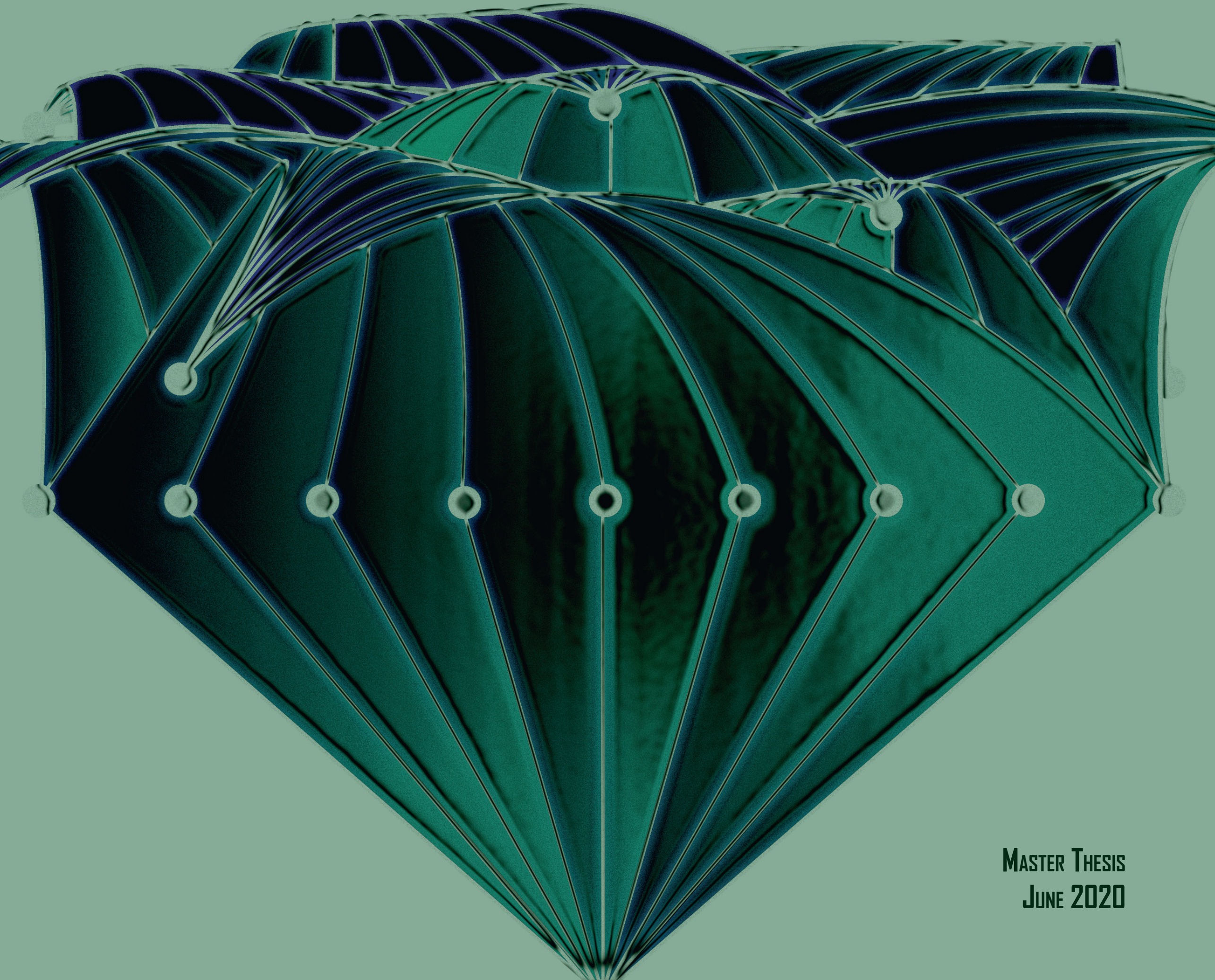
Drought Severity

Real-time evaluation of drought severity by means of
Artificial Neural Networks and damage functions

Mark Beltman



Master Thesis
June 2020

Drought Severity

Real-time evaluation of drought severity by means of
Artificial Neural Networks and Damage functions

Mark Beltman

COLOPHON

Student
M.R. Beltman
s1487795
Civil Engineering and Management

University
University of Twente
Drienerlolaan 5
7522 NB Enschede
The Netherlands

Involved organisation
Waterschap Vechtstromen
Kooikersweg 1
7609 PZ Almelo
The Netherlands

Supervision
University of Twente
Supervisor: Dr. Ir. M.J. Booij
2nd Supervisor: Prof. Dr. J.T.A. Bressers

Waterschap Vechtstromen
External supervisor: Ir. P.B. Worm

Date of publishing
June 17th 2020

Contact
mail: m.r.beltman@student.utwente.nl

PREFACE

This report belongs to a two-phased research project into the operationalisation of drought for the Vechtstromen Water Authority. A project that serves as a combined graduation project for the MSc. programmes Public Administration and Civil Engineering. This report is the second and last report produced in this context and serves specifically as graduation thesis for the master in Civil Engineering and Management, with a specialisation in Integrated Water Management. In the first phase of the project, a qualitative definition for the problem of drought from a water managing perspective has been formulated. The second phase, that is discussed in this report, focusses on the operationalisation of drought severity and builds upon the results from phase one. The work has been conducted under the supervision of the University of Twente and the Vechtstromen Water Authority.

Conducting this research would not have been possible without the help of many people. Therefore, I would like to express my sincere gratitude to anyone that helped in any way to making this project into a success. But there are a number of people I would like to thank specifically. Firstly, I would like to thank Bas Worm, who provided the opportunity and resources to work on this interesting topic and enabled me to conduct work that relates directly to the water managing practice. Also I would like to thank him for his open-mindedness towards my somewhat unorthodox approaches. I know I have been quite stubborn at times. Secondly I would like to thank my supervisors from the University of Twente, Martijn Booij and Hans Bressers. Your supervision provided a great support in putting my unconventional approaches into successful science. Finally, I want to express my appreciation for the support I received from Martin Mulder, the developer of the “Waterwijzer” Agriculture. Without even being involved in the project, he provided a lot of support in operating the “Waterwijzer”.

But it does not end here. Also in my private life I received an incredible amount of support that cannot stay unnoticed. Support that was there not only during my graduation, but during my whole study career. First I cannot thank my girlfriend Merlijn Smits enough for her tremendous support throughout the whole process and way before the process started. You were always there to discuss my thoughts no matter how exhausting your own working day had been. Never have you complained about me being distracted by my computer, training ANNs in the background, while we were watching one of our series. I truly admire your patience. Last but not least, the presentation of my work would not have looked so neatly without your help. Finally, I want to express my profound gratitude to my parents. They are the ones that have always encouraged me to discover who I am and what motivates me, both in my private as well as in my professional life. Without their unfailing support I would not have had the opportunity to study. Adding a second master’s program would have been even more impossible.

SUMMARY

As the climate changes and thereby the climatic extremes intensify, droughts occur more frequently. This holds also true for the Vechtstromen region in the Netherlands. To minimize the socio-economic drought impacts to the Vechtstromen region, adequate and effective crisis management is required. Yet, a lack of quick and reliable information regarding the socio-economic drought severity, limits the effectiveness of the crisis response in mitigating societal impacts. Instead, crisis management is based upon solely hydrological drought indicators, like precipitation deficits and surface water levels, that are far from linearly related with the water use impacts. To improve drought management in the Vechtstromen region, a quick and easy real-time evaluation of the socio-economic drought severity is, therefore, desirable.

Recently two tools have been developed that enable to evaluate the socio-economic impacts of hydrological conditions quick and easily: the "Waterwijzer" Agriculture and the "Waterwijzer" Nature. Applying these tools to evaluate drought severity in real-time is, however, limited by a lack of groundwater data. Only point measurements are available, while real-time spatial groundwater patterns are required. From a literature study it was found that Artificial Neural Networks (ANNs) are likely the best way to interpolate the point measurements into spatial groundwater patterns with sufficient accuracy and speed. This research, therefore, aims to operationalize the socio-economic drought severity in real time, by using Artificial Neural Networks to obtain daily spatial groundwater data as an input for drought impact models. For this it has been studied if and how accurate ANNs can interpolate groundwater depths and if this accuracy is sufficient for drought severity evaluation.

To study the ability of ANNs to accurately interpolate groundwater depths, two experiments have been setup: one in which the Vechtstromen region is interpolated by a single ANN and one in which two regional ANNs are used. This because the water systems of the northern and the southern region function differently. The northern region is predominantly a surface water controlled system, while the southern region is a free draining system. All three ANNs have been optimized individually by finding the optimal combination of input variables and number of hidden neurons. Their interpolation accuracy has subsequently been determined by testing the ANNs for an independent dataset that consisted of locations that were not used during model training and validation.

From these experiments it is found that ANNs provide spatial groundwater depths with higher accuracy than the currently available alternatives that require longer calculation times. This conclusion holds true regardless of the type of hydrological system the interpolation relates to. The second major finding was that, although ANNs can cope with different types of hydrological systems separately, ANNs are not well able to distinguish between different functioning systems in a single ANN. Yet, despite this limitation also the single ANN, trained to interpolate

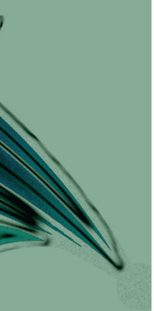
the full Vechtstromen region by one model, outperformed the traditional methods.

With these promising interpolation results, all elements to evaluate drought severity are in place. In the second research step, it has been studied if combining these elements results in sufficiently reliable severity evaluations, with a special focus on the effects of the uncertainty in the groundwater data to the severity evaluation. For this the socio-economic severity of 2019's drought in the Vechtstromen catchment area has been evaluated (in a code green, yellow or red) at 72 drought sensitive locations. These evaluations have been performed for both the upper and the lower confidence limits of the groundwater depth predictions, to see how the uncertainty affects the severity evaluation. This study revealed that for none of the locations the difference between the upper and lower confidence limit was more than one colour code. Even more, at 58 locations the colour code evaluation was consistent. For five locations, located at the eastern Twente moraine, the plausibility of the severity evaluation is, however, questioned as here the ANN provides too shallow groundwater depths. Yet, these plausibility issues are not expected to affect the difference in severity evaluation between the upper and lower confidence level. Therefore, despite these plausibility issues, it is concluded that the groundwater depth predictions are sufficiently accurate to reliably evaluate socio-economic severity.

With some minor improvements to the ANN for the eastern Twente moraine, the severity evaluation as presented in this report forms a solid basis to improve drought management. Nonetheless, there are also opportunities for further optimizations. Firstly, the informative strength of the severity evaluation to the drought management decision making process, can be enhanced when the severity evaluation links more closely to the qualitative drought severity definition, that is formulated in the first phase of this research project. For this more knowledge is required on the operationalisation of the qualitative drought severity definition in quantitative severity limits. Also for nature there needs to be found a way to separate the natural drought impacts from the human induced drought impacts. A second opportunity lies in providing drought severity predictions instead of evaluations. This will enable water managers to proactively mitigate drought severity. To enable severity predictions it is possible to combine the presented drought severity evaluation method with temporal groundwater depth predictions. The latter can be effectively done by ANNs.

All in all, it can be concluded that the combination of ANNs and damage models holds a lot of potential to evaluate, or even predict, drought severity quickly and easily. Water managers are, therefore, advised to further develop and explore the application of ANNs to operationalise drought severity. This will help them to manage droughts more effectively by putting more focus on their core responsibility: facilitating water use.

CONTENT



1

INTRODUCTION

BACKGROUND

For centuries the Dutch delta mostly had one water related problem, there was too much of it. To get rid of the water surplus the Dutch have built an ingenious system of pumps and dikes to keep their land and polders dry. But while improving and mastering this system towards perfection drought problems have intensified (Bressers et al., 2016; Tielrooij et al., 2000). This because the discharging practice was hardly limited by the drought problems that might occur on the other side of the water managing spectrum. For long the relevance of drought was underestimated, the country was believed to be water abundant.

But as global temperatures rise and thereby the climatic extremes intensify new and more severe drought problems occur (Trenberth, 2011). This also holds for water abundant North Western European countries like the Netherlands. This led the Dutch water managers to see that water management should focus more on balancing the water system between floods and droughts, instead of solely discharging water surpluses (Ritzema & Van loon-Steensma, 2017).

To manage droughts both on seasonal and structural time scale, and to be able to balance its impact to that of flooding, it must be known where and when droughts occur and how severe they are. This is currently not fully understood. The hydrological conditions are to certain extent known. But whether these conditions need to be considered as drought is understood limitedly. A drought dashboard that indicates for any given moment how severe the hydrological conditions are, can help the water authority to manage their droughts. In the short term it can provide information on where to take direct action. For the long term it provides insights in the spatial variation in drought vulnerability. This information can be used to design more structural drought preventing measures.

Drought severity, an indication for how extreme the drought conditions are, can be defined in two distinct ways, either statistically or in societal terms. Statistical definitions tend to define the drought severity relative to normal water conditions. Societal definitions define the drought severity in relation to the societal impact it causes. As regional water management is largely about enhancing society by facilitating water use, balancing floods and droughts is about weighing the impacts of floods and droughts to society. To do so, a society focused drought operationalization provides most valuable information.

The overall aim of this research, that comprises two phases, is, therefore, to obtain a real time insight in the societal severity of a drought. Here real-time insights are important to be able to adequately manage drought crises. Also the real time insights provide interesting insights for more structural drought management interventions. From an early literature review, that has been discussed in research phase one of this project, it became clear that two steps were required for such operationalization of drought. First the problem of drought needed to be defined from a water managing perspective. Thereafter a way to assess the hydrological conditions for their societal impact needs to be found.

The first step has already been performed in research phase one (Beltman, 2019). This second research phase focusses on the second step, the question of evaluating the hydrological conditions for their socio-economic effects.

RESEARCH GAP

To understand what knowledge gap withholds the assessment of hydrological conditions for their socio-economic impact, a literature study has been performed (Beltman, 2020) of which the conclusions will be summarized here. If you want access to the full report, you can email the author. This study focused on three aspects: (1) the conceptual relation between the hydrological system and the socio-economic response, (2) the way in which this conceptual relation can be operationalized and (3) the availability of data for this operationalization.

The literature review concluded that the relation between hydrology and society can best be conceptualized via soil moisture. This because this research focusses on land use related drought. In this conceptualization the “Waterwijzer” Nature (Witte et al., 2018) and the “Waterwijzer” Landbouw (Mulder et al., 2018) can be used to translate hydrological conditions to socio-economic effects. For this the Waterwijzer first models the soil moisture in the unsaturated zone, based upon groundwater levels, climate data and geological characteristics. Subsequently the calculated soil moisture conditions are related to damages to crops or nature. The main advantage of these tools is that they are designed for the modern Dutch agricultural and nature management context. Besides, they require relatively few computational power which makes them interesting for real time application. Therefore, the state of the art knowledge is believed to be sufficiently developed to translate hydrological conditions to the relevant socio-economic impacts.

Research gaps are, however, found in the input data that is required to run the “waterwijzer” tools for a real-time assessment. This because the waterwijzers require spatial groundwater patterns, which are not available in real time. Only point measurements, obtained by wells, are. To map spatial groundwater patterns, complex and time consuming groundwater models need to be used. Due to their complexity these models are not desirable to use for real-time purposes. Hence, to use the Waterwijzer’s potential to translate the real-time hydrological conditions in socio-economic terms, the spatial groundwater levels need to be available in real time more easily.

As the only groundwater information that is available in real time are the well measurements, literature has been studied to understand if there are possibilities to interpolate these point measurements to obtain spatial groundwater data. This study showed that traditional techniques are likely unable to interpolate accurately because they assume to some extent spatial linearity (Davis & Sampson, 1986). From the literature study, the most interesting option to interpolate the groundwater levels seems to be by means of using Artificial Neural Networks (ANNs). These have already proven their potential towards temporal predictions of groundwater levels (Chitsazan, Rahmani, & Neyamadpour, 2015; Daliakopoulos, Coulibaly, & Tsanis, 2005; Mohanty, Jha, Kumar, & Sudheer, 2010; Nayak, Rao, & Sudheer, 2006; Yoon, Jun, Hyun, Bae, & Lee, 2011) and have been used in other contexts for nonlinear interpolation of irregular spatial variables (Chowdhury, Alouani, & Hossain, 2010; Nourani, Mogaddam, & Nadiri, 2008; Rigol, Jarvis, & Stuart, 2001; Sun, Kang, Li, & Zhang, 2009). It is, therefore, likely that they are able to spatially interpolate groundwater depths sufficiently accurate.

RESEARCH OBJECTIVE AND QUESTIONS

Objective

Water managers benefit from a dashboard that evaluates the socio-economic drought severity in real-time (on a daily basis). This because it will enable them to improve their crisis response and their structural interventions to the water system. This real-time drought severity evaluation is mostly limited by the lack of spatial groundwater depth data. This second phase of the drought operationalization project therefore aims to:

Operationalize the land use related socio-economic drought severity in real time, by using Artificial Neural Networks to obtain daily *spatial groundwater data as an input for drought impact models.*

By operationalization it is meant to evaluate and define the severity of the hydrological drought conditions to the Water Authority in a way that it becomes meaningful to the water managing practice.

The relevant land use related socio-economic impacts are the impacts that are identified as problematic to the water authority. This relates to the results of the first research phase. Here the relevant impacts and the point at which they become problematic to the water authority are defined. For agriculture it are mostly the economic costs and the losses in nutritional values that are relevant indicators. Economic costs become problematic when there is a risk of large scale bankruptcy among agricultural enterprises due to the drought conditions. Reductions in nutritional values become problematic when they make reaching the legal self-sufficiency norms impossible. Regarding nature it are human induced impacts that are problematic to the water authority. These drought indicator and levels will be further elaborated in chapter, four.

Research questions

The research objective comprises two main elements, the interpolation of groundwater well data and defining the socio-economic impact that corresponds to the resulting hydrological conditions. For each of these two themes a research question has been defined:

1. *How accurate can ANNs spatially interpolate groundwater depths, based upon static spatial variables in combination with a limited number of reference groundwater depths?*

2. *Are the ANN interpolated groundwater depths sufficiently accurate to evaluate socio-economic drought severity with damage models?*

READING GUIDE

This report contains five chapters, of which the first is this introduction. The chapters are structured in a way that the scientifically relevant information and the practical water management information can be read separately.

The scientifically relevant information is predominantly provided in chapter two and three. Chapter two is written as an independently readable paper that discusses the design and testing of the ANNs to interpolate groundwater depths. The third chapter is also written in paper form. Here it is studied if the accuracy of the ANNs is sufficient to evaluate socio-economic drought severity. From these two chapters, one will obtain detailed insights in the methodology, results and conclusions regarding the individual research questions.

Water managers who are mostly interested to know if and how they should operationalize drought in terms of socio-economic severity, are advised to read chapter one, four and five. Chapter four will put the main conclusions of chapter two and three in a more practical water managing perspective. Based upon this practical perspective an advice is formulated regarding the usability of this drought operationalization and future steps are discussed to improve the practical usefulness of the operationalization. Finally, the main conclusions to the two research questions and the central objective, are summarized in the concluding chapter, chapter five.

2

SPATIAL INTERPOLATION OF GROUNDWATER DEPTHS WITH ARTIFICIAL NEURAL NETWORKS IN AN IRREGULAR CATCHMENT AREA

ABSTRACT

During water crises, like droughts, access to quick and reliable spatial groundwater level data is crucial to effectively mitigate socio-economic impacts. The currently used numerical groundwater models are, however, not able to quickly produce this data. The objective of this research is, therefore, to study whether reliable spatial groundwater data can be provided by using ANNs to interpolate well measurements, with a particular focus on non-linear catchment areas. For this a case study for the Vechtstromen catchment area is performed. Two experiments have been setup: one in which the region is interpolated by a single ANN and one in which two regional ANNs are used to separately interpolate the differently functioning water systems, a free draining system and a surface water controlled one, that are present in the study area. All three ANNs have been optimized individually by finding the optimal combination of input variables, learning epochs and number of hidden neurons. Their interpolation accuracy has subsequently been determined by testing the ANNs for an independent dataset that has not been used during model training and validation. From these experiments it is found that ANNs provide spatial groundwater depths with a higher accuracy than the currently available nummerical alternatives. This conclusion holds true regardless of the type of hydrological system the interpolation relates to. The second major finding was that, although ANNs can cope with different types of hydrological systems separately, ANNs are not well able to distinguish between different functioning systems in a single ANN. Yet also this single Vechtstromen ANN outperformed the traditional methods. Based upon these results, water managers are advised to start exploring the use of ANNs to provide real-time groundwater depth information during water crises.

INTRODUCTION

To evaluate and mitigate the socio-economic impacts of a drought in crises situations, it is important for water managers to have access to quick and reliable hydrological data. Herein, insights in spatial groundwater depth patterns are especially relevant, as socio-economic drought impacts are predominantly land bounded, like the damages to agricultural yields. Nevertheless, it is precisely this spatial groundwater data that is not easily available during crisis situations. This because spatial groundwater data is currently produced by complex numerical models that require relatively long computation times and relatively much human effort. To improve drought management it is, therefore, necessary to find a more easy way to obtain reliable spatial groundwater data.

An alternative approach to obtain this spatial groundwater data more quickly is to interpolate well measurements. These well measurements are often already available in real time and interpolation techniques require relatively short computation times. Yet, traditional interpolation techniques, like Inverse Distance Weighting or Kriging, are limitedly able to cope with the strong spatial variations that are present in many regions around the world, like that in the Netherlands. These catchments simply have too strongly varying abiotic conditions, are too heavily modified by the human being and they often have too complex geological characteristics. Thereby the groundwater levels are not likely to be spatially linear or second order stationary, which are assumptions that underly respectively Inverse Distance Weighting and Kriging (Davis & Sampson, 1986). An alternative interpolation method that can cope with spatial nonlinearity needs thus to be found.

The existing body of research on Artificial Neural Networks (ANNs) suggests that ANNs might bring a solution to the problem of non-linear groundwater level interpolation. In prior research ANNs have already been used for groundwater level interpolation in a relatively homogenous catchment area in Iran (Nourani et al., 2008) and China (Sun et al., 2009). Here, however, the ANNs have not been provided with spatial characteristics in addition to the groundwater dataset to improve the groundwater level prediction. Thereby, the ability of ANNs to combine multiple intercorrelated types of data, which is expected to be necessary in less homogenous catchment areas, was not exploited. This ability to combine a variety of datatypes to improve interpolation of non-linear patterns has already been demonstrated in adjacent research fields, like for example to spatially interpolate temperature data in a complex environment (Rigol et al., 2001) and in relation to interpolation of groundwater pollution. In the latter application ANNs outperformed Ordinary Kriging significantly because of the non-linearity in the contamination pattern (Chowdhury et al., 2010).

The aim of this research is, therefore, to study if ANNs are able to reliably interpolate groundwater depth measurements in catchments with spatially highly varying characteristics. For this, the Vechtstromen catchment area located in the Netherlands will be used as a case study.

The first section of this paper will further motivate the choice for this specific study area. Thereafter, the methodology will be discussed. Herein, first the research strategy will be elaborated and then the methodology to obtain the ANNs will be explained. The third section presents the results of the study. Here the optimal ANNs and their performances are presented. Section four of this paper, the discussion, delves into the methodological limitations, the comparison to other literature and explores the potential use and generalisation of the results. The fifth and last section concludes the paper by providing an answer to the central research question.

STUDY AREA

To study the ability of ANNs to interpolate groundwater well measurements a case study is performed for the Vechtstromen catchment area, see Figure 2.1. This area has two interesting characteristics. Firstly it contains two hydrologically different types of water systems, that of the Twente region, covering the southern halve of the Vechtstromen region and that of the Drenthe region covering the Northern half. Secondly, agriculture and nature are strongly interwoven in the Vechtstromen region. This provides an interesting interplay between natural and human induced effects on the hydrological cycle and vice versa.

The Twente region is predominantly shaped by moraines as a result of glacial deposits and has, therefore, relatively large elevation differences. Due to these moraines the hydrological system is rather a collection of little free-draining watersheds, that unite downstream in Twente, than a single connected system. The moraines also cause the geology to be highly complex. The soil types range from fine sand, to boulder clay, to peat and the aquifer thickness strongly varies throughout the region. The fragmented watershed combined with the complex geology cause the Twente region to be highly heterogeneous. This is expected to makes interpolation of the groundwater depths challenging.

The Drenthe region is a relatively flat region with a less complex geology. Here soil types and aquifer thickness are not as fragmented as in Twente. Also the water system is predominantly controlled by the surface water levels and is not a freely draining system like Twente. This because of the relatively flat landscape, the low elevation relative to the outflow point, and the dominant influence of multiple rivers and canals on the groundwater levels. Due to the more consistent geology and the flat surface water controlled water system, the Drenthe region is believed to be far more homogeneous than the Twente region. The spatial correlation between the groundwater levels is, therefore, expected to be relatively strong.

Having these two differently functioning hydrological systems in one study area allows to study the impact of this difference on the ANNs ability to interpolate groundwater depths. Here it is expected that ANNs perform better for relatively homogeneous surface water controlled systems, like Drenthe, than for heterogeneous free draining ones, like in Twente. The differently functioning systems also allow to study whether ANNs are able to cope with different types of hydrological behaviour in a single ANN, or if ANNs should only be applied to regions with similar functioning water systems.

Besides this difference in hydrology, the Vechtstromen region is also an interesting case study due to its strongly interwoven mixture of agriculture and nature in the rural areas and because of its heavily human modified hydrological system. Thereby, ANNs are also tested for their ability to deal with both natural and human factors that influence spatial groundwater level variability.

Since February 2019, the Vechtstromen region has a dense net of 187 wells at which groundwater depths are monitored in the rural areas. Data collection for these measurement locations is outsourced to a commercial business, that collects and filters the data (Wareco, 2020). Measurements are thus already adjusted for errors and noise. The time series that are collected by the 187 wells range between a full year and three quarters of a year.

A downside of the selected study area is that the data is only collected for

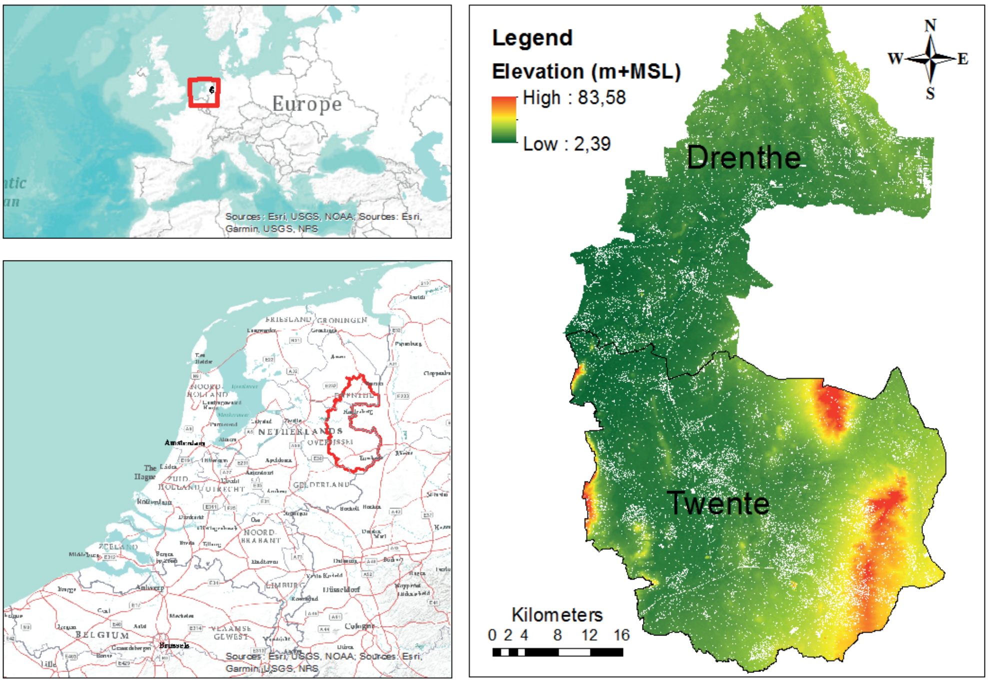


Figure 2.1: Elevation map of the Vechtstromen region with respect to the mean sea level (MSL)and indication of the catchment location in Europe and the Netherlands

a single year, 2019. As this was a year of substantial drought it suffices for this research objective. Nevertheless, one must be careful as the model can also be trained too much to this specific season and thereby not be applicable to other drought scenarios. Let alone to normal or wet conditions. The impact of this limited time span that is covered by the data will be discussed in the discussion chapter.

METHODOLOGY

To investigate the ability of ANNs to interpolate groundwater levels and to understand the influence of the hydrological functioning of a system on this interpolation ability, two experiments have been set up: one in which a single ANN is constructed for the full Vechtstromen region and one in which separate ANNs are constructed specifically for the Drenthe and Twente region. Subsequently the performance of the Vechtstromen model will be compared to the performance of the regional models. When the general Vechtstromen model performs at least equally well as the regional models, this proves that ANNs are able to cope with hydrologically differently functioning systems in a single dataset.

The model performance will be evaluated by the Kling Gupta efficiency (KGE) (Gupta, Kling, Yilmaz, & Martinez, 2009). This Kling Gupta efficiency is chosen as it accounts for three important performance indicators: the ANN's ability to describe the variation in the data, its ability to describe the average value and the extent to which the predictions correlate to the observed values.

This methodological section describes how the three ANNs are defined and how the results are compared. It will firstly explain what general ANN design has been taken from literature. Secondly, the relevant input and output variables and data will be derived from literature and the data will be prepared to fit the ANNs functioning. Hereafter, the methodology continues with discussing how the optimal sets of input variables and the optimal number of hidden neurons are defined for the three ANNs. Finally it concludes with stressing how the performance of the obtained optimal ANNs is studied in more detail to understand the usability and limitations of the ANNs.

Literature based ANN design

The general ANN designs, are based upon literature. A brief literature study, in which nine papers regarding ANN use have been studied, showed that in adjacent research topics, the ANN designs were to a large extent similar. Four of these papers related to spatial interpolation, of for example groundwater pollution or temperature data (Chowdhury et al., 2010; Nourani et al., 2008; Rigol et al., 2001; Sun et al., 2009), and five related to temporal groundwater interpolation by ANNs (Chitsazan et al., 2015; Daliakopoulos et al., 2005; Mohanty et al., 2010; Nayak et al., 2006; Yoon et al., 2011). All nine papers concluded that a standard feedforward backpropagation model (see Figure 2.2), a Levenberg Marquart learning strategy and a Sigmoid activation function provided accurate results for the modelling objective. The only design parameters for which the papers differed were the input and output variables, the number of hidden neurons and the learning epochs (the number of training iterations with the same dataset). This research will, therefore, build upon the same ANN model, learning strategy and activation function. The number of hidden neurons, the input and output variables and the number of learning epochs are customized.

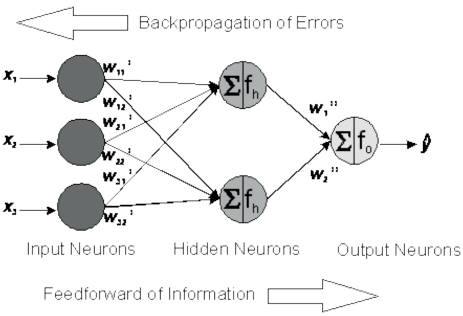


Figure 2.2. Feedforward backpropagation ANN model

Desired output unit

The goal of this study is to interpolate groundwater levels as a basis for the evaluation of socio-economic impacts. As it is the depth to groundwater that affects the soil moisture content and thereby is the most important factor to land use related effects, like the growth of vegetation (Remmelink, Blanken, van Middelkoop, Ouweltjes, & Wemmenhove, 2018), the depth to groundwater is the most meaningful way to define the groundwater levels. To obtain this depth to groundwater, all groundwater level measurements, which are measured relative to mean sea level (MSL), are subtracted from the ground elevation level of the wells.

Input variables and data

Spatial variability in groundwater depths is naturally determined by three predominant factors: climate, geology and topography (Condon & Maxwell, 2015; Devito et al., 2005; Freeze & Witherspoon, 1967; Haitjema & Mitchell-Bruker, 2005; Salvucci & Entekhabi, 1995; Tóth, 1970; Wolock, Winter, & McMahon, 2004). Besides also human interference has an effect on the groundwater level, by land use (Genxu, Lingyuan, Lin, & Kubota, 2005; Scanlon, Reedy, Stonestrom, Prudic, & Dennehy, 2005), water subtractions (Hoque, Hoque, & Ahmed, 2007) and surface water drainage (Bouwman, 1998). Each of these factors are included as an input to the ANN model.

The climatological conditions are reflected by including reference groundwater levels. These reference levels are the most direct reflection of the relevant recent climatic and hydrological history. The reference groundwater wells are determined by a correlation study between all groundwater wells. From these correlations, an optimal set of 3 reference wells has been derived. Herein the optimum was defined based upon two criteria: the number of strong correlations (>0.7) and the length of the combined measurements series. The latter is important because training, validation and testing samples can only be generated for moments at which all reference wells recorded a groundwater level. A set of three reference wells, shown in Figure 2.3, is found as the most optimal representation of the Vechtstromen area. These wells have a strong correlation with 92% of the other wells and their overlapping lengths cover 98,8% of the total timeseries.

In relation to geology it is predominantly the transmissivity that influences the groundwater level variations among different geological conditions. The transmissivity is, therefore, included as a variable that reflects the dependency of groundwater levels on geology. The transmissivity map for the Vechtstromen region is obtained from the BOFEK 2012 maps produced by Alterra (Wosten et al., 2013). The BOFEK maps are an extensively used soil data source in the Netherlands.

The topography is directly incorporated as an input variable by including the elevation relative to MSL. For this a raster dataset with a spatial resolution of 25 meters has been applied (AHN, 2019).

Land use has an effect on the local evaporation rates and thereby on the groundwater recharge and levels (Scanlon et al., 2005). The types of vegetation predominantly typify a specific land use and define the evaporation rate. (Beltman & Koerselman, 1998; Droogers, 2009). Therefore, to account for land use impacts, the Makkink reference evaporation factors of the dominant vegetation type per land use are used as a spatial variable. These reference evaporation factors are linked to a 2016 land use map for the Vechtstromen region. The reference evaporations are mostly taken from literature (Beltman & Koerselham, 1998; Droogers, 2009; Jansen, 1995; Moors, et al. 1996). However, for some hybrid land use types, like grassy heather fields, the reference evaporation is estimated based upon the values for related land use types, from which the hybrid land use is a combination.

Surface water drainage is included by adding the distance to drainage canals as an input variable, instead of the surface water levels. This because the inclusion of more temporal variables introduces a significant burden on water managers to collect reliable real time data. The distance to drainage systems is expected to be the best stationary drainage representation.

The effect of drinking water abstractions is included by including its impact on the average lowest groundwater level. Here the average lowest groundwater level, is calculated by averaging the three lowest groundwater levels in a year (of a two weakly measured time series) for eight years in a row. This is again a static variable. Here the same argumentation regarding collection of real time data holds as for the surface water drainage. The impact of abstractions on the average lowest groundwater levels is only known for the Twente region and thus not included in the Drenthe model.

Finally, two additional input variables have been added to this theoretically obtained set: the average lowest groundwater level and a classification for the ability to supply the specific location with water from downstream (value 1 if this is possible and value 0 if not). The ability to let water in from downstream locations is a characteristic that differs strongly over the Vechtstromen region. In almost the complete Drenthe region this supply of downstream water is possible, contrary to Twente where it is mostly impossible due to the elevation differences. During droughts the Twente regions will therefore be sooner confronted with water shortages. The second additional variable, the average lowest groundwater level, serves as an additional reference level to indicate the spatial variability in groundwater levels. Unlike the other spatial variables it does not reflect a physical process influencing the groundwater level.

Data preparation

All the data that is discussed so far is prepared for three reasons: to assure the correctness of the data, to shape the data in a way that it best suits the range of the ANN's activation functions and to obtain a reliable and representative training, validation and testing data set. For this three steps have been performed.

First, groundwater measurements that were equal to the depth of the measurement well, have been taken out of the dataset. These records are unreliable as it is possible that the actual groundwater depth was larger, but could not be recorded due to the limited depth of the well itself. In these situations the well sends its own depth as measured value.

Secondly, the temporal well measurements and the spatial variable maps have been scaled between -0.5 and 0.5. This scaling of the variables is useful as it best suits the ANN's sigmoidal activation function. Between this range the sigmoidal function has a relatively steep slope, that enhances the activation ability of the neurons. When values become too large or small the activation function does not differentiate as much anymore in its output signal. This diminishes the ability to selectively activate the neurons.

Finally, training, validation and testing datasets are constructed, for the Drenthe and Twente regions. For this, first data samples have been constructed in which each groundwater depth measurement is coupled to the spatial characteristics of the specific well location and to the three corresponding reference groundwater depths. Subsequently, these data samples have been divided in a training, validation and testing set with a size of respectively 70%, 15% and 15% of the complete dataset.

The testing dataset is formed by samples that relate to a set of 10 wells in Twente (this is about 15% of the total number of Twente wells) and 20 wells in Drenthe, (this is about 15% of the total number of wells located in Drenthe). These wells are not used in the training and validation phase. Thereby, they are completely independent and represent locations that are unknown to the ANN. To assure the representativeness of the testing set and to prevent a bias in the model, it has been made sure that the test set contains the second largest and second smallest values of each spatial input variable for each region. This led to the selection of 10 wells in Twente and of 9 wells in Drenthe. The remaining 11 wells for the Drenthe region are randomly selected.

For the remaining wells all samples

are constructed and stored in a single dataset per region. 18% of these regional datasets is selected randomly for validation purposes. The remaining 82% is used as training dataset. The resulting distribution of testing and training/validation wells is shown in Figure 2.3.

For the Vechtstromen model the regional datasets have been combined. The Vechtstromen model is thus trained, validated and tested by both the Drenthe and Twente datasets.

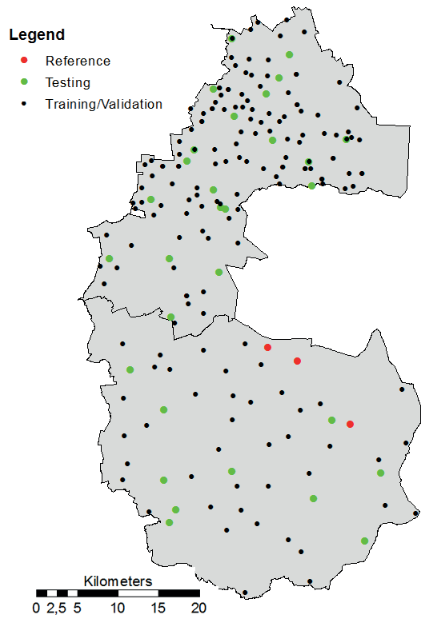


Figure 2.3. Groundwater well locations and their function for the construction of the ANNs.

Searching for candidate models

The ANN models, for both experiments, are defined in two phases. First candidate models are obtained by an explorative study. Afterwards better performing models near the candidate solutions are searched for by a neighbourhood analysis.

To define candidate models an extensive searching run has been performed in which the interpolation performances of 500 unique randomly generated model configurations have been calculated. Herein, the input variables, the number of hidden neurons and the number of learning epochs are all randomly selected.

For the input variables a set of one till ten input variables is randomly selected for the Vechtstromen and Twente model. For the Drenthe model a maximum of nine input variables is studied, as there is no abstraction data available for this region. This variation in the set of input variables provides an indication of the actual relevance of the theoretically defined input variables. The number of hidden neurons has been randomly set between 2 and 6. Test runs showed that more hidden neurons failed to obtain a generalized model that can cope with unknown locations. Finally, the maximum number of epochs is randomly set between 50 and 150 epochs, with a step size of 50 epochs. This because too many epochs might also overtrain the network, with consequential generalization problems.

Each randomly generated model has been assessed on its ability to predict groundwater depths at known and unknown locations. For this, each model is validated with the validation set, and tested with only 10% of the data for each location in the testing set. This ten percent is used to still have unused data samples for further model testing. Models that score a KGE above

0,85 are considered candidate models worth it to optimize. This threshold is chosen based upon explorative model runs. From these runs it appeared that the maximum performance was around this KGE. This rough approach is not problematic as in the end we are interested in the best scoring model. The other candidates only provide additional information on the importance of input variables.

Neighbourhood search for better performing models

For the best scoring candidate models per region, a neighbourhood analysis is performed to check if there is a better performing model design close to the candidate solution. For this the number of hidden neurons and the maximum number of learning epochs has been systematically varied. The number of hidden neurons has been gradually increased from 2 till 15. The number of epochs ranged between 50 and 300 epochs, with a step size of 50 epochs. For each combination of hidden neurons and epochs, 10 training iterations have been performed. The best training score, highest KGE, from these 10 iterations is considered the optimal ANN performance for that specific ANN configuration.

Finally, the optimal ANN design is the model with the highest average KGE (averaging the KGE scores for validation and testing), with a maximum difference between validation and testing of 0,05. This difference is included to prevent a bias towards any of the datasets.

Analysing the optimal models

The groundwater depths that are produced by the three optimal ANN's, that result from the neighbourhood analysis, are further studied to better understand the interpolation ability of ANNs. For this the performance, in terms of KGE value, for each individual well, as predicted by both the Vechtstromen and the regional models, is calculated. These insights in the individual performances provide more information regarding the performance differences between the Vechtstromen model and the two regional models.

Results

Candidate models

The models presented in Table 2.1, are the candidate models that resulted from the three extensive randomized runs. What is striking to see is that for the Twente and Vechtstromen region only two and four candidate models have been found respectively, contrary to a set of 23 candidates for the Drenthe region. Clearly the Drenthe dataset comprises more explanatory power than the Vechtstromen and Twente sets. Thereby, the interpolation accuracy is less dependent on the model structure.

What also stands out is that most candidate models use all the input variables that are identified by literature. Even though a wide variety of reduced combinations is tested. This confirms the importance of the spatial characteristics that are stressed in literature. But even though all input variables are used, most variables are not indispensable. In all three models there are candidate models that leave out some variables. Even more, the Drenthe model can still perform when any of the variables, not being the average lowest groundwater level, are left out. This proves the importance of the average lowest groundwater level for the interpolation. It also proves that the explanatory power of the input data lies in the combination of multiple relevant spatial characteristics, not necessarily in a single one. For the Twente and Vechtstromen model this cannot be concluded, as there are multiple variables that are present in all candidates. Here there is thus also a dependency on individual spatial characteristics.

Lastly it stands out that all possible numbers of hidden neurons show up in the candidate models for Drenthe, with four hidden neurons as dominant number. Also the Twente and Vechtstromen candidates differ in the number of hidden neurons that are used. Thereby, it can be concluded that the issue of overfitting is not solely a matter of picking the right number of hidden neurons. It is rather an interplay between the number of hidden neurons, the number of input variables and the number of learning epochs.

Neighbourhood search

The results of the neighbourhood study, for the three ANNs, are presented in Figure 2.4. This figure shows that all models score high KGE values for the validation dataset and keep improving when the number of hidden neurons increase. When 15 hidden neurons are applied, all three ANNs are able

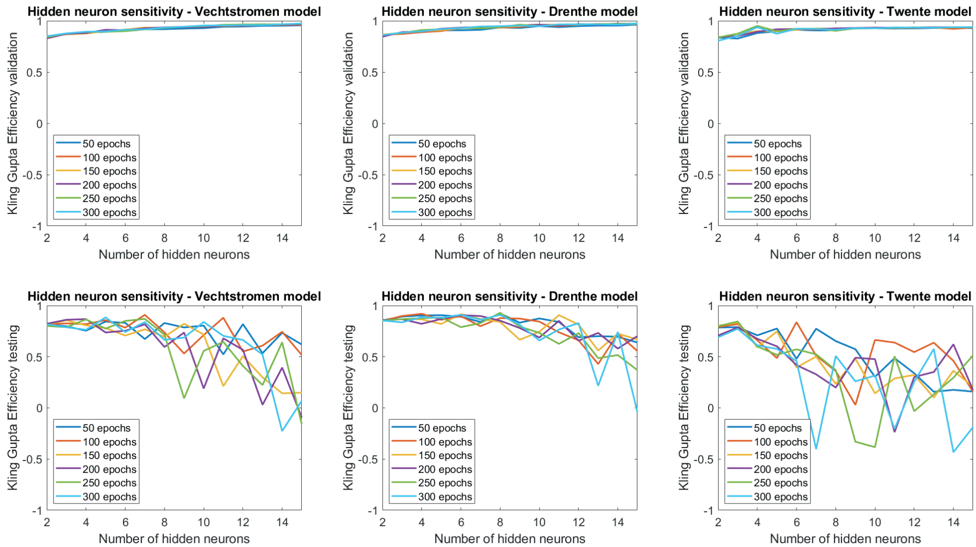


Figure 2.4: ANN sensitivity to hidden neurons and number of epochs plotted for the testing and validation set for the Vechtstromen and the regional ANNs.

Table 2.1: Variable use and corresponding KGE for the candidate Vechtstromen and regional ANNs

	Vechtstromen				Twente								
	1	2	3	4	1	2	1	2	3	4	5	6	7
Reference well 1	x	x	x	x	x		x		x	x	x	x	x
Reference well 2	x	x	x	x	x		x	x	x		x	x	x
Reference well 3			x		x	x	x	x	x	x	x	x	x
Elevation	x	x	x	x	x	x	x	x	x	x	x	x	x
Transmissivity	x	x	x		x	x	x	x	x	x	x	x	x
Evaporation	x		x	x	x		x	x	x	x	x	x	x
Distance to drainage	x		x	x	x		x	x	x	x	x	x	x
Abstraction impacts	x	x	x	x	x	x							
External supply	x	x	x	x	x	x	x	x	x	x	x	x	
Avgerage Lowest Depth	x	x	x	x	x	x	x	x	x	x	x	x	x
Hidden neurons	5	4	4	3	3	4	3	4	4	5	2	5	5
KGE validation	0,89	0,86	0,89	0,86	0,85	0,85	0,88	0,89	0,91	0,9	0,87	0,92	0,9
KGE testing	0,85	0,85	0,88	0,89	0,89	0,85	0,91	0,88	0,89	0,89	0,85	0,89	0,89

				Drenthe											
8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
x			x	x	x	x			x		x	x			x
x		x	x		x		x			x	x	x	x		x
	x		x	x		x	x	x	x	x	x		x	x	
x	x	x	x	x	x	x		x	x	x	x	x	x	x	
x	x	x		x	x		x		x	x	x				x
x	x	x		x	x			x	x	x	x		x	x	x
x	x	x		x	x	x	x	x	x			x	x		
x	x	x	x	x	x		x	x	x	x	x	x		x	
x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
4	4	5	6	3	4	6	4	4	6	3	3	5	3	4	4
0,91	0,89	0,89	0,86	0,87	0,88	0,87	0,86	0,87	0,91	0,88	0,88	0,86	0,86	0,86	0,87
0,90	0,90	0,86	0,86	0,86	0,87	0,87	0,86	0,86	0,85	0,86	0,85	0,88	0,91	0,89	0,86

to describe the validation data almost perfectly. These high performances are, however, a clear example of model overfitting, as can be seen in the corresponding testing scores. In this overfitting, substantial differences are visible between the three ANNs. The Drenthe model only starts to overfit when more than 8 hidden neurons are used, the Twente model on the other hand already overfits when more than 3 hidden neurons are applied and the Vechtstromen model starts to overfit when more than 7 hidden neurons are used. Also in the non-overfitted regions, the Drenthe model on average scores almost consistently higher for the model testing than the Twente and Vechtstromen model.

There is a less clear dependency on the number of learning epochs. None of the three models shows an evident relation. This can partly be explained by the early stopping mechanism, included in Matlab's Neural Network learning algorithm, to prevent network overfitting. Due to this early stopping, the training procedure is often stopped before the maximum number of learning epochs is reached.

The optimal number of hidden neurons and epochs is different for all three ANNs. The optimal configuration to describe the Vechtstromen region is found for a combination of 7 hidden neurons and 100 epochs. The Drenthe model scores best with a network that contains 8 hidden neurons and is trained by 50 epochs. The optimal model for the Twente region was found for 3 hidden neurons and 250 learning epochs. The performances of these three models are presented in Table 2.2.

Performance deconstruction

When the aggregated model performances, as obtained by the previous research steps, are decomposed in the performance per well it becomes clear that the aggregated performance and the individual well performance differ in multiple ways, see Figures 2.5 and 2.6.

Firstly, the individual performances are almost always lower than the aggregated performance. For the Vechtstromen model 128 of the 154 validation wells and 27 of the 30 testing wells score a lower KGE value than the aggregated KGE score. For the Twente model 40 out of 44 validation

Table 2.2: Aggregated ANN validation and testing performances, expressed in KGE and RMSE, for the Vechtstromen ANN and the regional ANNs

	Vechtstromen model			Regional models	
	Combined	Twente	Drenthe	Twente	Drenthe
KGE validation	0,92	0,95	0,90	0,89	0,94
KGE test	0,90	0,89	0,84	0,85	0,92
RMSE Validation (m)	0,25	0,20	0,27	0,31	0,20
RMSE Test (m)	0,44	0,41	0,45	0,48	0,30

wells and 9 out of 10 testing wells score worse than the aggregated score. For the Drenthe model 93 out of 110 validation wells and 19 out of 20 testing wells score below the aggregated model performance.

Secondly, when comparing the performance differences between Figure 2.5 and 2.6, with the performances presented in Table 2.2, it is found that the differences in aggregated performance and individual performance are inconsistent. The relatively high aggregated performance of the Drenthe model, for example, is not visible in the decomposed performances for each individual well.

Thirdly, Figure 2.5 shows that on an individual level, there is not much difference between the performance for the Twente and the Drenthe region. This contradicts with the aggregated results where the Drenthe model scores substantially better than the Twente model.

The low KGE values and the inconsistencies compared to the aggregated results are predominantly caused by a poor reflection of the standard deviation in the modelled groundwater depths for each individual well. This KGE component is the largest contributor to the low KGE values for 63% of the wells (validation and testing) that are interpolated by the Vechtstromen ANN and for 55% of the testing wells interpolated by the regional ANNs. Here there is no

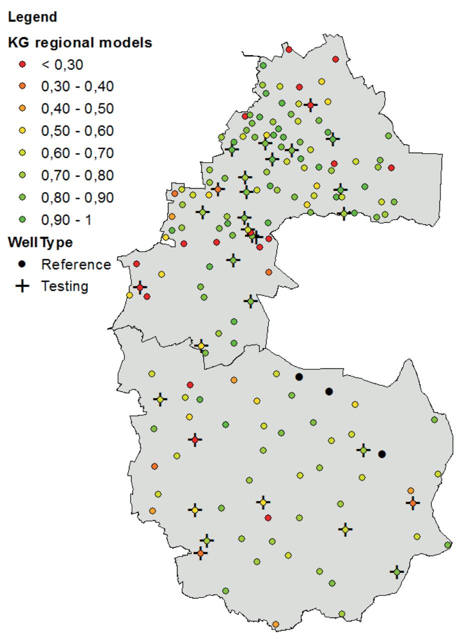


Figure 2.5: KGE scores for the individual validation and testing wells modelled by the regional ANNs

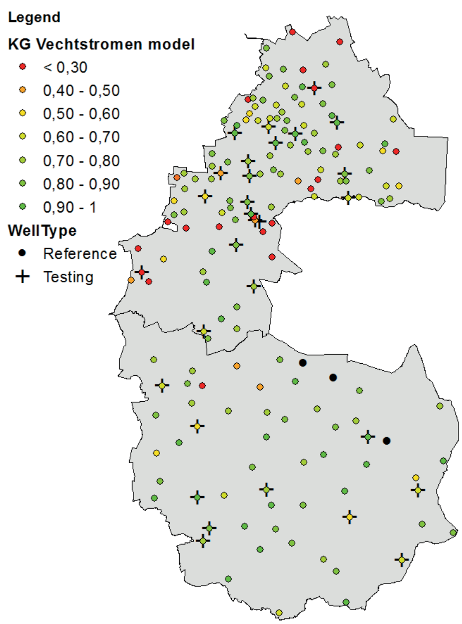


Figure 2.6: KGE scores for the individual validation and testing wells modelled by the Vechtstromen ANN

consistent under- or overestimation of the standard deviation, both occur equally often. Contrary, for the validation wells that are modelled by the regional models, the correlation is the largest contributor to low KGE scores. In 55% of these validation wells the correlation is the lowest scoring KGE parameter.

Commonalities have been found between the lowest scoring wells in Drenthe. Here the five wells that are interpolated the worst (KGE scores below -2), by both the Vechtstromen and the Drenthe model, are all located at a relatively high elevation in the Drenthe region, they belong to the 16% highest elevations in this specific region. At these elevations the groundwater depths are less dominantly controlled by surface water levels. Between the poorly scoring Twente wells no commonality in the input data has been found. As there is no clear variable that causes the problematic predictions it seems to be the heterogeneous character of the Twente region that limits the interpolation ability.

From these findings regarding the individual well performance, it can be concluded that although the model tends to be able to interpolate groundwater depths with acceptable accuracy, it does not do so by correctly describing the standard deviation in the individual well data. As this standard deviation relates to the temporal variation of the groundwater depths, it can be concluded that the model is better in describing spatial variability than in describing temporal patterns for individual wells.

DISCUSSION

This section will put the above presented results into perspective. First the methodological limitations and their effects on the results will be discussed. Then the ANN performances will be compared to traditional methods. Lastly, the potential use and the potential general applicability of the results will be discussed.

Methodological limitations

The ANNs that are constructed in this study are all trained by a single year of groundwater data. This limited time span requires some critical remarks on the results obtained by this study.

First, the interpolation ability of ANNs can only be considered to be demonstrated for moderate and drought conditions. It is expected that ANNs can also be trained for wet conditions, especially when the model is provided with the average highest instead of lowest groundwater levels. Moreover, it needs to be studied, if a single ANN can interpolate both dry and wet conditions.

Secondly, the limited time span might form a reason for the limited ability of ANNs to describe the standard deviation in the individual well data. Possibly when also relatively wet years are present in the data, the increasing temporal groundwater variations might improve the ANN's ability to describe these temporal variations on a local scale.

There are also some critical remarks to be made regarding the groundwater data that is used in this study. The piezometers used to measure the groundwater levels, actually measure the hydraulic head. At locations with perched groundwater levels, the measured hydraulic head can differ substantially from the actual unconfined perched groundwater level, with differences up to several decimetres (van der Gaast, Vroon, & Massop, 2006). This difference is caused by disturbing boulder clay layers. These perched groundwater levels, and thereby the possible measurement errors, are predominantly present at the moraines in Twente. Here it is thus not certain if the well data is reliable.

This potential inaccuracy in the data measured at the moraines of Twente poses questions to the reliability of the ANNs interpolation for the Twente region. It might also explain the difficulties ANNs had in predicting some wells, since the groundwater data might be incorrect. Thereby it might even be that the ANNs reflect the groundwater levels better than the piezometers do. If the piezometers are indeed inaccurate, then the impacts on the interpolation results needs to be further studied.

Comparison to traditional models

To put the final ANN performances in perspective, they are compared to the performance of two widely used numerical groundwater models in the Netherlands: the groundwater module (a Modflow model) of the Dutch National Hydrological Instruments and the groundwater model included in the Hydrology Stone V2.3 software (Knotters, Hoogland, & Brus, 2013). The latter is a software package that is predominantly designed to simulate nutrient washing in aquifers, but in doing so it also provides groundwater level predictions. Knotters, et al. (2013) have validated the accuracy of these models in predicting the average lowest groundwater levels. This study concluded that the Root Mean Square Errors of these models are respectively 120 and 67 cm (Knotters et al., 2013). Thereby, it can be concluded that in terms of RMSE all ANNs obtained in this study score better than these two extensively used groundwater models.

Potential use and generalization

Even though interpolated groundwater depths by ANNs seem to be outperforming alternative groundwater models, its use is not unlimited. This because of the significant difference between the aggregated and the decomposed performances. The ANNs mostly have problems incorporating the temporal variability for each individual well. Thereby, the ANNs can be used for questions that relate to spatial variability. Yet, one must be careful when applying these models to temporal questions. As the correlation and the mean terms in the KGE score are still relatively high, it is believed that ANNs are able to indicate if groundwater depths increase or decrease and around which means they vary. Yet, as the standard deviation is less well represented it can be said with less confidence how much the groundwater depths have changed over time.

Furthermore the results provide a nuanced view on the ability of ANNs to cope with differently functioning water systems in a single dataset. In the Vechtstromen model the ANN used information from the Drenthe region to improve the interpolation for Twente, because the Vechtstromen ANN performs better for the Twente region than the Twente model does. However, the ANNs has improved the predictions for the Twente region by compensating on the model fit for the Drenthe region. This indicates that the ANN is not able to simulate two differently functioning water systems in one model. Instead it seems to have found a balance between both systems. This disability to distinguish between differently functioning water systems, based upon the provided input variables, is supported by the poor performance of the Drenthe ANN regarding the higher located Drenthe wells. At these wells the surface water level control is less dominant than for the rest of Drenthe. The Drenthe ANN clearly fitted to the surface water controlled system and, thereby, failed to model the regions that are less dominantly controlled by surface water levels. What is interesting here is that no balance was found, contrary to the Vechtstromen model. Thereby, it can be concluded that when the data contains two hydrologically different systems but one is more dominantly present, the model does not balance between systems but fits to the dominant one.

The regional ANNs produced by this study are believed to be generally applicable to spatially interpolate groundwater depths for the Vechtstromen region. The small performance differences between the validation and the testing set indicate that the model has a strong general applicability. Thereby, there is confidence in the ANN's ability to also deal with other combinations of input values that are unknown to the trained ANNs. Only when the input values lie outside the ANN's training and testing range, the performance obtained in this study might not be an adequate indication.

Contrary, for other regions than the Vechtstromen region, the ANN models cannot be used. This because they are fitted to the hydrologic functioning of the Vechtstromen region, and the results appear to strongly depend on this fit. Nevertheless, the model structure is believed still to be valid. Especially for relatively flat surface water controlled regions interpolation by the obtained ANN configuration can provide highly accurate results.

CONCLUSION

This study aimed to investigate whether ANNs are able to interpolate real time groundwater depth measurements in spatially non-linear catchment areas. This ability has been studied for both a free draining and a surface water controlled water system within the Vechtstromen. These differently functioning systems have been interpolated both separately by two ANNs as well as combined by a single ANN. From these two experiments three major findings regarding the ability of ANNs to spatially interpolate groundwater depths are obtained.

Firstly, this study has found that ANNs are generally able to provide spatial groundwater depths with a higher accuracy than the currently available alternatives that require longer calculation times. This conclusion holds true regardless of the type of hydrological system the interpolation relates to. Yet, although both hydrological systems are interpolated with accuracies higher than available alternatives, the surface water controlled system is interpolated with significantly higher accuracy than the freely draining system. These score a KGE of respectively 0,92 and 0,85. A possible explanation for this might be that the groundwater depths in the surface water controlled systems are more connected and thereby more similar than they are in a free draining hydrological system.

The second major finding was that, although ANNs can cope with different types of hydrological systems separately, ANNs are limitedly able to deal with hydrologically differently functioning systems in a single model. This because ANNs appear to be unable to identify the different hydrological systems from the input variables. At least with the spatial input variables that are included in this study.

Finally, this study found that even though the interpolation results are sufficiently accurate, the ANNs do not reliably describe the individual well locations, mostly because they fail to adequately reflect temporal variations. Thereby, the use of an ANN that is trained for interpolation purposes is limited. It can provide information regarding the spatial variability in groundwater depths at a certain moment in time. However, temporal groundwater information that relates to a specific location, cannot be provided with sufficient certainty. This ability is expected to improve when the training dataset involves a longer timeseries with more temporal differences in groundwater depths. This would be a fruitful focus for further research.

Traditionally water management is based upon relatively complex numerical models. This water managing habit, however, limits the availability of crucial real-time day to day information during crisis situations. Based upon the results from this research, water managers are advised to start exploring the use of ANNs to provide real-time groundwater depth information during water crises. This will enable them to evaluate the land bounded socio-economic impacts more adequately and thereby improve their water managing responses.

3

USABILITY OF INTERPOLATED GROUNDWATER DEPTHS BY ANNs FOR DROUGHT IMPACT EVALUATION

ABSTRACT

The central goal of regional water management is to facilitate water use. Having insight in the socio-economic impact and severity of hydrological conditions is crucial to this objective. With the recent development of ANNs to obtain quick and easy spatial groundwater data, all elements are in place to evaluate this socio-economic severity, in a code green, yellow or red. This study, investigates if combining these elements will result in a sufficiently reliable severity evaluation, with a special focus on the uncertainty in groundwater characteristics. For this the socio-economic severity of 2019's drought in the Vechtstromen catchment area has been evaluated at 72 drought sensitive locations. This evaluation has been performed for both the upper and the lower confidence limits of the groundwater depth predictions, to see how the uncertainty affects the severity evaluation. This study concluded that none of the locations had more than one colour code difference between the two confidence limits and at 58 locations the colour code evaluation was the same for both confidence limits. From this result it is concluded that the groundwater depth predictions are sufficiently accurate to evaluate socio-economic drought severity.

INTRODUCTION

Water management is about facilitating water use, for both human and natural actors. Minimizing negative socio-economic impacts to these actors is, therefore, one of the main objectives. The translation from hydrology to socio-economic impact is, however, far from straight forward (Wilhite & Glantz, 1985). Locations with, for example, the lowest water availability during a drought not necessarily face the largest socio-economic consequences. Hence, to effectively manage water crises water management must not be based solely on hydrological parameters. Instead socio-economic severity evaluations must also be considered.

There are quite a number of tools around to translate hydrological conditions into measures for socio-economic impact. Yield reduction is for example an extensively used measure that can be calculated by different tools, like Aquacrop (Steduto, Hsiao, Raes, & Fereres, 2009) and WOFOST (Van Diepen, Wolf, Van Keulen, & Rappoldt, 1989). Most of these models, however, require relatively much input data and effort. In the Netherlands, therefore, two state of the art tools are being developed, the 'Waterwijzer' Nature (Witte et al., 2018) and the 'Waterwijzer' Agriculture (Mulder et al., 2018), that can quickly translate hydrological conditions into socio-economic impacts. These state of the art tools hold the potential to provide the basis for a real-time evaluation of socio-economic effects of water crises in a Dutch regional water management context.

Yet, until recently, utilizing the potential of the 'Waterwijzer' tools to evaluate drought severity in real time was limited by the availability of input data. This because the tools require spatial groundwater data, that is not easily obtained in real time. Phase one of this study focused on resolving this limitation, by studying whether ANNs are able to provide spatial groundwater patterns in a quick and easy way. The results are promising and, thereby, the input data issues seem to be resolved.

All elements to provide a real time socio-economic drought severity evaluation are thus in place. The question that rises is whether combining these separate models will result in reliable severity evaluations. Especially because of the uncertainty in the predicted groundwater data. The second phase of this research, that is discussed in this article, will therefore study if the ANN interpolated groundwater depths are sufficiently accurate to evaluate the socio-economic drought severity, in a code green (optimal), yellow (acceptable) or red (unacceptable). This three level severity evaluation fits the literature that has been used to operationalise the severity stages and will be discussed later on in the methodology. For this a case study for the Vechtstromen region will be conducted. This region is a relatively drought sensitive region in the Netherlands. Also the spatial differences in drought severity are relatively large. Some regions have access to substantial amounts of surface- and groundwater, others depend solely on rainwater. Besides ANNs have been constructed for this region to provide spatial groundwater data.

The first section of this paper delves into the methodology of the study. Here the general methodological outline will first be provided, here among others it will be defined when the outputs are considered to be reliable. Then the methodology will be elaborated in detail. The second section presents the results of the study, including a small validation of these results. The third section will reflect upon the results by discussing the limitations and the potential use and generalisation of the outcomes. The fourth and last section concludes the paper by providing an answer to the central research question

METHODOLOGY

To study if the ANNs are able to interpolate groundwater depths sufficiently accurate to evaluate drought severity, it is studied if a code green will not be evaluated as a code red, and vice versa, due to the uncertainty in the groundwater prediction. This is tested by evaluating drought impacts for both the upper and the lower confidence limit of the groundwater depth predictions, with a confidence interval of 95%. This test will be performed for a limited number of locations that combined represent the drought sensitivity of the Vechtstromen region sufficiently well.

The difference between a code green and red and vice versa, is considered as a hard measure for the usability of the ANN interpolated groundwater depths. This because when the uncertainty in the evaluation ranges from conditions being “optimal” to “unacceptable” there is no informative power left. While it was precisely the informative strength to the decision making process that formed the motivation to operationalise drought in socio-economic terms in the first place. Besides to further study the usability also the single colour code differences will be studied. These single colour code differences are not considered to make the information useless. Yet, it does provide some more insight in how reliable and thereby useful the ANN based evaluations are.

This section will first discuss what locations best reflect the drought sensitivity of the Vechtstromen region. Subsequently, the groundwater depths and the confidence limits of these predictions are defined. At last the methodology to translate these groundwater data into socio-economic impacts and finally in a colour code severity evaluation is elaborated.

Representative study locations

The allowable confidence range for which a code green is not evaluated as a code red is the smallest for drought sensitive locations. Therefore, to obtain a representative set of locations, the most drought sensitive locations in the Vechtstromen region are selected. Herein, two types of drought sensitivity are distinguished: (1) locations at which the groundwater depth is relatively sensitive to meteorological changes and (2) locations at which the land use is relatively sensitive to groundwater fluctuations.

Locations where the groundwater depths react relatively heavily to meteorological changes, are selected

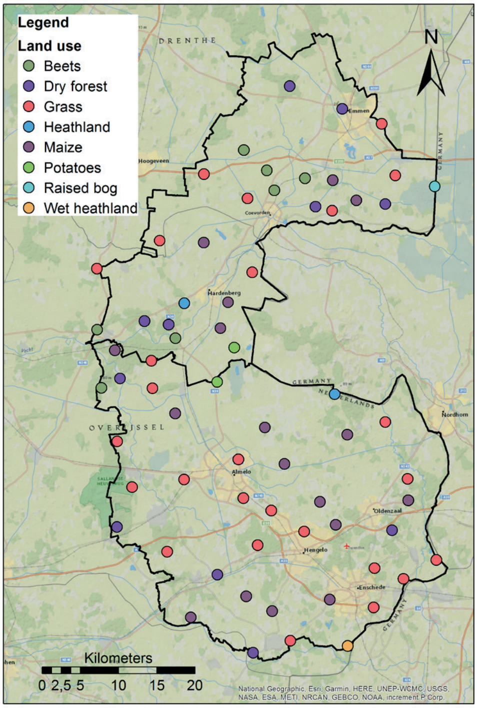


Figure 3.1: Land use for the drought sensitive locations

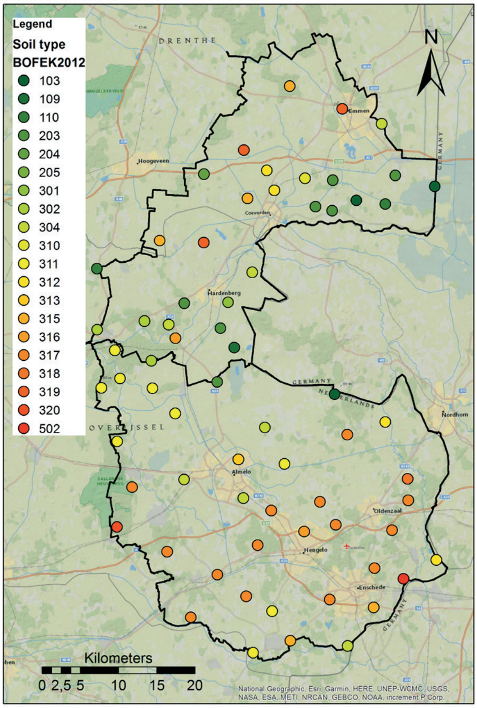


Figure 3.2: BOFEK soil type number for the drought sensitive locations

by using maps of the average lowest and highest groundwater levels in the Vechtstromen region. Sensitive locations are defined as the locations where the difference between the highest and lowest average groundwater level are the largest. Here local maxima have been used, to make sure that the selected drought sensitive points do not all relate to the same region. To do so, for each raster cell the maximum difference in a radius of 2500 meters has been searched for with ArcMap's focal statistics tool. Herein, urban regions are excluded, as the water authority is predominantly responsible for the rural areas.

Drought sensitive land use types are obtained from literature. Literature indicated that in relation to nature wet heathlands and raised bogs are the most drought sensitive nature types within the Vechtstromen region (Besse-Lototskaya et al., 2011). Contrary to for example dry woods or dry heathlands. In relation to the agricultural land uses, that are present in the Vechtstromen region, potatoes and maize are most prone to drought (Brouwer, Prins, & Heibloem, 1989). Although no sensitivity for grass was given by this source, grass is also included as it is the most commonly grown crop within the region. All these drought sensitive land use types where already included in the locations obtained by the previous groundwater criteria. Therefore, no additional locations are added.

The above described procedure resulted in a set of 72 locations that represent the drought sensitivity of the Vechtstromen region. These locations and their land use and soil type are presented in Figure 3.1 and 3.2. The presented numbers in Figure 3.2 refer to the soil type, here the hundreds reflect their soil type category. One hundreds relate to peat soil types, two hundreds to wetland soil types, three hundreds to sandy soil types, four hundreds to clay soil types and five hundreds to loam soil types.

Groundwater characteristics and confidence limits

To evaluate drought impacts, damage models require groundwater characteristics as hydrological input. For the agricultural impact models the Average Highest Groundwater level (AHG) and the Average Lowest Groundwater level (ALG) are needed. To evaluate impacts to nature the Average Spring Groundwater level (ASG) and the Average Groundwater level (AG) are also required next to the AHG and the ALG. All these groundwater characteristics are expressed in meters below the ground surface and derived from daily groundwater depth time series. This paragraph explains how these characteristics and their confidence intervals are determined in three steps.

Firstly, to obtain daily groundwater depth data the regional ANNs, constructed in the previous chapter, are used. The regional ANNs are chosen because the Twente and Drenthe ANN differ the most in accuracy. Thereby, it can be studied if this difference matters to the drought evaluation. For each location, the ANNs require two types of inputs: daily reference groundwater depths and a number of spatial characteristics for each location. The daily reference depths are obtained by averaging the hourly timeseries measured by the reference wells. Here, the full time span of the available groundwater dataset, starting at the first of February 2019 and ending at the fifteenth of January 2020, is used. Consequently the output groundwater depths span the same period. The spatial characteristics are extracted from the same datasets as are used for the construction of the ANN in the previous chapter.

Secondly, the confidence limits of these daily time series are calculated. These confidence limits are based upon the testing results of the ANNs. During the model testing the accuracy of the ANNs has been assessed by predicting depths at unknown well locations. Subsequently the differences between the predicted and the measured groundwater depths were calculated. To obtain an indication of the confidence limits of the ANN predictions, the standard deviation of the relative predictions errors have been used. This relative approach is chosen because in absolute terms the standard deviation will be dominated by the errors made for relatively deep groundwater levels. These deep groundwater levels are, however, less interesting as they limitedly influence the unsaturated zone. Taking relative deviations prevents this bias. The final confidence limits, to obtain a 95% confidence, are determined by adding and subtracting two times the standard deviation from the daily groundwater depth time series. This resulted in two time series per location, one that reflects the upper confidence limit and one that reflects the lower confidence limit.

Finally, the groundwater characteristics for each locations are calculated for both the upper and lower confidence limit time series. To define the AHG the three smallest daily groundwater depths within a year, that are at least 14 days apart, are averaged. The same has been done for the ALG only with the three largest groundwater depths. The ASG is obtained by averaging the groundwater levels at the 14th of March, the 28th of March and the 14th of April (Finke et al., 2004). Finally, to obtain the AG the complete time series are averaged per location. As these groundwater characteristics are averaged results of the daily time series, the confidence level is higher than the 95% for the individual daily predictions. This because the chance that the average of the daily predictions that are used in the calculation of the groundwater characteristics lies above the upper limit or below the lower limit is smaller than the chance that a single value exceeds this limit. What precisely is the confidence level is, however, not known.

Calculating drought damage

The impact of the hydrological characteristics on agriculture is calculated by means of the HELP tables (Van Bakel, Huinink, Prak, & van der Bolt, 2005). The HELP tool is a relatively old tool, but as the 'Waterwijzer' Agriculture is still under construction it is the best available method. The HELP tables were developed in 1987 to quantify the impacts of changing hydrological conditions due to spatial interventions on agricultural yields. To do so, the HELP tables provide goal reductions, expressed in terms of crop yield damages relative to the theoretical potential yield, as a function of the groundwater characteristics and soil type. Thereby, the HELP tables have the same goal as the state of the art 'Waterwijzer'. Only the relations are based on less advanced modelling principles, as they are defined by model simulations in the LAMOS model. A model that is derived from the MUST model, which at the time was the best available tool for unsaturated flow modelling (Van Bakel et al., 2005). The limitations will be discussed in more details in the validation paragraph.

To calculate the yield loss as a consequence of the groundwater characteristics, the HELP tables are provided with the crop type, soil type, AHG and the ALG for each of the 72 drought sensitive agricultural locations. The impacts are calculated for both the lower and the upper confidence limit of the AHG and ALG. At points where the groundwater characteristics exceeded the input range of the HELP tables, the characteristics are set to the maximum input limits. These were 200 cm and 320 cm for respectively the AHG and the ALG (Van Bakel et al., 2005). This adjustment to the input limits is believed not to be problematic as such deep groundwater levels have a negligible influence on the unsaturated zone.

To assess the hydrological impact to nature, the 'Waterwijzer' Nature is used. The 'Waterwijzer' Nature is a state of the art tool that enables to quantify the effects of changing abiotic conditions, like groundwater levels, to nature goal realisation (Witte et al., 2018). The 'Waterwijzer' model is based upon the same methodological principles as its predecessor, the WATERNOOD model that has been designed by Runhaar and Hennekens (2015). This method provides trapezoidal functions that for each nature type define the goal realization as a function of the four groundwater characteristics. To do so, for each characteristic and nature type four critical points have been defined, see Figure 3.3. Point A1 describes the critical minimum level of the groundwater characteristic below which there is no form of goal realisation anymore. Point A2 describes the critical maximum groundwater level above which the goal realisation becomes zero. Point B1 describes the critical minimal groundwater level for which there is 100% goal realisation

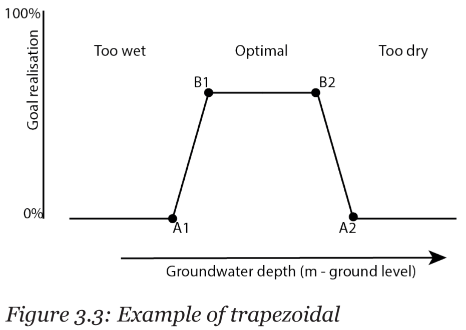


Figure 3.3: Example of trapezoidal WATERNOOD goal realisation function

and point B2 describes the critical maximum groundwater level to obtain the maximum goal realisation (Runhaar & Hennekens, 2015). Points A1 and B1 and points A2 and B2 are connected linearly. Thereby the goal realisation decay is assumed to relate linearly to the groundwater level.

The term goal realisation in relation to nature reflects the extent to which the abiotic condition, like the groundwater level, enable the realisation of the planned nature type (Runhaar & Hennekens, 2015). When groundwater levels are within the 100% range, groundwater is thus not a limiting factor for the development of the desired nature type. The other way round, when groundwater is in the 0% range, groundwater levels make it impossible for the desired nature types to develop. It is important to understand that goal realisation does not reflect the actual nature conditions. Only the extent to which the abiotic condition enables the nature type.

To evaluate the nature goal realisation, the 'Waterwijzer' nature model has been provided with the AHG, ALG, ASG, AG and the local nature type of each drought sensitive nature location. Herein the nature types are distinguished according to the management type characterization of the Index nature and landscape (Bij12, 2020). This characterization is currently most preferred in Dutch nature management. The impacts are calculated for both the lower and the upper confidence limit of the four input groundwater characteristics.

Evaluating damage severity by colour codes

To evaluate the severity of the reduced goal realisation for agriculture and nature, severity limits have been obtained from literature. For this the work of Bouwman et al. (1998) is used. This work formed the basis for the practical transition from surface water to groundwater oriented water management. In their work, Bouwman et al. provide an indication on when groundwater levels are acceptable or not. They distinguish three main categories: optimal, acceptable and unacceptable (Bouwman et al., 1998). The limits they proposed for these categories are presented in table 3.1. These limits have been used to evaluate the severity of the reduced goal realisation due to the hydrological conditions. The colour codes to indicate this severity are also presented in table 3.1.

Table 3.1: Severity limits and relation to colour codes

Severity evaluation	Goal reduction	Colour code
Optimal	<=10%	Green
Acceptable	>10% & <= 25%	Yellow
Unacceptable	>25%	Red

Validation of agricultural severity

The used HELP tables contain some serious limitations in calculating the goal reduction (Mulder et al., 2018b). First, the modelling of the unsaturated zone has improved significantly since then. Thereby, state of the art models are better able to simulate the relation between the input variables and the yield losses. The main improvement lies in the improved way of modelling transpiration reduction. Secondly, the relations are defined upon outdated climatological conditions. Conditions that due to climate change are not valid anymore. On top of that the way the meta-relations are defined is not climate robust. Thirdly, the farming operations have developed significantly over the past decades. The damage calculations, therefore, do not adhere to current farming practice. Fourth and finally, the inundation damages are not reproducible because they are mostly based upon expert judgement.

Because of these four seemingly influential limitations, the impact of these limitations on the results of this study are estimated. For this the 'Waterwijzer' Agriculture is used. Although it was not yet possible to use this tool to evaluate the goal reduction for all drought sensitive agricultural locations, a readily developed pilot model could be used for a small validation. This pilot model was made to evaluate drought impacts for the Rheezermaten region, which is located in the Vechtstromen catchment area. Within this pilot model locations have been searched that have the same combination of land use and soil type as any of the 72 drought sensitive locations for which HELP calculations have been performed. For eight locations there where matching locations in the pilot model. By studying the impact of the groundwater characteristics on these matching Rheezermaten locations the goal reductions for the eight drought sensitive agricultural locations are approximated. Comparing the severity evaluation of these approximated goal reductions with the severity of the reductions as predicted by the HELP tables, provides an indication of the impact of the limitations of the HELP tables on the conclusions of this study.

RESULTS

GROUNDWATER CHARACTERISTICS AND CONFIDENCE INTERVALS

The resulting groundwater characteristics that form the basis for the severity evaluation are presented in Figures 3.4, 3.5, 3.6 and 3.7. The confidence intervals are presented in Table 3.2. These confidence intervals apply to the underlying daily groundwater timeseries. Due to the way the groundwater characteristics are calculated these ranges also apply to the four groundwater characteristics, only the corresponding confidence level is higher than 95%, as is discussed in the methodology.

When studying the four groundwater characteristics and their confidence intervals the groundwater characteristics in Drenthe seem to be plausible. The relatively flat regions have shallower groundwater depths and locations with larger elevations have deeper groundwater depths. Hence there is no reason to distrust the results.

Table 3.2: Confidence limits of the ANN interpolated groundwater depths and

ANN model	Std.	Lower limit	Upper limit
Drenthe	6%	86%	110%
Twente	16%	70%	133%

For Twente the groundwater characteristics at the western and middle regions seem plausible. Even though, the groundwater characteristics located in the centre of Twente where expected to be a bit less deep. The predictions at the eastern moraines, however, raise some reliability questions. These seem to be too shallow. At these eastern moraine locations the long term ALG, that has been used as input to the ANN, does not fall within the confidence limits of the short term ALG prediction. While at the other locations this long term ALG does fall within the confidence limit. Hence it is questionable if the predicted groundwater depths at the eastern moraines can be trusted.

GOAL REDUCTIONS

The goal reductions as a consequence of the above presented groundwater characteristics, are shown in Figure 3.8 and 3.9. Three things stand out from these two figures.

Firstly, it is remarkable to see that for agricultural locations on average the goal reduction is larger in the Drenthe region. This contradicts with the differences in groundwater characteristics, where on average Twente had deeper groundwater levels. This contradicting translation from groundwater characteristics to goal reduction, shows that goal realisation is a function of more variables than solely

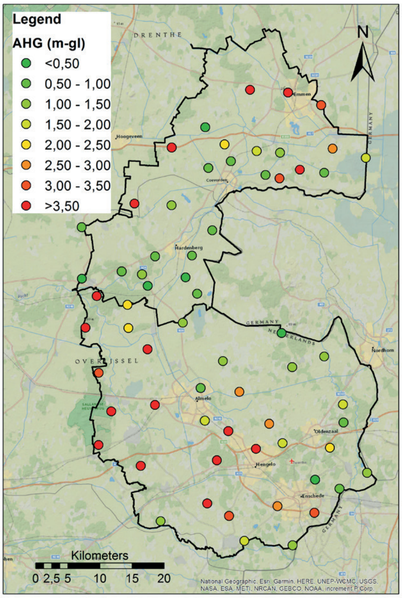


Figure 3.4: Average highest groundwater depths predicted by the ANNs

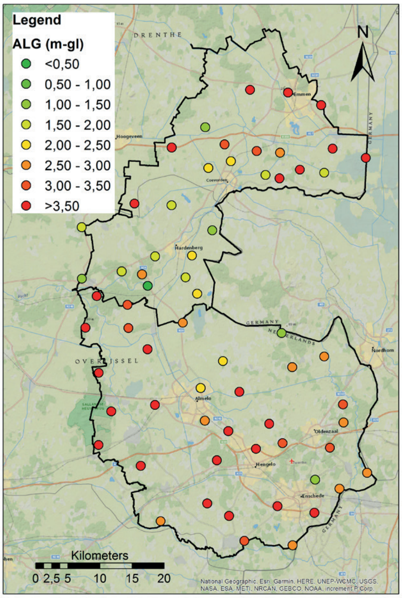


Figure 3.5: Average lowest groundwater depths predicted by the ANNs

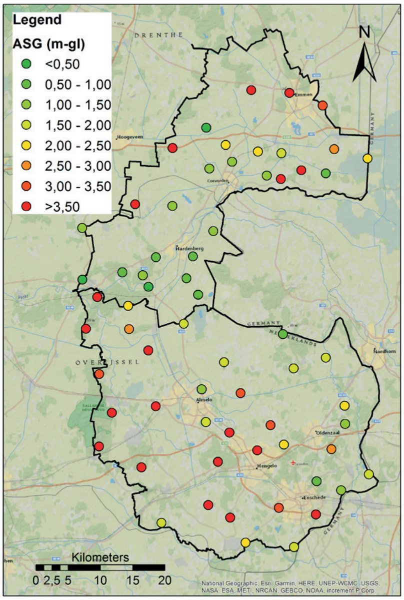


Figure 3.6: Average spring groundwater depths predicted by the ANNs

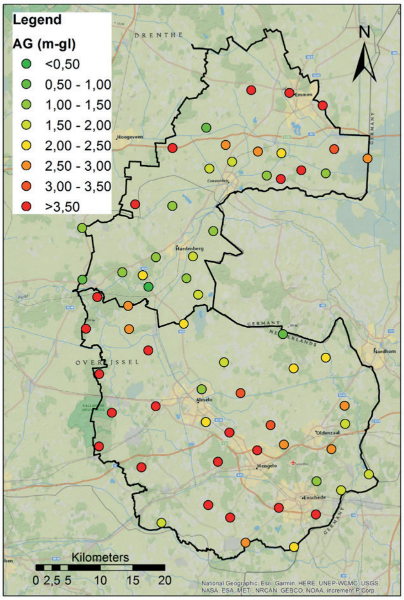


Figure 3.7: Average groundwater depths predicted by the ANNs

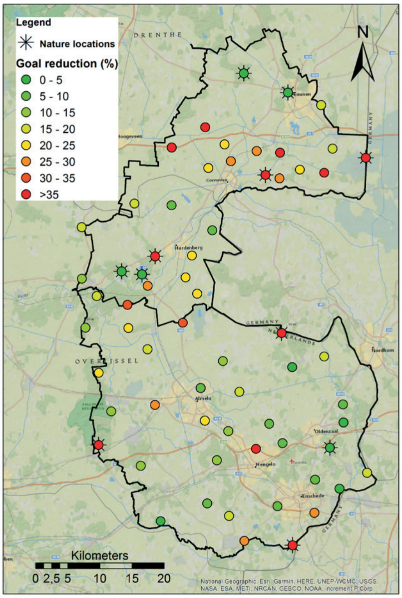


Figure 3.8: Goal reduction for the minimum scenario

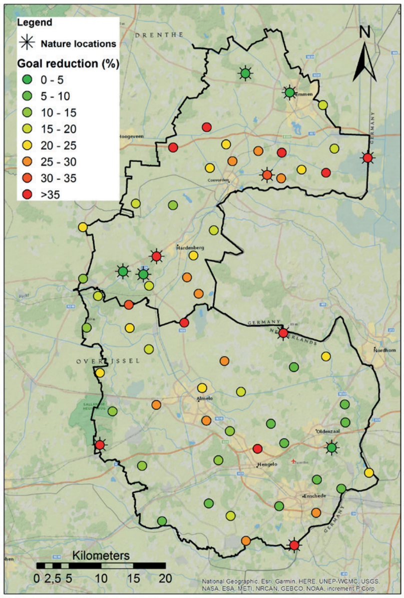


Figure 3.9: Goal reduction for the maximum scenario

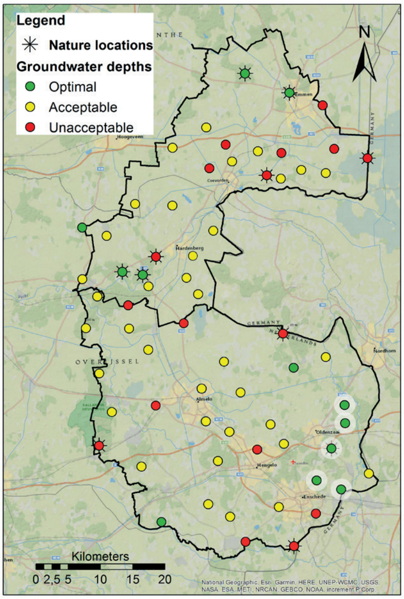


Figure 3.10: Severity evaluation for the minimum scenario, white marked points are improbable and discussed in the text

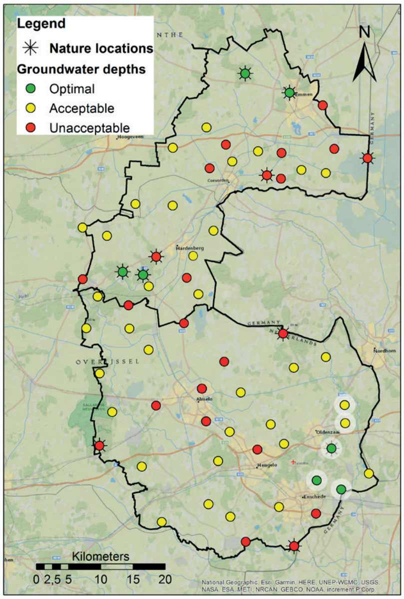


Figure 3.11: Severity evaluation for the maximum scenario, white marked points are improbable and discussed in the text

hydrology. Land use and soil type have a substantial contribution. This stresses the importance of operationalizing drought in socio-economic terms instead of solely hydrological ones. On the other hand it also leads to doubts regarding the reliability of the HELP tables for the sandy Twente region. It is well possible that the HELP tables underestimate the damages. These doubts are already present at the Vechtstromen water authority.

Second, the goal realisation regarding nature shows interesting extremes. Reductions are either zero or hundred percent. Here the hundred percent reductions relate to the drought sensitive nature types. The zero percent reduction relates to forest types of nature. Here one must bear in mind that hundred percent reduction does not mean that all nature disappeared in this season. It means that the groundwater levels do not contribute to the development of the specific nature type. The all or nothing result shows the strong differences between the different types of nature.

Finally, the results show that the differences between the lower and upper confidence limits are relatively small. On average the difference between the minimum and maximum is 10% of the maximum damage in Drenthe and 13% in Twente, even though the differences in groundwater characteristics where 24% for Drenthe and about 63% for Twente. The differences are thus strongly reduced in the translation from groundwater depth to goal reduction.

Severity evaluation

Figure 3.10 and 3.11 show how the goal reduction translates in severity evaluation. First and foremost it stands out that for none of the locations the severity evaluation differs more than one colour code between the upper and lower confidence limit. At 14 locations there was a difference of one colour code, these are all agricultural locations, and for the remaining 58 locations the colour code was the same for the upper and lower confidence limits. For the studied case the ANN predictions, therefore, seem sufficiently accurate.

When studying the Drenthe and Twente region separately there are no large differences in the evaluation performance, even though the confidence ranges differ substantially. Even more, with 31 out of 38 locations with the same colour code, Twente has more consistent evaluations than Drenthe, where 27 out of 34 locations where consistently evaluated for the two scenarios. Based upon these results the goal reduction in the Drenthe region seems to be more sensitive to groundwater depths. This is not unlikely because in Twente groundwater levels drop relatively quickly to a level below the wilting points of the different vegetation types. However, even though the differences between Drenthe and Twente can possibly be explained, the reliability of the Twente results can still be questioned. The obtained evaluations at Twente's eastern moraine do not correspond with the experience in practice and seem to be a result of inaccurate groundwater depth predictions in combination with limitations of the HELP tables.

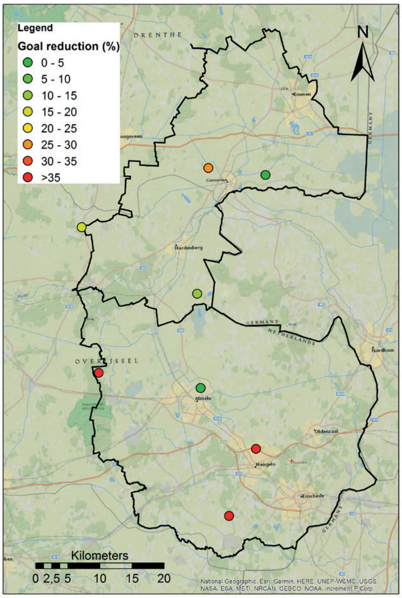


Figure 3.12: Goal reduction for the lower confidence limit calculated with the 'Waterwijzer' agriculture

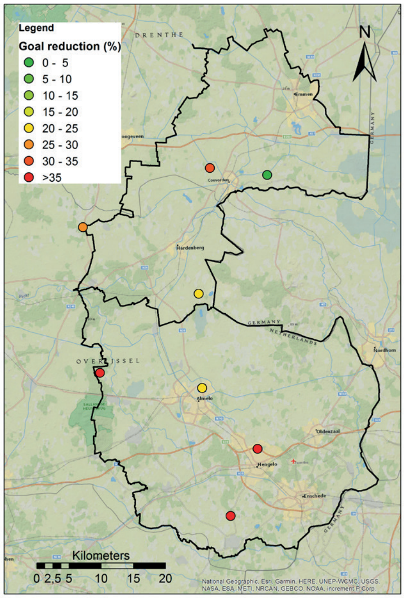


Figure 3.13: Goal reduction for the upper confidence limit calculated with the 'Waterwijzer' agriculture

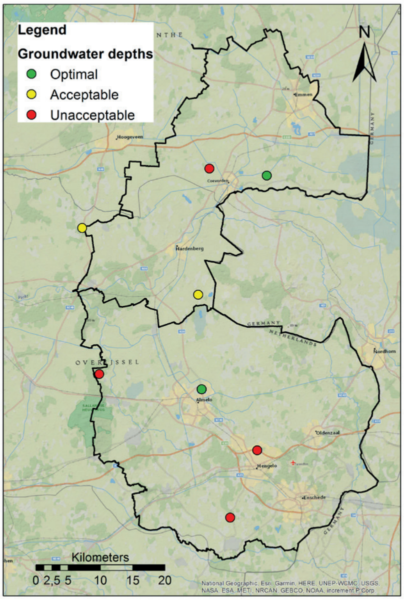


Figure 3.14: Severity evaluation for the lower confidence limit calculated with the 'Waterwijzer' agriculture

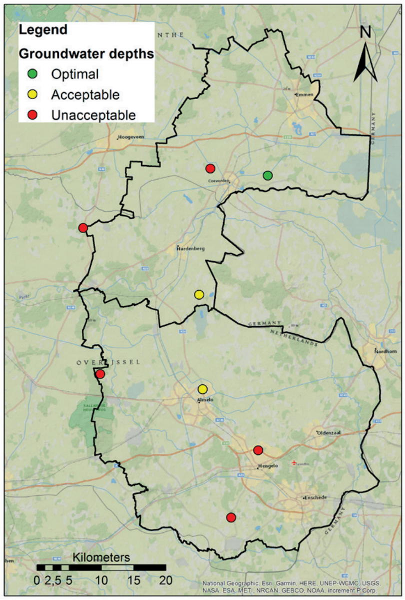


Figure 3.15: Severity evaluation for the upper confidence limit calculated with the 'Waterwijzer' agriculture

Validation of agricultural results

In Figure 3.12 and 3.13 the validation runs performed with the 'Waterwijzer' agriculture are shown. These validation results show some significant differences with the damages obtained from the HELP tables. In Twente the HELP tables have underestimated the goal reduction for three out of four locations. At all three locations the groundwater depths were relatively deep. In Drenthe where the groundwater depths are less extreme, two locations were underestimated and two were overestimated. On average for all eight locations the relative difference between the estimated maximum reduction by the HELP tables and the "Waterwijzer" Agriculture was 45%.

Contrary, when studying the differences between maximum and minimum severity evaluation, see Figure 3.14 and Figure 3.15, the results of the HELP tables and the 'Waterwijzer' agriculture are relatively similar. Just as for the results produced by the HELP tables, there was no "Waterwijzer" evaluation for which the colour code difference between the upper and lower confidence limit was more than one colour code. The only difference to the HELP tables based severity evaluation is that at some locations the minimum and maximum colour code shifted a code. This inaccuracy is, however, less relevant to the aim of this study. Hence, the differences in predicted goal reductions, do not necessarily mean that the conclusions regarding the consistency of the colour code evaluations are unreliable.

DISCUSSION

This section will put the above presented results into perspective. To do so, first the methodological limitations and their effects on the results will be stressed. Then the potential use and the generalisation of the results will be discussed, to explore the usefulness of the results.

Methodological limitations

The applied methodology in this research comprises three important limitations that might possibly affect the study's conclusions.

Firstly, it is to be questioned if 2019 was an appropriate year for this study. As 2019's drought was relatively extreme it seems that most of the locations ended up in a maximum impact, regardless of the confidence limits. This seems especially true for nature locations, that only produced goal reductions of zero or hundred percent. It might well be possible that for a less extreme year the upper boundary scenario would have caused maximum impact while the lower boundary scenario was not even near this state of maximum impact. In such scenario the differences between the two scenarios will be larger and thereby potentially also the difference in severity evaluation. This limitation was, however, unavoidable as data was only available for this specific year.

Secondly, limitations come with the use of the HELP table and the 'Waterwijzer' Nature tool. Both tools are designed to assess the impact of structural changes to hydrology, instead of deviations within a year that this study is interested in. For agricultural locations it is believed that this difference between structural and yearly deviation is not that large, as crop growth cycles are largely annual cycles. The impact of structural change is thereby largely the sum of individual years. The most problematic limitation of the HELP-table is, therefore, that the influence of precipitation on the crop growth is not included adequately. Yet, the validation of the HELP results by the 'Waterwijzer' Agriculture tool, that is able to evaluate single year impacts, does not provide an indication that this HELP limitation affects the conclusions of this study.

For nature this limitation of structural versus seasonal assessment might actually affect the conclusions of the study. This because nature development is not solely the result of individual annual cycles, like agriculture. Within nature there are also more gradual processes. These processes are often relatively resilient to short periods of drought, but heavily affected by structural drought. The impact of a single dry year is, therefore, substantially different from that of structural drought. The hundred percent goal reduction, that result from the 'Waterwijzer' Nature tool, are thus likely to be lower in practice as the provided groundwater characteristics related to a single year. Still, also within a single year more extreme results are expected than for agriculture. During a discussion on the results of this research with Bas Worm, a strategical water system advisor at the Vechtstromen water authority, he stressed that nature goal realisation follows a less gradual scale than agriculture does. Unfortunately, there is no alternative tool for nature that is able to evaluate drought impacts within a single year.

Finally, the ability of ANN interpolated groundwater depths to provide sufficiently accurate severity evaluations, is as good as the colour code limits that have been chosen. Herein lies a problem as the results appear to depend strongly on the limit between codes green and yellow. If this limit is increased by only 1% it appears that for one location the difference between the upper and lower limit becomes more than one colour code. The results are less dependent on the limit between codes yellow and red. Only when this limit is reduced from 25% to 13% the evaluations start to differ more than one colour code. However, despite this sensitivity to decreasing the yellow range by increasing the lower limit, it is not recommended to reduce the difference between codes green and red. A 15% difference between green and red seems to be in quite good proportions to the order of magnitude of the absolute impacts. Besides, making the yellow range smaller will result in less deviation between the locations and thereby less information on how to prioritize water managing decisions. The sensitivity of the results to these limits does affect the robustness of the conclusions. The severity limits are, however, obtained from an influential piece of literature in the Dutch regional water management. Therefore, the results are considered to be useful.

POTENTIAL USE AND GENERALIZATION

This study showed that operationalising drought in terms of socio-economic severity provides valuable water managing insights. The non-straightforward translation of groundwater characteristics into socio-economic impacts, shows that managing water solely based upon hydrological parameters does not necessarily result in the maximum socio-economic gains. This because the role of soil type, land use and meteorology is too significant.

The combination of ANNs and damage tools seems to be a promising way forward in this socio-economic operationalisation of drought. Even though the robustness of the results from this study to the colour code limits is limited, it is believed that there is sufficient potential for improvement in the individual components to improve the robustness of the outcomes. First of all, the new 'Waterwijzer' Agriculture is expected to resolve the issues that emerged from using the HELP tables. Secondly, the ANNs can probably be improved. It is believed that when the ANNs are trained to more specific conditions their accuracy will increase, especially in Twente. It is worth to try building an ANN specifically for the moraine region in the east of Twente. Also for some nature types, that have a fundamentally different interaction between land use and groundwater, like raised bogs, it might be interesting to design a separate ANN. With these more accurate ANNs the confidence range of the predictions will narrow down and thereby result in more robust predictions.

Finally, the conclusions of this study are expected to apply more generally than for the Vechtstromen region only. This because of two observations. First of all, there was a relatively wide range of land use and soil type combinations. The 72 locations contained 48 different combinations. For all these different combinations the difference was smaller than two colour codes. The general conclusion, therefore, seems not to be too sensitive to the land use and soil type combination. Secondly, it is expected that the confidence range of the Twente ANN is relatively large, especially within the Netherlands. This because of its complex geology, the Twente region is expected to be one of the more difficult regions to interpolate by using ANNs. ANNs for other locations will thus likely have smaller confidence ranges than the Twente ANN. Thereby, the chance that the upper and lower confidence severity evaluations differ more than one colour code reduces.

CONCLUSION

This research aimed to study if ANN interpolated groundwater data are sufficiently accurate to evaluate socio-economic drought severity. To do so, it has been studied if the difference between the severity evaluation for the upper and for the lower confidence limits of the groundwater characteristics, the average highest, average lowest, average spring and the average groundwater level, is no more than one colour code. For this, the drought severity for 72 drought sensitive rural locations, that combined reflect the Vechtstromen region, has been evaluated. These locations included both nature and agricultural land use types.

For none of the 72 locations the evaluation between the upper and lower confidence limit differed more than one colour code. Even more, at 58 locations the evaluation was constant for both confidence limits. Thereby, the groundwater data produced by the ANNs are considered to be sufficiently accurate.

The conclusion of this study is, however, not very robust. Shifting the boundary between code green and yellow by one percent already affects the conclusion. Yet, due to the possible improvements in the ANN models, it is believed that this limited robustness will not be problematic. When the individual components are improved, the confidence limits will narrow down and thereby also the sensitivity to the colour code boundaries will reduce.

4

DISCUSSION AND FUTURE STEPS

The previous two chapters delved into the possibilities of using state of the art tools to evaluate socio-economic drought severity. They showed that ANNs are able to interpolate groundwater depth data quickly and easily and with sufficient accuracy to evaluate socio-economic drought severity. In this chapter it will be discussed to what extent this state of the art operationalisation needs further improvement to better fit the drought severity definition that is defined in the first research phase. To do so, it will firstly be discussed how well this state of the art drought operationalisation adheres to the qualitative drought severity definition. As this qualitative severity definition embodies the responsibility of regional water managers, this comparison provides insights in how valuable the severity evaluation is to the water managing practice. Based upon these insights recommendations for further research and development are formulated to improve the usefulness of the operationalisation.

SOCIO-ECONOMIC DROUGHT DEFINITION

When drought becomes problematic to a water managing agent has been defined in the first phase of this research project (Beltman, 2019). This has been done by defining when societal impacts need to be considered problematic by a water manager. These impacts were defined for dairy farmers, wet heathlands and raised bogs. These water users were selected because of their special interest to the Vechtstromen water authority. Dairy farms are the most dominant land users in the region and the wet heathlands and the raised bogs are two of the most threatened and drought sensitive nature types located in the Vechtstromen region.

Dairy farmers experience four drought problems that should be considered problematic by water managers: reduced grass yield, reduced maize yield, emerging weeds and a harmed soil life. A reduction in grass yield needs to be considered problematic when it results in unlawful protein self-sufficiency, significant nitrogen surpluses or in profit losses to such extent that it risks large scale bankruptcy of dairy farms. Reduced maize yield and emerging weeds are mostly problematic because of the latter effect of risking large scale bankruptcy. Harmed soil life is problematic due to its inducing effect on the other three problems.

The importance of profit losses is likely to hold a broader applicability than only to dairy farmers. This because economic gains are the main focus of agricultural businesses. An

Table 4.1 Drought problems, indicators and problematic limits, obtained by research phase 1

Drought problem	Problem extent defined by	Fully problematic when
Reduced grass yield	Effect on profit	Risk of large scale bankruptcy
	Protein self-sufficiency	Self-sufficiency below legal requirement
Reduced maize yield	Effect on profit	Risk of large scale bankruptcy
Emerging weeds	Effect on profit	Risk of large scale bankruptcy
Harmed soil life	Effect on profit	Risk of large scale bankruptcy
	Protein self-sufficiency	Self-sufficiency below legal requirement
Reduced nature goal realisation	Human induced impact to goal reductions	Irreversible impacts

overview of the drought problems, their key indicator and their critical limit is provided in Table 4.1.

What makes impacts to nature problematic to regional water managers fundamentally differs from the dairy farmers' problems. Not the impacts themselves are problematic, but the human interference that intensified these natural problems. The moral responsibility of water managers not to induce negative impacts of hydrology to nature, makes that the conclusions of the first research phase are wider applicable than to the problems identified for wet heathlands and raised bogs only. Thereby all human induced drought impacts to nature need to be considered problematic.

STATE OF THE ART OPERATIONALISATION VS. DEFINITION

The qualitative drought severity parameters and limits, that are presented in Table 4.1, are not fully operationalised in the state of the art severity evaluation tool, that is constructed in this second research phase. For agricultural operationalisation the difference between the operationalised tool and the theoretical severity definition is limited. For nature, on the other hand, the differences are more fundamental.

Briefly summarized, it is mostly the water managing effect on profits that is problematic in relation to agriculture. Not because of the economic loss itself, but because the water systems needs to contribute to a stable economy. Additionally for dairy farmers the protein self-sufficiency level is a second key variable for drought severity. In the operationalisation provided in this research, the drought severity evaluation did not incorporate these two variables. It was only based upon the yield loss in terms of dry matter. This because this is the only information that the HELP tables provide. However, with the newly developed 'Waterwijzer' agriculture it becomes possible to evaluate drought severity based upon these two parameters (Mulder et al., 2018). The 'Waterwijzer' not only provides opportunities to defines yield losses in terms of kilograms of dry matter, but also in monetary terms. This translation to monetary value, is largely based upon the nutritional values of the yields. Hence both parameters are incorporated in the 'Waterwijzer'. The monetary damage that the 'Waterwijzer' agriculture can provide, however, do not incorporate the impact of weeds and harmed soil life.

To evaluate drought severity for nature it are the human induced drought impacts that are problematic. Unfortunately, the 'Waterwijzer' Nature does not include an option to differentiate between natural variation and human induced impacts. There is only one small correction possible that accounts for the limitations posed by the site characteristics. For this the 'Waterwijzer' provides an option to calculate the maximum goal realisation that can be reached given the spatial characteristics of the site (Witte et al., 2018). This function predominantly accounts for the limitations caused by elevation differences. This is however a static correction while it are the temporal variations that are most relevant to the drought evaluation. Another limitation of the 'Waterwijzer' nature, that is extensively stressed in the third chapter of this research, is that the 'Waterwijzer' nature evaluates the impacts of structural changes to hydrology, instead of seasonal changes. Thereby, the impacts to goal realisation are likely to be overestimated. All in all, the 'Waterwijzer' nature does not provide insights that correspond to the qualitative drought severity definition. Yet, it is the best tool that is available.

THEORETICAL VS. PRACTICAL SEVERITY LIMITS

As parameters by which drought is operationalised differ from the qualitatively defined parameters, also the drought severity limits differ. In this research, aggregated literature based limits, in terms of goal reduction, have been applied. However, when the 'Waterwijzer' agriculture will be applied to provide insights in the economic costs and in the self-sufficiency levels, these aggregated limits will not be applicable anymore. Instead, the operationalisation of the qualitative drought severity definition requires more conditional limits. The limit, both its unit and the level, varies for the problem that occurs. Only to nature static limits in terms of goal realisation still suffice, as here it is not the impact but the cause that is problematic to water management. Introducing these conditional limits instead of aggregated ones, will not only provide insights in the drought severity, but also in the problem that underlies this severity evaluation. Thereby, it provides more explanatory power and is more valuable to the decision making process.

PRACTICAL CONSIDERATIONS

Whether the state of the art operationalisation is useful to evaluate drought severity or not, not only depends on how well it fits the responsibility of regional water managers. Since the drought severity evaluation aims to support crisis management practical considerations are also important, especially in relation to workability and speed of the tools. Obtaining groundwater data with the ANNs will not form a bottleneck in the operation. Operating both 'Waterwijzers' requires a bit more attention.

The ANNs have specifically been designed to obtain reliable groundwater depths quickly and easily. To operate they only need to be provided with three well measurements. These wells are already in place and this data is already collected in real time. Next to these three well measurements, only static spatial data is required. This data can already be setup in advance. There is thus no data collection burden during a crisis situation. Besides the

ANNs calculate groundwater depths extremely fast. On an 8th generation i7 intel processor the ANNs produce about 1200 groundwater depth calculations per minute. This means that per minute the groundwater depths for 1200 grid cells are calculated. Finally, the testing of the ANNs, presented in chapter two, showed that these quick and easily obtained results are more accurate than traditional numerical groundwater models.

The 'Waterwijzers' are also designed for easy operation with as limited input data as possible. Yet, the agriculture tool does require more effort in setting up the model. Since the custom tool is used to calculate daily yield responses, there is relatively much site data needed. This can fortunately all be done in advance, since these are static parameters. Hence, during a crisis situation itself, there is no significant effort needed. The run time, 7 cells/minute, on the other hand might form a bottleneck for creating complete drought severity maps. It is currently not known how many grid cells there will be to operationalise drought severity in a fully covering map for the complete Vechtstromen region, so what the runtime will be cannot be said. Yet, it is likely that its order of magnitude is hours and not minutes. Finally, in the current tool the groundwater depths that are produced by the ANNs need to be manually implemented half way the calculation process. For operation this process needs to be automated. The nature tool on the other hand is not expected to form any operational bottleneck.

FUTURE STEPS

It is recommended to start operationalizing the socio-economic drought severity as is done by this study. This because the non-linear translation from hydrological parameters to severity, see figures 3.4 – 3.11 in chapter three, shows the importance of putting the hydrological influence in perspective. In the end it is not water that needs to be managed, but use that needs to be facilitated. Facilitating use solely based on hydrological parameters will provide suboptimal results. Despite the limitations it is, therefore, believed that the operationalised socio-economic drought severity will still improve water management. Yet, due to the limitations the results need to be treated with care and can on itself not form the basis for decision making. The insights in socio-economic severity complements hydrological insights, it does not replace them. Nevertheless, with the insights that the 'Waterwijzer' agriculture provide to the drought definition, the agricultural severity evaluations are considered to be strong indicators that may take substantial weight in decision making. The nature evaluations are however of more indicative nature. In the decision making process they need to take a less prominent role.

Before the operationalisation can be used in practice there are still some questions to be answered and improvements to be made. Firstly, the ANN predictions for the moraines in the eastern parts of Twente need to be improved. This can best be done by formulating a separate ANN for locations that are located relatively high at the moraines on a disturbing clay layer. Here the groundwater depths react differently than in other Twente locations. Therefore, they are poorly reflected by the current ANN. Whether formulating a separate ANN is possible depends on the amount of data that is available. Also it needs to be sure that the wells measure the groundwater depth above the disturbing clay layers. That is, after all, the groundwater level that influences soil moisture availability. To do so, possibly more wells need to be installed.

Secondly, it needs to be studied if full spatial mapping of the socio-economic severity is possible and how this should be done. Here there needs to be searched for a balance between calculation time and the added value of the information. It is advised first to study what would be the maximum allowable grid size to obtain reliable results. The number of grid cells can be reduced by changing the cell size of relatively continuous variables, like elevation. The boundaries of discrete variables, like land use, need to stay intact. If the minimum required number of cells still takes too much runtime it is advised to switch to point evaluations. Even though spatially covering maps where the reason to design ANNs, the model still provides added value when points are studied. This because it provides room to evaluate any point of interest, not only the points at which there is a well installed.

Thirdly, it is worth studying if it is possible to separate the human induced impacts to nature from the natural variability. This is, however, a complex question that can likely not be solved within the 'Waterwijzer' calculations. Therefore, a more pragmatic solution might be needed. If it is known what is the natural groundwater fluctuation due to the precipitation deficit and what is the human induced variation because of water use and management, separate calculations can be made in the "Waterwijzer". One in which only natural variations are assessed and the other in which the total variation is studied. The difference in goal realisation between the two scenario's is then the human induced goal realisation impact to the specific nature type. Separating groundwater fluctuations is probably possible by using the available hydrological models and has already been done by (van Loon et al., 2016). Obtaining the human induced impacts to nature this way is, however, a pragmatic solution that is not supported nor rejected by literature as it is not known if the resulting difference in socio-economic impact actually reflects the human induced impacts.

Finally, the qualitative severity boundaries, regarding economic costs and protein self-sufficiency, need to be quantified. It must be studied what profit losses will risk large scale bankruptcy of agricultural businesses. Herein there needs to be separately accounted for seasonal drought and consecutive dry years. Therefore, the eventual severity boundaries will likely become quite complex heuristics. In relation to protein self-sufficiency, the law and the planet proof hallmark already provide indications for severity limits. Yet, it needs to be explored how to deal with the different amounts of hectares that dairy farms have. This influences the self-sufficiency significantly and also poses the question whether the farm simply has too few grass land or if it is the water authority that insufficiently facilitated protein self-sufficiency.

Besides these improvements one opportunity also needs to be stressed as recommendation. The currently presented severity evaluation is a real time evaluation. Yet, since socio-economic impacts are at the end of the drought propagation process, managing water on these severity insights results in reactive management. The socio-economic severity of today is the result of precipitation deficits and surface water levels of days or weeks before. To improve management it would thus be desirable to be able to make socio-economic severity predictions. The severity evaluation approach that is presented in this report, can enable predictive evaluation. For this only the groundwater depths at the three reference wells, that serve as input for the ANNs need to be predicted and combined with the precipitation predictions. There are multiple studies that showed that ANNs are good in predicting groundwater depths (Chitsazan et al., 2015; Daliakopoulos et al., 2005; Mohanty et al., 2010; Nayak et al., 2006; Yoon et al., 2011). Hence, to enhance proactive water management based upon socio-economic severity predictions, it is advised to construct ANNs that are able to predict the groundwater levels at the three reference well locations.

5

CONCLUSION

Quick and easy daily evaluation of the socio-economic drought severity can significantly improve drought management, as it puts more emphasis on the water management objective: facilitating water use. This socio-economic evaluation is, however, limited due to a lack of spatial daily groundwater data that is needed as an input to damage models. In this study, an ANN based interpolation approach to provide this data is explored. It was found that due to its accuracy, speed and easy operability, this approach can enable quick and easy drought severity evaluation.

In the designing and testing phase of the ANN, that is presented in chapter two of this report, it is found that obtaining spatial groundwater depths by interpolating well data with ANNs outperforms the currently available groundwater models, in both accuracy and speed. This conclusion holds true regardless of the hydrological functioning of the region, either a free draining or surface water controlled systems. Yet, between these differently functioning systems there are performance differences. ANNs work best for surface water controlled regions. Here a Kling Gupta Efficiency score of 0,92 and a RMSE of 0,30m was reached. Where the ANN for the free draining Twente region scored a KGE score of 0,85 and an RMSE of 0,48m. A second major finding of this first phase was that although ANNs are able to interpolate both kinds of hydrologically differently functioning systems, they are less good in doing so within one model. Based upon the input variables that are provided to the ANN, they are not able to differentiate between the two hydrological systems. For accurate ANN interpolation separate networks are thus needed.

To further study the usability of the ANNs, they have been used to evaluate the drought severity of 2019's drought in the Vechtstromen region. In this study, the differences in severity evaluation between the upper and lower confidence level of the ANN interpolated groundwater depths have been investigated for 72 drought sensitive locations. Herein, severity was expressed in terms of code green, yellow and red. For none of the 72 locations the severity evaluation between the two confidence limits differed more than one colour code. Even more, for 58 locations the colour code remained constant. The ANN based severity evaluation is therefore considered to be sufficiently reliable.

However, the case study also introduced some doubts regarding the ANN's functioning for the eastern moraines in Twente. Here, some unexpected groundwater depths and consequently severity evaluations where obtained that do not match the observations in practice. The ANN seems to underestimate the groundwater depths. This unexpected result is likely caused by two reasons. Firstly the eastern moraines might be underrepresented in the dataset. Secondly, the hydrological functioning on these moraines might differ too substantially to fit within one network. It is, therefore, advised to collect more data to construct an ANN specifically for the eastern Twente moraine.

Finally, the ANN based severity evaluation has been compared with the qualitative drought definition that has been defined in the first phase of this research project. This to study how useful the operationalised evaluation is to the decision making process. From this comparison it was concluded that the agricultural operationalisation holds a lot of potential to provide directly relevant information to guide water managing responses. The severity evaluation for agricultural locations can, therefore, take much weight in the decision making process when this potential is exploited. Evaluations for nature locations, on the other hand, need to be considered more as indicative data. This because the operationalisation links less strongly to the responsibility of water managers as defined in research phase one. Natural and human induced impacts to nature are for example not separated, while it are mostly human induced impacts that are of interest to drought management. The 'Waterwijzer' Nature is also designed to evaluate structural changes in hydrology, instead of seasonal changes that are of interest to this study. To increase the usefulness of the severity evaluations for nature, it is therefore advised to study possibilities to separate natural and human induced effects and to assess the impact of seasonal deviations in hydrology. These are, however, recommendations that require a lot of new research and model building.

All in all, it can be concluded that the combination of ANNs and damage models holds a lot of potential to evaluate drought severity quickly and easily. With some minor improvements the tool can already be useful and operational. Besides, with some more effort there is a lot of potential to improve the usefulness by linking stronger to the qualitative drought definition. Water managers are, therefore, advised to further develop and explore the application of ANNs to operationalise drought severity. This will help them to manage droughts more effectively by putting more focus on their core responsibility: facilitating water use.

REFERENCES

CHAPTER 1

Bressers, H., Bleumink, K., Bressers, N., Browne, A., Larrue, C., Lijzenga, S., Lordkipanidze, M., Özerol, G., Stein, U. (2016). The Fragmentation-Coherence Paradox in Twente. In H. B. Bressers, N.; Larrue, C. (Ed.), *Governance for Drought Resilience* (pp. 193-213). Cham: Springer International Publishing.

Chitsazan, M., Rahmani, G., & Neyamadpour, A. (2015). Forecasting groundwater level by artificial neural networks as an alternative approach to groundwater modeling. *Journal of the Geological Society of India*, 85(1), 98-106.

Chowdhury, M., Alouani, A., & Hossain, F. (2010). Comparison of ordinary kriging and artificial neural network for spatial mapping of arsenic contamination of groundwater. *Stochastic Environmental Research and Risk Assessment*, 24(1), 1-7.

Daliakopoulos, I. N., Coulibaly, P., & Tsanis, I. K. (2005). Groundwater level forecasting using artificial neural networks. *Journal of Hydrology*, 309(1-4), 229-240.

Mohanty, S., Jha, M. K., Kumar, A., & Sudheer, K. (2010). Artificial neural network modeling for groundwater level forecasting in a river island of eastern India. *Water resources management*, 24(9), 1845-1865.

Mulder, M., Hack-ten Broeke, M., Bartholomeus, R., van Dam, J., Heinen, M., van Bakel, J., Walvoort, D., Kroes, J., Hoving, I., Holshof, G. (2018). *Waterwijzer Landbouw: instrumentarium voor kwantificeren van effecten van waterbeheer en klimaat op landbouwproductie (9057738120).*

Nayak, P. C., Rao, Y. S., & Sudheer, K. (2006). Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water resources management*, 20(1), 77-90.

Nourani, V., Mogaddam, A. A., & Nadiri, A. O. (2008). An ANN based model for spatiotemporal groundwater level forecasting. *Hydrological Processes: An International Journal*, 22(26), 5054-5066.

Rigol, J. P., Jarvis, C. H., & Stuart, N. (2001). Artificial neural networks as a tool for spatial interpolation. *International Journal of Geographical Information Science*, 15(4), 323-343.

Ritzema, H. P., & Van loon-Steensma, J. M. (2017). Coping with Climate Change in a densely Populated Delta: A paradigm Shift in Flood and Water Management in The Netherlands. *Irrigation and Drainage*, 67(S1), 52-65.

Sun, Y., Kang, S., Li, F., & Zhang, L. (2009). Comparison of interpolation methods for depth to groundwater and its temporal and spatial variations in the Minqin oasis of northwest China. *Environmental Modelling & Software*, 24(10), 1163-1170.

Tielrooij, F., van Dijk, J., Blécourt-Maas, J., van den Ende, A., Oosterbaan, G. A., & Overbeek, H. J. (2000). *Waterbeleid voor de 21e eeuw*.

Trenberth, K. E. (2011). Changes in precipitation with climate change. *Climate Research*, 47, 123-138.

Witte, J., Runhaar, J., Bartholomeus, R., Fujita, Y., Hoefsloot, P., Kros, J., Mol, J., de Vries, W. (2018). *De waterwijzer natuur: instrumentarium voor kwantificeren van effecten van waterbeheer en klimaat op terrestrische natuur:* Stowa.

Yoon, H., Jun, S.-C., Hyun, Y., Bae, G.-O., & Lee, K.-K. (2011). A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology*, 396(1-2), 128-138.

CHAPTER 2

AHN. (2019). AHN3. Retrieved from: https://www.pdok.nl/introductie/-/article/actueel-hoogtebestand-nederland-ahn3

Beltman, B., & Koerselman, W. (1998). *De verdamping van moerasvegetaties.*

Bouwman. (1998). *Grondwater als leidraad voor het oppervlaktewater; Een op het* *grondwater georiënteerde aanpak voor inrichting en beheer van oppervlaktewatersystemen*. In: Dienst Landelijk Gebied en Unie van Waterschappen, DLG-publicatie.

Chitsazan, M., Rahmani, G., & Neyamadpour, A. (2015). Forecasting groundwater level by artificial neural networks as an alternative approach to groundwater modeling. *Journal of the Geological Society of India*, 85(1), 98-106.

Chowdhury, M., Alouani, A., & Hossain, F. (2010). Comparison of ordinary kriging and artificial neural network for spatial mapping of arsenic contamination of groundwater. *Stochastic Environmental Research and Risk Assessment*, 24(1), 1-7.

Condon, L. E., & Maxwell, R. M. (2015). Evaluating the relationship between topography and groundwater using outputs from a continental-scale integrated hydrology model. *Water Resources Research*, 51(8), 6602-6621.

Daliakopoulos, I. N., Coulibaly, P., & Tsanis, I. K. (2005). Groundwater level forecasting using artificial neural networks. *Journal of Hydrology,* 309(1-4), 229-240.

Davis, J. C., & Sampson, R. J. (1986). *Statistics and data analysis in geology* (Vol. 646): Wiley New York et al.

Devito, K., Creed, I., Gan, T., Mendoza, C., Petrone, R., Silins, U., & Smerdon, B. (2005). A framework for broad-scale classification of hydrologic response units on the Boreal Plain: is topography the last thing to consider? *Hydrological Processes: An International Journal*, 19(8), 1705-1714.

Droogers, P. (2009). *Verbetering bepaling actuele verdamping voor het strategisch waterbeheer: definitiestudie:* Stowa.

Freeze, R. A., & Witherspoon, P. (1967). Theoretical analysis of regional groundwater flow: 2. Effect of water-table configuration and subsurface permeability variation. *Water Resources Research*, 3(2), 623-634.

Genxu, W., Lingyuan, Y., Lin, C., & Kubota, J. (2005). Impacts of land use changes on groundwater resources in the Heihe River Basin. *Journal of Geographical Sciences*, 15(4), 405-414.

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1-2), 80-91.

Haitjema, H. M., & Mitchell-Bruker, S. (2005). Are water tables a subdued replica of the topography? *Groundwater*, 43(6), 781-786.

Hoque, M. A., Hoque, M. M., & Ahmed, K. M. (2007). Declining groundwater level and aquifer dewatering in Dhaka metropolitan area, Bangladesh: causes and quantification. *Hydrogeology Journal*, 15(8), 1523-1534.

Jansen, P. (1995). Verdamping van korte vegetaties in natte natuurgebieden. *H2O: tijdschrift voor watervoorziening en* *afvalwaterbehandeling*, 28(15), 467-471.

Knotters, M., Hoogland, T., & Brus, D. J. (2013). *Validatie van grondwaterstandskaarten met de Landelijke Steekproef Kaarteenheden (1566-7197).*

Mohanty, S., Jha, M. K., Kumar, A., & Sudheer, K. (2010). Artificial neural network modeling for groundwater level forecasting in a river island of eastern India. *Water resources management*, 24(9), 1845-1865.

Moors, E., Dolman, A., Bouten, W., & Veen, A. (1996). De verdamping van bossen. *H2O: tijdschrift voor watervoorziening en afvalwaterbehandeling*, 29(16), 462-466.

Nayak, P. C., Rao, Y. S., & Sudheer, K. (2006). Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water resources management*, 20(1), 77-90.

Nourani, V., Mogaddam, A. A., & Nadiri, A. O. (2008). An ANN-based model for spatiotemporal groundwater level forecasting. *Hydrological Processes*, 22(26), 5054-5066.

Remmelink, G., Blanken, K., van Middelkoop, J., Ouweltjes, W., & Wemmenhove, H. (2018). *Handboek Melkveehouderij*.

Rigol, J. P., Jarvis, C. H., & Stuart, N. (2001). Artificial neural networks as a tool for spatial interpolation. *International Journal of Geographical Information Science*, 15(4), 323-343.

Salvucci, G. D., & Entekhabi, D. (1995). Hillslope and climatic controls on hydrologic fluxes. *Water Resources Research*, 31(7), 1725-1739.

Scanlon, B. R., Reedy, R. C., Stonestrom, D. A., Prudic, D. E., & Dennehy, K. F. (2005). Impact of land use and land cover change on groundwater recharge and quality in the southwestern US. *Global Change Biology*, 11(10), 1577-1593.

Sun, Y., Kang, S., Li, F., & Zhang, L. (2009). Comparison of interpolation methods for depth to groundwater and its temporal and spatial variations in the Minqin oasis of northwest China. *Environmental Modelling & Software*, 24(10), 1163-1170.

Tóth, J. (1970). A conceptual model of the groundwater regime and the hydrogeologic environment. *Journal of Hydrology*, 10(2), 164-176.

van der Gaast, J., Vroon, H., & Massop, H. (2006). Verdroging veelal systematisch overschat. *H2O: tijdschrift voor watervoorziening en afvalwaterbehandeling*, 39(21), 39-43.

Wareco. (2020). *Wareco Water Data*. Retrieved from https://warecowaterdata.nl/dashboard

Wolock, D. M., Winter, T. C., & McMahon, G. (2004). Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses. *Environmental management*, 34(1), S71-S88.

Wosten, J., de Vries, F.,

Hoogland, T., Massop, H., Veldhuizen, A., Vroon, H., Wesseling, J.G., Heijkers, J., Bolman, A. (2013). *BOFEK2012, de nieuwe bodemfysische schematisatie van Nederland (1566-7197).*

Yoon, H., Jun, S.-C., Hyun, Y., Bae, G.-O., & Lee, K.-K. (2011). A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology*, 396(1-2), 128-138.

CHAPTER 3

Besse-Lototskaya, A. A., Geertsema, W., Griffioen, A., van der Veen, M., & Verdonschot, P. F. M. (2011). *Natuurdoelen en klimaatverandering*.

Bij12 (2020). *Natuurtypen*. Retrieved from https://www.bij12.nl/onderwerpen/natuur-en-landschap/index-natuur-en-landschap/natuurtypen/

Bouwman. (1998). *Grondwater als leidraad voor het oppervlaktewater; Een op het grondwater georiënteerde aanpak voor inrichting en beheer van oppervlaktewatersystemen*. In: Dienst Landelijk Gebied en Unie van Waterschappen, DLG-publicatie.

Brouwer, C., Prins, K., & Heibloem, M. (1989). *Irrigation water management: irrigation scheduling. Training manual, 4.*

Finke, P., Brus, D., Bierkens, M., Hoogland, T., Knotters, M., & De Vries, F. (2004). Mapping groundwater dynamics using multiple sources of exhaustive high resolution data. *Geoderma*, 123(1-2), 23-39.

Mulder, M., Hack-ten Broeke, M., Bartholomeus, R., van Dam, J., Heinen, M., van Bakel, J., Walvoort, D., Kroes, J., Hoving, I., Holshof, G. (2018). *Waterwijzer Landbouw: instrumentarium voor kwantificeren van effecten van waterbeheer en klimaat op landbouwproductie:* Stowa.

Runhaar, H., & Hennekens, S. (2015). *Hydrologische randvoorwaarden natuur: gebruikershandleiding (waternoodapplicatie versie 3).*

Steduto, P., Hsiao, T. C., Raes, D., & Fereres, E. (2009). AquaCrop—The FAO crop model to simulate yield response to water: I. Concepts and underlying principles. *Agronomy Journal*, 101(3), 426-437.

Van Bakel, J., Huinink, J., Prak, H., & van der Bolt, F. (2005). *HELP-2005, uitbreiding en actualisering van de HELP-tabellen ten behoeve van het Waternood-instrumentarium:* Stowa.

Van Diepen, C. v., Wolf, J., Van Keulen, H., & Rappoldt, C. (1989). WOFOST: a simulation model of crop production. *Soil use and management*, 5(1), 16-24.

Wilhite, D. A., & Glantz, M. H. (1985). Understanding the Drought Phenomenon: The Role of Definitions. *Water International,* 10(3), 111-120.

Witte, J., Runhaar, J., Bartholomeus, R., Fujita, Y., Hoefsloot, P., Kros, J., Mol, J., de Vries, W. (2018). *De waterwijzer natuur: instrumentarium voor kwantificeren van effecten van waterbeheer en klimaat op terrestrische natuur:* Stowa.

CHAPTER 4

Beltman, M. R. (2019). *Defining Drought - towards a functional definition of drought for the Vechtstromen Water Authority.* (MSc.), University of Twente, Enschede.

Chitsazan, M., Rahmani, G., & Neyamadpour, A. (2015). Forecasting groundwater level by artificial neural networks as an alternative approach to groundwater modeling. *Journal of the Geological Society of India*, 85(1), 98-106.

Daliakopoulos, I. N., Coulibaly, P., & Tsanis, I. K. (2005). Groundwater level forecasting using artificial neural networks. *Journal of Hydrology*, 309(1-4), 229-240.

Mohanty, S., Jha, M. K., Kumar, A., & Sudheer, K. (2010). Artificial neural network modeling for groundwater level forecasting in a river island of eastern India. *Water resources management*, 24(9), 1845-1865.

Mulder, M., Hack-ten Broeke, M., Bartholomeus, R., van Dam, J., Heinen, M., van Bakel, J., Walvoort, D., Kroes, J., Hoving, I., Holshof, G. (2018). *Waterwijzer Landbouw: instrumentarium voor kwantificeren van effecten van waterbeheer en klimaat op landbouwproductie (9057738120).*

Nayak, P. C., Rao, Y. S., & Sudheer, K. (2006). Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water resources management*, 20(1), 77-90.

van Loon, A. F., Stahl, K., Baldassarre, G. D., Clark, J. R., S., Wanders, N., Gleeson, T., van Dijk, A.I.J.M., Tallaksen, L.M., Hannaford, J., Uijlenhoet, R., Teuling, A.J., Hannag, D.M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., van Lanen, H. A. J. (2016). Drought in a human-modified world: reframing drought definitions, understanding, and analysis approaches. *Hydrology and Earth System Sciences*, 20, 3631-3650.

Witte, J., Runhaar, J., Bartholomeus, R., Fujita, Y., Hoefslo ot, P., Kros, J., Mol, J., de Vries, W. (2018). *De waterwijzer natuur: instrumentarium voor kwantificeren van effecten van waterbeheer en klimaat op terrestrische natuur:* Stowa.

Yoon, H., Jun, S.-C., Hyun, Y., Bae, G.-O., & Lee, K.-K. (2011). A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology*, 396(1-2), 128-138.

Mark Beltman - 2020