



UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

APPLICATION OF MACHINE LEARNING TECHNIQUES IN PREDICTING GROUNDWATER LEVELS AND DISCHARGE RATES IN THE NORTH WEST AQUIFERS

A Master of Science Dissertation submitted to the Faculty of Science,
The University of the Witwatersrand, Johannesburg,
in fulfillment of the requirements for the degree of Master of Science by Dissertation
in Computer Science.

Yolanda Kanyama
2306572

Supervised by: Dr. Ritesh Ajoodha
and
Prof. Hlanganani Tutu

School of Computer Science and Applied Mathematics
The University of the Witwatersrand, Johannesburg

April 22, 2021

Abstract

The dolomite aquifers in the North West province have suffered from severe groundwater over-abstraction causing groundwater levels to decline. This has threatened water supply security and caused municipal and irrigation boreholes to dry up in the province. The continuous decline in groundwater levels in the region will inevitably lead to increased costs and risks to the sustainable development of groundwater resources.

The key principles underpinning water resource management in South Africa water law are sustainability of groundwater resources. Accurate prediction of groundwater levels and discharge rates is essential to the sustainable utilisation and management of groundwater resources in the Grootfontein Aquifer. In South Africa, there is a limited body of literature in the forecasting of both groundwater levels and discharge rates using machine learning models.

The objectives of this study were therefore (1) to present ranking by entropy determining the primary factors governing spring discharge patterns and groundwater levels using mutual information (2) to develop a robust model that can be used to forecast future discharge rates and groundwater levels in the Grootfontein aquifer, (3) to predict groundwater levels under adjusted inputs (rainfall and abstraction amounts) and (4) to test model generalizability by using the selected optimal model to predict groundwater levels in a different aquifer setting.

Results through mutual information showed that regional abstraction, discharge and precipitation were the highest contributing factors to predict both groundwater levels and discharge rates in the karst aquifer. An interesting observation came to light as the study highlighted decomposed signals of rainfall produced higher entropy scores (+0.30) and better prediction for the machine learning models as compared to actual measured rainfall. Temperature proved to have little to no influence in the prediction of both groundwater levels and discharge rates.

The research results obtained for the predictive models showed that the long-short term memory model was better suited to predict both groundwater levels and discharge rates in the aquifer. Scenario testing results of groundwater levels showed a decline in groundwater levels when abstraction amounts were doubled and a more subtle decline when rainfall peaks amounts were reduced. Moreover, during what can be conspired as drought years (worst case scenario) the groundwater levels in the compartment decreased by almost 2m. These results indicated the importance the chosen variables had on groundwater levels in the region further strengthening the entropy results. The results would also be useful to both hydrologists and water managers in aquifer management.

For discharge prediction, two case scenarios were tested. In both scenarios tested the gated recurrent unit and long-short term memory models were at par with one another. The results showed that including groundwater levels from surrounding boreholes drastically improved the prediction of discharge. In literature, the focus has been on using weather parameters and past discharge as model inputs, however this study showed that groundwater levels inclusion had a much better effect on the modelled results. Hence, we propose the use of groundwater levels, precipitation and abstraction to not only infill incomplete discharge records but to generate discharge data where there are no records at all.

Lastly, to test model generalization, the Steenkoppies aquifer was used as a case study to predict groundwater levels in the aquifer. The best performing model for groundwater level prediction was used with the same model pipeline as that used for the Grootfontein. The results from this experiment proved the capabilities of the long-short term memory model in predicting groundwater levels in a different setting. The model was able to capture the trends of the

observed dataset better than the other models tested. The results of the model simulations in this experiment highlighted the ability of the long-short term memory model to capture the broad patterns of the Steenkoppies data so we can assume, through this experiment that, the long-short term memory model is likely to be able to predict groundwater levels from other regions.

The results presented in this research study present a useful approach for discharge and groundwater prediction. The frameworks used could be used by water managers and municipalities to manage water resources more effectively during drier years while ensuring that groundwater levels do not deplete. Drought warnings and water restrictions can also be issued promptly.

Published Work

Some of the work in this dissertation appears in the following peer-reviewed publications:

[Kanyama *et al.* 2020b] **Yolanda Kanyama, Ritesh Ajoodha, Helen Seyler and Hlanganani Tutu.** *Application of Artificial Neural Networks to Forecast River Discharge Rates: A Case Study of the Grootfontein Aquifer.* In *7th Asia-Pacific Conference on Computer Science and Data Engineering*, Melbourne, Australia, 2020.

[Kanyama *et al.* 2020a] **Yolanda Kanyama, Ritesh Ajoodha, Helen Seyler, Ndivhuwo Makanda and Hlanganani Tutu.** *Application of Machine Learning Techniques in Forecasting Groundwater Levels in the Grootfontein Aquifer.* In *2020 2nd International Multidisciplinary Information Technology and Engineering Conference*, pages 167–174, South Africa, 2020.

[Kanyama *et al.* 2021] **Yolanda Kanyama, Ritesh Ajoodha, Helen Seyler, and Hlanganani Tutu.** *Groundwater Level Estimation using Recurrent Neural Networks: A Case Study of the Grootfontein Aquifer.* In *Interdisciplinary Research in Technology Management*, Kolkata, India, 2021.

Acknowledgements

I would like to thank God for his unfailing love that strengthened me to complete this dissertation.

Throughout the writing of this dissertation I have received a great deal of support and assistance.

I want like to thank my supervisors, Dr Ritesh Ajoodha and Prof. Hlanganani Tutu, for their support and guidance throughout the course of my dissertation. Their comments, remarks and engagement through our weekly meetings were a great learning experience for me. Their insightful feedback pushed me to sharpen my thinking and brought my work to a higher level.

I would like to acknowledge Helen Seyler who was of great help throughout the course of this dissertation. Her expertise in the field of hydrology was invaluable in formulating the research questions. I would also like to thank my direct mentor during my internship at IBM research lab, Ndivhuwo Makanda, whose comments and remarks were of great assistance during my internship.

In addition, I would like to recognize the WRC funding awarded for my master's degree, without which it would not be possible to do this research: the Water Research Commission project (K5/2879), which forms part of a Big Data and Transboundary Water Collaboration programme funded by the Department of Science and Technology, USAID and SADC-GMI, supported by IBM Africa Lab and USGS, and managed by the WRC and the Sustainable Water Partnership.

Finally, I would like to thank my mother, who has supported me throughout this entire process. I am forever grateful for her encouragement and support.

Thank you all

Declaration

I, Yolanda Kanyama (student number: 2306572), hereby declare the contents of this dissertation to be my own work. This dissertation is submitted for the degree of Master of Science in Computer Science at the University of the Witwatersrand. This work has not been submitted for any other degree or to any other institution.

Ykanyama

Yolanda Kanyama, April 22, 2021

Johannesburg, South Africa

Contents

1	Introduction	1
1.1	Introduction	1
1.2	Problem Statement	2
1.3	Overview of Literature	2
1.4	Research Methodology	3
1.5	Novelty and Contribution	5
1.6	Document Structure	6
2	Related Work	7
2.1	Introduction	7
2.2	Application of Groundwater Models in Modeling Groundwater Systems	8
2.3	Application of Machine learning Models in Hydrology	10
2.4	Discussion and Conclusion	12
3	Research Framework Overview	14
3.1	Introduction	14
3.2	Research Motivation	14
3.3	Study Area	15
3.4	Experimental Data Sources	15
3.4.1	River Discharge Data	16
3.4.2	Groundwater Level Data	16
3.4.3	Groundwater Usage Data	17
3.4.4	Weather Data	19
3.5	Experimental Design	20
3.5.1	Artificial Neural Networks	20
3.5.2	Support Vector Machines	23
3.5.3	Random Forest	24
3.6	Information Gain Ranking	24
3.7	Model Performance Analysis and Evaluation	25
3.8	Discussion and Conclusion	26
4	Groundwater Level Forecasting	27
4.1	Introduction	27
4.2	Model Setup	27
4.3	Experimental Results	29
4.3.1	Information Gain Ranking	29
4.3.2	Model Results	31
4.3.3	Discussion and Conclusion	38

4.4	Scenario Testing Results	39
4.4.1	Scenario 3c	44
4.4.2	Scenario 4	45
5	Spring Discharge Forecasting	47
5.1	Introduction	47
5.2	Model Setup	48
5.3	Experimental Results	48
5.3.1	Information Gain Ranking	48
5.3.2	Model Results	50
5.4	Discussion and Conclusion	51
6	Application of Learned Model	52
6.1	Introduction	52
6.2	Data Acquisition	52
6.3	Experimental Results	53
6.4	Discussion and Conclusion	53
7	Conclusion and Future Directions	55
7.1	Introduction	55
7.2	Literature Review	55
7.3	Research Framework	56
7.4	Research Findings	56
7.4.1	Groundwater Level Prediction	56
7.4.2	Discharge Prediction	57
7.4.3	Application of Learned Model	57
7.5	Future research	58
7.6	Conclusion	58

List of Figures

1.1	This diagram illustrates the study flow chart from attribute selection until the final predictive model is chosen for a) groundwater levels b) discharge rates. . . .	4
2.1	Schematic cross-section of a karst aquifer	7
3.1	Compartments of the North West dolomite aquifers	15
3.2	Monthly gauge data for stations D4H013 and D4H014	16
3.3	Time series of selected boreholes for modelling purposes	17
3.4	Analysis of linear trend for 24 boreholes selected for modelling, along with rainfall	18
3.5	Cumulative use based on WARMS licence registrations	19
3.6	(a) Monthly rainfall from three weather stations around Grootfontein (b) Monthly temperature from three weather stations around Grootfontein	20
3.7	An example of a Recurrent Neural Network. The hidden state is meant to carry important information from one input item in the series to others	21
3.8	General architecture of a FFNN	22
3.9	Support Vector Machine	23
4.1	Flowchart Summarising ANNs model approach	28
4.2	Bar graph showing model averages	35
4.3	Model prediction for scenario 4, long term prediction for borehole D4N0037 . . .	46
5.1	Hydrographs of the spring discharge at Molopo in relation to (a) precipitation and (b) subset of groundwater levels.	48
5.2	Foretasted discharge rate over a 2-year period	51

List of Tables

2.1	Features that need to be calculated to accurately solve the groundwater flow equation	9
2.2	Table showing authors' contributions using different machine models to predict groundwater levels and discharge rates.	11
3.1	Summary of pre-processing steps on groundwater level dataset	17
3.2	Summary of registered groundwater use in Grootfontein (WARMS, DWS)	19
3.3	Mean, minimum and maximum annual rainfall for stations around Grootfontein .	20
4.1	A ranking by entropy of each input feature against each target variable	30
4.2	Performance of developed models during testing phase for all 24 boreholes selected for modelling in the Grootfontein compartment	34
4.3	Time-series of actual vs predicted groundwater levels for 6 selected boreholes in the Grootfontein aquifer	36
4.4	Description of each case scenario	40
4.5	Model prediction for scenario 1, decrease rainfall peaks	41
4.6	Model prediction for scenario 2, Increase groundwater abstraction	42
4.7	Model prediction for the worst case scenario, increase abstraction + decrease rainfall peaks	43
4.8	Model prediction for average case scenario, decrease events by half	44
4.9	Model prediction for scenario 1b, increase rainfall peaks	45
5.1	A ranking of the mutual information for a subset of features to predict discharge	49
5.2	Experimental results of the different model and scenarios tested	50
6.2	Summary of model results	53
6.1	Time-series of actual vs predicted groundwater levels for 3 selected boreholes in the Steenkoppies aquifer	54

Chapter 1

Introduction

1.1 Introduction

In many countries, groundwater plays a crucial role in supplying water to a significant part of the population for industrial, agricultural and drinking purposes. One of the most invaluable resources in semi-arid countries such as South Africa is groundwater [Modisha 2017]. In South Africa, groundwater resources provide relatively drought-proof water supplies for the mining, agricultural and tourism industries as well as for domestic use [Villholth *et al.* 2013]. However, due to extreme changes in weather conditions and the increase in population resulting in higher groundwater abstraction rates (water taken out of the aquifer), groundwater resources have been under enormous stress [Villholth *et al.* 2013].

The Grootfontein dolomite aquifer located in the North West Province is one of the most important aquifers in South Africa [Cobbing 2018]. It serves as a key water supply source for irrigation and municipal water supplies. The aquifer is also ecologically important in supplying springs that feed important rivers. Several wetlands, ecosystems and towns such as Mahikeng, Grootpan, Lichtenburg, among many others rely on the aquifer as their main source of water [Cobbing 2018].

However, during 2017, the Grootfontein aquifer was said to have fallen more than 28m below ground level due to groundwater over-abstraction [Cobbing 2018]. Over-abstraction happens when the water demand exceeds the amount available during a certain period and as a result groundwater resources are over-exploited to meet the growing water demand. The over-abstraction of the Grootfontein aquifer led to some of the municipal boreholes failing and also threatened some of the irrigation boreholes in the North West [Cobbing 2018]. Several other water supplies for both domestic and environmental services were severely threatened.

Aquifers represent a critically important water resource world-wide. They act as natural reservoirs of water, storing large quantities of water over a long period. This makes them important sources of water especially during drought years. Prediction of groundwater levels and discharge rates therefore becomes an imperative step to achieving sustainable groundwater management. Predictions will not only provide means of future preparation but will also assist in understanding how changes in the different parameters can affect groundwater levels and discharge rates in the region further strengthening the course of groundwater sustainability.

1.2 Problem Statement

According to a report published by the Department of Environmental Affairs of South Africa, the country will experience an increase in the number of heat-wave days and low rainfall days [Department of Environmental Affairs 2019]. Furthermore, South Africa was also recently placed 30th as the driest country in the world [Basson *et al.* 2018]. It is considered a high water-stressed country, with extreme temperature and rainfall fluctuations [Luo *et al.* 2015]. The Department of Environmental Affairs published a report in 2019 outlining the vulnerability of key socio-economic sectors in South Africa due to weather changes including high water demand which is highly problematic as current water usage in South Africa already exceeds reliable yield or water available [DWS 2017].

Due to the dynamic nature of groundwater systems they continually adjust to long-term and short-term changes in land use, groundwater abstraction and climate change. Traditionally, groundwater flow models have been the basis for the quantitative analysis of groundwater resources [Coppola *et al.* 2005]. These models are used to predict the effects of hydrological changes on the behaviour of the aquifer system [Coppola *et al.* 2005].

The computation done in groundwater flow models is based on the mathematical relationship used to describe the flow of groundwater through an aquifer system [Chen *et al.* 2013]. To solve the equation inputs such as the hydraulic head, flow direction, storage capacity and hydraulic conductivity must be predefined (further discussed in chapter 2). Not only is the information challenging to obtain but it is also an expensive and time consuming process [Coulibaly *et al.* 2001]. Extensive physical observations will need to be performed by the hydrologist to collect data for all the different parameters that will lead to solving the equation. Calibrating the groundwater flow model therefore requires a vast knowledge of a wide variety of fields such as geology, hydraulics, hydrology, mathematics, and physics to construct it [Martin *et al.* 2016].

An alternative to these models is the use of data-driven methods with machine learning techniques. Using historical data, machine learning models can identify patterns and make better decisions by learning information directly from the data. Data driven models can capture the trends and time-variant behaviour of hydrological systems without the deep knowledge of the underlying physical attributes that groundwater flow models would need [Steyn and Brink 2018].

Machine learning models, therefore, serve as attractive alternatives to process-based models since they have the potential to be less laborious and also the ability to provide useful results even with fewer data samples. Proper water resource planning and management require accurate predictions of both discharge and groundwater levels. This study therefore proposes the use of machine learning models as opposed to groundwater flow models to predict both groundwater levels and discharge rates in the Grootfontein dolomite aquifer.

1.3 Overview of Literature

As highlighted previously, groundwater flow models have proved to be challenging to develop, mainly due to their large hydrogeological and meteorological data requirements. Researchers have proved the capabilities of machine learning models against groundwater flow models. Some of these studies are outlined below.

Recently, Chen *et al.* [2020] conducted a comparative study between machine learning models and groundwater flow models to simulate groundwater dynamics in the Heihe River Basin in northwestern China. MODFLOW (a computer code that solves the groundwater flow equation)

was used to simulate the groundwater dynamics together with three different machine learning algorithms, namely; support vector machines, multiplayer perceptron and, radial basis function network. Results revealed that the accuracy of the machine learning models was significantly better than the groundwater flow model. While the groundwater flow model achieved a coefficient of determination score of 0.51 the multilayer perceptron, radial basis function network and support vector machine achieved R^2 scores of 0.71, 0.75 and 0.76 respectively. The study therefore concluded by recommending machine learning models as a more cost effective method for accurate prediction of groundwater resources.

Furthermore, Chitsazan *et al.* [2015] also did a comparative study between a feed forward neural network model and the groundwater flow model to simulate groundwater levels in Aghili plain, urban area of Gotvand in southwest Iran. A groundwater flow model had previously been developed to predict groundwater levels in the area therefore the accuracy of their machine learning model was quantified by comparing it to the groundwater flow model. Within this context, the results of the artificial neural network were compared with the results of the groundwater flow model. The plots/graphs in the study showed that the predicted groundwater levels by the artificial neural network better fitted the observational data trend than the groundwater flow model. Hence, the study concluded that artificial neural networks were better suited to predict groundwater levels in the region than the previously computed groundwater flow model.

Other studies have reported impressive predictive accuracy of different machine learning models in modelling hydrological systems when compared to groundwater flow models. This has resulted in an increase of machine learning models in modelling hydrological systems since these models offered vast benefits when compared to groundwater flow models.

Models such as support vector machines [Liong and Sivapragasam 2002; Behzad *et al.* 2010], artificial neural networks [Sreekanth *et al.* 2009; Mirzavanda *et al.* 2014], decision trees [Al-Aboodi *et al.* 2016; Gizaw and Gan 2016], multiple linear regression models [Tsakiri *et al.* 2018; Sahoo and Jha 2013] among several others have been widely used in the study of hydrological systems. Neural networks and support vector machines were favoured due to their good generalization capabilities and ability to handle non-linear data. On the other hand, multiple linear regression and decision tree models were recommended as a more cost-effective alternative in predictive modelling of hydrological systems considering the practical advantages these models offered such as the easy interpretation of the results between the parameters.

In this study, we therefore, test the use of multiple data driven models to select the optimal model in forecasting groundwater levels and discharge rates in the Grootfontein aquifer.

1.4 Research Methodology

This study made use of data obtained from the Department of Water and Sanitation (DWS), The Water Authorization and Registration Management System (WARMS) and The South African Weather Service (SAWS). Monthly groundwater levels and discharge rates were obtained from DWS while the abstraction dataset was obtained from WARMS and monthly rainfall and temperature datasets were obtained from SAWS. Not all boreholes that were used to measure groundwater levels and abstraction rates are still active today and the same applies for stations used to measure discharge rates. Hence, the datasets were filtered to select boreholes, discharge stations and weather stations with the longest continuous time-periods and with minimal data missing.

Furthermore, input parameters are not equally informative in predicting the target variable.

This is because some parameters may be noisy, correlated or have an insignificant relationship with the target variable. To understand the importance of each input driver in relation to groundwater levels and discharge rates, this study investigated the impact of groundwater abstraction through pumping and climate change on both groundwater levels and discharge rates through Mutual Information (MI). Using MI this study quantified the relationship between each input variable and the target variable. MI proved to be a useful framework for feature selection and ultimately improved model results.

Using the top features selected from MI, different machine learning models were tested to find the optimum model that best modelled our data and made the most accurate predictions. The feed forward neural network (FFNN) was used as the base model while the results from the random forest (RF), support vector regression (SVR), gated recurrent unit (GRU) and long short term memory (LSTM) models were compared against the base model. It should be noted that the top 3 models for groundwater level prediction were selected for discharge prediction since those models proved to have great capabilities in predicting hydrological systems in the region.

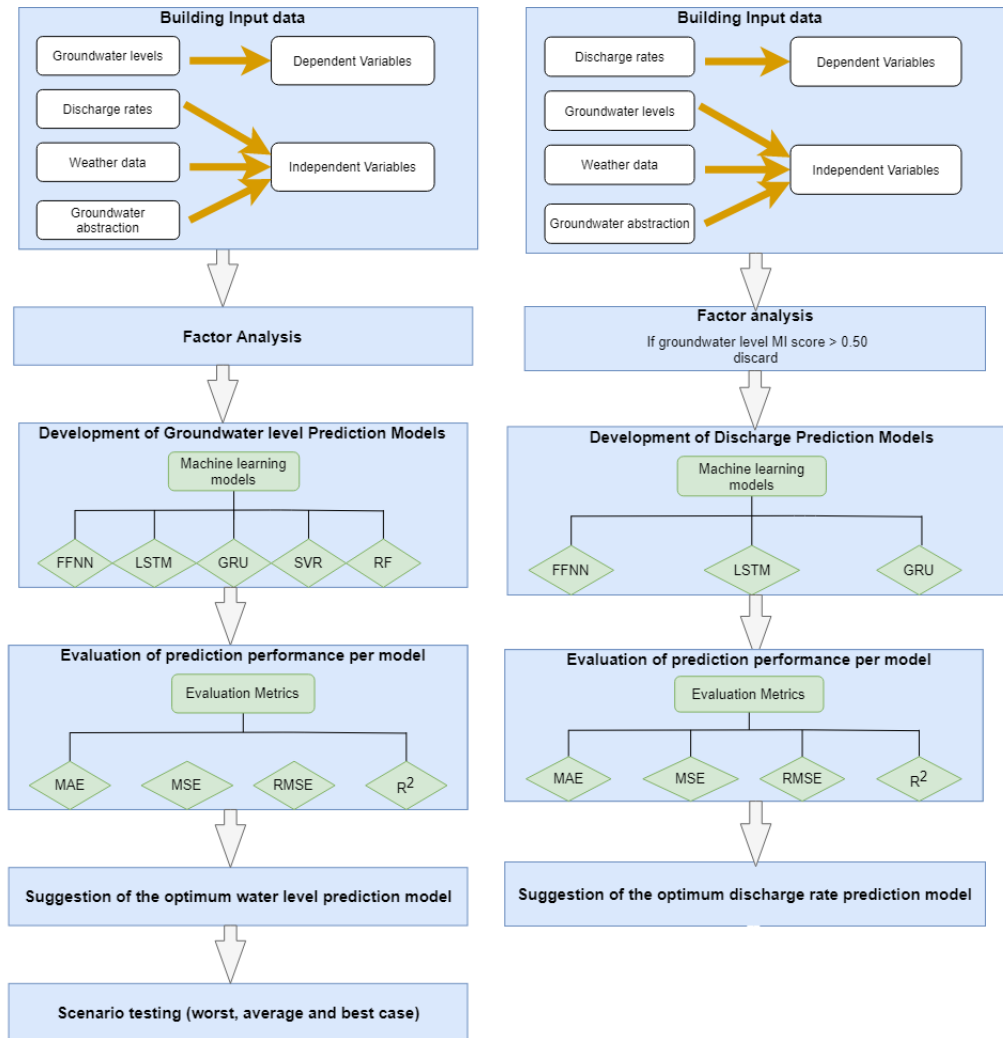


Figure 1.1: This diagram illustrates the study flow chart from attribute selection until the final predictive model is chosen for a) groundwater levels b) discharge rates.

Using the optimal model selected for groundwater level prediction, we also tested different sce-

narios to represent the worst case, average case and best case scenario. In South Africa, scenario testing is mainly performed using numerical models. However, this method is time consuming and costly hence we suggested the use of a machine learning models to perform scenario tests. Numerous studies have also successfully predicted groundwater levels using different machine learning models, however, the way to take into account different case scenarios under specific weather conditions i.e drought conditions or using specific groundwater abstraction rates i.e. over-demand is not well apprehended using machine learning models. Necessities are more into simply predicting groundwater levels. Hence, we suggest the use of the modelling framework used in this study as an alternate approach to simulating groundwater level change specifically in areas where subsurface properties are not known.

Finally, this study also conducted a model generalizability test. To prove model generalizability, the Steenkoppies aquifer located in the Gauteng region was selected. Three boreholes were chosen for the case study which had distinct characteristics from those in the Grootfontein aquifer. This was done to further strengthen the results of the modelling technique recommended in the study.

A summary of steps followed for each prediction is provided in figure 1.1.

1.5 Novelty and Contribution

Accurate prediction of groundwater levels is essential to the sustainable utilisation and management of groundwater resources in the Grootfontein Aquifer. In South Africa, there is a limited body of literature in forecasting of both groundwater levels and discharge rates using machine learning models. The aim of this study is therefore to determine the primary factors governing spring discharge patterns and groundwater levels using MI and to develop a model that can be used to forecast future discharge rates and groundwater levels in the Grootfontein aquifer.

More specifically, this research aims at making the following three contributions:

1. Predict groundwater levels in Grootfontein dolomite aquifer
 - present a ranking by entropy of which feature has the most influence on groundwater levels in the dolomite aquifer
 - present an optimal predictive model that can be used to forecast groundwater levels in the aquifer
 - test different scenarios to represent worst, average and base case
2. Predict discharge rates in Grootfontein dolomite aquifer
 - present a ranking by entropy of which feature has the most influence on spring discharge in the dolomite aquifer
 - present an optimal predictive model that can be used to forecast spring discharge
 - test multi-step forecasting of spring discharge
3. Provide evidence of the effectiveness of our chosen method by applying the model to predict groundwater levels in a different karst aquifer to test model generalizability.

As previously highlighted, robust groundwater level estimates and spring discharge estimates are primary requirements for effective management of water resources and sustainable use of groundwater.

In literature several studies have successfully predicted groundwater levels using different machine learning models. However, it is also evident in literature that machine learning models face several drawbacks. Therefore, this study aims at training a generalized model (through rigorous testing) for prediction of groundwater levels by applying the recommended model in a different aquifer setting to prove the generalizability and effectiveness of the model. To do this, we firstly, need to identify the factors that influence groundwater level change and how the machine learning models adjust to the different input parameters fed. This will assist us in identifying the most prominent feature sets and also the most prominent parameter settings for the model to accurately predict groundwater levels.

Moreover, In South Africa, there is a limited body of literature in the forecasting of discharge using machine learning models. This could be because there are limited gauge stations recording discharge data hence data is also limited or in some cases data is not available at all. In literature the way to take into account the temporal behaviors of the groundwater levels in the surrounding area is not well apprehended and necessitates are more into using weather parameters and past discharge rates as inputs. We therefore propose the use of groundwater levels to not only predict discharge data but to generate spring discharge where there are insufficient records in an area.

These aforementioned methodologies will allow for the effective management of groundwater resources in the area. For example, appropriate mitigation measures such as: issuing timely drought warnings or water restriction declarations can be taken in a timely manner effectively leading to sustainable use of the groundwater resources in the province.

1.6 Document Structure

This dissertation is structured as follows. Chapter 2 provides relevant related work which served as the foundation for the machine learning models chosen in this study. The chapter discusses the definition of the groundwater flow models, how they are used, why they are used and also the pitfalls to using such models and why machine learning models would be a better alternative. Chapter 3 presents the research framework followed including a description of the study area, data source, models chosen and evaluation metrics used. Chapter 4 provides the results and discussion of the groundwater level predictions. Chapter 5 provides the results and discussion of the discharge predictions and Chapter 6 discussed the application of the learned model in a different aquifer setting. Chapter 7 rounds off with concluding remarks and future directions.

Chapter 2

Related Work

2.1 Introduction

Groundwater is water that is stored beneath the earth's surface. It is stored in pores and fractures of soil and rock. Rocks with well connected spaces allow water to easily flow through them. Such rocks are often referred to as permeable rocks and these rocks make aquifers [Barnett *et al.* 2012]. There are also different rock formations. Rocks can be fractured, dissolved, folded and cemented differently depending on the materials they contain [WRC 2003]. This means that rocks store and allow the flow of groundwater in different ways.

In this study, we made use of data from a karst aquifer (relatively soft rock). Karst aquifers are a special type of fractured rock made of limestone and magnesite [Barnett *et al.* 2012]. Water present in karst aquifers will dissolve the soft rock thereby significantly increasing the size of fractures making groundwater flow more concentrated than other aquifers. Figure 2.1 illustrates a cross section example of a karst aquifer.

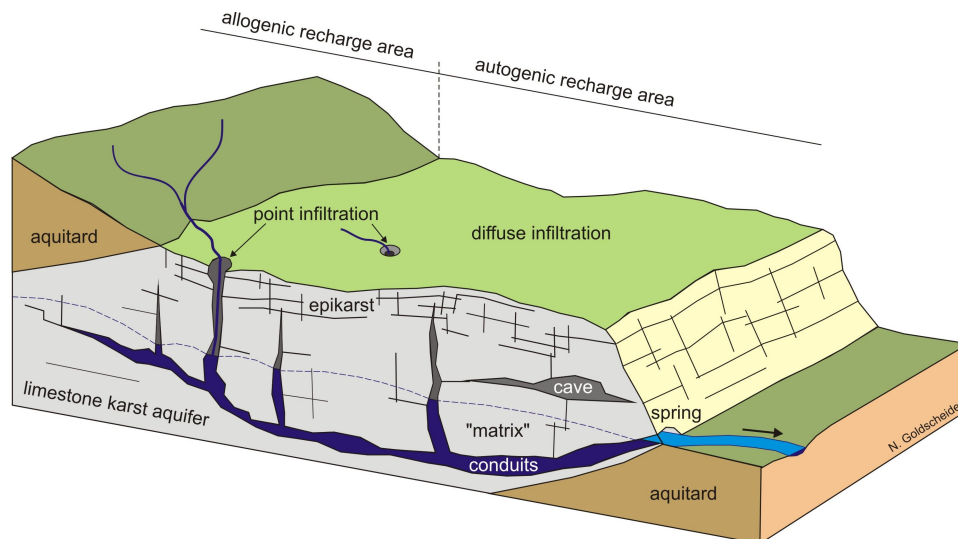


Figure 2.1: Schematic cross-section of a karst aquifer

As evidenced in figure 2.1, the aquifer is recharged when water seeps through the aquifer by infiltration. Recharge can be allogenic (recharge sourced from runoff) or autogenic (recharge sourced entirely from rainfall that falls within the karst area). Infiltration will continue until

the aquifer is fully saturated meaning relatively all pores and fractures within the aquifer are filled with water. Outflows from the aquifer can be through spring discharge points, evapotranspiration or groundwater pumping.

The goal of this chapter is to review and highlight the main concepts, theories and empirical findings relating to the prediction of discharge rates and groundwater levels in current literature. Literature from different sources including journals, books, research reports and online articles were reviewed. The sections below will discuss how groundwater flow models have been used to model hydrological systems, the disadvantages of using these models and why machine learning techniques would be a more viable solution to model hydrological systems. The chapter will focus on the current machine learning trends that are being used to model hydrological systems and end of with a conclusion and discussion section which summarises the chapter.

2.2 Application of Groundwater Models in Modeling Groundwater Systems

Groundwater flow models are computer models of groundwater flow systems that are used to simulate and predict aquifer conditions [Chen *et al.* 2013]. For example, the models can be used to predict the effects of hydrological changes (like groundwater use) on the behaviour of the aquifer [Chen *et al.* 2013]. Groundwater flow models have played an important role in the development and management of groundwater systems in different parts of the world by monitoring groundwater changes and analysing subsurface hydrodynamics. By providing insight into the complex groundwater system behaviour a conceptual model of the aquifer system is developed.

The conceptual model is the idealised representation (a picture) of the hydrologists hydrological understanding of the key flow processes of the system [Barnett *et al.* 2012]. Generally the area to be modelled is divided into smaller regions to get as much information about the aquifer as possible [Barnett *et al.* 2012]. The development of the conceptual model alone needs to include people with a range of skills including hydrology, geology, modelling, environmental systems and climate [Barnett *et al.* 2012].

During model conceptualization, the hydrologist will pre-define different parameters (mentioned in table 2.1) which will assist in solving the groundwater flow mathematical equation through a series of extensive physical tests of the aquifer system. It should be noted that expert judgement is needed to arrive at representative values hence the hydrologist is tasked with collecting this data. Only after these parameters have been calculated can the groundwater flow equation be solved. Furthermore, each parameter indicated in table 2.1 is calculated based on certain mathematical equations and derivations of it's own [Hiscock and Bense 2005]. Table 2.1 shows some of the most important parameters as mentioned by Hiscock and Bense [2005] and Rushton [2005].

Two fundamental principles govern the groundwater flow mathematical equation: Darcy's law, and the equation of continuity that describes the conservation of fluid mass flow through a porous material. The groundwater flow equation shown in equation 2.1 is used to estimate the direction of groundwater flow (or the distribution of hydraulic heads).

$$\frac{\partial}{\partial x} \left(K_x \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_y \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_z \frac{\partial h}{\partial z} \right) - W = S_s \left(\frac{\partial h}{\partial t} \right), \quad (2.1)$$

Table 2.1: Features that need to be calculated to accurately solve the groundwater flow equation

Parameter	Explanation
Rock type and porosity	Different rock types have different porosity volumes. The amount of water a rock can hold is directly related to the porosity since water will try and fill up the empty spaces in the rock. Groundwater moves through the porosity available to flow hence these parameters are used to calculate the hydraulic conductivity.
Location of aquifer	This will assist in knowing where the aquifer starts and ends which is important when determining the boundary conditions (boundary conditions represent the physical boundaries of the aquifer).
Depth (thickness) and size of aquifer	By Knowing the depth and size the hydrologist can estimate the storage capacity of the aquifer.
Permeability	This is a quantitative description of the ability of rock structures or other materials to transmit water. This is also used to calculate the hydraulic conductivity.
Inflow and outflows	Inflow is water entering the aquifer and outflows is water leaving the aquifer. Measuring this will assist in knowing the height at which water will rise or fall (groundwater head, h).
Groundwater head (h)	This is the height with which the water will rise in an aquifer relative to the consistent datum. The groundwater head will also be used to calculate the direction of flow (water flows from locations of high h to locations of low h).
Hydraulic conductivity (K)	This is a quantitative measure of the ability of a rocks to transmit water.
Storativity (S_s)	The amount and distribution of water in the aquifer. This tells us how much water is stored in the aquifer.

where K_x , K_y and K_z represent values of hydraulic conductivity along the x, y, and z coordinate axes; h is the hydraulic head which can be converted to groundwater level; W represents source and/or sink with $W < 0$ representing water flowing out of the system (outflows), and $W > 0$ representing water flowing into the system (inflows); S_s denotes the specific storage of the aquifer which is multiplied by the rate of change of the groundwater head (h) with time (t).

The left side of the equation represents flows in the three coordinate directions, while due to the change in groundwater head over time, the right side represents the storage contribution. The solution to this equation describes the value of the hydraulic head at any point in a 3 dimensional flow field [Ayenew *et al.* 2008].

Groundwater flow models have been widely used to model complex groundwater systems [Ayenew *et al.* 2008]. Successful assessment of groundwater systems heavily depends on the reliable groundwater model simulation [Zadeh 2005]. However, constructing these models largely relies on the hydrologists knowledge of the aquifer system [Wu and Zeng 2013]. It requires them to define many additional spatial data on the hydrological and geological properties of the aquifer under study. Collecting this data is time consuming and expensive because of the complexity of groundwater systems [Gumbrecht and Thunvik 1997].

In addition, hydrogeological data is always limited in the field because not all parameters can be measured due to the complexity of aquifer systems. This at times forces modellers to use literature values to calculate variables such as porosity, storativity, inflows and outflows of the system [Wu and Zeng 2013]. This alone places a great responsibility on the modeller as they will have to evaluate the implications of adopting values that are not site specific. The model accuracy therefore heavily depends on the hydrologists ability to model the groundwater system parameters accurately. Failure to do so will inevitably result in model prediction errors. Due to these drawbacks machine learning models have served as an attractive alternative to groundwater flow models.

2.3 Application of Machine learning Models in Hydrology

Machine learning is an application of artificial intelligence that focuses on the development of computer programs that can assess data and use it to learn themselves [Mirzavanda *et al.* 2014]. Using historical data machine learning models can identify patterns and make better decisions by learning information directly from the data. The introduction of big data and machine learning in hydrology has resulted in several new advances in the sustainable development of groundwater resources (discussed later in the chapter) in different parts of the world. As highlighted in section 2.2 groundwater flow models have their pitfalls. An alternate solution to these major drawbacks is the use of data driven models with machine learning techniques.

Chen *et al.* [2020] conducted a comparative study between machine learning models and groundwater flow models to simulate groundwater dynamics in the Heihe River Basin in northwestern China. MODFLOW (a computer code that solves the groundwater flow equation) was used to simulate the groundwater dynamics together with three different machine learning algorithms namely; the support vector machine, multiplayer perceptron and the radial basis function network. Results revealed that the accuracy of the machine learning models was significantly better than the groundwater flow model. The groundwater flow model achieved a coefficient of determination score of 0.51 while the multilayer perceptron, radial basis function network and support vector machine achieved scores of 0.71, 0.75 and 0.76 respectively, proving the vast capabilities of machine learning techniques as compared to groundwater flow models.

Furthermore, Chitsazan *et al.* [2015] also did a comparative study between artificial neural networks and a groundwater flow model to simulate groundwater levels in Aghili plain, urban area of Gotvand in southwest Iran. A groundwater flow model had previously been developed to predict groundwater levels in the area therefore the accuracy of their machine learning model was quantified by comparing it to the groundwater flow model. Within this context, the results of the artificial neural network were compared to the results of the groundwater flow model. The plots/graphs in the study show that the predicted groundwater levels by the artificial neural network better fitted the observational data trend than the groundwater flow model. Hence, the study concluded that artificial neural networks were better suited to predict groundwater levels in the region than the previously computed groundwater flow model.

These recent studies show how machine learning models can recognize patterns hidden in historical data and then apply those patterns to make future predictions without a deep knowledge of physical area under study. Researchers have applied a variety of machine learning models including artificial neural networks, support vector machines, multiple linear regression, decision trees and many more to model groundwater systems. Table 2.2 illustrates how different machine learning models have been successfully applied in the field of hydrology to model both groundwater levels and discharge rates.

Table 2.2: Table showing authors' contributions using different machine models to predict groundwater levels and discharge rates.

Author	Models	Target & Input variables	Temporal resolution	Sample size	Prediction length	Accuracy
Zealand et al., [1999]	FFNN	Target: Discharge Input: Rainfall, temperature, river discharge	Weekly	28 years	8 years	2-step ahead: RMSE = 12.1cms 4-step ahead: RMSE = 5cms
Nayak et al., [2006]	FFNN	Target: Groundwater levels Input: Rainfall, canal releases	Monthly	8 years	4 months	RMSE = 0.32 - 0.58 m
Le et al., [2019]	RNN	Target: Discharge Input: discharge, rainfall	Daily	Station dependent	1 – 3 days	RMSE = 152.4 – 594m
Maheshwara Babu et al., [2018]	RNN	Target: Groundwater levels Input: Previous water levels, rainfall, evapotranspiration, wind speed	Monthly	17 years	25 months	$R^2 = 0.8311$ RMSE = 0.9896
Nagesh Kumar et al., [2004]	RNN FFNN	Target: Discharge Input: previous discharge	Monthly	57 years	7 years	RNN: $R^2 = 0.825$ FFNN: $R^2 = 0.697$
Supreetha et al., [2019]	RNN FFNN	Target: Groundwater levels Input: Previous groundwater levels, rainfall	Monthly	Not highlighted	12 months	RNN: RMSE= 2.4 FFNN: RMSE =3.1
Goyal et al., [2017]	SVM	Target: Discharge Input: Discharge rate, rainfall	Daily	4 years	3 months	RMSE = 0.77
Behzad et al., [2010]	SVM ANN	Target: Groundwater levels Input: abstraction rates, rainfall, temperature and groundwater level	Multiple	Not highlighted	28 months	SVM: RMSE = 0.13 ANN: RMSE = 0.16
Lin et al., [2006]	SVM ANN ARMA	Target: Discharge Input:	Monthly	30 years	60 months	SVM: RMSE = 351.81 ANN: RMSE = 368.49 ARMA: RMSE = 353.06,
Steyn et al., [2018]	SVR MLR	Target: Discharge Input: Discharge, rainfall	Daily	14 years	6 months	SVR: $R^2 = 0.89$ MLP: $R^2 = 0.92$
Kenda et al., [2018]	LR DC RF GB	Target: Groundwater levels Input: Temperature, groundwater levels, abstraction rates	Daily	23 years	N/A	LR: $R^2 = 0.62$ DC: $R^2 = 0.42$ RF: $R^2 = 0.61$ GB: $R^2 = 0.64$
Wang et al., [2018]	LS-SVR CCF-CRF RF	Target: Groundwater levels Input: Rainfall, temperature, groundwater level	Daily	17 years	Various horizons	LS-SVR: $R^2 = 0.75$ CCF-CRF: $R^2 = 0.81$ RF: $R^2 = 0.76$
Al-Abadi et al., [2016]	MLP DC	Target: Discharge Input: Discharge, river stage	Daily	1 year	4 months	MLP: $R^2 = 0.82$ DC: $R^2 = 0.88$
Sahoo et al., [2013]	MLP MLR	Target: Groundwater levels Input: Groundwater levels, rainfall, temperature and discharge	Daily	6 years	11 months	Site dependent (17 sites modelled)

Where RNN: Recurrent Neural Network, FFNN: Feed Forward Neural Network, ANN: Artificial Neural Network, SVM: Support Vector Machine, SVR: Support Vector Regression, ARMA: Autoregressive Moving Average, MLP: Multiple Layer Perceptron, DC: Decision Tree, RF: Random Forest, GB: Gradient Boosting, MLR: Multiple Linear Regression, LR: Linear Regression,

LS-SVR: Least squares support vector regression, CCF-CRF: Canonical Correlation Forests with a Combination of Random Features

It is evident in literature that recurrent neural networks (RNNs) and feedforward neural networks (FFNNs) have been the most widely used architectures when it comes to modelling groundwater systems [Coulibaly and Baldwin 2008]. The results from the studies in table 2.2 prove how machine learning models are promising tools in modelling groundwater systems.

FFNNs were able to accurately predict future values for a short period while RNN were able to predict future values over an indefinite lead time. While FFNNs are more popular in modelling hydrological systems, they lack the feedback connections necessary to model dynamic systems making it difficult for them to model time dependant patterns [Brezak *et al.* 2011]. On the other hand, RNN have feedback connections that make them inherently dynamic in nature [Haykin 1999]. In several studies mentioned in table 2.2, RNNs were able to cope with the seasonality trend and time-varying behaviour of the semi-arid aquifer systems better than FFNNs. Hence, most of the studies concluded by recommending RNNs as useful tools for predicting hydrological systems.

Moreover, several studies cited in table 2.2 presented SVR and SVM models as promising tools to predict river discharge and groundwater levels over a long-term period. These studies compared the accuracy of prediction on SVMs against ANNs and ARMA models. Through these comparisons, the studies demonstrated that the SVM model performance was a reliable candidate for prediction of long-term river discharge rates and groundwater levels. In some cases as evidenced in table 2.2 SVMs performed just as good or even better than ANNs in terms of accuracy and robustness when predicting groundwater levels and discharge rates. Behzad *et al.* [2010] study in particular concluded that SVMs were indeed a better alternative to ANNs because of its global optimization algorithm.

Table 2.2 also shows satisfactory results from different regression models including random forest, decision trees, linear regression among several others. Regression models have exhibited strong predictive capabilities and high robustness, with a lower complexity of parameterization and shorter training time for optimization when compared to SVMs and ANNs [Rodriguez-Galiano *et al.* 2012]. The studies in table 2.2 highlighted that these models were a more cost-effective alternative to ANNs and SVMs. Moreover, these models offer easy interpretation of the results between the parameters, hence they have also been favoured in modelling hydrological systems [Sahoo and Jha 2013].

However, other studies such as Kenda *et al.* [2018] study highlighted some of the implications of using regression models. Although their study achieved “satisfactory” results as evidenced in table 2.2, they noted how they encountered some anomalies as regression models were unable to model the underlying dynamics of the aquifer such as sharp declines in groundwater levels. Their study showed how a better understanding of the aquifer could help with interpreting the results especially if the results were not what was expected.

2.4 Discussion and Conclusion

Machine learning models popularity has dramatically increased among hydrologists due to the vast benefits and potential in modelling hydrological systems as evidenced above. The most advantageous aspect of machine learning models is their ability to capture and represent complex input and output relationships whether linear or nonlinear using historical data. Although FFNNs can be used to model non-linear dependencies and are also easier to maintain they lack

the feedback connections necessary to model dynamic systems making it difficult for them to model time dependant patterns [Brezak *et al.* 2011].

On the other hand, RNNs have feedback connections that make them inherently dynamic in nature allowing them to predict future values over an indefinite lead time are able. However, RNN also face challenges such as exploding gradients and vanishing gradients [Chiang *et al.* 2004]. Support vector machines and regression models have instead been applied as alternative models to ANN due to their easy interpretability and shorter computation time.

The main aim of this chapter was to demonstrate the drawbacks groundwater models face and why machine learning models would work as a better alternative when predicting groundwater levels and discharge rates. It is evident that machine learning models can be successfully be applied to model hydrological processes. In the next chapter the research framework will be explained in more detail.

Chapter 3

Research Framework Overview

3.1 Introduction

In this chapter we discuss the research methodology followed in this study. The chapter aims to explain the motivation for the chosen research methodology, explain data acquisition and pre-processing steps and the models and evaluation techniques used. The chapter will end off with a summary section of the proposed methodology. The following section begins by explaining the need for the research.

3.2 Research Motivation

South Africa is a country that suffers from severe water scarcity [Donnenfeld *et al.* 2018]. Only 1/3 of South Africa's main rivers are in good condition while the remaining 2/3 of the rivers have been over-exploited [Donnenfeld *et al.* 2018]. Moreover, over-abstraction of groundwater pumping in certain regions of South Africa has led to a decline in groundwater levels to the point where pumps were no longer able to access groundwater [The Conversation 2018].

Groundwater plays an important role in the hydrological cycle and can affect the natural environment making it a vital resource in the country. Correctly forecasting groundwater level fluctuations and river discharge rates can assist water resource managers and engineers in efficient allocation of the regional water needs in the North West province leading to sustainable water use. Groundwater flow models have been developed in the past to monitor groundwater water levels and discharge rates in different parts of the world [Song *et al.* 2007; Upadhyaya and Chauhan 2001; McCarthy and Skaggs 1991; Roulet 1991]. However, these models needed extensive physical observations to perform the modelling making them time consuming and costly to develop.

Machine learning models have been the alternate solution to predict groundwater levels and discharge rates in different areas around the world and good results were obtained using these models. The present study research therefore seeks to train a machine learning model to predict future discharge rates and groundwater levels in the North West. The accurate prediction of groundwater levels and discharge rates is essential for sustainable utilization and management of groundwater resources in the province.

3.3 Study Area

The Grootfontein aquifer is among South Africa's most important aquifers. It is located in the northern part of South Africa [Cobbing 2018]. The aquifer is highly productive containing good quality water that supplies the domestic and agricultural water needs of people in the region [WRC 2003]. Several towns and settlements such as, Itsoseng, Lichtenburg and Mahikeng (also the capital of the North West) also largely rely on groundwater as their main source of water. Average temperatures in the region range between 22 to 34 degrees in summer (August to March) and 2 to 16 degrees in winter (May to July) [SA-Venues 2019]. The province also experiences an average rainfall of 300 to 700 mm annually. Figure 3.1 shows the location of the area under study.

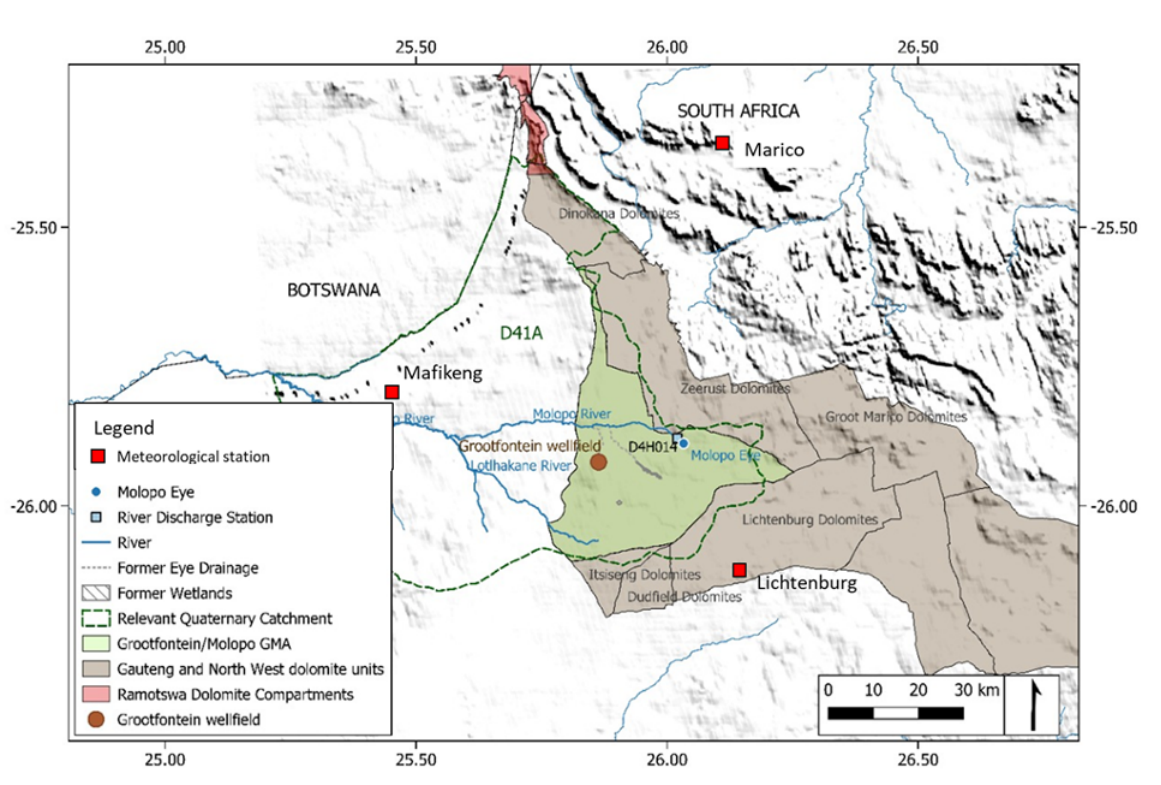


Figure 3.1: Compartments of the North West dolomite aquifers

3.4 Experimental Data Sources

The first step to obtaining data was to get the coordinate grid for the North West dolomitic region which was done by comparing shape files for the dolomite compartments with Quaternary catchments using QGIS software. From that all the Quaternary catchments within the study area were obtained. The data required for the study is all freely available online through the Department of Water and Sanitation therefore ethics was not required. Emails were sent to the relevant departments requesting for data within the Quaternary catchments obtained from QGIS. The next section will discuss each dataset obtained.

3.4.1 River Discharge Data

Historical data for two flow-gauge stations along the Molopo River recorded at monthly intervals was sourced from the DWS, namely: D4H013 (1964 – 2017) and D4H014 (1982 – 2019). Gauge D4H014 is located at Molopo eye (focus of this reserach), which is the source of the Molopo River. Gauge D4H013 is located further downstream (west of D4H014) on the Molopo River (see figure 3.1). The Grootfontein eye is located more centrally in the Grootfontein compartment. The spring discharge dried in 1981; which Cobbing [2018] attributed to the drilling of the Grootfontein wellfield boreholes in the vicinity of the spring, and to agricultural abstraction. Figure 3.2 shows the monthly discharge data from both gauge stations over a common period. Discharge station D4H014 was the main focus of this case study. The station had less than 5% missing data. Missing data points were patched using data from station D4H013.

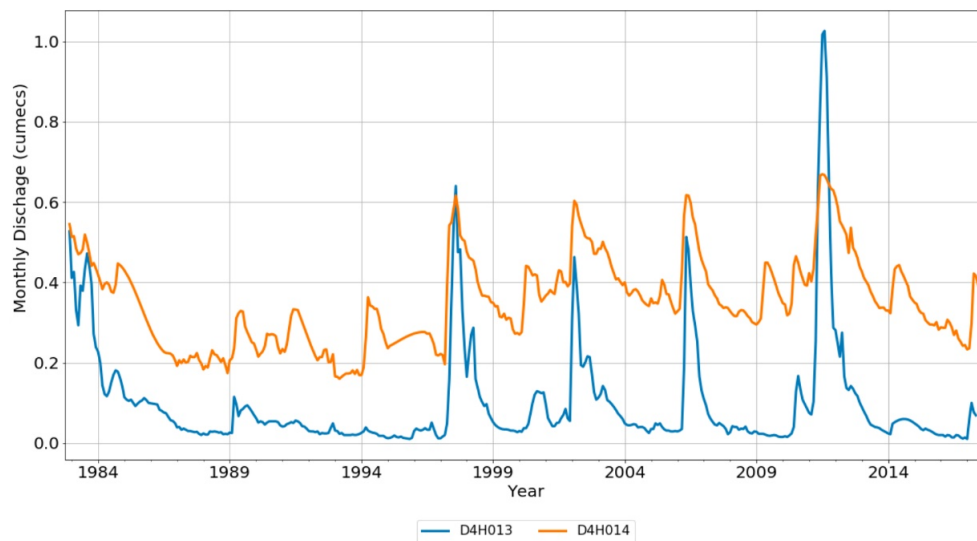


Figure 3.2: Monthly gauge data for stations D4H013 and D4H014

3.4.2 Groundwater Level Data

Groundwater level data was obtained from the HYDSTRA and NGA which are sub departments of DWS. A total of 105 boreholes were found within the vicinity of the compartment. However, not all boreholes were still active and the majority of them had significant missing data. It was therefore necessary to filter the datasets to select boreholes that would give us the best results.

To begin, boreholes that had less than 20 years worth of data and more than 30% missing data were excluded from analyses. This provided a subset of 24 boreholes across the compartment for modelling. All 24 borehole stations were either currently active or have monitoring records that extend over a 20-year period. Table 3.1 shows a summary of the pre-processing steps taken to arrive at the representative boreholes and figure 3.3 shows the time series of the selected boreholes.

Cobbing [2018] analysed groundwater levels in Grootfontein and reported a long-term decline in groundwater levels in his study. Several of the boreholes used by Cobbing [2018] are included in the 24 boreholes selected for this case study. After calculating the trend gradient and annual drop in groundwater levels (shown in table 3.4) results show that 19 of the 24 boreholes selected show a decline in groundwater levels, with an average gradient (or rate of decline in meters per month) of 0.005905; equivalent to a fall in water level of 2.2 m per decade. The rate of decline

Table 3.1: Summary of pre-processing steps on groundwater level dataset

Step Number	Action	Boreholes left after pre-processing step
1	Merged NGA & HYDSTRA	105
2	Remove Duplicates	81
3	Drop boreholes with less than 20 years data	73
4	Drop boreholes with more than 30% missing data points	30
5	Total boreholes left	24

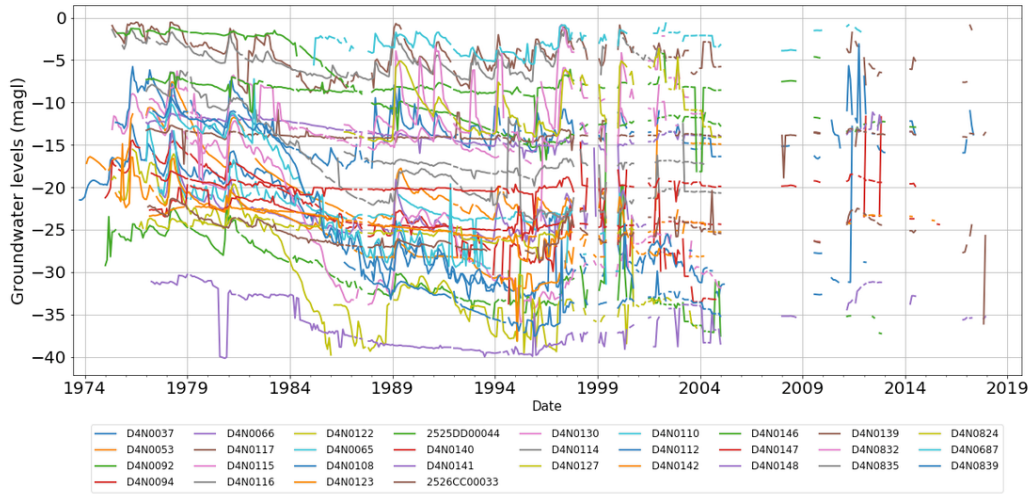


Figure 3.3: Time series of selected boreholes for modelling purposes

varies by an order of magnitude across the boreholes from 0.0002 for D4N0108 and 0.002 for D4N0115. This is expected due to the heterogeneity of the karst aquifer and the potential local effects of abstraction or recharge. Of the 24 water level records shown in the table 3.4, all but 5 show a declining water level.

Records of the selected boreholes began in 1973, during which time irrigation increased. Figure 3.3 shows a time series of the boreholes selected for modelling. As can be seen in figure 3.3 all boreholes in the compartment have a large data gap from late 2005 to early 2008. Furthermore, only a few water level measurements are available after 2008. This can be due to several reasons such as water meters failing or workers being on strike. Linear interpolation was performed to replace missing data for small data gaps (before 2005 period) in order to generate a continuous time-series dataset for modelling. Linear interpolation is achieved by geometrically rendering a straight line between the adjacent points on a graph or plane. The larger data gaps (from 2005 and 2019) were not used for modelling.

3.4.3 Groundwater Usage Data

The Grootfontein aquifer supports extensive groundwater abstractions for irrigated agriculture and domestic water supply [Cobbing 2018]. Table 3.2 illustrates the registered groundwater ab-

Borehole	Latitude	Longitude	Start date	Start wl	End date	End wl	Years Active	Trend	Trend Gradient	Annual drop
D4N0037	-25.982	25.9225	1973-08-01	20.620	2012-02-01	28.85	39	decline	0.001200	0.211026
D4N0053	-25.9094	25.8931	1974-02-01	17.175	2005-01-01	14.92	31	incline	0.000700	-0.072742
D4N0065	-26.0031	25.9	1975-06-01	18.500	2000-02-01	20.76	25	decline	0.000700	0.090400
D4N0066	-26.0192	25.9206	1975-03-01	23.620	2002-04-01	24.50	27	decline	0.000400	0.032593
D4N0092	-26.0357	25.972	1975-01-01	29.220	2013-04-01	36.67	38	decline	0.000900	0.196053
D4N0094	-26.0356	25.9722	1975-01-01	21.240	2004-12-01	33.21	29	decline	0.001300	0.412759
D4N0108	-25.9084	25.8863	1975-07-01	12.350	2019-06-01	14.42	44	decline	0.000200	0.047045
D4N0110	-25.9522	25.9103	1976-11-01	11.575	2000-01-01	29.90	24	decline	0.003200	0.763542
D4N0112	-25.9367	25.9094	1977-01-01	11.020	2003-08-01	27.48	26	decline	0.002500	0.633077
D4N0115	-25.918	25.8943	1975-05-01	13.180	2005-03-01	31.89	30	decline	0.002000	0.623667
D4N0116	-26.0072	25.855	1975-05-01	1.410	2005-01-01	5.08	30	decline	0.000100	0.122333
D4N0117	-25.966	25.843	1975-05-01	0.930	2019-04-01	6.27	44	decline	0.000200	0.121364
D4N0123	-26.0281	25.9736	1975-07-01	20.680	2005-01-01	28.14	30	decline	0.001000	0.248667
D4N0126	-25.9642	25.9811	1976-08-01	17.500	2018-04-01	33.70	42	decline	0.001200	0.385714
D4N0127	-25.8742	26.0031	1976-09-01	24.640	2005-01-01	25.41	29	decline	0.000300	0.026552
D4N0130	-25.9175	25.9781	1976-04-01	6.300	2002-12-01	12.70	26	decline	0.000200	0.246154
D4N0139	-25.9511	26.0837	1977-02-01	22.310	2019-02-01	26.87	42	decline	0.000080	0.108571
D4N0140	-25.977	26.0912	1976-01-01	22.950	2017-09-01	24.46	41	decline	0.000200	0.036829
D4N0141	-25.9596	26.1348	1976-02-01	10.000	2014-05-01	13.46	38	decline	0.000200	0.091053
D4N0142	-25.8973	26.1719	1977-01-01	23.690	2017-08-01	24.38	40	decline	0.000200	0.017250
D4N0146	-25.8855	26.0197	1977-01-01	7.310	2011-11-01	20.00	34	decline	0.000050	0.373235
D4N0147	-25.8942	26.0573	1977-01-01	18.680	2015-06-01	10.93	38	incline	0.000010	-0.203947
D4N0687	-25.8928	25.9658	1985-03-01	5.550	2012-01-01	1.53	27	incline	0.000100	-0.148889
D4N0824	-25.9317	25.8356	1983-11-01	13.490	2005-01-01	14.10	22	decline	0.000500	0.027727
D4N0832	-25.9353	25.9797	1983-10-01	12.620	2004-03-01	11.30	21	incline	0.000700	-0.062857
D4N0835	-25.9558	25.9919	1983-10-01	15.830	2005-01-01	17.51	22	decline	0.000090	0.076364
2526CC00033	-25.894	26.029	1976-03-01	12.000	2019-09-01	14.04	43	decline	0.000009	0.047442
2625BB00028	-26.0288	25.9735	1975-07-01	20.680	2004-03-01	28.14	29	decline	0.001000	0.257241
Mean Precipitation	-	-	1950-01-01	95.900	2019-01-01	95.20	69	decline	0.000600	-0.010145

Figure 3.4: Analysis of linear trend for 24 boreholes selected for modelling, along with rainfall

straction rates in the Grootfontein aquifer based on data received from WARMS as of November 2016. As illustrated in table 3.2, irrigated agriculture accounts for nearly 95% of groundwater abstraction in the region. Apart this, groundwater is also used for mining and other industrial-based activities. The sum of the registered abstractions per water use sector is shown in the table 3.2.

There were several challenges with the obtained abstraction dataset. In addition to the challenge of determining actual water use rather than registered use there are other known inaccuracies with the WARMS dataset including incomplete entries, inaccurate coordinates and duplicate entries. Furthermore, the registrations may be under-estimates to secure a license or over-estimate in an attempt to secure a supply. Moreover, several users may not be registered (i.e. unlicensed abstraction).

It was therefore necessary to remove all duplicate entries and discard registrations with missing registered volume amounts as these would be of no use. To generate the cumulative use shown in figure 3.5, the “registered volume start date” and “registered volume end date” was used in

Table 3.2: Summary of registered groundwater use in Grootfontein (WARMS, DWS)

Water use sector	Number of Registrations	Total Use (m3/a)
Agriculture: Irrigation	261	21 220 316
Agriculture: watering livestock	11	33 708
Mining	10	38 927
Industry	5	951 614
Total	287	22 244 565

conjunction with the “register status” column. If the registered status column had a value of “active” then the close date for those boreholes was set to 01/11/2016 as this is when the data was acquired. Moreover, the earliest registration start date in the compartment was used to replace the blank values that were found in the registered volume start date column. The total sum per year was then calculated using the aforementioned columns. Following this process, the cumulative registered abstraction within the Molopo/Grootfontein GMA boundary totals 22 244 565 m3/a as of November 2016. This appears in line with Cobbing’s (2016) lower estimate of 18.3 m3/a for the Grootfontein GMU boundary.

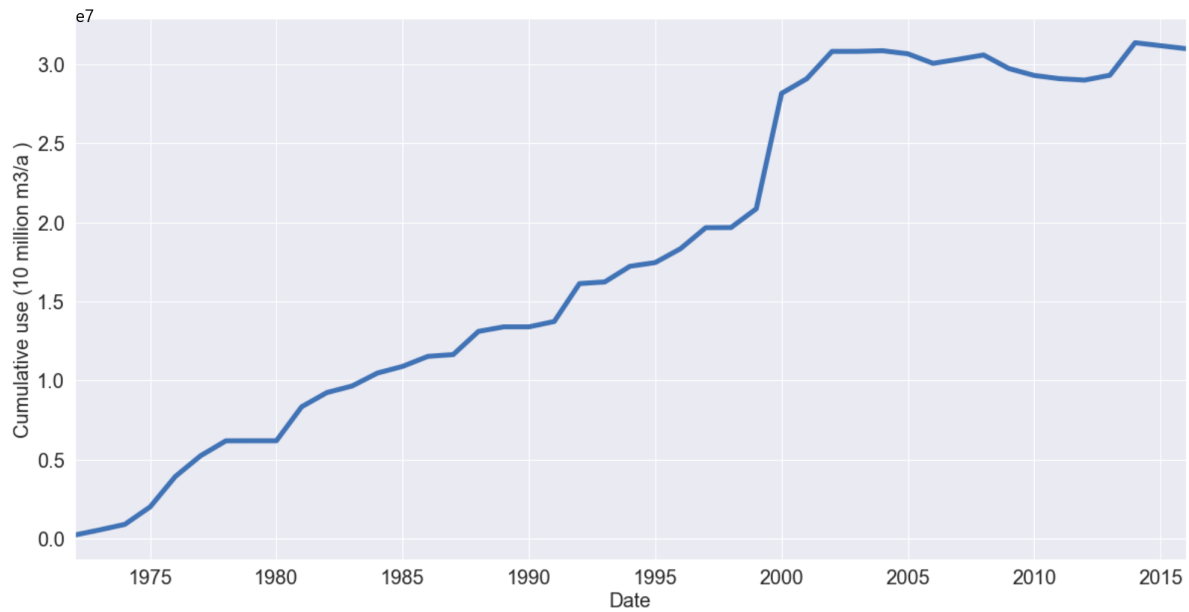


Figure 3.5: Cumulative use based on WARMS licence registrations

3.4.4 Weather Data

As seen in figure 3.1 there are numerous weather stations that are in close proximity with the dolomite aquifer. It should be noted that only rainfall and temperature datasets were obtained from SAWS. There was limited data for evapotranspiration rates in the Grootfontein area hence the variable was omitted as a model input for predicting both water levels and discharge rates in the region.

Three weather stations surround the compartment; one within the Mafikeng region (North West of the compartment), one in Marico (North of the compartment), and another in Litchenburg

(South of the compartment). The location of each station is shown in figure 3.1. The three stations were selected due to the availability of longer historical data and their positions relative to the compartment boundaries to indicate the rainfall and temperature variability across the compartment. The timeframe over which the data extends varies from station to station and is highlighted in table 3.3. The minimum, maximum and mean annual rainfall of each station is also shown in table 3.3. Figure 3.6 illustrates the monthly rainfall and temperature values for each station over a common period.

Table 3.3: Mean, minimum and maximum annual rainfall for stations around Grootfontein

Station	Data Availability	Min Annual Rainfall	Max Annual rainfall	Mean Annual Rainfall
Mafikeng	1994-2018	85,6 (1994)	777,4 (2010)	524,51
Litchenburg	1984-2018	152,7 (1984)	873,8 (1997)	507,86
Marico	1950-2019	159,8 (1994)	902,5 (1956)	609,43

The station located south-east of the compartment (Litchenburg) appears to have higher rainfall events between the period 1984 to 2009. Rainfall stations in the northern part of the compartment (Mafikeng and Marico) appear to have lower rainfall events as compared to the southern area. High rainfall events usually occur in summer (from August to March). Moreover, the Marico area experiences higher temperatures as compared to Mafikeng and Lichtenburg areas. A time series of monthly temperature and rainfall was compiled by averaging the datasets from all three stations.

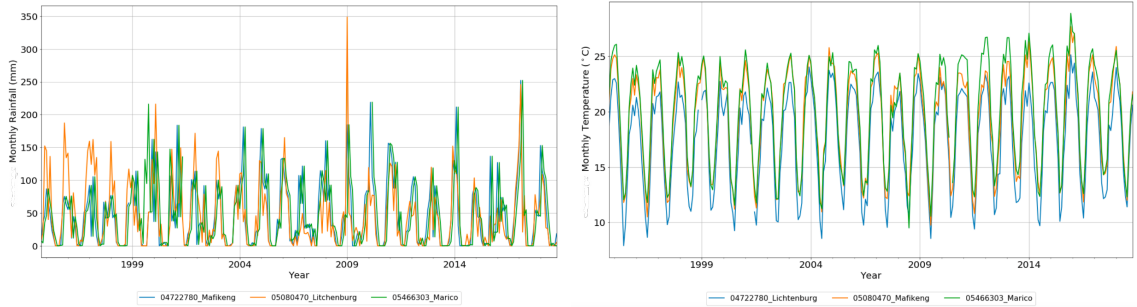


Figure 3.6: (a) Monthly rainfall from three weather stations around Grootfontein (b) Monthly temperature from three weather stations around Grootfontein

3.5 Experimental Design

3.5.1 Artificial Neural Networks

The brain can memorize, learn and recognize patterns and yet still generalize in a wide variety of circumstances. The ability of the brain to do this was the driving factor behind the development of neural networks which attempt to mimic biological neural systems [Engelbrecht 2007]. Artificial neural networks (ANN) are built out of densely interconnected set of nodes that are inspired by the biological learning process. As evidenced in Chapter 2, neural networks namely the FFNN and the RNNs, have been successfully applied in the field of hydrology. In this section we discuss the two networks in more detail.

RNNs take two sources of input to determine how they will respond to new data. The RNN

model combines the current input example it sees (the present) and the previously computed output (the past). RNNs can be distinguished from FFNNs by the feedback loop that connects their present to their past decisions. The sequential information stored in the RNNs hidden state spans several time stamps as it cascades forward to affect the processing of each new example [Taver *et al.* 2015]. Therefore an event downstream in time depends upon, and is a function of one or more events that came before.

Figure 3.7 shows an example of a RNN where W_x represents the weight vector for the hidden layer, W_y represents the weight vector for the output layer, W_h represents the same weight vector at different Timesteps, x_n and y_t are the input vector and output result at time t and h_n represents the activation function.

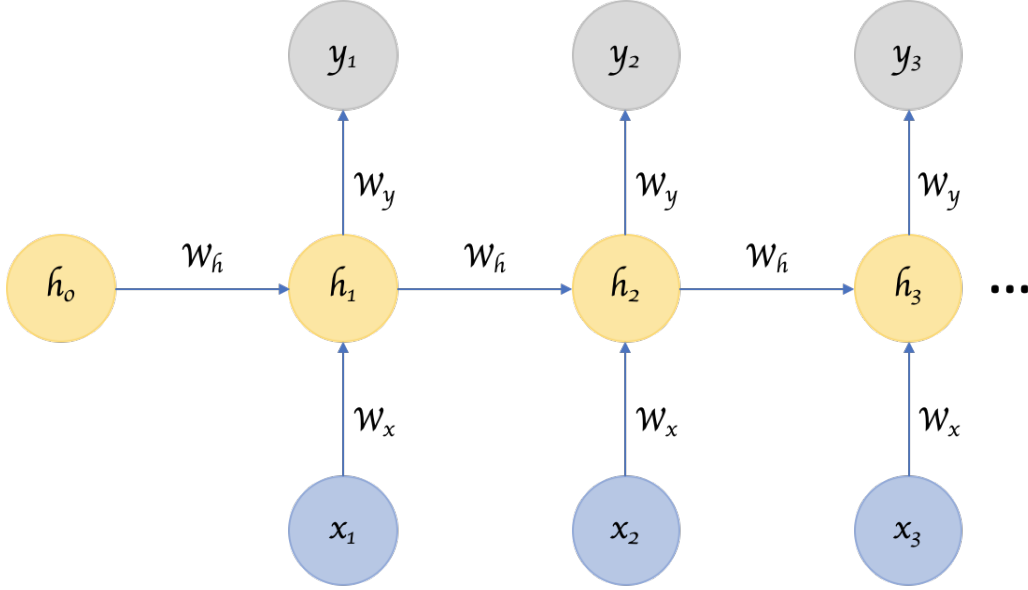


Figure 3.7: An example of a Recurrent Neural Network. The hidden state is meant to carry important information from one input item in the series to others

To calculate the hidden layer at Timestamp (t) equation 3.1 is used and to calculate the output equation 3.2 is used.

$$h(t) = f(h_{t-1}, x_t), \quad (3.1)$$

where h_t represents the current state, $h_t - 1$ represents previous state and x_t represents the input state

$$y(t) = W_{h_y} h_t, \quad (3.2)$$

where y_t represents the output and W_{h_y} represents the weight at the output layer

In this study two of the most popular RNNs will be employed. Namely, the Gated recurrent unit (GRU) and the Long-short term memory (LSTM).

FFNNs on the other hand do not contain a feedback loop and only take 1 source of input. With FFNNs the outputs of the nodes on one layer are inputs to the next layer. A typical multilayer

FFNN contains an input layer, one or more hidden layers and an output layer. Information in the network flows from left to right through connections between nodes hence it is referred to as a feedforward network [Mitchell 1997]. Each neuron receives inputs from neurons on its left and multiplies the input with the weights of the connections they travel along. The formula for this is shown in equation 3.3.

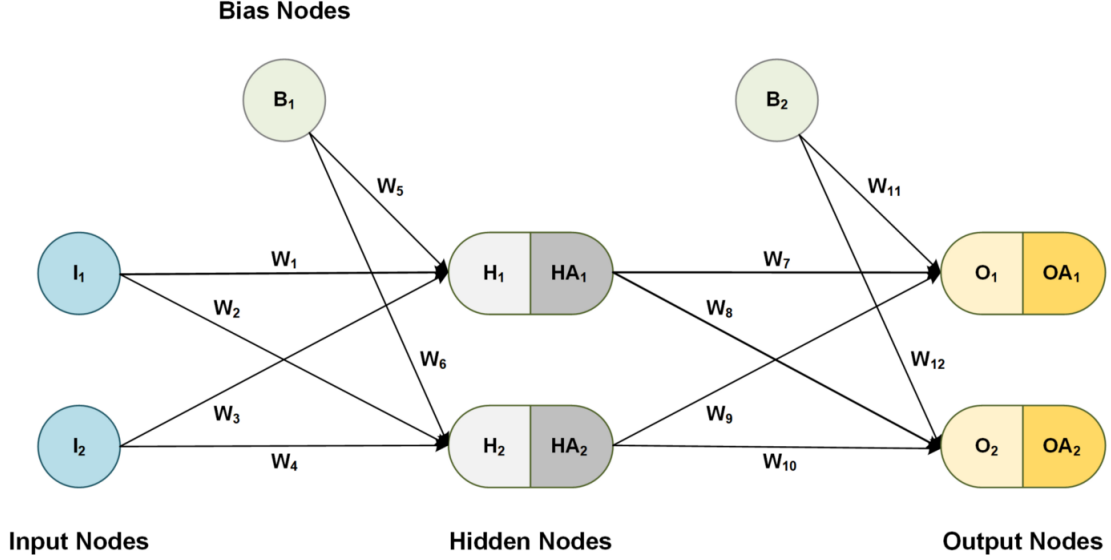


Figure 3.8: General architecture of a FFNN

Figure 3.8 shows the typical structure of a FFNN where: w represents the weight connection, I represents the input node, H represents the hidden node which is the weighted sum of inputs layers or previous hidden layers, HA represents an activated node which means the node is passed through a predefined function such as sigmoid so HA is the resulting value of that node after it is passed through the activation function, O represents the output node which is the weighted sum of the last hidden layer, OA represents output an activated node and B represents the bias node.

$$\theta = b + \sum_{i=1}^n x_i \times w_i, \quad (3.3)$$

where b is the bias of the node, x_i is the i th nodal value in the previous layer, w_i is the weight that connects x_i to θ and N is the number of nodes. So, each neuron calculates a single real-valued output based on the weighted sum of the inputs, adds a bias term and passes all this through an activation function.

The weights and bias of all neurons can be learned by the backpropagation algorithm which attempts to minimize the error or loss function. Backpropagation is the practice of fine tuning the weights of the neural network based on the error rate that was obtained from the previous iteration. In essence, the algorithm seeks to minimize the error function by updating the weights and biases of all neurons using the gradient descent approach. Proper tuning of the weights and bias ensures low error rates making the model reliable because of its increased generalization.

Additional information about the two networks can be found in [DiPietro and Hager 2019].

3.5.2 Support Vector Machines

Support vector machines (SVMs) are a set of supervised learning methods which are known for performing regression and classification analysis using pattern recognition and data analytics [Gove and Faytong 2012]. The popularity of SVMs applied to regression problems can be attributed to 1) good generalization performance because rather than minimizing the training error (done by ANNs), SVMs seek to minimize an upper-bound of the generalization performance, 2) while several machine learning tools such as ANNs are subject to local minima, the solution of SVMs is always globally optimal and 3) solution is represented sparsely by support vectors which are typically a small subset of all training examples [Cao *et al.* 2003].

With SVMs, data is classified into two different classes using a hyperplane. Figure 3.9 shows the structure of SVMs. Each data point in the training set is known and marked as a member in either of the two classes. The SVM has two lines other than the hyperplane which create a margin and these lines are referred to as boundary lines. The support vectors can be on the boundary line or outside it. The basic idea behind SVMs is to try separate the data points into different classes by constructing an optimal hyperplane which maximizes distance between the plane and the nearest data point of each class (also referred to as margin).

Support vector regression (SVR) uses the same concept. The difference between the SVR and SVMs is that SVR is a regression problem used to predict continuous values while SVM is a classification problem used to predict categorical output. In SVR the hyperplane is a line that will help us predict the continuous values or target value. The aim is to try fit the error within a certain threshold (ϵ).

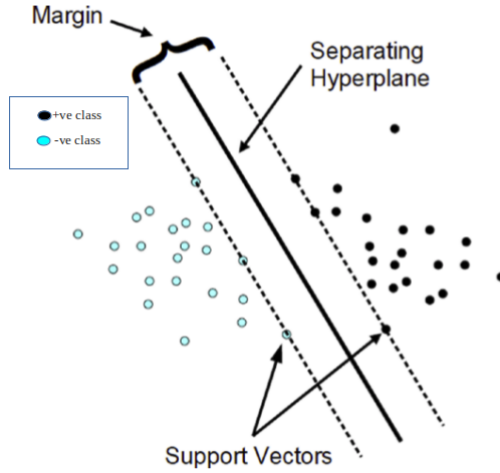


Figure 3.9: Support Vector Machine

The SVR method is designed to maximize the margin between the two data classes shown in figure 3.9, which in turn maximizes the generalization ability of the model. Mathematically this output expression is equivalent to

$$f(x) = w \cdot \phi(x) + b, \quad (3.4)$$

where f denotes an SVR estimator, w denotes the weight vector and ϕ is a non-linear transfer function that maps input vectors x_i into a higher dimensional feature space. The coefficients of w and b are estimated by solving the following optimization problem for SVR:

$$\begin{aligned}
& \underset{w, b, \xi, \xi^*}{\text{minimize}} && \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \\
& \text{subject to} && w^T \phi(x_i) + b - y_i \leq \epsilon + \xi_i^* \\
& && y_i - w^T \phi(x_i) - b \leq \epsilon + \xi_i \\
& && \xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, N
\end{aligned} \tag{3.5}$$

where ξ^* and ξ are slack variables that penalize errors of estimated values over the error tolerance ϵ , x_i is the input vector in the training stage and C is the regularization parameter which determines the penalties to estimation errors during the model building process. The input x_i is first projected to a higher dimensional feature space by mapping $\phi: X \Rightarrow F$. Usually this mapping function is implemented implicitly via kernels functions such as radial basis, linear and polynomial kernels. A detailed explanation of these kernel functions can be found in references [Afonja 2017; Chiroma *et al.* 2014]. A brief overview of SVR is provided here. For further information, readers are referred to Xu *et al.* [2012] and Bhattacharyya [2018].

3.5.3 Random Forest

Random forest (RF) is the most popular bagging model that is used to achieve high variance and low bias in datasets [Breiman 2001]. The RF model is a collection of unpruned classification and regression trees (CART). RF generates multiple decision trees during the training phase and generates an output by taking the majority vote (for classification) and an average of a single tree method (for regression) [Breiman 2001].

The two user-defined parameters of a RF model are: B which represents the number of trees in the forest and D which represents the number of features used to split the nodes. The final results of the regression can be calculated using the following equation 3.6

$$\hat{y}(x_i) = \frac{1}{B} \sum_{b=1}^B T_b(x_i), \tag{3.6}$$

where B represents the number of trees in the forest and T_b represents each tree.

3.6 Information Gain Ranking

Generally, input parameters are not equally informative in predicting the target variable. This is because some parameters may be noisy, correlated or have an insignificant relationship with the target variable. To quantify the influence of each of the input variables, feature analysis was carried out. Specifically, mutual information (MI) was selected as the appropriate feature analysis method for this study. MI is a statistical method that is used to measure the relationship between two random variables that are sampled simultaneously [Sudharsan 2018]. In particular, it measures how much information is communicated, on average, in one variable about another. The MI between two random variables X and Y can be stated as shown in the following equation,

$$I(X; Y) = H(X) - H(X|Y), \tag{3.7}$$

where $I(X; Y)$ is the Information Gain Ranking for X and Y , $H(X)$ is the entropy for X and $H(X|Y)$ is the conditional entropy for X given Y . Entropy quantifies how much information

there is in a random variable. MI will be greater than zero when X and Y exhibit mutual dependence regardless of how linear or nonlinear that dependence is. MI will be 0 if and the two variables are statistically independent. So, the stronger the relatedness or the higher the dependency the higher the MI value. For this study, MI was calculated in python using “mutual_info_regression” class from the sklearn library.

It should be noted that the MI does not model the influence of each feature on the target value. Instead MI, was used as an indicator of which variables are the most correlated with the target variable.

3.7 Model Performance Analysis and Evaluation

Building a robust machine learning model is based on the principle of constructive feedback. Firstly, a machine learning model is built, then model performance is evaluated using a set of metrics and based on the feedback received from the metrics, improvements are made to the model until a desired accuracy is achieved [Engelbrecht 2007]. The performance of machine learning models is defined by how well the models can map the input space to the output space and how well the models can generalize to new points that were previously not seen during training [Engelbrecht 2007].

To test model accuracy the data is usually split into training and testing datasets. The model is trained using the training set and tested by making predictions against the test set. Model accuracy needs to be considered for both values in the training set and values not included in the training set. The training error and generalization error will need to be calculated. The end goal is to use a model that gives a high accuracy for both training and test data.

The following loss functions can be used to evaluate the prediction ability of different machine learning techniques.

O_i and $\bar{O}$ denote the observed value and the mean observed value respectively, P_i and $\bar{P}$ denote the predicted values and mean predicted value respectively and n is the number of samples available.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2}, \quad (3.8)$$

Equation 3.8 measures the average magnitude of the error between observed (o) and predicted (p) values. RMSE is a good measure of error that can serve as a loss function to minimize [Engelbrecht 2007].

$$MSE = \frac{1}{n} \sum_{t=1}^n (O_i - P_i)^2, \quad (3.9)$$

Equation 3.9 measures the average of errors squared. Error in this case means the difference between the observed values (O_i) and the predicted values P_i . The difference is squared so that negative and positive values do not cancel each other out.

$$MAE = \frac{1}{n} \sum_{t=1}^n (y_i - \bar{y}_i)^2, \quad (3.10)$$

Equation 3.10 measures the average magnitude of the errors without considering their direction. It calculates the averages of the absolute differences between prediction and actual observation where all individual differences have equal weight.

$$R^2 = \left(\frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^n (P_i - \bar{P})^2}} \right)^2, \quad (3.11)$$

Equation 3.11 explains how much variability of one factor can be caused by its relationship to another factor (co-linearity between the observed and predicted data). R^2 values range from 0 to 1, with higher values indicating less error variance (i.e. the closer the value is to 1, the better the fit, or relationship, between the two factors) [Krause *et al.* 2005].

3.8 Discussion and Conclusion

This chapter described the methodology of the research undertaken. It defined the process followed in obtaining the data for the study, described characteristics of the study area, the research method and the evaluation metrics used. In summary, the research will utilize data obtained from DWS (groundwater levels, river discharge rates), WARMS (groundwater abstraction) and from the SAWS (precipitation and temperature). Different machine learning models will be tested including FFNNs, RNNs, SVR and RF models. Lastly, RMSE, R^2 MSE and MAE will be used as model evaluation metrics.

Chapter 4

Groundwater Level Forecasting

4.1 Introduction

The first aim of this study was to accurately predict groundwater levels in the Grootfontein aquifer. Under this aim were several objectives. Firstly, to present the ranking by entropy of which features have the most influence on groundwater levels. To measure this MI was used. Next aim was to present an optimal predictive model that can be used to forecast groundwater levels in the aquifer. To do this, 5 models mentioned in section 3.5 were trained and tested to select the optimal model. Next aim was to use the final selected optimal model to test the fluctuations of groundwater levels under different case scenarios including worst case, average case and best case scenarios. Final aim was then to provide evidence of the effectiveness of our chosen method by applying the selected optimal model to predict groundwater levels in a different karst aquifer to test model generalizability. The Steenkoppies aquifer located in the Gauteng region was used as a case study.

In the following section we discuss the model setup process for predicting groundwater levels.

4.2 Model Setup

For groundwater level prediction, we employed three different neural network architectures namely FFNN, GRU and LSTM model together with 2 other regression models namely the RF and SVR models. In this section we explain how the models for this research were defined.

The tensorflow library in Python was used to model all the neural network architectures. Pandas, NumPy and Matplotlib were libraries imported for management, processing and visualisation of data. Figure 4.1 summarises the methodological approach taken to build the ANN models.

Firstly, lag testing was performed to find the appropriate lag time between rainfall and groundwater levels. Hydrogeologically there is a time lag between the occurrence of rainfall and groundwater level response [Mcphie and Post 2010], hence accounting for the lag could potentially increase the model's ability to pick up the influence of rainfall on groundwater levels. The optimum time lag between rainfall and groundwater levels in the Grootfontein compartment was calculated using cross correlation. An optimal time lag of 2 months was selected.

Next we had to stationarize the datasets as ANNs are unable to efficiently capture trend variations or seasonality. Data which is stationary is also easier for ANNs to model which could

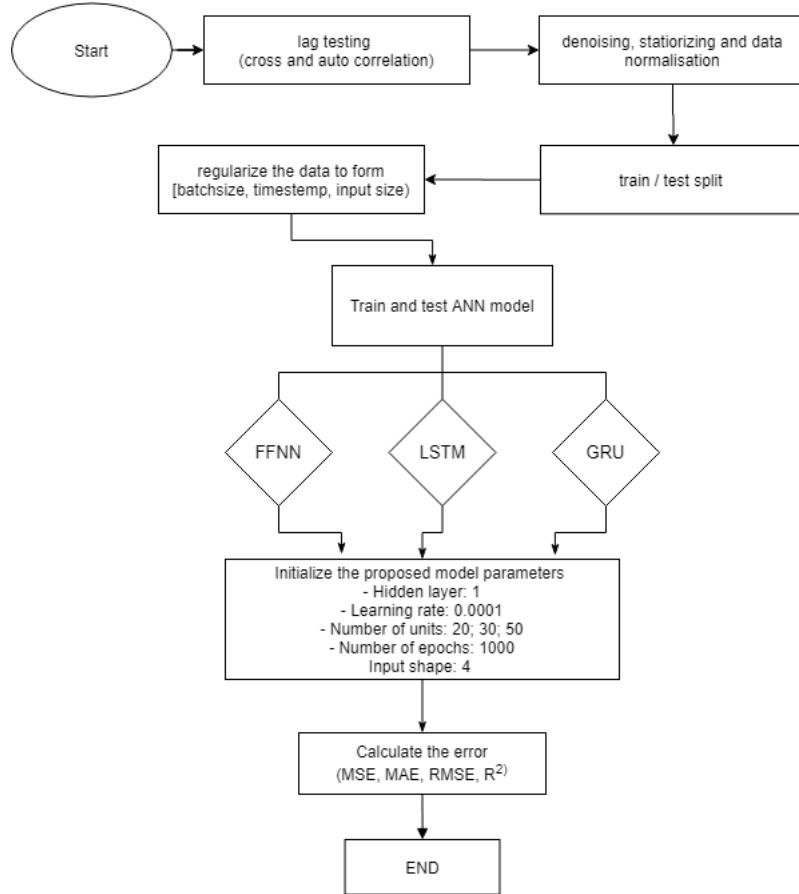


Figure 4.1: Flowchart Summarising ANNs model approach

possibly lead to better predictions. Stationarity implies, in the most intuitive sense that the statistical properties of a process producing a time series does not change over time. This however, does not imply that the sequence does not change over time but rather the way in which it changes does not itself change overtime. To stationarize the features we removed the trend by differencing the data. Differencing is achieved by taking the current observation (t) and subtracting the previous step ($t-1$) from it. The differenced time series is then inverted to its original scale to make predictions. [Zhang and Qi 2005] study found that applying such techniques could dramatically reduce forecasting errors.

Subsequently the time series datasets were transformed to scale using normalisation. ANNs anticipate data to be in scale of the activation function in this case tanh (has a range of -1 to 1). To avoid contaminating the experiment with information from the test dataset the scaling coefficients (max and min) values was calculated on the training dataset and applied to scale the test dataset. This also made the experiment fair. Normalization was done using the MinMaxScaler from sklearn.

Next was to divide the data into training and test sets. 20% of the data was used as the test set and 80% as the train set. After this, the data was regularized. Since the ANN layer expects input fed to be in a three-dimensional matrix where the first dimension represents the samples which are the total of time steps collected, the second dimension represents the time steps which are the past observations for the feature in this case 5 and the final dimension represents features which are the data columns or number of features which in this case was 4. To match

the architecture of the neural network model, the input layer had to be reshaped into a 3D vector.

An important step in constructing ANNs is hyperparameter selection specifically choosing the number of epochs, number of units and the learning rate. In this study we made use of the Adam optimization algorithm which is an extension to the stochastic gradient descent procedure. A learning rate of 0.00001 was chosen instead of the default 0.001. This was done to make the training process slower and the curve of the loss function smoother. The number of units selected was 20. Since there is no particular way for choosing the number of units in each cell, different values were tested (20, 30, 40) and evaluated on the forecast results. Finally, to record the required information during training and validation, the maximum number of epochs was set to 1000.

To improve the effectiveness of the model and prevent overfitting, early stopping and dropout techniques were employed. When a model is said to have overfitted, it would have learned the noise and detail in the training data to the extent that it negatively impacts its generalisation to new data. Early stopping is used to stop the model when output begins to decrease or remains constant on the validation dataset. If the performance on the validation dataset did not improve after 100 epochs, the developed neural network models were set to halt the training phase.

Finally, neural networks give different results each time the model is run. To get a more robust result, the experiment for each borehole was replicated ten times in order to achieve a more accurate outcome and the average estimate was taken as the final simulated result. The error metrics were calculated on the averaged values to indicate how well the model is expected to perform on unseen data on average.

4.3 Experimental Results

4.3.1 Information Gain Ranking

The contribution of each feature against the target variables (groundwater levels) is shown in table 4.1. The first column indicates the borehole ID (BHID), and the remaining 5 columns indicate the feature names and the corresponding entropy scores to the target variables.

Table 4.1 shows that on average regional abstraction correlates with groundwater the strongest in each borehole as entropy ranges between 0.47 to 0.69. Again, correlation does not imply causality, but these results essentially show that the method used detects a strong correlation between the curve of increasing groundwater use over time with the groundwater levels at each borehole site. It should be noted that the results obtained are surprising given that the groundwater level dataset is significantly noisier than that of groundwater use (seen in figure 3.5). However, the results also support the assertion that the compartment is highly impacted by regional abstraction as recorded in previous studies.

The second most correlating feature with groundwater levels was discharge with an entropy average of 0.42. A strong relationship was expected between the two variables, given that the hydraulic gradient between the surrounding aquifer and the eye is what drives discharge. Hence, changes in groundwater levels will directly translate to changes in discharge rate.

The rainfall trend (decomposed time series dataset mentioned in 4.2) is the feature with third strongest correlation with entropy values ranging between 0.286 to 0.453 and on average a value of 0.36. A high correlation was expected between rainfall and groundwater levels since high groundwater levels respond to high rainfall events. On the other hand, the actual observed

Table 4.1: A ranking by entropy of each input feature against each target variable

	BHID	Temperature	Rainfall trend	Actual rainfall	River Discharge	Groundwater Abstraction
	D4N0037	0,01	0,36	0,05	0,45	0,61
	D4N0053	0,02	0,35	0,02	0,40	0,56
	D4N0065	0,01	0,29	0,05	0,38	0,57
	D4N0066	0,01	0,32	0,01	0,39	0,68
	D4N0092	0,04	0,39	0,02	0,43	0,53
	D4N0094	0,02	0,38	0,01	0,49	0,55
	D4N0108	0,05	0,33	0,04	0,37	0,58
	D4N0110	0,05	0,39	0,03	0,43	0,58
	D4N0112	0,02	0,37	0,05	0,38	0,57
	D4N0115	0,01	0,36	0,05	0,47	0,69
	D4N0116	0,04	0,33	0,03	0,32	0,48
	D4N0117	0,03	0,37	0,04	0,36	0,60
	D4N0123	0,03	0,37	0,05	0,38	0,67
	D4N0126	0,02	0,39	0,04	0,48	0,62
	D4N0127	0,01	0,32	0,05	0,35	0,69
	D4N0130	0,02	0,38	0,02	0,45	0,65
	D4N0139	0,02	0,32	0,01	0,49	0,61
	D4N0140	0,02	0,32	0,03	0,35	0,66
	D4N0141	0,04	0,37	0,06	0,32	0,65
	D4N0142	0,03	0,36	0,06	0,32	0,62
	D4N0146	0,00	0,40	0,06	0,45	0,60
	D4N0147	0,05	0,35	0,02	0,46	0,62
	D4N0687	0,06	0,37	0,06	0,52	0,60
	D4N0824	0,06	0,35	0,05	0,45	0,59
	D4N0832	0,06	0,29	0,03	0,42	0,54
	D4N0835	0,02	0,36	0,03	0,46	0,63
	2526CC00033	0,05	0,44	0,03	0,49	0,64
	2625BB00028	0,04	0,34	0,03	0,41	0,61
	Average	0,03	0,36	0,04	0,42	0,61

rainfall dataset exhibited low entropy scores seen in table 4.1 with an average value of 0.04. The low scores from the actual rainfall value were attributed to the dataset being too noisy. A study by Xu *et al.* [2019] looked at the decomposed signals of rainfall as inputs for forecasting rainfall. The study found that the decomposed signals of rainfall produced better predictions compared to actual measured rainfall. A similar observation is seen in this study.

Temperature had the lowest correlation (and therefore least influence in predicting water levels) with an entropy ranging between 0.0004 to 0,055 and an average of 0.03. A decomposed time series signal of temperature was also tested and the score was found to be similar with that obtained from the actual dataset. Therefore, we conclude that for this study, temperature has little to almost no correlation on groundwater levels in the region.

It should be noted that MI only measures the mutual dependence between two random variables by identifying how much information of one of the features can be obtained from the other features. Hence, this method does not tell us what drives groundwater levels in the region. But instead, MI provides a useful framework for feature selection by indicating the contribution of each feature relative to all the other features.

4.3.2 Model Results

The results obtained by the predictive models are evaluated using metrics discussed in chapter 3.7. The metrics are used to evaluate the performance of the model when comparing the simulated values made during the training and testing process and the actual observed values. Depending on the period length of the boreholes, the size of the training and testing datasets varies. The results obtained from the 24 boreholes selected for modelling are summarised in table 4.2.

BHID	Model	R ²	MSE	RMSE	MAE
DN40037	LSTM	0.866	0.008	0.091	0.064
	GRU	0.853	0.009	0.095	0.070
	FFNN	0.670	0.551	0.742	0.583
	SVR	0.694	0.588	0.763	0.604
	RF	0.584	0.798	0.893	0.721
D4N0053	LSTM	0.799	0.252	0.502	0.343
	GRU	0.676	0.763	0.873	0.353
	FFNN	0.682	2.382	1.543	0.247
	SVR	0.660	2.624	1.620	0.851
	RF	0.631	2.945	1.716	0.736
DN40065	LSTM	0.620	0.025	0.157	0.123
	GRU	0.615	0.024	0.155	0.120
	FFNN	0.600	0.025	0.158	0.123
	SVR	0.511	0.234	0.484	0.321
	RF	0.408	0.345	0.588	0.386
	LSTM	0.593	0.597	0.773	0.574
	GRU	0.598	0.598	0.598	0.566

DN40066	FFNN	0.578	0.619	0.786	0.576
	SVR	0.574	3.443	1.856	1.266
	RF	0.455	4.404	2.098	1.516
D4N0092	LSTM	0.570	0.770	0.880	0.334
	GRU	0.530	0.845	0.920	0.308
	FFNN	0.330	1.220	1.100	0.527
	SVR	0.270	1.310	1.140	0.506
	RF	0.180	1.590	1.260	0.831
	LSTM	0.629	0.023	0.153	0.111
D4N0094	GRU	0.578	0.029	0.17	0.135
	FFNN	0.541	0.192	0.438	0.358
	SVR	0.538	0.182	0.426	0.354
	RF	0.440	0.215	0.463	0.379
	LSTM	0.578	0.197	0.443	0.295
D4N0108	GRU	0.574	0.195	0.442	0.298
	FFNN	0.425	1.026	1.013	0.759
	SVR	0.363	1.468	1.212	0.931
	RF	0.304	1.969	1.403	1.137
	LSTM	0.719	0.016	0.126	0.103
DN40110	GRU	0.691	0.033	0.182	0.157
	FFNN	0.385	1.055	1.027	0.775
	SVR	0.354	1.589	1.260	0.932
	RF	0.288	2.104	1.434	1.183
	LSTM	0.276	0.012	0.108	0.088
D4N0112	GRU	0.253	0.012	0.109	0.088
	FFNN	0.123	0.017	0.126	0.108
	SVR	0.145	2.624	1.607	1.309
	RF	0.153	2.751	1.658	1.410
	LSTM	0.685	0.026	0.161	0.137
DN40115	GRU	0.688	0.026	0.159	0.133
	FFNN	0.585	1.113	1.054	0.737
	SVR	0.657	1.541	1.241	0.892
	RF	0.788	0.953	0.976	0.621
	LSTM	0.488	0.272	0.520	0.464
	GRU	0.599	0.213	0.459	0.407
	FFNN	0.386	0.326	0.571	0.511
	SVR	0.313	0.971	0.984	0.660

DN40116

	RF	0.273	1.026	1.013	0.759
D4N0117	LSTM	0.628	0.148	0.390	0.230
	GRU	0.615	0.145	0.385	0.230
	FFNN	0.552	0.213	0.461	0.270
	SVR	0.522	0.234	0.483	0.283
	RF	0.443	0.266	0.519	0.303
D4N0123	LSTM	0.788	0.596	0.772	0.787
	GRU	0.741	0.656	0.810	0.843
	FFNN	0.700	1.832	1.354	0.246
	SVR	0.687	1.913	1.383	0.851
	RF	0.606	2.409	1.552	0.243
DN40126	LSTM	0.656	0.025	0.158	0.094
	GRU	0.667	0.024	0.155	0.092
	FFNN	0.476	1.060	1.030	0.687
	SVR	0.444	1.123	1.059	0.727
	RF	0.407	1.199	1.094	0.786
D4N0127	LSTM	0.806	0.324	0.569	0.364
	GRU	0.762	0.398	0.631	0.405
	FFNN	0.674	0.428	0.654	0.504
	SVR	0.631	0.485	0.692	0.572
	RF	0.532	0.614	0.782	0.620
DN40130	LSTM	0.634	0.224	0.474	0.343
	GRU	0.613	0.235	0.485	0.353
	FFNN	0.571	1.800	1.342	0.736
	SVR	0.537	1.785	1.330	0.851
	RF	0.431	2.381	1.540	1.065
DN40139	LSTM	0.672	0.125	0.354	0.243
	GRU	0.671	0.126	0.355	0.246
	FFNN	0.605	0.394	0.638	0.451
	SVR	0.549	0.450	0.670	0.511
	RF	0.421	0.577	0.759	0.569
DN40142	LSTM	0.851	0.032	0.179	0.128
	GRU	0.804	0.035	0.186	0.139
	FFNN	0.813	0.033	0.182	0.132
	SVR	0.722	0.159	0.399	0.263
	RF	0.544	0.453	0.673	0.441
	LSTM	0.443	0.081	0.285	0.196

D4N0147	GRU	0.424	0.085	0.292	0.207
	FFNN	0.413	0.098	0.312	0.188
	SVR	0.328	0.255	0.505	0.324
	RF	0.231	0.491	0.697	0.599
D4N0146	LSTM	0.546	0.715	0.846	0.566
	GRU	0.381	0.975	0.984	0.815
	FFNN	0.502	0.387	0.622	0.313
	SVR	0.422	0.450	0.671	0.319
	RF	0.435	0.443	0.665	0.323
DN40687	LSTM	0.768	0.633	0.795	0.515
	GRU	0.778	0.607	0.779	0.486
	FFNN	0.726	0.656	0.810	0.518
	SVR	0.506	2.561	1.600	0.853
	RF	0.392	3.150	1.774	1.048
2625BB00028	LSTM	0.854	0.069	0.262	0.161
	GRU	0.825	0.083	0.287	0.198
	FFNN	0.717	0.134	0.365	0.282
	SVR	0.695	1.401	1.182	0.755
	RF	0.643	2.067	1.412	0.919
2526CC00033	LSTM	0.265	0.204	0.451	0.213
	GRU	0.254	0.207	0.455	0.211
	FFNN	0.220	0.280	0.530	0.343
	SVR	0.314	0.456	0.675	0.247
	RF	0.140	1.670	1.300	0.950

Table 4.2: Performance of developed models during testing phase for all 24 boreholes selected for modelling in the Grootfontein compartment

Table 4.2 shows how the LSTM and GRU model were at par with one another for several boreholes when comparing both the error metrics and R^2 values. The LSTM model however slightly outperformed the GRU model for boreholes DN40037, DN40065, DN40066, DN40108, DN40115 and DN40127. Similarly the FFNN and the SVR were also at par with one another with slight differences in error. The random forest had the highest error rates for all 24 boreholes modelled. The RF model was unable to capture the trend of the observed groundwater levels on several instances. This comes as no surprise considering the model is unable to capture non-linear patterns as well as ANNs or SVR models.

As evidenced in table 4.2 the LSTM outperformed all other models tested for most boreholes (17 out of 24). The GRU model on the other hand, slightly outperformed the LSTM model for 6 out of the 24 boreholes modelled. For borehole 2526CC00033 the SVR model outperformed all other models including the LSTM and GRU, however it should be noted that it also had the

highest MSE, RMSE and MAE compared to the RNN's models.

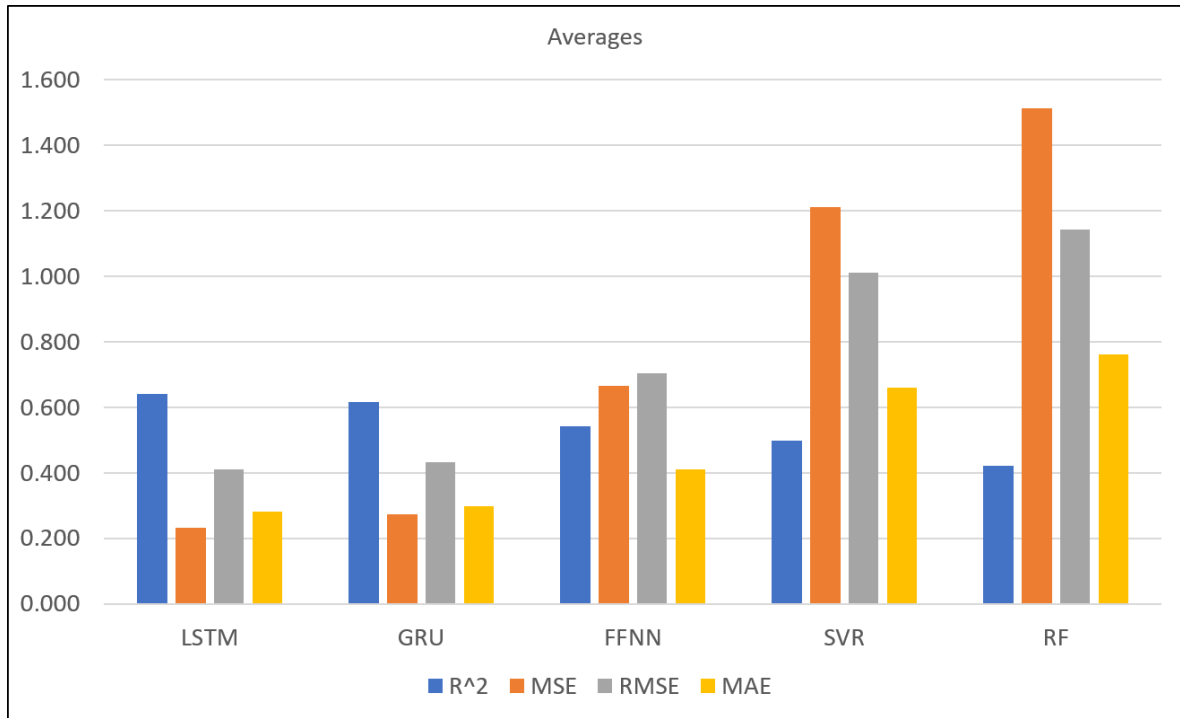
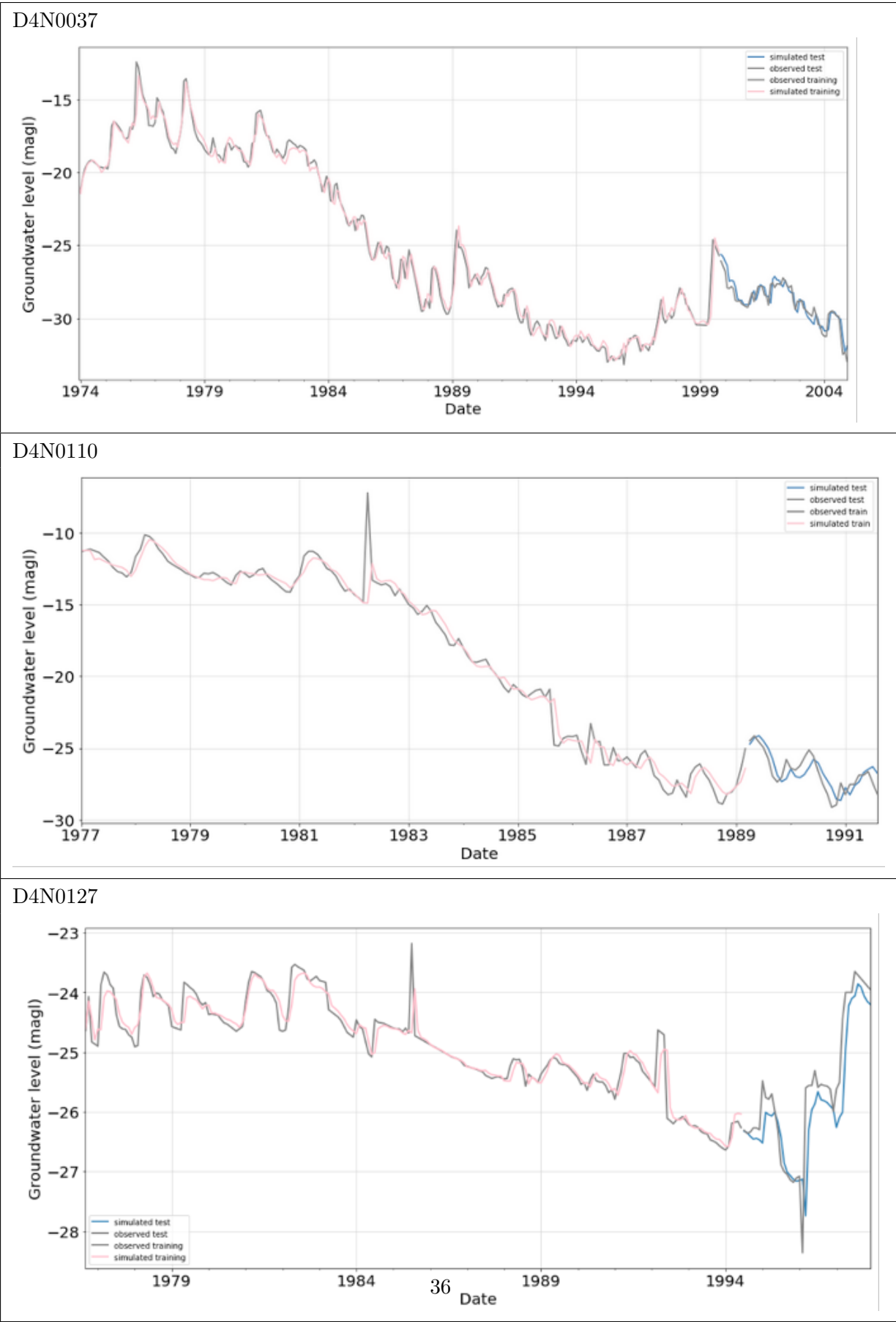


Figure 4.2: Bar graph showing model averages

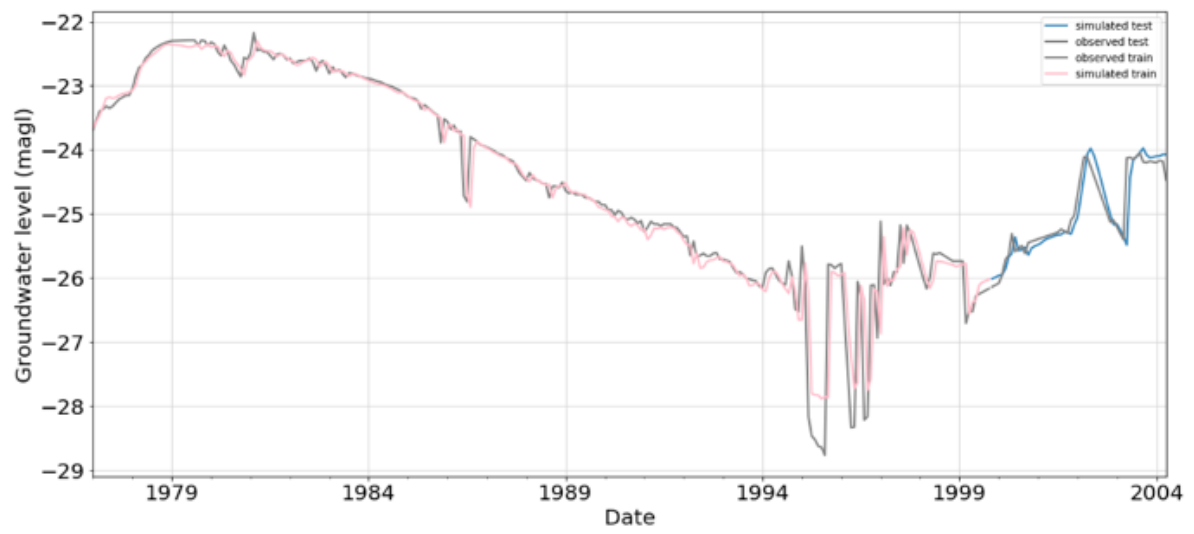
Figure 4.2 shows the evaluation metric averages of each model. First, the LSTM and GRU models produced the lowest MAE (0.282m and 0.298m respectively) and RMSE (0.411m and 0.433m) values on average. The two models also produced the highest R^2 values on average with the LSTM model (0.641) outperforming the GRU model (0.601). Secondly, the FFNN performed better compared to the other models with an average MAE value of 0.413m and RMSE value of 0.705m. The error metrics of the SVR and RF models were the highest recording an average MAE value of 0.660 and 0.763 respectively. The RMSE values of the two models were also high with an average of 1.011 for SVR model and 1.148 for RF model. From the results obtained in table 4.2 and figure 4.2 it is evident that the LSTM model achieved the best results compared to the 4 models tested.

Since the LSTM was the best performing model, we decided to focus more on the model for the prediction of groundwater levels in the compartment. A subset of some of the best and average performing boreholes shown in table 4.2 were selected to show a comparison of the predicted results versus the observed results for the LSTM model in table 4.3.

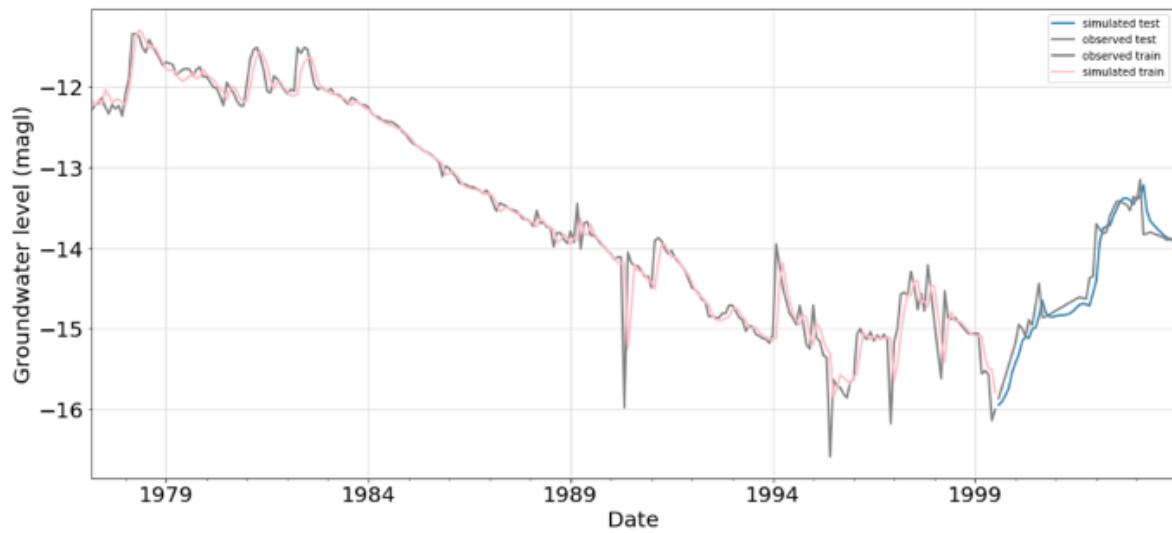
Table 4.3: Time-series of actual vs predicted groundwater levels for 6 selected boreholes in the Grootfontein aquifer



D4N0142



D4N0141



D4N0146

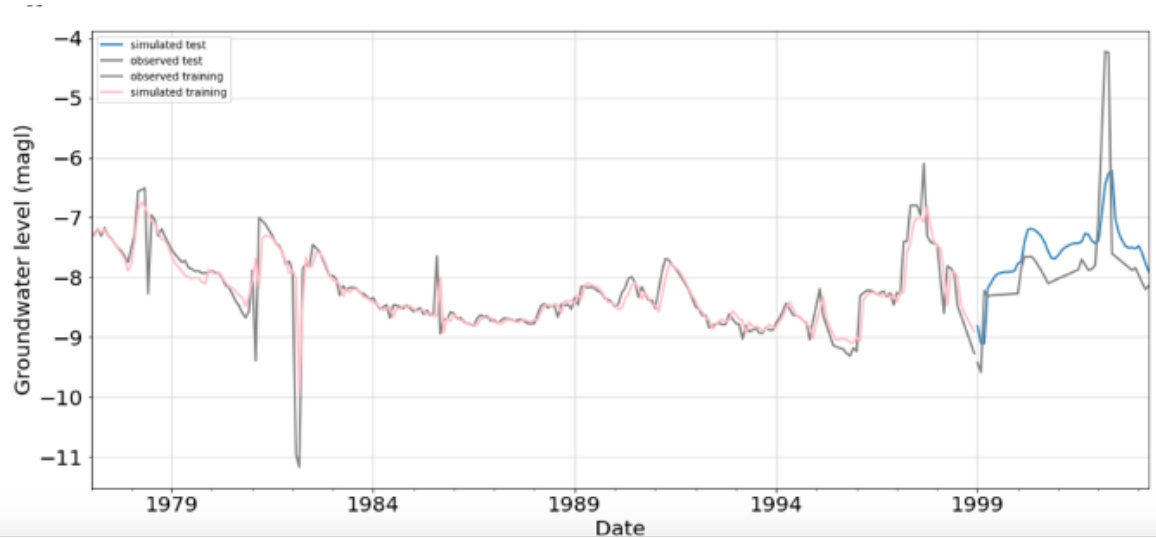


Table 4.3 suggests a good correlation between the predicted and observed values for boreholes D4N0141 and D4N0142 with key patterns replicated by the model. It should be noted that both boreholes also exhibited high R^2 values with borehole D4N0141 recording a score of 0.851 and D4N0142 a score of 0.821. The RMSE scores for both boreholes were low, showing the ability of the LSTM model to capture the underlying trends of the observed water levels.

Boreholes D4N0037 and D4N0127 performed better. For borehole D4N0127, the model seemed to under-predict values in the training set as compared to the test set. On the other hand, borehole D4N0146 seemed to over-predict groundwater level values which also explains why the borehole exhibited high error rates (MSE:0.715, RMSE: 0.864 and MAE:0.566). This could be due to precipitation and/or discharge influencing this borehole more than it did with other boreholes. Borehole D4N0110 seemed to under and over-predict groundwater level values throughout the test phase.

Furthermore, table 4.3 illustrates how outliers affect model performance. An observation is defined as an outlier if its values are outside the range of the normal values in the dataset. The model learns patterns from data during the training phase hence patterns which deviate from patterns previously seen decrease the overall performance of the model. This can be seen during the testing phase of boreholes D4N0146 and D4N0127. During the test phase groundwater levels for borehole D4N0146 had a sudden peak which occurred during 2003. Such a peak was previously not seen by the model during the training phase hence the model was unable to accurately predict the peak which in-turn decreased model performance.

RNNs generally perform better the more variability it sees during the training phase. This is evidenced by the three troughs seen in borehole D4N0146 during 1978, 1981 and 1982. The model under-predicts the first two troughs but by the time it reaches the third trough the model can better predict the trough. In addition, the model is unable to pick up the variability seen in the test set for borehole D4N0127. This is because the training data is not representative of the test data (variability happening in test set is different from the train set). Due to this high variability model errors also increase significantly.

What is common across most boreholes for both the training and testing phases is that the LSTM model is often unable to capture the minimum and extreme groundwater level peaks and troughs whenever they occur within a short timeframe. This is illustrated in table 4.3 where we see that across all boreholes the model failed to reproduce the extreme minimums and maximum values of each borehole. For example, the minimum groundwater levels in 1990, 1995 and 1997 for borehole D4N0141, 1995 - 1997 for borehole D4N0142 and 1981 and 1982 for borehole D4N0146 are not well captured. Similarly, extreme peaks are also not well replicated as seen in 1982 for borehole D4N0110, 2002 for borehole D4N0146 and 1985 for borehole D4N0127.

Generally speaking, the observed and modelled groundwater levels are closely matched for boreholes D4N0037, D4N0142 and D4N0146, with key features of the observed datasets replicated in the modelled data. On the other hand, boreholes D4N0110, D4N0127 and D4N0146 the LSTM model over and underestimated the observed groundwater levels.

4.3.3 Discussion and Conclusion

To predict groundwater levels in the dolomite aquifer, a total of 24 boreholes were tested. Firstly, a ranking by entropy was presented for each borehole against the 4 features to quantify the influence of each feature on the groundwater levels. Results revealed that abstraction, rainfall and discharge played a major role in the groundwater fluctuations in the compartment. This information will later be used as a basis for scenario testing. Next, we tested 5 different

models to select the optimal predictive model for groundwater levels. Out of the models tested the two RNNs performed well with the LSTM model slightly outperforming the GRU model.

To further analyse the results and select the optimum predictive model, average scores were computed. The LSTM model had the lowest error metrics and the highest R^2 score hence it was selected as the optimal predictive model that can be used to forecast groundwater levels in the aquifer. We then took 6 boreholes as a case study to visually analyse model results of the LSTM model. From these we noted that although the LSTM model was performing well, it failed to capture the minimum and extreme groundwater level peaks and troughs whenever they occurred within a short time frame. Overall, the results proved the vast capabilities of ANN models in modelling groundwater systems in the North West aquifers.

4.4 Scenario Testing Results

As previously highlighted, the LSTM model proved to be the optimal model in predicting groundwater levels in the Grootfontein compartment. Further investigation was therefore taken to explore the future trends of groundwater behaviour under different scenarios for the 6 boreholes chosen in section 4.3.2. Different precipitation and abstraction conditions were considered for the different case scenarios including worst case, average case and best case scenario. The results from this paper stem from Kanyama *et al.* [2021] study. Table 4.4 shows a summary of each case scenario. Results obtained from each case presented were compared against the base case seen in table 4.3.

The scenarios in table 4.4 were created to mimic possible changes in groundwater abstraction and climate conditions. Amplifying and decreasing rainfall peaks are both case scenarios seen in drought conditions and doubling abstraction amounts is inevitable due to increase in population. When running model simulations the training data was left unchanged and the test dataset for the rainfall or abstraction variable was altered according to the case scenario to be modelled. For example, for scenario 1 the test dataset of the rainfall time series was altered from the actual measured values to the values that mimic the case scenario (decrease rainfall peaks above the 100mm threshold by 100mm) and that was fed as model input with the rest of the other variables remaining unchanged.

Table 4.4: Description of each case scenario

Scenario ID	Name	Description
1	Decrease rainfall peaks	If monthly rainfall peak $> 100\text{mm}$ then minus 100mm
2	Increase groundwater abstraction	Double the monthly abstraction
3a	Simulate worst case scenario	Combination of scenario 1 and 2
3b	Simulate average case scenario	Decrease rainfall: Half all monthly rainfall data-points
3c	Simulate best case scenario	Increase rainfall peaks: If monthly rainfall peak $> 100\text{mm}$ then add 200mm
4	Long term prediction (15 years)	Use 100% data for training and predict groundwater levels for the next 15 years using long term averages of input variables

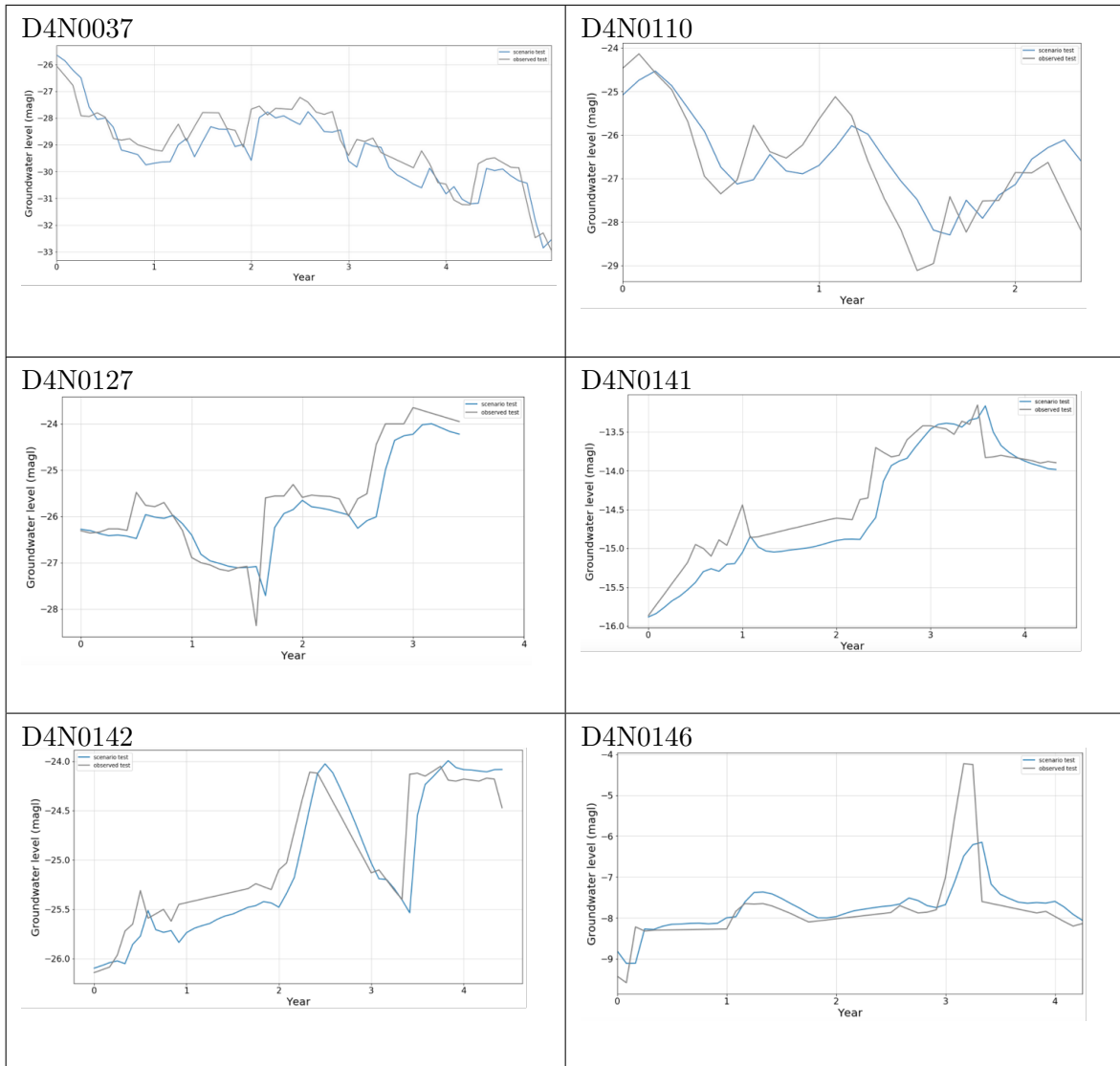
In the next sections we discuss the results obtained from each case scenario.

Scenario 1

Subtracting 100mm from rainfall events that had peaks higher than a 100mm threshold resulted in a drop of groundwater levels at the specific month where precipitation was reduced. It is important to note that precipitation peaks above the 100mm threshold were not abundant in the testing phase. Instead only a few high rainfall events occurred during 2000 – 2004 (duration of test data for most boreholes) that passed the required threshold for the case. For this reason, the drop in groundwater levels is only seen at the particular months where recharge was reduced by 100mm .

For boreholes D4N0142 and D4N0142, the decrease in water levels between year 0 to year 2 is observed as recharge events were reduced during this period. For borehole D4N0037, a decline in water levels is observed during year 2 to 3 as compared to table ?? where we have an over-predicted value of groundwater levels in 2002. As for borehole D4N0127 there was only a single rainfall event that was above the threshold during the testing phase hence only one negative change was observed with a drop of 0.2m at the beginning of year 3. The general trend for the rest of the years remained similar to that shown in our base case (table ??). For boreholes D4N0110 and D4N0146, a decline in water levels (0.2m to roughly 0.4m) can be observed from year 0.

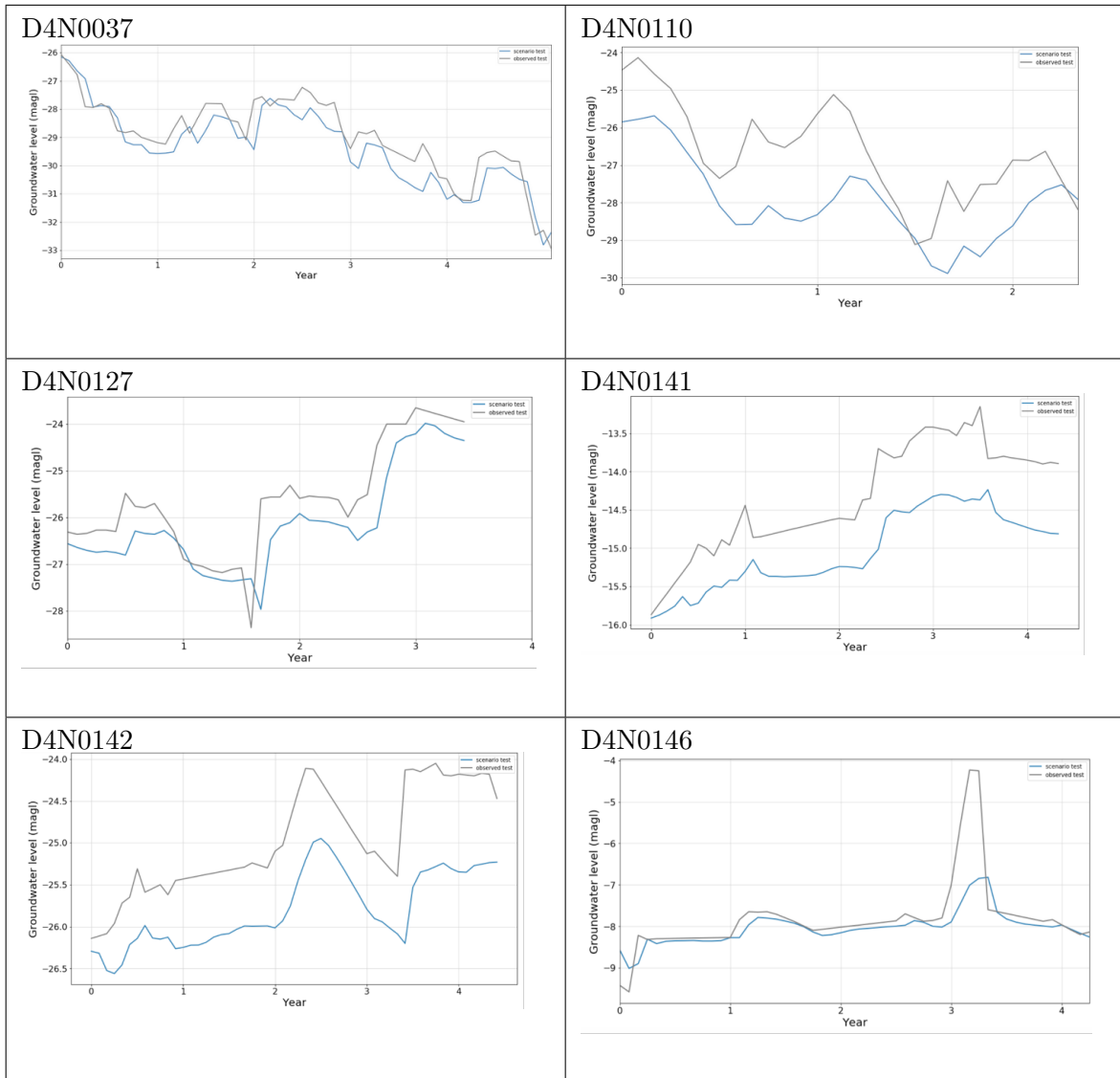
Table 4.5: Model prediction for scenario 1, decrease rainfall peaks



Scenario 2

Doubling the regional abstraction rates in the compartment resulted in a significant decrease of groundwater levels (ranging between 0.5 to 1.7m). As can be seen in figure 4.6, all boreholes had a significant decline in water levels as compared to the base case. Although the doubled abstraction rates seemed to have a higher effect on boreholes D4N0110, D4N0141 and D4N0142, the magnitude for boreholes D4N0127, D4N0037 and D4N0146 was much smaller.

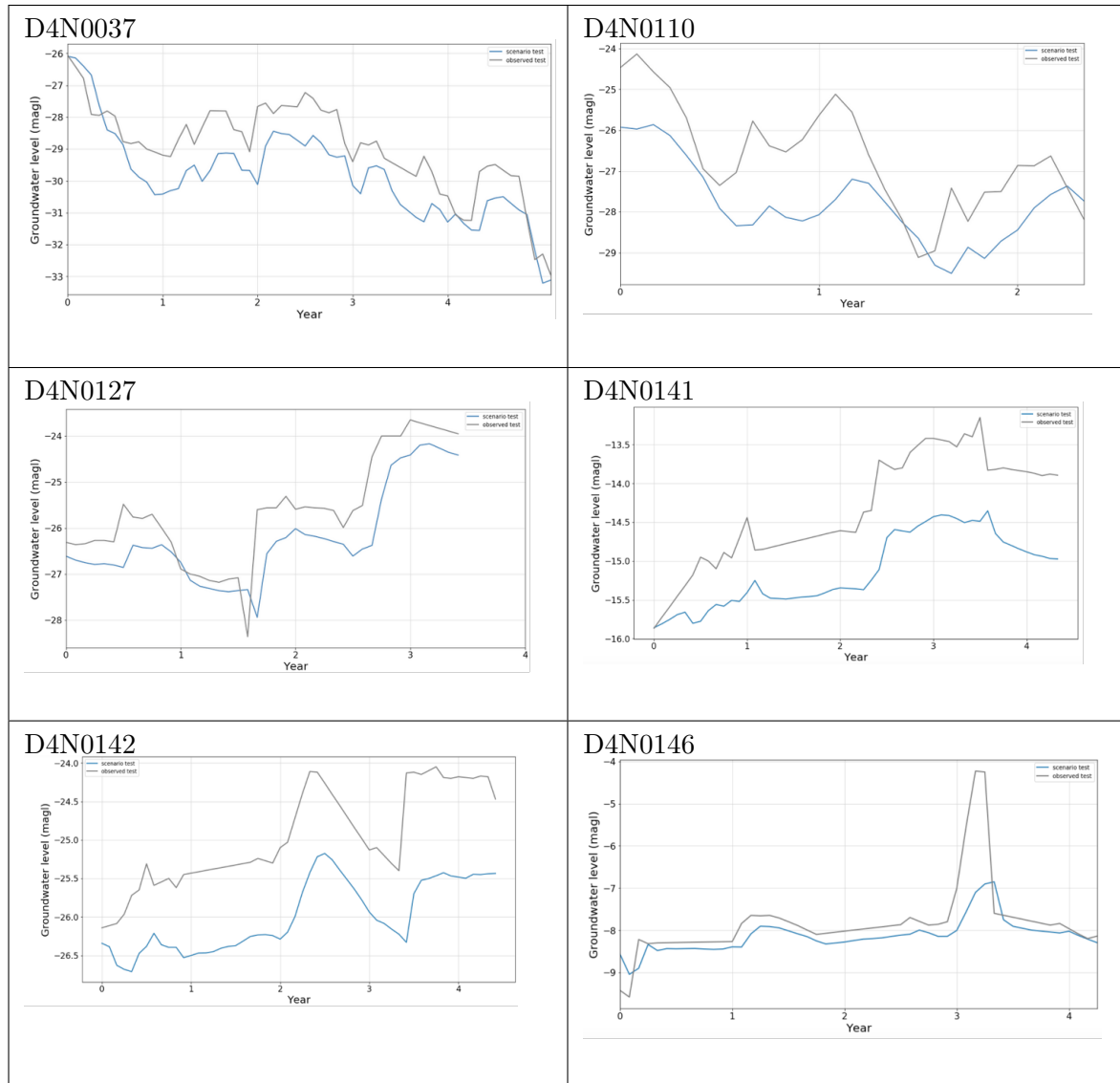
Table 4.6: Model prediction for scenario 2, Increase groundwater abstraction



Scenario 3a

The highest negative change was observed from the third scenario were regional abstraction was doubled and recharge was reduced by 100mm. As seen in figure 4.7, all selected boreholes show a deep decline in groundwater levels. This indicates that if groundwater pumping were to increase and recharge to decrease then groundwater resources could be at risk of diminishing. Groundwater levels for boreholes D4N0141 and D4N0142 dropped by almost 1m as observed in figure 4.7. For boreholes D4N0037, D4N0127, D4N0110 and D4N0146 groundwater levels dropped by at least 0.5m.

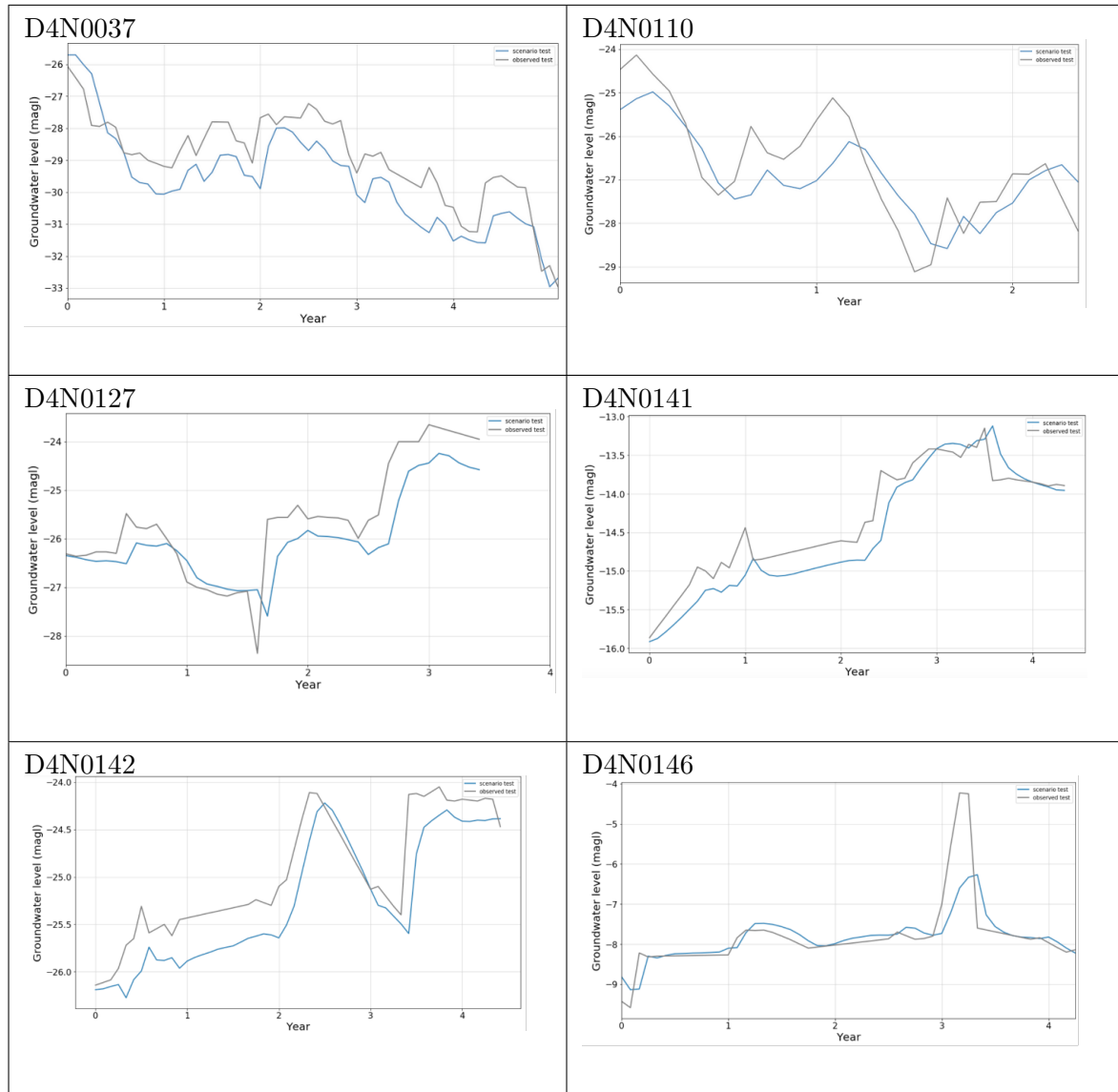
Table 4.7: Model prediction for the worst case scenario, increase abstraction + decrease rainfall peaks



Scenario 3b

By halving all precipitation values in the test set, we can see a clear decline in groundwater levels for all boreholes presented in figure 4.8. The decline in groundwater levels is consistent throughout the years since all precipitation values were halved unlike in scenario 1 where only specific rainfall events were tweaked hence the water levels only declined at particular months. The LSTM model is clearly able to predict the influence of precipitation on groundwater levels in the compartment evidenced by the constant drop in water levels for all boreholes when compared to the base case.

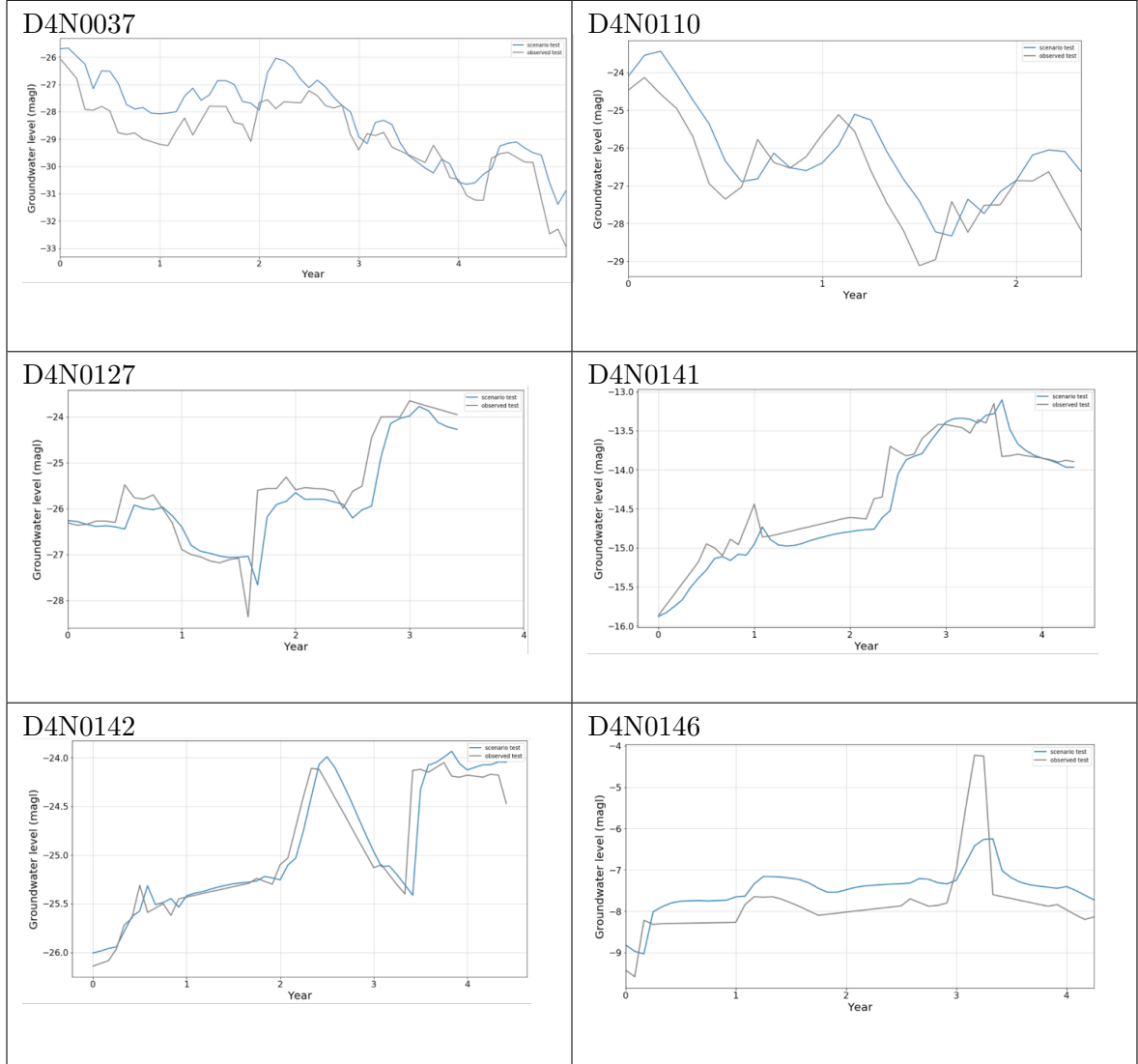
Table 4.8: Model prediction for average case scenario, decrease events by half



4.4.1 Scenario 3c

Similar to scenario 1 boreholes in this scenario showed an increase in groundwater levels in particular months were the rainfall event meet the required threshold for increase. The increase in precipitation resulted in an increase in groundwater levels. This can be seen more clearly in figure 4.9 for boreholes D4N0037, D4N0110 and D4N0146. In each of these boreholes the increase in recharge resulted in an increase in groundwater levels a lot higher than compared to boreholes D4N0142 and D4N0141. For borehole D4N0127 we can see an increase in water levels for year 3 while groundwater levels in the rest of the years remained almost stable. Generally, in this scenario the groundwater levels increased by 0.6 to 1.3m when compared to the base case.

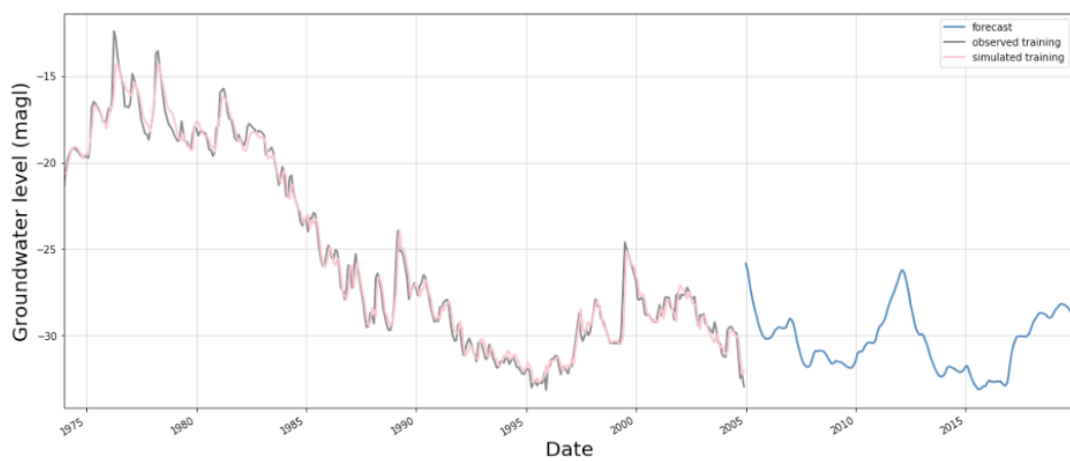
Table 4.9: Model prediction for scenario 1b, increase rainfall peaks



4.4.2 Scenario 4

For the final scenario, precipitation, abstraction, temperature and discharge datasets were formulated by averaging the 30years of data for each variable and predicting 15years into the future. Figure 4.3 shows a better prediction of groundwater levels by simply feeding the model averaged values of precipitation, temperature, discharge and abstraction. Due to fact that the model performed adequately for the base case were the model had actual data for each variable, it is not a far-fetched assumption to imply that the LSTM model would be able to predict long term trends up to 15years given sufficient training data.

Figure 4.3: Model prediction for scenario 4, long term prediction for borehole D4N0037



Conclusion

The purpose of scenario testing was to simulate the effects of dynamic groundwater abstraction and precipitation on groundwater levels in the Grootfontein aquifer, so that groundwater resources can be managed more effectively. As highlighted previously, South Africa is expected to experience more low rainfall days and high water demand hence groundwater availability becomes a concern for water managers and city councils including those in the North West. Four case scenarios were considered to investigate the effects of spatial and temporal variability of precipitation and groundwater abstraction on groundwater levels in the compartment using the LSTM model.

Scenario testing results showed a deep decline in groundwater levels when abstraction amounts were doubled and a more subtle decline when rainfall peaks amounts were reduced. Moreover, during what can be conspired as drought years the groundwater levels in the compartment decreased by almost 2 m. This shows that increasing groundwater abstractions during dry years may result in adverse reductions on groundwater levels in the compartment. These results indicated the importance the chosen variables had on groundwater levels in the region further strengthening our entropy results in section 4.3.1.

The framework used here could be used by water managers and city councils to establish sustainable groundwater allocation limits in certain areas and manage water resources more effectively during drier years while ensuring that groundwater levels do not deplete. In the next section we will discuss spring discharge forecasting results

Chapter 5

Spring Discharge Forecasting

5.1 Introduction

The second part of this study focused on using data driven approaches for modelling discharge rates in a groundwater-fed spring emanating from the Grootfontein dolomite aquifer, within the North West province. Springs emanating from the Grootfontein compartment are important sources of water as a significant amount of the population in the Mafikeng and other surrounding areas heavily depend on water obtained from these karst springs. In this context, it becomes essential to accurately predict discharge from the compartment so that water resources can be managed more effectively.

In South Africa, there is a limited body of literature in the forecasting of discharge using machine learning models. This could be because there are limited gauge stations recording discharge data hence data is also limited or in some cases data is not available at all. The objectives of this section is therefore to determine the primary factors governing spring discharge patterns using MI and to develop a model that can be used to forecast future discharge rates in the compartment and other surrounding areas.

Numerous studies have produced satisfactory results for discharge prediction using neural networks; however, the way to take into account the temporal behaviors of the groundwater levels in the surrounding area is not well apprehended and necessitates are more into using weather parameters and past discharge rates as inputs. We therefore propose the use of groundwater levels to not only predict discharge data but to generate spring discharge where there are insufficient records in an area.

Temporal data including temperature, precipitation, groundwater levels and licensed regional abstraction were used as model inputs and natural groundwater discharge from the Molopo spring as the model output. Feature analysis was performed to select the most prominent feature set. Using selected feature sets, two scenarios (explained in section 5.2 of this chapter) were tested to identify the best scenario /feature sets for long-term prediction of discharge rates. The results of this section stem from Kanyama *et al.* [2020a] study.

In the next sections we discuss model setup and model results.

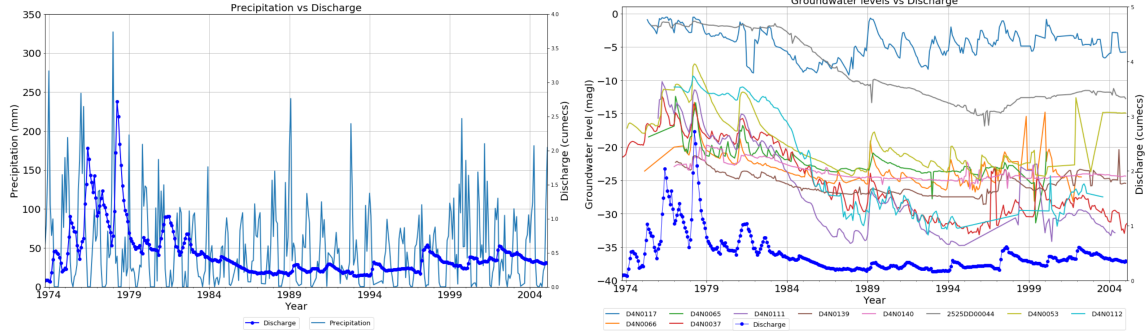


Figure 5.1: Hydrographs of the spring discharge at Molopo in relation to (a) precipitation and (b) subset of groundwater levels.

5.2 Model Setup

The spring discharge at Molopo is considered to be hydraulically connected to groundwater across the Grootfontein compartment. Therefore, characteristics of the feature sets which may influence model prediction were also of interest. As a result, two different combinations of input data sets were chosen for the study.

In the first scenario, the dataset consisted of 16 input features (13 groundwater level boreholes, precipitation, temperature and regional abstraction) and in the second scenario the dataset consisted of 3 input features (precipitation, temperature and regional abstraction). MI was used for feature selection and three different machine learning models namely: FFNN, LSTM and GRU models were trained and tested to predict discharge rates at the Molopo station. These networks were selected due to their ability to model groundwater dynamics well in chapter 4. Similar evaluation metrics mentioned in section 3.7 were used to gauge model performance.

Similar data pre-processing and model setup techniques adopted in section 4.2 were also applied in discharge forecasting including the architecture of the neural network models.

5.3 Experimental Results

5.3.1 Information Gain Ranking

Using MI, the study deduced the contribution of each feature to the target variable. The results of applying MI are summarized in table 5.1. The first column indicates the feature name and the next column indicates the corresponding scores of each feature to the target variable. Features that begin with "D4N..." and "2525..." are all borehole IDs.

Table 5.1 shows only a subset of the boreholes tested. A total of 24 boreholes were run using MI but only those boreholes that had a correlation of above 0.5 were selected for display and modelling. A total of 13 boreholes met this criterion. As can be seen from the results presented groundwater levels do indeed exhibit a high correlation with discharge in the compartment with 5 boreholes namely: D4N0037, D4N0111, D4N0140, D4N0115, D4N0112 recording entropy values of above 0.7. These results support the assumption that discharge in the compartment is driven by the hydraulic gradient between the surrounding aquifer and the eye, hence, high groundwater levels result in high discharge and low groundwater levels will also correlate with low discharge rates.

The rainfall feature recorded an entropy score of 0.41 with discharge. Spring discharge is

Table 5.1: A ranking of the mutual information for a subset of features to predict discharge

	Feature	Entropy
0	D4N0117	0.541
1	D4N0130	0.567
2	D4N0066	0.631
3	D4N0065	0.659
4	D4N0037	0.733
5	D4N0111	0.772
6	D4N0139	0.588
7	D4N0140	0.710
8	2525DD00044	0.601
9	D4N0053	0.688
10	D4N0115	0.714
11	D4N0108	0.525
12	D4N0112	0.724
13	Precipitation	0.417
14	Temperature	0.002
15	Abstraction	0.394

ultimately a consequence of rainfall, which infiltrates to groundwater, increasing groundwater levels and subsequently increasing discharge. The long-term decrease in spring discharge also correlates with the decline in regional precipitation as seen in figure 5.1. Therefore, a strong correlation between the two features was to be expected. The weak score of precipitation (as compared to groundwater level scores) could be due to the fact that it is only the larger rainfall events (above a 100mm threshold) that generate recharge and cause an increase in discharge as shown in figure 5.1. A clear response between discharge and intense rainfall events (above 230mm threshold) can also be seen in figure 5.1.

The temperature could influence evapotranspiration and therefore groundwater recharge rates which could influence temporal variation in spring discharge. However, for this case, temperature exhibited a low correlation score (0.002) which meant the variable was either too noisy or just irrelevant in relation to discharge. To factor out the reasons for this low correlation, the temperature dataset was decomposed using time series decomposition in python. The decomposed signal was then used as input in MI procedure. Although the entropy result improved (0.12) the correlation was still low. This meant that the feature was irrelevant in predicting discharge.

Lastly, abstraction exhibited a score of 0.394 in relation to discharge. Regional abstraction could result in discharge declines due to the groundwater levels lowering hence the relationship between the two was expected.

MI proved to be a useful framework for feature elimination in discharge forecasting as 11 boreholes were discarded as input features after feature analysis.

5.3.2 Model Results

This section will present the results obtained by the different networks for the two case scenarios using MAE and R^2 as evaluation metrics. Scenario 1 consists of 16 input features, namely: precipitation, temperature, regional abstraction and 13 boreholes that were selected during feature elimination. Scenario 2 consists of only 3 input features, namely: precipitation, temperature and regional abstraction.

To validate the effectiveness of the proposed models, the neural network models are tested for both scenarios. The model results are then compared with our baseline model, a FFNN which is the most widely used in literature. Table 5.2 summarizes the results of the different models on the test dataset.

Table 5.2: Experimental results of the different model and scenarios tested

Scenario	Models	R^2	MAE
Scenario 1	FFNN	0.661	0.058
	LSTM	0.855	0.046
	GRU	0.811	0.047
Scenario 2	FFNN	0.134	0.421
	LSTM	0.291	0.219
	GRU	0.223	0.334

The results presented in table 5.2 show that by including previous groundwater levels, all networks significantly improved in their performance as compared to without, recording MAE scores of between 0.047 - 0.058m³/s. The results obtained during the testing phase reveal that RNNs can outperform our standard baseline FFNN model which performed poorly for both case scenarios. As illustrated in table 5.2, the baseline model had the highest error scores for both scenario 1 (0.058) and scenario 2 (0.421).

The poor results obtained by the FFNN as compared to the RNNs can be explained by the fact FFNN models lack feedback connections found in RNNs. RNNs have feedback connections that make them inherently dynamic in nature Haykin [1999]. This feedback connection allows RNNs to predict future values over an indefinite lead time allowing the network to perform tasks that FFNN cannot. As a network which remembers previous inputs, a RNN has greater success in determining time dependent patterns Brezak *et al.* [2011].

Results presented in table 5.2 show that the two RNNs performance were at par with one another. Although the performance of the GRU (0.047) was close to the LSTM (0.046) in terms of the MAE, the LSTM model recorded the highest R^2 score of 0.855 as compared to the GRU which recorded 0.811.

Since the LSTM model proved to be a promising tool in predicting future discharge rates over an indefinite lead time, the model was then used for multi-step forecasting of discharge rates for the next 2 years using 32 years of observed data (see figure 5.2). The results presented in

the graph show a reasonable prediction of future discharge rates as over/under prediction of observed discharge remained below 2% during the training phase.

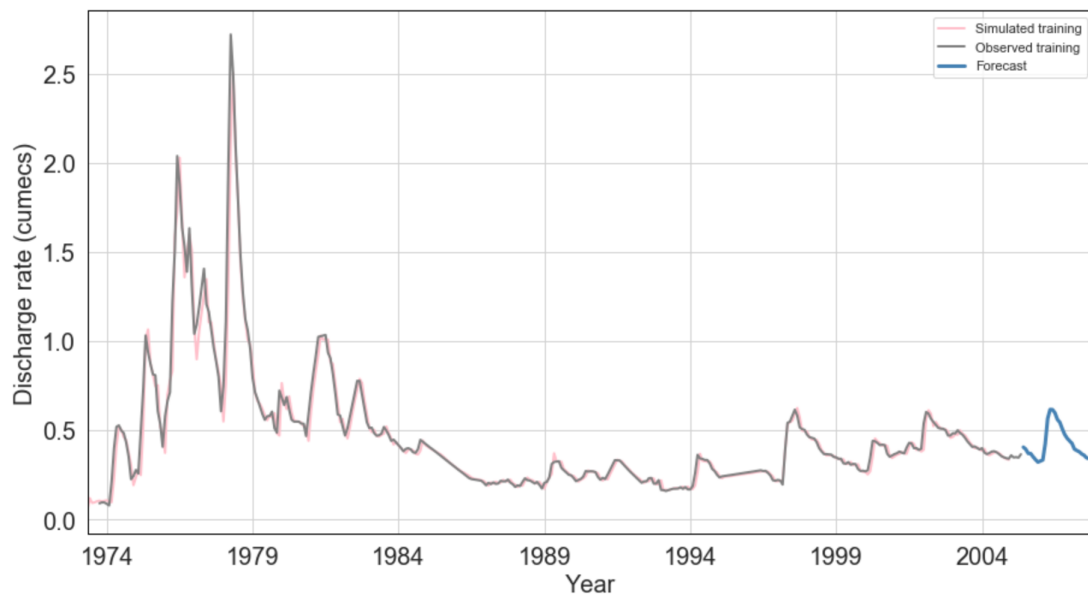


Figure 5.2: Foretasted discharge rate over a 2-year period

5.4 Discussion and Conclusion

The utility of time series records for spring discharge analyses is often dependent on uninterrupted, continuous observations. However, interruption is often unavoidable but such interruptions could negatively impact the sustainable management of water resources. This study proposes the use of groundwater levels, precipitation and abstraction to infill incomplete discharge records or generate discharge data.

In section 5.3.2 we demonstrated that feature selection techniques are useful tools to reduce the number of features. 13 out of 24 groundwater levels were used based on a cut off criteria value of 0.5 obtained from using MI. The target of this work was to successfully predict discharge using nearby groundwater level measurements and precipitation. In South Africa, not many gauge stations are present in each region (in some cases there are none) and in most cases gauge data is limited (data gaps are significant).

Our proposed methodology demonstrated a technique for not only infilling discharge gaps but for generating an entire discharge dataset using only a few surrounding boreholes, abstraction and a rainfall datasets available in the region as model inputs. Furthermore, our results revealed that groundwater levels, abstraction and precipitation contribute more to the success of machine learning models compared to only using weather parameters and abstraction data as model inputs.

However, it should be noted that the significance of this work to the hydrology field rests on the ability to find meaningful groundwater level and rainfall datasets within the study area. As a result, the suggested methodology is only as good as the data fed into it. The model presented in this study could prove as a useful tool for predicting spring discharge in karst regions of South Africa. The modelling results could also potentially help other researchers to obtain accurate results in other karst regions.

Chapter 6

Application of Learned Model

6.1 Introduction

In machine learning, the generalizability of models is an important factor. The generalizability (or robustness) of a model is a measure of its successful application to data sets other than the one used for training and testing. The final research objective of this study was to provide evidence of the effectiveness of our chosen methodology by applying the optimal model chosen to predict groundwater levels and discharge rates in a different karst aquifer to test model generalizability.

The developed LSTM model which outperformed all other models when predicting groundwater levels and discharge rates was used to model three boreholes from the Steenkoppies compartment. In doing this, the generalizability of the model was tested, albeit in similar settings (similar rapid response to rainfall events), and with similar data intervals (monthly datasets). It should be noted that the same model approach or pipeline followed to model groundwater levels in the Grootfontein compartment (mentioned in chapter 4) was used to model groundwater levels in the Steenkoppies.

6.2 Data Acquisition

The Steenkoppies aquifer is situated in the Gauteng province, west of Tarlton. The aquifer provides good quality water primarily for local agriculture [Cobbing 2017]. Groundwater level and discharge data were sourced from DWS. The discharge station (Maloney’s eye) serves as the only natural significant outlet for groundwater in the compartment.

Similar to the Grootfontein aquifer, abstraction data was obtained from WARMS. Rainfall data was sourced from a study conducted by Holland *et al.* [2009] where the authors combined rainfall data from four weather stations to form a representative time series (from 1908 - 2015) of the weather conditions in the Gauteng region. To create a time series dataset to extend until 2019, the dataset from Cobbing [2017] was then augmented with data from the Tarlton station maintained by the Agricultural Research Council (ARC). Temperature data on the other hand was sourced from the SAWS and ARC. A time series of temperature was formed by averaging the two datasets to form one continual time series. Discharge data was also sourced from DWS. Same pre-processing steps as those carried out in the Grootfontein aquifer in section 4.2 were applied. The results obtained from the LSTM model for the 3 boreholes will be discussed in the following section.

6.3 Experimental Results

To test the generalizability of the LSTM model, we attempted to use the same model pipeline used on the Grootfontein aquifer in the Steenkoppies compartment. Boreholes chosen show different characteristics as compared to the Grootfontein boreholes. For example, two of the chosen boreholes show less spikes when compared to the Grootfontein boreholes and only one borehole shows similar spikes to those in the Grootfontein aquifer. The most distinct characteristic between the 2 datasets is that the Steenkoppies generally shows much higher groundwater levels as opposed to the lower groundwater levels shown in the Grootfontein aquifer caused by high abstraction rates (mainly for irrigation purposes).

Table 6.2 shows a summary of the LSTM model results when modelling Steenkoppies data. When looking at the error scores for the model, the scores were considerably higher than the error scores received for boreholes in the Grootfontein compartment. This means the groundwater levels in the Steenkoppies compartment could potentially be influenced by other factors other than the ones used to model. Borehole A2N0553 had the lowest error scores and the highest R^2 of the tested boreholes. It should be noted that the results in the train set were mostly driven by discharge in comparison to the other variables. This was verified through a series of tests of matching each input variable against the predicted model values. On average the 3 boreholes recorded R^2 values of 0.61m. RMSE values of 0.27m and MAE values of 0.22m.

Table 6.2: Summary of model results

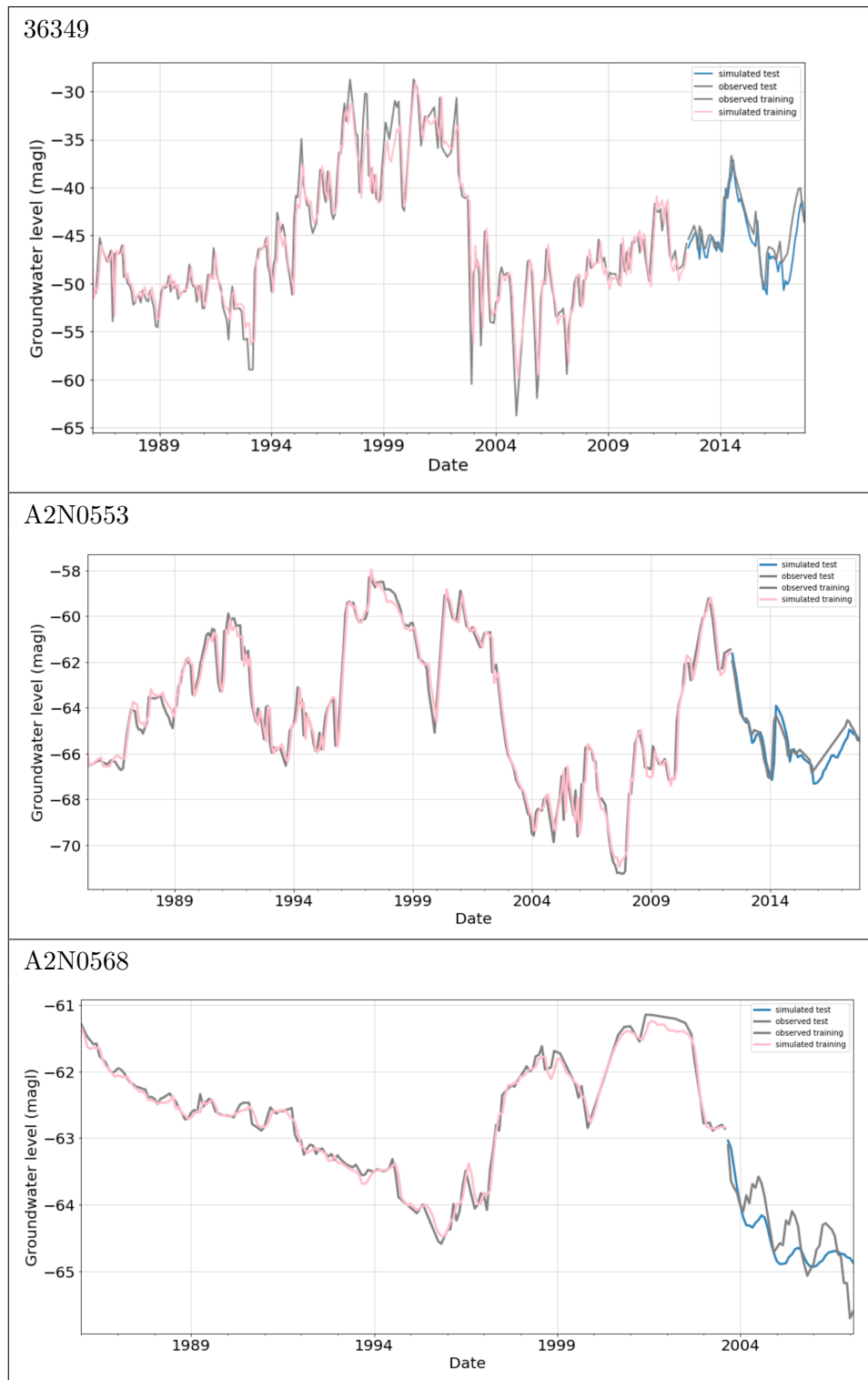
	BHID	R2	MSE	RMSE	MAE
1	36349	0.63	0,08	0,28	0,21
2	A2N0568	0,45	0,14	0,37	0,31
3	A2N0553	0,74	0,03	0,17	0,14
	Average	0,61	0,08	0,27	0,22

Moreover, as shown in table 6.1, the LSTM model was still not able to model the sudden peaks and troughs present in the Steenkoppies, similar to the results we saw in the Grootfontein compartment. The declining trend shown in borehole A2N0568 was unable to be modelled well by the LSTM model. This is since the declining trend is different from the declining trend presented in the training phase hence the model errors were also significantly high.

6.4 Discussion and Conclusion

Generalisability of a model is important as a model becomes more valuable when it can make predictions on a different dataset than it was intended. To test the model generalisability of the developed LSTM model, predictions were made on groundwater level data from a different compartment using the same model pipeline. The LSTM model was able predict groundwater levels in both the Steenkoppies and Grootfontein compartment with high accuracies for 2 out of the 3 chosen boreholes. The LSTM model was however unable to model the unexpected peaks and troughs present in the datasets. However, the results of the LSTM model simulations show the ability of the model to capture the broad patterns of the Steenkoppies dataset so we assume that the LSTM model is likely to be able to generalize datasets from different regions.

Table 6.1: Time-series of actual vs predicted groundwater levels for 3 selected boreholes in the Steenkoppies aquifer



Chapter 7

Conclusion and Future Directions

7.1 Introduction

This research aimed at demonstrating what conditions, variables and feature sets would be necessary for machine learning models to better predict groundwater levels and discharge rates particularly in the Grootfontein aquifer. As highlighted previously, prediction of groundwater levels and discharge rates is essential for the sustainable management of groundwater resources. The preceding chapters have laid out the context of the aquifer under study, the literature and the research methodology in detail. The preceding chapters have also, in detail, examined the use of different machine learning models and presented an optimal model for the prediction of groundwater levels and discharge rates. This chapter provides a summarized overview of the chapters mentioned previously and the results obtained from the study in relation to the research aims mentioned in chapter 1.5.

7.2 Literature Review

Due to their dynamic nature, groundwater systems continually adjust to long-term and short-term changes in land use, groundwater withdrawal and climate change (including temperature, precipitation and evapotranspiration). Traditionally, groundwater flow models have been the basis for the quantitative analysis of groundwater resources. However as discussed in chapter 2 these models require extensive physical observations for data collection which are both time consuming and costly. An alternate to these models is the use of data driven methods with machine learning techniques.

The literature review presented in chapter 2 highlighted the potential of several machine learning models in the prediction of groundwater systems. The review presented several studies where comparisons between machine learning models and groundwater flow models were made and in each research machine learning models outperformed the traditional models. However, the review also highlighted the limitations to which machine learning models could predict both groundwater levels and discharge rates to which this research would address. Therefore this chapter formed the basis of the framework used in this study.

7.3 Research Framework

This study made use of data obtained from the SAWS, WARMS and DWS. Datasets were filtered to select boreholes, discharge stations and weather stations with the longest continuous time-periods and with minimal data missing to avoid inaccurate results. Furthermore, to understand the importance of each input driver in relation to groundwater levels and discharge rates, the study investigated the impact of groundwater abstraction through pumping and climate change on both groundwater levels and discharge rates through mutual information. Using MI this study quantified the relationship between each input variable and the target variable. MI proved to be a useful framework for feature selection and ultimately improved model results.

Using the top features selected from MI, different machine learning models were tested to find the optimum model that best modelled our data and made the most accurate predictions. The LSTM, GRU, SVR and RF models were compared against the base model (FFNN). The top 3 models for groundwater level prediction were then selected for discharge prediction since those models proved to have great capabilities in predicting hydrological systems in the region. This study also conducted a model generalizability test to further strengthen the results of the modelling technique recommended in the study.

Lastly, RMSE, R^2 MSE and MAE loss functions were used to evaluate the prediction ability of different machine learning techniques.

In summary, to address the research aims, different scenarios, variables and feature sets were tested to identify the most appropriate features and variables that would improve machine learning models performance in groundwater levels and discharge rates prediction.

7.4 Research Findings

The primary objective of this study was to determine the primary factors governing spring discharge patterns and groundwater levels using mutual information (MI) and to develop a model that can be used to forecast future discharge rates and groundwater levels in the Grootfontein aquifer. Most studies mentioned in chapter 2.3 focused on the parameters that could be obtained without analysing whether these parameters had any influence in the prediction of the groundwater resource. Therefore we aimed at first performing feature analysis to select the most prominent features in (1) groundwater level prediction, (2) discharge prediction.

7.4.1 Groundwater Level Prediction

The results obtained through MI showed that regional abstraction, discharge and precipitation were the highest contributing factors to predict groundwater levels. An interesting observation came to light as the study through MI highlighted how decomposed signals of rainfall produced higher entropy scores (+0.30) compared to the actual measured rainfall values. Using these decomposed signals as an input to the machine learning algorithms selected significantly improved the results as the models became more sensitive to the decomposed signal compared to actual measured rainfall (which most studies in literature used).

The results obtained for the predictive models showed that the LSTM model was better suited for the prediction of groundwater levels compared to the other models tested. The detailed model setup where different parameters were tested among other things can be found in chapter 4.2. The observed and modelled groundwater levels were closely matched, with key features of

the observed datasets replicated in the modelled data highlighting the effectiveness of the chosen parameters of the LSTM model.

As evidenced in the problem statement in chapter 1, South Africa is expected to experience more low rainfall days and high water demand due to population growth in the next few years. Hence, our next objective was to attempt to predict groundwater levels under different weather conditions and groundwater abstraction amounts. The results showed the deep decline in groundwater levels when abstraction amounts were doubled and a more subtle decline when rainfall peaks amounts were reduced. These results could be used by water managers to limit groundwater usage in certain regions and manage water resources more effectively. Moreover, the results indicate that the chosen variables do indeed influence water levels in the region further strengthening our entropy results.

7.4.2 Discharge Prediction

The results obtained through MI showed that groundwater levels were the strongest feature in predicting discharge rates in the area. This was to be expected as discharge in the compartment was expected to be driven by the hydraulic gradient between the surrounding aquifer and the eye. Precipitation had a higher entropy compared to abstraction. This again was to be expected as spring discharge is ultimately a consequence of rainfall, which infiltrates to groundwater, increasing groundwater levels subsequent to recharge events and in-turn driving discharge. Therefore, discharge was expected to correlate with rainfall changes. Temperature again proved to have an insignificant relationship with the prediction of discharge rates hence through the results of MI the variable was entirely dropped from model predictions.

The results obtained from the predictive models highlighted two important facts. Firstly, the results showed the promising role of RNNs in the prediction of discharge rates. In both the two case scenarios tested the LSTM and GRU models were at par with one another. Secondly, the results showed that including groundwater levels from surrounding boreholes drastically improved the prediction of discharge. In literature the focus has been on using weather parameters and past discharge as model inputs, however this study showed that groundwater levels inclusion had a much better effect on the modelled results. Hence, we propose the use of groundwater levels, precipitation and abstraction to not only infill incomplete discharge records but to generate discharge data where there are no records at all.

7.4.3 Application of Learned Model

The final aim of this study was to test the model generalizability. We therefore tested the generalizability of the LSTM model in a different aquifer setting using the optimized LSTM model. The parameters used in the LSTM model when predicting the groundwater levels in the Grootfontein aquifer were unchanged to thoroughly test the effectiveness of the model in a different aquifer setting. The Steenkoppies aquifer was used as a case study. The results from this experiment proved the capabilities of the optimized LSTM model in predicting groundwater levels in a different setting. The LSTM model was able to capture the trends of the observed dataset well. But similar to the Grootfontein the model failed to capture the unexpected peaks and troughs present in the dataset. The results of the LSTM model simulations in this experiment highlighted the ability of the model to capture the broad patterns of the Steenkoppies data. We could therefore assume, that the LSTM model with the trained parameter settings from this study is likely to be able to predict groundwater levels from other regions.

7.5 Future research

This research focused mainly on the Grootfontein aquifer with limited data. Future research on this topic would look at the following:

- Increasing the amount of input data and iterations (for ANN models) could potentially increase model accuracy. However this will also require additional data sources and increased computing power. Increasing both these parameters would allow the model to train for longer with more information which could ultimately increase performance.
- The research can also be extended to other areas in South Africa not affected by man-made structures. This would limit the influences on the aquifer fluxes. Information on groundwater usage is not readily available. Therefore experimenting on boreholes that are unaffected by human abstraction could yield more accurate results.
- Extend the research of discharge predictions to other rivers. While the spring discharge at the Molopo eye plays an important role in the North west, prediction of other stations would help to assess the performance of the machine learning models on the entire catchment area.
- While this research only looked at generalising the model with a few boreholes, testing model generalization on a larger scale would help create a more robust model.
- This study used mutual information to quantify the relationship between input and output variables. Using probabilistic graphical modelling such as the Bayesian network to learn the conditional probability distribution between the input variable and the target variable could help the researcher better understand the exact influence of each input variable on the target variable.

These future research ideas would look at building on the findings obtained from this research and as such would allow the development of a machine learning model that could be used in real world application and help improve water management in South Africa.

7.6 Conclusion

This research evaluated the use of different machine learning models to predict groundwater levels and discharge rates in the Grootfontein aquifer. The research presented a ranking by entropy of features affecting groundwater levels and discharge rates in the Grootfontein compartment. Once the most important features were highlighted, different machine learning models from different paradigms were tested to select the optimal model to predict (1) groundwater level and (2) discharge rates in the compartment.

In both scenarios the LSTM model was presented as the best model for predictions, yielding good results as compared to other models tested. Using this model different real world case scenarios were tested to see their effect on groundwater levels. For discharge a multi-step forecasting was conducted to forecast discharge rates for the next two years using averaged data. Lastly to provide evidence of the effectiveness of the LSTM model on groundwater level prediction we tested model generalizability in a different aquifer setting.

The results presented in research study present a useful approach for discharge and groundwater prediction. The ever evolving field of machine learning allows the approach used in this study to continually be improved upon and become more usable in real world application that assists

hydrologists and water management in the sustainable use of groundwater resources in South Africa.

References

- [Afonja 2017] Tejumade Afonja. *Kernel Functions*, 2017.
- [Al-Aboodi *et al.* 2016] Ali H Al-Aboodi, Kifah M Khudhair, and Ali S Al-Aidani. Prediction of Groundwater Level in Safwan-Zubair Area Using Artificial Neural Networks. *Basrah Journal for Engineering Science*, 16(1):42–50, 2016.
- [Ayenew *et al.* 2008] Tenalem Ayenew, Molla Demlie, and Stefan Wohnlich. Application of Numerical Modeling for Groundwater Flow System Analysis in the Akaki Catchment, Central Ethiopia. *Math Geosciences*, 40:887–906, 2008.
- [Barnett *et al.* 2012] B. Barnett, L.R. Townley, V. Post, R.E. Evans, R.J. Hunt, L. Peeters, S. Richardson, A.D. Werner, A. Knapton, and A. Boronkay. *Australian groundwater modelling guidelines*. Technical report, 2012.
- [Basson *et al.* 2018] Lauren Basson, Salomé Bronkhorst, Nicholas Fordyce, and Claire Pengelly. *Water-2018 Market Intelligence Report-18 Roeland Street, Cape Town, 8001, South Africa Editorial and review*. Technical report, 2018.
- [Behzad *et al.* 2010] Mohsen Behzad, Keyvan Asghari, and Emery A Coppola Jr. Comparative Study of SVMs and ANNs in Aquifer Water Level Prediction. *Journal of Computing In Civil Engineering*, 24(5):408–413, 2010.
- [Bhattacharyya 2018] Indresh Bhattacharyya. *Support Vector Regression Or SVR*, 2018.
- [Breiman 2001] Leo Breiman. Random forests. *Machine learning*, 45:5–35, 2001.
- [Brezak *et al.* 2011] Danko Brezak, Bacek Tomislav, Majetic Dubravko, Kasac Josip, and Branko Novakovic. *A Comparison of Feed-forward and Recurrent Neural Networks in Time Series Forecasting*. Technical report, 2011.
- [Cao *et al.* 2003] L. J. Cao, K. S. Chua, W. K. Chong, H. P. Lee, and Q. M. Gu. A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. *Neurocomputing*, 55(1):321–336, 2003.
- [Chen *et al.* 2013] Xi Chen, Yanfang Zhang, Yanyu Zhou, and Zhicai Zhang. Analysis of hydrogeological parameters and numerical modeling groundwater in a karst watershed, south-west China. *Carbonates and Evaporites*, 2013.
- [Chen *et al.* 2020] Chong Chen, Wei He, Han Zhou, Yaru Xue, and Mingda Zhu. A comparative study among machine learning and numerical models for simulating groundwater dynamics in the Heihe River Basin, northwestern China. *Scientific Reports*, 10(1):1–13, 2020.
- [Chiang *et al.* 2004] Yen Ming Chiang, Li Chiu Chang, and Fi John Chang. Comparison of static-feedforward and dynamic-feedback neural networks for rainfall-runoff modeling. *Journal of Hydrology*, 290:297–311, 2004.

- [Chiroma *et al.* 2014] Haruna Chiroma, Sameem Abdulkareem, Adamu I. Abubakar, and Tutut Herawan. Kernel functions for the support vector machine: Comparing performances on crude oil price data. In *Advances in Intelligent Systems and Computing*, volume 287, pages 273–282. Springer Verlag, 2014.
- [Chitsazan *et al.* 2015] Manouchehr Chitsazan, Gholamreza Rahmani, and Ahmad Neyamadpour. Forecasting groundwater level by artificial neural networks as an alternative approach to groundwater modeling. *Journal of the Geological Society of India*, 2015.
- [Cobbing 2017] Jude Edmund Cobbing. *The Grootfontein aquifer at Mahikeng, South Africa as a hydro-social system*. PhD thesis, Nelson Mandela Metropolitan University, 2017.
- [Cobbing 2018] J. E. Cobbing. The North West Dolomite Aquifers, South Africa: A Stalled Opportunity for Water Security and Development. *Water SA*, 44(1):54–64, 2018.
- [Coppola *et al.* 2005] Emery A. Coppola, Anthony J. Rana, Mary M. Poulton, Ferenc Szidarovszky, and Vincent W. Uhl. A neural network model for predicting aquifer water level elevations. *Ground Water*, 2005.
- [Coulibaly and Baldwin 2008] P Coulibaly and C.K Baldwin. Dynamic neural networks for non-stationary hydrological time series modeling. *Journal of Hydrology*, 67:71–75, 2008.
- [Coulibaly *et al.* 2001] Paulin Coulibaly, François Anctil, Ramon Aravena, and Bernard Bobée. Artificial neural network modeling of water table depth fluctuations. *Water Resources Research*, 37(4):885–896, 2001.
- [Department of Environmental Affairs 2019] Department of Environmental Affairs. *National Climate Change Adaptation Strategy*. Technical report, 2019.
- [DiPietro and Hager 2019] Robert DiPietro and Gregory D. Hager. Deep learning: RNNs and LSTM. In *Handbook of Medical Image Computing and Computer Assisted Intervention*, chapter 21, pages 503–519. Elsevier, 2019.
- [Donnenfeld *et al.* 2018] Zachary Donnenfeld, Courtney Crookes, and Steve Hedden. *A delicate balance: water scarcity in South Africa*. Technical report, 2018.
- [DWS 2017] DWS. *Benchmarking of Water Loss, Water Use Efficiency and Non-Revenue Water in South African Municipalities*. Technical report, 2017.
- [Engelbrecht 2007] Andries P. Engelbrecht. *Wiley: Computational Intelligence: An Introduction, 2nd Edition - Andries P. Engelbrecht*. Wiley, 2007.
- [Gizaw and Gan 2016] Mesgana Seyoum Gizaw and Thian Yew Gan. Regional Flood Frequency Analysis using Support Vector Regression under historical and future climate. *Journal of Hydrology*, 538:387–398, 2016.
- [Gove and Faytong 2012] Robert Gove and Jorge Faytong. Machine Learning and Event-Based Software Testing: Classifiers for Identifying Infeasible GUI Event Sequences. In *Advances in Computers*, volume 86, pages 109–135. Academic Press Inc., 2012.
- [Gumbricht and Thunvik 1997] T Gumbricht and R Thunvik. Hydrogeological modelling with an expert GIS interface. *Hydrogeology Journal*, 28:329–338, 1997.
- [Haykin 1999] Simon Haykin. *Neural networks: a comprehensive foundation*, 1999.
- [Hiscock and Bense 2005] K. M. (Kevin M.) Hiscock and V. F. (Victor Franciscus) Bense. *Hydrogeology : principles and practice*. 2005.

- [Holland *et al.* 2009] M Holland, F Wiegmanns, J Cobbing, and K.T Witthüser.). *Geohydrological assessment of the Steenkoppies dolomite compartment. Project no. 14/14/5/2. Activity 25*. Technical report, Water Geosciences Consulting for Department of Water Affairs, Pretoria, 2009.
- [Kanyama *et al.* 2020a] Yolanda Kanyama, Ritesh Ajoodha, Helen Seyler, Ndivhuwo Makanda, and Hlanganani Tutu. Application of Machine Learning Techniques in Forecasting Groundwater Levels in the Grootfontein Aquifer. In *2020 2nd International Multidisciplinary Information Technology and Engineering Conference*, pages 167–174, South Africa, 2020.
- [Kanyama *et al.* 2020b] Yolanda Kanyama, Ritesh Ajoodha, Helen Seyler, and Hlanganani Tutu. Application of Artificial Neural Networks to Forecast River Discharge Rates: A Case Study of the Grootfontein Aquifer. In *The 7th Asia-Pacific Conference on Computer Science and Data Engineering*, Melbourne, Australia, 2020.
- [Kanyama *et al.* 2021] Yolanda Kanyama, Ritesh Ajoodha, Helen Seyler, and Hlanganani Tutu. Groundwater Level Estimation using Recurrent Neural Networks : A Case Study of the Grootfontein Aquifer. In *Interdisciplinary Research in Technology Management*, Kolkata, India, 2021.
- [Kenda *et al.* 2018] Klemen Kenda, Matej Čerin, Mark Bogataj, Matej Senožetnik, Kristina Klemen, Petra Pergar, Chrysi Laspidou, and Dunja Mladenčić. Groundwater Modeling with Machine Learning Techniques: Ljubljana polje Aquifer. *Proceedings*, 2(11):697, 2018.
- [Krause *et al.* 2005] P. Krause, D. P. Boyle, and F. Base. Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 5:89–97, 2005.
- [Liong and Sivapragasam 2002] Shie Yui Liong and Chandrasekaran Sivapragasam. Flood stage forecasting with support vector machines. *Journal of the American Water Resources Association*, 38(1):173–186, 2002.
- [Luo *et al.* 2015] T. Luo, R. Young, and P. Reig. Aqueduct Projected Water Stress Country Rankings. *World Resources Institute*, 2015.
- [Martin *et al.* 2016] Eric Martin, Simon Gascoin, Youen Grusson, Clément Murgue, Mélanie Bardeau, François Anctil, Sylvain Ferrant, Romain Lardy, Patrick Le Moigne, Delphine Leenhardt, Vincent Rivalland, José Miguel Sánchez Pérez, Sabine Sauvage, and Olivier Therond. *On the Use of Hydrological Models and Satellite Data to Study the Water Budget of River Basins Affected by Human Activities: Examples from the Garonne Basin of France*, 2016.
- [McCarthy and Skaggs 1991] E. J. McCarthy and R. W. Skaggs. Simplified model for predicting drainage rates for changing boundary conditions. *Transactions of the American Society of Agricultural Engineers*, 34(2):443–448, 1991.
- [Mirzavanda *et al.* 2014] M Mirzavanda, S Sadatinejad, H Ghasemiehc, M Akbarid, and H Shariatie. Groundwater level fluctuation forecasting using artificial neural network in arid and semi-arid environment. *Journal of Applied Hydrology*, 1(2):43–53, 2014.
- [Mitchell 1997] Tom M Mitchell. *Machine Learning*. McGraw-Hill Science, 1997.
- [Modisha 2017] Reshoketswe Caroline Oudi Modisha. *Investigation of the Ramotswa Trans-boundary Aquifer Area, groundwater flow and pollution*. PhD thesis, Witswatersrand, 2017.

- [Rodriguez-Galiano *et al.* 2012] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, (67):93–104, 2012.
- [Roulet 1991] N. T. Roulet. Surface level and water table fluctuations in a subarctic fen. *Arctic & Alpine Research*, 233(3):310, 1991.
- [Rushton 2005] K. R. Rushton. *Groundwater hydrology: Conceptual and computational models*. John Wiley & Sons Ltd, 2005.
- [SA-Venues 2019] SA-Venues. *North West Province Climate and Weather, South Africa*, 2019.
- [Sahoo and Jha 2013] Sasmita Sahoo and Madan K. Jha. Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment. *Hydrogeology Journal*, 21(8):1865–1887, 2013.
- [Song *et al.* 2007] Zhiyao Song, Ling Li, Jun Kong, and Honggui Zhang. A new analytical solution of tidal water table fluctuations in a coastal unconfined aquifer. *Journal of Hydrology*, (340):256–260, 2007.
- [Sreekanth *et al.* 2009] P. D. Sreekanth, N. Geethanjali, P. D. Sreedevi, Shakeel Ahmed, N. Ravi Kumar, and P. D. Kamala Jayanthi. Forecasting groundwater level using artificial neural networks. *Current Science*, 96(7):933–939, 2009.
- [Steyn and Brink 2018] Melise Steyn and W H Brink. *Short-term stream flow forecasting and downstream gap infilling using machine learning techniques*. PhD thesis, Stellenbosch University, 2018.
- [Sudharsan 2018] Asaithambi Sudharsan. *Why, How and When to apply Feature Selection*, 2018.
- [Taver *et al.* 2015] V. Taver, A. Johannet, V. Borrell-Estupina, and S. Pistre. Feed-forward vs recurrent neural network models for non-stationarity modelling using data assimilation and adaptivity. *Hydrological Sciences Journal*, 60(7):1242–1265, 2015.
- [The Coversation 2018] The Coversation. *Explainer: what there is to know about South Africa’s aquifers*, 2018.
- [Tsakiri *et al.* 2018] Katerina Tsakiri, Antonios Marsellos, and Stelios Kapetanakis. Artificial neural network and multiple linear regression for flood prediction in Mohawk River, New York. *Water (Switzerland)*, pages 1–20, 2018.
- [Upadhyaya and Chauhan 2001] A. Upadhyaya and H. S. Chauhan. Water table fluctuations due to canal seepage and time varying recharge. *Journal of Hydrology*, (244):1–8, 2001.
- [Villholth *et al.* 2013] Karen G. Villholth, Christian Tøttrup, Martin Stendel, and Ashton Maherry. Integrated mapping of groundwater drought risk in the Southern African Development Community (SADC) region. *Hydrogeology Journal*, 2013.
- [WRC 2003] WRC. *Understanding groundwater: A stakeholders guide to the North West Dolomite Aquifer*. Technical report, 2003.
- [Wu and Zeng 2013] JiChun C. Wu and XianKui K. Zeng. *Review of the uncertainty analysis of groundwater numerical simulation*, 2013.

- [Xu *et al.* 2012] Tianfang Xu, Albert J. Valocchi, Jaesik Choi, and Eyal Amir. Improving groundwater flow model prediction using complementary data-driven models. In *XIX International Conference on Water Resources CMWR*, 2012.
- [Xu *et al.* 2019] Lizhong Xu, Jia Zhao, Chenming Li, Changli Li, Xin Wang, and Zhifeng Xie. Simulation and prediction of hydrological processes based on firefly algorithm with deep learning and support vector for regression. *International Journal of Parallel, Emergent and Distributed Systems*, 2019.
- [Zadeh 2005] Lotfi A. Zadeh. Toward a generalized theory of uncertainty (GTU)- An outline. *Information Sciences*, (172):140, 2005.
- [Zhang and Qi 2005] G. Peter Zhang and Min Qi. Neural network forecasting for seasonal and trend time series. In *European Journal of Operational Research*, volume 160, pages 501–514, 2005.