few time steps. The short period that the hydrological variable is larger than the threshold was called "the inter-event time" by Zelenhasić and Salvai (1987) who studied steamflow droughts.

To address this problem, pooling is used which combines events happened closely timewise to larger events and the three most common approaches are inter-event time method, moving average, and sequent peak algorithm (Fleig et al., 2006). Furthermore, there may be many events of short duration which increase the skewness of duration distribution; therefore, omission of them is suggested. Regarding inter-event time method, Tallaksen et al. (1997) considered not only the period between previous and following drought event but also the surplus during the inter-event time. More specifically, in case the ratio of surplus during inter-event to the deficit of the previous drought event was lower than a predefined threshold, pooling was applied.

2.4.3 Frequency and Spatial Extent

Regarding frequency distribution analysis, standardized indices have been used in many studies revealing drought temporal patterns. It is common that frequency and areal extent be analyzed together since they are inextricably interwoven. McKetie et al. (1993) have showed that the frequency of droughts is reduced as time scale goes up using SPI; droughts are fewer but longer at large time scales. Weighted cumulative and average Severity-Areal Extent-Frequency (SAF) curves have been estimated using SPI for Thessaly, Greece but firstly the study area has been split into grid cells and SPI has been determined for each one of them (Loukas and Vasiliades, 2004). Spatial interpolation was applied before the estimation of SAF curves. Kolmogorov-Smirnov and Chi-Square tests have been employed so that the best fitted distribution could be selected; Extreme Value I has been selected even though other distributions like Gamma, Generalized Extreme Value (GEV), and Log-Pearson have passed the tests. The same distribution has been selected by Mishra and Desai (2005) who studied the spatio-termporal analysis of droughts in Kansabati, India showing that long-term droughts did not occur very frequently and occupy a limited areal extent. The fact that the severity is high for drought events with low frequency has been also in line with the study by Dalezios et al. (2000) for dry and wet periods at a large scale. However, the uncertainty is higher for the extreme events with lower frequency due to limited sample size.

Two common approaches different than SAF curves to analyze drought frequency from time series of recharge, discharge, and GWL data are: (i) Partial Duration Series (PDS) where only the values lower than a threshold are analyzed (Peters, 2003; Gurwin, 2014), (ii) Annual Minimum Series (AMS) where the minimum value for each year is selected (Bloomfield et al., 2003). A significantly low threshold value for PDS can lead to a limited number of drought events and a statistical analysis cannot be performed. Additionally, in PDS method, a large number of droughts of short duration and deficit is not desirable since it may influence the skewness of the distribution (Hisdal and Tallaksen, 2000). The threshold varies but values between the 50^{th} and 95^{th} percentiles of hydrological variables tend to be selected (Peters, 2003). The main drawbacks of PDS are: (i) there is subjectivity for the threshold selection, and (ii) the events may not be independent. On the other hand, for AMS the main drawback is: some low values are missed since only one value can be selected in a year (or per block if an annual period is not used) (Peters, 2003). Therefore, multi-year drought events cannot be recognized using AMS compared to PDS. After PDS or AMS is created, a distribution is fitted (Pareto and GEV are the most common ones respectively).

As for the spatial analysis, interpolation techniques are often implemented based on the positions of the measurement points (e.g. rain gauges, groundwater sensors). The most common method is Kriging for drought interpolation but Loukas and Vasiliades (2004) and Mahajan and Dodamani (2016) used multiple linear regression and Inverse Distance Weighting (IDW) accordingly. Mahajan and Dodamani (2016) employed SPI and Percent of Normal Precipitation (PNP) in order to analyze the spatial dimension of droughts applying IDW.

2.4.4 Copulas

Drought characteristics have strong interrelationships and copulas are used to describe their dependence. There are different ways to analyze the characteristics of a drought: one by one (uni-variate analysis), in pairs (bi-variate analysis) and many of them using more than two characteristics (multi-variate analysis) (Mishra and Singh, 2011). The last two approaches can be conducted with the help of copulas which convert univariate distributions to multi-variate ones; families of copula such as Clayton and Gumbel-Hougaard can be used for that purpose. Copula parameters determination is conducted after maximum likelihood estimation has been performed for the marginal distributions (Shiau, 2006). Copula approach is gaining momentum regarding drought analysis showing the relationships of drought variables but more attention to multi-variate copula needs to be paid.

Summary for drought characteristics

- Deficit and duration can be estimated using threshold method and standardized indices. Threshold method provide more advantages than indices even though the latter ones are common.
- Pooling is a technique to combine drought events which happened closely timewise. It can only be applied for threshold method (not for standardized indices).
- It is possible frequency and areal extent to be analyzed together. Annual Minimum Series (AMS), and Partial Duration Series (PDS) are two common techniques for frequency distribution analysis. Both of them have their pitfalls (issues regarding (i) omission of values and (ii) subjectivity, for AMS and PDS respectively).
- Long droughts tend to occur less frequently and occupy a limited areal extent (Mishra and Desai, 2005).
- Copulas can analyze the interrelationships among two or more drought characteristics. This approach seems promising but more research is required.

2.5 Soil Moisture Droughts

Iñiguez et al. (2016) have studied soil moisture droughts at catchment scale using Probability Distribution Moisture (PDM) model, and more specifically, Nash-Sutcliffe was the objective function for calibration and validation. It is worth noting measurements of soil water content have been conducted using time domain reflectrometry (TDR) probes. Threshold method has been employed for drought identification. Based on sensitivity analysis, precipitation influenced more the drought recovery compared to other variables.

Seasonality and climate zone of study area are two parameters which influence soil moisture drought and drought propagation through the hydrological cycle. Soil moisture droughts in dry periods are limited since they cannot be expanded over the permanent wilting point (Van Loon et al., 2014). Soil moisture varies between field capacity and permanent wilting point and its value influences percolation.

Soil Moisture Index, which was proposed by Hunt et al. (2009), takes into account the permanent wilting point, and the field capacity of the soil; the scale varies from -5 to 5, and the value of 0 means that the water content is half full to its full capacity. The main disadvantages of this approach is that measurements are required (e.g. using soil probes) but are unavailable in most cases. Additionally, the field capacity and permanent wilting point are difficult to be estimated especially when there is not a long time series of data (Hunt et al., 2009). Two other indices which have been used for soil moisture droughts are: Soil Moisture Deficit Index (SMDI) (Narasimhan and Srinivasan, 2005), and Soil Moisture Anomaly (SMA) (European Drought Observatory (EDO), 2019).

2.6 Groundwater Droughts

Mishra and Singh (2010) suggested that groundwater drought ought to be the fifth category in the main drought classification of droughts based on which droughts are classified into meteorological, hydrological, agricultural, and socio-economic ones (Wilhite and Glantz, 1985). Its features are different than those of the aforementioned drought types. The impacts of groundwater droughts can be apparent in three components: GWLs, groundwater recharge, and groundwater discharge.

The causes of a groundwater drought are the reduced recharge or increased baseflow but the most common is the first reason. Higher evaporation and reduced effective precipitation lead to less infiltration which means less groundwater recharge (Peters, 2003). In case of regions where groundwater is shallow such as the Netherlands, groundwater uptake by vegetation can occur for transpiration, resulting in reduced reserve of groundwater.

Hydrogeological conditions of the aquifer influence groundwater droughts and their ignorance can lead to misleading results (Bloomfield and Marchant, 2013). Groundwater reservoir characteristics (such as the reservoir coefficient) affect the deficit of a recharge drought to a discharge one (Peters, 2003).

Regarding droughts based on GWLs, three methods which have been used to evaluate the deficit of a groundwater drought are: Standardized Groundwater Index (SGI) (Bloomfield and Marchant, 2013), Base Flow Index (BSI) (Gustard et al., 1992), and the threshold method (Yevjevich, 1967) which is based on run theory. SGI is based on the way Standardized Precipitation Index (SPI) is calculated (transformation from a gamma distribution to a normal one). As for threshold method, the threshold can be constant or variable as mentioned in Section 2.4.1. Heudorfer and Stahl (2017) have showed that the number of short and long groundwater droughts has risen and decreased respectively when the variable threshold was used compared to the constant one. Threshold method was designed for monthly time series but it has also been used with daily data.

The relationship between groundwater and meteorological droughts is a common topic of interest since the second ones can be monitored without difficulties almost everywhere. Correlation between SPI and SGI is dependent on the inherent characteristics of the aquifer such as transmissivity and storage. Kumar et al. (2016) have showed that the meteorological index SPI is not able to show the existence of a groundwater drought since the correlation among the two indices was limited (21-66%). However, there are several meteorological indices such as Rainfall Anomaly Index (RAI), Standardized Anomaly Index (SAI), and Deciles which have not been tested for their relationship with groundwater droughts. Additionally, mapping the cumulative deficit for GWLs below a threshold and comparing it to SPI based on the study of Shahid and Hazarika (2010) in Bangladesh provided a deeper understanding of the relationship between meteorological and groundwater droughts. Autocorrelation cannot be executed in drought events extracted from threshold method since this technique does not create a continuous index which is required. Standardized techniques and threshold method can complement each other underlining groundwater drought characteristics.

Summary for soil moisture and groundwater droughts

- There is a variety of indices available for soil moisture droughts.
- The impacts of groundwater droughts can be apparent in three components: GWLs, groundwater recharge, and groundwater discharge.
- Reduced recharge (due to high evaporation and less effective precipitation) is one common reason for groundwater droughts.
- Hydrogeological conditions of the aquifer influence groundwater droughts.
- Both standardized indices and threshold method have been used for groundwater drought analysis.
- Research regarding the relationship between groundwater and meteorological droughts can help

in predicting groundwater droughts. Meteorological index SPI is not able to show the existence of a groundwater drought (Kumar et al., 2016).

Summary Literature Review - Research Gap

- As underlined in this literature review, the number of studies regarding vulnerability of cities to droughts is limited. Most research has focused on agricultural droughts or droughts in general. Even though, there are some studies about the urban environment, they compare vulnerability among cities and not among different parts of one city. The current study is one of the first steps to cover this gap.
- Urban drought is a complex phenomenon and lack of categorization hinders a thorough analysis with narrow scope. Although, some attempts have been made to categorize urban droughts, there is no clear description of the procedure which needs to be followed. That is another gap which this research addresses.
- For the estimation of vulnerability of cities to droughts, analyzing drought characteristics is required. There is a variety of techniques available to identify the characteristics of droughts in general but their performance have not been tested for urban droughts (regarding groundwater and soil moisture specifically).

Chapter 3

Methodology

3.1 Urban Drought Categorization Framework

In order to determine the vulnerability of a city to droughts, urban drought definition is a prerequisite so that the scope can be narrow. For that purpose, a framework of urban drought classification was created where water functions of cities were considered. The Urban Drought Categorization Framework provided a definition for an urban drought and a step-by-step process so that urban droughts could be categorized. The main goals of the framework are:

- Providing a generic definition of urban droughts.
- Determination of a method based on which urban droughts with different characteristics can be categorized. In that way, policy makers and stakeholders can determine which type of urban drought takes place in a city which is catalytic for planning and urban drought mitigation.
- Determining the hydrological variables which contribute to the disruption of a water function, and are necessary for the analysis of each urban drought category.
- Determining the exposure and sensitivity of a city to the specific urban drought which the city faces.
- Determining the time scale for each urban drought category.

Before presenting the methodology, it is important to highlight the definitions of key terms for this analysis. Most definitions are based on the framework suggested by Fritzsche et al. (2014). Exposure refers to the damaging event in itself. In the present case, it is a drought. Sensitivity shows how the system changes due to increased exposure; this term is mostly related to the inherent characteristics of the system. Coping capacity is the ability of the system to manage adverse conditions based on its available resources. Coping capacity is mostly a short-term characteristic whereas adaptive capacity is a long-term one including learning of past events. Adaptive capacity is more about taking protection of future hazard events. According to Intergovermental Panel on Climate Change (2007)[p.27] "Vulnerability is the degree to which a system is susceptible to, and unable to cope with, adverse effects of climate change, including climate variability and extremes".

3.1.1 Drought Definition

As mentioned above, it is very difficult to define droughts let alone urban ones but a definition is necessary in that framework enhancing clarity even though it may not be the best one. The formulated definition is: "Urban drought occurs when one specific part of the water cycle from the following ones (drinking water supply via reservoirs or/and groundwater which services the city, soil moisture, groundwater, and canals/ponds/wetland) has less water than 'normal' conditions and at least one of the city functions which are directly related to this water cycle part is 'disrupted' ". Policy makers of a city who want to categorize drought events need to determine the words in quotes ('') in the urban

drought definition. 'Normal' conditions for water availability in a water cycle part is considered the 30-year average. They can be either measured or estimated (it depends on the water cycle part). The term 'disrupted' means there is a deviation for a water function from normal conditions (expressed in terms of probability, e.g. percentage of normal values) to such a degree that the city starts facing losses regarding society, economy, and environment (at least in one of those aspects). For the estimation of when each water function is disrupted, a workshop with policy makers and water managers of the city studied is suggested.

Tipping points or thresholds for each water function of a city need to be determined after a thorough discussion/workshop with its policy makers. The term 'tipping point' was used for first time in sociology but has been used in other disciplines such as meteorology, environmental science, and especially climate change research. Tipping point is a threshold and when is exceeded leads to a rapid change in a short period of time. It is not sure that the water functions have tipping points; they may just have thresholds which leads to significant water function disruption. For instance, if the water level in a canal is lower than a specific value, recreation boats cannot float which disrupts the water function about recreation in a city. Only if the city has suffered a drought, its policy makers would be able to provide valuable insight into the thresholds or tipping points.

Urban droughts are region dependent since an urban drought in one city influences different water functions than another one. Different water functions are related to different urban drought types. Drought categorization is the topic of the next section.

3.1.2 Drought Categorization

In this part, the steps for drought categorization are presented (see Figure 3.1). Firstly, policy makers need to record the water functions of a city. There is no limitation on the type of water; functions of surface water, groundwater and so on ought to be included given that the type of water is present on the city studied. Water functions can be categorized into the following classes: water landscape, water ecology, water resources, water security, water economy, and water culture based on Yu et al. (2018). Water functions can belong to more than one out of the six water categories as can be seen in Figure 3.1. This categorization is helpful for identifying all water functions by policy makers.

After that, drought disrupted water functions need to be categorized based on the affected part of the water cycle. There are four water cycle categories which can be affected: soil moisture within city's boundaries, groundwater within city's boundaries, canals\ponds\wetland if the city has one of the three ones, and reservoirs and groundwater from which the city receives water. The latter water cycle part (reservoirs and groundwater) can be in the catchment where a city is located but it is quite uncommon since most cities worldwide derive their main water resources from rural environment often even located in another watershed. According to the four affected water cycle parts, there are four urban drought categories: Soil Moisture Urban Drought (SMUD), Groundwater Urban Drought (GUD), Open Water Urban Drought (OWUD), and Water Supply Urban Drought (WSUD). These four types can be the result of a shortage of water (quantity), but also due to poor quality, so that the water can no longer be used for the functions it is supposed to serve. It is possible that in one city more than one of these urban drought types occur simultaneously.

Some examples of disrupted water functions are provided so that the framework becomes more tangible and easier to be applied. Regarding SMUD, the functions of (i) preservation of urban parks and urban agriculture, and (ii) keeping bearable temperature for residents can be disrupted. The second one is an indirect water use since lower soil moisture can lead to less evaporation and transpiration, and as a result, cooling is reduced. Transferring geothermal energy and keeping stable buildings are two water functions which may be disrupted due to GUD. Wooden pole foundations of buildings can be extremely sensitive to low GWLs, as they start rotting when exposed to air in an unsaturated zone. If the city is located in a wetland or canals go though it or it contains a pond, water has more functions which may be hindered due to a OWUD such as recreation, shipping goods, and preservation of aqua

flora and fauna. Last but not least, the urban drought category which has the most significant impact on a city is WSUD and some water function examples are providing drinking water supply to residents, and smooth operation of water-related industries.

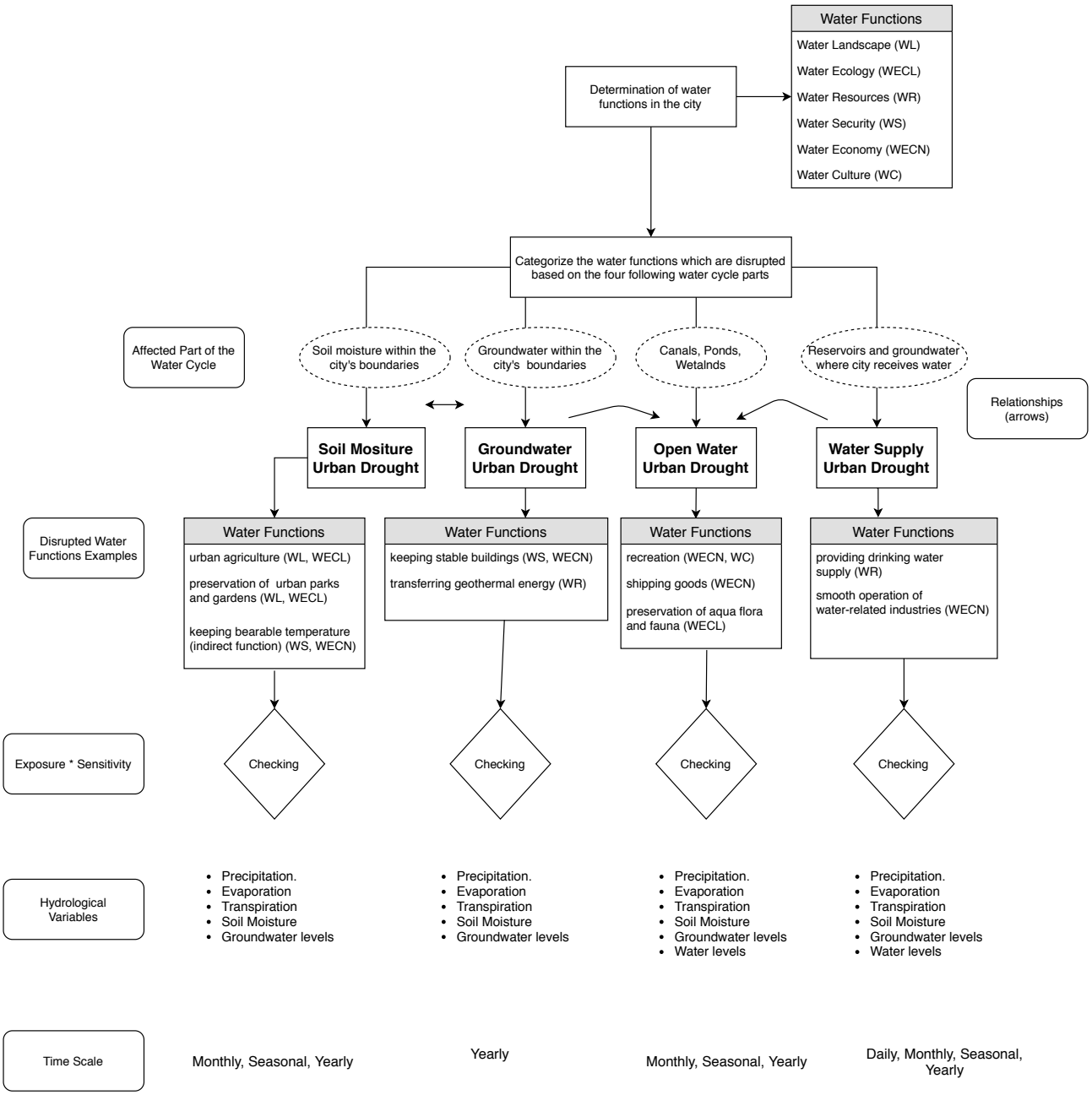


Figure 3.1: Urban Drought Categorization Framework

The degree to which each urban drought category has an influence on a city is determined by the product of exposure by sensitivity. For instance, regarding GUD, if the city is not related to any water function which can be disrupted due to lower GWLs, there would be no impact. To elaborate it, if there is exposure to GUD but the sensitivity related to this specific drought type is zero, the city would not sustain any damage (e.g. if GWL in a city in lowlands goes down but buildings' poles are not constructed with wood, the function of keeping stable buildings is not disrupted).

The indicators for exposure and sensitivity vary over the cities in the world since cities have different physical, economic, and social characteristics. The indicators used for one city may not be applicable

to other cities. For instance, indicators regarding illiteracy level and gender inequality are not useful in developed countries. Exposure and sensitivity can be estimated using one out of many available vulnerability frameworks.

Hydrological variables which are required for analyzing each specific urban drought is another part of this framework. These variables may also contribute to the disruption of water functions directly or not. The hydrological variables used for SMUD are the same as those of GUD since unsaturated zone and groundwater reservoirs have a high interrelationship. Those variables are: precipitation, evaporation, transpiration, soil moisture, and GWLs. For the estimation of potential evaporation, other variables such as temperature, humidity, and wind speed can be useful. Regarding OWUD, an added variable to the previous ones is canals or pond levels. The variables used for WSUD are similar as those as OWUD; the spatial scale is different for WSUD though. For all other categories, it is the area where city is located but for WSUD can be different since the location of groundwater wells and reservoirs from which a city takes water can be far away from the city. This makes the analysis more complicated especially when the water is derived from a reservoir and water is transferred through another catchment in order to reach the city and the treatment plant for drinking water. It is worth noting that in this phase of the framework, there is confusion since many hydrological variables are indirectly connected to all types of drought; therefore having data about all of them is recommended but not required.

Last but not least, time scale is important and varies over the different types of urban drought. Time scale for soil moisture is suggested to be from monthly to yearly whereas for groundwater to be yearly. Groundwater reservoirs require a lot of time for recharging; therefore suitable time scales are long (more specifically, years but shorter time periods can be applied too). A medium time scale is appropriate for OWUD (months and seasons) but it can be longer (years) to identify patterns as well. Daily analysis is suitable for WSUD since dramatic changes happen rapidly and a larger time scale cannot control the situation (smoothing takes place for long time scale which is undesirable for water supply analysis). However, larger scales can be useful in extreme conditions when WSUD occurs for a long time.

Water functions determination of a city is a key process in this framework and haphazard attitude towards it can derail the whole analysis. If a specific water function of a city is ignored and not included in the first step of this analysis, then it is not going to be tested regarding drought disruption. Consequently, exposure and sensitivity for this specific water function is not going to be determined. Policy makers and stakeholders play a catalytic role in the process; a fruitful discussion among them is a prerequisite making sure all water functions of a city are included.

Coping capacity which is a component of vulnerability (based on some frameworks) is not included in this framework of urban drought categorization since coping capacity can be considered after an urban drought event has started occurring. Goal of this framework is to identify the conditions at which each different urban drought category occurs and not to define the vulnerability to this specific urban drought type. Vulnerability is determined in a subsequent part of this methodology in the current research.

3.1.3 Relationship among urban drought categories

The different types of urban droughts are interrelated since one type can cause another one. SMUD can cause GUD since less percolation occurs due to lower soil moisture. In addition, GUD can cause SMUD when less capillary rise takes place but the degree of influence is much less compared to the other way around since the amount of water that moves from groundwater to unsaturated zone reservoir via capillary rise is not large.

A GUD can cause an OWUD in case there are canals or a pond in a city. GWLs decrease and as a result less water flows into the canals or pond, leading to an OWUD. Consequently, the levels of

pond or canals drop, but this process takes time given that groundwater flows with a low velocity (it can take months, seasons, or even years; permeability is a key factor and is dependent on soil types). Additionally, WSUD can cause an OWUD in case the supply comes from groundwater within a city's boundaries, but this case is quite uncommon. The context is similar as before that GWLs go down. A WSUD can cause a GUD and the latter one can cause an OWUD.

This framework includes a comprehensive inventory of indicators which can be used for exposure and sensitivity estimation of SMUDs and GUDs (see Table 3.1). These indicators apply only for developed countries. Exposure, sensitivity, and coping capacity are related to each other, so determining in which group an indicator fits best, is often confusing and unclear but context can be helpful (Polsky et al., 2007).

Table 3.1: Exposure and sensitivity indicators to soil moisture and groundwater urban droughts

Soil Moisture	Groundwater
Exposure	Exposure
Physical characteristics • Potential evaporation based on the geographical location (indirect) • Deficit of oil moisture droughts based on past events • Duration of soil moisture droughts based on past events • Frequency of soil moisture drought events	Physical characteristics • Deficit of groundwater droughts based on past events • Duration of groundwater droughts based on past events • Frequency of Standardized Groundwater Index (SGI) • Population density
Sensitivity	Sensitivity
Physical characteristics • Percentage of soils with high permeability • Percentage of urban agricultural area to total city area • Percentage of urban gardens, parks, forests to total city's area • Percentage of parks which are irrigated • Percentage of parks whose plants\trees require little water • Distance from the closest canal • Existence of risk-reduced services	Physical characteristics • Percentage of soils with low permeability • Building density • Percentage of buildings on wooden piles • Percentage of cultural heritage buildings on wooden poles • Locations of critical infrastructure (e.g. hospital, electricity station) • Location of enterprises • Groundwater extraction permits
Social characteristics • Population density	Social characteristics • Population density • Percentage of minority population whose house is on wooden piles • Financial status of residents whose house is on wooden piles (this could be in adaptive capacity as well)
Political characteristics • Existence of regulations regarding where parks, urban gardens, urban forests, and urban agricultural areas can be created and what kind of vegetation, crops, and tress can be planted	Political characteristics • Existence of regulations regarding the way buildings ought to be erected and type of materials ought to be used eradicating the use of wooden poles

This study focuses only on GUD and SMUD of that framework; a scope of urban droughts in general can be chaotic due to the complexity of urban droughts. In hydrological analysis, components of other urban drought categories may influence SMUD and GUD. For instance, drinking water supply (which is related to a WSUD) may influence a SMUD via leakages of the drinking water network or by assigning a specific amount of drinking water for urban irrigation. However, there was a lack of this data type; therefore it was ignored in the study.

3.2 Study Area

Case Study

The framework was applied to Leiden in the Netherlands for the two following reasons: (i) it is a well monitored city, and (ii) it may face drought challenges in the future. Only some administrative units (districts) were used and not the whole city since a high resolution analysis at a limited spatial scale could provide a profound understanding of vulnerability. A city's characteristics vary over space; thus, studying it as a whole requires assumptions and oversimplification, missing the heterogeneity. The latter one is useful to indicate drought differences among regions. The current study was focused on the following districts (wijken in Dutch): Binnenstad-Zuid, Binnenstad-Noord, Bos en Gasthuis, and Boerhaave (Figure 3.2). These districts were selected because they were quite heterogeneous (they were built in different periods and their soil type is different too). Additionally, at 'wijk' level there is homogeneity regrading hydrological conditions. Another administrative level is 'buurt' which is at street scale but does not ensure homogeneity; this extra division was created for management purposes and is not related to the hydrological context of the area. In the current analysis, 'wijk' level was employed. A second study area was planned to be the city of Athens in Greece (more specifically a few municipalities) but lack of data was the main deterrent. A limited analysis was conducted and vulnerability was not estimated (its results can be found in Appendix D).

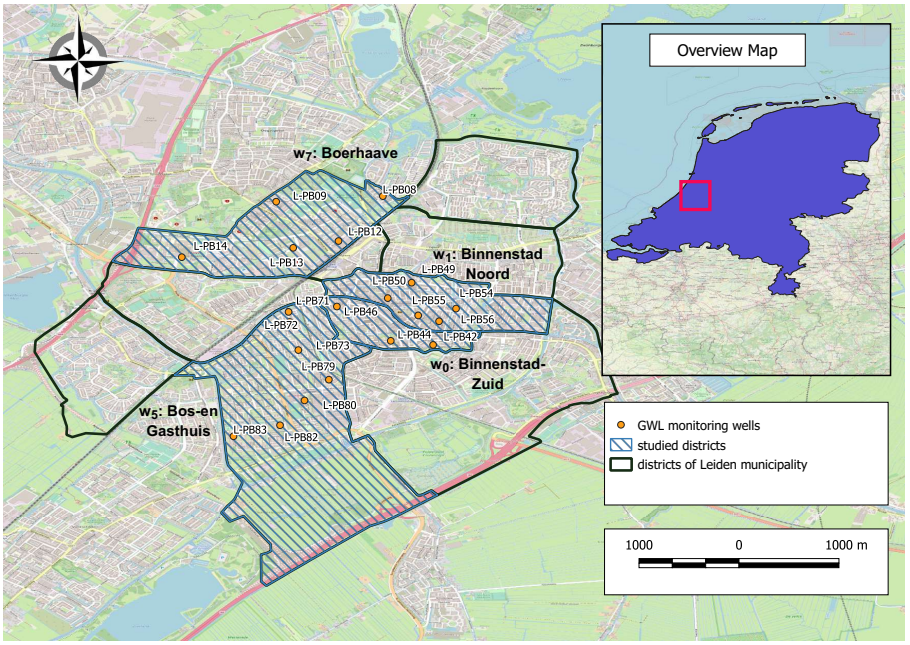


Figure 3.2: Study area in Leiden and the groundwater monitoring wells

Due to the fact that a meeting with the stakeholders of Leiden was not arranged, a detailed record of water functions of this city was not created. For soil moisture droughts, the main function is preservation of urban parks, whereas for groundwater ones, it is keeping buildings stable. Exposure and sensitivity are parts of (i) the urban drought categorization framework, and (ii) the vulnerability framework suggested in the current research. To save space, the results of exposure and sensitivity are presented in the vulnerability part. Even though it was not tested explicitly, the product exposure by sensitivity is larger than zero for the two drought types studied; therefore, groundwater and soil moisture urban droughts affect Leiden.

Data description

Regarding exposure to groundwater droughts in Leiden, actual measurements were mostly used. Companies Wareco, ireal, and Acacia monitor that area but for all of their monitoring wells, the measuring period was not quite long (for Acacia it was about two years, for ireal about four years, whereas for Wareco about six months to one year). More specifically, for the four districts: Binnenstad-Zuid, Binnenstad-Noord, Bos en Gasthuis, and Boerhaave, there were seven, 11, 41, and eight Acacia wells,

two, one, six, and four ireal wells, and three, one, one, and one Wareco wells respectively. Out of those, only the wells with data for the hydrological year 2018-2019 were selected (i.e. three Acacia wells in Binnenstad-Zuid (w_0 district), five ones in Binnenstad-Noord (w_1 district), seven ones in Bos en Gasthuis (w_5 district), and five ones in Boerhaave (w_7 district). Figure 3.2 illustrates the location of the selected groundwater wells where there was available data for the period April 2018 - March 2019. In the Netherlands, a hydrological year starts when the growing season in agriculture begins, i.e. in April. Hydrological year differs based on the country and climate conditions. For example in Mediterranean countries, it starts in October since the system is empty and there are no large fluxes. However, this does not hold in the Netherlands and April is the start of the hydrological year since it is the best time to measure water balance. Thus, hydrological year for Leiden is different than that of Athens.

Hydrological year 2018-2019 was selected in Leiden since it was the one with the most available data. Some wells had data for another hydrological year (e.g. 2019-2020) but they were dismissed since these wells were limited. Maximization of wells, for a common hydrological year, was an objective when wells were selected. That was another factor which influenced the selection of the studied districts (wijken) as well.

Wareco wells had logical measurements but were rejected due to their low duration; they did not cover a full hydrological year. On the other hand, ireal company had wells with long measurements (around four years) but many time series were erratic with high fluctuations or abrupt drops and therefore were dismissed. Both companies measure groundwater using sensors but also conduct manual measurements to calibrate the raw ones. Even though the calibrated measurements of ireal company were of poor quality; therefore, its wells were not used. Wells of Acacia were also dismissed when they had large gaps, or their values were illogical. It is worth mentioning that raw data was collected on an hourly basis but it was resampled to a daily one. GWL is a state and not a flux variable with a delay in its response; therefore a high resolution time scale (such as hourly) was not suitable for the current study.

Regarding missing values, linear interpolation over short periods (maximum 1.5 months) was used since it is a quick method which can provide satisfactory results. Neural network was used for one well only (L-PB08 in in w_7 district). Data driven methods tend not to work when there was lack of data. That was the case in the current study. A data driven approach was used to verify the former statement. A detailed explanation about the Sequential Model used and its results are attached in Appendix B.2. In addition, regarding Acacia wells, some unrealistic values in groundwater time series were corrected manually; they were substituted by interpolated values only when wrong values did not span a long time period (30 days). In case that period was exceeded, no correction was conducted and the well was dismissed.

3.3 Exposure to groundwater droughts

Exposure is a critical component of vulnerability and a modified version of the framework, suggested by Fritzsche et al. (2014), was applied in the current study. This framework is suitable for soil moisture and groundwater urban droughts (in general, engineering projects). The modifications to the present study compared to the original framework were: (i) exposure and sensitivity were not aggregated but used directly for the estimation of vulnerability (ii) Analytical Hierarchy Process (AHP) was used to assign weights, (iii) adaptive capacity was substituted by coping capacity, and (iv) coping capacity was not considered.

Figure 3.3 illustrates how vulnerability of Leiden to groundwater droughts was estimated. It highlights the techniques used for the estimation of drought characteristics. A chart flow regarding the methodology for both groundwater and soil moisture droughts is presented in Figure A.1.

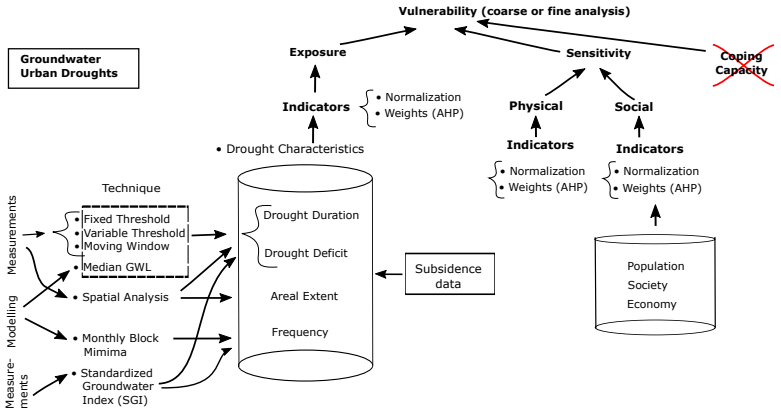


Figure 3.3: Roadmap of Methodology regarding Groundwater Urban Droughts

3.3.1 Estimating drought deficit and duration

As for groundwater drought's deficit and duration, threshold method and a standardized index were used. More specifically, as regards threshold method, four different approaches were employed: (i) fixed threshold, (ii) moving average of monthly quantile (variable threshold for short), (iii) 30-day moving window quantile (moving window for short), and (iv) median GWL. Selected threshold approach is one critical parameter for drought recognition (Beyene et al., 2014; Stahl, 2001). For this reason, four approaches were employed in this study and compared to each other. In that way, the most suitable technique for identifying drought events, was determined based on the output. The applied technique influences drought characteristics and those in turn the exposure to drought and as a result the vulnerability. There is no physical superiority of one method compared to another (when the threshold is not fixed) for spotting drought events; hence, visual comparison is critical (Beyene et al., 2014).

Fixed, Variable, and Moving Window threshold

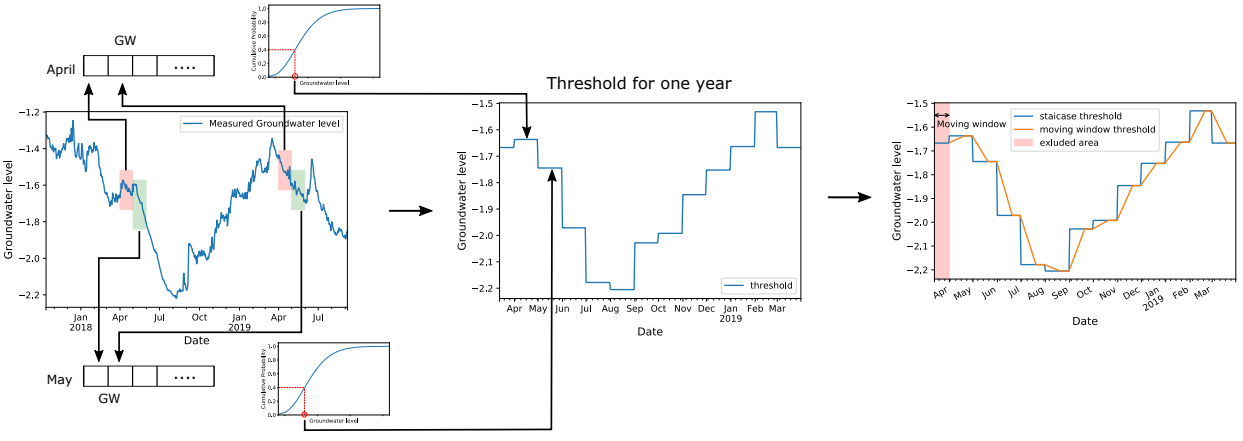


Figure 3.4: Schematic regarding the estimation of variable threshold (GW: Groundwater level)

The three out of the four techniques used to estimate a threshold are analyzed below. Fixed threshold is a constant value based on a percentile considering the whole time series. To elaborate the 2nd technique, moving average of monthly quantile technique (e.g. for 20th percentile), the following steps (see Figure 3.4) were applied: (i) for each month of the year, the 20th percentile was determined with the help of the cumulative distribution function (CDF) from all daily values in that month over all years, (ii) the value of the 20th monthly percentile was assigned to all days of the month, and (iii) backwards moving average of 20 days was applied to the whole year to smooth the 'staircase

differences' and extinguish the abrupt jumps in threshold function between consecutive months. For this specific approach, the input data period was from the 13th of March 2018 to the 31st of March 2019 due to moving average application. The start date was not the 1st of April since in that case the resulting time series (after applying moving average) would start from the 20th of April. This could lead to inconsistencies on the comparison with other methods where the study period was exactly one year. It is worth highlighting that using this technique, the output threshold function had a span of one calendar year. Therefore, for studies (not the present one) with duration more than one year, the threshold function would repeat itself.

As for the 30-day moving window quantile approach, for each time step a specific percentile (e.g. 20th one) was estimated based on the GWLs which were measured 15 days before and 14 days subsequently to that time step (the size of the window was 30 days). In that case, a centered moving average approach was applied and not a backwards one compared to variable threshold. For that reason, the input data for estimating the 30-day moving window quantile function was from the 17th of March 2018 to the 15th of April 2019. In that way, the threshold function spanned a hydrological year and comparison period was the same as to the ones of other approaches (fixed and variable threshold). Another technique is daily moving window quantile which was not estimated in the current analysis due to lack of data. It requires long time series in order to estimate the CDF for each day of the year. Another popular technique for drought identification is Sequential Peak Algorithm but it was not used since it requires recharge and discharge information which was not available; it is also more suitable for fluxes and not states.

There is subjectivity on the selection of threshold values which are based on the purpose of study and the involved stakeholders. To eliminate it, three percentiles (20th, 30th, and 40th) were applied for each of the three aforementioned techniques and their results were evaluated. The most common percentiles in literature are 10th, 20th, and 30th (e.g. studies of Heudorfer and Stahl (2017) and Hisdal and Tallaksen (2000)). Tallaksen et al. (2009) have used only 20th percentile whereas Gurwin (2014) has empoyed percentiles based on the standard deviation (50% of standard deviation). In the current study, given the range of groundwater data in study area, a low percentile such as 10% would lead to a limited number of drought events, impeding further analysis. For that reason, percentiles were shifted upwards by 10%.

The scripts of the functions for identification of drought events and the calculation of their characteristics (deficit and duration) are attached in Appendix C.1. The algorithm considered the following cases as well; (i) when one drought event started before the first time step of the time series, and (ii) when one drought event did not finish at the end of the given time series.

Calculation of drought characteristics

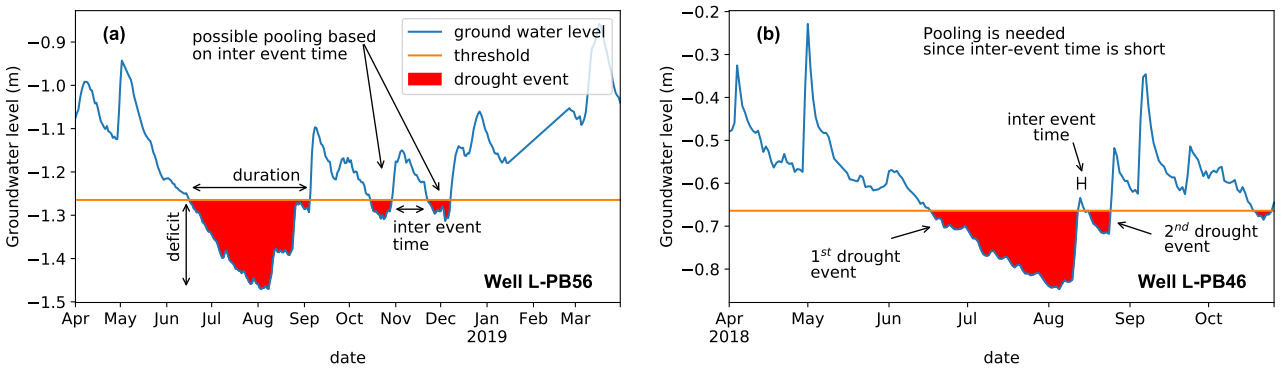


Figure 3.5: Illustration of basic drought concepts using fixed threshold method for wells L-PB56 and L-PB46

As for drought characteristics, for a flux variable, the severity metric is the sum of deficits for a drought

event where deficit is the subtraction of the hydrological flux from the threshold. On the other hand, regarding a state variable, severity is the maximum variance from the threshold during the drought event (Peters, 2003; Van Loon and Van Lanen, 2012) (the threshold function may be estimated based on one out of the four threshold techniques: (i) fixed threshold, (ii) variable threshold, (iii) moving window, and (iv) median GWL). GWL is a state variable; therefore the second case for severity estimation was applied. A drought event starts when the GWL is lower than the threshold and ends when it is higher than that. Their difference defines drought duration. In Figure 3.5, fixed threshold technique was used but the same concepts (regarding drought duration and deficit) hold for the other techniques of threshold estimation as well.

Pooling

Pooling is the process of merging drought events which occurred closely time-wise. It is common that one large drought event is split into two or more events (of less magnitude) since GWL exceeded the threshold function for a limited number of time steps (see Figure 3.5-b). Therefore, instead of using one large drought deficit and duration, two (or even more) lower values of deficit and duration are used. This can lead to underestimation of deficit and duration in case the average values of a specific period are estimated.

Pooling was applied in all cases of threshold except median GWL. Drought events with duration of one time step were removed. The impact of pooling of drought events which happened closely time-wise was investigated. There are many pooling techniques available but only inter-event time method was applied in the current study. When the difference between the start time of i+1 drought event and the end time of i drought event was lower than a specific threshold, pooling was conducted. In this study, five, seven, 10, 15, 20, 25 days were used as inter-event time and compared to the case where no pooling was applied. Standardization implemented using z-score; the latter is a normalization technique which shows how many standard deviations an element differs from the mean (Formula 3.1). Standardization allows the comparison of scores from different normal distributions. In the current study, dataset of one well consisted of the mean deficits of drought events where pooling was conducted (6 cases), and the mean deficit of the case where 'no pooling' was applied. Regarding $\overline{x}$, instead of the mean drought deficit of the whole dataset in one well, the mean deficit of no pooling was used (see Formula 3.2). Standard deviation in that Formula was derived using the mean deficit values of all pooling inter-event periods and 'no pooling' mean deficit as well.

The procedure mentioned above concerned one well for one specific threshold method technique and percentile. This process was executed for the rest of threshold derivation techniques and percentiles for at least one well in all districts to identify a pattern. A threshold derivation technique and one percentile with satisfactory results regarding drought identification were selected and based on those, standardized mean deficit was estimated for all wells in each district. The same procedure applied for drought duration as well. Additionally, the influence of pooling in the distribution of drought events was examined inspecting the changes in histograms visually. The sequent peak algorithm (another technique for pooling) was not applied since it requires flux and not state data.

Statistics of drought deficit and duration over different threshold methods and percentiles were determined without the application of pooling. Pearson coefficient was used to evaluate the relationship between drought duration and deficit. Regression was not applied since there was no cause-effect relationship between these two drought characteristics.

$$z = \frac{x - \overline{x}}{\sigma} \tag{3.1}$$

z: standardized score,
x: each value in the dataset,
$\overline{x}$: mean of dataset
σ: standard deviation of the dataset

$$Std\ mean\ defic_{pool=i} = \frac{Mean\ defic_{pool=i} - Mean\ defic_{no\ pool}}{\sigma_{all\ pool\ and\ no\ pool\ mean\ defic\ values}} \tag{3.2}$$

i: inter-event time for pooling (five, seven, 10, 15, 20, 25 days)
defic: deficit of drought events
pool: pooling

Using the median GWL as threshold

Another threshold for the estimation of drought deficit and duration was the median GWL for the period April-September. Drought deficit and duration were estimated using the technique mentioned above. Besides drought deficit, GWL deficit was estimated. It is the difference between one daily GWL value of one year and the corresponding median value. When the difference was negative, zero was set.

Most GWL measurements lasted only for one year; therefore, the following steps were applied to create data for more years via which the median could be estimated. Two weather stations were used and their data (hourly time scale) were merged. They were: Valkenburg and Advances stations; their data period was 1988-2015 and 2016-2019 accordingly. Valkenburg station had long data but it did not include the year 2018 when a severe drought occurred. For the common period of the two weather stations (2014-07-16 - 2016-05-02), Pearson correlation coefficient was calculated for precipitation and Makkink potential evaporation (it is estimated by KNMI) and their values were quite high (0.93 and 0.91 respectively with p-value lower than 0.01). That means that weather stations acted the same way. Merging of weather stations is not generally suggested but in this case was conducted since a rough estimation of groundwater deficit was the objective and not the exact values. Thus, the robustness of those results was quite low. Based on the merged measurements of weather parameters, potential evaporation was estimated. Additionally, reference crop grass transpiration was computed based on the former. These two datasets and precipitation were used as inputs for Urban Water Balance Model (UWBM) (see Section 3.4). One of the model outputs was GWL for 32 years. Its median value was used for the determination of groundwater deficit on the period April-September.

Furthermore, precipitation deficit was estimated using the merged data of the two weather stations. It is the cumulative difference between precipitation and potential evaporation for the period April-September; for negative values of cumulative difference, zero value is set (Beersma, 2007). The two weather stations mentioned above were used. KNMI provides data regarding daily Makkink evaporation for all stations; therefore, Makkink evaporation was used instead of Penman one for the estimation of precipitation deficit.

Changing the reference system of GWL allowed a district comparison. GWL results of UWBM were measured in respect to ground surface; the latter was different in each district. By adding the elevation of the well used in calibration to modeled GWL, the reference system changed from the ground surface to NAP. This process was repeated for each district. In that way comparison among districts was feasible.

3.3.2 Spatial identification of drought events

Another dimension of droughts is areal extent and it can be estimated by spatial analysis. This method can also be used for the identification of drought duration and deficit. By applying threshold method, spatial extent of droughts was ignored. More specifically, for all techniques except median GWL, threshold method was applied only at monitoring wells, whereas for median GWL, one value covered each district, which was not right either. In order to incorporate drought spatial extent as a parameter of drought identification, distributed groundwater data was required. The method suggested by Tallaksen et al. (2009) was followed in the current study but the input that was used was different. More specifically, the authors used SWAP and MODFLOW models which provided distributed outputs; thus a spatial analysis was suitable. On the contrary, in this study, these models

were not applied and therefore interpolation was used to create spatial information over the area. The gridded cells of the raster output from the interpolation were used as input. The spatial analysis was implemented only to one district (w_0) and only for GWL since it was computationally expensive.

Interpolation was conducted for 365 days (April 2018 - March 2019) for one time per day, using the daily measurements of the three monitoring wells L-PB42, L-PB44, L-PB46. In that way, 365 raster images were created and their areal extent was the same as the boundaries of w_0 district. The resolution was coarse (23 m x 23 m) since this method is computationally expensive; higher resolution would require web cloud services with higher capacity than a usual computer. Each raster image consisted of 26 rows and 72 columns.

Regarding the estimation of drought duration, Formula (3.3) was used at the beginning to determine if a drought occurs or not in each time step, where: i: spatial unit (-), t: time step (-), X[i,t]: GWL (L), and τ[i]: groundwater threshold for unit i (L). This formula can be used for flux variables as well without any change.

$$\mathbf{1}_{\{X[i,t]<\tau_0[i]\}} \tag{3.3}$$

Formula (3.3) is an indicator function which equals one in case GWL of a unit is lower than the threshold, otherwise the indicator function equals zero. It is worth noting that the threshold varied over space but not over time. Thus, there was one fixed threshold for each unit. The fixed threshold was created based on the 365 GWLs of each unit over the period April 2018 - March 2019. The used percentiles were the same as those employed before (in the not spatial analysis) (20^{th}, 30^{th}, and 40^{th} one).

Afterwards, Formula (3.4) was applied to estimate drought duration for each drought event where, j: drought event (-), $L_T[j]$: duration of drought event j (T), L[j]: total number of of time steps of drought event j (-), and k: number of units for the study area (-). Thus, the duration of a drought event is the number of times steps in a row with at least one unit (grid cell) below the threshold.

$$L_T[j] = \sum_{t=1}^{L[J]} \max\left\{\mathbf{1}_{\{x[i,t]<\tau_0[i]\}}, \quad i = 1, k\right\} \tag{3.4}$$

Regarding drought area, firstly, the influence (weight) of each unit (u_i) to the total area was estimated by just dividing the area of one grid cell over the total study area. Because the study area (i.e. w_0) did not cover the whole rectangular frame of the raster interpolation image, units which were outside of w_0 were considered null. Those values were not taken into account for the estimation of total study area. Additionally, all the grid cells had the same area; thus, the influence (weight) of each unit was determined by the ratio one to total number of not null cells.

After that step, Formula (3.5) was used to determine the areal extent of each drought event, where: $A_T[j]$: percentage of area suffered from drought to total area of region, u[i]: area fraction (weight) of each unit, and the other variables were described above. Regrading the areal fraction, the sum of all weights is equal to one as can be seen in Formula (3.6).

$$A_T[j] = \frac{\sum_{t=1}^{L[j]} \sum_{i=1}^{k} u(i) \left(1_{\{X[i,t]<\tau_0[i]\}}\right)}{L[j]} \tag{3.5}$$

$$\sum_{i=1}^{k} u[i] = 1 \tag{3.6}$$

More specifically, the following process was followed: (i) firstly the sum area fraction u[i] of units which had GWL below threshold was estimated for the first time step of a drought event, (ii) the previous

step was repeated for all time steps of a drought event, and (iii) these values over all time steps of a drought event were aggregated and divided by the number of time steps of one drought event.

As for drought deficit, the average deviation (i) over area using Formula (3.7), and (ii) over the drought period were calculated using Formula (3.8). In Formula (3.7), D[i,t] is deficit volume of GWL for unit i and time step t, whereas the other variables have been described previously. In Formula (3.8), $D_T^*[j]$ is total groundwater deficit of a drought event j, whereas the other variables are known.

Weighted sum of units' deviation was estimated in each time step where the weight was the areal fraction of the unit. Deviation was the difference between GWL and threshold for each unit. Afterwards, the results of all time steps were summed and divided by the number of time steps of one drought event. Figure 3.6 illustrates basic concepts of spatial droughts and how drought deficit of one drought event was determined.

$$D^*[i,t] = \begin{cases} u[i]\,(\tau_o[i] - X[i,t]) & \text{for } X[i,t] < \tau_0[i] \\ 0 & \text{for } X[i,t] \geq \tau_0[i] \end{cases} \tag{3.7}$$

$$D_T^*[j] = \frac{\sum_{t=1}^{L[j]} \sum_{i=1}^{k} D^*[i,t]}{L[j]} \tag{3.8}$$

These formulas can also be used for flux variables (e.g. rainfall, recharge) instead of state ones with the difference that in Formula (3.7), the value is multiplied by the time step duration Dt, and in Formula (3.8), the total deficit is not divided by the number of time steps of the drought event.

The algorithms of these functions are attached in Appendix C.2. The aforementioned Formulas were not directly converted to code since it was not possible given that (i) pre-processing was required, and (ii) data was not stored in the format these Formulas suggest.

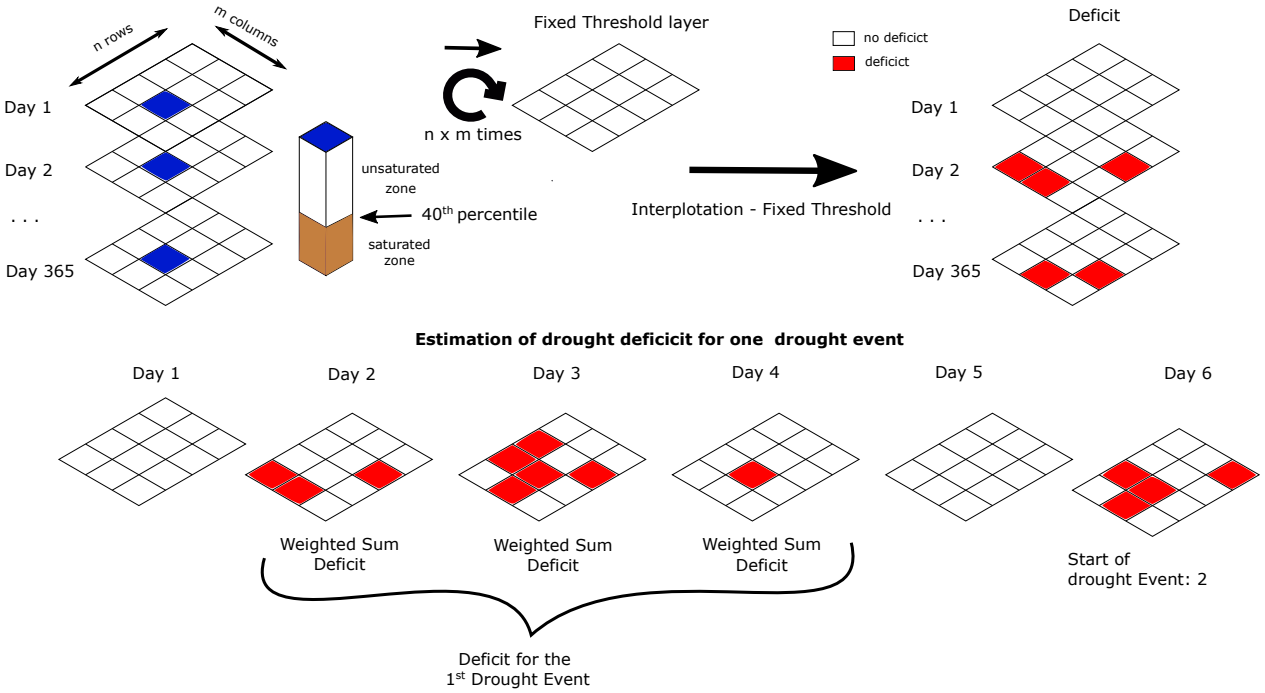


Figure 3.6: Schematic of the procedure for the estimation of spatial drought deficit

3.3.3 Frequency Analysis

GWL frequency distribution analysis was conducted for all districts based on modeled GWL for a period of 32 years. GWLs were negative and most theoretical distributions require only positive values in the sample. That means that the sample was unsuitable. In order to overcome this obstacle,

it was assumed that GWLs were positive (zero value was the same as NAP but negative values were considered positive).

Regarding the analysis, Block Minimum (Maximum) was used. More specifically, minimum GWL per month was estimated (based on the assumption mentioned before, that was not the minimum but the maximum). Besides monthly block, an analysis using annual blocks was performed as well. The fit of histogram (created by the minimum values of the previous blocks) to the following distribution functions were examined: uniform, normal, beta, weibull, t, generalized extreme value (GEV), log-normal, Pareto, exponential, and gamma. Residual Sum of Squares (RSS) was used as metric to determine how well the data fitted to the distribution. Using the best fitted theoretical distributions, the lowest GWL for return periods of one, 10, and 20 years were computed.

Drought frequency can also be estimated using SGI. This index can be used to estimate deficit and duration too. The method for estimating SGI is presented in Appendix B.6.1.

3.4 Modeling for the estimation of soil moisture

In order to estimate soil moisture, Urban Water Balance Model (UWBM) was employed which is a lumped conceptual model used for limited scale study areas and its forcing data consisted of precipitation, potential open water evaporation, and potential reference crop evaporation (see Figure 3.7). A thorough description of the steps for the estimation of Penman potential evaporation is presented in Appendix B.1. Based on Dryers (2009), reference crop grass transpiration was assumed to be 0.8982 * Penman evaporation (as cited in (Vergroesen, 2020)). UWBM was applied to each studied district in Leiden and to one region in Athens. There were differences among districts in Leiden which influenced the parameters' values of the model. There was not only one study period but three ones since each threshold method approach required a slightly different input period of data. For instance, regarding 30-day moving window quantile, input precipitation and evaporation spanned from the 17th March 2018 to the 15th April 2019. Modeled soil moisture via UWBM was used as input for the threshold method approaches to determine soil moisture drought events.

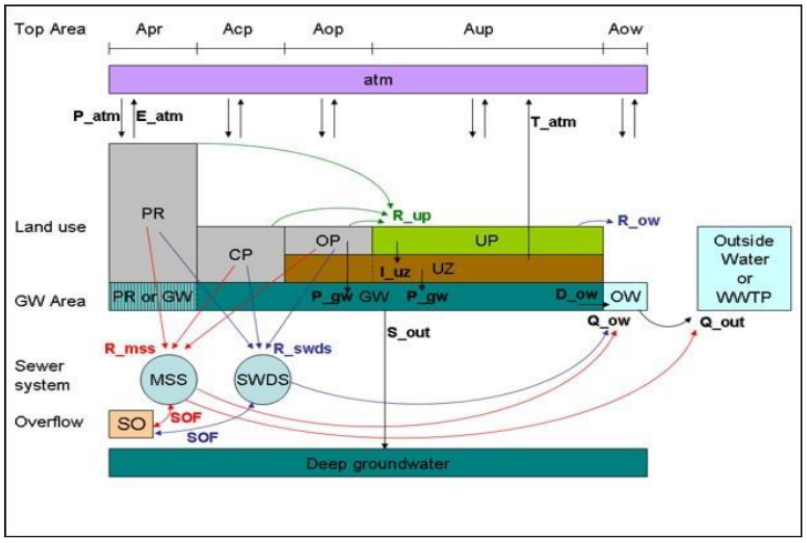


Figure 3.7: Overview of Urban Water Balance Model including its elements and involved fluxes (Vergroesen, 2020)

The model could handle hourly and daily data but hourly time scale was used. Almost half of the parameters of the model regarding fluxes needed to be provided on hourly time scale. That time scale is more suitable given that many hydrological processes last only some hours. Besides, more

parameters of the model were known at hourly scale and for their estimation at daily scale, saturation should have been considered. In that way, the possibility of using wrong parameter values increased. This was one of the reasons why daily scale was not employed.

Parameters of the model were based on the characteristics of the study area regarding (i) land cover, (ii) hydrology, and (iii) sewerage and storm water drainage system. Firstly, land use parameters are presented followed by the other ones.

Land use Parameters

In UWBM, land use is divided into Paved Roof, Closed Paved, Open Paved, Unpaved, and Open Water. Percentages of these land uses were estimated for each one of the four studied districts. Paved Roof is the area covered by buildings regardless their type. Data for it was derived from Wareco company. Closed Paved are the areas which are covered with impermeable materials and no infiltration occurs such as roads, and parking lots. Open Paved category includes pathways, permeable pavements, and roads with bricks where water can infiltrate. Regarding Unpaved Area, this class includes parks, urban forests, pitches, and in general green areas where infiltration occurs. OpenStreetMap was used to extract this data. Open Water category includes all canals, ponds, and ditches in study area. The Water Board of Rhineland, which is responsible for the city of Leiden, provided this data.

The difference of paved roof, unpaved, and open water from the total area is the sum of closed and open paved areas. However, it is not that straightforward to discern closed and open paved areas in an automatic way. Roads' area was available via TOP10NL (the cadastre agency in the Netherlands) but it was not useful since some roads were constructed by bricks and others by concrete; hence, they did not belong to one single category. For that reason, inspection, using Google Maps Streetview (Google, nd), was conducted and based on that, an assumption of the open paved share was made. For the closed paved, its percentage was one minus the previous percentage. For instance, regarding w_0 and w_1 districts, many roads were constructed with bricks; therefore, it was assumed the percentage of open paved to the sum of open and closed paved was 65%. District w_5 does not contain many brick roads, thus, the percentage of open paved areas decreased to 50%. That percentage decreased even more in w_7 district reaching to 40%. Additionally, for district w_5, its south part which is covered with agricultural fields was excluded of the land uses' percentage calculation since it could create problems to the application of UWBM (see Figure 3.8). Percentages of the four studied districts in Leiden are tabulated in Table 3.2.

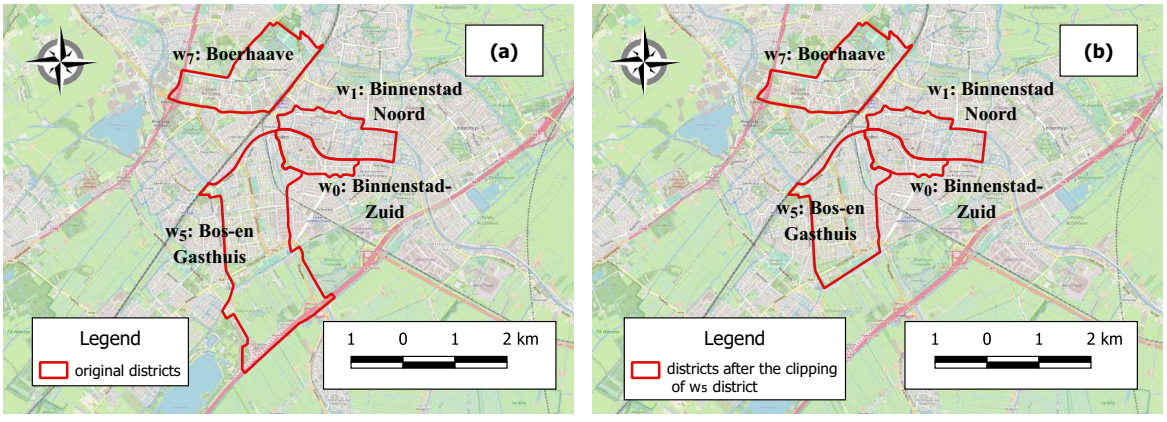


Figure 3.8: Removal of agricultural area from w_5 district so that UWBM could be applied; (a) before the removal, (b) after the removal

Table 3.2: Land uses for the four studies districts in Leiden based on the categorization which UWBM requires

	w_0	w_1	w_5	w_7
Paved Roof (%)	35	34	13	12
Closed Paved (%)	14	17	20	29
Open Paved (%)	21	25	21	19
Unpaved (%)	21	12	23	35
Open water (%)	9	12	23	5

Hydrological Parameters

Based on the land uses mentioned above, there are hydrological parameters related to them. More specifically, parameters regarding interception capacity of paved roofs, closed, open, and unpaved areas were used in that model (intstorcap_pr, intstorcap_cp, intstorcap_op, and intstorcap_up accordingly). For open paved and unpaved classes, there were also parameters regarding infiltration (infilcap_op and infilcap_up respectively). The unsaturated zone was modeled internally in UWBM; therefore, there were only two parameters to influence that (soil and crop type). There are 21 types of soil and category 9 was selected which corresponded to weakly loamy and fine sand to represent the study area. Concerning crop types, there are 16 available crop types in UWBM and in the current study grass was used. Based on soil and crop type, UWBM estimates soil moisture and capillary rise, using interpolation based on look-up tables of GWLs. There are also parameters regarding initial storage for all land uses but they were considered as zero.

Groundwater table and open water bodies are two critical components of UWBM. Regarding the former one, its parameters were: drainage resistance from groundwater to open water (w), seepage to deep groundwater flux (down_seepage_flux), hydraulic head of deep groundwater (head_deep_gw), and vertical flow resistance from shallow groundwater to deep groundwater (vc). Seepage to deep groundwater flux could be constant or variable. In the second case, it is estimated based on (i) the head difference between deep and shallow GWL, and (ii) vc. In the present study, constant downward flux was opted for. Another point is that initial GWL for the model (at t=0) was considered the first measured value to increase model performance. The parameters regarding open water were: (i) storage capacity of open water (storcap_ow), (ii) predefined discharge capacity from open water (internal) to outside water (external) (q_ow_out_cap), and (iii) percentage of open water which is above groundwater table. The third parameter was set as zero since GWL was higher than open water bodies.

Sewerage and storm water drainage system Parameters

Regarding sewer characteristics, part of area disconnected from sewer system is a parameter for paved, closed, and open paved areas. There was no information about it; therefore, these variables were not considered. Other parameters were: percentage of urban paved area with storm water drainage system (swds_frac), storage capacity of storm water drainage system (SWDS) (storcap_swds), storage capacity of mixed sewer system (MSS) (storcap_mss), rainfall intensity when SWDS overflow occurs on street (rainfall_swds_so), and rainfall intensity when combined overflow to open water occurs (rainfall_mss_so). For those parameters except the first one, their default values from UWBM were used; they are applicable to areas in the Netherlands. In addition, there are parameters regarding initial storage of sewerage and storm water drainage systems but they were considered as zero.

Calibration

Time series from April 2018 to March 2019 was used as input to the model (the first 8 months for calibration and the rest ones for validation). Spin-up period was not employed since available data was limited. Calibration of the model was conducted using (i) each monitoring well in each district, and (ii) combination of those ones (monitoring wells). It is worth noting that the reference for all used elevations in the model is the ground surface and not NAP. GWL from some wells showed erratic

behavior and UWBM was not able to reproduce their irregular pattern. There were abrupt drops or large fluctuations and none of the parameters were able to handle those ones. Consequently, wells with poor performance were dismissed. For each district, the parameters which led to highest performance or to the best approximation of reality were used for the estimation of soil moisture. Regarding the details of calibration, Monte Carlo approach was used for parameters' sampling using 1000 iterations.

Trial and error approach was employed before Monte-Carlo simulation to test how the model reacted to different parameters' ranges. Nash-Sutcliffe Efficiency of log values (NSE_{log}) (see Formula 3.9) was used as objective function since the focus of this research was only on low values and not peaks.

$$NSE_{log} = 1 - \frac{\sum_{t=0}^{T}(logGWL_o^t - logGWL_m^t)^2}{\sum_{t=0}^{T}(logGWL_o^t - log\overline{GWL_o})^2} \tag{3.9}$$

where: GWL_o^t: observed GWL at time step t, GWL_m^t: modeled GWL at time step t, and $\overline{GWL_o}$: mean observed GWL.

Not all the aforementioned parameters were calibrated but only four of them: w, down_seepage_flux, storcap_ow, and q_ow_out_cap. The rest ones were considered as fixed for all districts and tabulated in Table 3.3. Those parameters were not calibrated since most of them were not related directly to groundwater regime and they could not affect it considerably. Regarding parameter head_deep_gw, in Table 3.3 its value for well L-PB46 is presented. This value was different for each well. The ranges of the calibrated parameters were not constant over studied regions and are included in Results section 4.2.

The values of the not calibrated parameters and ranges of values for the calibrated ones were based on literature, the creator of UWBM (Toine Vergoesen), but also on trial and error approach in the UWBM. For instance, regarding the infiltration capacity of open paved areas, area covered with permeable pavement was considered representative for this land use category. Based on the experiments of Boogaard et al. (2014), infiltration capacity of permeable pavements varies from 29 to 342 mm/h and the elapsed time of the last maintenance or first installment plays an important role to that value. In the current analysis, an average of 100 mm/h was used as infilcap_op (Table 3.3). Assumptions were made regarding parameters' range in case of lack of relevant information.

Table 3.3: Fixed Parameters of UWBM for district w_0

Distrct w0			
Parameters	Value	Parameters	Value
intstorcap_pr (mm)	4	frac_ow_aboveGW (-)	0
intstorcap_cp (mm)	5	swds_frac (-)	0.2
intstorcap_op (mm)	10	storcap_swds (mm)	2
intstorcap_up (mm)	20	storcap_mss (mm)	9
infilcap_op (mm/d)	100	rainfall_swds_so (mm/d)	16.8
infilcap_up (mm/d)	30	rainfall_mss_ow (mm/d)	6.7
head_deep_gw (m)	0.920	vc (d)	1000

Another point about calibration is that only for district w_1, upsampling using the mean values (from hourly scale to daily one) was conducted before the estimation of NSE_{log} since the results were much better that way. Additionally, daily and not hourly GWLs were used as measurements for that specific district. For the other districts, no upsampling was applied and hourly measurements were employed.

Instead of providing just one solution, the uncertainty interval of 80% was estimated as well which is quite useful to the evaluation of model performance. This method was applied only for feasible solutions where $NSE_{log} > 0.2$ except for the wells in district w_1, two wells in district w_5 (L-PB71 and L-PB82), and one well in district w_7 (L-PB14). For those wells, a lower NSE_{log} value was used since

their model performance was poor and there was no solution with $NSE_{log} > 0.2$. For each feasible solution the ratio of its NSE_{log} to the sum NSE_{log}, derived from all feasible solutions, was estimated. These ratios were called 'weights' in the following part of this method. Next, the solutions were sorted in ascending order by GWL at a specific time step. The weights were sorted as well based on the GWLs. Then, the cumulative sum of weights were estimated. Using 10% and 90% markers of the accumulated sum, feasible solutions were filtered out and only solutions within the following percentiles ($> 10\%$ and $< 90\%$) were included in the uncertainty interval. Not the exact values of 10% and 90% regarding the cumulative weight were used but the values which were closer to them. The aforementioned steps of this method were only for one time step. This process was repeated for all time steps. In the current study, time step was hourly; therefore, this process was time-consuming.

Modeling Sensitivity

Sensitivity analysis was applied using three techniques. In the first approach (Method A), all parameters except one were kept constant (a parameter set with the highest performance or the best fit in calibration was used). For the non constant parameter, Monte Carlo sampling was conducted and NSE_{log} values were computed. Regarding the second approach (Method B), for the one solution either with highest performance or best fit, NSE_{log} was estimated for the cases that each calibrated parameter increased and decreased by 10%. Regarding the third approach (Method C), how the performance of model changed based on the range of parameters was investigated. More specifically, two conditions were examined: when the (i) upper range, and (ii) lower range of each parameter increased and decreased respectively by 10%. Sensitivity analysis was conducted only either for the monitoring well with the highest performance or that one which approximated reality best for each district. It is worth noting that 500 iterations were used for sensitivity for the methods (A and C) since they were computationally expensive.

3.5 Exposure to soil moisture droughts

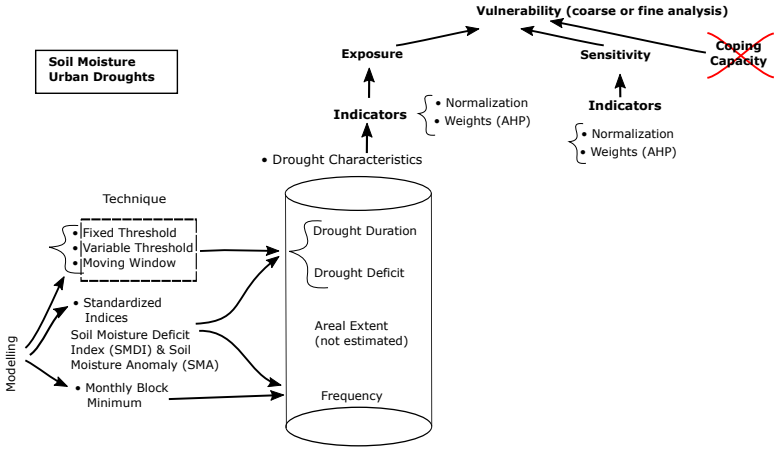


Figure 3.9: Roadmap of Methodology regarding Soil Moisture Urban Droughts

The techniques (fixed, variable, and moving window threshold) and percentiles (20^{th}, 30^{th}, and 40^{th} percentile) used for estimation of threshold function for GWL were applied for the case of soil moisture as well. Correlation between drought duration and deficit was examined using Pearson coefficient. The influence of pooling to both deficit and duration was investigated too. Spatial analysis was not conducted since soil moisture values represented the whole district; therefore, no distributed data was available. There were no point measurements of soil moisture content as the case of GWL; thus, spatial interpolation was not applied for the creation of distributed data. As for frequency distribution analysis of soil moisture droughts, it was estimated using the minimum of monthly blocks. Long soil

moisture data was produced via UWBM. The left part of Figure 3.9 summarizes the techniques used to determine drought characteristics for soil moisture droughts.

3.5.1 Standardized Indices

Besides threshold method techniques, two indices were computed to estimate soil moisture drought severity: (i) Soil Moisture Deficit Index (SMDI), and (ii) Soil Moisture Anomaly (SMA). SMDI (Narasimhan and Srinivasan, 2005) was calculated using Formula (3.11). Estimation of soil moisture deficit via Formula (3.10) was required beforehand.

$$SD_{i,j} = \begin{cases} \frac{SW_{i,j} - MSW_j}{MSW_j - \min SW_j} \times 100 & \text{if } SW_{i,j} <= MSW_j \\ \\ \frac{SW_{i,j} - MSW_j}{\max SW_j - MSW_j} \times 100 & \text{if } SW_{i,j} > MSW_j \end{cases} \tag{3.10}$$

where: $SD_{i,j}$ is soil moisture deficit for a 7-day period for the year i and week j (%), $SW_{i,j}$ is the mean weekly soil moisture (mm), MSW_j is the long-term median available soil moisture of week j over all years (mm), and max (min) SW_j is the long-term maximum (minimum) soil moisture of week j over all years (mm).

$$SMDI_j = \begin{cases} \frac{SD_j}{50} & \text{if } j = 1 \\ 0.5 \times SMDI_{j-1} + \frac{SD_j}{50} & \text{if } j > 1 \end{cases} \tag{3.11}$$

$SMDI_j$ is Soil Moisture Deficit Index for week j and the rest variables are known. The second estimated index was Soil Moisture Anomaly (SMA) (European Drought Observatory (EDO), 2019) via Formulas (3.12), and (3.13) expressing the difference of normal conditions in units of standard deviation.

$$\text{SMI} = 1 - \frac{1}{1 + \left(\frac{\theta}{\theta_{50}}\right)^6} \tag{3.12}$$

where θ is daily soil moisture, θ_{50} is the mean between field capacity and permanent wilting point. The later ones were known based on crop and soil type of UWBM.

$$\text{SMA} = \frac{\text{SMI}_t - \overline{\text{SMI}}}{\delta_{\text{SMI}}} \tag{3.13}$$

where: SMI_t is the average SMI per 10 days, t is the 10-day interval, $\overline{\text{SMI}}$ is the long-term average, and δ_{SMI} is the standard deviation.

Other available techniques to estimate soil moisture droughts were not applied since they did not fit to the current analysis. More specifically, spatial techniques which could calculate spatial soil moisture anomalies were not considered since they are not appropriate for urban environment. Cities are not homogeneous and spatial resolution of satellite soil moisture products is still coarse impeding urban environment analysis. Additionally, Palmer Drought Severity Index is a popular index but was not used since it applies only to the U.S.A. Soil Water Deficit Index (SWDI) (Martínez-Fernández et al., 2015) was not estimated since it is mostly used for agricultural droughts where soil moisture less than field capacity is considered as drought. It is often used in cases where the objective was to maximize crop production. However, that did not fit to the goals of this analysis where the crop type is mostly grass and trees which do not bear fruit. SWDI was similar as Soil moisture Index by Hunt et al. (2009) (the differences were some coefficients and intercepts to change the range values).

3.6 Indicators

The following steps were applied regarding indicators:

- Collecting Data
- Pre-processing so that available indicators can be produced
- Checking missing values
- Checking outliers (Inter-Quartile Range -1.5 IQR)
- Multi-collinearity among indicators (Kendall's Tau for small samples)
- Normalization (it is required since it ensures that different variables can be used at the same scale)

Multi-collinearity applied only for for social sensitivity indicators since the scale of recording (buurt (neighborhood)) allowed that. If data was not recorded in that level but only in the coarser one (i.e. wijk), data had a limited sample size not letting multi-collinearity be tested. A detailed description of the social sensitivity indicators is attached in Section B.9.1 in Appendix B.9. The correlation of social sensitivity indicators are presented in Appendix B.9.2.

Outliers may distort the min-max normalization since the range (min-max) would be quite large and most normalized indicator values would be similar (Commission et al., 2008). Some techniques to handle outliers are: removing or keeping them, exporting findings with them and without, transformation (e.g. log function), winsorization (transforming the outliers to a specific percentile of data), and robust techniques such as bootstrapping. This was not a complete list; there are also non-parametric approaches. Social sensitivity indicators were transformed using log function but there were still some outliers. For that reason, keeping the outliers and using the non transformed data was the final option.

Checking the outliers was a way to ensure that there was no wrongly recorded or measured value. Dropping one value since it is an outlier is not suggested; more investigation is required before that since this outlier may provide useful information. In the current study, just dropping an outlier was not possible in social indicators since all administrative units (buurten) needed to have values before normalization. However, an outlier in modeled subsidence was excluded but it was not certain whether it was an error.

The selected indicators were also S.M.A.R.T. This term was introduced as a management tool for setting objectives by Doran (1981) but it has been incorporated in other components of management such as monitoring and evaluation using indicators. S stands for Specific (the indicators should not be broad but as narrow as possible), M for Measurable (the indicators should be quantified and measured in a consistent way so that comparison can be possible), A for Achievable or Attainable (the data for this indicator can be collected or measured effortlessly), R stands for Relevant (it should be highly related to the parameter for which is used), and T for Time-bound (the frequency which the indicator is measured or collected should be known) (Global Environment Facility, 2010, pp. 28–29). The S (Specific) component tends to be considered as the most difficult to be met out of the five principles in S.M.A.R.T. approach which described above.

The used indicators in this study were split into metric and categorical ones since the method for normalization was not the same. Categorical indicators were classified into two categories; ordinal and nominal ones. The former were ranked classes whereas the latter were descriptive ones. In the current analysis, there were no ordinal indicators. For metric indicators' normalization, min-max method was used; maximum vulnerability was represented by one and lowest vulnerability by zero. This process applied for both exposure and sensitivity indicators. The direction of the indicators were considered ensuring that high values are related to negative conditions and increased vulnerability. Therefore, regarding indicators with positive relationship with vulnerability (e.g. buildings constructed before 1960 which were probably constructed on wooden poles) Formula (3.14 left part) was used, whereas for variables with negative relationship (e.g. mean income per resident) Formula (3.14 right part).

$$X_{Norm_i} = \frac{X_i - X_{Min}}{X_{Max} - X_{Min}} \qquad X_{Norm_i} = \frac{X_{Max} - X_i}{X_{Max} - X_{Min}} \tag{3.14}$$

where: X_i: is an individual element of the indicator sample, X_{Min}: the lowest value of the indicator, X_{Max}: the highest value of the indicator, X_{Norm_i}: the normalized value of the individual element.

As for nominal indicators, min-max method could not be applied; thus they were converted to ordinal classes and then transformation of their classes to a range between zero and one was conducted. For instance, soil was a nominal sensitivity indicator and was transformed to an ordinal one; peat soil was classified as 'rather negative', deposits as 'negative', and old dunes (consisted of sand) as 'positive' regarding groundwater urban droughts. Then, values in the range [0-1] were assigned based on the classes.

3.7 Estimating Vulnerability

One of the main objectives of this study was the estimation of vulnerability. For both groundwater and soil moisture droughts, over two different spatial levels of estimation were performed (coarse and fine one). Not all vulnerability indicators were spatially distributed in Leiden; for those which were not, interpolated values were used in fine resolution analysis. Differences between the two types of analysis (coarse and fine) were investigated. For coarse resolution, the mean value of each indicator standardized by its district area was used. On the contrary, for fine resolution, spatial overlay was implemented for layer indicators. The indicators' weights for vulnerability aggregation were estimated via AHP (see Appendix E). Their ranges and the selected values are presented in Figure 3.10-a, b. The histograms of indicators' weights are attached in Section E.2 in Appendix E, highlighting how much the drought experts agree or contradict each other. Besides the weights presented in Figure 3.10, vulnerability was also estimated using extreme cases of weights (see Appendix G).

It is worth noting that experts assigned a higher weight to duration than deficit for groundwater droughts which was logical. Number of dry days determines the height difference between top of timber poles and GLG (minimum GWL) (National Institute for Public Health and the Environment, 2019b); therefore, it influences oxidation and subsidence. For the rest indicators, there was no related research.

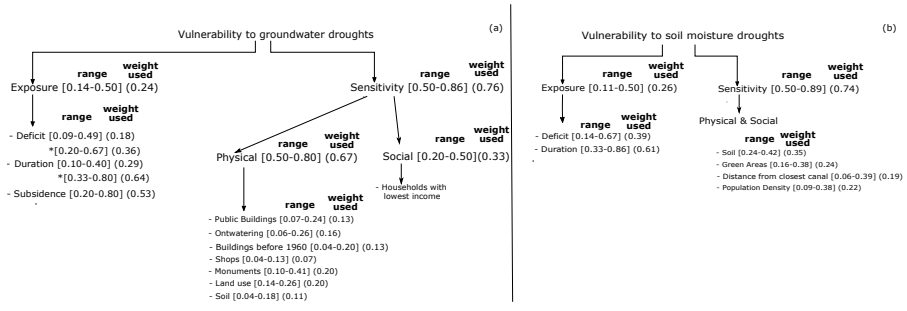


Figure 3.10: Vulnerability Aggregation for (a) groundwater, and (b) soil moisture droughts with indicators' weights. The brackets show the range of the weight and the parentheses the value used for vulnerability estimation. The values with * are about the case for which subsidence was excluded from exposure in groundwater droughts vulnerability

The vulnerability was also estimated re-distributing the priorities (i.e. the indicators' weights) in Appendix H.

Rejecting techniques to determine indicators' weights

Entropy evaluation and Garrett ranking were not selected for indicators' weights estimation in the current analysis. In the case of entropy evaluation, subjectivity is minimized since the weights are based

on data variation and drought experts' contribution is not required. This method could be applied using only the social sensitivity indicator (i.e. percentage of households belonging to households with the lowest 40% income nationwide). It would not be suitable for the other indicators since their number of districts (spatial units) was insufficient for that analysis and therefore, not applied. Nonetheless it could have been used in case the analysis was based on buurten (neighborhoods) and not wijks (districts) since the number of study regions was high. Garrett ranking is a commonplace technique for indicators comparison. Based on it, factors' weights can be created. However, it requires a large number of expert respondents which was not met in the current research leading to dismissal of this method. Therefore, indicators' weights were determined by AHP only.

Aggregation

Vulnerability estimation was just based on exposure and sensitivity. Coping capacity was excluded from this process due to the lack of appropriate variables. Coping capacity is mostly related to knowledge, technology, and institutions (Fritzsche et al., 2014) but given that all districts belong to the same municipality and are located close to each other, all three elements were similar in all districts. After normalizing the indicators of each vulnerability component (exposure, and sensitivity), aggregation was applied. Then, these aggregated components were combined to determine the vulnerability composite map. Weighted arithmetic aggregation (weighted average) was used due to its simplicity and ability to provide accurate and intuitive results. Another common technique is weighted geometric aggregation but it was not selected since (i) individual indicators with zero values are not allowed because the whole composite would be zero, and (ii) there is a strong bias towards low values.

Another worthwhile point about aggregation is that upsampling was applied for indicators with coarse resolution such as subsidence. In that way, the finer sensitivity resolution of other indicators was not missed in the aggregation part to determine vulnerability. When spatial aggregation was performed, result resolution was the coarsest of the aggregated ones.

3.7.1 Groundwater Droughts

Exposure

Regarding exposure, ideally the four drought characteristics should be included but only two of them (deficit and duration) were incorporated in the current analysis. Except those two indicators, the third one was subsidence. It is a peculiar variable. On the one hand, it can be included since it can be considered as an indirect measure of drought severity. On the other hand, comparison of subsidence to deficit and duration is not straightforward since the droughts can cause subsidence. Therefore, subsidence could be excluded. Drought is not the only reason for subsidence, though. For the aforementioned reasons, vulnerability was estimated based on the following cases: (i) using modeled, (ii) and measured subsidence, and (iii) excluding it. Data regarding subsidence is presented in Appendix B.4. Vulnerability was also estimated based on different techniques to calculate drought deficit and duration. Whether and to what extent the selection of drought estimation method affects vulnerability was investigated. More specifically, duration and deficit were derived by (i) variable threshold of 30^{th} percentile as threshold, (ii) the same as (i) in conjunction with 10-day pooling, and (iii) using the median GWL as threshold.

Sensitivity

As for sensitivity, it was separated into physical and social categories. A detailed description about the social sensitivity indicators tested for groundwater droughts is attached in Appendix B.9. Seven variables were used for physical one: public buildings; ontwatering which is the difference between ground surface and GLG (minimum GWL); buildings erected before 1960; shops; monuments; land use; and soil. For the coarse analysis, the mean normalized values of those variables (over the area of each district) were used. As for land use and soil, their classes and the relevant weights are presented in Table 3.4. These weights are not based on drought experts.

Table 3.4: Weights for 'land use' and 'soil' classes

Land use		Soil	
Class	Weight	Class	Weight
built up	1	trench and trench deposits	0.4
water	0	old dunes	0.2
company sites	0.8	Holland peat	1
highway	0.8	deposits	0.6
recreation	0.2	tidal basin / tidal deposits	0.6
semi-built	0.5	river bed	0
agriculture	0.2		
railway	0.9		
forest	0.2		

Social sensitivity consisted of only one indicator, percentage of households with the lowest 40% income nationwide (low_inc_H). Shops, monuments, and public buildings were derived from openstreetmap.org whereas buildings before 1960, and soil from Wareco company. More specifically, the following were considered to be public buildings: schools, university and colleges, hospitals, and town halls. Regarding shops, what was used was the whole building where a shop was placed, not the area of the shop individually, a method which is not necessarily valid. Ontwatering was estimated based on minimum measured GWL and ground surface (see Appendix B.5.1). For this variable, no values were estimated for the area covered by canals. To comply with the other sensitivity indicator layers, zero values were assigned to the canals' locations. In addition, pixels with negative ontwatering (errors) were set to 0, and pixels with value higher than 2 m were set to 2 m. In that way, normalization provided satisfactory results since the range did not include extreme values. Otherwise, there would be no difference over the study area due to extreme smoothing. Land use data was derived by Publieke Dienstverlening Op de Kaart (PDOK) (2015). Soil and land use were the only nominal sensitivity indicators for groundwater droughts analysis.

Ignored Sensitivity Indicators

Some physical sensitivity indicators for groundwater droughts which were ignored in the current analysis are mentioned below. Firstly, the length of drainage pipes and the frequency of their maintenance were not considered. The higher the length and the less frequent the maintenance, the more vulnerable the city is. Drainage pipes are clogged when GWL is below them since oxygen enters the pipes and iron oxide can be created (Hoogvliet et al., 2012). Secondly, high voltage grid power lines could also be a sensitivity indicator. An explanation is required since it is not straightforward. The power lines release heat when GWL is low and overheating can occur; in order to avoid it, transport of electricity is reduced leading to power shortage (Hoogvliet et al., 2012). There was no relevant information; hence, this indicator was ignored. Thirdly, distinction of roads was not considered; construction and foundation of highways are different than those of local roads. The social damage of the two road classes differs too.

Indicators' Sensitivity Analysis

Sensitivity analysis for the case of removing indicators was also estimated (i.e. the degree to which the removal of each indicator influenced the vulnerability output). The current sensitivity analysis was not the same as that Wu et al. (2017) performed in Guanzhong Plain, China for agricultural droughts; they focused only on the statistics. Spatial differences by removing each indicator were underlined in the current analysis.

For each indicator removal, the vulnerability output including all indicators was used as reference. More specifically, the rate of change was estimated using Formula (3.15) where Vulner: Vulnerability. For all vulnerability outputs, exposure indicators were determined using variable threshold of 30^{th} percentile with 10-day pooling. The indicators' weights were estimated afresh for each removal of indicator by removing the corresponding row and column from the corresponding combined pair-wise comparison matrix.

$$Rate\ of\ change = \frac{Vulner_{excluding\ an\ indicator} - Vulner_{all\ indicators\ included}}{Vulner_{all\ indicators\ included}} * 100 \quad (3.15)$$

In addition, how much the weights of soil and land use classes affect the final vulnerability product was investigated. Increases by 0.1, 0.2, and 0.4 were used for each class in soil except peat whose vulnerability was the maximum possible (i.e. 1). The same approach applied to all classes in land use and at this case 'built-up' class was exempted. A difference compared to 'soil' indicator was that an increase of 0.4 for the following classes: company site, highway, and railway were not possible since the total weight would exceed 1; therefore, a decrease by 0.4 was applied.

Sensitivity of vulnerability indicators' and their aggregated components' weights was also estimated; increases by 0.1, 0.2, and 0.3 were investigated. When the weight of one indicator (or component) increased, the weights of the rest ones were recalculated so that their sum weight equaled 1. For instance, if one indicator increased by 0.1, each of the rest ones at that level decreased by 0.1 over the their number; the reduction was equally distributed over the rest indicators.

3.7.2 Soil Moisture Droughts

For soil moisture, there were no distributed results for drought characteristics (deficit and duration) since UWBM was used (it is a lumped model). Therefore, exposure resolution was coarse but this did not apply for sensitivity since its sensitivity indicators were distributed.

The four used sensitivity indicators were: soil, distance from canals, green areas, and population density. Firstly, regarding soil, some soil types may be more permeable than others (e.g. old dunes than peat); thus, they cannot hold water for the plants. That means that parks that lie on permeable soil types are more sensitive. Weights for soil classes are presented in Table 3.5. Secondly, green areas were highly sensitive to soil moisture droughts than the rest area. A combination of parks and trees was used. For trees a buffer of 10 m was used in order to convert point data to polygon one. Thirdly, distance to the closest canal is another sensitivity indicator. The further it is, the shallower the GWL since there is a GWL bulge between two adjacent canals. As a result, more groundwater evaporation can occur. Groundwater evaporation is the process when capillary rise tales place from a shallow aquifer to the unsaturated zone in order to satisfy the increased demand of evaporation and transpiration in dry periods (Yeh and Famiglietti, 2009). In that way, it is less possible for the plants/trees to wilt out. Groundwater evaporation does not occur when the distance between root zone and groundwater table is quite large. Fourthly, parks are important for social interaction. The higher the population density, the higher the need for green areas (from the residents) and the more sensitive the area is.

Table 3.5: Weights for soil classes regarding soil moisture droughts

Class	Trench	Old dunes	Holland peat	Deposits	Tidal Basin	River bed
Weight	0.5	1	0.2	0.5	0.5	0

In addition, how much the selected drought identification method influences exposure and vulnerability to soil moisture droughts was investigated. In particular, the following three techniques: fixed threshold of 30^{th} percentile in conjunction with 15-day pooling, SMDI, and SMA were compared to fixed threshold of 40^{th} percentile without pooling. It is worth stressing that for the case of 30^{th} percentile and pooling of 15 days, the maximum deficit and duration were used and not the mean ones. On the other cases, the mean values were employed.

Indicators' Sensitivity Analysis

Sensitivity analysis for the case of removing indicators was also applied for the vulnerability to soil moisture droughts. The technique was as the same as that mentioned for the case of groundwater droughts. In addition, the impact of soil classes' weights to vulnerability were explored. Increases or

decreases by 0.1, 0.2, and 0.4 were used for each soil class's weight. The difference with groundwater drought analysis is that even the class with the maximum possible weight included in the analysis (i.e. the class of 'old dunes' for this case having a value of one). The increase or decrease of the weight was determined by the class weight proximity to maximum and minimum possible value. For instance, if a weight increase could lead to exceedance of one, it was dismissed and a decrease was applied instead.

Last but not least, sensitivity of vulnerability indicators' and their aggregated components' weights were estimated.

Chapter 4

Results

Section 4.1 refers to results of groundwater dry conditions in Leiden stressing the differences among the techniques used to determine drought characteristics. Calibration of UWBM for the studied districts and its sensitivity are presented in Section 4.2. Section 4.3 highlights the results of soil moisture dry conditions. The vulnerability of Leiden to groundwater and soil moisture urban droughts are presented in Section 4.4. Last but not least, an evaluation of the results is presented in Section 4.5.

4.1 Groundwater Droughts

This part is separated into the results about (i) drought deficit and duration (Section 4.1.1), (ii) spatial analysis (Section 4.1.2), and (iii) drought frequency estimation (I.1). Regarding deficit and duration, there is further categorization. The results using fixed, variable, and moving window threshold based on the data of monitoring wells are presented in Section 4.1.1.1. Then, the influence of pooling on drought duration and deficit is highlighted (Section 4.1.1.3). The results using median GWL as threshold is presented in 4.1.1.4. The outcomes of SGI are attached in Appendix B.6.2.

Key results are summarized at the end of each subsection. Sometimes, the key results are derived based on plots included in the Appendix. Additionally, it is important for the reader to have in mind the definitions of important terms in order to follow the arguments presented in this thesis. These terms are: exposure, sensitivity, coping capacity, and vulnerability (see this part here).

4.1.1 Drought deficit and duration

4.1.1.1 Districts w_0 and w_1

Table 4.1: Statistics for the wells of w_0 district using fixed, variable, and moving window threshold for groundwater drought identification (n: number of identified drought events)

		Fixed Threshold					Variable Threshold					Moving Threshold				
		duration (days)			deficit (m)		duration (days)			deficit (m)		duration (days)			deficit (m)	
Well	percentile	n	mean	std dev	mean	std dev	n	mean	std dev	mean	std dev	n	mean	std dev	mean	std dev
L-PB42	20	4	18.3	23.9	0.064	0.050	10	8.3	6.3	0.045	0.029	10	4.4	1.4	0.038	0.024
	30	6	18.0	23.5	0.108	0.070	13	9.3	6.8	0.053	0.037	15	5.1	2.8	0.050	0.038
	40	8	18.0	21.4	0.137	0.087	15	10.5	13.3	0.070	0.049	13	8.5	3.5	0.080	0.046
L-PB44	20	2	36.5	46.0	0.110	0.147	8	10.5	5.3	0.033	0.012	8	5.0	1.4	0.017	0.014
	30	3	36.7	39.3	0.133	0.122	10	14.0	6.8	0.052	0.021	9	7.0	3.0	0.023	0.017
	40	5	29.2	40.7	0.114	0.125	9	20.2	26.4	0.060	0.032	11	8.5	4.0	0.031	0.018
L-PB46	20	3	24.0	28.6	0.086	0.086	15	5.3	3.2	0.017	0.012	15	5.3	3.2	0.017	0.012
	30	6	18.0	30.2	0.062	0.089	12	10.2	6.4	0.033	0.016	14	5.4	2.7	0.022	0.013
	40	8	18.0	26.2	0.068	0.083	14	10.9	8.9	0.041	0.025	18	6.2	3.7	0.029	0.018

Table 4.1 presents statistics of identified drought events for the wells which lie within w_0 district (Binnenstad Zuid) employing (i) fixed, (ii) moving window monthly percentile (variable threshold), and (iii) moving window threshold. GWLs of well L-PB44 and its three threshold functions using three different percentiles (20^{th}, 30^{th}, and 40^{th}) are illustrated in Figure 4.1. Firstly, the number of drought events was limited using fixed threshold for the three wells weakening a statistical interpretation of the data. The findings may not hold for other cases where the number of drought events is larger; the possibility of error is high.

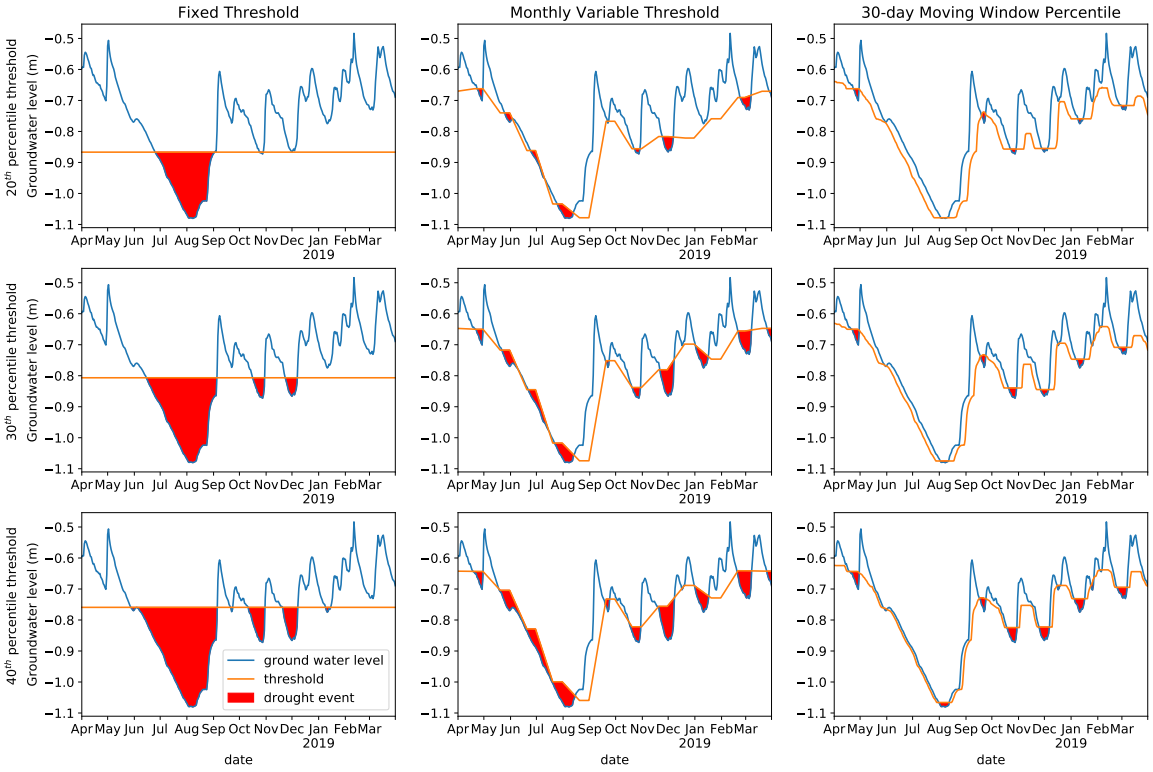


Figure 4.1: Identified groundwater drought events using fixed, variable, and moving window threshold for three different percentiles for well L-PB44

The mean drought duration employing fixed threshold was almost constant for well L-PB42 for the three percentiles, whereas it decreased for the well L-PB46 as percentile increased from 20^{th} to 40^{th} one. More minor drought events identified as percentile rose reducing the mean value. Standard deviation for all wells for all percentiles was larger than the corresponding mean which is an indication that there was high uncertainty. Regarding mean drought deficit of GWL for fixed method, it was much higher in absolute value than the other two methods due to the extreme dry conditions in summer 2018. In addition, there was no clear relationship between deficit from fixed threshold and percentile. From 30^{th} to 40^{th} percentile for L-PB44, mean deficit decreased but for the other two wells increased. It is worth noting that standard deviation for deficit in all wells and percentiles was almost as large as the mean value, which reduced the robustness of the results.

Regarding the use of variable threshold, the number of identified drought events was larger than using the fixed one but there was not a clear relationship with the increase of percentile. In wells L-PB44 of w_0 district (see Table 4.1), and L-PB56 of w_1 district (see Table B.2 in Appendix B.3), more drought events were recognized with the 30^{th} percentile and not the 40^{th} one compared to the rest wells at these percentiles. Mean drought duration using the variable threshold increased in all wells of w_0 and w_1 district (except L-PB49) as percentile did. The maximum increase from 30^{th} to 40^{th} percentile happened for L-PB56 well of w_1 district (from 12.0 to 25.3 days). However, this well had the highest standard deviation for 40^{th} percentile. Another point is that in all wells the relationship of standard

deviation of duration and percentile was positive.

The detailed results of districts w_1, w_5, w_7 are presented in Appendix B.3.

4.1.1.2 Summary of fixed, variable, and moving window thresholds

This part is about the main results regarding the three different techniques applied for drought identification. They are based on the four districts. Fixed threshold was able to identify the severity of the drought event of summer 2018 much more satisfactorily compared to the other two techniques. However, the limited number of identified drought events led to high standard deviation for drought deficit and duration reducing the robustness of the results. The identified drought events were not many since this method is suitable only for dry season; it missed drought events in wet season. Variable method did not have this drawback and identified drought events over the whole study period.

Regarding variable threshold, the fact that only data of one year data was used decreased the performance of that technique. If more data was available for summer periods before 2018 when GWL was higher, this method would have used this data and as a result variable threshold function would be higher in summer 2018. In that way, both duration and deficit of the identified summer drought in 2018 would be much greater. Additionally, higher consistency over percentiles is observed using variable threshold (e.g. the most severely drought stricken well regarding deficit was constant over percentiles in districts w_5 and w_7).

Moving window threshold identified drought events when GWL was much lower than the closely neighboring time steps of the time-window. As a result these drought events were short. In general, the results were not satisfactory (this threshold identified a minor drought in summer 2018 which does not represent the reality). This technique would work satisfactorily for hydrological variables which change rapidly (e.g. streamflow) but it is not suitable for groundwater since changes regarding groundwater status take much more time. Furthermore, using only data from one year did not influence the performance of this technique since this technique is only based on the data which time window covers (in this case, 30 days). Consequently, even if there was long data available, the results regarding the summer drought in 2018 would be the same.

For the reasons mentioned above, variable threshold was selected as the most suitable approach (compared to the other two ones) to identify drought events. Regarding percentiles, the 30^{th} one was selected. However, a combination of fixed and variable threshold can provide a deeper insight into the drought status in one area. Drought characteristics derived only by variable threshold were used in the following steps: (i) mapping of monitoring wells, (ii) pooling over all districts, and (iii) vulnerability estimation. The reasons why the other percentiles were rejected were: (i) results regarding 40^{th} percentile had large standard deviation, and (ii) drought characteristics (duration and deficit) derived by 20^{th} percentile had quite low values; hence this percentile was not able to incorporate long drought events.

Key Results

- Using fixed threshold method, the number of drought events was limited which hindered a statistical analysis.
- There is no clear positive relationship between each of the following elements (drought deficit, duration and number of events) and percentile for all drought identification techniques used (fixed, variable, and moving window threshold). Increase of the percentile did not entail increase of mean deficit and duration for all monitoring wells.
- Variable threshold showed higher consistency for drought deficit over thresholds than the other techniques (e.g. the monitoring well with the highest deficit was the same over the three percentiles in many wells).

- Variable threshold performed poorly in summer missing the severity of the drought event in 2018; the reason was lack of data.
- Moving window is not a suitable technique for groundwater droughts identification.

4.1.1.3 Influence of pooling

The following part focuses on the differences of mean drought deficit and duration for the three approaches used (fixed, variable threshold, and moving window) applying pooling for a series of inter-event time periods. Standardized mean deficits using pooling of five, seven, 10, 15, 20, 25 days for monitoring well L-PB42 are presented in Figure 4.2. Regarding fixed threshold method and 30th percentile, for inter-event time of seven days, standardized mean deficit was negative. The explanation is how pooling regarding drought deficit was applied. Only the maximum deficit was kept from two or more large pooled events. This led to slightly lower mean deficit of that inter-event time compared to the case where no pooling was applied. In that way, standardized mean deficit was negative.

In fixed method there was a significant rise from seven to 10 days for 20th, and 30th percentile whereas for 40th percentile the increase was not that steep. For the same method, the mean standardized deficit was constant after 10 days for 20th, and 30th percentile but not for the 40th (it kept increasing).

Applying pooling for variable threshold, the behavior of mean standardized mean deficit was almost similar as that for fixed threshold but there was a slight increase from 20 to 25 days for all percentiles. There was also convergence for 20th and 40th percentile at 25 days. Another striking result is that for moving window, the cases of all percentiles increased smoothly compared to fixed and variable threshold. In that method, there were droughts to be pooled at different inter-event times compared to the other two techniques (this is why a large plateau was not observed).

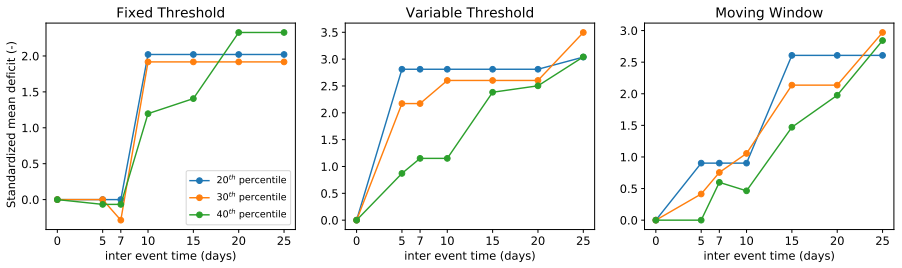


Figure 4.2: Standardized mean deficit of groundwater drought events derived from no pooling and pooling using five, seven, 10, 15, 20, 25 days as inter-event time for well L-PB42

As for standardized mean duration (see Figure 4.3), it increased abruptly for 20th and 30th percentile using fixed threshold from a 7-day to a 10-day inter-event time. 10 days onwards it remained constant until 25 days. On the contrary, for fixed threshold method and 40th percentile, standardized mean duration increased more smoothly and remained constant (around 2.5) between 20 and 25 days exceeding the mean of the other two percentiles. The plot of the mean duration of 40th percentile showed an even smoother increase for variable threshold compared to the fixed one. Standardized duration mean for 20th percentile reached a plateau earlier than that in fixed threshold (in five days) but there was a rise from 20 to 25 days. The identified droughts in that method were either too closely or too far away timewise which could explain the intermediate plateau. In general, using a higher percentile, more drought events were identified and the time distance between them decreased letting pooling take place at various inter-event times.

Regarding moving window, there was smoothing for 30th and 40th percentile but not that much for 20th percentile. Its plateau was split into two smaller ones of five (5-10) and 10 (15-25) days. In addition, it is worth noting that for both variable and moving window approach, the highest increase of standardized mean duration between no pooling and pooling of five days was noticed for 20th percentile. A similar pattern was observed for standardized mean deficit. A possible reason is that

when the percentile was low, the majority of drought events identified were closely timewise; therefore, most drought events were pooled at a short inter-event time. From that inter-event time onwards, there were not many drought events left to be pooled. As for the general pattern of pooling, most of the wells showed a similar pattern as that of L-PB42 but there were some wells (such as L-PB44) whose relationships between deficit, duration and inter-event time were different. GWL for that well was quite smoother and pooling did not occur for short inter-event times (see Figures B.7 and B.8 in Appendix B.3).

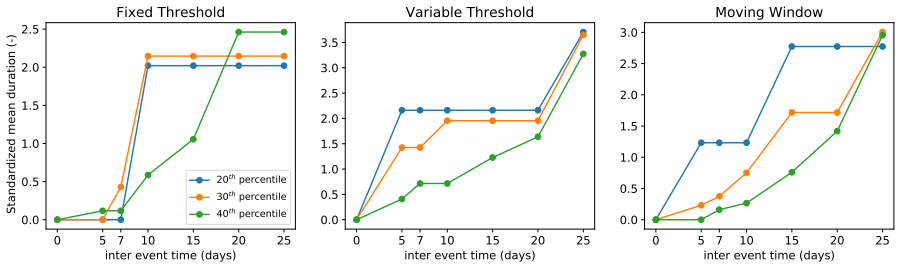


Figure 4.3: Standardized mean duration of groundwater drought events derived from no pooling and pooling using five, seven, 10, 15, 20, 25 days as inter-event time for the well L-PB42

Figure 4.4 depicts the standardized mean deficit over different inter-event time periods used for pooling for all wells of each district of the study area. The results were derived based on the variable threshold method using 30th percentile. Most wells showed: (i) a sharp increase for a five-day inter-event time, (ii) a plateau or slight increase from five days onwrards until 20 days, and (iii) another sharp increase from 20 to 25 days. Except well L-PB49, wells of w_1 did not have large differences; their standardized mean deficit for all wells varied between one and just above two for the plateau period (from seven to 20 days). On the other hand, excluding well L-PB82 of w_5, standardized mean deficit for the wells in that district varied between zero and two and a half for the same time period as before (from seven to 20 days) suggesting that drought deficit varied over wells in that district. The fact that the wells are not located close to each other in w_5 could be the reason why there was a high deviation. The GWL of the wells were quite heterogeneous; therefore, GWL regime varied over that district and so did the results of pooling.

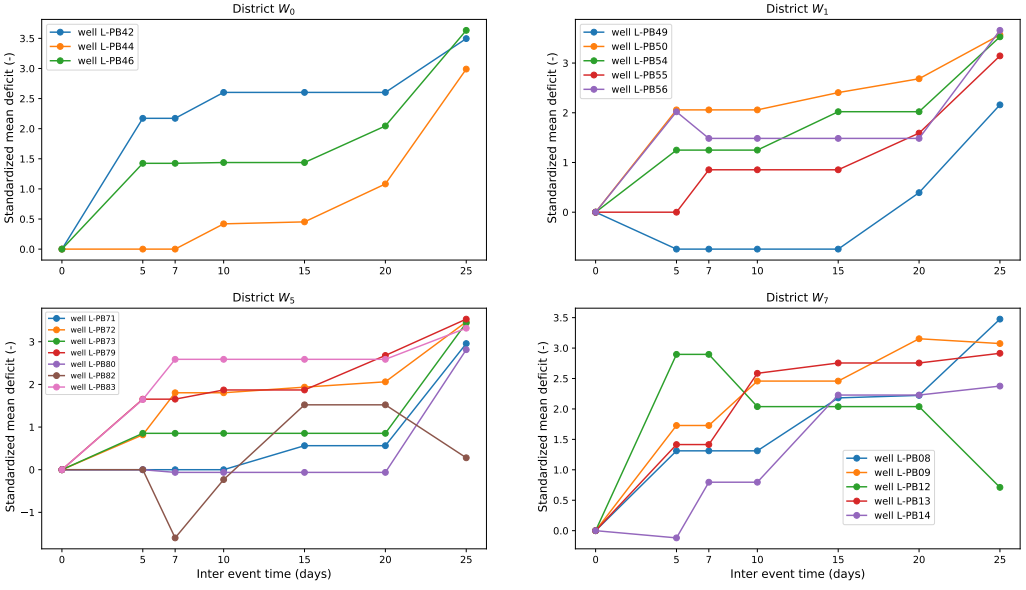


Figure 4.4: Relationship between standardized mean deficit and inter-event time for pooling for the wells in the four districts (w_0, w_1, w_5, and w_7) of the study area (variable threshold and 30th percentile was used for the identification of groundwater droughts)

Regarding well L-PB82 of w_5 district, its erratic behavior can be explained by the fact that large

drought events which occurred closely timewise pooled (merged). In this way, the mean deficit was lower compared to the case where no pooling applied. More specifically, the nominator decreased considerably since many large drought events were pooled and only one value for each merging was saved (only the largest deficit was considered for each merging). On the other hand, the denominator decreased due to smaller sample size. The influence of nominator was larger than that from the denominator.

As for w_7 district, all wells except L-PB12 showed a smoother increase of standardized mean deficit. For well L-PB12, deficit distribution was left skewed which means that the majority of drought events had large deficit. As a result, when minor drought events (which were closely timewise) pooled to a larger one for the five-day inter-event time, the increase was abrupt (the denominator decreased but the nominator did slightly). This explains the high peak at the beginning. The reduction after that peak is explained by the fact that large drought events were merged but not minor ones as inter-event time increased.

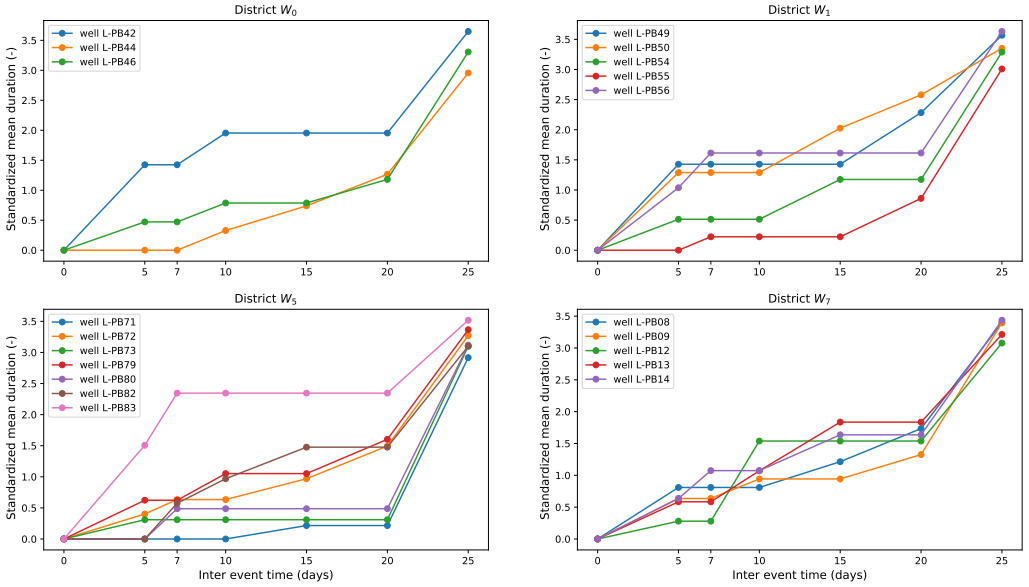


Figure 4.5: Relationship between standardized mean duration and inter-event time for pooling for the wells in the four districts (w_0, w_1, w_5, and w_7) of the study area (variable threshold and 30^{th} percentile was used for groundwater droughts)

As regards drought duration (see Figure 4.5), there was no decrease from one inter-event time to the following one since not the maximum duration of the pooled events was used but their sum. It is worth noting that for the wells of w_1 and w_5 districts, there was a distinct plateau but not for the case of w_7 district where the relationship was smoother. Additionally, the deviation (among the wells) was significantly lower for w_7 district compared to the other ones indicating that its wells had a similar relationship between mean duration and inter-event time.

Key results

- 40^{th} percentile showed the smoothest increase of deficit and duration as inter-event time rose for the three threshold techniques used. The lower the threshold, the sharper the increase for both deficit and duration regarding variable threshold and moving window.
- An inter-event time of around 10 days can be considered as satisfactory for both deficit and duration regarding variable threshold of 30^{th} percentile. From 10 to 20 days, there was a plateau which means that most minor drought events have been merged before that. When inter-event time increases or decreases slightly, an abrupt change is not desirable. The goal is to identify an inter-event time on which many droughts can be merged for most of the wells in all districts. Based on that, the minimum value of the plateau (i.e around 10 days) can be a suitable inter-event

time. In addition, more than 20 days is a long period which is not suitable since independent events may be merged; therefore, applying pooling using inter-event time larger than 20 days may distort the data. However, it is worth highlighting that there is not a unique right answer regarding inter-event time of pooling.

- For most wells, using variable threshold of 30th percentile, deficit increased using an inter-event time between zero and five or 10 days; it was almost constant from 10 to 20 days; and between 20 and 25 days, it increased slightly.
- An irregular behavior was noticed for the relationship between deficit and inter-event time for a few wells. The reason was how merging of drought deficit was implemented. That irregular pattern was not observed in duration given that another approach is used. The sum is used for deficit whereas the maximum for duration when pooling occurs.
- The relationship between between duration and inter-event time showed lower variation over the wells for each district than that between deficit and inter-event time.

4.1.1.4 GWL Deficit using the median as threshold

Median GWL value was also used as threshold function for the estimation of deficit which is one out of the four drought characteristics analyzed in the current study. It was separated from the other three threshold functions since it was based on modeling and not measurements. For the case of 5% driest year in district w_5 (upper bound of gray area) (2nd row in Figure 4.6), its average GWL deficit was considerably lower than the other districts; its maximum value just exceeded 120 mm in April. This result was quite unexpected. District w_5 was close to agricultural fields with lower elevation; therefore, drainage gradient could be higher and consequently deficit could be larger. However, the results were not in line with the previous argument. The poor calibration performance could also have affected the results.

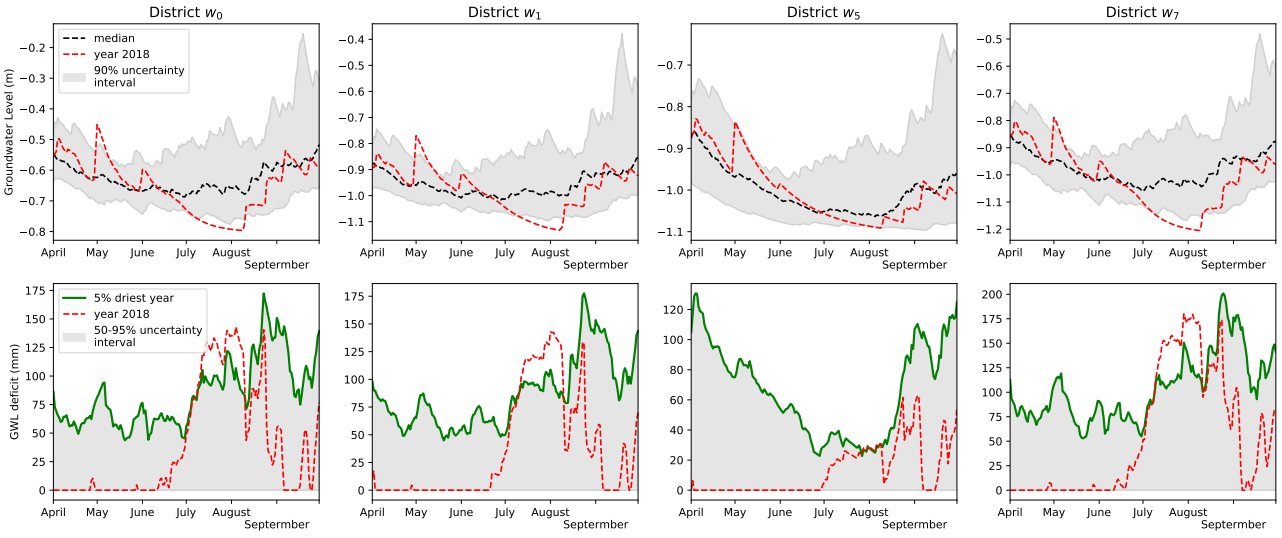
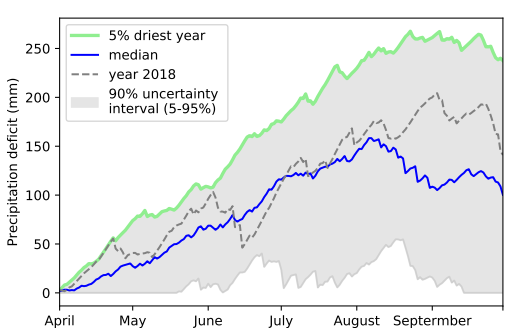


Figure 4.6: Uncertainty interval of modeled GWL in April-September for all districts derived by the period 1988-2019 (1st row); GWL deficit based on the median GWL for the same period for all districts (2nd row)

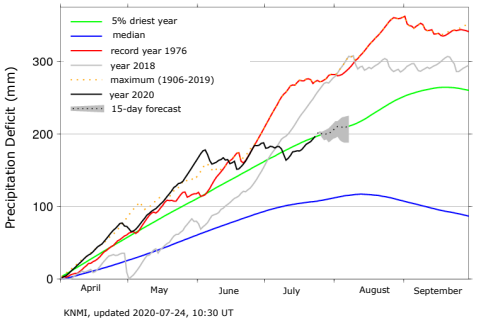
The pattern of GWL deficit in district w_5 differed from the other ones. For the 5% driest year, the deficit was high in April-May but was not for the other districts. Additionally, in July, there was a moderately increase for all districts except w_5. Another point for that district is that GWL deficit of the year 2018 did not exceed the 5% driest year unlike the other districts. The exceedance occurred in July-August for those districts. At that period in w_5, the difference GWL between 5% driest year and 2018 was the lowest, though.

As for the estimation of water amount needed to hinder GWL lowering, it is often underestimated. It is not what the 2nd row in Figure 4.6 presents but more than that. By adding extra water to the system, extra transpiration occurs boosting water consumption. However, this fraction was not considered in the current analysis.

Comparing GWL deficit to precipitation one for Leiden, two main points are worthwhile. These are: (i) GWL deficit was lower than precipitation deficit, and (ii) precipitation deficit of year 2018 did not exceed the 5% driest year (Figure 4.7-a) but GWL deficit did in most districts. However, the method used for deficit estimation was not the same for both cases; therefore, the comparison may not be reliable. Additionally, year 2018 was much more extreme to other locations than Leiden since the national maximum precipitation deficit of that year reached 300 mm (Figure 4.7-b) but in Leiden its peak was almost 200 mm.



(a) Precipitation deficit of study area in Leiden for the period 1988-2019 based on two weather stations with a different timespan



(b) National average of precipitation deficit in the Netherlands based on 13 weather stations (Royal Netherlands Meteorological Institute (KNMI), 2020)

Figure 4.7: Precipitation deficit

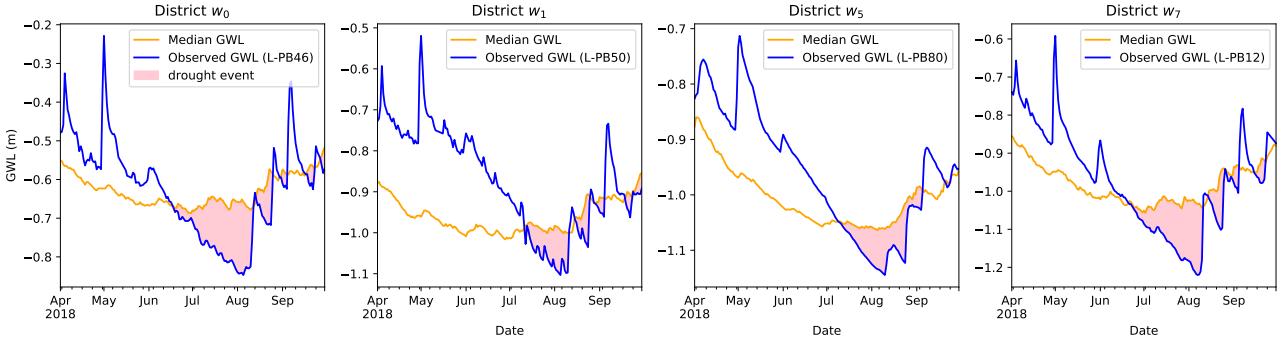


Figure 4.8: Drought events identified using modeled median GWL of the period 1988-2019 as threshold from April to September for all districts in Leiden (GWL of the well with the highest performance in calibration or best fit is presented for each district)

Figure 4.8 depicts the drought events identified for the period April-September in 2018 using as threshold function the median modeled GWL of the period 1988-2019. Visually inspecting the threshold, median GWL decreased from April to July in district w_5 but this drop was minor for the rest ones. No drought was identified before summer for all districts. However, in the current analysis, GWL median value was dependent on the performance of UWBM; performance of w_1 was not high, therefore, the real GWL median could be different. Consequently, the identified drought events could be different. The pattern of this threshold (median GWL) seems to be quite sensible and outperform the previous ones created by fixed, variable, and moving window threshold but these results are indicative. Median GWL was based on modeling whereas the other thresholds on measurements; hence, their comparison

is not well grounded. For median GWL threshold, data of 32 years were used whereas for the other techniques just one year.

Key Results

- Using median GWL as threshold seems promising (based on its pattern) but deriving GWL by modeling increases the possibility of errors since it depends on the performance of the model.
- Based on precipitation deficit, Leiden suffered much less than the average status in the Netherlands.

4.1.2 Spatial Drought Analysis

Areal extent is another component of droughts and was quantified using spatial analysis. Figure 4.9 depicts the relationship between duration and areal extent of drought events identified via fixed threshold approach for three percentiles (20^{th}, 30^{th}, and 40^{th} one). The threshold was fixed over time but varied over space. Areal extent of droughts rose rapidly and covered almost the whole district as drought duration increased for 30^{th} and 40^{th} percentile. For 20^{th} percentile, the number of identified drought events were limited; therefore, no clear pattern could be observed. It is worth noting that the average areal extent of each drought event is depicted ofr each point.

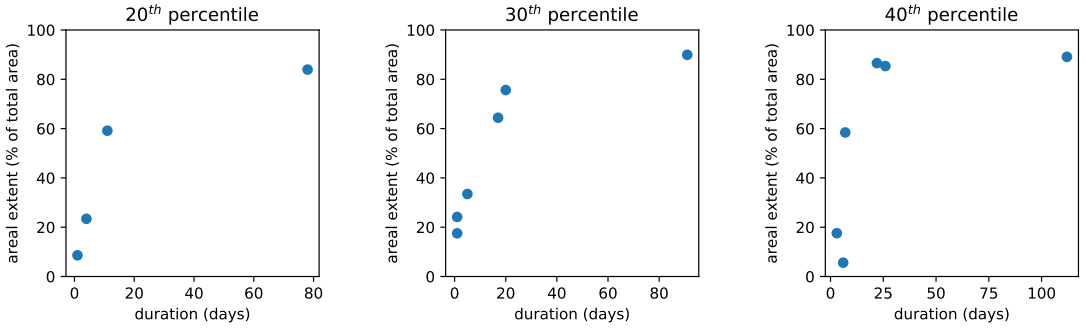


Figure 4.9: Relationship between groundwater drought duration and areal extent for drought events derived by spatial analysis for fixed threshold method over all percentiles in w_0 district

Droughts with mediocre duration (around 25 days) covered almost 80% of study area for 40^{th} percentile which means that for a large proportion of these drought events' duration, almost the whole area suffered from drought. For areal extent, there was an upper boundary (100%) which did not hold for drought duration. In addition for the extreme event (upper right corner for all percentiles), its areal extent was not 100% of which means that there were some time steps for which that drought event did not cover the whole area.

Regarding deficit, a similar pattern as that of duration was observed as can be seen in Figure 4.10. Drought events with high deficit were linked with larger areal extent but there was a threshold. When the deficit exceeded five cm for the case of 40^{th} percentile, droughts had an average areal extent of at least 80%. For a deficit from five cm onward, there was saturation for areal extent. The areal extent for the drought event with the largest deficit was 89%, whereas for the two events with the second and third largest deficit were 85%, and 86% respectively.

Regarding drought duration and deficit, most drought events had low values for both variables (see Figure 4.11). As percentile increased, both variables of all drought events increased which is sensible. A worthwhile point is that for 40^{th} it seems that there was a linear relationship between duration and deficit (excluding the outlier) but the number of identified drought events was only six; therefore, no conclusion could be drawn since it could be quite baseless.

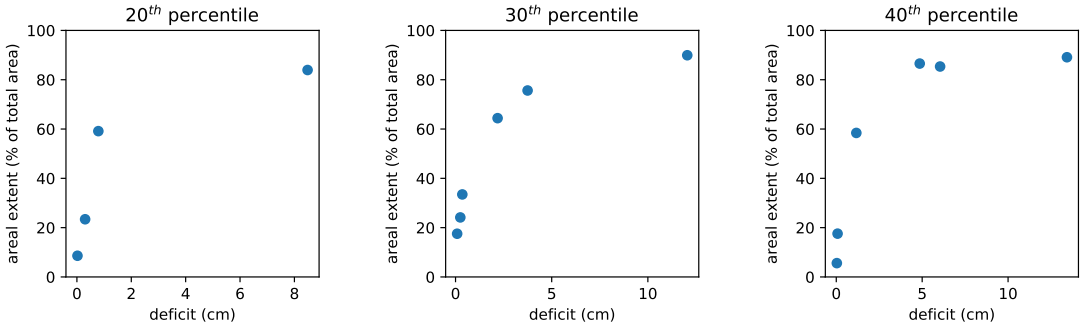


Figure 4.10: Relationship between groundwater drought deficit and areal extent for drought events derived by spatial analysis for fixed threshold method over all percentiles in w_0 district

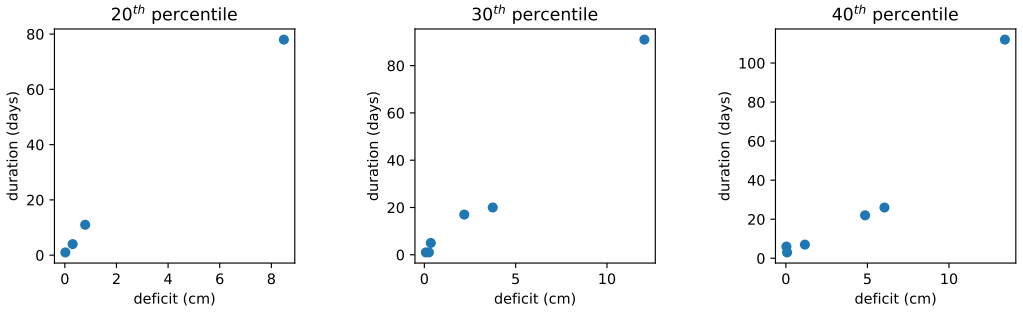


Figure 4.11: Relationship between groundwater drought duration and deficit for drought events derived by spatial analysis for fixed threshold method over all percentiles in w_0 district

Key results

- It seems that there is a threshold relationship between deficit, duration and areal extent in pairs. When a certain deficit or duration value was exceeded, around 80% of the study area (w_1) suffered from a groundwater drought.

4.2 Calibration of Urban Water Balance Model

The results of calibration in w_0 district are presented in Section 4.2.1. Next, Section 4.2.2 refers to the sensitivity of the model in the same district. The results of the other districts are attached in Appendix B.7.

4.2.1 District w_0

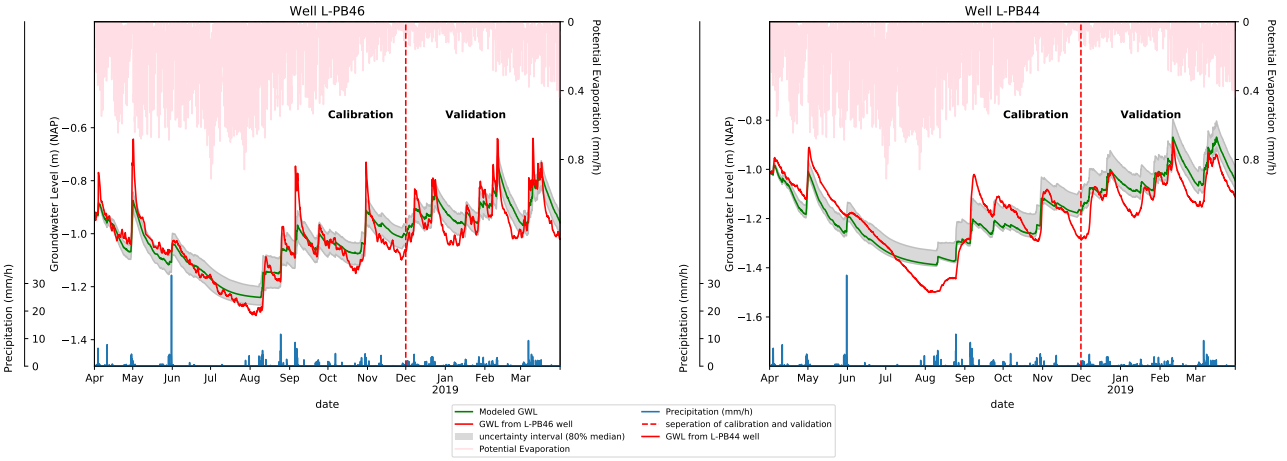


Figure 4.12: Calibration and validation of UWBM for w_0 district using the GWL of L-PB44 and L-PB46 individually

Wells L-PB44 and L-PB46 were used for calibration in w_0 district (see Figure 4.12). GWL in UWBM is measured with respect to ground surface and not NAP. Additionally, in UWBM, the negative GWLs are considered as positive. In Figure 4.12 are measured in respect to NAP. The ground levels for L-PB44 and L-PB46 were 0.418 m and 0.458 with respect to NAP accordingly. The calibration performance using L-PB46 was better than that using L-PB44 but in both cases the model was quite conservative and could not approximate reality to a high degree. The drop of GWL in summer of 2018 could be explained by low precipitation and high potential evaporation but in period December-February when potential evaporation was low, not a significant rise in GWL was observed. The peak of precipitation in June did not have an important influence on GWL and a possible explanation is that most of it ran off and collected to storm water drainage system. On the contrary, events with lower intensity had a greater impact on GWL. This indicates poor recording of input data (precipitation). Another interesting point regarding the input data to the model is that there was some potential evaporation at night (white spots on top of Figure 4.12). Penman formula allows it since aerodynamic resistance is one of the parameters; wind in itself could crate evaporation without net radiation.

Not only were measurements of individual monitoring wells used but also combination of them given that there was a high variation. The mean groundwater of L-PB44 and L-PB46 wells was used for calibration and the results are illustrated in Figure 4.13; it was like a fictional monitoring well which represented the whole district. Monitoring well L-PB42 was excluded from this averaging since it had large fluctuations and it could disrupt the GWLs. The model fit was not as satisfactory as the two previous cases but at least for most of study period, modeled GWL was between the range of the two monitoring wells. Its performance was poor in summer (calibration period) but it was better in winter period (validation period). The model was conservative as happened in the previous two cases; it was not able to lower GWL (not that abruptly as the measurements indicated). Using average measurements, extremes were smoothed but extreme low values were important for this analysis. Therefore, averaging of measurements is not suggested.

Based on aforementioned Figure 4.13, it is worth noting that there was heterogeneity in GWLs. The difference between the two monitoring wells was more than 0.1 m most of the study period and this difference maximized in summer 2018 (around 0.2 m). Even though (i) the areal extent of study area was limited (less than one km^2), and (ii) the distance between those two wells was about 800 m, their GWLs were different. Therefore, the assumption of homogeneity in the study area was not absolutely valid. That means that UWBM may not be a suitable model for these conditions. Location of the wells, their surrounding land use, and their distance from canals were three factors which may played a role in GWL behavior.

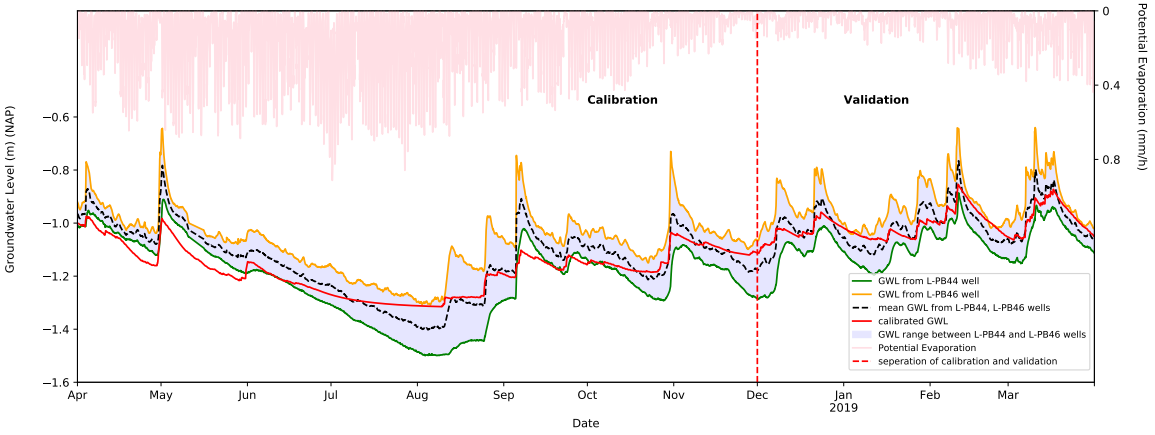


Figure 4.13: Calibration of UWBM for w_0 district using the mean GWL of L-PB44 and L-PB46

Regarding the model results based on the objective function, calibration performance was satisfactory using the three datasets to calibrate the model. However, NSE_{log} results regarding validation were poor (their median was lower than 0) (see Figure 4.14). The change in variation from calibration to validation was the largest for L-PB44 well (based on the interquartile range). Additionally, using the mean GWL of L-PB44 and L-PB46, NSE_{log} for calibration was almost equal to that using only L-PB44 (based on the median values). Nonetheless, this did not hold for validation where the results using L-PB44 and L-PB46 outperformed that one using L-PB44 only.

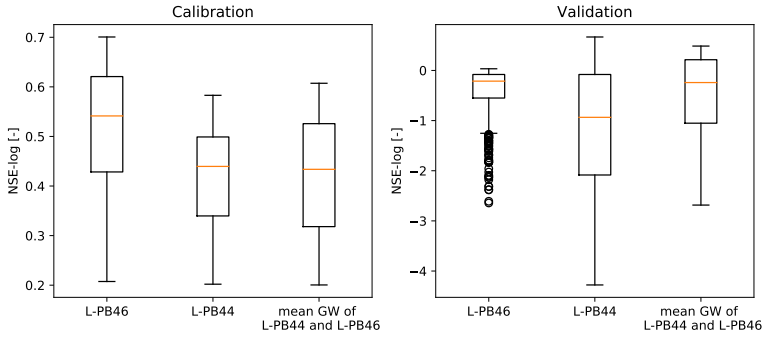


Figure 4.14: Boxplots of NSE_{log} for calibration and validation using GWL of L-PB44, L-PB46 wells individually, and the mean of them for w_0 district

The parameters' ranges used are presented in Table 4.2. They were the same for the cases of calibration using L-PB46, L-PB44 individually, and their mean. For w there was a prominent peak for values close to its upper limit (150 d); using these values, the model performance was higher (see Figure 4.15). That seemed to apply for down_seepage_flux for values closely to its lower limit (1 mm/d) but it was not that clear and the difference was not great. On the contrary, there was no pattern for the rest two calibrated variables.

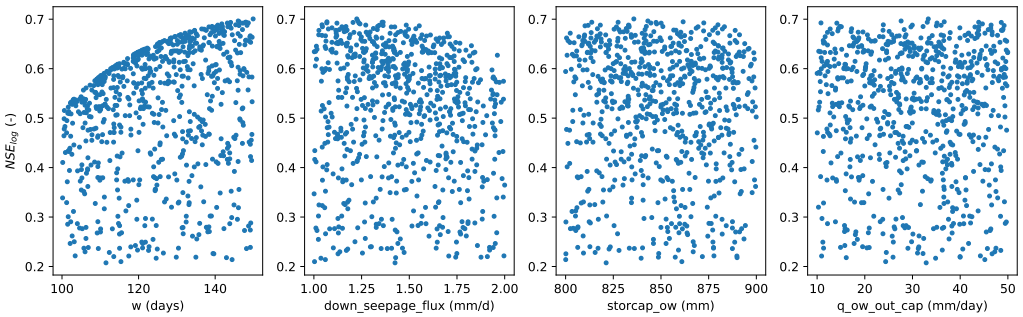


Figure 4.15: NSE_{log} values for the parameters' range using monitoring well L-PB46 for calibration

Table 4.2: Parameters' range for calibration of district w_0

	Parameters			
district	w (days)	down_seepage_flux (mm/d)	storcap_ow (mm)	q_ow_out_cap (mm/d)
w_0	100-150	1-2	800-900	10-50

4.2.2 Model Sensitivity

Regarding sensitivity, using the first approach (Method A), only the cases where $NSE_{log} > 0.2$ in calibration were considered (feasible solutions). Based on the parameters of feasible solutions used for calibration, validation results were estimated. Parameter storcap_ow had a large variation for validation (Figure 4.16 right part) and changes to that parameter could lead to poor performance. On the contrary, q_ow_out_cap had the least variation (its interquantile range was about 0.01 and 0.02 for calibration and validation accordingly), showing that its sensitivity was limited. Parameter w showed the highest median NSE_{log} out of all parameters in calibration but this did not occur for validation. Additionally, its variation (based on the interquantile range) in validation was the second lowest after q_ow_out_cap; its difference from down_seepage_flux was low, though. In calibration, another point is that variation of w regarding the low values was larger than that regarding the high ones. The median and the 75^{th} percentile were quite close. For all parameters, their difference between maximum and 75^{th} percentile were low indicating that the variation of solutions with high performance were limited. A similar pattern held for validation.

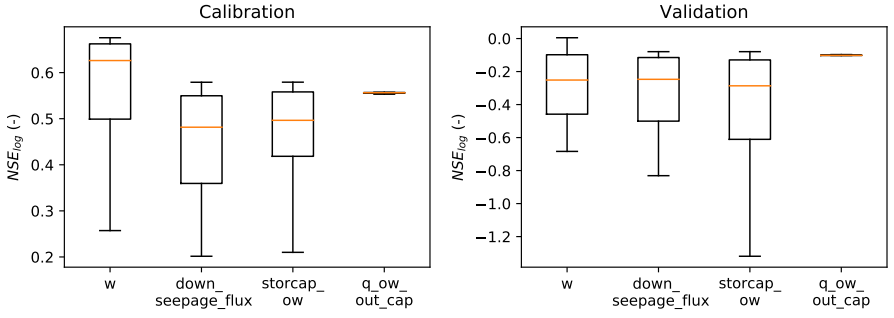


Figure 4.16: Sensitivity analysis (method A) for each calibrated parameter (using well L-PB46 for the calibration in district w_0)

Using Method B for estimating sensitivity (see Table 4.3), the results (explicit changes in model performance) were in line with those in Method A. The largest deviation from the normal conditions (the case of no change, 'default') was observed for parameter storcap_ow (especially in validation). Regarding q_ow_out_cap, it showed the least sensitivity as was observed using Method A (the difference was minor and was not noticeable using three significant digits). The changes for parameters w and

down_seepage_flux were much larger for validation than calibration as noticed using Method A for sensitivity analysis.

Table 4.3: NSE_{log} values when each parameter of the selected model using well L-PB46 (in district w_0) increased or decreased by 10% (method B)

	Calibration - NSE_{log} (-)				Validation - NSE_{log}			
	w	down_seepage_flux	storcap_ow	q_ow_out_cap	w	down_seepage_flux	storcap_ow	q_ow_out_cap
+10%	0.546	0.549	-0.181	0.666	0.031	0.004	-1.445	-0.086
-10%	0.638	0.682	0.231	0.666	-0.281	-0.448	-2.381	-0.086
default	0.666	0.666	0.666	0.666	-0.086	-0.086	-0.086	-0.086

Regarding Method C, boxplots in Figure 4.17 present the variability in model performance based on a change in each parameter's range by a 10% increase or a 10% decrease for calibration and validation respectively. Variation of w and down_seepage_flux in calibration increased slightly for both increase and decrease cases compared to the 'no change' status. Another point is that, in calibration, the performance (based on median values) of storcap_ow decreased for both changes but this did not happen for q_ow_out_cap. For that parameter in calibration, a 10% reduction of its range increased the median NSE_{log} whereas a 10% increase of its range lowered that performance reducing slightly the variation as well.

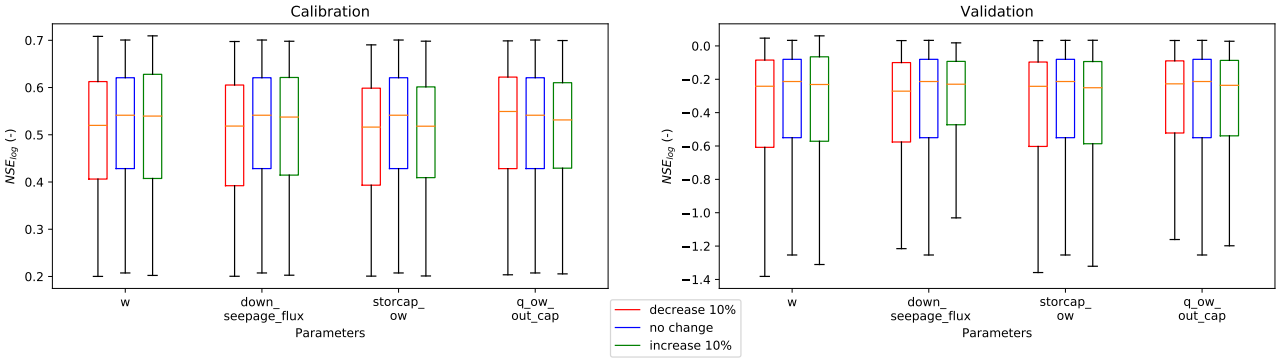


Figure 4.17: NSE_{log} values of calibration and validation period for the cases of 10% increase and decrease for each calibrated parameter's range (using well L-PB46 for the calibration in district w_0)

Regarding validation, an increase of 10% in down_seepage_flux's range led to a much lower variation compared to the 'no change' status. Variation was also reduced by a decrease of 10% in q_ow_out_cap's range to a lesser extent though. For the other two parameters, increase and decrease of 10% in their range led to a greater variation. Another point about validation is for all parameters, low NSE_{log} values' variation was larger than that of high values as was observed using method A for sensitivity analysis. The difference between median and 25^{th} percentile was larger than that between median and 75^{th} one.

Calibration was conducted for the other 3 districts (w_1, w_5, and w_7). For w_1, NSE_{log} was not estimated based on hourly values but based on daily ones, since the performance was higher. Therefore, down-sampling was conducted for both measured and modeled GWLs. The results of calibration and sensitivity analysis for w_1 using Method B and C are attached in Appendix B.7.1.2.

Calibration results of the other 2 districts (w_5 and w_7) and their sensitivity using Method B and C are presented in Appendix B.7.2 and B.7.3 accordingly. Performance of different monitoring wells in calibration and validation for districts w_1, w_5, w_7 before selecting the model for soil moisture estimation are presented in Appendix B.7.4. Furthermore, that part contains the sensitivity analysis using Method A for the aforementioned districts using the well for each district which represented the

reality best. The wells used for (i) for the estimation of model parameters, and (ii) for soil moisture analysis were: L-PB46 for w_0, L-PB50 for w_1, L-PB80 for w_5, and L-PB12 for w_7.

Key results
In general

- There was a high spatial heterogeneity regarding GWL in Leiden even though wells were located close to each other. It hindered a high performance of UWBM.
- UWBM was able to predict the general pattern of GWL for some monitoring wells in each district.
- Poor recording of rainfall was one of the reasons why UWBM performance was not that high.
- Using a fictional well (averaging the GWL of two real monitoring wells) did not enhance the performance of the model.

Sensitivity

- Discharge capacity from open water (internal) to outside water (external) (q_ow_out_cap) is the parameter which showed the least sensitivity among the calibrated parameters in UWBM.
- Storage capacity of open water (storcap_ow) and seepage to deep groundwater (down_seepage_flux) were sensitive parameters of UWBM; the former especially in validation.
- Sensitivity was larger for low NSE_{log} values than high ones for all parameters.
- Surrounding land use was not able to explain the wells' GWL behavior.

4.3 Soil Moisture Droughts

Quantifying the characteristics of soil moisture droughts was part of the research questions. The parameters which led to best fit in calibration for each district were used to estimate soil moisture via UWBM. Figure 4.18 depicts soil moisture for each district and the observed GWL (in hourly time scale) from the monitoring well used for calibration. The well which led to the highest performance in calibration was used for each district.

The patterns of soil moisture in districts w_0, w_5, and w_7 were similar. There was a significant drop in volumetric soil moisture from July to August and both its beginning and end of the drop were quite abrupt. On the contrary, soil moisture was quite smooth in winter especially for district w_5. The difference from district w_1 was that the rise of soil moisture in August was smoother; it had a staircase pattern. Regarding the magnitude of decrease, it was not the same for all districts; district w_5 had the largest drop (60 mm in July) but it recovered in September. For that district, after the drop occurred in July, soil moisture was almost constant until the middle of August. In general, based on drought propagation, as a drought moves through the hydrological cycle, attenuation and lengthening occur. Attenuation from soil moisture to groundwater droughts was not observed. On the other hand, it seems that lengthening took place. Soil moisture droughts had a shorter duration than that of groundwater droughts.

Additionally, fluctuations of soil moisture during the day were greater than GWL ones. For most wells, GWL decreased smoothly compared to soil moisture. As for the high values, most peaks of GWL in winter were not observed for soil moisture except those in February and March. The model used for each district was quite conservative which could explain the low variation in soil moisture in winter. Many peaks and valleys occurred the same timing for both GWL and soil moisture indicating that there was no delay for the infiltration of unsaturated zone to groundwater table. However, their difference in magnitude in winter was noticeable (with respect to the whole time series). Another point is that in dry periods, decrease of soil moisture is high when the groundwater table is deep since capillary rise cannot occur; only precipitation re-wets the unsaturated zone but this amount cannot limit the dryness (Wiesner et al., 2016). That applied to the current results.

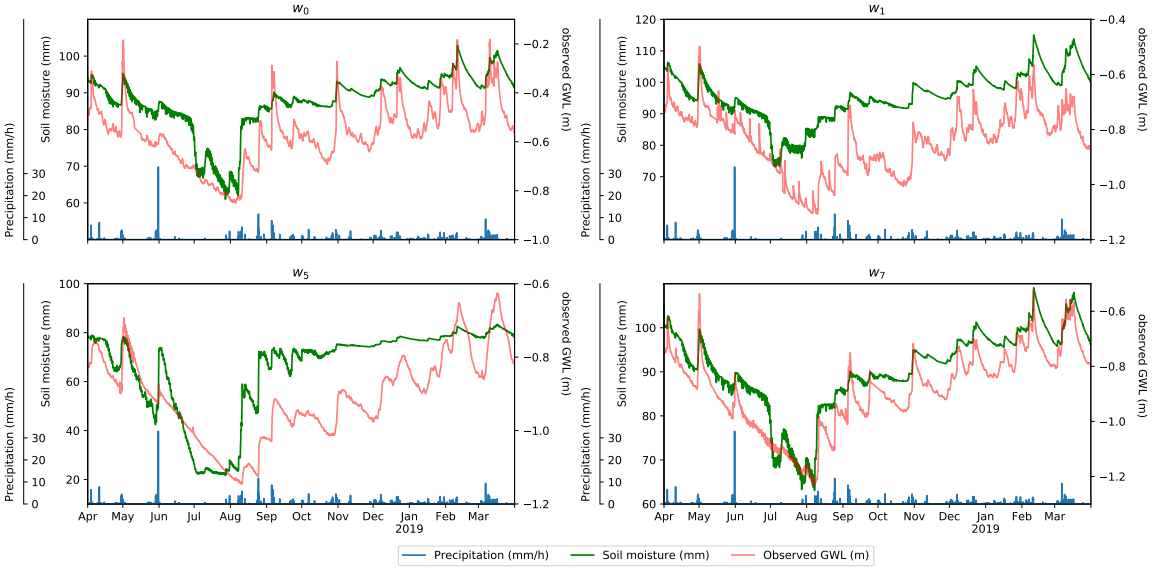


Figure 4.18: Modeled soil moisture and observed GWL for the four districts in Leiden (the following wells regarding observed GWL were used: L-PB46 for w_0, L-PB50 for w_1, L-PB80 for w_5, and L-PB12 for w_7)

Soil moisture fluctuated during the day from April to September in all districts except w_5 (the time

series line was thicker in that period for w_0, w_1, and w_7). The fluctuations of GWL were less than those in soil moisture. District w_1 showed large daily fluctuations on both variables, though. Fluxes related to unsaturated zone are much faster than those of shallow groundwater; therefore, it was expected that soil moisture would have more fluctuations. It is worth noting that local conditions such as vegetation and root depth influence the fluctuations of soil moisture content at short time scales based on the study of Wiesner et al. (2016) who used observations (via tensiometer probes).

4.3.1 Drought deficit and duration

Figure 4.19: Soil moisture drought events identified using fixed, variable, and moving window threshold for three different percentiles for well L-PB46 in district w_0

Deficit and duration are two main characteristics of soil moisture droughts and their analysis was one of the main goals in this thesis. Fixed, variable, and moving window threshold showed similar results as those when applied for GWL. Fixed threshold method identified well the dry conditions regarding soil moisture in summer of 2018 but not those in winter. Performance of variable threshold was better than that of moving window but for neither of the two techniques, dry conditions in summer of 2018 were considered as one event. Deficit and duration of the event in summer 2018 were underestimated based on those techniques. Due to the poor performance of moving window, no further analysis was performed regarding its mean deficit and duration.

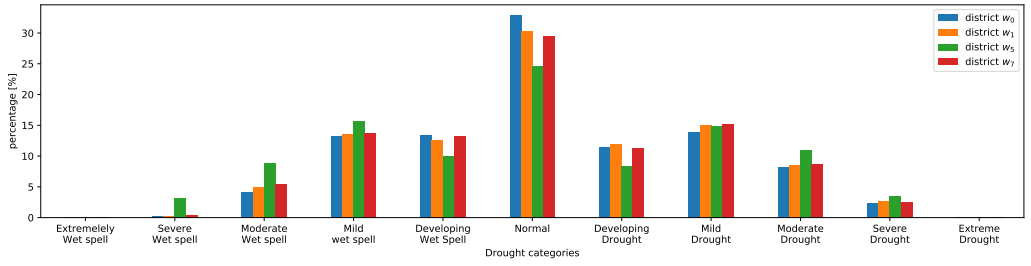


Figure 4.20: Percentages of different drought classes derived by SMDI for all districts studied in Leiden

Besides threshold method, deficit and duration for soil moisture can be estimated using standardized indices. In this research, SMDI and SMA were applied. Using SMDI, soil moisture drought categorization was almost the same over all districts except w_5 (see Figure 4.20). Ranges of the classes are presented in Table 4.4. Percentage of w_5 was larger for moderate droughts and lower for normal and developing droughts than that of the other districts. The fact that the time series of data was not measured but modeled via UWBM increases the uncertainty of the outcome. The results of the other standardized index used in this thesis (SMA) are attached in Appendix B.8. An unexpected behavior for district w_5 was observed too.

Table 4.4: Soil Moisture Deficit Index (SMDI) classes (Keshavarz et al., 2014)

Values	Class	Values	Class
SMDI⩾ 4	Extremelely Wet spell	-1.0⩽SMDI<-0.5	Developing Drought
3⩽ SMDI<4	Severe Wet spell	-2⩽SMDI<-1	Mild Drought
2⩽SMDI<3	Moderate Wet spell	-3⩽SMDI<-2	Moderate Drought
1⩽SMDI<2	Mild wet spell	-4⩽SMDI<-3	Severe Drought
0.5⩽SMDI<1	Developing Wet Spell	SMDI< -4	Extreme Drought
-0.5⩽SMDI<0.5	Normal		

Key results

- Soil moisture dropped abruptly in summer 2018 whereas GWL decreased smoothly.
- Timing of variation between GWL and soil moisture was the same but amplitude was much larger for GWL especially in winter period.
- Soil moisture fluctuated during the day compared to GWL which remained more constant.
- Fixed threshold identified the severity of the large soil moisture drought in summer 2018.
- Using variable threshold to identify soil moisture droughts, deficit and duration were highly related.
- There is high uncertainty regarding the results of SMDI and SMA since they are based on the performance of UWBM.

4.4 Vulnerability

The main goal of this thesis was to estimate vulnerability of cities to soil moisture and groundwater droughts and specifically to Leiden. It was computed at two levels (fine and coarse); the fine resolution is presented in this chapter whereas the coarse one is attached in Appendix F. The analysis for groundwater and soil moisture droughts are presented in Sections 4.4.1 and 4.4.2 respectively.

4.4.1 Groundwater Droughts

Fine Resolution

Vulnerability at downscale level provided thorough insight about the districts which were highly vulnerable. Firstly, the components of exposure and sensitivity were analyzed. Only one social sensitivity indicator was used to limit confusion about assigning weights by drought experts as mentioned in Methodology section.

The final products of exposure, sensitivity, and vulnerability were compared having used variable threshold of 30^{th} percentile for drought identification. Afterwards, the influence of including subsidence on vulnerability was analyzed. Next, how much the selection of drought identification technique affects vulnerability was mentioned. The three compared techniques were: (i) variable threshold of 30^{th} percentile without pooling, (ii) the same as (i) with the addition of 10-day pooling, (iii) using the 30-year modeled median GWL as threshold. At the end, sensitivity analysis excluding indicators were estimated.

Duration values were more pronounced than deficit ones but their pattern in general had some similarities (see Figure 4.21). More specifically, these were located in the following areas: east of w_0 district, west of w_1, north and west of w_5, and east of w_7. Regarding modeled and measured subsidence, their pattern differed due to resolution; resolution of measured subsidence was quite coarse. A high value in district w_1 led to smoothing in modeled subsidence. That point was dismissed. At that location, historical landmark Burcht is located which could explain the high subsidence since its elevation is quite high. However, there is a possibility of error given that the value is quite different than its surroundings. Regarding measured subsidence, the north side of w_1 suffered the most.

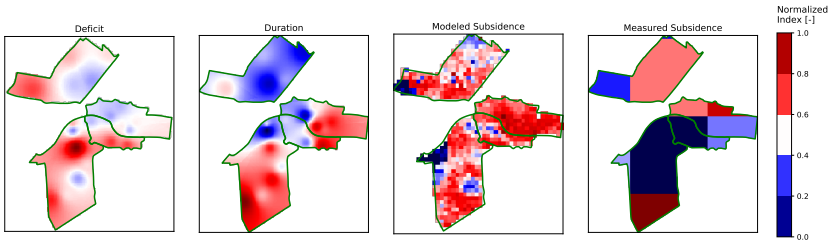


Figure 4.21: Exposure indicators for groundwater droughts at downscale level for Leiden (duration and deficit were derived by variable threshold of 30^{th} percentile without pooling; both available subsidence formats were used separately for vulnerability estimation)

Regarding sensitivity, two indicators - land use and difference between ground level and minimum GWL - showed high values over a large area; the values were larger for land use indicator (see Figure 4.22). On the contrary, the following indicators: buildings before 1960, monuments, public buildings, and shops mostly covered by low values. The objects of these indicators (e.g. public buildings in themselves) covered a tiny percentage of the total area and their sensitivity values were the maximum (i.e. 1) whereas the rest areas had the minimum value (i.e. 0). Consequently, the sensitivity was the minimum for most of the area regarding these indicators. It is also worth noting that 'low income household' indicator was not as much fine (in resolution) as the other ones since its data was collected at neighborhood level. The largest value was observed on the east side of w_7 district. A cluster of high values were also noticed in w_0, and w_1.

Figure 4.23 illustrates vulnerability and its main components. Its pattern followed total sensitivity component at fine spatial resolution; that was noticed in coarse resolution too. The difference between physical and total sensitivity was that the values were slightly larger for the latter one. There were also two hotspots in total sensitivity: the first one in the center of Leiden between w_0 and w_1 districts, and the second one on the south-east of w_7. In the latter district, the lowest vulnerability values were noticed on its north-west part. Regarding exposure, it is worth noting that outliers of subsidence (such as the pixel with the highest value in w_1) were ignored to avoid smoothing.

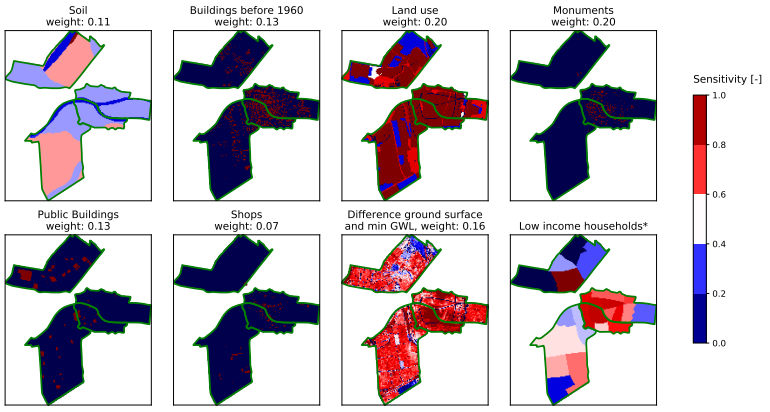


Figure 4.22: Sensitivity indicators for groundwater droughts at downscale level for Leiden; their weights are attached; *the weight of 'low income household' indicator is not attached to avoid confusion since this indicator was not used for the aggregation of physical sensitivity

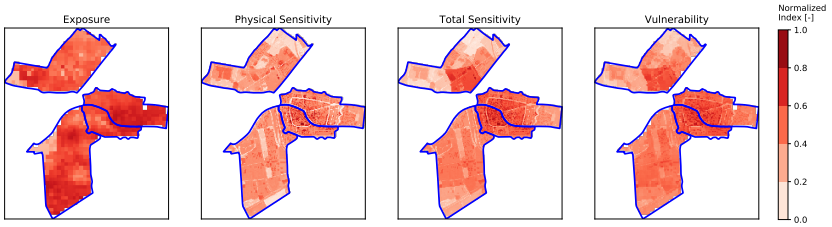


Figure 4.23: Exposure, sensitivity, and vulnerability to groundwater droughts at downscale level for Leiden ((i) exposure was derived by variable threshold of 30th percentile without pooling; (ii) modeled subsidence was employed as exposure indicator)

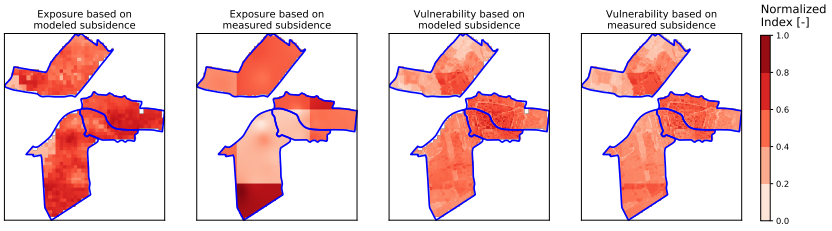


Figure 4.24: Comparing exposure and vulnerability using modeled and measured subsidence (exposure was derived by variable threshold of 30th percentile without pooling)

Comparing the results using modeled and measured subsidence (Figure 4.24), the main point is that their difference was noticeable in exposure but it was less pronounced in vulnerability; the latter can be noticed though. This can be explained by their weights. Subsidence weight was larger than deficit and duration ones in exposure aggregation (almost triple and double respectively). Regarding vulnerability, the influence of subsidence was minor given that sensitivity weight was three times as much as that for exposure.

Comparison including subsidence and not

Subsidence is a peculiar indicator which is not clear whether it should be included in the analysis. Excluding modeled subsidence led to (i) slightly larger variation (based on the interquantile range) in exposure over the study area, and (ii) lower median exposure but it did not affect the vulnerability (see Figure 4.25). This result was in agreement with the coarse analysis. More specifically regarding exposure, w_1, and w_5 showed large variation for the case of not including subsidence (see the 2nd subplot of Figure 4.25) but these areas with peaks and valleys were not observed in vulnerability distinctly.

It is worth noting that there are some gaps near the district boundaries for the case of including subsidence (see the 1st and the 3rd subplot of Figure 4.25). These are explained by the coarser resolution of subsidence than deficit and duration one. Despite those gaps, the resolution for the case 'exposure including subsidence' was the same as that for sensitivity since upsampling applied for subsidence.

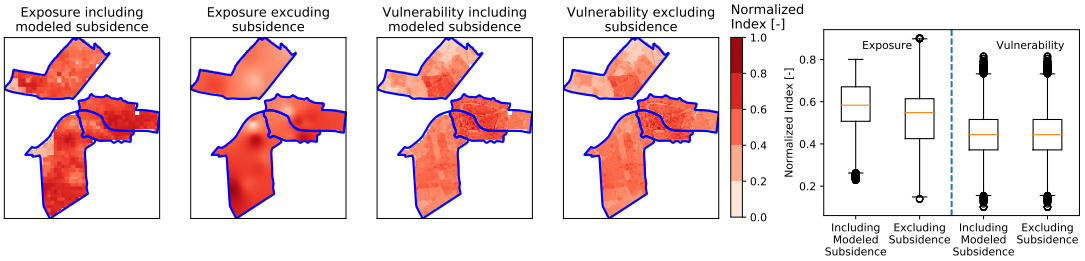


Figure 4.25: Comparing exposure and vulnerability using modeled subsidence and not (exposure was derived by variable threshold of 30th percentile without pooling); (the number of pixels used for boxplots were from left to right: 23,581; 24,652; 23,531; 24,557; their divergence are due to different vulnerability indicators' spatial extent)

Comparing exposure and vulnerability based on drought identification techniques

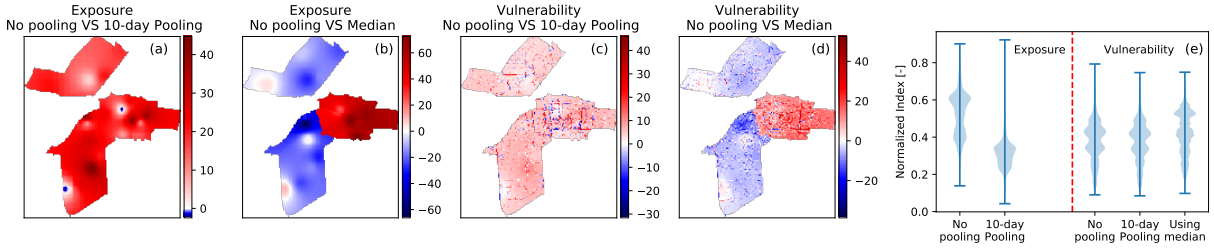


Figure 4.26: Difference (10^{-2} unitless) in exposure between variable threshold of 30th percentile in conjunction with 10-day pooling and not (a); difference in exposure between variable threshold of 30th percentile without pooling and using median GWL as threshold (b); difference in vulnerability for the aforementioned comparisons (c, d); violin plots for exposure, and vulnerability using (i) variable threshold of 30th percentile without pooling, (ii) the same as (i) with the addition of 10-day pooling, (iii) using median GWL. Case (iii) was applied only for the period April-September of 2018 (e)

How much a selected technique for drought identification affected exposure and vulnerability is presented below. The three techniques used were: (i) variable threshold of 30th percentile without pooling, (ii) the same as (i) in conjunction with 10-day pooling, and (iii) the 30-year GWL median as threshold. The first one was used as reference; therefore, two comparisons were conducted. It is worth noting that exposure from median GWL had a coarse resolution given that UWBM was used to estimate the median GWL. For deficit, there was one value per district. The same held for duration. Furthermore,

for the aforementioned results in Figure 4.26, subsidence was not included for exposure estimation due to high uncertainty of the data.

As for the spatial pattern of the differences, they were more pronounced for the cases of exposure than vulnerability for both comparisons. For example, there were exposure hotpots for the case pooling-no pooling but there was no clear pattern in vulnerability (Figure 4.26-a,c). For the same case, the largest positive exposure differences were noticed on the south part of w_5, and east side of w_1 (Figure 4.26-a). Large differences were also observed for the latter region comparing no pooling and median GWL cases (Figure 4.26-b). The smoothing in vulnerability was not as high as that for the case comparing 'no pooling' to '10-day pooling'. The reason why the differences between no pooling and median were not smoothed significantly from exposure to vulnerability is not clear. It may be related to the fact that deficit and duration had coarse resolution for the case of median GWL (the whole district had the same value).

Variation of the three techniques used in exposure and vulnerability showed interesting results. Exposure violin plot was not estimated for the case of 'median GWL' since deficit and duration had only four values (one for each district) given that coarse resolution was used. Even though 'no pooling' case had larger exposure variation than '10-day pooling', their kernel density distribution in vulnerability was highly similar (Figure 4.26-e). On the other hand, their vulnerability kernel density distribution were slightly different than the one using 'median GWL'. In general, all techniques had similar results regarding vulnerability. Therefore, out of the three techniques for drought estimation, the simplest one is preferred. This is variable threshold without pooling. The simplest one is not 'median GWL' since it would require long data via modeling.

Sensitivity Excluding Indicators

How much each indicator affects the final vulnerability product was estimated removing each indicator and computing the rate of change in vulnerability. The most important findings are highlighted below. Not only were their spatial differences analyzed but also the vulnerability variation.

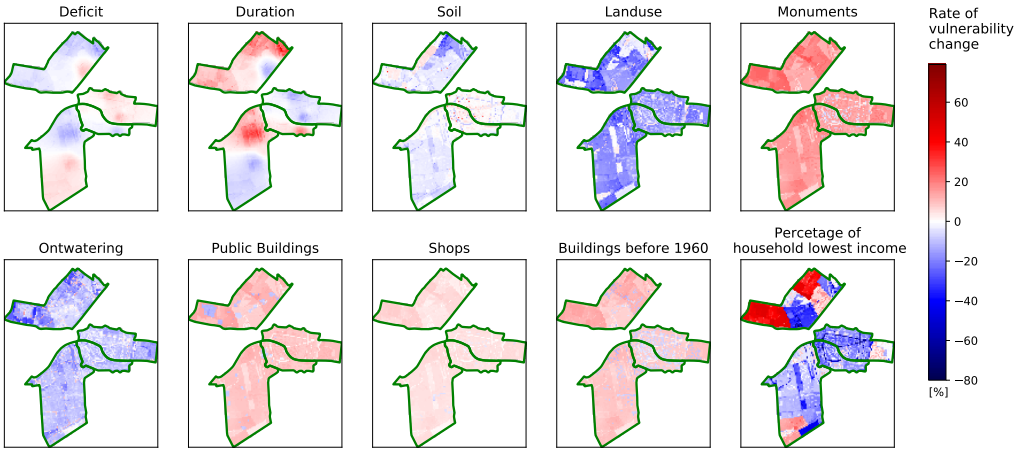


Figure 4.27: Rate of vulnerability change (%) excluding indicators regarding groundwater droughts (exposure was derived by variable threshold of 30^{th} percentile without pooling)

Removal of monuments led to a large increase of vulnerability over the whole study area (see Figure 4.27). Vulnerability also rose considerably for removal of duration, and percentage of lowest household income but only at specific regions. These were: north of w_5 and w_7 for duration; northwest and southwest of w_7 for lowest household income. On the contrary, vulnerability variation decreased for the majority of study area when 'land use' indicator was not included. In this case, the canals in districts w_0 and w_1, and the north side of w_7 showed almost no change though. 'Ontwatering' removal seemed to have a negative influence on vulnerability but the drop was lower than that removing land use. Regarding the indicators: public buildings, shops, and buildings before 1960, there was

a minor decrease at the location of these objects but for the rest area, there was a slight increase (the lowest increase was noticed for 'shops' indicator). This vulnerability drop was expected given that the normalized index was maximum (i.e. 1) at objects' location whereas it was minimum (i.e. 0) anywhere else in the study area. The low weights of these indicators also led to an insignificant decrease. Furthermore, soil was expected to influence more vulnerability but the results did not show that; its weight was not high either.

The results of maps were verified by vulnerability boxplots in Figure 4.28. Monuments had the largest increase in median value followed by public buildings, shops, and buildings before 1960. Land use showed the largest median decrease; its variation also dropped based on the interquantile range. The lowest variation was observed for 'percentage of lowest household income' but there were more outliers. Regarding soil, the differences in vulnerability compared to the original one including all indicators were that minimum value and the 25^{th} percentile decreased slightly.

To determine the useful indicators for vulnerability estimation, the change in vulnerability was tested when each indicator was excluded. Large variation is desirable so that differences over the city can be spotted. In case of a larger variation (based on interquantile range) compared to the vulnerability derived using all indicators, the indicator should not be included. On the contrary, when variation decreased the indicator needs to be included since its removal led to high smoothing. Based on the aforementioned points, land use, and household income need to be included (see Figure 4.28).

A slight increase of variation was noticed for 'historical monuments' and 'public buildings'. This means that these indicators need may be excluded. Therefore, excluding these indicators would lead to a different vulnerability map. However, the fact that the differences are not pronounced decreases the certainty of the statement regarding exclusion. Based on the aforementioned points, the recommended indicators are: 'land use', and 'percentage of households belonging to lowest 40% income nationwide', whereas 'historical monuments' and 'public buildings' indicators are not suggested. It is worth highlighting that the results indicate that the two aforementioned variables would stress vulnerability differences; they do not state that only these two variables need to be used.

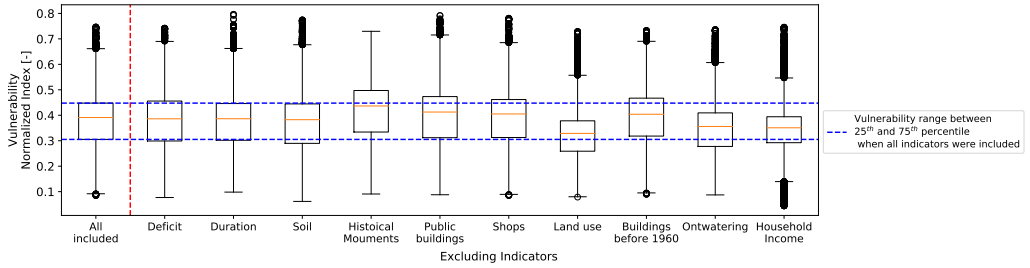


Figure 4.28: Boxplots for vulnerability to groundwater droughts when each indicator was excluded ('household income' represents the percentage of households belonging to the lowest 40% income nationwide)

Sensitivity for indicators' weights

Sensitivity of weights for vulnerability estimation showed interesting results (see Figure 4.29). Deficit and duration were not sensitive. For the physical sensitivity indicators derived by points (such as monuments, public buildings, shops), the median dropped as weight increased but the median increased for the other physical sensitivity indicators derived by polygon data (not points) (such as soil, land use and ontwatering). This was logical given that for indicators derived by points, the majority of the study area was no sensitive (value=0). Therefore, increasing their weights led to lower median. The variation kept almost constant for all vulnerability indicators. In general, the majority of one indicator's values in the study area affected how a change in its weight influenced vulnerability (positively or negatively).

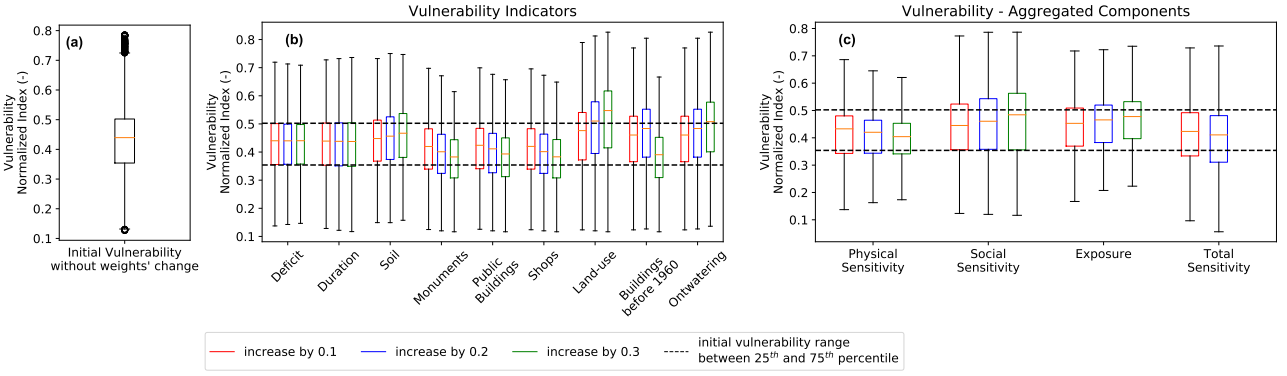


Figure 4.29: Original vulnerability to groundwater droughts without changing indicators' weights (a), sensitivity of vulnerability by changing the weight of each indicator (increase by 0.1, 0.2, 0.3) (b), sensitivity of vulnerability by changing the weight of each aggregated component (increase by 0.1, 0.2, 0.3); the case of total sensitivity and increase by 0.3 was not included since it exceeded one (c). Exposure was derived by variable threshold of 30$^{\text{th}}$ percentile without pooling.

Regarding the aggregated components (see 4.29-c), variation of physical and social sensitivity decreased and increased respectively as their weights rose. The explanation is that physical sensitivity was derived by a host of indicators, therefore extremes averaged out. In this way, sensitivity had smooth values. By increasing its weight, the smoothing of vulnerability rises. Additionally, by increasing the weight of physical sensitivity, the weight of social sensitivity decreased which included extremes given that it consisted of only one indicator. That explains the higher variation for social sensitivity when its weight rose.

Key results

- Sensitivity compared to exposure influenced more vulnerability spatial patterns; therefore more attention needs to be paid to collecting and analyzing sensitivity indicators than exposure ones.
- Coarse and fine resolution vulnerability results were in agreement but fine resolution provides more valuable information on local vulnerability.
- Including or not subsidence for exposure estimation led to noticeable difference in exposure spatial pattern but not in vulnerability given that sensitivity weight was much higher than exposure.
- The selection of drought identification technique for groundwater droughts affected spatial pattern of exposure considerably but this degree of influence decreased regarding vulnerability.
- The difference between 'no pooling' and 'median GWL' cases seemed to keep its spatial pattern from exposure to vulnerability at a higher degree than the difference between variable threshold of 30$^{\text{th}}$ percentile with and without 10-day pooling.
- Excluding 'land use' and 'lowest household income' decreased variation; therefore, these indicators are important to highlight differences in vulnerability. On the contrary, 'historical monuments' and 'public buildings' indicators may be excluded since their removal led to larger variation. There is high uncertainty regarding these results since the change in variation was not great.
- The majority of one indicator's normalized values in the study area affected how a change in its weight influenced vulnerability (positively or negatively). More specifically, an increase of physical sensitivity indicators' weights decreased vulnerability only for those indicators derived by sensitive elements with a limited coverage such as shops, and monuments.
- Vulnerability is not sensitive to the weights of soil and land use classes.

4.4.2 Soil Moisture Droughts

As mentioned above, the fine resolution is presented in this Section whereas the coarse one is attached in Appendix F. Firstly, vulnerability to soil moisture droughts over the study area is presented. Next, how the selected technique for drought identification affected exposure and vulnerability is highlighted. Sensitivity analysis for the indicators were also investigated to underline the importance of each indicator to vulnerability estimation. Last but not least, how the AHP weights of indicators influence vulnerability was examined.

Fine resolution

Regarding fine resolution, exposure indicators (duration and deficit) were in coarse resolution (see Figure 4.30) since the computed soil moisture by UWBM concerned the whole district. These indicators are based on district level whereas the sensitivity indicators had a finer resolution. Population density was not as fine as the other sensitivity indicators since data was recorded at neighborhood level (per buurt). Its highest values were observed in the center of Leiden as was expected. 'Distance to canals' indicator had high values all over the study area, whereas for soil, they were located on a north west stripe of w_7 which is covered by old dunes. The relationship between vulnerability and distance to canals was negative which means that areas further from the canals were less vulnerable. As for 'green areas' indicator, high values were observed at parks' location which are scattered over the study area. The largest continuous area of high values were on the north side of w_7 district.

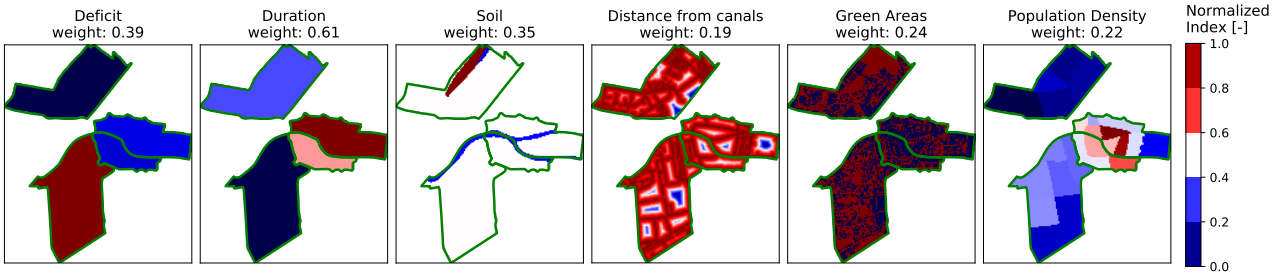


Figure 4.30: Exposure and sensitivity indicators for soil moisture droughts at downscale level for Leiden (the unit is dimensionless); their weights are also presented; the sum of weights for deficit and duration equals one and the sum of the rest indicators' weights equals one

Regarding vulnerability (see Figure 4.31), there was no clear cluster of high values; two large continuous areas were observed on the north side of w_7, and at the center of w_5. The rest high values were scattered. Another point is that the pattern of sensitivity component affected vulnerability as was noticed for groundwater droughts; the reason was the much higher sensitivity weight over exposure (almost threefold).

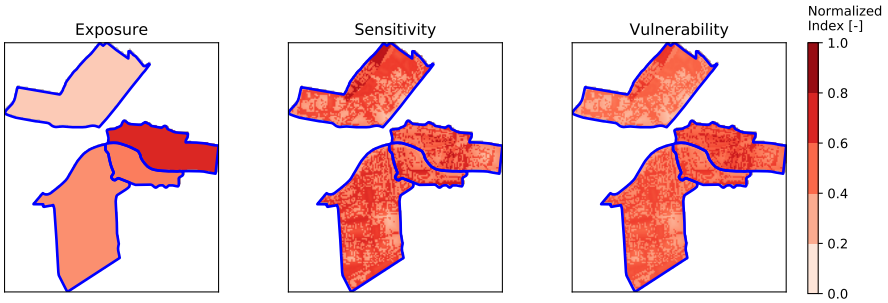


Figure 4.31: Exposure, sensitivity, and vulnerability for soil moisture droughts at downscale level using fixed threshold and 40^{th} percentile for the estimation of exposure indicators (the unit is dimensionless for all cases)

Comparing exposure and vulnerability based on drought identification techniques

The spatial differences in exposure and vulnerability among four different drought identification techniques are illustrated in Figure 4.32. More specifically, the following three techniques (fixed threshold of 30^{th} percentile with 15-day pooling, SMDI, and SMA) were compared to fixed threshold of 40^{th} percentile without pooling. Differences in vulnerability were coarse (per district) since (i) sensitivity was constant, and (ii) exposure resolution was coarse because deficit and duration were coarse. The latter ones were provided per district.

Vulnerability difference between the drought identification techniques were smooth compared to exposure difference (the maximum range in exposure was [-60, 60] whereas in vulnerability was [-15, 15]). The spatial pattern of exposure difference was similar as the vulnerability one for each of the pairs: (i), (ii) (see Figure 4.32-a,d and b,e respectively). These pairs were: (i) fixed threshold of 30^{th} percentile including 15-day pooling and fixed threshold of 40^{th} percentile without pooling, and (ii) SMDI and fixed threshold of 40^{th} percentile without pooling. The aforementioned relationship between exposure and vulnerability did not hold comparing SMA and fixed threshold of 40^{th} percentile (see Figure 4.32-c,f). It was not clear why the latter happened.

Another point is that SMDI and SMA had similar spatial differences in vulnerability (see Figure 4.32-e,f). Their vulnerability patterns were different than the one using fixed threshold of 30^{th} percentile including 15-day pooling (see Figure 4.32-d). The main explanations are: (i) the different methods to estimate drought indicators, and (ii) the fact that maximum duration and deficit were used for the case of fixed threshold 30^{th} percentile including 15-day pooling whereas for the other cases the mean ones were used.

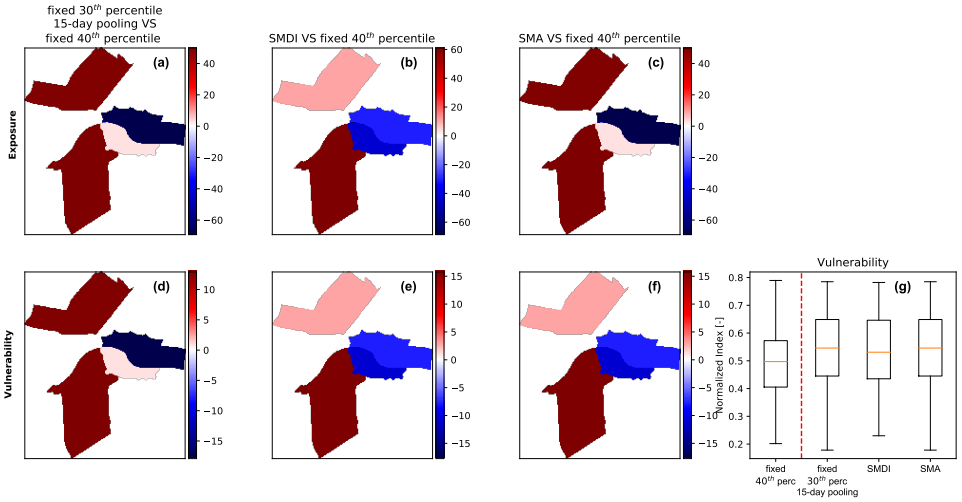


Figure 4.32: Difference in exposure and vulnerability (10^{-2} unitless) using the following techniques for the estimation of exposure indicators (fixed threshold of 30^{th} percentile and 15-day pooling, fixed threshold of 40^{th} percentile without pooling, SMDI, and SMA)

In general, the drought identification techniques which led to larger variation are preferred. However, the fact that vulnerability was not validated creates high uncertainty. An inventory of damages to houses, infrastructure could be useful. However, it would not be that straightforward given that there is a variety of reasons which could cause the same damage. Hence, the most effortless drought identification technique is suggested for the time being since the differences in vulnerability among the techniques are minor and validation was not performed.

Comparing SMDI to SMA regarding exposure, SMDI is preferred for the two following reasons: (i) it does not consider permanent wilting point and field capacity values which may be wrong (in most cases assumptions are made about these values), and (ii) it considers the mean soil moisture value of each week of the year; therefore, the estimation of deviation from normal conditions is more accurate. On the contrary, for SMA estimation, the mean of the whole time series is used. Comparing the soil

moisture indices (SMDI and SMA) to fixed threshold, the drawback of the fixed threshold is that it can be used satisfactorily only for summer period. Based on the aforementioned points, SMDI is suggested for soil moisture drought identification.

Sensitivity Excluding Indicators

By removing all indicators (one at a time) except 'green areas' and one district of 'duration' and 'population density' (i.e. w_5, and w_7 respectively) vulnerability decreased for the majority of the study area (see Figure 4.33). More specifically, a large drop over the whole study area was noticed removing 'distance from canals' indicator. Removing 'green areas' indicator, vulnerability decreased at the locations of green areas but for the rest area increased slightly as was expected. Soil had a neutral effect on vulnerability for a significant part of the study area whereas for the rest one vulnerability decreased slightly. Soil did not show to influence vulnerability considerably for groundwater droughts either.

Soil needs to be excluded for vulnerability estimation since vulnerability variance rose (based on the interquantile range) (see Figure 4.34) when soil was removed meaning that this indicator smoothed the differences. Green areas indicator was quite useful since without it there were no large differences in vulnerability. For the rest indicators, vulnerability variation increased compared to the case 'all indicators included' but not as much as soil. The variances for deficit and duration had almost the same magnitude even though it seemed that duration's variation was larger based on Figure 4.33 (2$^{\text{nd}}$ sub-figure). As for the case excluding population density, its median was the maximum compared to those of other indicators; therefore, this indicator may be excluded.

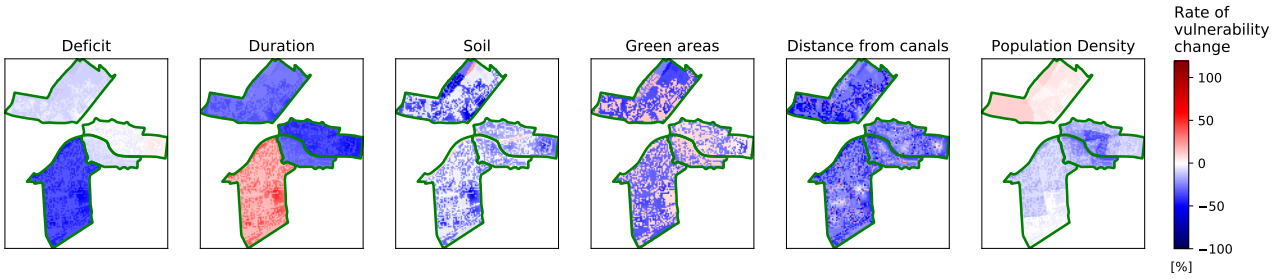


Figure 4.33: Rate of vulnerability change (%) excluding indicators regarding soil moisture droughts (exposure was derived by fixed threshold of 40$^{\text{th}}$ percentile without pooling)

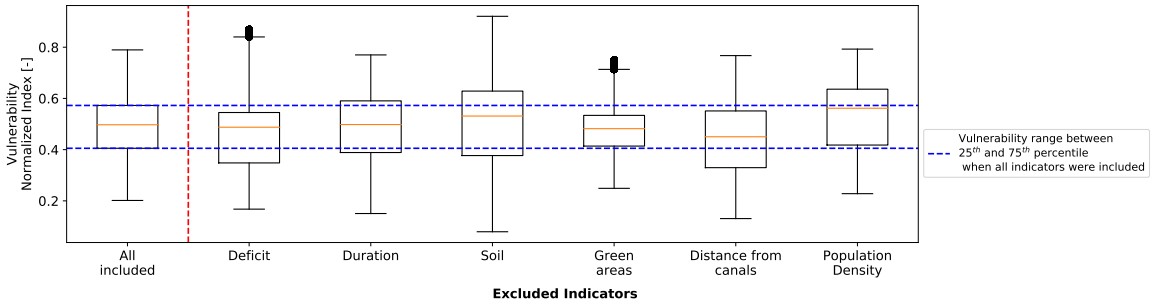


Figure 4.34: Boxplots for vulnerability to soil moisture droughts when each indicator was excluded

Sensitivity for indicators' weights

Figure 4.35 presents the sensitivity of vulnerability indicators' and aggregated components' weights for soil moisture analysis. They showed a larger sensitivity than those for groundwater droughts. The reason is that fewer indicators were used; hence, smoothing was not that high. The most sensitive indicator was green areas based on interquantile range. The only values used for green areas were minimum and maximum (i.e. zero and one respectively); therefore, increasing its weight led to higher variation. It is interesting that variation increased mostly between 25$^{\text{th}}$ percentile and its median.

Variation for soil decreased as its weight rose. As for soil classes, most of the study area is covered by trench and its vulnerability value was equal to 0.5. Therefore, by increasing the soil weight, the smoothing was higher. Consequently, 'green areas', and 'soil' indicators had the most sensitive weights.

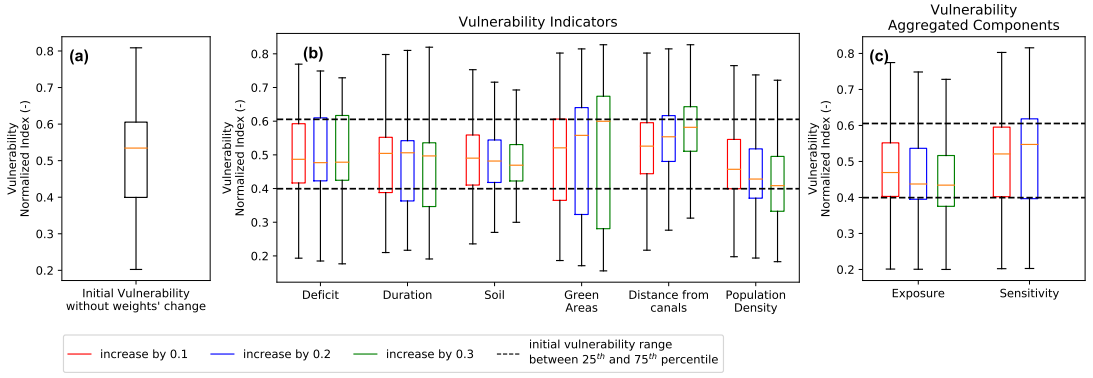


Figure 4.35: Original vulnerability to soil moisture droughts without changing indicators' weights (a), sensitivity of vulnerability by changing the weight of each indicator (increase by 0.1, 0.2, 0.3) (b), sensitivity of vulnerability by changing the weight of each aggregated component (increase by 0.1, 0.2, 0.3) (c). The boxplot for the case of sensitivity and increase by 0.3 in sub-figure (c) was not calculated since its weight exceeded 1. Exposure was derived using fixed threshold of 40^{th} percentile without pooling)

In addition for the indicators, 'distance from canals', and 'population density', their median values were increased and decreased accordingly as their weights increased. The explanation is that for 'distance from canals' indicator, the majority of the study area was covered by large values; thus, increasing its weight, vulnerability rose. A similar explanation applies for population density for which most area had low values. As a conclusion, the majority of one indicator's normalized values in the study area affected how a change in its weight influenced vulnerability (positively or negatively). That held for groundwater droughts as well. In case there was not a clear majority of high or low values for a vulnerability indicator, the vulnerability difference in median value due to a change in its weight was not noticeable (e.g. deficit).

Key Results (only for soil moisture)

- In fine resolution, vulnerability pattern followed the sensitivity one. High values of vulnerability were scattered and no hot-spot was identified.
- Differences in vulnerability using different drought identification techniques were smooth compared to those in exposure. The fact that validation was not conducted decreases robustness.
- The spatial pattern of differences between SMA and fixed threshold of 40^{th} percentile changed from exposure to vulnerability.
- Using SMDI and SMA for soil moisture drought identification, vulnerability showed a similar spatial pattern but their pattern was different than using fixed threshold of 30^{th} percentile including 15-day pooling.
- 'Green areas' and 'soil' indicators need to be included and excluded respectively for soil moisture vulnerability estimation given their vulnerability variation. Population density is a another indicator which may be excluded since its removal led to larger vulnerability median and variation than the case where all indicators were included. As highlighted earlier for groundwater droughts, these results do not indicate that only one indicator ought to be used.
- The weights of all used indicators (in particular 'green areas' and 'soil') were highly sensitive to vulnerability estimation for soil moisture droughts.

Comparing groundwater and soil moisture drought vulnerability

- In coarse analysis, vulnerability followed the pattern of sensitivity regrading groundwater droughts but there were slight differences for the case of soil moisture droughts between sensitivity and

vulnerability.

- In fine resolution, vulnerability followed the pattern of sensitivity for both groundwater and soil moisture droughts.
- Using different techniques to identify droughts affected exposure noticeably but not vulnerability for both groundwater and soil moisture drought analyses.
- Soil indicator needs to be excluded for soil moisture analysis. For groundwater analysis, neither strong inclusion nor exclusion was suggested for this indicator. The other sensitivity indicators used were different; therefore, no comparison regarding them was possible.
- Vulnerability estimation was more sensitive to AHP indicators' weights for the case of soil moisture droughts than GWL ones.
- The majority of one indicator's normalized values in the study area affected how a change in its weight influenced vulnerability (positively or negatively) for both groundwater and soil moisture droughts. In case there was no clear pattern of high or low values for the indicator, the vulnerability difference in median value due to a change in its weight was not noticeable.
- Indicators derived from non continuous elements showed a drop (in average) in vulnerability as their weights increased; that applied to monuments, public buildings, and shops for groundwater droughts. However, it did not hold for green areas in soil moisture droughts analysis since they covered a significant amount of area.
- Changes in the weights of soil classes regarding soil moisture droughts could affect vulnerability more than those in groundwater drought analysis. The reason was that fewer indicators were used for vulnerability to soil moisture droughts. As a result, the weight of soil indicator was higher. This conclusion is based on Section F.2 in Appendix F.

4.5 Evaluation

The objective of this thesis is (i) to define urban drought in a relevant and operational way and (ii) to estimate the vulnerability of Leiden to groundwater and soil moisture droughts. To fulfill this goal, the analysis focuses on answering the following research questions.

• How to define an urban drought?

Due to the complexity of urban droughts, a universal definition applicable to all cases seems not to exist in current literature. For that reason, an urban drought categorization framework was created where droughts are classified into four classes based on the affected water cycle parts. These four categories are: groundwater urban drought, soil moisture urban drought, open water urban drought, and water supply urban drought. Only the two first ones were investigated in the current analysis. A long, generic definition was created incorporating all four classes, however, there are still issues to be clarified.

• How can an urban drought be quantified? How can the inherent characteristics of groundwater and soil moisture urban drought be analyzed?

The drought characteristics analyzed were: deficit, duration, spatial extent and frequency. High computational cost and lack of data hindered the estimation of the latter two ones respectively. The analysis consists of groundwater and soil moisture droughts. Identification methods for droughts in general were tested whether they could quantify a groundwater and soil moisture urban drought.

- **Groundwater droughts**

  Groundwater level deficit and duration are evaluated using four types of threshold methods: (i) fixed, (ii) variable, (iii) moving window, (iv) and median GWL. Deficit and duration identified via variable threshold provided satisfactory results even though lack of data in the current study hindered a much higher performance. On top of that, variable threshold showed the highest

consistency for drought deficit over all percentiles compared to fixed and moving window thresholds. A drawback of variable threshold is that the statistical process is biased. The reason is that moving average is included for the calculation of variable threshold which increases the interdependency of drought events. Fixed threshold could also identify droughts events satisfactorily where the main goal is related to construction, and foundation of buildings. Moving window was the technique with the worst performance. Using the median GWL as threshold provided the more realistic results but long data is required which is not often available; therefore, modeling is required to predict these values but uncertainty may be high. In general, a combination of fixed and variable threshold is appropriate to characterize the drought status in a specific area; it provides a more profound understanding than using only one of those. Another way to estimate the threshold for droughts identification is with the help of stakeholders based on the water functions of the city. However, this approach is more complicated than it looks like. Different water functions would require a different threshold but only one threshold can be applied.

To deal with large drought events which were split into more than one, pooling was considered. For the identification of the ideal pooling, the relationship between deficit, duration and inter-event time was established. A period of around 10 days was considered as a satisfactory inter-event time of pooling for both deficit and duration for variable threshold of 30^{th} percentile. 10 days onwards, there was a plateau regarding deficit and duration which lasted around 10 days. A inter-event time more than 20 days is too long which is not suitable since independent events may be merged. The latter case could distort reality. The larger the inter-event time, the higher the possibility of merging droughts which are independent and should not be merged. Therefore, an inter-event time which satisfies the following conditions was desirable: (i) most drought events which happened closely time-wise were merged, (ii) slight increase or decrease of the inter-event time did not change the results significantly, and (iii) the inter-event time in itself is not that long. The relationship between inter-event time and deficit could be weird for some wells; this did not happen for duration, though. This is explained by the different ways pooling was performed for deficit and duration. Additionally, there is not one right inter-event time which works perfectly in all conditions. This research tries to find the inter-event time for which most minor drought events are merged for the groundwater wells used in the study area.

Areal extent of groundwater drought was also analyzed using spatial analysis. The results suggest that there is a threshold relationship between deficit, duration and areal extent in pairs. When a specific deficit or duration value is exceeded, the majority of the study area suffers from droughts. Deficit and duration were calculated from spatial analysis. Applying spatial analysis only to one district is not enough to draw a valid conclusion, though. It is worth noting that these results apply only for the spacial scale used (i.e. district level). They cannot be generalized. Spatial scale may affect the relationship between drought characteristics. In general what holds regarding droughts at a large spatial scale may not hold at a fine spatial resolution by default. Further research is required to verify or refute that.

Frequency distribution analysis using the minimum of monthly blocks can be promising but a peculiar pattern in modeled GWL hindered this analysis. Uncertainty was high due to the mediocre performance of UWBM calibration. GWL did not decrease below a specific threshold for each district throughout the study period. In general, its high dependency on modeling (given that in most cases there is no long measured data) is a critical drawback. No robust conclusions can be drawn for that drought characteristic since any conclusion is related to the model used.

- **Soil moisture droughts**

  Soil moisture was modeled via UWBM. Its performance influenced the values of soil moisture. There was high heterogeneity in GWL measurements which were used for calibration; therefore, UWBM performance was not ideal. Regarding deficit and duration estimation, a combination of fixed threshold in summer in conjunction with variable threshold could provide deeper un-

derstanding. Vegetation suffers more for shorter amount of time when there is not enough soil moisture; consequently variable method is important for that case. Moving window was unsuccessful as was for groundwater droughts. Regarding pooling, the relationship between deficit and inter-event time for variable threshold showed a peculiar behavior; therefore, pooling is not suggested.

No conclusions were derived regarding spatial extent and frequency of soil moisture droughts. Areal extent of soil moisture droughts was not estimated given that the output soil moisture from UWBM was lumped per district. Regarding the frequency distribution analysis of soil moisture droughts, an irregular pattern was noticed and the analysis was dismissed. The modeled soil moisture did not decrease below a specific threshold for district w_5 over all years. The performance of UWBM in that region was not satisfactory. This could explain that behavior. In general, the fact that there is no long data available for soil moisture leads to modeling so that the long data can be created. However, the modeled soil moisture depends on the model used. This is a critical drawback.

• How can the vulnerability of cities to soil moisture and groundwater droughts be determined?

This was the main research question and the case study consisted of four districts in Leiden, the Netherlands. The analysis consists of two parts: groundwater and soil moisture urban droughts. The suggested method was via aggregation of exposure, sensitivity, and coping capacity (the latter was omitted) following the framework which Fritzsche et al. (2014) suggested. AHP was used to determine indicators' weights from the expert judgments of eight senior experts in this field.

- **Groundwater droughts**

  The selection of drought identification technique for groundwater droughts affected exposure considerably but this degree of influence decreased regarding vulnerability. Therefore, even applying a sophisticated technique to identify drought events, will not lead to significant improvement regarding vulnerability estimation. Hence, a technique which is straightforward and effortlessly applicable can be selected since the changes in vulnerability are minor. In that case, variable threshold of 30^{th} percentile without pooling is suggested.

  Regarding the vulnerability indicators, 'land use', and 'percentage of households belonging to the lowest 40% income nationwide' are useful for vulnerability estimation since their removal decreased the variation in vulnerability. The second indicator shows whether the residents can afford possible repairs in their houses. Large variation in vulnerability is desirable since differences over area can be more noticeable; therefore, hot spots of high vulnerability can be pinpointed. In that way, priorities can be set and protection measures can be more effective. Another point is that 'historical monuments'and 'public buildings' indicators are not suggested since when they were excluded, vulnerability variation increased. Consequently, places which need more and less protection can be highlighted. Last but not least, it is worth noting that including subsidence (as exposure indicator) or not did not seem to affect vulnerability despite its large weight (for the case of including it); the explanation is that exposure weight is much lower than sensitivity one.

- **Soil moisture droughts**

  Vulnerability followed sensitivity pattern in fine spatial resolution analysis but there were minor differences in coarse one. Regarding the suggested indicators for soil moisture vulnerability determination, 'green areas' needs to be included whereas 'soil', and 'population density' to be excluded. The two latter ones increased the vulnerability variation and their removal could help in underlining differences. Another point is that the AHP weights of 'green areas' and 'soil' were highly sensitive to vulnerability estimation.

As for drought identification techniques, their difference in exposure was pronounced but smooth in vulnerability. Therefore, there is no need for a complex technique.

- **Comments applicable on both groundwater and soil moisture droughts**

  By estimating vulnerability at fine spatial resolution, water managers would be able to determine the priority hierarchy of different regional drought mitigation measures; they will not distribute the resources equally over the city, either. In this way, resources are not misused, planning could be effective, and policy makers would be able to make concrete decisions.

  Regarding the AHP weights, the majority of one indicator's normalized values in the study area affected how a change in its weight influenced vulnerability (positively or negatively) for both groundwater and soil moisture droughts. For instance, for those indicators derived by sensitive elements with a limited coverage (such as monuments in case of groundwater droughts), most area had low values and an increase of their weights could lead to lower vulnerability. This result is important since it shows how indicators' weights influence vulnerability and in general raises awareness about how vulnerability is estimated. This technique is not like a black-box method. Therefore, policy makers can have a strong understanding about it which can lead to better decisions and more critical assessment of the inputs and results.

Chapter 5

Discussion

This chapter includes the meaning of the current study stressing how it contributes to the expansion of knowledge regarding urban droughts. Firstly, the results of the current research are compared to those from other studies (Section 5.1). Studies in urban and rural environment are considered. Secondly, reflections on some critical points are underlined in Section 5.2. Thirdly, the limitations of this research are highlighted (in general and for UWBM specifically) (Section 5.3).

5.1 Comparing the results of this research to current literature

Most literature regarding vulnerability estimation to droughts is about rural areas; there are not many studies regarding urban ones. This is the main gap which this research tried to cover. In this part, the present study is compared to other studies (i) whose study areas were cities, and (ii) whose study areas were in rural environment.

Comparison with studies regarding urban areas

One important outcome of the present analysis is that the weight for social sensitivity was half as much as physical sensitivity for groundwater droughts However, this result was not in line with the study by Chang et al. (2015) who found that in some Chinese cities social sensitivity contributed the most to vulnerability. The results are region specific; therefore, spatial variations are expected. Another explanation is that entropy evaluation was applied to estimate weights in that study but for the current one, Analytical Hierarchy Process (AHP) was employed.

Another difference is about the population density which was used as vulnerability indicator. It was the indicator with the second lowest weight for the estimation of vulnerability to soil moisture droughts in the current research. This result is not in line with the study by Chang et al. (2015) based on which this indicator had the highest weight. However, they studied the vulnerability of cities to droughts in general; not to a specific drought category. This could be an explanation. On the contrary, the study by Yuan et al. (2015) is in line with the present research since the weight for population density was also low. It is worth highlighting that for the study by Yuan et al. (2015), the weights were estimated using AHP but Chang et al. (2015) used entropy evaluation.

Two other important differences among the current study and others are related to: (i) the level of analysis in the current research, and (ii) the fact the other studies were about droughts in general. Regarding the element (i), in the studies by Chang et al. (2015); Yuan et al. (2015), vulnerability among different cities were compared. On the contrary, in the current analysis, regions over one city were compared. As for the (ii) element, the previous two studies were not specific to a drought category; therefore, indicators such as water supply and rate of effective irrigation were used which are not applicable to the current analysis.

Comparison with studies regarding rural areas

One main difference between the current study and those focused on rural areas conducted by Gurwin (2014); Kim et al. (2015); Thomas et al. (2016) is the number of aggregation levels. In the current analysis for groundwater droughts, four levels of aggregation were applied (exposure, physical sensitivity, total sensitivity, and vulnerability). Social sensitivity consisted of one indicator; therefore, aggregation was not performed. Regarding the studies focused on rural areas, only one aggregation was performed including all indicators. This difference highlights the importance of the weights for 'exposure' and 'sensitivity' components in the present research. For instance, if the weight for one exposure indicator is high but the weight of 'exposure' component over sensitivity is not high, the indicator cannot affect the vulnerability significantly. If one aggregation level was considered, then its influence to vulnerability would be larger. On the other hand, it is less chaotic to use more aggregation levels and uncomplicated for the experts to compare the indicators in AHP.

Another difference is that for the majority of the studies regarding vulnerability to agricultural droughts (e.g. Gurwin (2014); Kim et al. (2015); Nasrollahi et al. (2018); Shahid and Behrawan (2008); Thomas et al. (2016)), standardized indices were considered for exposure estimation. The frequency of meteorological droughts in combination with their severity were used as exposure indicators. None of these studies used threshold method. Therefore, a comparison regarding the performance of the latter technique cannot be performed between rural and urban environment.

A disagreement on the weights of exposure and sensitivity was noticed between the present study and current literature regarding agricultural droughts. Frequency of droughts was the indicator with the largest weight (50%) in the study performed by Gurwin (2014); there was only one aggregation level. However, regarding the current analysis, sensitivity's weight was three times as much as that of exposure. It is worth noting that Gurwin (2014) followed a different approach for vulnerability estimation (a different framework); therefore, the comparison may not be completely valid. However, it can explain the general pattern of their differences.

Regarding the indicators, those which were used for droughts in both urban and rural environment were: land use (Gurwin, 2014), economic status (Shahid and Behrawan, 2008), and population density (Kim et al., 2015; Nasrollahi et al., 2018). However, the majority of the variables used for agricultural droughts could not be used for urban droughts.

Another point is that the current study focuses on a developed country whereas the majority of studies focus on a developing one. This is an indication that in developing countries, they may be mostly interested in agricultural droughts since their economy is highly dependent on the primary sector. However, this priority needs to be changed since cities in these countries are exposed to urban droughts, too. The sensitivity of these cities would probably be greater than that of developed countries which means that they may suffer severe damages from droughts.

Importance of this research

The main contribution of this thesis are: (i) the urban drought categorization framework, and (ii) the suggested method to determine vulnerability of cities to groundwater and soil moisture droughts at a fine spatial scale. Firstly, concerning the importance of the urban drought categorization framework, researchers would be able to focus only on one drought class providing more accurate results. Policy makers and stakeholders can determine which type of urban drought takes place in a city which is catalytic for planning and urban drought mitigation. It may also be the incentive to raise awareness in cities to evaluate their vulnerability to types of droughts for which they do not face challenges at the time being but they will due to climate change. Secondly, the current research not only determines the vulnerability of cities to groundwater and soil moisture droughts but also provides a better understanding into (i) which vulnerability indicators are important, (ii) how the drought identification techniques affect vulnerability, (iii) how changes in the indicators' weights influence vulnerability. In that way, this research provides a detailed steering regarding vulnerability estimation

of cities.

5.2 Reflections

Evaluation of the proposed vulnerability method

The main advantages of the technique to determine vulnerability in cities applied in this thesis are: (i) simplicity, (ii) low computational cost, and (iii) applicability at two spatial levels (fine and coarse resolution). The technique suggested is intuitive; therefore, policy makers would be able to understand it and cooperate with water management engineers. Furthermore, this method is not computationally expensive given that only deficit and duration are used as exposure indicators. However, if areal extent and frequency of droughts are used as exposure indicators, the spatial analysis and frequency distribution analysis can increase the computational time. Another critical advantage is that the technique suggested can be performed at two different spatial levels (coarse and fine resolution). Regarding coarse resolution, the analysis can be performed for administrative units (such as districts, or municipalities, etc). As for fine resolution, raster analysis is performed which can be as fine as the indicators' resolution allows. The most appropriate resolution can be selected based on the goal of the analysis.

As validation using damage data cannot be done, there is high uncertainty in these results. Therefore the suggested methodology cannot be evaluated thoroughly. Another drawback is that this technique requires much sensitivity data but generally speaking every technique would. Availability and quality of the indicators' data influences vulnerability results but this is a generic conclusion. Given that the necessity to model exposure in detail is less important than to have detailed information on sensitivity, the UWBM can be considered accurate enough for this type of vulnerability analysis.

Reflections on Urban Water Balance Model

UWBM was able to represent the general pattern of GWL for some monitoring wells in each district despite the fact that the representation of the groundwater system in the UWBM is simplistic; the UWBM was never built to model GWLs. There was high GWL spatial heterogeneity in Leiden even though the wells were located close to each other. Large variation was unexpected and the surrounding land use could not explain the wells' pattern. Human intervention to the water system in Leiden could be one of the reasons but related data was not available. In general, lumped models may not be the most appropriate option and distributed models could provide better results given the high GWL variation.

Out of the four parameters (w, down_seepage_flux, storcap_ow, q_ow_out_cap) used for calibration, q_ow_out_cap showed the least sensitivity. This was expected since this parameter was not related directly to GWL changes. On the contrary, water balance in Leiden was highly sensitive to down_seepage_flux.

Even though the performance of UWBM could be improved, this model is suggested to give the first idea of similar studies. It is not computationally expensive and the only inputs are basically precipitation and evaporation which can be found effortlessly in developed countries. Some drawbacks of UWBM are highlighted in Limitations (Section 5.3).

Reflections on the case study of Athens

Vulnerability was not estimated for the case study of Athens. In general, the fact that data was found for Leiden and not for Athens highlights the linkage of water and society for the two cities. On one hand, Leiden is a city in lowlands where there is high awareness of water conditions (as opportunity and threat). On the contrary, in Athens, in the peoples' perception, water is linked to drinking use and its other links were and are ignored. The focus is only on water supply urban drought based on the suggested framework. This explains why there is lack of data. However, it does not mean that

Athens is not vulnerable to groundwater and soil moisture urban droughts. It may be but it was not tested. Data recording could be the first step towards the estimation of vulnerability. It is not only about soil moisture and GWLs though. Precipitation and other weather data need to be recorded at a finer resolution too.

The social sensitivity component of vulnerability

Regarding social sensitivity, only one indicator was used. A variety of variables was available regarding population, economy, and society but after redundancy analysis only five ones were saved; two for population, two for economy, and one for society. Drought experts found it difficult to compare them in AHP method. For that reason, only one indicator was used. The number of drought experts in the AHP was limited. Therefore, there is high uncertainty about rejecting the indicators. Furthermore, dismissal of four out of the five indicators may be considered as an extreme approach. Using two or three indicators could lead to a better representation of social sensitivity. Besides 'percentage of households belonging to the lowest 40% income nationwide', the two most likely ones are: percentage of population older than 65 years old, and percentage of the disabled people.

Using several indicators is undesirable for vulnerability estimation since it complicates the pair-comparison regarding weights in the AHP method. Additionally, the fewer indicators there are, the lower the cost and effort required for vulnerability determination. Minimizing the number of indicators without losing information could be a recommendation for further research. It is analyzed in Section 6.2 (click here). On the other hand, using only one indicator for social sensitivity may be extreme and may lead to misleading results since critical data is ignored.

Density of Sensitivity Indicators

A point which raises discussion is the density of observations which is required to represent sensitivity at a satisfactory degree. The case of groundwater droughts is discussed. A fine resolution may not be worthwhile at least at this first phase. First of all, there are two assumptions which increase uncertainty: (i) residents own the houses they live in; therefore, they bear the financial burden regarding repairs related to damages caused by droughts, (ii) the age of the building implies the material of the poles (i.e. for buildings erected before 1960, their poles were made of wood). It is possible that these two assumptions may not hold. Therefore, the results regarding vulnerability may be wrong. Consequently, even if sensitivity data is recorded at a fine resolution, the aforementioned assumptions can disrupt the vulnerability results. Another point which was not considered in the current analysis is that it is possible that some buildings have been repaired before. Hence, they will not suffer any damage related to droughts. In that way the vulnerability would be much lower. Additionally, neither the styles of timber pile foundation nor their wood types were considered. There are three types of timber pile foundations in the Netherlands: the Rotterdam foundation, the Amsterdam foundation and a Rotterdam foundation for which the pole is connected with a concrete beam instead of a wooden one (Schreurs, 2017). Amsterdam foundations are more robust since more piles are used (Schreurs, 2017). The wood type and the pile foundation style are two characteristics which affect the sensitivity of the piles to groundwater droughts.

Given the points mentioned above and the fact that a deep understanding of vulnerability of cities to groundwater droughts has not been established yet, a coarse resolution is suggested for the time being. However, a finer resolution could be useful in the future when (i) fewer assumptions would be considered, and (ii) the relationship between drought characteristics and damages would be more clear.

Quality of the current monitoring network regarding GWLs in Leiden

The weight of sensitivity was much larger than that of exposure. This implies that not much attention needs to be paid to monitoring the network of GWLs. However, this does not mean that the existing network should not be improved or even expanded. There are three different companies (Acacia, ireal,

and Wareco) which monitor GWL in Leiden but more emphasis needs to be laid on Wareco network since the quality of the measurements was better than the other two companies. However, its duration was not long enough; therefore, this dataset was not used in this study.

Regarding soil moisture analysis, increasing the density of the GWL monitoring network will not lead to a better representation of reality given that UWBM is used. There were no measurements of soil moisture. Consequently, soil moisture was modeled via UWBM. The calibration was conducted using measured GWL. Using more than one monitoring well to calibrate the model did not provide satisfactory results. This is why increasing the density of the GWL monitoring network will not benefit soil moisture analysis. A physically based distributed model may be preferred in case the monitoring network is denser. There are drawbacks about this type of model, though. In physically based models, the number of calibrated parameters is large. Therefore, there is a high risk of equifinality (i.e. many parameter sets can provide satisfactory results). The possibility that a not suitable parameter set may be selected is high. It may misrepresent the dynamics of the hydrological system even though its performance in calibration is high.

For the time being, it is recommended that Leiden should enhance the quality of the current network (not increase its density yet). Acacia and ireal need to improve their in-situ measurements. The recorded GWLs by sensors are calibrated using in-situ measurements. However, the calibrated GWLs contain errors (i.e. unrealistic values) too. Increasing the frequency of in situ measurements can be the first step of handling this issue. In this way, irregular patterns and gaps in monitoring wells would be identified faster. Introducing an automated way to identify large rises or drops of GWL regarding the neighboring time steps can be useful too. The majority of these values tend to be errors. In case the aforementioned measures do not provide satisfactory results, it is recommended that the city of Leiden should extend the monitoring network of Wareco to the southern.

Regarding the case of using fewer wells for the fine vulnerability analysis, there would be some loss in exposure resolution given that GWL is heterogeneous in Leiden. The monitoring may not be that satisfactory. Therefore, there would be differences in vulnerability too. However, given that the weight of sensitivity is much larger than that of exposure, the differences would not be pronounced. In general, emphasis should be laid on the assumptions mentioned before (see Density of Sensitivity Indicators) since they can have a large influence on vulnerability. The improvement in vulnerability determination would not be noticeable, even if more wells were available, unless an optimized technique to identity droughts was used.

The method applied to fill in the gaps of GWL is another subject of discussion. In general, filling missing values can be a topic on its own. Linear interpolation was implemented but in many cases its results were conservative. Besides linear interpolation, a data-driven technique was used (i.e. sequential algorithm) but failed due to the lack of data. With more data (i.e. after some years), the performance of that technique would be higher.

Reflection on weights' estimation

In this study, the weights of vulnerability indicators were estimated using AHP. A series of judgments comparing the indicators in pairs helps to decompose a complex problem like the determination of vulnerability. The weights of indicators play the most important role in vulnerability estimation. If the weights are different, then the vulnerability output would be changed. Therefore, not paying attention to AHP could lead to misleading results.

The following points underline how reliable the weights estimated in this research are. First of all, the low number of droughts experts (eight) who participated in AHP decreases the credibility of the results. On top of that, for most of the indicators, the experts' weights diverged; the ranges of weights were large. A workshop at which drought experts could discuss their judgment would be helpful; it could lead to convergence regarding indicators' weights.

Secondly, using the 'extreme' weights (which varied the most than the average ones), the output vulnerability was different. This highlights the high uncertainty of the results, too. That holds for the vulnerability to groundwater and soil moisture droughts. For the latter type of droughts, the vulnerability differences were lower, though. This may be related to the fewer aggregation levels for soil moisture droughts.

Comparing the current procedure to determine vulnerability in Leiden to the one suggested in this thesis

Until two years ago, Leiden municipality just monitored the damages to buildings caused by droughts. It is highly likely that not all damages are recorded given that most buildings are privately owned. Leiden municipality outsourced the determination of vulnerability to Wareco two years ago. The indicators used were: (i) soil composition/structure, (ii) age of buildings, (iii) type of foundation (yes/no wooden poles), and (iv) groundwater levels but there were no weights attached to them. The analysis was applied at neighborhood level ('buurt'). In general, there is no clear methodology which the municipality of Leiden follows to determine its vulnerability to groundwater droughts. The current practice consists of three components: (i) recording the damages, (ii) analyzing the results of Wareco, and (iii) exploring different options to create new policies for the future.

Comparing the current practice in Leiden to the method suggested in this thesis, there are two benefits for the latter case. First of all, the resolution is finer; therefore hot spots of vulnerability may be identified which is important for planning mitigation measures. For instance, if the Drainage Infiltration Transport (DIT) System is applied, its implementation can be specific to a limited areal extent. In that way, the municipality spends less financial resources. DIT can drain and transport water; therefore GWL can be more stable. The second benefit is that assigning weights to vulnerability indicators may be the incentive for the municipality of Leiden to prioritize the assets of the city. A thorough discussion between drought experts and the municipality of Leiden can be conducted and the priorities of the city can be evaluated. Consequently, this thesis can enhance the current procedure applied for vulnerability estimation of Leiden to groundwater droughts .

Regarding soil moisture droughts, there is no procedure which the municipality of Leiden follows to determine the vulnerability of the city to that drought type.

Creating a Decision Support System (DSS)

The method suggested in this thesis, to determine vulnerability, has the potential to be converted to a Decision Support System (DSS). To be more accurate, it can be converted to a Spatial Decision Support System (SDSS). In SDSS, modeling is integrated via GIS and both spatial data and operations are used (Keenan, 1995). Data which is not spatial may be employed too.

Based on Power (2002); Sprague Jr and Carlson (1982), a DSS consists of three components: (i) the database, (ii) the model, and (iii) the graphical user interface. In this case, the database is about the indicators, and the model is the aggregation process to determine vulnerability (considering the indicators' weights derived by AHP). The whole process needs to be automated. The model may include the determination of drought characteristics but this is not a prerequisite. The most important component of the system is to provide the user the capability to modify indicators' weights. The user (decision maker) should be able to intervene in the system by changing only its parameters without encountering long series of commands. Regarding the model, given that (i) UWBM is used in Python, and (ii) ArcPy is mostly used in GIS projects, Python is suggested for the creation of the SDSS. As for the third component of a DSS (the graphical user interface), its creation is important given that some stakeholders and policy makers may not be familiar with scripting. Therefore, the process should be as straightforward as possible. Engagement of the policy makers during the design process of the user interface can be beneficial.

It is worth noting that a DSS may be expanded. If the current method is used for the creation of

a DSS, then the DSS will be considered as passive. A passive system helps the decision making but not explicitly since not direct solutions are the outputs of the DSS. However, it has the potential to become an active DSS providing suggestions and solutions. The system should include mitigation solutions and the impact thereof (i.e. how much each solution reduces vulnerability). To conclude, it is possible for the current approach to be converted to a DSS.

Coarse VS Fine resolution in vulnerability analysis

In the current method to determine vulnerability, two levels of spatial analysis are provided - coarse and fine. The coarse analysis was at district level. If the output resolution of vulnerability needs to be coarse, then the indicators may be coarse. In that case, there is no reason for the monitoring network to be dense since averaging is performed (for each district). For instance, even if three monitoring wells are available, their average can be used for coarse analysis. Coarse analysis could be useful where (i) there is a lack of data, or (ii) the goal is to obtain a first and quick idea of vulnerability.

On the contrary, if fine resolution is desired for the vulnerability output, distributed indicators are required. A fine resolution of indicators is not suggested for the time being as mentioned before. The financial cost should not be ignored either. A fine spatial resolution of vulnerability is desired when detailed spatial mitigation solutions are going to be implemented. Another case for which fine resolution would be useful is when the administrative units are large; therefore, coarse analysis is not that useful. Regarding the outputs of fine and coarse resolution, they almost followed the same pattern for groundwater droughts but there were some slight differences in soil moisture droughts vulnerability.

Ordered Weighted Average (OWA)

Ordered Weighted Average (OWA) is a method which could be used for the aggregation of vulnerability indicators. It is based on fuzzy set theory. Two sets of weights are included in that approach: the weights of criterion importance, and order weights (Malczewski et al., 2003). These weights are based on functions; therefore bias of drought experts is eliminated. For the determination of ordered weights, min and max operations can be applied which are based on the logical operators AND and OR respectively. There are other methods available to determine the weights, too.

5.3 Limitations

The limitations are categorized into two classes: (i) limitations for the whole study and, (ii) limitations for the UWBM specifically.

General Limitations

- Regarding the social sensitivity indicator which was used (percentage of households belonging to the lowest 40% income nationwide), it was assumed that the residents of the properties own them which is not everywhere true given that there are many rental housing agencies. Finding the owners was difficult, and as a result there is high uncertainty about the validity of this indicator.
- Validation of the vulnerability was not performed; vulnerability could be approximated by damage given that sensitivity could not be quantified. Exposure can be measured; the problem to validating vulnerability is in measuring sensitivity.
- Only eight drought experts participated to assign indicators' weights; a larger number of them was desirable. This reduces the credibility of vulnerability analysis. A workshop to seek more consensus on the weights would also help improve credibility.

Limitations of UWBM

- UWBM is a lumped model which means that GWL needs to be representative for the whole area but this has a cost on the accuracy; the bulge of groundwater which occurs between two drains (and adjacent canals) was not considered in this step of the analysis. However, this bulging

exists. To include it, another type of modeling (distributed and physically based models) would be required.

- Soil moisture content in the unsaturated zone is different at each depth below the ground surface. These differences were not considered given that soil moisture is considered as one component (entity) in UWBM.
- Vegetation affects soil moisture content. Grass shows higher short-term variability of soil moisture compared to tree vegetation but the deficit of soil moisture can be higher for tree vegetation (Wiesner et al., 2016). Leiden is covered by both trees and grass but only one vegetation type in UWBM (grass) was used. Therefore, the previous relationship between trees and grass was ignored and differences were averaged out. Including more than one vegetation type could be a useful addition to UWBM leading to more realistic results.

Chapter 6

Conclusion & Recommendations

6.1 Conclusions

The purpose of this thesis is to estimate vulnerability of cities to groundwater and soil moisture urban droughts. This topic is of great importance since the consequences of droughts on cities can be devastating for their residents in a variety of sectors such as economy and health. The focus is only on cities from developed countries and Leiden is selected as a case study given the availability of data.

An urban drought categorization framework is proposed which categorizes urban drought into four classes but in the current analysis, only two of them are studied: groundwater and soil moisture droughts. The drought characteristics (deficit, duration, areal extent, and frequency) for groundwater and soil moisture droughts are computed using a variety of techniques. More specifically, to determine drought deficit and duration for groundwater droughts four threshold methods are used. These are: (i) fixed threshold, (ii) variable threshold, (iii) moving window method, and (iv) median groundwater level as threshold.

The results show that variable threshold identifies the deficit and duration of drought events in a more realistic way throughout the year than fixed threshold and moving window. However, a combination of fixed and variable threshold can provide deeper insight into the drought status in an area. The combination of the two thresholds is suggested for soil moisture droughts, too. In order not to underestimate deficit and duration, pooling (merging of drought events) is required. A useful inter-event time for pooling is established for groundwater droughts (around 10 days) but not for soil moisture droughts. In that way, calculated deficit and duration for groundwater droughts can represent exposure more accurately.

For the other two drought characteristics (areal extent and frequency), spatial analysis and frequency distribution analysis are applied respectively. The most striking results are presented below. Areal extent seems to be positively related to both drought duration and deficit but spatial analysis was applied only to one district; thus, uncertainty is high. This characteristic was not estimated for soil moisture droughts since soil moisture was modeled for each district by an urban water balance model at district scale. The last drought characteristic is drought frequency but no conclusion is drawn for it given its high uncertainty.

As for vulnerability estimation, the framework suggested by Fritzsche et al. (2014) is used considering both exposure, and sensitivity components. Weights of the different components are assessed using the Analytical Hierarchical Process to evaluate the opinions of eight independent experts in the field of urban water management. The main result of this part is that vulnerability follows the spatial pattern of sensitivity given its larger weight than exposure; therefore, more emphasis needs to be laid on damage sensitivity data collection than to assessing exposure. That applies for both groundwater and soil moisture droughts at coarse and fine spatial resolution. In general, the results of fine and

coarse resolution were in agreement but vulnerability in fine one can provide a deeper insight in the need for local adaptation measures.

In addition, for both studied drought types, using different drought identification techniques, exposure varies considerably but their differences are subtle in vulnerability. That means that there is no need for a sophisticated drought identification technique. A technique which is effortlessly applicable can be used for exposure assessment. However, the following elements ((i) the fact that validation of vulnerability is not performed, and (ii) the large variation of indicators' weights) decrease the credibility of the aforementioned results. Therefore, there is high uncertainty about the results and strong conclusions cannot be drawn.

The implications of this research are large. Estimating sensitivity at fine spatial resolution can lead to timely drought adaptation and an effective preparedness plan on the cities. Hot spots of vulnerability can be pinpointed and consequently adaptation solutions can be to the point; in that way, no resources are wasted since the protection measures would vary spatially based on the local needs.

6.2 Recommendations for further research

What this thesis adds to the current literature are: (i) an urban drought categorization framework, (ii) and estimation of vulnerability to groundwater and soil moisture at urban environment at fine spatial resolution. This research is one of the first steps exploring vulnerability of cities as the majority of literature has focused on rural areas. The topic of urban drought is quite unexplored; therefore there is plenty of room for research.

Closely related to the current research

The main recommendations for further research are highlighted below:

- Applying validation of vulnerability for Leiden is the most significant step of further research. It would provide useful answers whether this technique used in the current study can be applied for vulnerability determination. On top of that, it would clarify which methods to identify groundwater and soil moisture droughts work best for vulnerability determination. However, it is not an simple task. An inventory of damages to houses and infrastructure would be useful. Social damage, which is quite difficult to be quantified, should not be ignored.
- Including spatial extent and frequency distribution analysis for the estimation of exposure could provide a more concrete understanding of the drought events. Using the other techniques (variable threshold and moving window) in spatial analysis of droughts could add more knowledge about spatial extent of droughts. Additionally, the influence of those techniques on vulnerability could be investigated.
- Large variation was noticed applying pooling for inter-event time periods less than 10 days. Pooling could be studied at the range [0-10 days] to provide deeper insight into the link between deficit (or duration) and inter-event time.
- Applying a hybrid method of fixed and variable method to identify droughts is recommended. The benefits of the two approaches are combined. In the suggested approach, for the summer months, instead of applying the monthly percentiles as threshold, the fixed threshold would be applied which is based on the whole time series.
- Research about 'open water urban drought' class from the urban drought categorization framework is another important topic for further research. The following step would be to determine the interrelationship of groundwater, soil moisture, and open water urban droughts (water supply urban drought is excluded since it differs considerably and requires a different way of analysis). Prioritization of these drought categories on cities could be an additional part of that research.
- Minimizing the number of physical sensitivity indicators without losing information in vulnerability estimation could be a suggestion for further research. Removing the redundant indicators

could be useful. The indicators with low weights may be removed too.

- In order to strengthen the component of social vulnerability, including two or three social sensitivity indicators instead of one is suggested for vulnerability estimation. The two extra ones can be: (i) percentage of residents over 60 years old, and (ii) percentage of disabled people. The reasoning is as follows. Relocation may be possible for residents whose houses suffered severe damage (e.g. due to subsidence, problems with utilities connections may occur). However, this would be an agonizing process for groups with mobility limitations (such as the elderly and the disabled). Therefore, these groups are vulnerable to groundwater droughts.

  It is worth noting that more explanation about how the social sensitivity indicators are related to vulnerability to groundwater and soil moisture droughts need to be provided to drought experts. The links may not be that straightforward. In that way, drought experts would be able to better understand why each social indicator was selected.

Generic ones

- While the current research identified that vulnerability of a city from a developed country, there is still lack of information regarding these types of droughts in developing countries. The main difference to determine vulnerability is that different indicators need to be used based on local conditions.
- This analysis is focused on the current status but an analysis for the future is more important. Climate change needs to be considered regarding precipitation and evaporation which are used as inputs in UWBM for soil moisture calculation. If climate change is not considered, vulnerability would be conservative and probably underestimated. KNMI has created four scenarios for climate change which could be used.
- An important sub-product which could be derived by a similar analysis is the water supply needed to avoid groundwater or soil moisture droughts. When extra water is added to the system to prevent groundwater droughts, extra evaporation and transpiration occurs by vegetation. Therefore, the quantity of water needed is larger than the water deficit. This question was not addressed in the current analysis. On top of that, there is no variable regarding irrigation currently available in UWBM. A study focused on that would be helpful given that (i) being aware of the water quantity needed to avoid groundwater droughts would be beneficial to protection planning by water managers, and (ii) estimation of water balance would be more accurate incorporating irrigation.
- Creating a Decision Support Scheme (DSS) regarding the estimation of drought vulnerability is a recommendation. For the short term, it can be passive but there is potential for a DSS that provides the user with an evaluation of the suggested solutions.

Bibliography

Adamo, S. B. (2010). Environmental migration and cities in the context of global environmental change. *Current Opinion in Environmental Sustainability*, 2(3):161–165.

Ahmed, B. (2015). Landslide susceptibility mapping using multi-criteria evaluation techniques in chittagong metropolitan area, bangladesh. *Landslides*, 12(6):1077–1095.

Alcamo, J., Acosta-Michlik, L., Carius, A., Eierdanz, F., Klein, R., Krömker, D., and Tänzler, D. (2008). A new approach to quantifying and comparing vulnerability to drought. *Regional Environmental Change*, 8(4):137.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. (1998). Fao irrigation and drainage paper no. 56. *Rome: Food and Agriculture Organization of the United Nations*, 56(97):e156.

Attrill, M. J. and Power, M. (2000). Effects on invertebrate populations of drought-induced changes in estuarine water quality. *Marine Ecology Progress Series*, 203:133–143.

Bechini, L., Ducco, G., Donatelli, M., and Stein, A. (2000). Modelling, interpolation and stochastic simulation in space and time of global solar radiation. *Agriculture, ecosystems & environment*, 81(1):29–42.

Beersma, J. (2007). *Extreme hydro-meteorological events and their probabilities= Extreme hydrometeorologische gebeurtenissen en de kans daarop*. PhD thesis, Wageningen University.

Beyene, B., Van Loon, A., Van Lanen, H., and Torfs, P. (2014). Investigation of variable threshold level approaches for hydrological drought identification. *Hydrology and Earth System Sciences Discussions*, 11(11):12765–12797.

Birkmann, J., Cardona, O. D., Carreño, M. L., Barbat, A. H., Pelling, M., Schneiderbauer, S., Kienberger, S., Keiler, M., Alexander, D., Zeil, P., et al. (2013). Framing vulnerability, risk and societal responses: the move framework. *Natural hazards*, 67(2):193–211.

Birkmann, J., Welle, T., Solecki, W., Lwasa, S., and Garschagen, M. (2016). Boost resilience of small and mid-sized cities. *Nature News*, 537(7622):605.

Bloomfield, J., Gaus, I., and Wade, S. (2003). A method for investigating the potential impacts of climate-change scenarios on annual minimum groundwater levels. *Water and Environment Journal*, 17(2):86–91.

Bloomfield, J. and Marchant, B. (2013). Analysis of groundwater drought building on the standardised precipitation index approach. *Hydrology and Earth System Sciences*, 17:4769–4787.

Boar, R., Lister, D., and Clough, W. (1995). Phosphorus loads in a small groundwater-fed river during the 1989–1992 east anglian drought. *Water Research*, 29(9):2167–2173.

Bohle, H. G., Downing, T. E., and Watts, M. J. (1994). Climate change and social vulnerability: toward a sociology and geography of food insecurity. *Global environmental change*, 4(1):37–48.

Boogaard, F., Lucke, T., Van de Giesen, N., and Van de Ven, F. (2014). Evaluating the infiltration performance of eight dutch permeable pavements using a new full-scale infiltration testing method. *Water*, 6(7):2070–2083.

Caruso, B. (2002). Temporal and spatial patterns of extreme low flows and effects on stream ecosystems in otago, new zealand. *Journal of Hydrology*, 257(1-4):115–133.

Centraal Bureau voor de Statistiek (CBS) (2015). Wijk- en buurtkaart 2015 [District and neighborhood map 2015]. Derived June, 24, 2020, from: https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/wijk-en-buurtkaart-2015.

Centraal Bureau voor de Statistiek (CBS) (2017). Wijk- en buurtkaart 2017 [District and neighborhood map 2017]. Derived June, 24, 2020, from: https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/wijk-en-buurtkaart-2017.

Chang, L., Yu, D., and Zheng, J. (2015). Assessment of urban vulnerability to drought in northern slope of tianshan mountains. In *2015 23rd International Conference on Geoinformatics*, pages 1–7. IEEE.

Chen, J. and Chan, F. (2007). A preliminary drought mitigation plan for hong kong. *MODSIM07-Land, Water and Environmental Management: Integrated Systems for Sustainability, Proceedings*, page 1494.

Chok, N. S. (2010). *Pearson's versus Spearman's and Kendall's correlation coefficients for continuous data*. PhD thesis, University of Pittsburgh.

Commission, J. R. C.-E. et al. (2008). *Handbook on constructing composite indicators: methodology and user guide*. OECD publishing.

Corti, T., Muccione, V., Köllner-Heck, P., Bresch, D., and Seneviratne, S. I. (2009). Simulating past droughts and associated building damages in france. *Hydrology and Earth System Sciences*, 13(9):1739–1747.

Corti, T., Wüest, M., Bresch, D., and Seneviratne, S. I. (2011). Drought-induced building damages from simulations at regional scale. *Natural Hazards and Earth System Sciences*, 11(12):3335–3342.

Costa, A. L., Kok, S., and Korff, M. (2020). Systematic assessment of damage to buildings due to groundwater lowering-induced subsidence: methodology for large scale application in the netherlands. *Proceedings of the International Association of Hydrological Sciences*, 382:577–582.

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

Dai, M., Guo, X., Zhai, W., Yuan, L., Wang, B., Wang, L., Cai, P., Tang, T., and Cai, W.-J. (2006). Oxygen depletion in the upper reach of the pearl river estuary during a winter drought. *Marine chemistry*, 102(1-2):159–169.

Dalezios, N. R., Loukas, A., Vasiliades, L., and Liakopoulos, E. (2000). Severity-duration-frequency analysis of droughts and wet periods in greece. *Hydrological sciences journal*, 45(5):751–769.

De Sherbinin, A., Schiller, A., and Pulsipher, A. (2007). The vulnerability of global cities to climate hazards. *Environment and urbanization*, 19(1):39–64.

Desbureaux, S. and Rodella, A.-S. (2019). Drought in the city: The economic impact of water scarcity in latin american metropolitan areas. *World Development*, 114:13–27.

DG Environment European Commission and others (2007). Water scarcity and droughts. http://ec.europa.eu/environment/water/quantity/pdf/comm_droughts/2nd_int_report.pdf.

Dinoloket (n.d.). Subsurface models. Derived June, 1, 2020, from: https://www.dinoloket.nl/en/subsurface-models.

Doran, G. T. (1981). There'sa smart way to write management's goals and objectives. *Management review*, 70(11):35–36.

Dryers, P. (2009). *Verbetering bepaling actuele verdamping voor het strategisch waterbeheer : definitiestudie [Improvement determination current evaporation for strategic water management: definition study]*. Stowa.

Duffie, J. A. and Beckman, W. A. (1991). *Solar engineering of thermal processes*. Wiley New York.

Dumitraşcu, M., Mocanu, I., Mitrică, B., Dragotă, C., Grigorescu, I., and Dumitrică, C. (2018). The assessment of socio-economic vulnerability to drought in southern romania (oltenia plain). *International journal of disaster risk reduction*, 27:142–154.

earth.org (2020). 'sponge cities' could be the answer to china's impending water crisis. Derived October, 22, 2020, from: https://earth.org/sponge-cities-could-be-the-answer-to-impending-water-crisis-in-china/.

European Drought Observatory (EDO) (2019). Soil moisture anomaly (sma). Derived July, 26, 2020, from:https://edo.jrc.ec.europa.eu/documents/factsheets/factsheet_soilmoisture.pdf.

Fleig, A. K., Tallaksen, L. M., Hisdal, H., and Demuth, S. (2006). A global evaluation of streamflow drought characteristics.

Fritzsche, K., Schneiderbauer, S., Bubeck, P., Kienberger, S., Buth, M., Zebisch, M., and Kahlenborn, W. (2014). The vulnerability sourcebook: Concept and guidelines for standardised vulnerability assessments.

García-Prieto, J. C., Cachaza, J. M., Pérez-Galende, P., and Roig, M. G. (2012). Impact of drought on the ecological and chemical status of surface water and on the content of arsenic and fluoride pollutants of groundwater in the province of salamanca (western spain). *Chemistry and Ecology*, 28(6):545–560.

Gleick, P. H. (2014). Water, drought, climate change, and conflict in syria. *Weather, Climate, and Society*, 6(3):331–340.

Global Environment Facility (2010). *The GEF Monitoring and Evaluation Policy 2010*. Global Environment Facility, Evaluation Office, Washington, DC.

Goepel, K. D. (2018). Implementation of an online software tool for the analytic hierarchy process (ahp-os). *International Journal of the Analytic Hierarchy Process*, 10(3).

Google (n.d.). Leiden. Derived June, 24, 2020, from: https://www.google.nl/maps.

Granemann, A. R. B., Mine, M. R. M., and Kaviski, E. (2018). Frequency analysis of minimum flows. *RBRH*, 23.

Greening, H. and Gerritsen, J. (1987). Changes in macrophyte community structure following drought in the okefenokee swamp, georgia, usa. *Aquatic Botany*, 28(2):113–128.

Güneralp, B., Güneralp, İ., and Liu, Y. (2015). Changing global patterns of urban exposure to flood and drought hazards. *Global environmental change*, 31:217–225.

Gurwin, J. (2014). Long-term monitoring and gis based determination of groundwater drought propagation, the lower silesia region, sw poland. *Episodes*, 37(3):172–181.

Gustard, A., Bullock, A., and Dixon, J. (1992). *Low flow estimation in the United Kingdom*. Institute of Hydrology.

Ha, K., Cho, E.-A., Kim, H.-W., and Joo, G.-J. (1999). Microcystis bloom formation in the lower nakdong river, south korea: importance of hydrodynamics and nutrient loading. *Marine and Freshwater Research*, 50(1):89–94.

Hargreaves, G. H. and Samani, Z. A. (1982). Estimating potential evapotranspiration. *Journal of the irrigation and Drainage Division*, 108(3):225–230.

Heim Jr, R. R. (2002). A review of twentieth-century drought indices used in the united states. *Bulletin of the American Meteorological Society*, 83(8):1149–1166.

Henderson, J. V., Storeygard, A., and Deichmann, U. (2017). Has climate change driven urbanization in africa? *Journal of development economics*, 124:60–82.

Heudorfer, B. and Stahl, K. (2017). Comparison of different threshold level methods for drought propagation analysis in germany. *Hydrology Research*, 48(5):1311–1326.

Hisdal, H. and Tallaksen, L. (2000). Technical report no. 6 drought event definition. *Assessment of the Regional Impact of Droughts in Europe*, 1(6):41.

Hoogvliet, M., Van de Ven, F., Buma, J., Van Oostrom, N., Brolsma, R., Filatova, T., Verheijen, J., and Bosch, P. (2012). *Schades door watertekorten en overschotten in stedelijk gebied: Quick scan van beschikbaarheid schadegetallen en mogelijkheden om schades te bepalen [Damages due to water shortages and surpluses in urban areas: quick scan of the availability of damage numbers and possibilities to determine damage]*. Deltares. http://publications.deltares.nl/1205463_000a.pdf.

Hoque, M. A.-A., Pradhan, B., and Ahmed, N. (2020). Assessing drought vulnerability using geospatial techniques in northwestern part of bangladesh. *Science of The Total Environment*, 705:135957.

Hunt, E. D., Hubbard, K. G., Wilhite, D. A., Arkebauer, T. J., and Dutcher, A. L. (2009). The development and evaluation of a soil moisture index. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 29(5):747–759.

Iñiguez, V., Morales, O., Cisneros, F., Bauwens, W., and Wyseure, G. (2016). Analysis of the drought recovery of andosols on southern ecuadorian andean páramos. *Hydrology and Earth System Sciences*, 20(6):2421–2435.

Intergovermental Panel on Climate Change (2007). Ar4 climate change 2007: Impacts, adaptation, and vulnerability. https://www.ipcc.ch/site/assets/uploads/2018/03/ar4_wg2_full_report.pdf.

Intergovermental Panel on Climate Change (2014). Annex II: Glossary Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

Intergovermental Panel on Climate Change (2018). Global warming of 1.5°, summary for policymakers. https://wg1.ipcc.ch/presentations/Sbsta_drought.pdf.

Iyengar, N. S. and Sudarshan, P. (1982). A method of classifying regions from multivariate data. *Economic and political weekly*, pages 2047–2052.

Keenan, P. (1995). *Using a GIS as a DSS Generator*. University College Dublin, Department of Management Information Systems.

Keshavarz, M. R., Vazifedoust, M., and Alizadeh, A. (2014). Drought monitoring using a soil wetness deficit index (swdi) derived from modis satellite data. *Agricultural Water Management*, 132:37–45.

Kim, H., Park, J., Yoo, J., and Kim, T.-W. (2015). Assessment of drought hazard, vulnerability, and risk: A case study for administrative districts in south korea. *Journal of Hydro-environment Research*, 9(1):28–35.

Knutson, C., Hayes, M., and Phillips, T. (1998). How to reduce drought risk.

Kovats, R. S., Bouma, M. J., Hajat, S., Worrall, E., and Haines, A. (2003). El niño and health. *The Lancet*, 362(9394):1481–1489.

Kumar, R., Musuuza, J. L., Loon, A. F. V., Teuling, A. J., Barthel, R., Ten Broek, J., Mai, J., Samaniego, L., and Attinger, S. (2016). Multiscale evaluation of the standardized precipitation index as a groundwater drought indicator. *Hydrology and Earth System Sciences*, 20(3):1117–1131.

Larimore, R. W., Childers, W. F., and Heckrotte, C. (1959). Destruction and re-establishment of stream fish and invertebrates affected by drought. *Transactions of the American Fisheries Society*, 88(4):261–285.

Lloyd-Hughes, B. (2014). The impracticality of a universal drought definition. *Theoretical and Applied Climatology*, 117(3-4):607–611.

Loukas, A. and Vasiliades, L. (2004). Probabilistic analysis of drought spatiotemporal characteristics in thessaly region, greece.

Mahajan, D. R. and Dodamani, B. M. (2016). Spatial and temporal drought analysis in the krishna river basin of maharashtra, india. *Cogent Engineering*, 3(1):1185926.

Malczewski, J., Chapman, T., Flegel, C., Walters, D., Shrubsole, D., and Healy, M. A. (2003). Gis–multicriteria evaluation with ordered weighted averaging (owa): case study of developing watershed management strategies. *Environment and Planning A*, 35(10):1769–1784.

Martínez-Fernández, J., González-Zamora, A., Sánchez, N., and Gumuzzio, A. (2015). A soil water based index as a suitable agricultural drought indicator. *Journal of Hydrology*, 522:265–273.

McGowan, S., Leavitt, P. R., and Hall, R. I. (2005). A whole-lake experiment to determine the effects of winter droughts on shallow lakes. *Ecosystems*, 8(6):694–708.

McKetie, T. B., Doesken, N. J., Kleist, J., et al. (1993). The relaonship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology*, volume 17, pages 179–183. Boston.

Mishra, A. and Desai, V. (2005). Spatial and temporal drought analysis in the kansabati river basin, india. *International Journal of River Basin Management*, 3(1):31–41.

Mishra, A. K. and Singh, V. P. (2010). A review of drought concepts. *Journal of hydrology*, 391(1-2):202–216.

Mishra, A. K. and Singh, V. P. (2011). Drought modeling–a review. *Journal of Hydrology*, 403(1-2):157–175.

Mosley, L. M., Zammit, B., Leyden, E., Heneker, T. M., Hipsey, M. R., Skinner, D., and Aldridge, K. T. (2012). The impact of extreme low flows on the water quality of the lower murray river and lakes (south australia). *Water Resources Management*, 26(13):3923–3946.

Murthy, C., Singh, J., Kumar, P., and Sai, M. S. (2017). A composite index for drought hazard assessment using cpc rainfall time series data. *International Journal of Environmental Science and Technology*, 14(9):1981–1988.

Narasimhan, B. and Srinivasan, R. (2005). Development and evaluation of soil moisture deficit index (smdi) and evapotranspiration deficit index (etdi) for agricultural drought monitoring. *Agricultural and forest meteorology*, 133(1-4):69–88.

Nasrollahi, M., Khosravi, H., Moghaddamnia, A., Malekian, A., and Shahid, S. (2018). Assessment of drought risk index using drought hazard and vulnerability indices. *Arabian Journal of Geosciences*, 11(20):606.

National Institute for Public Health and the Environment (2019a). Klimaatschadeschatter report 2019, klimaatbestedndige stad [climate damage estimation report 2019, climate-proof city]. Derived July, 30, 2020, from:https://ruimtelijkeadaptatie.nl/publish/pages/178373/klimaatschadeschatter_rapportage_2019.pdf.

National Institute for Public Health and the Environment (2019b). Schade aan wegen en riolering door droogte, technische beschrijving en stappenplan [damage to roads and sewage by droughts, technical description and step-by-step plan]. Derived July, 30, 2020, from: https://cas.klantsite.net/l/nl/library/download/urn:uuid:ddbac95b-85a0-4622-a465-98c9b3e0bea6/factsheet+-+wegen+droogte.pdf?format=save_to_disk&ext=.pdf.

Netherlands Center for Geodesy and Geo-Informatics (NCG) (2018). Soil subsidence map. Derived June, 22, 2020, from: https://bodemdalingskaart.nl/kennis-datacentrum/download-bodembewegingsdata/.

Nikam, B. R., Kumar, P., Garg, V., Thakur, P. K., and Aggarwal, S. (2014). Comparative evaluation of different potential evapotranspiration estimation approaches. *International Journal of Research in Engineering and Technology*, 3(6):544–552.

NOAA, National Centers Environmental Information (NCEI) (2019). U.S. Billion-Dollar Weather and Climate Disasters. https://www.ncdc.noaa.gov/billions/.

Panahi, M., Rezaie, F., and Meshkani, S. (2014). Seismic vulnerability assessment of school buildings in tehran city based on ahp and gis. *Natural Hazards and Earth System Sciences*, 14(4):969.

Pandey, S., Behura, D. D., Villano, R., and Naik, D. (2000). Economic cost of drought and farmers' coping mechanisms: a study of rainfed rice systems in eastern india. Technical report.

Peters, E. (2003). *Propagation of drought through groundwater systems*. PhD thesis, Wageningen University.

Polsky, C., Neff, R., and Yarnal, B. (2007). Building comparable global change vulnerability assessments: The vulnerability scoping diagram. *Global Environmental Change*, 17(3-4):472–485.

Power, D. J. (2002). *Decision support systems: concepts and resources for managers*. Greenwood Publishing Group.

Publieke Dienstverlening Op de Kaart (PDOK) (2015). Dataset: CBS Bestand Bodemgebruik [Dataset: Statistics Netherlands File Land Use]. Derived June, 22, 2020, from: https://www.pdok.nl/downloads/-/article/cbs-bestand-bodemgebruik.

Rahman, M. R. and Lateh, H. (2016). Meteorological drought in bangladesh: assessing, analysing and hazard mapping using spi, gis and monthly rainfall data. *Environmental Earth Sciences*, 75(12):1026.

Rossi, G., Benedini, M., Tsakiris, G., and Giakoumakis, S. (1992). On regional drought estimation and analysis. *Water resources management*, 6(4):249–277.

Royal Netherlands Meteorological Institute (KNMI) (2020). Neerslagtekort/droogte [precipitation deficit/drought]. Derived July, 24, 2020, from: https://www.knmi.nl/nederland-nu/klimatologie/geografische-overzichten/neerslagtekort_droogte.

Satty, T. L. (1980). *The Analytic Hierarchy Process*. Mcgraw-Hill, New York.

Schmidt, D. H. and Garland, K. A. (2012). Bone dry in texas: resilience to drought on the upper texas gulf coast. *Journal of Planning Literature*, 27(4):434–445.

Schreurs, E. (2017). Deterioration of timber pile foundations in rotterdam. Master's thesis, TU Delft.

Schröter, D., Polsky, C., and Patt, A. G. (2005). Assessing vulnerabilities to the effects of global change: an eight step approach. *Mitigation and Adaptation Strategies for Global Change*, 10(4):573–595.

Shahid, S. and Behrawan, H. (2008). Drought risk assessment in the western part of bangladesh. *Natural hazards*, 46(3):391–413.

Shahid, S. and Hazarika, M. K. (2010). Groundwater drought in the northwestern districts of bangladesh. *Water resources management*, 24(10):1989–2006.

Sheffield, J. and Wood, E. F. (2012). *Drought: past problems and future scenarios*. Routledge.

Shiau, J. (2006). Fitting drought duration and severity with two-dimensional copulas. *Water resources management*, 20(5):795–815.

Sprague, L. A. (2005). Drought effects on water quality in the south platte river basin, colorado 1. *JAWRA Journal of the American Water Resources Association*, 41(1):11–24.

Sprague Jr, R. H. and Carlson, E. D. (1982). *Building effective decision support systems*. Prentice Hall Professional Technical Reference.

Stahl, K. (2001). *Hydrological drought: A study across Europe*. PhD thesis, Institut für Hydrologie der Universität.

Swearingen, W. D. (1992). Drought hazard in morocco. *Geographical Review*, pages 401–412.

Tallaksen, L. M., Hisdal, H., and Van Lanen, H. A. (2009). Space–time modelling of catchment scale drought characteristics. *Journal of Hydrology*, 375(3-4):363–372.

Tallaksen, L. M., Madsen, H., and Clausen, B. (1997). On the definition and modelling of streamflow drought duration and deficit volume. *Hydrological Sciences Journal*, 42(1):15–33.

Thomas, T., Jaiswal, R., Galkate, R., Nayak, P., and Ghosh, N. (2016). Drought indicators-based integrated assessment of drought vulnerability: a case study of bundelkhand droughts in central india. *Natural Hazards*, 81(3):1627–1652.

Turner, B. L., Kasperson, R. E., Matson, P. A., McCarthy, J. J., Corell, R. W., Christensen, L., Eckley, N., Kasperson, J. X., Luers, A., Martello, M. L., et al. (2003). A framework for vulnerability analysis in sustainability science. *Proceedings of the national academy of sciences*, 100(14):8074–8079.

United Nations (2014). World urbanization prospects: The 2014 revision, highlights. department of economic and social affairs. *Population Division, United Nations*, 32.

Van Loon, A., Tijdeman, E., Wanders, N., Van Lanen, H. J., Teuling, A., and Uijlenhoet, R. (2014). How climate seasonality modifies drought duration and deficit. *Journal of Geophysical Research: Atmospheres*, 119(8):4640–4656.

van Loon, A. F. (2013). On the propagation of drought: how climate and catchment characteristics influence hydrological drought development and recovery.

Van Loon, A. F. (2015). Hydrological drought explained. *Wiley Interdisciplinary Reviews: Water*, 2(4):359–392.

Van Loon, A. F. and Van Lanen, H. A. (2012). A process-based typology of hydrological drought. *Hydrology and Earth System Sciences*, 16(7):1915.

Van Vliet, M. and Zwolsman, J. (2008). Impact of summer droughts on the water quality of the meuse river. *Journal of Hydrology*, 353(1-2):1–17.

Ven, F., Nieuwkerk, E. v., Stone, K., Veerbeek, W., Rijke, J., Herk, S. v., and Zevenbergen, C. (2011). *Building the Netherlands climate proof: urban areas*. Utrecht: Programme office Knowledge for Climate.

Vergroesen, T. (2020). Urban water balance model. Derived June, 22, 2020, from: https://publicwiki.deltares.nl/display/AST/Urban+Water+balance+model.

Wang, P., Qiao, W., Wang, Y., Cao, S., and Zhang, Y. (2019). Urban drought vulnerability assessment-a framework to integrate socio-economic, physical, and policy index in a vulnerability contribution analysis. *Sustainable Cities and Society*, page 102004.

Wiesner, S., Gröngröft, A., Ament, F., and Eschenbach, A. (2016). Spatial and temporal variability of urban soil water dynamics observed by a soil monitoring network. *Journal of Soils and Sediments*, 16(11):2523–2537.

Wilbers, G.-J., Zwolsman, G., Klaver, G., and Hendriks, A. J. (2009). Effects of a drought period on physico-chemical surface water quality in a regional catchment area. *Journal of Environmental Monitoring*, 11(6):1298–1302.

Wilhite, D. A. (1994). *Preparing for drought: A guidebook for developing countries*. Diane Publishing.

Wilhite, D. A. (2000). Drought as a natural hazard: concepts and definitions.

Wilhite, D. A. and Glantz, M. H. (1985). Understanding: the drought phenomenon: the role of definitions. *Water international*, 10(3):111–120.

Wilhite, D. A. and Vanyarkho, O. V. (2000). Drought: Pervasive impacts of a creeping phenomenon.

Wolski, P. (2018). How severe is cape town's "day zero" drought? *Significance*, 15(2):24–27.

Wright, J., Gunn, R., Winder, J., Wiggers, R., Vowles, K., Clarke, R., and Harris, I. (2002). A comparison of the macrophyte cover and macroinvertebrate fauna at three sites on the river kennet in the mid 1970s and late 1990s. *Science of the Total Environment*, 282:121–142.

Wu, H., Qian, H., Chen, J., and Huo, C. (2017). Assessment of agricultural drought vulnerability in the guanzhong plain, china. *Water resources management*, 31(5):1557–1574.

Yeh, P. J. and Famiglietti, J. (2009). Regional groundwater evapotranspiration in illinois. *Journal of Hydrometeorology*, 10(2):464–478.

Yevjevich, V. M. (1967). Objective approach to definitions and investigations of continental hydrologic droughts, an. *Hydrology papers (Colorado State University); no. 23*.

Yu, H., Song, Y., Chang, X., Gao, H., and Peng, J. (2018). A scheme for a sustainable urban water environmental system during the urbanization process in china. *Engineering*, 4(2):190–193.

Yuan, X.-C., Wang, Q., Wang, K., Wang, B., Jin, J.-L., and Wei, Y.-M. (2015). China's regional vulnerability to drought and its mitigation strategies under climate change: data envelopment analysis and analytic hierarchy process integrated approach. *Mitigation and Adaptation Strategies for Global Change*, 20(3):341–359.

Zelenhasić, E. and Salvai, A. (1987). A method of streamflow drought analysis. *Water resources research*, 23(1):156–168.

Zhang, Q., Sun, P., Li, J., Xiao, M., and Singh, V. P. (2015). Assessment of drought vulnerability of the tarim river basin, xinjiang, china. *Theoretical and applied climatology*, 121(1-2):337–347.

Zhang, X., Chen, N., Sheng, H., Ip, C., Yang, L., Chen, Y., Sang, Z., Tadesse, T., Lim, T. P. Y., Rajabifard, A., et al. (2019). Urban drought challenge to 2030 sustainable development goals. *Science of the Total Environment*, 693(13):133536.

Zwolsman, J. and Van Bokhoven, A. (2007). Impact of summer droughts on water quality of the rhine river-a preview of climate change? *Water Science and Technology*, 56(4):45–55.

Appendices

Appendix A

Flow chart for Methodology

Figure A.1 presents the flow chart of the steps followed in this thesis; vulnerability to groundwater and soil moisture droughts are considered.

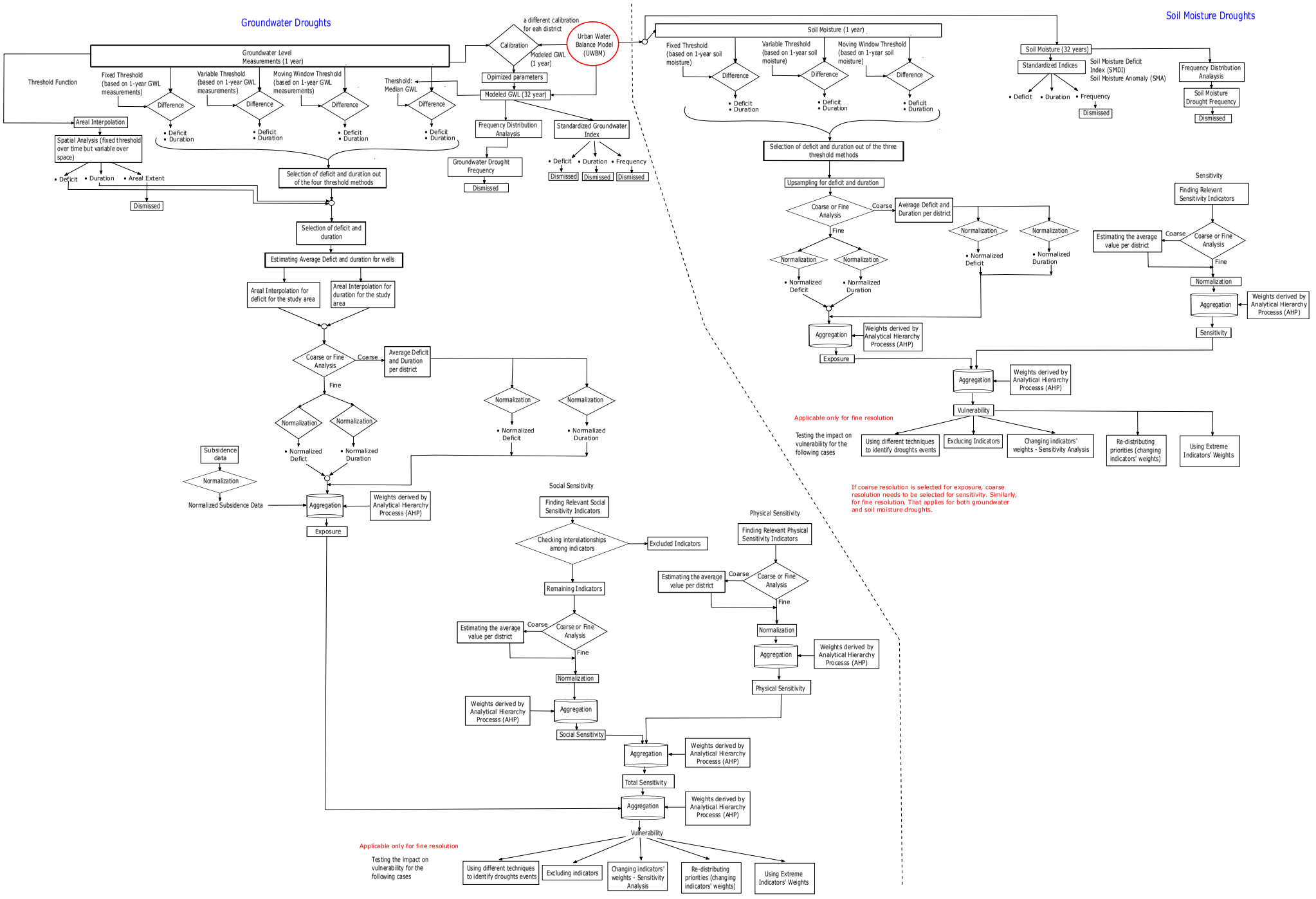


Figure A.1: Flow chart for Methodology

Appendix B

Calculations

B.1 Estimation of Penman Evaporation for Leiden

In order to determine potential evaporation which was one of the inputs to the UWBM, Penman formula was employed for the case of Leiden. The four required measured parameters were: mean wind speed (m/s), mean temperature (^{o}C), incoming radiation (J m^{-2}), and relative humidity (%). Measurements were derived by the KNMI weather station Advances (serial number: 215) which is close to Leiden (south of it) and the time scale was hourly. KNMI weather station Valkenburg (serial number: 210) is closer but it was not used since it lacked data; there were no measurements after May of 2016. There were neither missing nor uncommon values which would require filling gaps and removal respectively. Thus, no technique of filling gaps was applied to the data.

The below mentioned formulas are based on Allen et al. (1998). Penman Formula (B.1) was used for the estimation of open water evaporation and is quite accurate since it is physically-based. The unit of the output is in m/s but other relevant units are used as well.

$$E_p = \frac{\frac{s(R_n - G)}{\rho\lambda} + \frac{c_p \rho_a}{\rho\lambda}\frac{(e_s - e_a)}{r_a}}{s + \gamma} \tag{B.1}$$

where:

E_p: Penman potential evaporation (m/s)
R_n: net radiation on the earth's surface (J s^{-1} m^{-2})
G: ground heat flux (J s^{-1} m^{-2})
λ: latent heat of vaporization (2.45 MJ/kg)
s: slope of saturation vapor pressure-temperature curve (kPa/^{o}C)
c_p: specific heat of air (1004 J kg^{-1}) K^{-1}
ρ_α: density of air (1.205 kg/m^3)
ρ: density of water (1000 kg/m^3)
e_a: actual vapor pressure in the air at 2m height (kPa)
e_s: saturation vapor pressure in the air at 2m height (kPa)
γ: psychrometric constant (0.066 kPa/^{o}C)
r_a: aerodynamic resistance (s/m)

Before applying Formula (B.1), the following calculations preceded. Aerodynamic resistance was estimated based on Formula (B.2) and (B.3), where f(u) is an intermediate parameter (J/m^2sPa), and u is wind velocity (m/s). The constants ρ_α, c_p, and γ were defined above. In Formula (B.3), the unit of γ was Pa and not KPa.

$$f(u) = 7.4 \times 10^{-2}(0.5 + 0.54u) \tag{B.2}$$

$$r_a = \frac{\rho_\alpha c_\mathrm{p}}{\gamma f(u)} \tag{B.3}$$

Subsequently, saturation vapor pressure (e_s) was estimated based on Formula (B.4), where T is temperature (°C). Combining it with relative humidity (RH in %), the actual vapor pressure was estimated (Formula B.5). Then, the slope of 'saturation vapor pressure' - 'temperature' curve was calculated which showed how quickly the dew point (where condensation occurs) could be reached considering the current air temperature. The temperature in Formula (B.6) was in °C.

$$e_s(T) = 0.61 \exp\left(\frac{19,9T}{273+T}\right) \tag{B.4}$$

$$e_a(T) = \frac{RH}{e_s} \tag{B.5}$$

$$s = \frac{5430 e_s}{(273+T)^2} \tag{B.6}$$

One of the main parameters which determine evaporation is net radiation which can be calculated by Formula (B.7) where, $R_{s,in}$ is shortwave incoming radiation, $R_{s,out}$ is shortwave reflected radiation from the Earth's surface to the atmosphere, $R_{l,in}$ is the longwave incoming radiation which is scattered from the atmosphere (mainly clouds, and greenhouse gases) back to the Earth, and $R_{l,out}$ is the longwave radiation which is not scattered back but is released into the atmosphere. Albedo (r) determines the ratio of $R_{s,out}$ to $R_{s,in}$ and depends on brightness of the land use; it is dimensionless and varies from 0 (black surface) to 1 (white surface). However, Formula (B.7) was not applied since net longwave radiation ($R_{l,n}$) was not estimated. For the calculation of $R_{l,n}$, the percentage of sunshine per hour is required but there was lack of information regarding it. Consequently, an assumption was made that $R_n = 0.65 R_{s,in}$

$$R_n = (R_{s,in} - R_{s,out}) - (R_{l,out} - R_{l,in}) = R_{s,in} - rR_{s,in} - R_{l,n} = (1-r)R_{s,in} - R_{l,n} \tag{B.7}$$

Another assumption was applied too. Considering that cities are built by materials which absorb more radiation compared to rural environment, it was assumed that $G = 0.1 R_n$.

B.2 Filling gaps using Sequential Algorithm

Neural networks have gained momentum recently due to advance of computer performance and one of their capabilities is to fill in data gaps. Well L-PB08 missed GWL data from the 24th of August 2018 to the 3rd of October 2018 (41 days). Data from wells L-PB09 and L-PB12 were used to fill in the gaps of well L-PB08. Those wells were selected since they were close to well L-PB08 and they did not have large data gaps. They were some minor gaps (no more than 10 days) in wells L-PB09 and L-PB12 which were filled in using linear interpolation. The study period was from 01-04-2018 to 31-03-2019; hence, only data at that period was used for the three wells. Pearson correlation of wells L-PB09 and L-PB12 with well L-PB08 was high (0.858 and 0.941 respectively) (see Table B.1) which is an indication that GWL of L-PB09 and L-PB12 wells could be potential predictors of GWL in L-PB08 well. Pearson coefficient ranges from -1 to 1 (-1 shows a perfect negative relationship whereas 1 show a perfect positive one). Correlation coefficients were derived based only on the training period (from 01-04-2018 to 23-08-2018 and from 04-10-2018 to 31-03-2019).

Table B.1: Pearson correlation coefficients among well L-PB08 (which lacks data) and wells L-PB09, L-PB12 which do not have data gaps

	L-PB08	L-PB09	L-PB12
L-PB08	1.000	0.858	0.941
L-PB09	0.858	1.000	0.952
L-PB12	0.941	0.952	1.000

A sequential model was employed to predict the missing values which is derived from biology (it is used to predict the conformation changes of enzymes). The model was executed using Python library Keras and its predictors were the GWL from wells L-PB09 and L-PB12, whereas GWL from well L-PB08 was the dependent variable. Predictor variables were scaled for both training and test period before applying the model. Some configurations of the model are: (i) Adam optimization was employed which is a first-order gradient-based optimization algorithm; (ii) mean square error was selected as loss function which Adam algorithm minimized. In addition, 200 epochs were used in that model; epoch shows how many times the algorithm passes through the whole training data to train the model so that the latter one can understand the data pattern. The period from 01-04-2018 to 23-08-2018 and 04-10-2018 to 31-03-2018 was used for training whereas the remaining part with the missing data (from 24-08-2018 to 03-10-2018) was predicted.

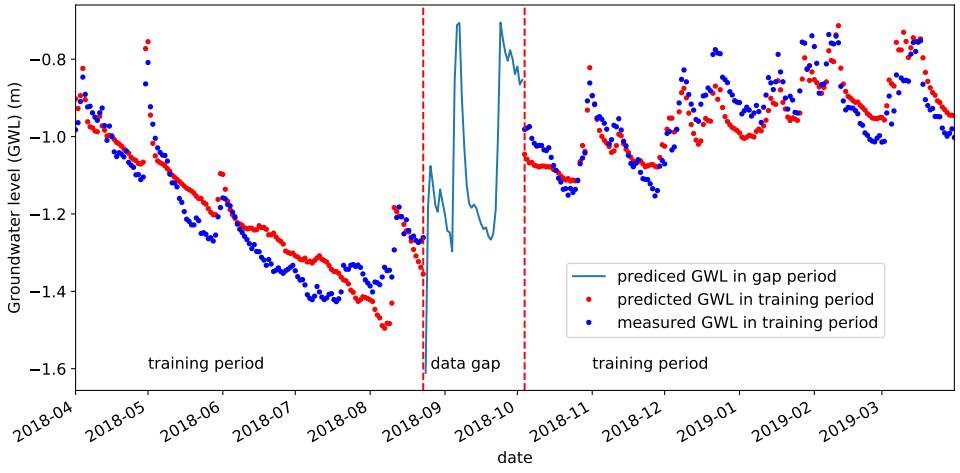


Figure B.1: Comparison of measured and modeled GWL using Sequential algorithm, and prediction of data gaps for well L-PB08

Validation was not conducted since the training period was quite short; including validation period

would decrease training period even further. However, validation is of great importance since it shows the performance of the model and if a model can be trusted or not. Lack of data weakens data driven approaches for filling gaps which was verified by this example.

Even though the mean square error in the training period is quite low (0.03) indicating that the fit of the model was good, the visual inspection was not in line with that. The model missed the valleys before 24-08-2018 and there was a delay as well. On the contrary, the fit was better after 04-10-2018, especially for the high peaks but they were not useful for drought analysis. Regarding the predictions, the results were not satisfactory since the model predicted two peaks of GWL which were impossible to have occurred. For that reason, sequential model was not selected as a method to fill in the data gaps of GWL.

B.3 Results for Groundwater Droughts

District w_1

Table B.2: Statistics for the wells of w_1 district using fixed, variable, and moving window threshold for groundwater drought identification (n: number of identified drought events)

		Fixed Threshold					Variable Threshold					Moving Window Threshold				
			duration (days)		deficit (m)			duration (days)		deficit (m)			duration (days)		deficit (m)	
well	percentile	n	mean	std deviation	mean	std deviation	n	mean	std deviation	mean	std deviation	n	mean	std deviation	mean	std deviation
L-PB49	20	2	36.0	48.1	0.118	0.134	5	15.2	10.3	0.038	0.023	9	3.2	1.3	0.018	0.014
	30	5	21.6	29.7	0.081	0.098	9	10.8	11.0	0.033	0.031	13	4.1	2.6	0.018	0.014
	40	2	73.0	96.2	0.200	0.188	14	10.3	21.9	0.031	0.033	18	4.5	3.3	0.019	0.014
L-PB50	20	5	14.4	13.0	0.065	0.060	17	4.8	3.2	0.019	0.015	13	3.5	1.3	0.017	0.013
	30	5	21.6	22.6	0.085	0.073	18	7.1	4.7	0.031	0.020	15	4.5	2.3	0.025	0.011
	40	6	24.3	30.2	0.080	0.092	18	9.1	5.8	0.040	0.024	21	5.2	3.3	0.026	0.016
L-PB54	20	3	24.3	35.3	0.042	0.064	9	9.2	6.1	0.026	0.012	11	4.5	1.6	0.015	0.010
	30	3	36.3	43.0	0.080	0.069	9	12.9	7.9	0.039	0.020	15	5.2	2.7	0.020	0.011
	40	6	24.3	33.3	0.071	0.067	11	15.2	11.5	0.048	0.027	17	6.9	3.8	0.029	0.017
L-PB55	20	2	36.5	40.3	0.100	0.111	9	10.2	3.7	0.027	0.009	8	5.9	0.4	0.018	0.010
	30	3	36.7	53.1	0.082	0.118	10	13.7	6.8	0.039	0.012	9	7.8	2.4	0.025	0.016
	40	7	20.9	36.1	0.056	0.086	11	16.0	14.8	0.045	0.022	10	9.8	3.9	0.030	0.019
L-PB56	20	3	24.3	34.4	0.068	0.092	7	11.7	5.0	0.026	0.014	10	3.6	1.4	0.012	0.007
	30	3	36.7	38.4	0.100	0.092	10	12.0	9.6	0.033	0.028	12	4.6	2.2	0.016	0.009
	40	5	29.2	38.2	0.091	0.103	6	25.3	29.1	0.055	0.036	12	7.1	3.5	0.022	0.013

District w_5

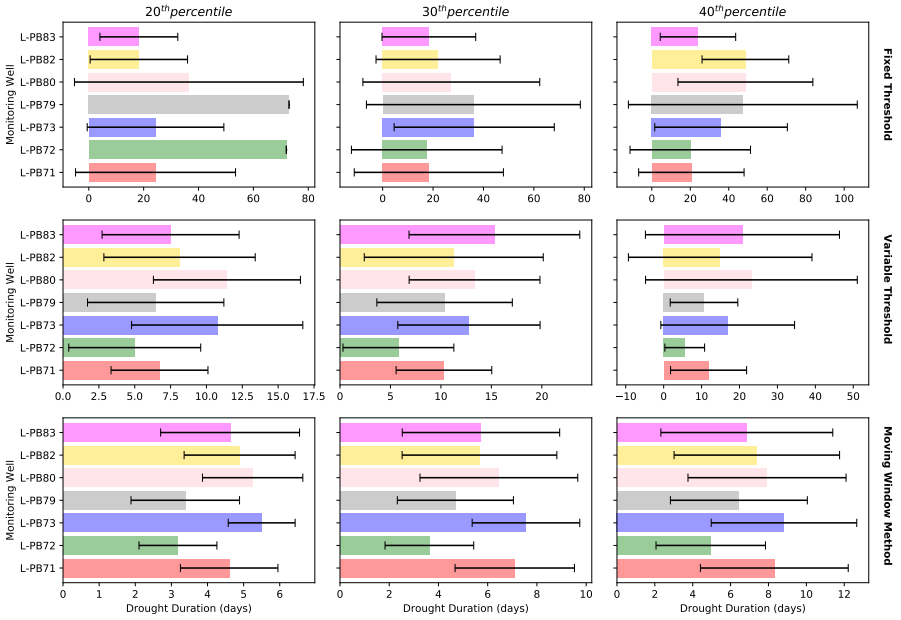


Figure B.2: Drought duration for monitoring wells in district w_5 using fixed, variable, and moving window threshold for three different percentiles

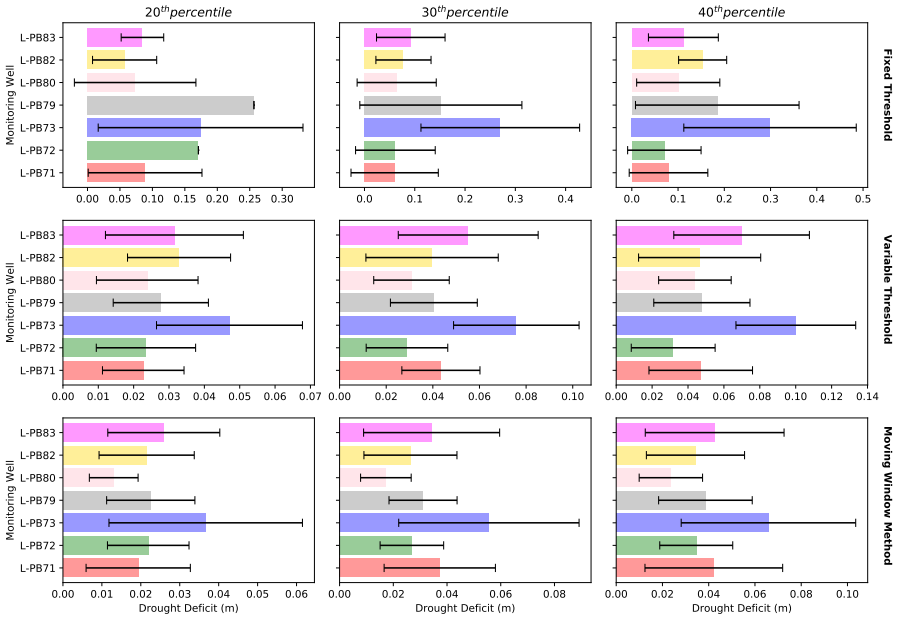


Figure B.3: Drought deficit for monitoring wells in district w_5 using fixed, variable, and moving window threshold for three different percentiles

District w_7

Table B.3: Statistics for the wells of w_7 district using fixed, variable, and moving window threshold for groundwater drought identification (n: number of identified drought events)

		Fixed Threshold					Variable Threshold					Moving Threshold				
			duration (days)		deficit (m)			duration (days)		deficit (m)			duration (days)		deficit (m)	
well	percentile	n	mean	std dev	mean	std dev	n	mean	std dev	mean	std dev	n	mean	std dev	mean	std dev
L-PB08	20	3	24.3	33.5	0.069	0.089	9	6.7	5.1	0.027	0.017	14	3.2	1.7	0.020	0.012
	30	2	55.0	53.7	0.170	0.111	12	7.7	7.0	0.041	0.024	14	5.1	2.7	0.032	0.020
	40	3	48.0	69.3	0.141	0.158	18	7.2	6.7	0.039	0.031	19	6.1	3.3	0.036	0.027
L-PB09	20	3	24.3	27.4	0.087	0.086	12	6.9	4.7	0.030	0.013	11	3.5	1.7	0.022	0.010
	30	4	27.3	32.3	0.101	0.107	13	9.9	7.9	0.044	0.021	11	6.2	2.5	0.032	0.016
	40	6	24.2	32.5	0.096	0.108	13	13.2	10.3	0.058	0.027	15	6.9	3.9	0.038	0.017
L-PB12	20	1	73.0	-	0.215	-	10	7.5	4.6	0.020	0.013	12	3.7	1.9	0.012	0.007
	30	5	22.0	33.0	0.078	0.108	12	9.7	5.0	0.031	0.018	13	6.1	2.8	0.026	0.014
	40	5	29.2	36.7	0.103	0.119	14	11.9	16.8	0.034	0.029	16	7.1	3.8	0.032	0.021
L-PB13	20	4	18.0	22.8	0.048	0.042	12	7.0	4.2	0.024	0.011	13	3.5	1.3	0.018	0.010
	30	7	15.7	26.1	0.042	0.050	15	7.7	4.6	0.034	0.018	14	5.6	2.6	0.025	0.015
	40	5	28.6	34.9	0.085	0.067	19	9.0	8.5	0.036	0.027	16	7.1	3.4	0.032	0.017
L-PB14	20	4	18.3	26.7	0.075	0.112	9	8.1	5.8	0.035	0.023	7	4.3	1.4	0.026	0.012
	30	3	36.3	38.5	0.141	0.154	11	11.0	8.6	0.052	0.034	12	4.8	2.7	0.028	0.019
	40	3	46.7	41.7	0.179	0.154	10	16.9	26.7	0.065	0.043	12	6.9	4.1	0.040	0.022

Relationship between deficit and duration

Figure B.4 depicts the relationship between drought duration and deficit of well L-PB42 from w_0 district. As can be seen, there are outliers in the sample (mostly for fixed threshold). The largest outlier was the drought in summer of 2018. The number of outliers was lower for moving window method compared to the other two approaches. The requirement of a large sample size (larger than 25) to apply Pearson coefficient was not valid in most cases; thus, the probability of error was high. For that reason, statistics of fixed method were not included in Table B.4.

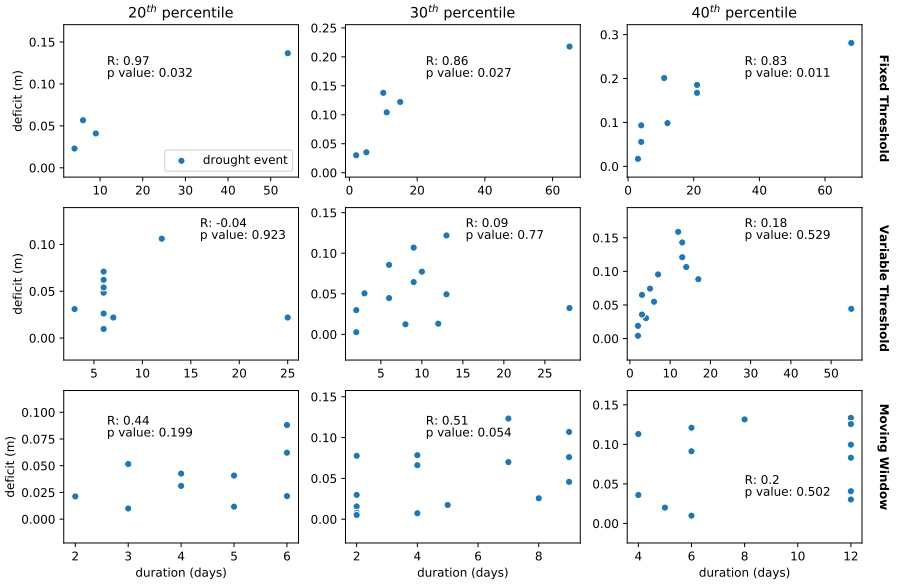


Figure B.4: Scatterplots of groundwater drought duration and deficit using fixed, variable, and moving window threshold for three different percentiles for monitoring well L-PB42 (Pearson coefficient and p value are attached for each case)

Table B.4: Pearson coefficients for the relationship between groundwater drought duration and deficit with their corresponding p-values for wells located in w_0, and w_1 districts (p-values lower than 0.05 are highlighted with pink color). The drought events were identified via variable threshold and moving window.

Method	Variable Method						Moving Window					
percentile	20		30		40		20		30		40	
W_0 District	**R**	**p value**	**R**	**p value**	**R**	**p value**	**R**	**p value**	**R**	**p value**	**R**	**p value**
L-PB42	-0.04	0.923	0.09	0.770	0.18	0.529	0.44	0.199	0.51	0.054	0.20	0.502
L-PB44	0.52	0.189	0.40	0.257	0.31	0.421	0.40	0.329	0.49	0.182	0.71	0.014
L-PB46	0.71	0.003	0.55	0.066	0.67	0.009	0.47	0.148	0.69	0.007	0.58	0.011
	20		30		40		20		30		40	
W_1 District	**R**	**p value**	**R**	**p value**	**R**	**p value**	**R**	**p value**	**R**	**p value**	**R**	**p value**
L-PB49	0.95	0.011	0.91	0.001	0.68	0.007	-0.01	0.713	0.36	0.226	0.43	0.078
L-PB50	0.60	0.011	0.52	0.028	0.49	0.040	0.45	0.126	0.18	0.513	0.41	0.066
L-PB54	0.28	0.463	0.24	0.534	0.24	0.486	0.80	0.003	0.57	0.028	0.55	0.022
L-PB55	-0.29	0.450	-0.02	0.966	0.44	0.172	0.63	0.093	0.76	0.018	0.75	0.012
L-PB56	0.84	0.017	0.86	0.002	0.65	0.161	0.43	0.210	0.58	0.050	0.52	0.081

In district w_0, there was a statistical significant relationship for two wells (L-PB44 and L-PB46) out of the three ones. L-PB46 well showed a relationship in four different cases (variable threshold in conjunction with 20^{th} and 40^{th} percentile ($R = 0.71$ and 0.67 respectively), and moving window in conjunction with 30^{th} and 40^{th} percentile ($R = 0.69$ and 0.58 accordingly). On the contrary, L-PB44 showed a significant relationship only in one case; it was moving window in conjunction with 40^{th} percentile ($R = 0.71$).

In w_1 district, it is surprising that wells L-PB49, L-PB50,and L-PB56 had a statistically significant relationship using variable threshold but not using the moving window. On the other hand, wells L-PB54 and L-PB55 showed a positive relationship between deficit and duration using moving window but not employing variable threshold. The two wells with the highest Pearson coefficients were: L-PB49 ($R = 0.95$ and $R = 0.91$ for 20^{th} and 30^{th} percentile respectively using variable threshold), and L-PB56 ($R = 0.84$ and $R = 0.86$ for 20^{th} and 30^{th} percentile respectively using variable threshold).

Regarding w_7 district, more statistical significant relationships were identified for variable threshold and not moving window (see Table B.5). The other way around applied for w_5 district. Last but not least, L-PB12 and L-PB13 had a statistically significant relationship using variable threshold for all

percentiles but none for moving window; a similar result was mentioned above for the wells L-PB49, L-PB50,and L-PB56 of w_1 district.

Table B.5: Pearson coefficients for the relationship between groundwater drought duration and deficit with their corresponding p-values for wells located in w_5, and w_7 districts (p-values lower than 0.05 are highlighted with pink color). The drought events were identified via variable threshold and moving window.

Method	Variable Method						Moving Window					
percentile	20		30		40		20		30		40	
W_5 District	**R**	**p value**	**R**	**p value**	**R**	**p value**	**R**	**p value**	**R**	**p value**	**R**	**p value**
L-PB71	0.26	0.439	-0.04	0.903	0.37	0.213	0.53	0.114	-0.08	0.833	0.14	0.656
L-PB72	0.80	0.001	0.69	0.001	0.71	0.000	0.70	0.016	0.52	0.021	0.41	0.061
L-PB73	0.24	0.560	0.24	0.542	0.26	0.496	-0.07	0.872	0.11	0.779	0.44	0.176
L-PB79	0.61	0.027	0.58	0.059	0.66	0.010	0.61	0.027	0.39	0.133	0.75	0.001
L-PB80	0.37	0.414	0.43	0.246	0.08	0.870	0.07	0.868	0.63	0.036	0.73	0.005
L-PB82	0.57	0.140	0.83	0.001	0.67	0.230	0.24	0.535	0.58	0.049	0.66	0.015
L-PB83	0.55	0.066	0.04	0.916	0.25	0.516	0.46	0.248	0.58	0.061	0.64	0.014
	20		30		40		20		30		40	
W_7 District	**R**	**p value**	**R**	**p value**	**R**	**p value**	**R**	**p value**	**R**	**p value**	**R**	**p value**
L-PB08	0.70	0.034	0.52	0.082	0.68	0.002	0.48	0.083	0.64	0.014	0.79	0.000
L-PB09	0.42	0.176	0.57	0.043	0.65	0.016	0.66	0.026	0.56	0.071	0.62	0.014
L-PB12	0.70	0.023	0.76	0.004	0.83	0.000	0.42	0.176	0.11	0.725	0.38	0.142
L-PB13	0.71	0.009	0.73	0.002	0.6	0.007	0.47	0.104	0.47	0.092	0.490	0.056
L-PB14	0.59	0.093	0.46	0.151	0.41	0.234	0.76	0.050	0.67	0.018	0.55	0.065

Key results

- No clear connection between drought identification technique and deficit-duration relationship. Some wells showed a positive relationship between deficit and duration only using variable threshold; others using both techniques (variable and moving window threshold).

Pooling

The differences in histograms for groundwater drought deficit of well L-PB50 from w_1 district applying 5-day pooling or not are highlighted in Figure B.5. The peak of the histogram was shifted slightly to the right for the 20^{th}, and 30^{th} percentile in variable threshold method when pooling was applied. The number of drought events with a low deficit decreased since minor droughts merged to larger ones. It is worth noting the scale on the y axis is different for each threshold.

Additionally, reduction in the amplitude of the peak was noticed in variable threshold for all percentiles. For other wells such as L-PB44 (see Figure B.6), the histograms of deficit for applying 5-day pooling or not were the same. The reason was that using five days as inter-event time, no drought event was merged.

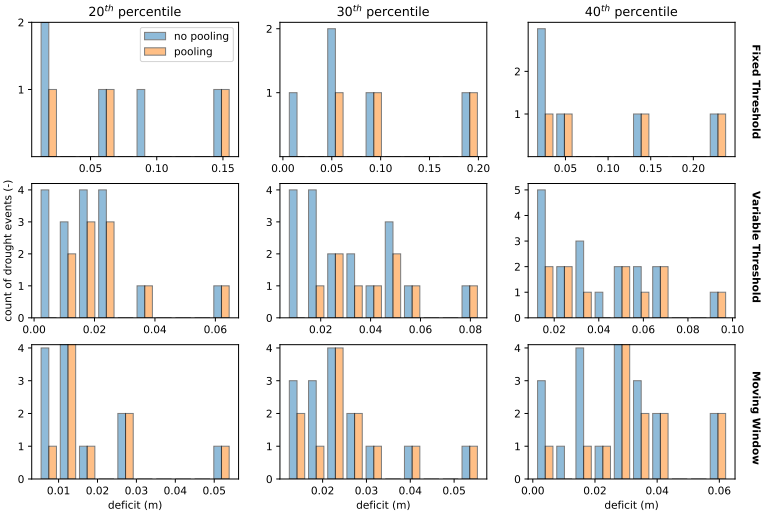


Figure B.5: Groundwater drought deficit histograms of well L-PB50 applying 5-day pooling or not, for three percentiles and three thresholds

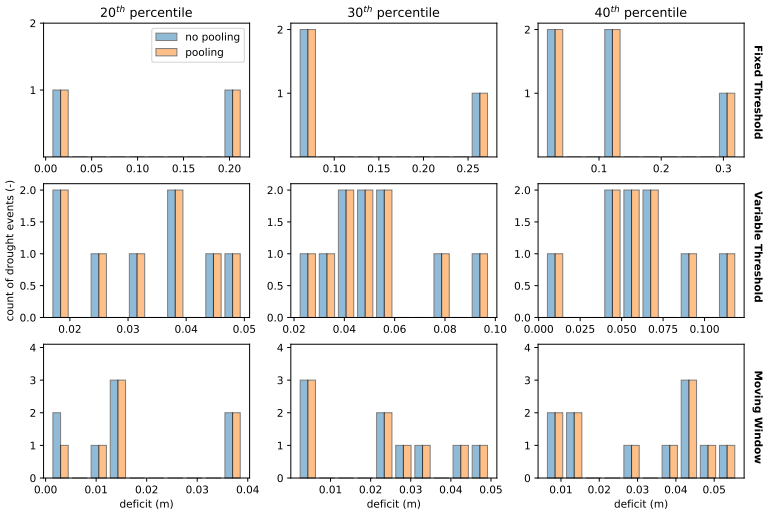


Figure B.6: Groundwater drought deficit histograms of well L-PB44 applying 5-day pooling or not, for three percentiles and three thresholds

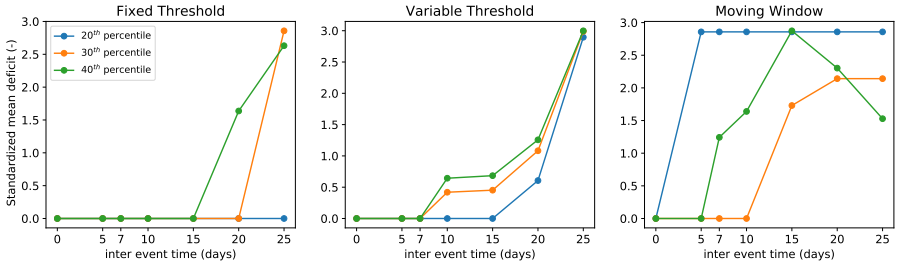


Figure B.7: Standardized mean deficit of groundwater drought events derived from no pooling and pooling using five, seven, 10, 15, 20, 25 days as inter-event time for well L-PB44

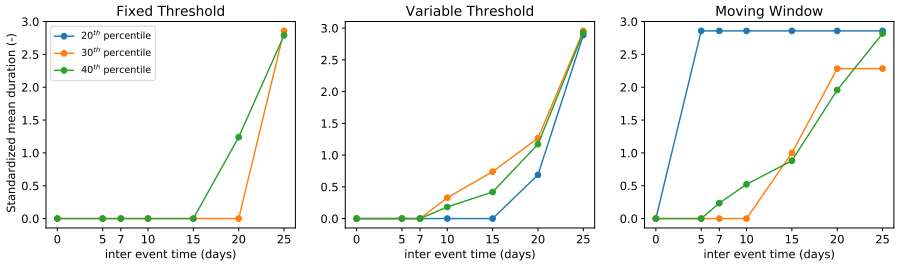


Figure B.8: Standardized mean duration of groundwater drought events derived from no pooling and pooling using five, seven, 10, 15, 20, 25 days as inter-event time for well L-PB44

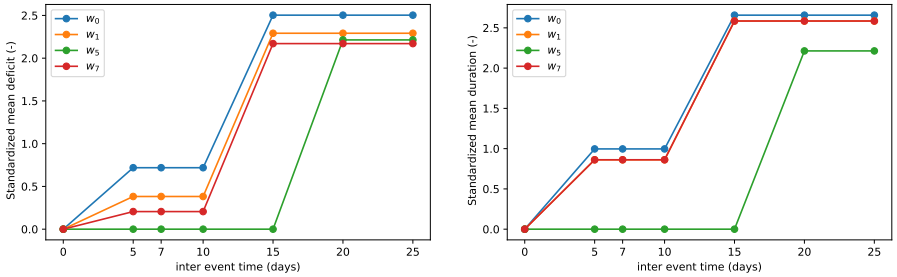


Figure B.9: Pooling for groundwater droughts identified using median GWL for all districts studied in Leiden (w_1 and w_7 districts showed the same standardization line for duration since the onset and end of their drought events occurred almost at the same time steps)

Deficit and Duration (spatial distribution)

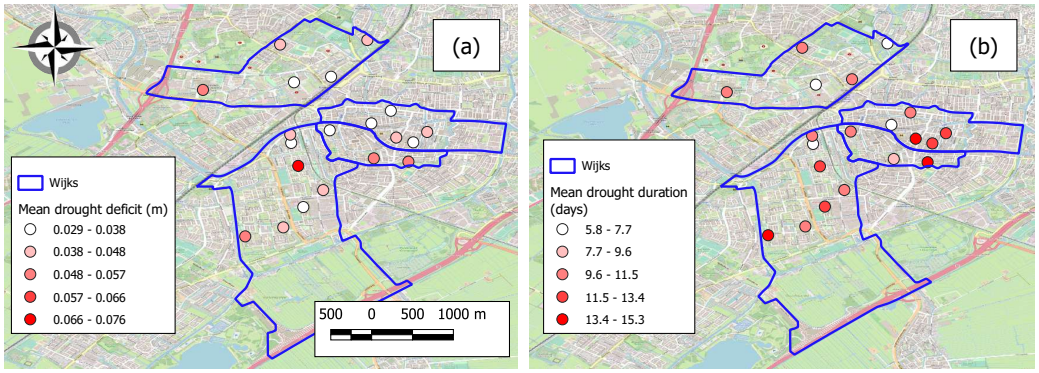


Figure B.10: Mean groundwater drought deficit (a) and duration (b) of all drought events occurred in April 2018 - March 2019 for each well in the study area (variable threshold method and 30th percentile was used to identify droughts)

B.4 Subsidence

B.4.1 Data Description

Two datasets were analyzed (the first one based on measurements and the second one based on modeling). More specifically, as for the first dataset, data from vertical deformation based on gravity and GNSS measurements was analyzed (Netherlands Center for Geodesy and Geo-Informatics (NCG), 2018). Data for the whole territory of the Netherlands was available. The grid resolution was 2 km x 2 km and measurements were taken in February of 2018. Total deformation consisted of shallow and deep subsidence (reasons which may have led to shallow subsidence were settling of peat soil and lowering of GWL, whereas reasons such as mineral extraction and coal mining could cause deep subsidence). Not the total ground surface movement is important but the movement compared to surface water level. This difference is mostly explained by shallow compaction on the top layer of soil. For that reason, the focus was on that part of subsidence.

Regarding the second dataset, data was derived by modeling subsidence for the period 2020-2050. Wareco company conducted this research and subsidence was considered as the sum of consolidation (compaction of soil) and oxidation (when organic matters oxidizes). The grid resolution of that dataset was 53 m x 54 m.

Regarding subsidence of infrastructure, there are two main mechanisms: (i) differential settlement on shallow foundations, and (ii) timber pile degradation. The current analysis is mostly focused on the latter. Some parameters which were ignored and could be used in vulnerability estimation) are mentioned below:

- Data regarding basements and cellars was not available. Differential settlement of a building is highly likely when a cellar is below it since the construction is uneven; the probability is lower for basements.
- There is no database about the buildings which have been repaired in the past.
- There are three types of timber pile foundations in the Netherlands: the Rotterdam foundation, the Amsterdam foundation and a Rotterdam foundation for which the pole is connected with a concrete beam instead of a wooden one (Schreurs, 2017). About 80% of the foundations nationwide follow Rotterdam style (Costa et al., 2020). Rotterdam style is more vulnerable than Amsterdam one. This distinction in types of timber pile foundations was not considered in the current analysis.
- Damage from past years to piles was unknown; it is cumulative.
- Another ignored parameter was the distance of the top of pole heads to ground surface; the lower it is, the higher the probability that pole heads are not covered with groundwater leading to poles' oxidation.
- The relationship between dryness and damage was not considered. It is quite uncertain since limited research has been conducted; it can be exponential when poles are dry for a long period of time (National Institute for Public Health and the Environment, 2019a).

In general, regarding subsidence, there are many factors which influence it; therefore, vulnerability of cities to subsidence could be a research topic on its own.

B.4.2 Subsidence (measured and modeled)

Measured Subsidence

There are regions which suffer more from subsidence due to their soil characteristics or even other causes such as extraction of gas, salt, and coal. In the wider region of study area, oil and gas were extracted in the past which may lead to subsidence with delay (see Figure B.11). However, this kind of information was not useful for this analysis; shallow subsidence is important for this research. Figure

B.12 depicts the shallow and total movement of ground surface (negative values mean subsidence). Deep movement was not presented since (i) it was not useful not being related to water level, and (ii) peculiar behavior was observed (positive values which could not be explained).

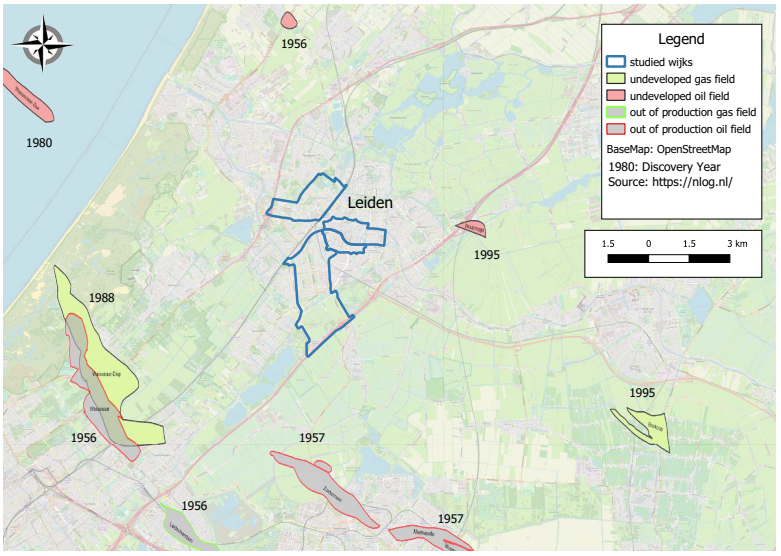


Figure B.11: Extraction fields in the wider area of Leiden with the years of discovery

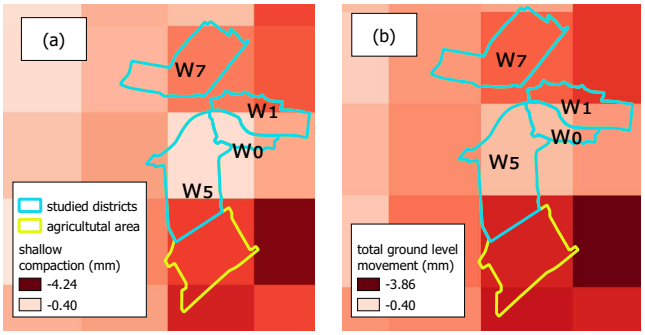


Figure B.12: Movement of ground level (shallow, and total) in Leiden based on satellite measurements of 2018 (Netherlands Center for Geodesy and Geo-Informatics (NCG), 2018)

Modeled Subsidence

As for the prediction of subsidence in Leiden for the period 2020-2050, its two components (consolidation and oxidation) are depicted in Figure B.13. There was no information for the majority of studied districts regarding oxidation. White pixels represent missing data. Consolidation was higher in w_1 district compared to the other ones; its mean value was 0.015 mm/year (see Table B.6).

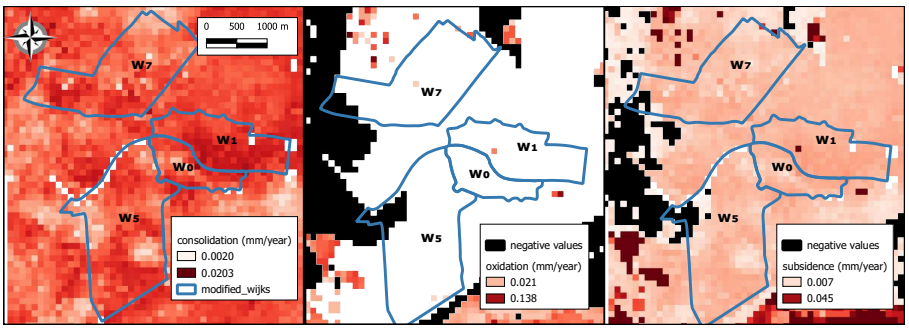


Figure B.13: Consolidation, oxidation, and subsidence in Leiden for the period 2020-2050 based on modeling conducted by Wareco company (negative values are probably modeling errors)

Table B.6: Descriptive statistics of modeled subsidence for all districts in Leiden (unit is 10^{-3} mm/year except skewness which is dimensionless)

	Subsidence			
	w_0	w_1	w_5	w_7
mean	13	15	12	12
std dev	3	5	3	3
max	19	70	18	36
min	5	11	0	2
skewness	-0.6	10.0	-1.3	0.9

Key results

- Modeled subsidence followed the pattern of consolidation.
- There was high uncertainty in both modeled and measured subsidence; the possibility of errors in data and modeling was high.

B.5 Difference between ground surface and minimum GWL (ontwatering)

B.5.1 Method

The difference between ground level and minimum GWL was estimated. Pre-processing was required for both cases. Ground level data was derived from the AHN website (www.ahn.nl/distributie) and specifically from the AHN_2 batch with accuracy of 0.5 m where all non-ground level objects such as buildings, and trees were removed and not filled in via an algorithm from the side of the AHN. Three map sheets were used to cover all the area of the selected districts (wijken). The null values were filled in via interpolation but water bodies were filled in as well which was undesirable. This issue was handled at the last step.

As for minimum GWL at spatial scale, the minimum GWL was selected for each well for the period April 2018 - March 2019 and then those values were interpolated via Inverse Weighted Interpolation (IWI) algorithm. Afterwards, the difference of interpolated GWL from the interpolated ground level was calculated. Water bodies were excluded from this raster having been transformed from polygons to raster beforehand. It is also worth noting that re-projections to coordinate system Amersfoor (EPSG:28992) were conducted applying geoprocessing tools.

B.5.2 Results

Figure B.14 illustrates the spatial variation of the maximum level difference between ground level and GWL for the study area. As can be seen, there are gaps due to water bodies removal as mentioned in Methodology Section B.5.1. The differences range from -1.6 m to 12.57 m but all negative values are wrong. Regarding the high values, they are explained by the historical landmark Burcht van Leiden which is located at the center of the city and is elevated compared to the surrounding area.

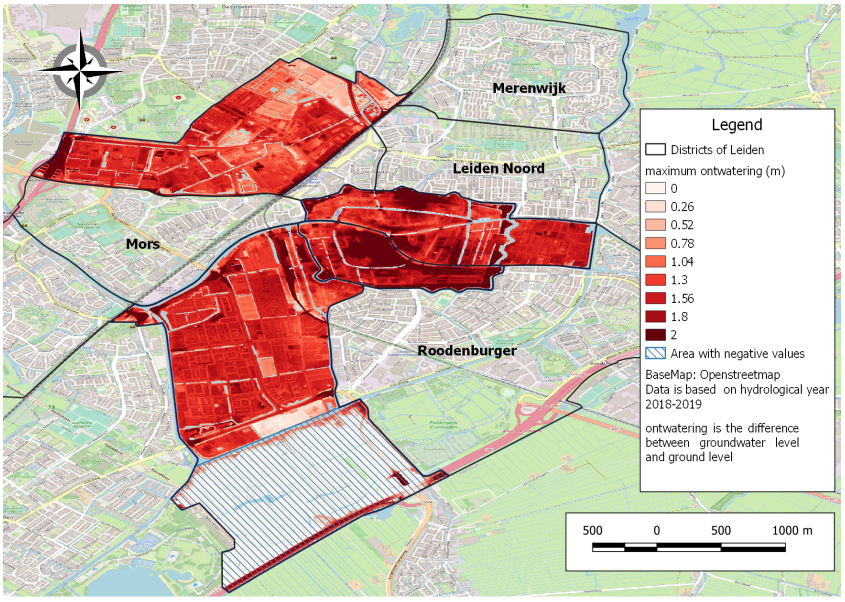


Figure B.14: Maximum ontwatering (difference between ground and GWL) using the minimum observed GWL of the hydrological year 2018-2019

The largest variation was observed on Bos-en Gasthuis district w_5 as can be seen in the blue hatched area on the south which is part of that district. In that specific region, all values were negative which is illogical. The reason is that the ground level in that part is quite low and the estimated GWL which was based on the nearest wells was high. There was no well at the blue hatched area; therefore GWLs were determined based on the interpolation of wells which were not situated there. As a result,

estimated GWL was higher than the ground surface which does not hold. The actual GWL in that area was lower than that of its neighboring area but the lack of measurements led to wrong results.

Temporal analysis of the minimum GWL over the hydrological year 2018-2019 was also analyzed (Figure B.15). The noticeable difference started in June and it peaked in August when GWL was considerably low. It is worth noting that the GWL remained low after August and it did not recover back to the values before May. Regarding the region with high GWL (white circular shape in w_1 district), its GWL remained almost constant throughout the year. In that spot, monument "Burcht van Leiden" is situated which is a medieval castle at elevated ground level. GWL tends to follow the topography; hence, this can be a possible explanation regarding that spot.

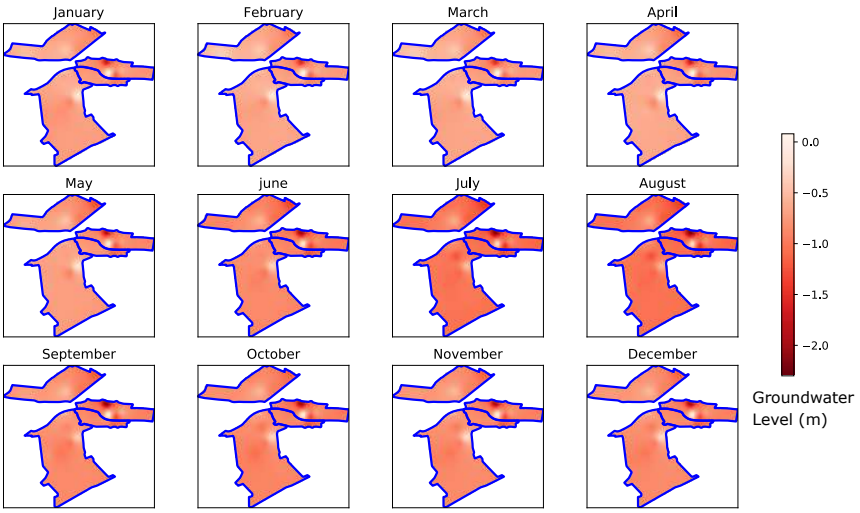


Figure B.15: Minimum GWL for each month of the period April 2018 - March 2019 for all districts

Key Results

- The largest ontwatering (difference between surface and minimum GWL) was noticed near the center of Leiden.

B.6 Standardized Groundwater Index (SGI)

B.6.1 Method

Besides threshold method, standardized indices can be used for the determination of drought severity and duration. Standardized Groundwater Index (SGI) for monthly data was used; SGI is based on the Standardized Precipitation Index (SPI). The CDF for each month of a year was transformed to a Gamma distribution and then to a normal one (the mean value equals zero and the variance equals one) (see Figure B.16). Gamma distribution requires positive values; for that reason GWL with respect to ground level (and not NAP) was used and all ground levels below ground surface were considered positive. The applied time steps (windows) on the current case study were one, three, six, and 12 months. For instance, in order to determine the three-month SGI, the steps were: (i) the rolling sum of a three-month window was estimated, (ii) calculation of CDF for each month, (iii) transformation to Gamma distribution and (iv) then to normal one. Additionally, droughts were categorized into seven classes based on SGI values (see Table B.7 which introduced by McKetie et al. (1993) for SPI). This categorization is valid for groundwater droughts as well.

Table B.7: Drought categories severity based on SGI values

SGI	Drought Category
2.0 or more	Extremely wet
1.5 to 1.99	Very wet
1.0 to 1.49	Moderately wet
0 to 0.99	Mildly wet
-0.99 to 0	Mild drought
-1.00 to -1.49	Moderate drought
-1.5 to -1.99	Severe drought
-2.00 or less	Extreme drought

In the study area, there were three monitoring wells with long data which did not have long gaps. However, only one out of those was used (B30F0222001, monitoring well '222' for short) since the other monitoring wells were not located in the city. They were situated in the w_5 district, specifically in the eastern part where there are some agricultural fields; thereforee, groundwater regime would be completely different than that one in the city. The gaps in monitoring well '222' were filled in using interpolation. Monitoring well '222' measured deep GWL in a confined aquifer and regarding the filter setting, it was placed from -22.81 m to -43.81 m (NAP).

Concerning the lithology of monitoring well '222', the first 4 m below surface level, there is loam; fine sand from 4.0 m to 7.5 m, and from 7.5 m to 14.4 m clay and loam (the majority in the latter part is clay). Then, from 14.4 m to 23.6 m, there is fine sand which becomes coarse at deeper levels (Dinoloket, nd). The time scale of measurements at that monitoring well was not constant (once or twice per month). Periods with high fluctuations were dismissed; the selected period was from January 1961 to December 1989. The requirement of SGI for time series of a period with more than 30 years was not met. However, the analysis continued due to lack of better data.

The source of the data was Dinoloket which is a repository with long groundwater data in the Netherlands (dinoloket.nl/en/subsurface-data). Data was provided with respect to the AHN reference system, and to ground level but the latter was chosen for SGI estimation. CDF plots were calculated to show the fit of a gamma distribution to the data for each month for the monitoring well '222'. Furthermore, a mosaic plot was created for the SGI highlighting how groundwater droughts evolved over the years.

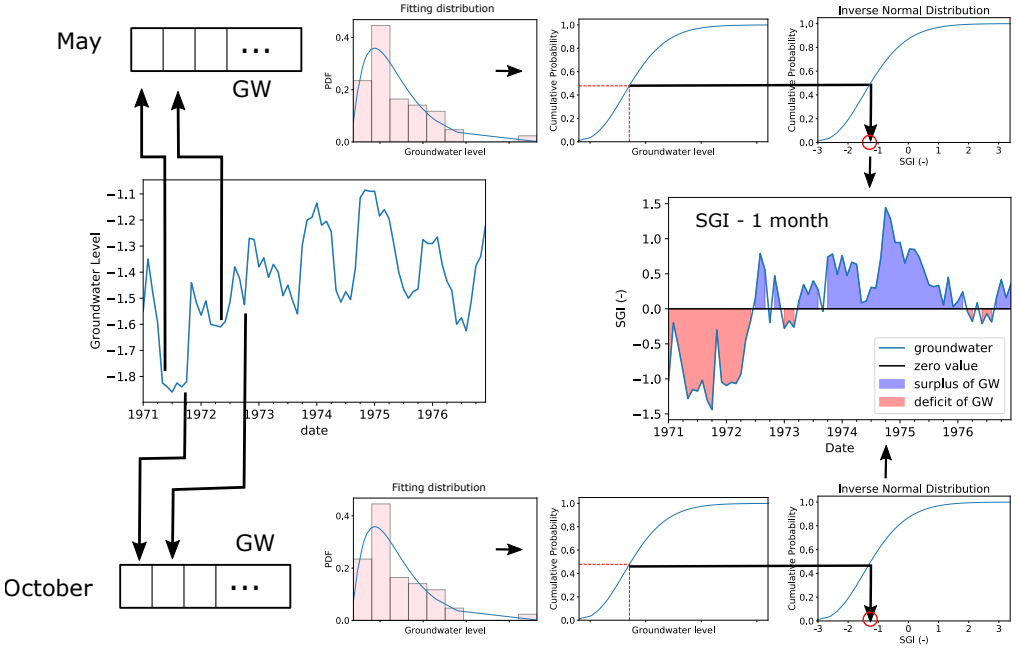


Figure B.16: Schematic of how SGI works (GW: groundwater level)

B.6.2 Results

Figure B.17 depicts the Gamma CDF of each month for '222' monitoring well which was used for the estimation of SGI. The differences among months were pronounced; the months which were close timewise such as September and July had similar CDFs as can be seen on the right part of Figure B.17. The difference between January and July, as well as between January and September were larger than that between September and July. As for the curvature of the CDFs, some months showed an abrupt rise but others not; for the latter, it may be an indication that gamma distribution may be not the distribution which fitted the data best. However, fit of other distributions was not tested.

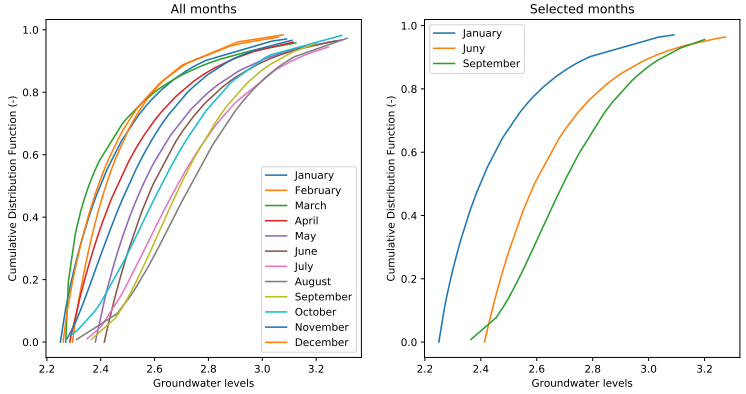


Figure B.17: CDF of gamma distribution which used for each month for the estimation of SGI for well '222'

SGI for four different time windows is presented in Figure B.18 and its unexpected results were highlighted. As time window increased, more smoothing occurred and the highest one was for time window of 12 months. There was a small surplus of groundwater in 1983 between two large periods of groundwater deficits but it attenuated as time window increased and almost disappeared at time window of 12 months. After 1977, there was almost a continuous drought period with some extreme SGI values (about -5) for some years such as the end of 1981, and the period 1987-1989. These drought events could not be explained. On top of that, the drop of SGI for the period 1987 to 1989 was lower

for the 3-month time window compared to the other time windows; this was unexpected as well.

Figure B.19 illustrates the SGI value of each month over the studied years in more detail. The most severe drought occurred from June 1987 to March 1988. On the contrary, there were some months with quite low values such as May of 1979, 1981, and April 1982 but their previous and following months had higher values; therefore the duration of extreme drought was limited to 1 month. Regarding the wettest years, they were 1969 and 1970. SGI values of year 1976 were higher (meaning less dry conditions) than its surrounding years (timewise) which was another unexpected point given that it is known that a severe drought happened in the Netherlands that year.

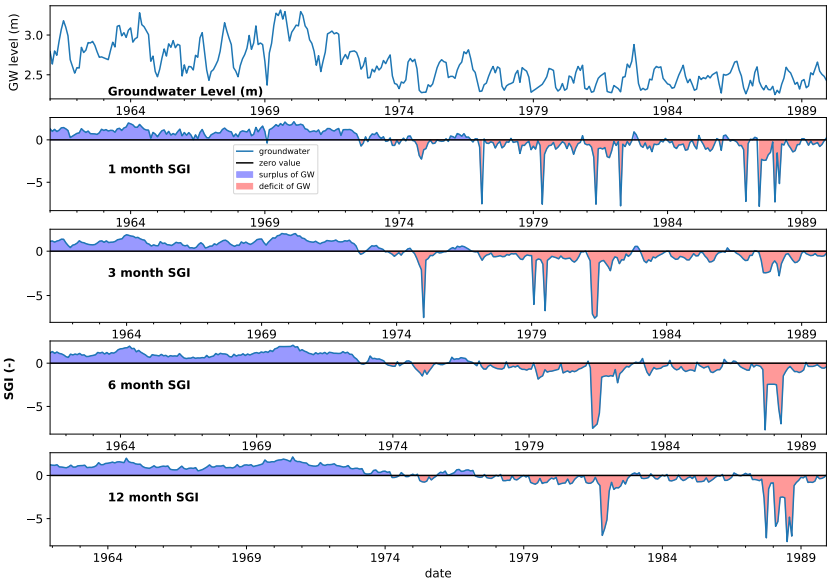


Figure B.18: SGI for well '222' over the period 1961-1989 using time windows of one, three, six, and 12 months

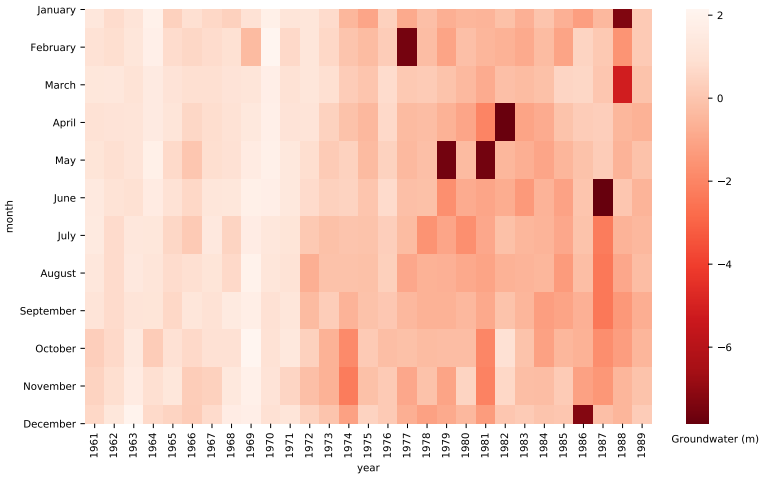


Figure B.19: Mosaic Plot of 1-month SGI based on month and year for the well '222' over the period 1961-1989

Based on SGI values, most of study period was categorized as normal (from 64% to almost 67%). The percentage increased from 1-month to 3-month, and from 6-month to 12-month time window but it decreased from 3-month to 6-month one (see Table B.8). Extreme droughts for all time windows did not exceeded 5% and there was neither clear positive nor negative relationship between time window and percentage value for that drought category. Regarding moderate and severe droughts,

their average percentages were around 6% and 1% respectively and their maximum percentage was noticed at different time windows (3-month one for moderate drought (8.67%) and 1-month one for severe drought (2.01%)).

Table B.8: Drought categorization of each time step based on SGI values for well '222'

	222 Station [%]			
	1-month	3-month	6-month	12-month
extremely wet	0.57	0.00	0.29	0.30
severely wet	5.46	5.49	5.54	6.53
moderately wet	16.09	15.32	18.66	19.58
normal	64.94	66.18	65.60	66.47
moderate drought	6.90	8.67	4.37	2.37
severe drought	2.01	0.58	1.75	0.89
extreme drought	4.02	3.76	3.79	3.86

The fact that the severe drought which happened in 1976 in the Netherlands was not recognized satisfactorily is an indication that data may contain errors. Furthermore, period 1987-1988 was quite dry compared to the rest years which is weird. A possible explanation is human intervention by Water Board in the region, e.g. construction works including excavation occurred; consequently GWL was lowered to reduce nuisance to workers (groundwater drawdawn). However, there was no further information in order to validate it. On top of that, it is highly unlikely that construction works took so much time. Consequently, there would be probably errors in measurements and for that reason, no further analysis was conducted using this dataset.

Key results

- SGI results were illogical and were dismissed for further analysis.

B.7 Calibration of UWBM

B.7.1 District w_1

B.7.1.1 Explaining GWL variation for w_1 district

Figure B.20 shows the measured GWLs of the wells with respect to NAP and ground surface. There was a high variation among the wells which revealed spatial heterogeneity in GWL. In the period of January to February for wells L-PB55 and L-PB56, the straight lines in GWL were due to interpolation which was applied to fill in data gaps.

Surrounding land use was not able to explain wells' GWL behavior in district w_1 satisfactorily. In gardens and parks, transpiration is high, consequently not much water is infiltrated to shallow groundwater table which can lead to large drops of GWL. However, this was not the case regarding L-PB55 which was located close to a garden. The garden was probably over irrigated; as a result much water was infiltrated to shallow groundwater table smoothing the variations. The rest wells were located under an open paved area (pavement with bricks) and higher recharge of groundwater table was expected since infiltrated water could not go back to the atmosphere via transpiration or soil evaporation. However, no clear pattern about this land use was determined.

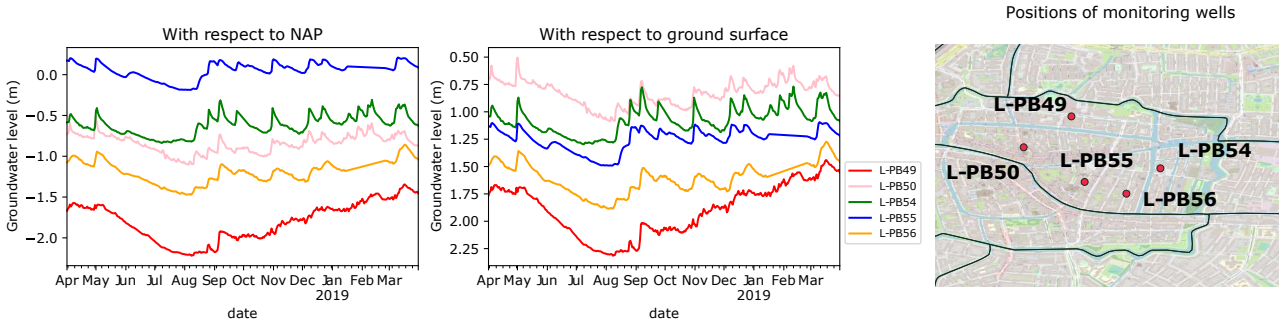


Figure B.20: GWL time series of monitoring wells in district w_1 with respect to NAP and surface level

Distance to open water bodies played an important role regarding the reaction of the monitoring well to forcing. Well L-PB54 was the closest to a canal compared to the other wells and its peaks were quite pronounced. After rainfall events, drainage of groundwater was quite fast since the GWL gradient was high due to short distance between the well and the canal.

B.7.1.2 Calibration results

Well L-PB49 was not included in Figure B.22 since its calibration was quite poor. Additionally, the behavior of this well was quite different compared to the others in the same district.

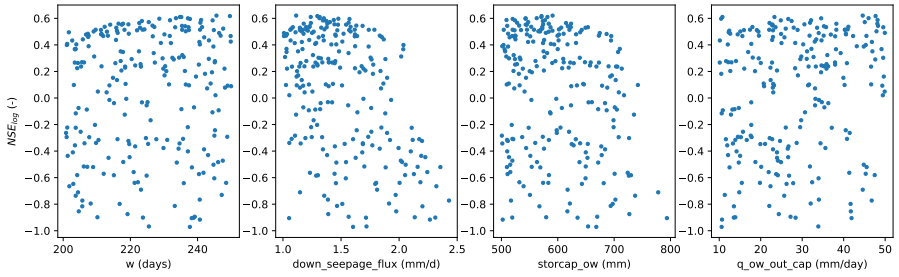


Figure B.21: NSE_{log} values for parameters' range using monitoring well L-PB50 for calibration in district w_1

Table B.9: Parameters' range for calibration of wells in district w_1

	Parameters			
well	w (days)	down_seepage_flux (mm/d)	storcap_ow (mm)	q_ow_out_cap (mm/d)
L_PB50	100-150	0-2	500-1000	10-50
L_PB54	100-150	0-2	800-1500	10-50
L_PB55	100-150	0-0.5	1000-1500	10-50
L_PB56	200-250	1-3	500-1000	10-50

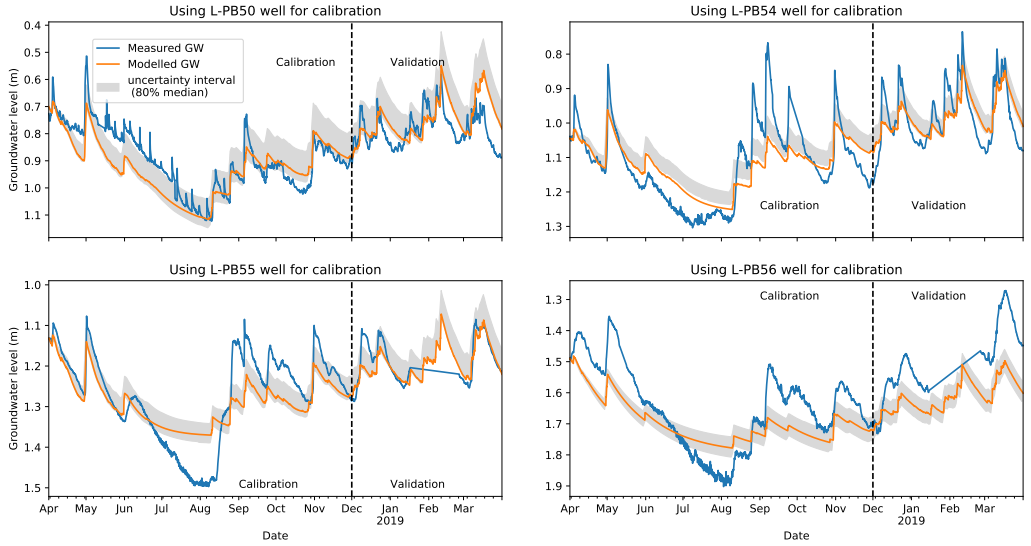


Figure B.22: Performance of UWBM calibration using four monitoring wells of district w_1 (the uncertainty interval was estimated based on the cases where NSE_{log}>-1)

Table B.10: NSE_{log} values when each calibrated parameter of the selected model using well L-PB50 (in district w_1) increased or decreased by 10%; not the model with highest performance was used but that one which approximated reality best

	Calibration - NSE_{log}(-)				Validation - NSE_{log}(-)			
	w	down_seepage_flux	storcap_ow	q_ow_out_cap	w	down_seepage_flux	storcap_ow	q_ow_out_cap
+10%	-0.044	-0.182	-0.748	0.138	0.131	0.019	-0.834	0.171
-10%	0.250	0.359	0.469	0.138	0.154	0.028	-0.528	0.171
default	0.138	0.138	0.138	0.138	0.171	0.171	0.171	0.171

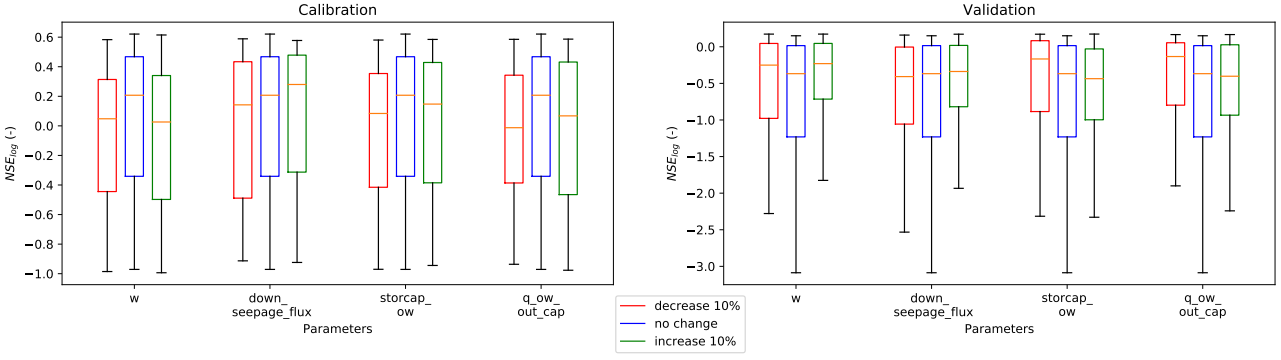


Figure B.23: NSE_{log} values of calibration and validation period for the cases of 10% increase and decrease for each calibrated parameter's range (using well L-PB50 in district w_1)

B.7.2 District w_5

Calibration using each of the monitoring wells in w_5 district except L-P73 was conducted (see Figure B.24). Well L-P73 was excluded due to its erratic behavior. Additionally, the uncertainty interval for wells L-PB72 and L-PB83 was estimated based on solutions with NSE_{log}>-2 and NSE_{log}>-4 respectively (not 0.2 which was used in the rest cases). There were not solutions with NSE_{log}>0.2 for those wells; thus, creation of uncertainty interval would not be possible. It is worth noting that the depicted modeled GWLs in Figure B.24 were not always those ones with the highest performance; for some cases, models which represented best the dynamics of GWL were used.

Uncertainty interval was not estimated for all cases. Due to poor performance of calibration using wells L-PB71 and L-PB82, their uncertainty interval was not estimated. Overall, none of the parameter sets of Figure B.24 could be used to predict GWL at a specific spot accurately but the best ones were using L-PB79, and L-PB80. The former outperformed in calibration and the latter in validation. Well L-PB80 was used for the estimation of soil moisture content; the other wells were rejected. The model using L-PB80 underestimated GWL from April to July but during July-August, it overestimated it.

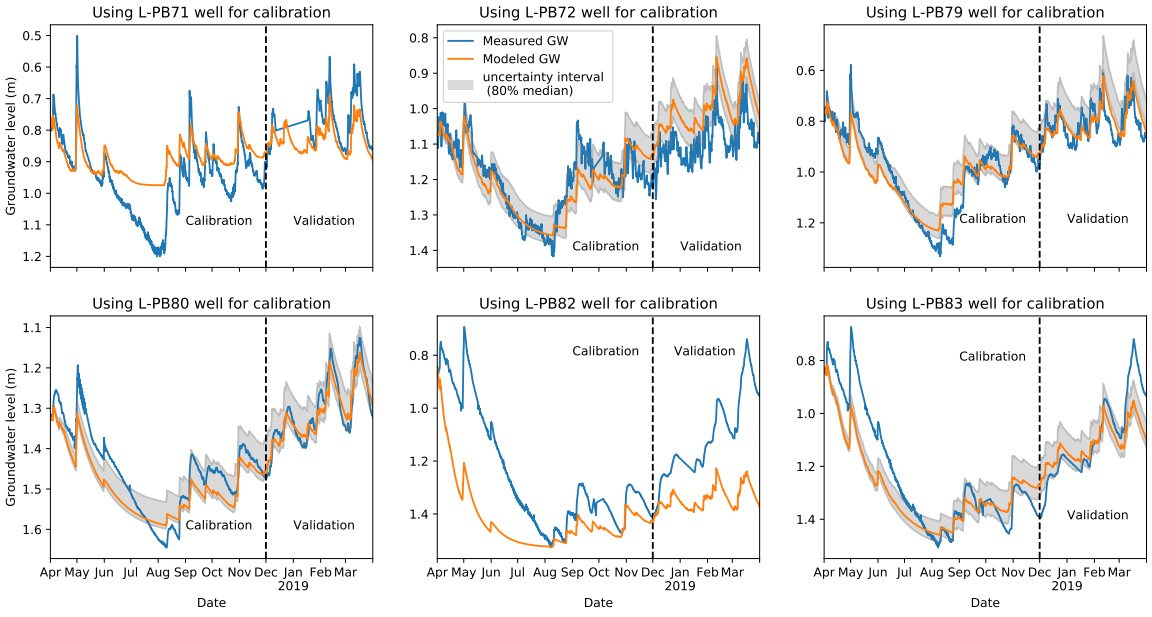


Figure B.24: Performance of UWBM calibration using six monitoring wells of district w_5 (uncertainty interval was not estimated for wells L-PB71 and L-PB82)

Table B.11: Parameters' range for calibration of wells in district w_5

	Parameters			
well	w (days)	down_seepage_flux (mm/d)	storcap_ow (mm)	q_ow_out_cap (mm/d)
L_PB71	50-80	1.5-2.5	640-860	10-50
L_PB72	140-200	1-2	800-900	10-50
L_PB79	180-250	0-0.5	800-900	10-50
L_PB80	200-250	1.5-2.0	1000-1200	10-50
L_PB82	100-120	1.6-1.8	1200-1800	30-50
L_PB83	200-300	1.5-2.0	800-900	30-40

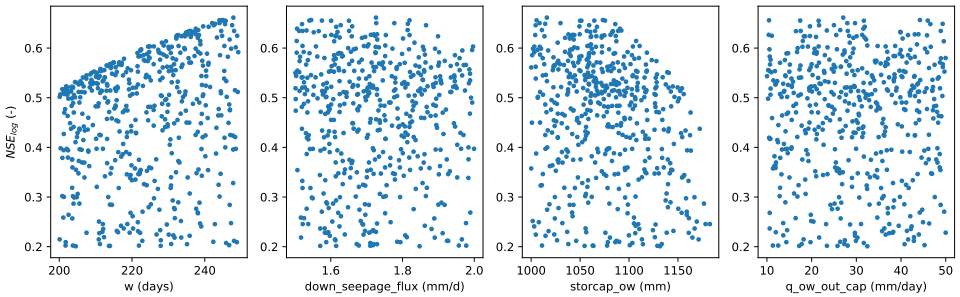


Figure B.25: NSE_{log} values for parameters' range using monitoring well L-PB80 for calibration in district w_5

Table B.12: NSE_{log} values when each parameter of the selected model using well L-PB80 (in district w_5) increased or decreased by 10%

	Calibration - NSE_{log}(-)				Validation - NSE_{log}(-)			
	w	down_seepage_flux	storcap_ow	q_ow_out_cap	w	down_seepage_flux	storcap_ow	q_ow_out_cap
+10%	0.202	0.128	-0.688	0.509	0.560	0.292	-2.512	0.846
-10%	0.587	0.656	0.130	0.509	0.668	0.581	-1.183	0.846
default	0.509	0.509	0.509	0.509	0.846	0.846	0.846	0.846

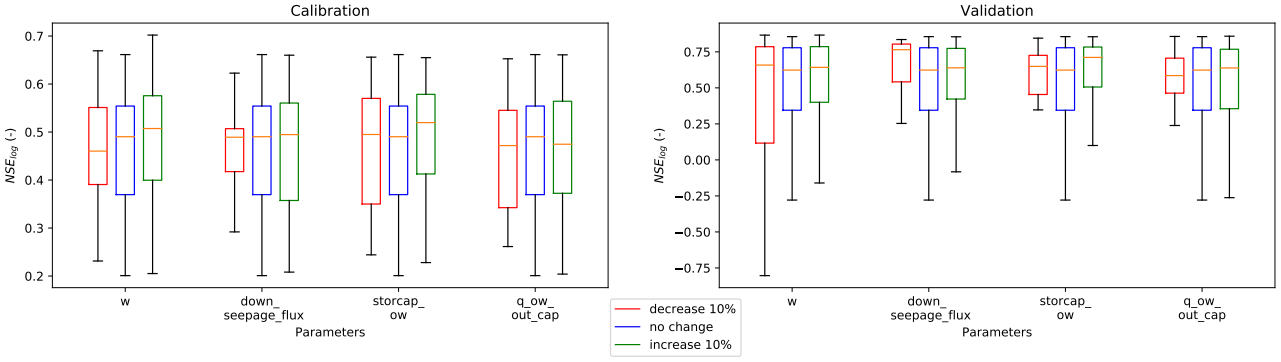


Figure B.26: NSE_{log} values of calibration and validation period for the cases of 10% increase and decrease for each calibrated parameter' range (using well L-PB80 in district w_5)

B.7.3 District w_7

The majority of models were not able to represent the dynamics of GWL. It is worth mentioning that in cases of (i) L-PB08 and L-PB14, and (ii) L-PB09 and L-PB13, the best models were the upper and lower bound of uncertainty interval respectively. Regarding well L-PB14, its uncertainty interval for GWL was estimated based on the solutions where NSE_{log} > -2.

Table B.13: Parameters' range for calibration of wells in district w_7

	Parameters			
well	w (days)	down_seepage_flux (mm/d)	storcap_ow (mm)	q_ow_out_cap (mm/d)
L_PB08	180-220	0-0.5	800-900	10-50
L_PB09	180-220	0-0.5	800-1200	10-50
L_PB12	180-220	0-0.5	800-1200	10-50
L_PB13	220-250	0-0.5	800-1000	20-40
L_PB14	100-500	0-0.5	1000-1200	10-40
mean L_PB09 & L_PB13	200-250	0-1	800-1200	10-50

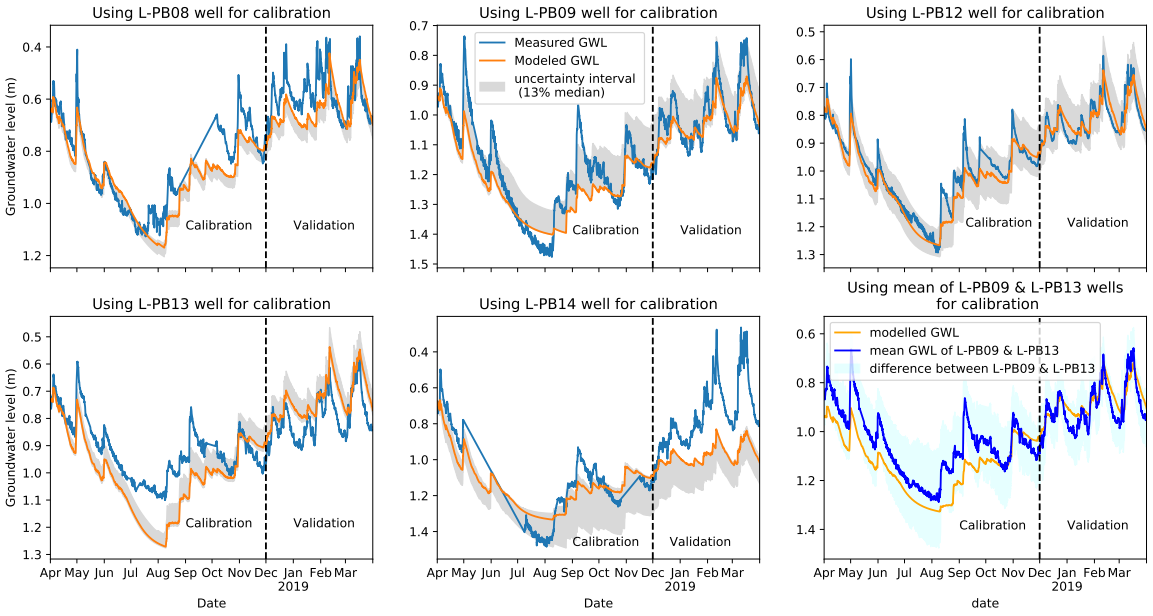


Figure B.27: Performance of UWBM calibration using five monitoring wells of district w_7 and the mean of L-PB09 and L-PB12 (in the last calibration plot)

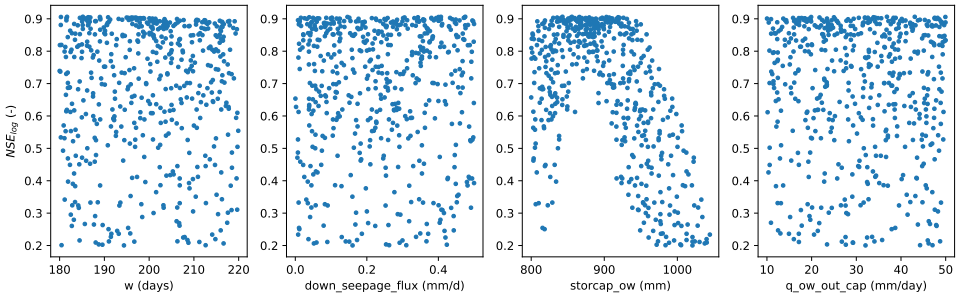


Figure B.28: NSE_{log} values for parameters' range using monitoring well L-PB12 for calibration in district w_7

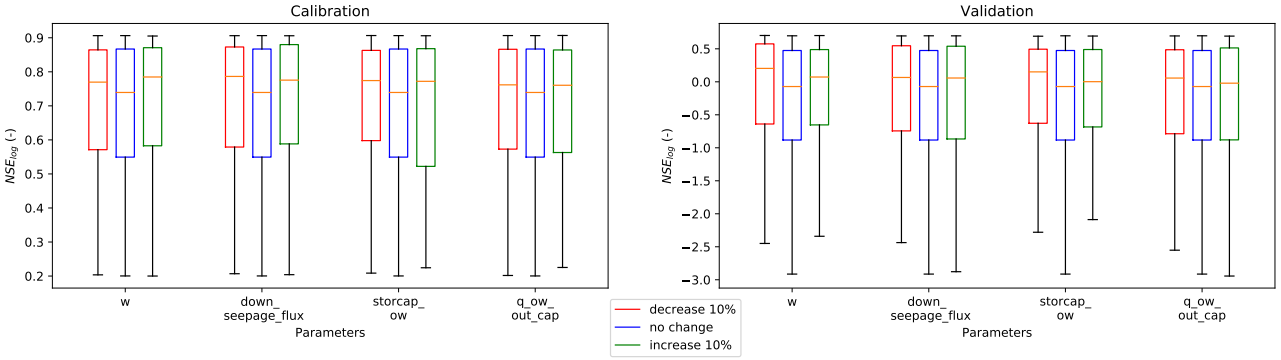


Figure B.29: NSE_{log} values of calibration and validation period for the cases of 10% increase and decrease for each calibrated parameter' range (using well L-PB12 in district w_7)

Table B.14: NSE_{log} values when each parameter of the selected model using well L-PB80 (in district w_7) increased or decreased by 10%

	Calibration - NSE_{log}(-)				Validation - NSE_{log}(-)			
	w	down_seepage_flux	storcap_ow	q_ow_out_cap	w	down_seepage_flux	storcap_ow	q_ow_out_cap
+10%	0.804	0.839	0.191	0.852	0.666	0.667	-1.955	0.664
-10%	0.885	0.864	0.701	0.853	0.621	0.654	-1.240	0.664
default	0.852	0.852	0.852	0.852	0.664	0.664	0.664	0.664

B.7.4 Models' Performance and Sensitivity Analysis (Method A) for districts w_1, w_5, and w_7

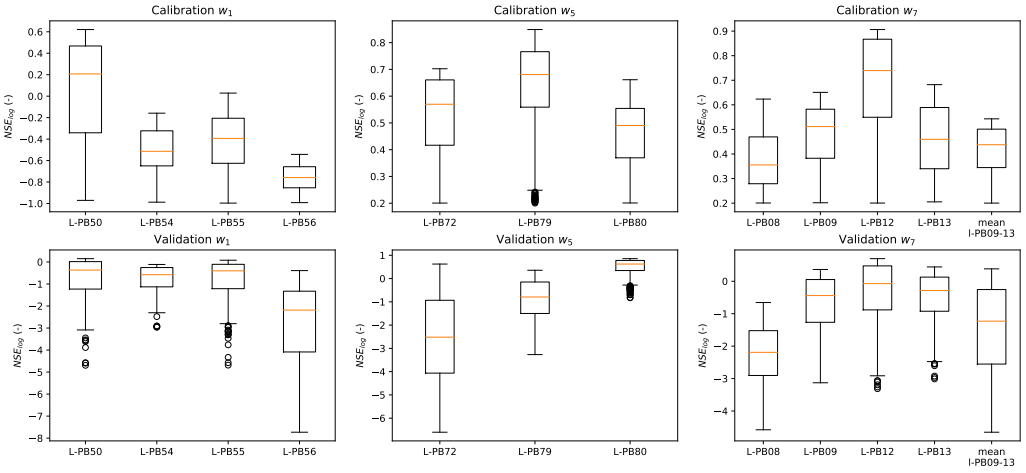


Figure B.30: NSE_{log} values of calibration and validation period for used models based on wells which are located in districts w_1, w_5, w_7

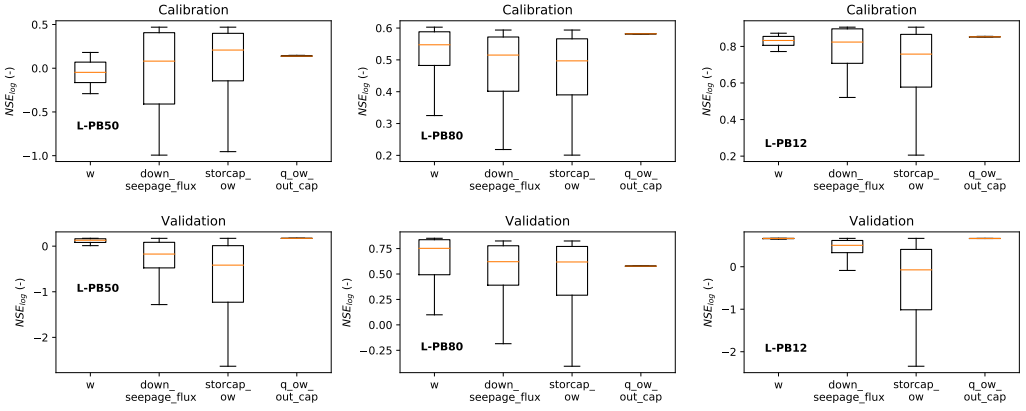


Figure B.31: NSE_{log} values for sensitivity analysis using Method A for w_1 (1^{st} column: well L-PB50), w_5 (2^{nd} column: well L-PB80), and w_7 (3^{rd} column: well L-PB12) districts (keeping fixed the selected parameter set except one of its parameter each time)

Solutions with NSE_{log}> 0.2 were used in districts w_5 and w_7 but those with NSE_{log}> -1 in district w_1. Validation results were based on the calibrated parameters which satisfied that condition.

Regarding sensitivity using Method A, it is worth highlighting that parameter q_ow_out_cap had the lowest sensitivity in districts w_1, w_5, and w_7 for both calibration and validation. On the contrary, storcap_ow had large variation for all districts. Large variation was also observed for parameter down_seepage_flux but not as much as storcap_ow. However for calibration in district w_0, down_seepage_flux variation exceeded that one of storcap_ow (based on interquantile range). Finally, the models were not highly sensitive to parameter w especially in districts w_1 and w_7.

B.8 Results for Soil Moisture Droughts

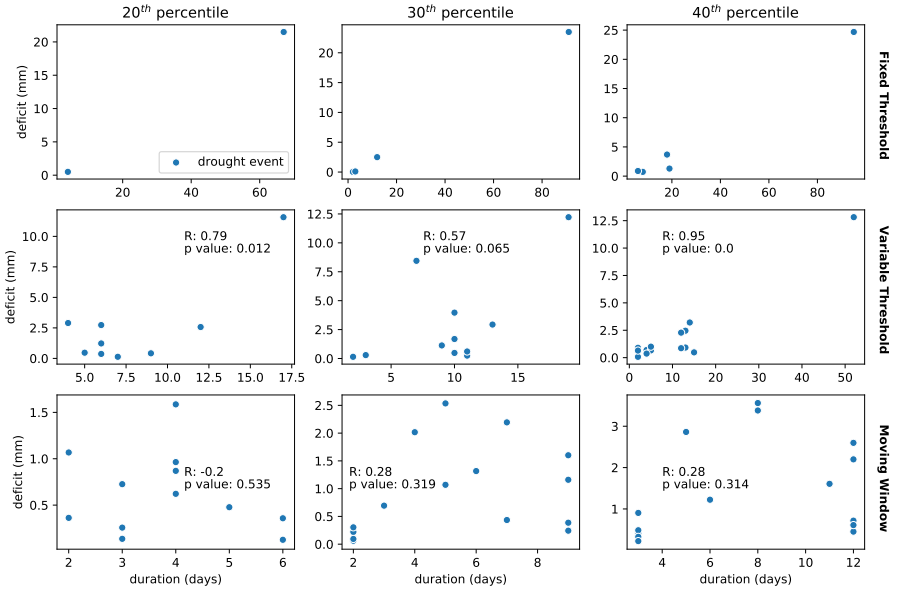


Figure B.32: Scatterplots of soil moisture drought duration and deficit using fixed, variable, and moving window threshold for three different percentiles (20^{th}, 30^{th}, and 40^{th} one) for district w_0 (Pearson coefficient and p value are attached only for variable and moving window method)

Table B.15: Pearson coefficients for soil moisture drought duration and deficit with their corresponding p-value for all districts (for drought events identified via variable threshold and moving window approach)

Method	Variable Method						Moving Window					
percentile	20^{th}		30^{th}		40^{th}		20^{th}		30^{th}		40^{th}	
District	R	p value	R	p-value	R	p-value	R	p-value	R	p-value	R	p-value
w_0	0.79	0.012	0.57	0.065	0.95	0.000	-0.20	0.535	0.28	0.319	0.28	0.314
w_1	0.91	0.002	0.87	0.001	0.87	0.000	0.21	0.551	0.55	0.033	0.49	0.054
w_5	0.74	0.036	0.81	0.002	0.83	0.000	0.13	0.711	0.17	0.553	0.19	0.489
w_7	0.82	0.006	0.88	0.000	0.86	0.000	0.05	0.867	0.32	0.270	0.21	0.044

High statistically significant ($p_{value} < 0.05$) relationship was found between soil moisture duration and deficit using variable threshold for almost all percentiles and districts (see Table B.15). An exception was district w_0 and 30^{th} percentile. More specifically, regarding the 40^{th} one, the relationship was statistically significant for all districts with a p_{value} less than 0.01 indicating that null hypothesis could be rejected and these results did not occur by chance. Visual inspection (see Figure B.32) and the fact that sample size was limited decreased the robustness of the relationship between soil moisture duration and and deficit. No statistically significant relationship was found using moving window ($p_{value} < 0.05$) except two cases; one for w_1 and the other one for w_7 but their Pearson coefficient (R) were not high (0.55 and 0.21 accordingly).

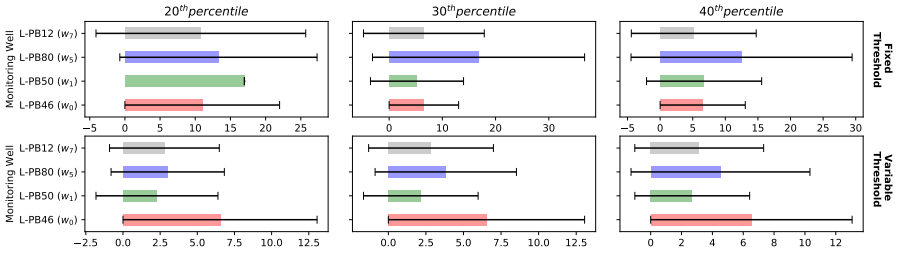


Figure B.33: Mean soil moisture drought deficit for districts studied in Leiden (w_0, w_1, w_5, and w_7) using fixed, and variable threshold for three different percentiles

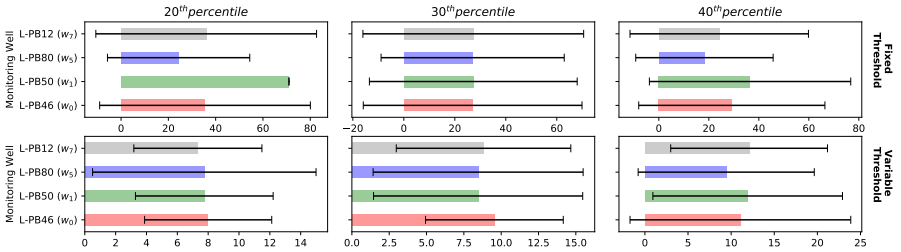


Figure B.34: Mean soil moisture drought duration for districts studied in Leiden (w_0, w_1, w_5, and w_7) using fixed, and variable threshold for three different percentiles

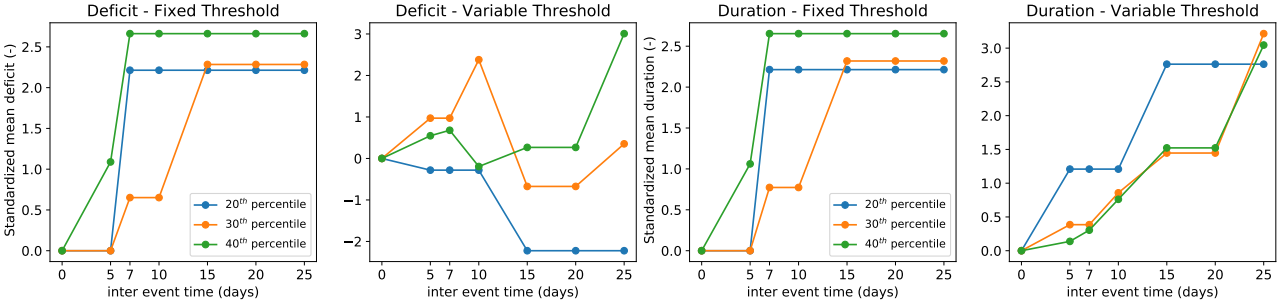


Figure B.35: Standardized mean deficit and duration of soil moisture drought events derived from no pooling and pooling using five, seven, 10, 15, 20, 25 days as inter-event time for district w_0 (fixed and variable threshold)

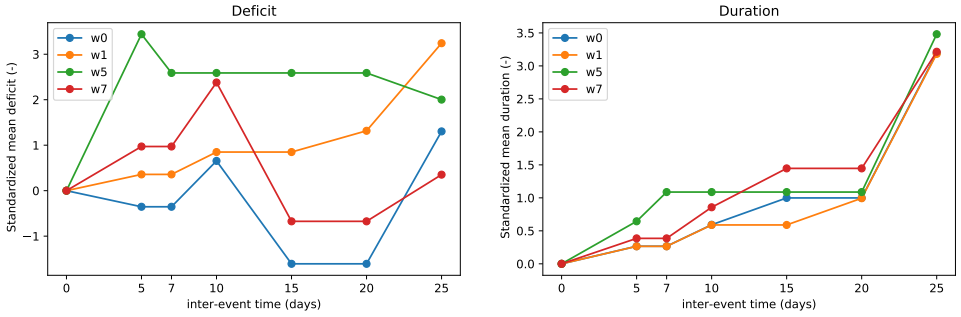


Figure B.36: Relationship between standardized mean deficit (left part), duration (right part) and inter-event time for pooling in four districts of Leiden regarding soil moisture droughts (variable threshold and 30$^{\text{th}}$ percentile was used to identify droughts)

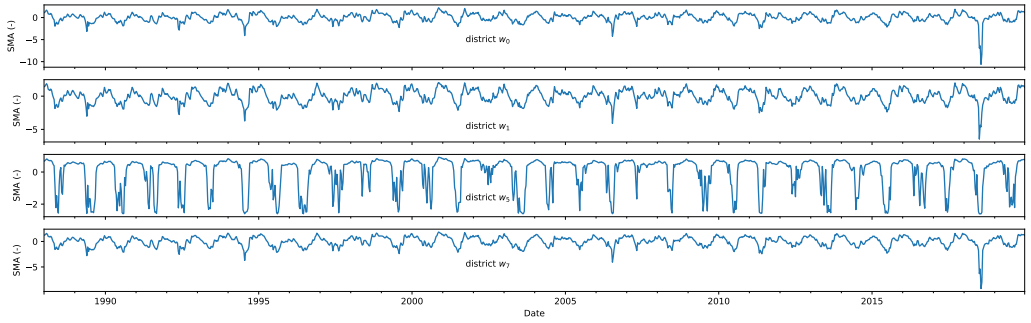


Figure B.37: Soil Moisture Anomaly (SMA) for districts studied in Leiden (period: 1988-2019)

B.9 Social Sensitivity indicators for Groundwater Droughts

B.9.1 Data Description

In this part, social sensitivity indicators for groundwater droughts are presented which are categorized based on population, economy, and social security. Regarding population, those were: percentage of minorities to the total population (Minor), average household size (number of residents over the number of households) (Avg_Hous), percentage of people aged 65 and over (pop_65), and population density per km^2 (pop_dens). As for the minority percentage, the total number of immigrants with non-Western background was considered.

Regarding economy, the indicators used were: percentage of residents with the lowest 40% income nationwide (low_inc_P), percentage of households with the lowest 40% income nationwide (low_inc_H), number of income recipients (Inc_recip), average income per income recipient (Avg_Inv_re), and average income per resident (Avg_inc_p). Indicators Inc_recip, Avg_inc_re, and Avg_inc_p had a negative relationship with sensitivity; thus Formula (3.14 right part) was used for normalization. The percentage of disabled people and unemployed per 100 inhabitants (disabl_perc and unempl_perc respectively) were used as social security indicators. For the two latter variables, raw data was absolute, hence it was normalized using the total population of each district. The use of absolute numbers could lead to misleading results since the number of unemployed people would be higher in districts with large population. For that reason, unemployment rate was estimated per 100 residents. All aforementioned social indicators were derived by Centraal Bureau voor de Statistiek (CBS) (2017) in 2017 except two variables (Avg_Inv_re and Avg_inc_p) for which data from 2015 (Centraal Bureau voor de Statistiek (CBS), 2015) was used since data in 2017 contained many missing values. Besides, it is worth noting that data was collected at neighborhood level.

This paragraph explains why social sensitivity indicators are important for vulnerability estimation to groundwater droughts. One of the impacts of groundwater droughts is subsidence; in that case, cords or pipes (used for electricity or drinking water transport) in the household may over-extend and as a result they may malfunction or even be cut off which makes living conditions in the house unbearable. Drinking water and gas pipes are more fragile compared to electricity and tele-communication ones which are more flexible. Low classes cannot probably afford repair costs increasing their sensitivity. This is the reason why that aspect was considered. Relocation may be possible for residents whose houses suffered severe damage. However, this would be an agonizing process for groups with mobility limitations (such as the elderly and the disabled).

Many variables were quite similar; hence, they did not contribute to the vulnerability index. For that reason, their correlation was investigated and out of indicators with high correlation (positive or negative), only one was saved for further analysis (the other ones were dismissed). Kendall Tau correlation coefficient and p-value were estimated for that step. Kendall Tau was preferred since (i) the sample size was small, and (ii) there were outliers. Pearson coefficient was not selected since it works poorly for data with outliers (Chok, 2010).

Neighborhood (buurt) Leeuwenhoek in w_7 district did not have values for demographic and socio-economic indicators since it did not meet the requirements of the minimum number of inhabitants. In case the number of inhabitants in a neighborhood is less than 100, statistics are not estimated by Centraal Bureau voor de Statistick. This neighborhood was not excluded from sensitivity since the latter one is not only related to demographics and socio-economics. Moreover, removal from sensitivity estimation would mean removal from vulnerability estimation. Replacing its missing values with very high or low ones (based on the indicator) was not straightforward and could have influenced the variation of data leading to misleading normalization. For that reason, after normalization applied (without including this neighborhood) for each socio-economic indicator, a zero value was assigned to this neighborhood meaning that there was no sensitivity regarding socio-economic factors.

Regarding social sensitivity indicators, after preliminary testing of the questionnaires for AHP, drought experts faced difficulties comparing them. It was hard to find reasons to value one variable (indicator) higher than the others concerning droughts; for that reason, only one social indicator was used. Indicator low_inc_H was considered the leading social sensitivity factor and the other social variables were dismissed.

B.9.2 Correlation of Social Sensitivity Indicators

The left part of Figure B.38 shows the Kendall Tau correlation coefficient among social indicators in Leiden whereas the right part of that Figure presents their p-values. Indicator Avg_Hous showed a negative correlation with low_inc_P, and low_inc_H (-0.82 and -0.57 respectively); therefore, only one of the three variable was used in following analysis (low_inc_P was selected). The correlation of those variables was also statistically significant ($p_{values} < 0.1$).

Additionally, Avg_Hous was correlated with pop_dens and pop_65 negatively and positively respectively. However, Avg_Hous had already been dismissed. Parameter low_inc_P was also correlated with pop_dens and pop_65 but its correlation was lower compared to that of low_inc_H with those two population variables. That strengthened the argumentation of selecting low_inc_P instead of low_inc_H. Indicator pop_dens was not dismissed since the correlation with low_inc_P was not that high. In the same way, some other variables were dismissed. Therefore, the remaining variables were: Minor, pop_65, low_inc_H, unempl_per, and disabl_perc

pop_65

disabl_perc

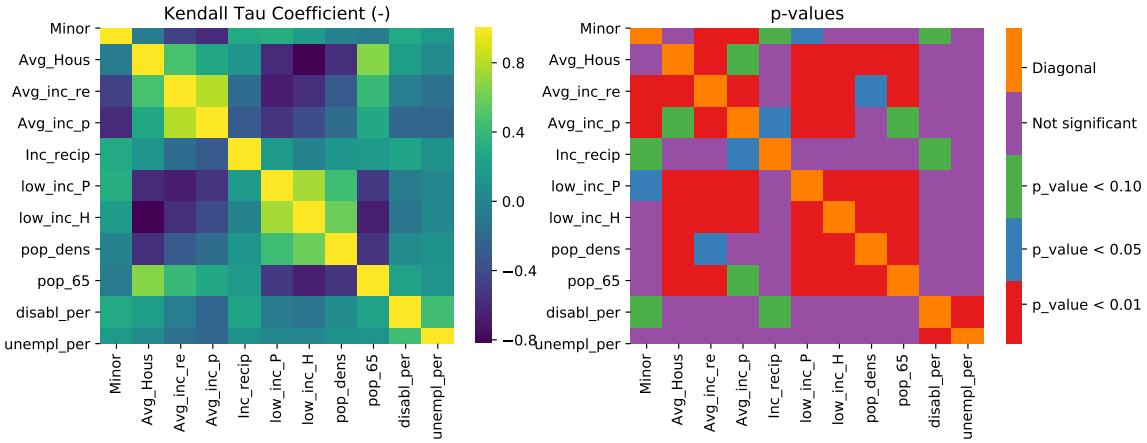


Figure B.38: Kendall Tau correlation coefficient and p-values for social sensitivity indicators

Appendix C

Python Scripts

C.1 Identifying droughts events and estimating their characteristics

This part includes the main functions used in the current study for drought identification and calculation of drought characteristics (deficit and duration). The algorithm ignores the first time step in case a drought event has started before the beginning of the time series. Comparison with the previous time step is not possible; therefore, the function starts counting from the second time step.

The scripts can be found at github.com/iliasmachairas/MSc_Thesis_Scripts

C.2 Spatial droughts events

Scripts of spatial droughts were not uploaded on Github since they required many large files and space is limited. For that reason, the most important code snippets have been attached.

```

# Function to convert the interpolated raster to np.array
def Reading_img(path):
    import imageio
    im = imageio.imread(path)

    im_ar = np.array(im, dtype='float')
    im_ar[im_ar <-99] = 99
    # After converting the null values on the edges to np.NaN, ArcGIS use a very  low number to represent null values
    im_ar[im_ar > 98] = np.NaN

    return im_ar

# Reading data for June.
list_img_June = []
for i in range(1,31): # for July it would be 30
    path = "Arc_GIS_Files/june_"+str(i)+".tif"
    list_img_June.append(Reading_img(path))

# Converting 2D arrays in one 3D array including all days of June
array_3d_June = np.zeros((30, 26, 72), dtype='float')
for i in range(30):
    array_3d_June[i] = list_img_June[i]

# Reading data for February
# February of year 2017-2018 was not a leap one
list_img_Feb = []
for i in range(1,29):
    path = "Arc_GIS_Files/feb_"+str(i)+".tif"
    list_img_Feb.append(Reading_img(path))

# Converting 2D arrays in one 3D array including all days of February
array_3d_Feb = np.zeros((28, 26, 72), dtype='float')
for i in range(28):
    array_3d_Feb[i] = list_img_Feb[i]

# The other months are not presented for space saving reasons. However, the code is the same; the difference is that the loop
    end for reading the interpolated images was either 31 or 32, and the first dimension of the 3D array was either 30 or 31
    based on the month.

# Creating the 3D array for the whole year
array_3d_merge = np.concatenate([array_3d_Apr, array_3d_May, array_3d_June, array_3d_July, array_3d_Aug, array_3d_Sep,
    array_3d_Oct, array_3d_Nov, array_3d_Dec, array_3d_Jan, array_3d_Feb, array_3d_Mar])

# Estimating areal threshold for 40th percentile
```

```
thres_ar = np.zeros((26,72)) # 26, 72 are the rows and columns of each interpolated raster
for j in range(26):
    for k in range(72):
        x = array_3d_merge[:,j,k]
        thres = np.nanpercentile(x,40)
        thres_ar[j,k] = thres
```

Listing C.1: Pre-processing for Spatial Drought Identification

```
tmax = array_3d_merge.shape[0]
k = 0 # drought events
onset = np.zeros(tmax)
end = np.zeros(tmax)
j_max = int(tmax/2)
Dur = np.zeros((j_max,tmax))
Def = np.zeros((j_max,tmax))

for i in range(1, tmax-1):
    Dif_before = array_3d_merge[i-1,:,:] - thres_ar
    Dif_current = array_3d_merge[i,:,:] - thres_ar
    Dif_after = array_3d_merge[i+1,:,:] - thres_ar
    # Removing NaN
    Dif_before_clip = Dif_before[np.isfinite(Dif_before)]
    Dif_current_clip = Dif_current[np.isfinite(Dif_current)]
    Dif_after_clip = Dif_after[np.isfinite(Dif_after)]

    if (np.all(Dif_before_clip >= 0) & np.any(Dif_current_clip < 0) & np.all(Dif_after_clip >= 0)):
        # Drought of 1 time step
        start=i
        onset[k]=start
        end[k]=start
        Dur[k,i-start]=np.count_nonzero(Dif_current_clip < 0)
        Def[k,i-start]=np.sum(Dif_current_clip[Dif_current_clip < 0])
        k+=1
    elif (np.all(Dif_before_clip >= 0) & np.any(Dif_current_clip < 0) & (i != tmax - 2)):
        # Start
        start = i
        onset[k] = start
        Dur[k,i-start]=np.count_nonzero(Dif_current_clip < 0)
        Def[k,i-start]=np.sum(Dif_current_clip[Dif_current_clip < 0])
    elif (np.any(Dif_current_clip < 0) & np.all(Dif_after_clip >= 0) & (i!=1)):
        # End
        end[k] = i
        Dur[k,i-start]=np.count_nonzero(Dif_current_clip < 0)
        Def[k,i-start]=np.sum(Dif_current_clip[Dif_current_clip < 0])
        k+=1
    elif (np.any(Dif_current_clip < 0)):
        if (i==1 & np.all(Dif_after_clip < 0)):
            # Start phase of a drought event longer than 1 time step'
            start=1
            Dur[k,i-start]=np.count_nonzero(Dif_current_clip < 0)
            Def[k,i-start]=np.sum(Dif_current_clip[Dif_current_clip < 0])
        elif (i==1 & np.all(Dif_after_clip > 0)):
            # Case when time series starts with a drought event of 1 time step
            start=1
            onset[k]=start
            end[k]=start
            Dur[k,i-start]=np.count_nonzero(Dif_current_clip < 0)
            Def[k,i-start]=np.sum(Dif_current_clip[Dif_current_clip < 0])
            k+=1
        elif ((i == tmax - 2) & np.all(Dif_before_clip > 0)):
            # Case when the time series ends with a drought event of 2 time steps
            start=i
            onset[k]=start
            end[k]=i+1
            Dur[k,i-start]=np.count_nonzero(Dif_current_clip < 0)
            Dur[k,(i-start+1)]=np.count_nonzero(Dif_after_clip < 0)
            Def[k,i-start]=np.sum(Dif_current_clip[Dif_current_clip < 0])
            Def[k,(i-start+1)]=np.sum(Dif_current_clip[Dif_after_clip < 0])
            k+=1
        elif (i == tmax - 2):
            # End when the time series finishes with incomplete drought but
            # has started earlier than 2 time steps
            Dur[k,i-start]=np.count_nonzero(Dif_current_clip < 0)
            Dur[k,(i-start+1)]=np.count_nonzero(Dif_after_clip < 0)
            Def[k,i-start]=np.sum(Dif_current_clip[Dif_current_clip < 0])
            Def[k,(i-start+1)]=np.sum(Dif_current_clip[Dif_after_clip < 0])
            end[k]=i+1
            k+=1
        else:
            # Normal intermediate step
            Dur[k,i-start]=np.count_nonzero(Dif_current_clip < 0)
            Def[k,i-start]=np.sum(Dif_current_clip[Dif_current_clip < 0])
```

Listing C.2: Drought Identification using a spatially aggregation approach based on fixed threshold method

```
# Onsets and ends of drought events
onset_clear=onset[:k]
end_clear=end[:k]

# Drought Duration

# Max of elements per row
maxim_rows=np.max((Dur!=0).sum(axis=1))
# Reduced Dur array based on drought events (in order to remove the 0 values from it)
```

```
Dur_final=Dur[:k,:maxim_rows]
# durations
durations=(Dur_final != 0).sum(axis = 1)

# Estimating sum of negative cells per drought over area for each time step
count_neg_cells =np.zeros(k)
for i in range(k):
    count_neg_cells[i] = np.sum(Dur_final[i])
    # the rest values are negative - in this way, I estimate the count

total_count_pixels = (array_3d_merge.shape[1]*array_3d_merge.shape[2])
non_null_cells_count = np.count_nonzero(~np.isnan(array_3d_merge[56,:,:]))
# the number 56 is random, it can be whatever
# the null values are constant over the first dimension of array_3d_merge
null_cells = total_count_pixels - non_null_cells_count

# Areal fraction of cell
areal_fraction = 1/(non_null_cells_count)
areal_fraction
# Drought Area per drought event
dr_area = (areal_fraction * count_neg_cells) / durations

# Drought Deficit

# Max of elements per row
maxim_rows = np.max((Def!=0).sum(axis=1))

# Reduced Def array based on drought events (in order to remove the 0 values from it)
Def_final = Def[:k,:maxim_rows]

# Estimating deficit over the drought duration
def_droughts = np.zeros(k)
for i in range(k):
    def_droughts[i] = np.sum(Def_final[i])

def_droughts_final = -1 * def_droughts * areal_fraction / durations
```

Listing C.3: Estimating drought duration and deficit spatially

Appendix D

Case Study Athens

D.1 UWBM Implementation

Another case study outside the Netherlands could provide a useful insight into groundwater and soil moisture urban droughts. Athens was selected but analysis was not conducted due to lack of data. There was no agency responsible for recording groundwater measurements and if some measurements were taken, their duration was limited. The only data which was found was about soil moisture for the National Garden but its duration was quite low. This is an indication that the policy makers and water managers do not recognize groundwater and soil moisture droughts; it does not mean it does not not experience them though. Not being able to estimate the water balance in Athens showed that the city is quite vulnerable and unable to monitor its water needs especially in the future since summers would be more extreme in Mediterranean region. Water balance was not estimated for the National Park since (i) it does not represent the whole residential area, and (ii) it is irrigated and there were no measurements about the water supply.

Using UWBM for a study area different than lowlands and without calibration is highly risky for mistakes. For that reason, no results of drought severity and duration are presented. UWBM without calibration was just implemented to test whether the pattern is logical or not. The district selected was located in the center of Athens consisting of Koukaki, Acropolis, and Petralona neighborhoods (Figure D.1-a). Districts in Leiden are different than Athens size-wise. This region was of interest since it includes parks used as tourist destinations. Potential Evaporation was estimated using empirical formula of Hargreaves (Section D.2) since there was lack of weather parameters data in Vyronas station and Penman was not able to be applied. The used period was old (1984-1991) since recent data had many gaps. Vyronas station was selected since it was the closest to study area. Daily potential evaporation was estimated via Hargreaves formula but UWBM requires hourly one. To overcome it, it was assumed that evaporation occurred only daytime hours between 06:00 and 18:00. Thus, daily evaporation was split into 12 time spans. It is also worth mentioning that daily precipitation was only available and was converted to hourly which led to removal of extremes.

The parameters of UWBM were configured for the case of Athens so that it can be more suitable for the conditions which hold in that city. The most important changes were the following ones: w increased dramatically (in order to deactivate drainage of groundwater to open water), fraction of open water was minimized to 1%, the fraction of storm water sewer system was set 100%, deep GWL was considered as 5 m (assumption), storcap_ow and q_ow_out_cap were minimized. Regarding the interception capacity of different land use classes, the same values as those for Leiden were used. As for land uses' areas, paved roofs and unpaved classes were measured via openstreetmap. For the percentages of the rest land use classes, some assumptions were made. About geology, most study area is covered by schist (relatively impermeable) but this category is not present in UWBM. The

same soil type as Leiden was used but this was not right. This parameter affected the results since it was related to soil moisture characteristics such as field capacity and permanent wilting point.

Despite the parameter changes and manipulation of UWBM, some unexpected trend was observed for GWL (drop from 1989 to 1991) (Figure D.1-c). On the contrary, the pattern of soil moisture seemed more realistic with long dry conditions for the period April-August meaning that soil water content reached its permanent wilting point (maybe w was set too high). Soil moisture seems to follow a sinusoidal annual cycle and this pattern is affected by meteorological characteristics of the region. Even though this pattern seems logical, it does not mean that soil moisture was right since the model was uncalibrated. UWBM might work (under conditions) for the estimation of soil moisture but not for GWL. There is an input flux to groundwater storage but not output given that GWL is quite deep and disconnected from the rest fluxes; therefore, UWBM is not suitable.

Given that exposure was not able to be estimated in Athens, vulnerability was not determined. Exposure and sensitivity are two components of vulnerability. Even if sensitivity was estimated, vulnerability would not be assessed. For that reason, sensitivity was not computed.

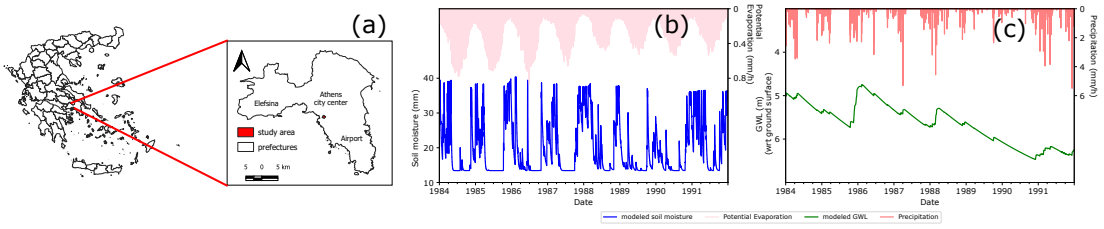


Figure D.1: Map of study Area in Athens (a), modeled soil moisture for the period 1984-1991 (b), modeled GWL for the same period (c)

D.2 Estimation of Evaporation using Hargreaves formula

Potential Evaporation in Athens was estimated based on Hargreaves formula (Hargreaves and Samani, 1982). Thornthwaite and Hargreaves approaches are commonly used for the empirical estimation of evaporation where there is a lack of measured hydrological data and Penman formula cannot be applied. One of their difference is that temporal scale is monthly for Thornthwaite whereas for Hargreaves, it can be daily or monthly. In the current study, hourly data was required as input to the UWBM. Hargreaves method was closer to that format. Hargreaves method is reliable and performs sufficiently in summer period compared to other methods such as Thornthwaite, and Priestley-Taylor based on the study of Nikam et al. (2014) in India. Thus, it was suitable for the case study in Athens where the temperatures can be high as well. The input data are minimum and maximum temperature, and the geographical latitude of the station. The formula is as follows:

$$ET_0 = 0.0023 R_a \left(\frac{T_{\max} + T_{\min}}{2} + 17.8 \right) \sqrt{T_{\max} - T_{\min}} \tag{D.1}$$

where:

ET_o: reference evaporation and transpiration (MJ/m^2/d), T_{min}: minimum temperature (^{o}C), T_{max}: maximum temperature (^{o}C), and R_a: daily extraterrestrial radiation (MJ/m^2/d)

R_a is dependent on the geographical latitude of the area and their relationship is negative (lower geographical locations receive more radiation). Based on Duffie and Beckman (1991), its formula is as follows:

$$R_a = \frac{1440}{\pi} G_s d_r \left[\omega_s \sin(\phi) \sin(\delta) + \sin(\omega_s) \cos(\phi) \cos(\delta) \right] \tag{D.2}$$

where:

G_s: solar constant (0.082 MJ/m^2/d), d_r: inverse relative distance from the Earth to the sun (-), δ: solar declination angle (rad), ω_s: sunset hour angle (rad), ϕ: well geographical latitude (rad), and J: number of day in calendar year (from 1 to 365 or 366)

$$d_r = 1 + 0.033 \cos\left(\frac{2\pi}{365} J\right) \tag{D.3}$$

$$\delta = 0.409 \sin\left(\frac{2\pi}{365} J - 1.39\right) \tag{D.4}$$

$$\omega_s = \arccos[-\tan(\phi)\tan(\delta)] \tag{D.5}$$

The coefficient 0.023 in Formula (D.1) is used to transform the extraterrestrial solar radiation at the top of the atmosphere to the radiation at the surface of the Earth. An older version of the formula included the global radiation at the surface of the Earth and not the extraterrestrial radiation but later the formula changed. The updated formula estimates the global radiation at the surface of the Earth considering the extraterrestrial radiation and possible sunshine hours based on the geographical latitude.

Appendix E

Analytical Hierarchy Process

E.1 Method

AHP was used for assigning weights to vulnerability indicators and to aggregated components. Eight drought experts, who worked either in public or private sector, filled in a questionnaire comparing indicators or vulnerability components in pairs. For each expert, a comparison matrix was created and and then based on all experts' matrix a combined one. Each element of the combined matrix was the geometric mean (see Formula E.1) of that element on the individual matrices. For the cases where there is high volatility, geometric mean provides more accurate results than arithmetic one.

More specifically regarding the comparison matrix, a scale of judgment [1-9] that represented how much one element is more important than another one was used (see Table E.1). Inverse values of those in the scale can be used in the symmetrical positions of the matrix elements. For instance, if the importance of indicator i to indicator j is extreme, then 9 is placed in position i, j of the matrix but in position j, i the inverse value was used (i.e. 1/9).

$$\left(\prod_{i=1}^{n} a_i\right)^{\frac{1}{n}} \tag{E.1}$$

where, n: number of matrices, a_i: comparison value between two indicators

Table E.1: AHP scale for pair comparison (Satty, 1980)

Intensity of Importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
3	Moderate importance	Experience and judgment slightly favor one activity over another
5	Essential importance	Experience and judgment strongly favor one activity over another
7	Very strong importance	An activity is strongly favored and its dominance demonstrated in practice
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation
2,4,6,8	Intermediate values	When compromise is needed

Next, indicators' weights were synthesized; the process includes the estimation of eigenvalues and eigenvectors but a simplified approach was used. This process applied not only on the combined pairwise comparison matrix but also on the individual ones. Regarding one matrix, the values in each column were summed and each element of the matrix was divided by its column sum. The output is called normalized pairwise matrix. The average for each row of the latter matrix represented the weight of each indicator; these weights are also called relative priority vector.

Consistency of the judgment is important when pairwise comparison matrix is filled in. For that reason, degree of consistency was estimated based on all judgments which each expert made. Firstly, each value of one column was multiplied by its column corresponding weight (this weight was determined before the consistency). Next, the weighted sum was estimated per row and was divided by its corresponding row weight. The average of the latter values was denoted as λ_{max} (this is a single value for all indicators). The following steps were the estimation of consistency index (CI) (Formula E.2) and consistency ratio (CR) (Formula E.3).

$$CI = \frac{\lambda_{max} - n}{n - 1} \tag{E.2}$$

$$CR = \frac{CI}{RI} \tag{E.3}$$

where: n is the number of indicators\criteria, RI is the random index and the other variables were mentioned above. Random Index is the the consistency index of randomly generated pairwise matrix and its value depends on the number of indicators (see Table E.2).

If consistency ratio is greater than 0.1, the judgment is inconsistent and the results are not reliable. In that case, experts changed some pairwise comparison values to increase consistency. The online tool created by Goepel (2018) for the weights' estimation was used to correct inconsistencies. Consistency was not only estimated for each individual matrix but also for the combined one. Another important point is that geometric mean ensures that the combined pairwise comparison matrix (of all experts) is consistent, given that pairwise comparison matrices of individual experts are consistent.

Table E.2: Random index of consistency regarding pair comparison in AHP method (Satty, 1980)

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
RI	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51	1.48	1.56	1.57	1.59

E.2 Estimated weights

E.2.1 Groundwater droughts

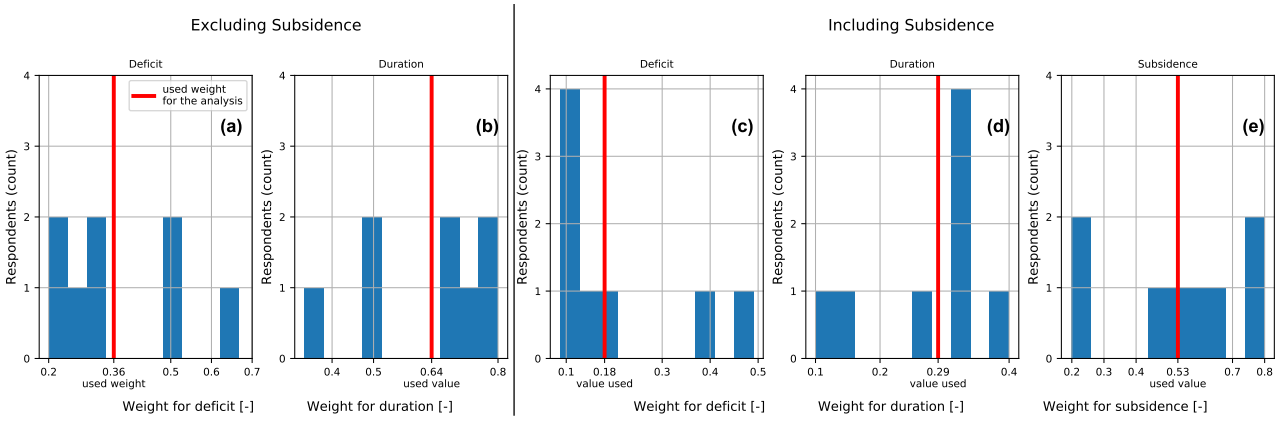


Figure E.1: Weight histogram for exposure indicators regarding vulnerability to groundwater droughts

Regarding the weights of deficit and duration when subsidence was excluded (see Figure E.1), there was no clear agreement among the drought experts. Their variations were large. On the contrary when subsidence was included, there was high agreement regarding the weights of deficit and duration but not for subsidence. The range of subsidence was from 0.2 to 0.8. This indicates that vulnerability results to groundwater droughts are highly uncertain when subsidence is included. This result supports the

decision of rejecting subsidence regarding vulnerability estimation. Even rejecting subsidence, there is still uncertainty regarding the weights of deficict and duration.

As for sensitivity indicators (see Figure E.2), historical monuments showed a large variation compared to the others. On the other hand, the lowest variation was observed for shops. As for the aggregated components (see Figure E.3), there is no clear agreement regarding the weights of (i) exposure aggregation (deficit versus duration), and (ii) sensitivity aggregation (physical versus social sensitivity aggregation). Their ranges are about 0.3.

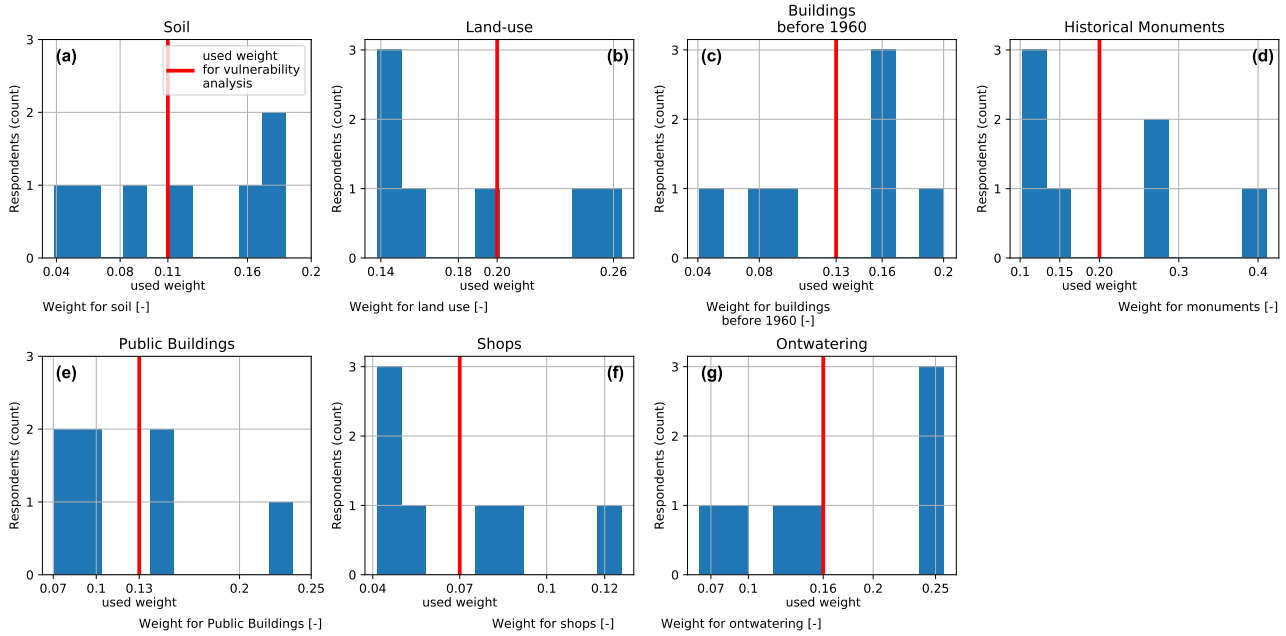


Figure E.2: Weight histogram for physical sensitivity indicators regarding vulnerability to groundwater droughts

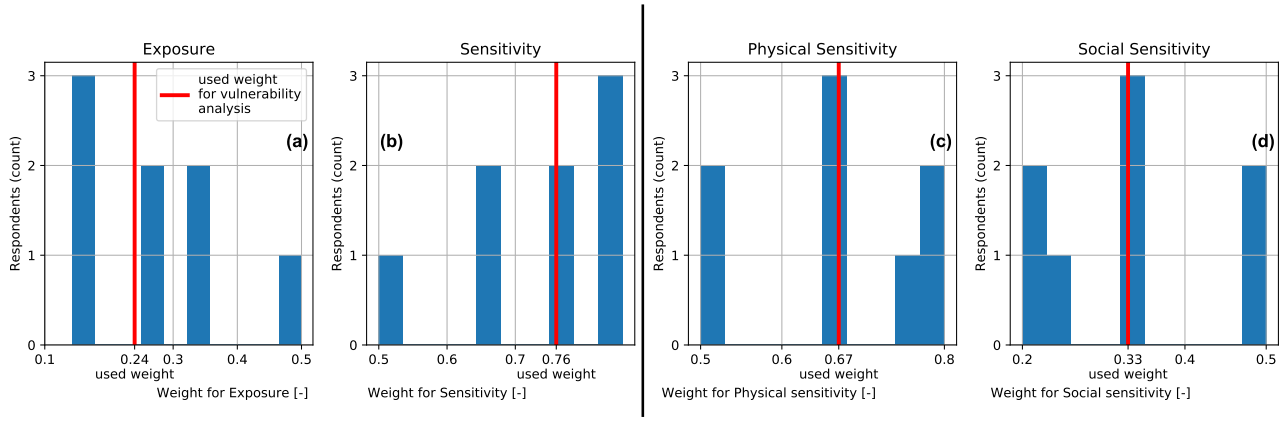


Figure E.3: Weight histogram for aggregated components regarding vulnerability to groundwater droughts

E.2.2 Soil Moisture droughts

Regarding soil moisture droughts, there is high uncertainty for the four sensitivity indicators (see Figure E.4); for most of them, the range of weight is 0.3. Large variation was noticed for aggregated components, too; it is more than 0.4 in all cases. As for the aggregated components, based on Figure E.5, the variation is larger for vulnerability aggregation (exposure versus sensitivity) than physical sensitivity aggregation for soil moisture droughts.

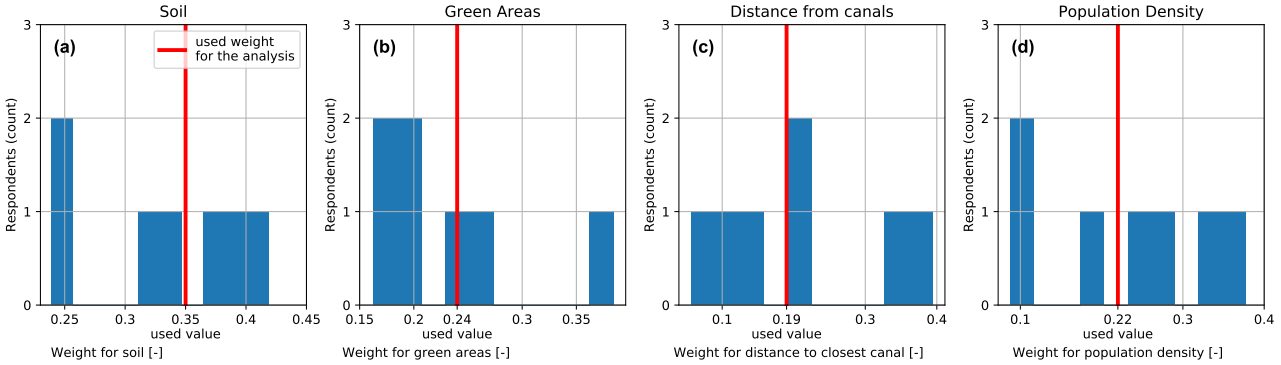


Figure E.4: Weight histogram for sensitivity indicators regarding vulnerability to soil moisture droughts

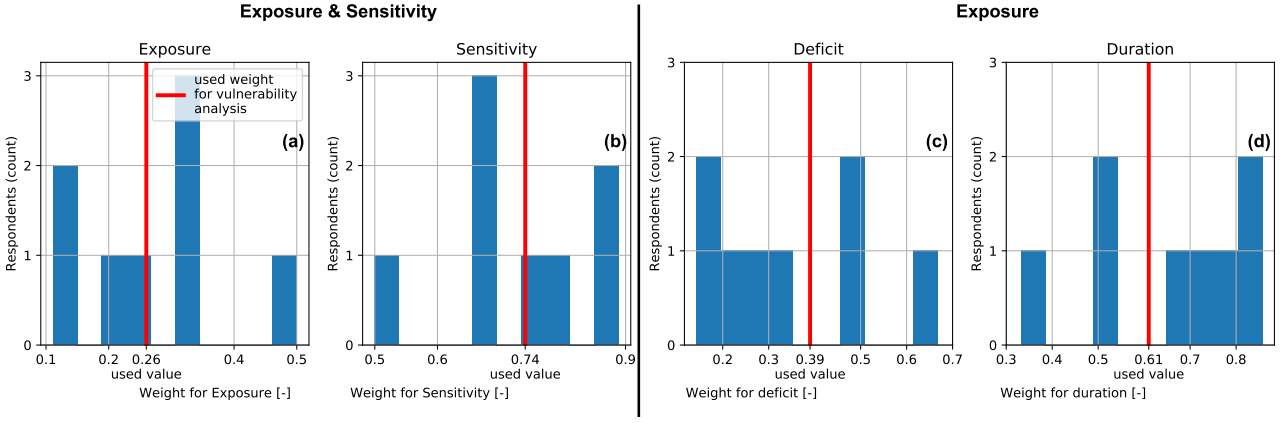


Figure E.5: Weight histogram for aggregated components regarding vulnerability to soil moisture droughts

Appendix F

Vulnerability

F.1 Groundwater Droughts

Coarse analysis

Districts located at the center of Leiden seemed to be more vulnerable to groundwater droughts than the rest ones. Both exposure and sensitivity were high for these districts. Given that sensitivity's weight was higher than exposure, vulnerability followed the spatial pattern of sensitivity. More specifically, district w_1 was the most exposed district but it did not hold for sensitivity and vulnerability where w_0 had the largest values. The aforementioned case included modeled subsidence for the estimation of exposure and vulnerability.

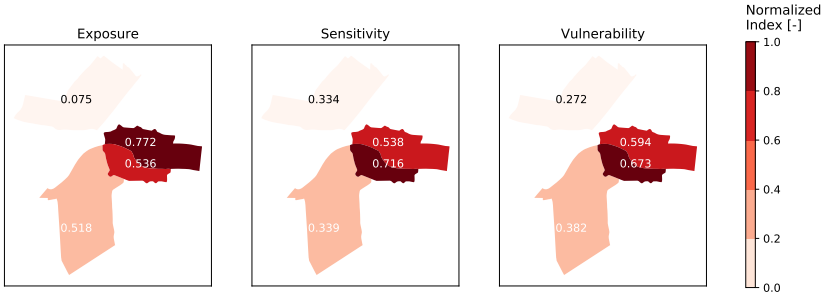


Figure F.1: Exposure, sensitivity, and vulnerability for groundwater droughts at district level using variable threshold of 30^{th} percentile for the estimation of exposure indicators (the unit is dimensionless for all cases)

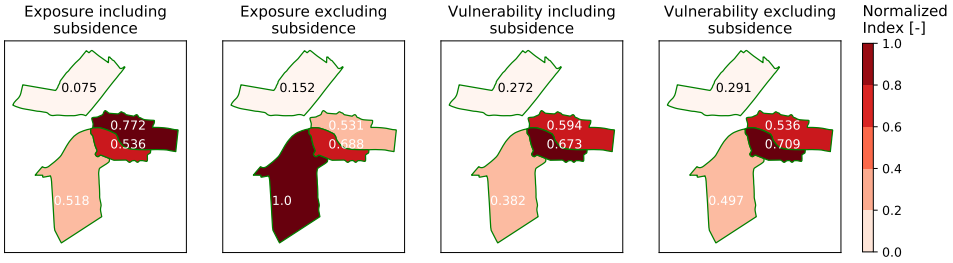


Figure F.2: Comparing exposure and vulnerability using (modeled) subsidence and not in exposure part (exposure was derived by variable threshold of 30^{th} percentile without pooling)

Vulnerability was calculated including subsidence and not (see Figure F.2). Excluding subsidence led to a different pattern in exposure but not in vulnerability; the latter is explained by higher sensitivity

weight than exposure for the estimation of vulnerability. More specifically, w_5 district was the most exposed for the case of excluding subsidence but w_1 including it. On the contrary, w_0 was the most vulnerable district including subsidence or not. Vulnerability values excluding subsidence were slightly larger for or all districts except w_1, though.

Sensitivity Soil Classes

Only the weights of the following soil classes: (i) deposits, (ii) tidal deposits, (iii) and trench deposits seemed to affect physical sensitivity but this influence smoothed to total sensitivity and vulnerability (see Figure F.3). The difference for 'deposits' class in physical sensitivity was noticeable only for the case of 0.4 increase. For the other two classes, all magnitudes of change (0.1, 0.2, and 0.4) led to slightly different results. The fact that (i) changes on the weights in two out of the five soil classes tested did not show variation in vulnerability, and (ii) variation of the rest classes was subtle indicate that all classes' weights may not be that significant. Therefore, spending resources for their estimation may not worth it; even applying a thorough method with drought experts would not change the vulnerability outcome noticeably.

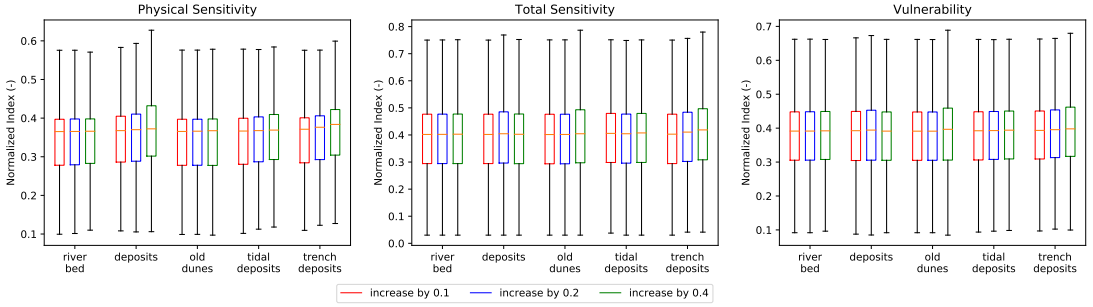


Figure F.3: Sensitivity for soil classes for the cases of increase by 0.1, 0.2, and 0.4 in each class (exposure was derived by variable threshold of 30$^{\text{th}}$ percentile with 10-day pooling)

Sensitivity Land use Classes

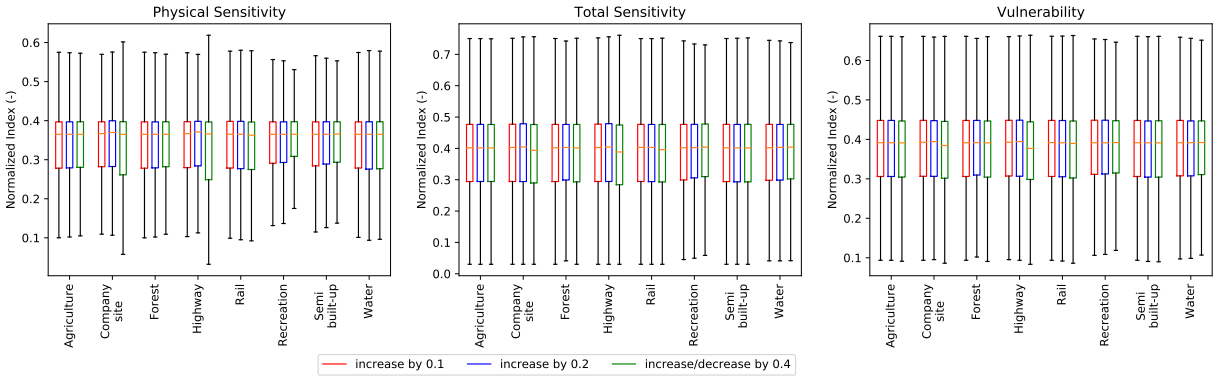


Figure F.4: Sensitivity for land use classes for the cases of increase by 0.1, 0.2, and 0.4* in each class (exposure was derived by variable threshold of 30$^{\text{th}}$ percentile with 10-day pooling). For the classes company site, highways, and railways, their percentages did not increase but decreased by 0.4 since their percentages could not increase more by 0.2 (their rankings were 0.8, 0.8, and 0.9 respectively)

Regarding sensitivity of land use classes, a similar pattern as that in soil was observed. The classes 'company site' and 'highway' showed the largest physical sensitivity variation for the case of 0.4 change. Another point is that recreation was the only class which led to reduced variation in physical sensitivity. There is no clear explanation about the latter one given that recreation class covers a limited area. For all classes, total sensitivity, and vulnerability were almost constant over all weight changes.

The weights of indicators determined by AHP played the most critical role to vulnerability aggregation. Therefore any change to them even without changing indicators' spatial pattern would alter the aforementioned outcomes for sensitivity regarding soil and land use classes. On top of that, how much area is covered by each class in soil or land-use certainly influenced the vulnerability results and their corresponding sensitivity.

F.2 Soil Moisture Droughts

Coarse Analysis

The most vulnerable district based on coarse analysis for soil moisture droughts differed from that for groundwater droughts. They were w_5 and w_0 for soil moisture and groundwater droughts respectively. It is also worth mentioning that the pattern of vulnerability changed slightly from sensitivity. This did not hold for the case of groundwater droughts in coarse resolution.

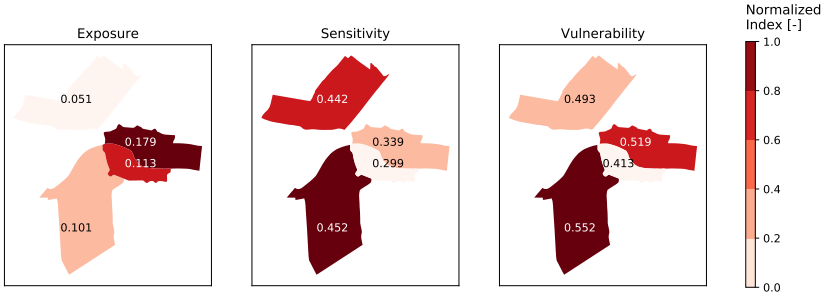


Figure F.5: Exposure, sensitivity, and vulnerability for soil moisture droughts at district level (fixed threshold and 40th percentile was used for the estimation of exposure indicators) (the unit is dimensionless for all cases)

Sensitivity Soil Classes

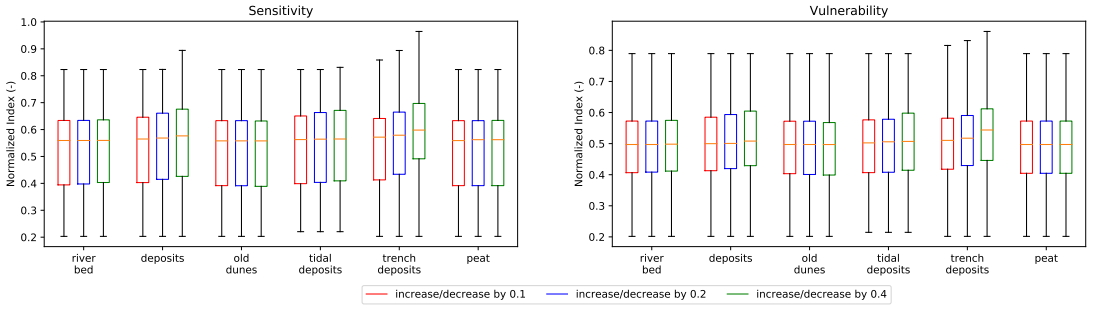


Figure F.6: Sensitivity for soil classes for the cases of increase/decrease by 0.1, 0.2, and 0.4 in each class regarding vulnerability to soil moisture droughts (exposure was derived by fixed threshold of 40th percentile without pooling)

The last part was sensitivity analysis of vulnerability based on the soil classes' weights (see Figure F.6). Trench deposits had the largest sensitivity (for the case of 0.4 increase/decrease) which was logical given that the majority of study area is covered by this soil class. This class showed large sensitivity for both exposure and vulnerability. On the contrary, for groundwater droughts, there was high smoothing from sensitivity to vulnerability for all classes. One reason is that for soil moisture analysis fewer indicators were used. Consequently, soil's weight was larger compared to the case of groundwater droughts vulnerability. For the rest soil classes in soil moisture analysis, changes to their weights seemed to affect marginally both sensitivity and vulnerability.

Key Results for groundwater and soil moisture droughts

- In coarse analysis there were slight differences between sensitivity and vulnerability patterns in soil moisture droughts. This did not hold for groundwater droughts.
- Changes in the weights of soil and land use classes regarding groundwater droughts did not affect vulnerability.
- 'Trench deposits' soil class showed high sensitivity for both sensitivity and vulnerability components when its weight changed by 0.4 regarding soil moisture droughts. The reason was that fewer indicators were used for vulnerability aggregation; therefore, soil had a larger weight affecting more both sensitivity and vulnerability.

Appendix G

Vulnerability results using extreme cases of the indicators' weights

G.1 Method

The weights used in this analysis are based on AHP. More specifically, the average (geometric mean) opinion of eight drought experts were used. Given that the weights affect considerably the vulnerability, cases of extreme weights were studied in this research, too. Extreme weights are considered those which differ the most from the average weights.

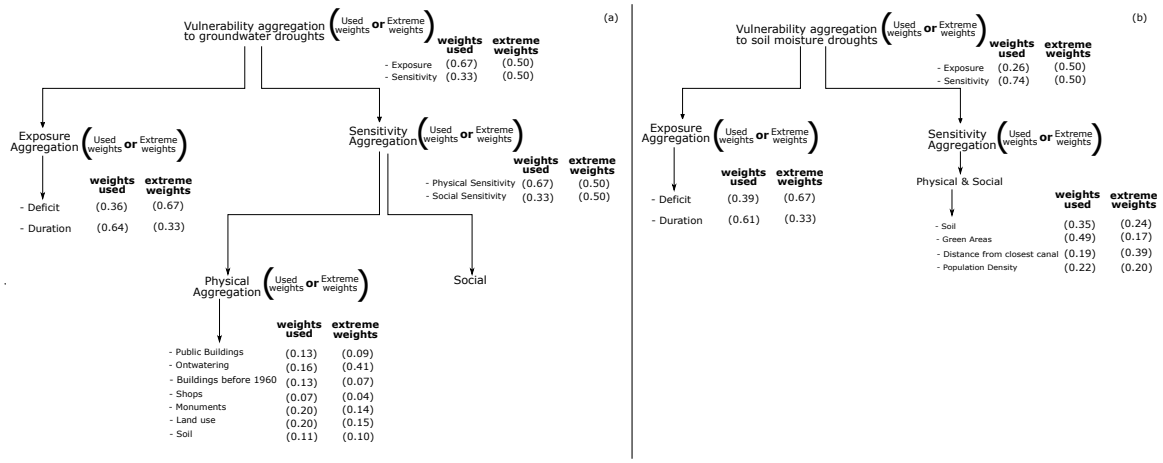


Figure G.1: The average indicators' weights ('used weights') which were used for vulnerability estimation and those ones ('extreme weights') which differed the most from the former ones. The latter ones ('extreme weights') are based on the opinion of one drought expert. This drought expert is not necessarily the same at all aggregation levels. The weights at all aggregation levels for both groundwater and soil moisture droughts are presented.

Two cases of extreme weights were applied. In the first one, for each indicator the two series of weights which led to its minimal and maximum weight value were used separately. To be more specific, for each indicator, the weights of two drought experts were used (one for the minimum and the other for the maximum value). These weights are not presented in Figure G.1

Regarding the second case, the approach used is more complicated. For each aggregation level, the weights of the drought expert which differed the most from the average weights were selected. The sum of square differences was used as a metric to determine the drought expert who assigned the most

'extreme' weights. In Figure G.1, the average and extreme weights are presented for both groundwater and soil moisture droughts. The average weights are annotated as 'weights used'. The following step was the estimation of all possible combinations of extreme and average weights at the aggregation levels. For the case of groundwater droughts, the number of combinations is 16 (2^4) since there are 2 options (Average VS Extreme weights) and 4 levels of aggregation ((i)Exposure VS Sensitivity, (ii) Deficit VS Duration, (iii) Physical Sensitivity Indicators, and (iv) Physical VS Social Sensitivity). All possible combinations regarding groundwater droughts are tabulated in Table G.1. Each case (row) in that table represents a different combination. It is worth highlighting that the first combination (Case 1) is the case for which the average weights were used at all aggregation levels; it is excluded from Table G.1.

Table G.1: All possible combination using extreme and average weights at different aggregation levels regarding groundwater droughts. The four aggregation levels used were: (i) Exposure VS Deficit, (ii) Deficit VS Duration, (iii) Physical Sensitivity Indicators, (iv) Physical VS Social Sensitivity Indicators; avg stand for average

	Exp_VS_Sens	Def_VS_Dur	Physic_sens indicators	Physic_VS_Social Sens_Indicators
Case-2	avg	avg	avg	extreme
Case-3	avg	avg	extreme	avg
Case-4	avg	avg	extreme	extreme
Case-5	avg	extreme	avg	avg
Case-6	avg	extreme	avg	extreme
Case-7	avg	extreme	extreme	avg
Case-8	avg	extreme	extreme	extreme
Case-9	extreme	avg	avg	avg
Case-10	extreme	avg	avg	extreme
Case-11	extreme	avg	extreme	avg
Case-12	extreme	avg	extreme	extreme
Case-13	extreme	extreme	avg	avg
Case-14	extreme	extreme	avg	extreme
Case-15	extreme	extreme	extreme	avg
Case-16	extreme	extreme	extreme	extreme

The combinations regarding soil moisture droughts are tabulated in Table G.2. The combinations are fewer compared to the case of groundwater droughts since the aggregation levels were fewer (three instead of four). The number of combinations is 8 (2^3) since there are 2 options (Average VS Extreme) and 3 levels of aggregation. Case 1 has been excluded from Table G.2 since it included the average weights at the three aggregation levels.

Table G.2: All possible combinations using extreme and average weights at different aggregation levels regarding soil moisture droughts. The three aggregation levels used were: (i) Deficit VS Duration, (ii) Deficit VS Duration regarding Exposure, and (iii) Physical Sensitivity Indicators

	Exp_VS_Sens	Exp	Physical_sens_indicators
Case 2	avg	avg	extreme
Case 3	avg	extreme	avg
Case 4	avg	extreme	extreme
Case 5	extreme	avg	avg
Case 6	extreme	avg	extreme
Case 7	extreme	extreme	avg
Case 8	extreme	extreme	extreme

Formula G.1 was used to determine the rate of change in vulnerability when the weights were changed. $Vulner_{using\ initial\ weights}$ is the vulnerability when the average weights of the droughts experts were used.

$$Rate\ of\ change = \frac{Vulner_{using\ changed\ weights} - Vulner_{using\ initial\ weights}}{Vulner_{using\ initial\ weights}} * 100 \qquad \text{(G.1)}$$

G.2 Results

G.2.1 Groundwater Droughts

Using the minimum and maximum weight for each indicator

- The differences in vulnerability were noticeable when minimum and maximum weights were used regarding the following physical sensitivity indicators: land-use, historical monuments, and ontwatering (see Figure G.2)
- Minor differences in vulnerability were observed when the minimum and maximum weights for the aggregated components were used.

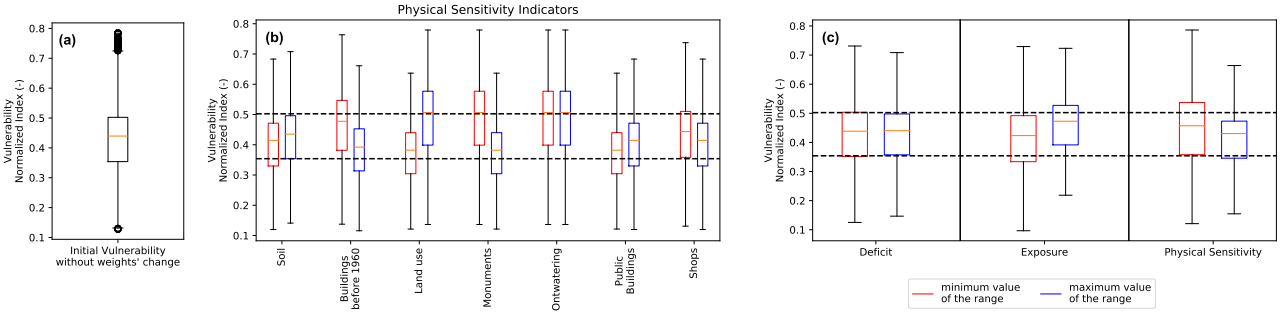


Figure G.2: Original vulnerability to groundwater droughts without changing indicators' weights (a); vulnerability when for each physical sensitivity indicator, its minimum and maximum weight value of its range were used; for each case, the weights of all indicators were changed (in that aggregation level). The weights used were based on the opinion of the drought expert which assigned the lowest or highest weight to that indicator (b); the same as sub-figure (b) but in this case the components are deficit, exposure, and physical sensitivity (c). Their complementary components (duration, sensitivity, and social sensitivity accordingly) were not estimated since they had the reverse boxplots) (e.g. the boxplot based on minimum weight for deficit would be the boxplot based on maximum weight for duration and vice versa). The same applies between (i) exposure and sensitivity, and (ii) between physical and social sensitivity (e.g. the boxplot based on minimum weight for exposure would be the boxplot based on maximum weight for sensitivity and vice versa)

Using the extreme weights compared to the average ones

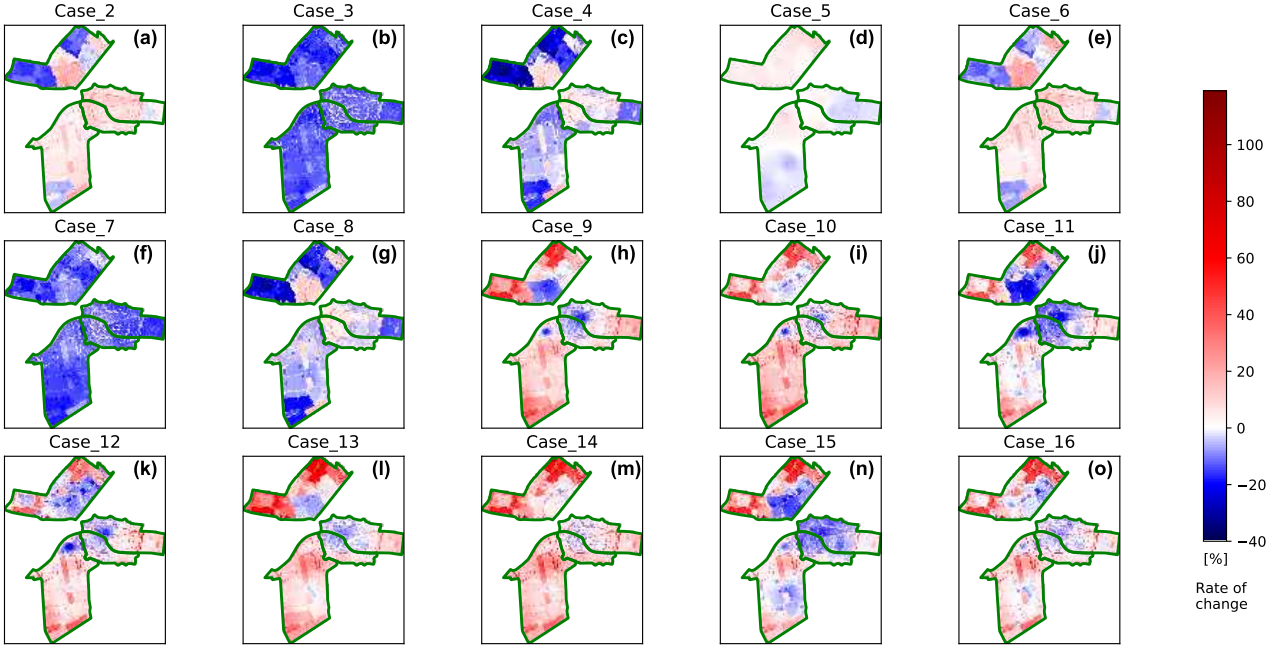


Figure G.3: Rate of change regarding vulnerability to groundwater droughts when 'extreme' indicators' weights were used at four aggregation levels (15 possible combinations were considered). The rate of change was estimated using the vulnerability with the average weights at all aggregation levels ($Vulner_{using\ initial\ weights}$) as reference.

It is interesting that for the case 9 (see Figure G.3-h)) and onwards, there are both positive and

negative rates of change over the study area. Backwards from case 9, the rates of change are mostly negative. This can be explained by the fact that for case 9 and onwards, the weights of aggregation 'Exposure VS Sensitivity' changed from the average values to extreme ones. Additionally, the variation of the case 9 and onwards decreased as can be seen in Figure G.4 (based on mix-max difference and not the interquantile range). The 'extreme' weights for exposure and sensitivity were 0.5, and 0.5 which smoothed the variation. Another worthwhile point is that the case 16, for which all aggregation levels used extreme weights, did not show the largest spatial variation but cases 11 and 15 did (Figure G.3-j,n).

The rate of change was high for most of the cases studied; it even exceeded 100% at some locations for three of the 15 cases. The fact that vulnerability changed considerably when "extreme" indicators' weights were used indicates that there is high uncertainty about the results when the weights are different. Therefore, no credible conclusions can be drawn regarding hot-spots of vulnerability.

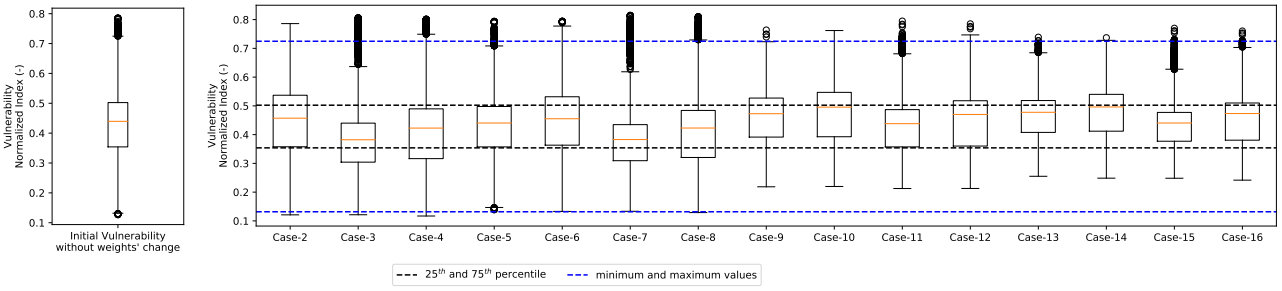


Figure G.4: Boxplots regarding vulnerability to groundwater droughts when 'extreme' indicators' weights were used at four aggregation levels (those weights differed the most from the average values) (all possible combinations were considered)

G.2.2 Soil Moisture droughts

Using the minimum and maximum weight for each indicator

There are differences in vulnerability variation when the minimum and maximum weight were used for 'green areas' and 'population density' indicators (see Figure G.2-b). On the contrary, the vulnerability changes were less pronounced when extreme values for the aggregated components' weights were used.

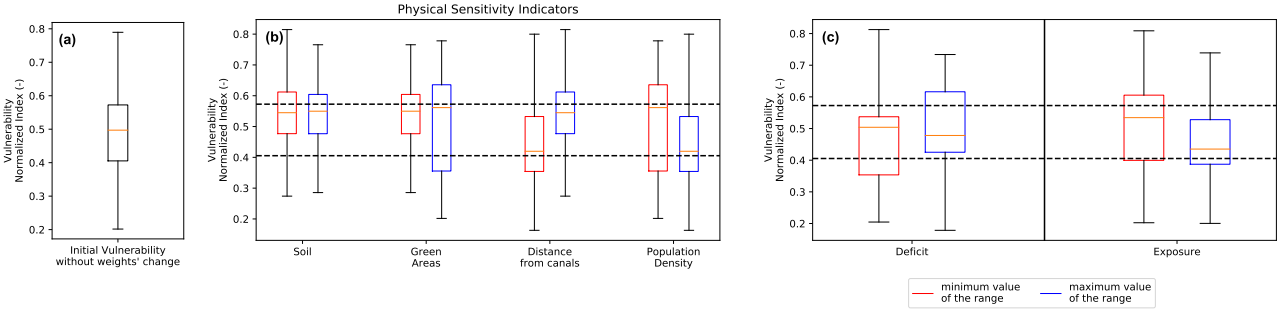


Figure G.5: Original vulnerability to groundwater droughts without changing indicators' weights (a); vulnerability when for each physical sensitivity indicator, its minimum and maximum weight value of its range was used. For each case, the weights of all indicators were changed (in that aggregation level). The weights used were based on the opinion of the drought expert which assigned the lowest or highest weight to that indicator (b); the same as sub-figure (b) but in this case the components are deficit, and exposure (c). Their complementary components (duration, and sensitivity accordingly) were not estimated since they had the reverse boxplots) (e.g. the boxplot based on minimum weight for deficit would be the boxplot based on maximum weight for duration and vice versa). The same applies between exposure and sensitivity (e.g. the boxplot based on minimum weight for exposure would be the boxplot based on maximum weight for sensitivity and vice versa)

Using the extreme weights compared to the average ones

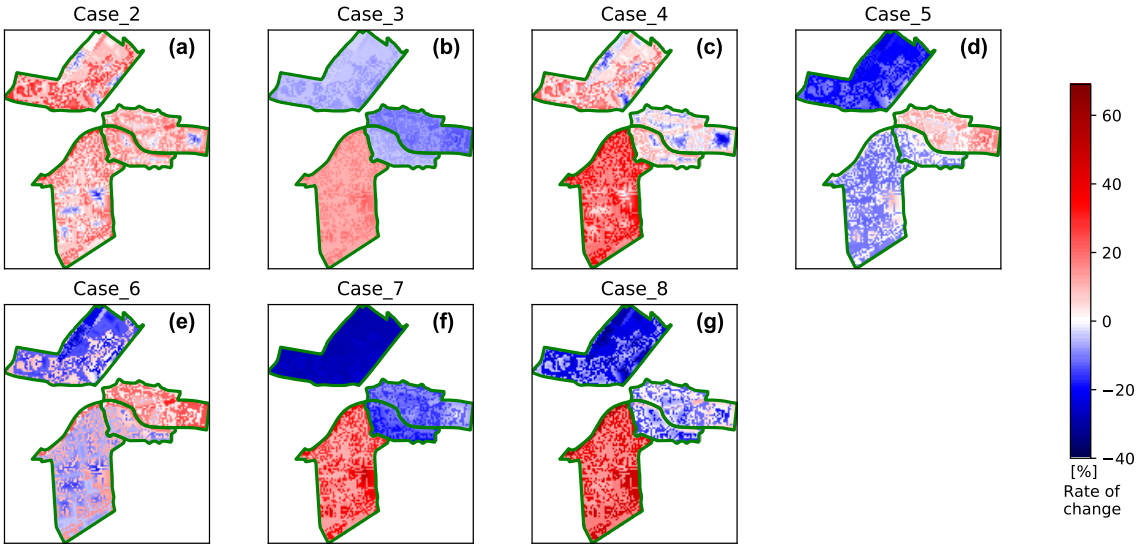


Figure G.6: Rate of change regarding vulnerability to soil moisture droughts when"extreme" indicators' weights were used at three aggregation levels (7 possible combinations were considered)

Figure G.6 presents the rate of vulnerability change when extreme weights at three aggregation levels were used. Seven different cases were tested. The rate of change had a lower range ([-40, 70]) for soil moisture droughts than groundwater ones ([-40, 110]). That implies that the results are more credible. Another point is that the variation of the vulnerability (based on the interquantile range) increased when the case 7 and 8 were used (see Figure G.7).

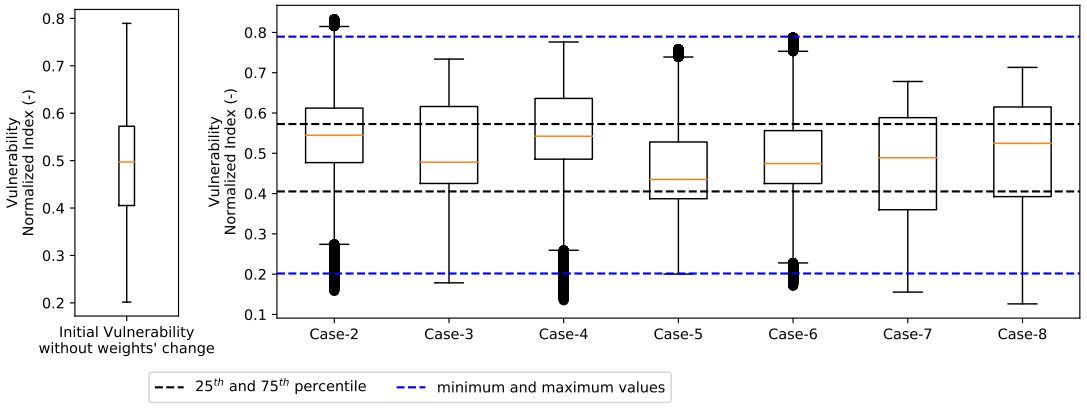


Figure G.7: Boxplots regarding vulnerability to soil moisture droughts when "extreme" indicators' weights were used at three aggregation levels (those weights differed the most from the average values) (all possible combinations were considered).

Key results

- Using "extreme" indicators' weights instead of the average ones lead to significant changes in vulnerability. This underlines that the results derived by the method suggested in this thesis are highly uncertain. The uncertainty is lower for soil moisture droughts.

Appendix H

Redistributing the priorities

H.1 Method

Decision makers may not have the same priorities regarding vulnerability. For that reason, the priorities were re-distributed (i.e. their weights were changed) and the analysis was re-applied. More specifically, the indicators' weights at aggregation levels were switched. More specifically, regarding groundwater droughts, the following components were switched: (i) Exposure and Sensitivity, (ii) Physical and Social Sensitivity, and (iii) both (i) and (ii) cases. Regarding soil moisture droughts, only one switch was applied this was for the weights of exposure and sensitivity. Formula G.1 was used to determine the rate of change in vulnerability when the weights were changed.

Regarding changing the priorities of physical sensitivity indicators, the following approach was followed. First of all, how the vulnerability changes was tested in case that all indicators of physical sensitivity had the same weights. Secondly, it was also tested the assumption that one indicator had a large weight and the other indications shared the rest weight. More specifically, regarding groundwater droughts, it was assumed that one indicator had 30% weight and the others shared the rest 70% equally. 30% was a random value; it could be larger. To elaborate this method, in Figure H.2-c, land use's weight was 30% and for the rest indicators their weights were 11.67% ($\frac{70\%}{6}$) whereas in figure H.2-d, monuments' weight was 30% and for the rest indicators their weights were 11.67%. Similarly, regarding soil droughts, one indicator had 50% weight and the others shared the rest 50% equally. 50% was used instead of 30% since the indicators were fewer. Formula G.1 was applied to determine the rate of change in vulnerability.

H.2 Results

H.2.1 Groundwater droughts

Switching the weights of the aggregated components

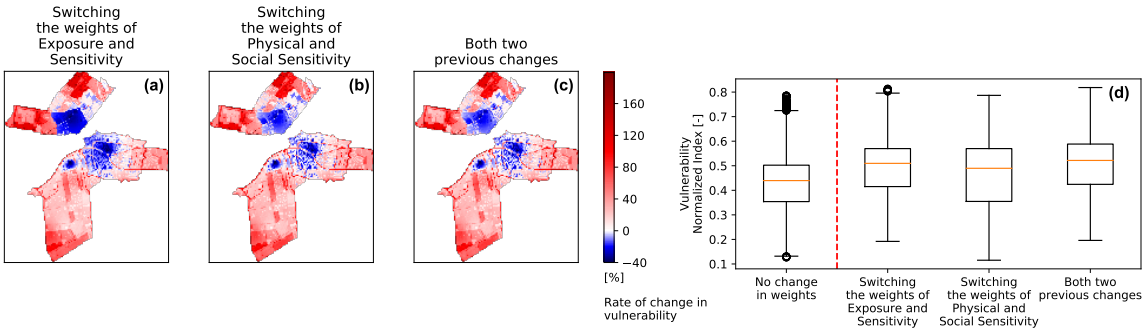


Figure H.1: Rate of vulnerability change (regarding groundwater droughts) over the study area for the three following cases: (i) switching the weights of exposure and sensitivity (subplot a), (ii) switching the weights of physical and social sensitivity (subplot b), and (iii) applying both (i) and (ii). The boxplots of vulnerability for these three cases are presented in subplot d.

Redistributing the weights of physical sensitivity indicators

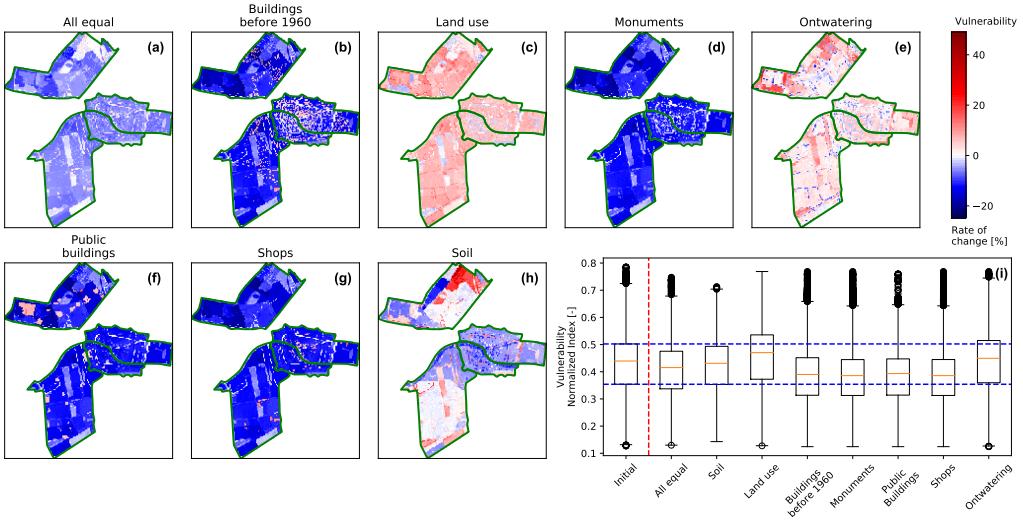


Figure H.2: Rate of vulnerability change to groundwater droughts redistributing the indicators' weights (a-h); for case (a), all indicators had equal weights. For each of the cases (b-h), the weight of one indicator was 30% and the others shared the rest 70% equally. 30% was a random value. For the estimation of the rate of change, the vulnerability with the average weights were used.

H.2.2 Soil Moisture Droughts

Switching the weights of the aggregated components

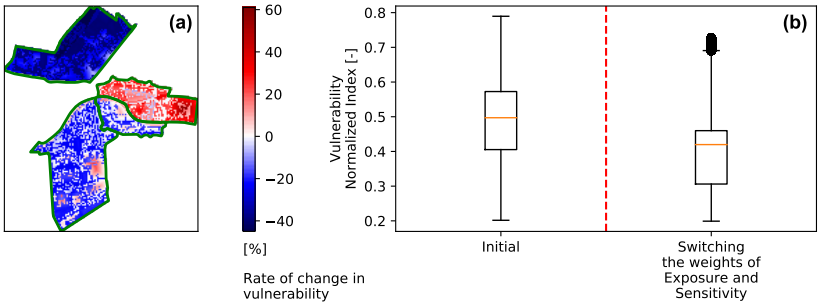


Figure H.3: Rate of vulnerability change (regarding soil moisture droughts) over the study study area for the case of switching the weights of exposure and sensitivity (a); boxplots of vulnerability (b)

Redistributing the weights of physical sensitivity indicators

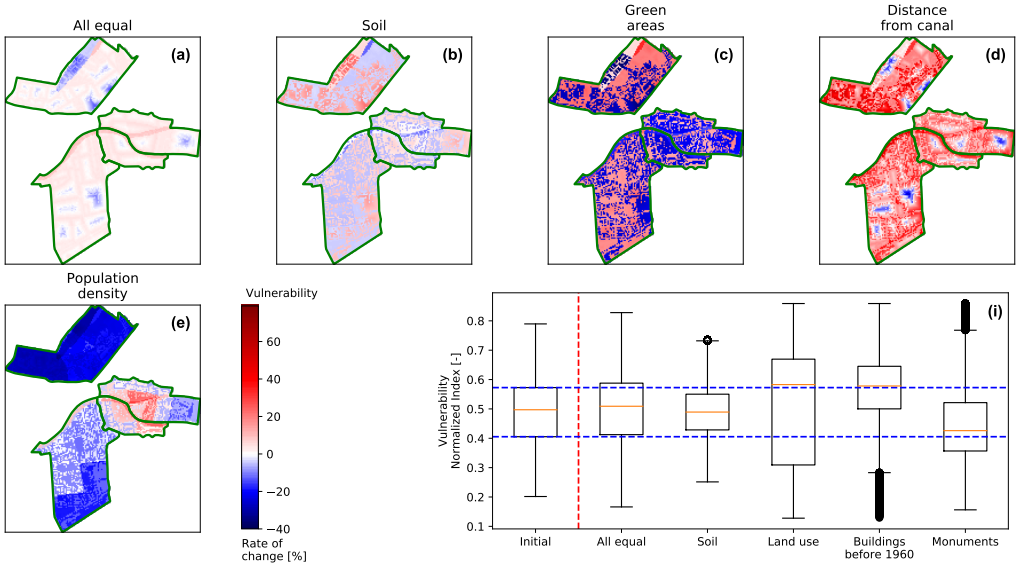


Figure H.4: Rate of vulnerability change to soil moisture droughts redistributing the indicators' weights (a-e); for case (a), all indicators had equal weights. For each of the cases (b-e), the weight of one indicator was 50% and the others shared the rest 50% equally. 50% was a random value. For the estimation of the rate of change, the vulnerability with the average weights were used.

Key results

- Switching the weights of aggregated components regarding groundwater droughts led to a distinct pattern; the vulnerability increased near the boundaries of the study area but decreased in the city center. This is explained by the fact that in the city center, sensitivity values are high whereas exposure ones are low or at least not that high as those in sensitivity.
- The majority of one indicator's normalized values in the study area affect its influence (positive or negative) on vulnerability when this indicator is prioritized (i.e. it weight increases). That holds for groundwater and soil moisture droughts. This result was noticed in sensitivity analysis, too.

Appendix I

Frequency Distribution Analysis

I.1 Groundwater Droughts

Frequency of droughts is another drought characteristic and was examined via modeling due to lack of long data in the current study. The upper part of Figure I.1 depicts the modeled GWL for the period 1983-2019. The most striking finding is that GWL did not decrease below a specific threshold for each district (e.g -0.8 m for w_0 district). This is quite irregular and indicates the mediocre performance of UWBM to represent low values based on calibration results.

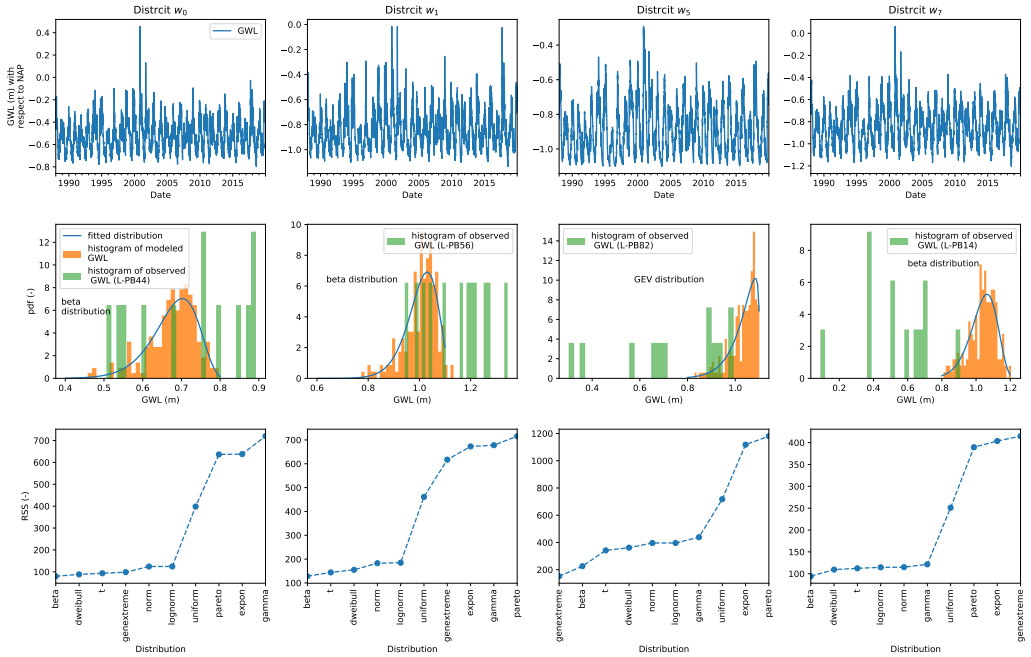


Figure I.1: Modeled GWL based on UWBM for all districts (1983-2019) (1st row); observed and modeled GWL histograms, and the best fitted distribution functions (2nd row); performance of each distribution out of the following ones (normal, weibull, beta, t, generalized extreme value (GEV), log-normal, uniform, pareto, exponential, and gamma) to fit the data (GWL) using RSS as metric (3rd row); (in the 2nd row, GWLs were converted to positive values so that fitting of distributions could be possible)

Modeled and measured GWL histograms did not follow the same pattern in all districts. In the second row of figures, the distribution with the best fit to modeled GWL histogram is depicted for each district. The distributions showed the lowest RSS were the following ones: beta (for w_0, w_1, and w_7), and General Extreme Value (GEV) (for w_5). The histograms of observed GWL were added to

highlight the pattern and not the values themselves. There was a mismatch in the histograms and one of the reasons were: (i) a different well than that in calibration was used since longer data was available (two years than one), and (ii) two years of measured data was not sufficient for the production of a histogram. For districts w_0 and w_5, the histogram based on measurements were left skewed which was in line with the histogram of modeled GWL. That was not verified in district w_1 nor w_7.

Drought frequency was not used as exposure indicator in the current analysis. A right skewed histogram was expected but none of the districts followed that pattern. Hence, even though minimum GWL was estimated for return periods of one, 10, and 20 years (using the best fitted distribution), these results were highly unreliable and were not used to the following step as an exposure indicator. These results are tabulated in Table I.1. The difference of this Table from Table I.2 is that yearly instead of monthly blocks were used.

Table I.1: GWL (m) (NAP) for return periods of one, 10, and 20 years for each district in Leiden based on the function of the distribution which fitted the data best using monthly block maxima method

	T=1 year	T=10 years	T=20 years
w_0	-0.735	-0.777	-0.784
w_1	-1.064	-1.107	-1.114
w_5	-0.855	-1.039	-1.050
w_7	-1.113	-1.173	-1.184

Table I.2: GWL (m) (NAP) for return periods of 10 and 20 years for each district in Leiden based on the function of the distribution which fitted the data best using yearly block maxima method

	T=10 years	T=20 years
w_0	-0.775	-0.783
w_1	-1.099	-1.109
w_5	-1.101	-1.110
w_7	-1.170	-1.183

Key results

- Modeled GWL followed an unexpected pattern and for half of the districts histograms based on measured and modeled GWL did not match. Therefore, drought frequency was rejected and were not used as exposure indicator to determine vulnerability.
- UWBM was not able to predict low values satisfactorily (one of the reasons was mediocre performance in calibration).

I.2 Soil Moisture Droughts

Drought frequency is another characteristic of soil moisture droughts which could be helpful to drought exposure estimation. Frequency distribution analysis was performed using block monthly minima and its results are presented in Figure I.2. More specifically, soil moisture dropped considerably in the summer of 2019 compared to the other years only in districts w_0 and w_7. Values of district w_5 were probably wrong given that minimum soil moisture decreased the same way every year. Soil moisture probably reached close to the permanent wilting point (13.4 mm); therefore, it could not decrease more. This is highly unlikely for areas in the Netherlands, though. Poor calibration performance in that district could be the reason for not representing reality satisfactorily.

Weibull distribution fitted soil moisture the best for three out of the four districts. The high performance of this distribution was expected. Gumbel for Minima, Weibull, Pearson Type III and Lognormal are the most frequently distributions for minima in hydrology (Granemann et al., 2018). The return periods for the minimum soil moisture for one, 10, and 20 years are attached in the Table I.3. Due to high uncertainty, drought frequency was not considered as an exposure indicator.

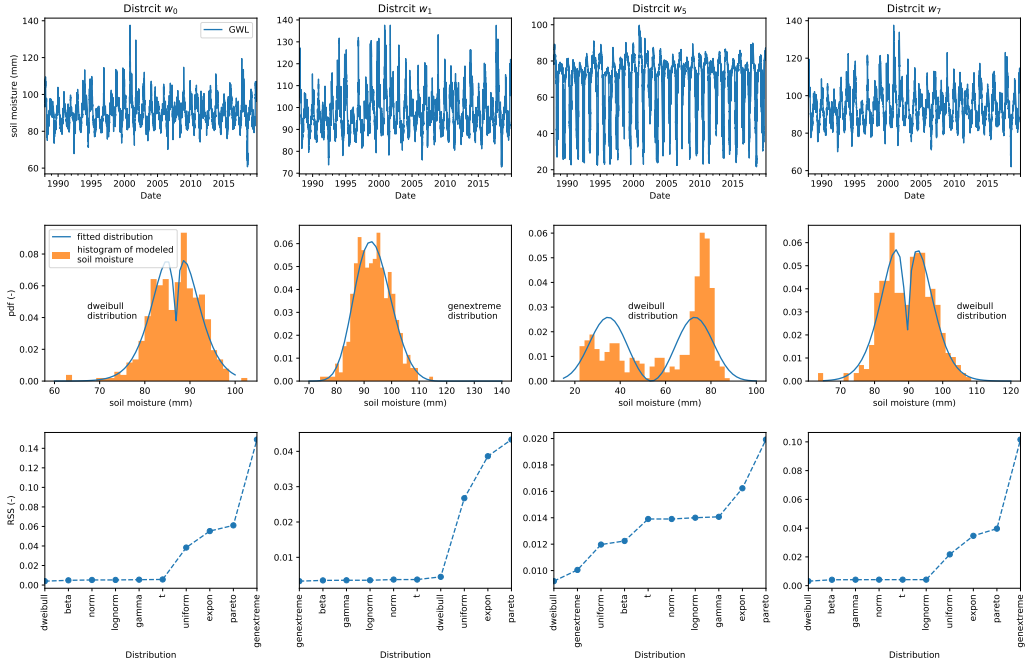


Figure I.2: Modeled soil moisture based on UWBM for all districts (1983-2019) (1st row); modeled soil moisture histograms, and their best fitted distribution functions (2nd row); performance of each distribution out of the following ones (normal, weibull, beta, t, generalized extreme value (GEV), log-normal, uniform, pareto, exponential, and gamma) to fit the soil moisture data using RSS as metric (3rd row)

Table I.3: Soil moisture (mm) for return periods of one, 10, and 20 years for each district in Leiden based on the function of the distribution which fitted the data best using monthly block minima method

	T=1 year	T=10 year	T=20 year
w_0	81.6	75.0	73.3
w_1	87.1	80.7	79.4
w_5	30.9	20.1	17.9
w_7	82.5	74.6	72.6

Key results

- In district w_5, an irregular pattern was noticed for soil moisture; this led to rejection of using the drought frequency to vulnerability estimation.

Appendix J

Glossary

This glossary defines some specific terms as author intends them to be interpreted in the context of this thesis.

Exposure
Exposure refers to the damaging event in itself. In the present case, it is a drought.

Sensitivity
It shows how the system changes due to increased exposure; this term is mostly related to the inherent characteristics of the system.

Coping Capacity
It is the ability of the system to manage adverse conditions based on its available resources.

Analytical Hierarchy Process (AHP)
It is a technique used in decision making when the environment is quite complex. It estimates the weights of the criteria and prioritizes the available alternatives. In the present research, using AHP, only the first part (calculating the weights of criteria) is considered.

Entropy evaluation method
A technique to estimate weights for a series of variables based on their variation.

Tipping point
A level of change in system properties beyond which a system reorganizes, often abruptly, and does not return to the initial state even if the drivers of the change are abated (Intergovermental Panel on Climate Change, 2014).

Coping capacity
Coping capacity is the ability of the system to manage adverse conditions based on its available resources. Coping capacity is considered as a short-term characteristic.

Adaptive capacity
Adaptive capacity is a characteristic like coping capacity but it is long-term and includes learning of past events. It is more about taking protection for future hazard events. It shows the flexibility of a system to adapt to new conditions (Fritzsche et al., 2014).

Vulnerability
Vulnerability is the degree to which a system is susceptible to, and unable to cope with, adverse effects of climate change, including climate variability and extremes based on Intergovermental Panel on Climate Change (2007)[p.27]