International Conference on "Interdisciplinary Research in Technology & Management" (IRTM 2021)

Conference Proceedings

Edited by

Satyajit Chakrabarti
Director, IEM Kolkata, India

Rintu Nath
Scientist – F, Vigyan Prasar, Department of Science and Technology, Govt. of India

Pradipta Kumar Banerji
Dean (Management Studies), Institute of Engineering and Management, Kolkata

Sujit Datta
Institute of Engineering & Management, Kolkata, India

Sanghamitra Poddar
Institute of Engineering & Management, Kolkata, India

Malay Gangopadhyaya
Institute of Engineering & Management, Kolkata, India

First edition published 2022
by CRC Press
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

and by CRC Press
6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742

CRC Press is an imprint of Informa UK Limited

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalogue entry has been requested.

ISBN: 978-1-003-20224-0 (ebk)

DOI: 10.1201/9781003202240

Chapter 93

Groundwater Level Estimation using Recurrent Neural Networks: A Case Study of the Grootfontein Aquifer

Yolanda Kanyama
School of Computer Science and Applied Mathematics
The University of the Witwatersrand, Johannesburg, South Africa

Ritesh Ajoodha
School of Computer Science and Applied Mathematics
The University of the Witwatersrand, Johannesburg, South Africa

Helen Seyler
Delta-H Groundwater Systems Pty Ltd, Pretoria, South Africa

Hlanganani Tutu
Molecular Sciences Institute School of Chemistry
The University of the Witwatersrand, Johannesburg, South Africa

Abstract—Precise estimation of groundwater levels is essential for the management and sustainability of groundwater resources. The main objectives of this study are therefore (1) to quantify the influence of groundwater extraction, precipitation, temperature and discharge on the prediction of groundwater levels in the Grootfontein aquifer, (2) predict groundwater levels under different climate and groundwater extraction conditions using a recurrent neural network and (3) compare results of each case scenario with the base case for analysis. Selected datasets from feature analysis were fed into a recurrent neural network architecture to simulate the seasonal groundwater level changes. Feature analysis results revealed that the variables selected indeed had a strong influence on the prediction of groundwater levels on the selected boreholes. Discharge, groundwater abstraction and precipitation were the highest contributing factors to groundwater level fluctuation. A recurrent neural network model was used to simulate different case scenarios. The model results reveal that the neural network model was able to predict groundwater level change under the adjusted input variables. Groundwater level fluctuations were no more than 2 m below the base case for all scenarios tested. The deep learning techniques introduced in this study to estimate groundwater level change under different case scenarios can be convenient for groundwater management as drought warnings and/or water restrictions could be issued in a timely manner. We therefore suggest the use of the modelling framework used as an alternate approach to simulating groundwater level change specifically in areas where subsurface properties are not known.

Index Terms—recurrent neural networks, long-short term memory, groundwater levels, prediction, scenario testing, ground-water sustainability

I. 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 [1]. Groundwater supplies are often less susceptible to drought than surface water due to the larger reservoir. Where surface water, such as lakes and rivers, are scarce or inaccessible, groundwater supplies many of the hydrologic needs of people. In South Africa, groundwater resources not only provides water supplies for domestic use but also for the mining, agricultural and tourism industries [2]. However, due to extreme changes in weather conditions and the increase in population resulting in higher groundwater extraction rates, groundwater resources have been under enormous stress [2].

The Grootfontein dolomite aquifer is located in the North West Province of South Africa. The aquifer serves as a key source of water supply for municipal and irrigation purposes. Several ecosystems, wetlands and towns such as Mahikeng, Litchenburg and Grootpan rely on the aquifer as their main source of water [3]. The aquifer is also ecologically important in supplying springs that feed important rivers.

In 2017, groundwater levels in the Grootfontein aquifer had fallen by more than 28m compared to their pre-abstraction levels. This drop has been attributed to high groundwater extraction rates in the region [3]. In this context, it is evident how improved understating of groundwater level response to pumping and climate change are essential for sustainable planning and management of groundwater resources particularly in areas of changing weather conditions and increasing groundwater demands such as South Africa.

Machine learning models based on non-linear interdependence's serve as an attractive alternative to the traditional physical process models. One of the challenges in using physical process based models is their large data requirements for model development and calibration [4]. Machine learning models on the other hand are able to predict groundwater level changes without the deep knowledge of the underlying physical processes making them a more appealing choice. These models recognize patterns hidden in historical data and then uses those patterns for future prediction [5].

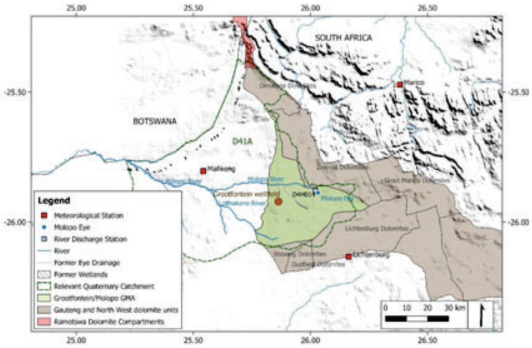

Fig. 1. Grootfontein location map (for credit of image see acknowledgements).

The Grootfontein aquifer was selected as the focus of this study. Temporal data including monthly temperature, precipitation, natural groundwater discharge from a groundwater-fed spring and groundwater extraction were used as model inputs and groundwater levels from four monitoring boreholes as model outputs. Root mean squared error (RMSE), mean absolute error (MAE) and r-squared (R2) were used as model evaluation metrics.

The main objectives of this study are therefore (1) to quantify the influence of groundwater extraction, precipitation, temperature and discharge on the prediction of groundwater levels in the aquifer, (2) predict groundwater levels under different climate and groundwater extraction conditions using a recurrent neural network and (3) compare results of each case scenario with the base case.

To quantify the relationship between each feature and the groundwater levels mutual information was performed. Mutual information results revealed that the variables selected indeed had an impact on the groundwater levels at the borehole sites selected for modelling with abstraction, precipitation and discharge carrying the highest weights. Using a recurrent neural network model specifically the Long Short-Term Memory (LSTM), different case scenarios were tested and each case scenario modelled was compared to the base case for analysis. Numerous studies have successfully predicted groundwater levels using recurrent neural networks; however, the way to take into account various case scenarios under specific weather conditions or using specific groundwater extraction cases is not well apprehended. Necessities are more into simply predicting groundwater levels. We therefore 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.

Table I Description of each case scenario

Scenario ID	Defintion	Description
1	Decrease rainfall peaks	If monthly rainfall peak > 100mm 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 datapoints
3c	Increase rainfall peaks	Increase rainfall peaks: If monthly rainfall peak > 100mm 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

This document is structured as follows: Section II discusses the contributions of machine learning in the domain hydro-geology; Section III will focus on data acquisition, data pre-processing, and model development; Section IV will outline and discuss the findings of the study and Section V will conclude this paper and also offer recommendations for future work.

II. LITERATURE REVIEW

The introduction of big data and machine learning in hydrology has brought about significant new advances in the sustainable management of groundwater resources around the world. By assessing data and using it to learn themselves, machine learning models can identify complex patterns present in the data for future analysis. Due to their ability to model complex nonlinear relationships, artificial neural networks (ANNs) have been the most widely used algorithms when studying hydrological systems particularly feed forward neural networks (FFNNs) and recurrent neural networks (RNNs) [6]. As previously highlighted machine learning models offer several advantages over the process based models. Recent studies suggest that data driven models can achieve performance comparable to or at times even more accurate than process based models. [7] 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; support vector machines, multi-player perceptron and, 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 (R^2) score of 0.51 while the multilayer perceptron, radial basis function network and support vector machine achieved R^2 scores of 0.71, 0.75 and 0.76 respectively, proving capabilities of machine learning techniques in groundwater level prediction.

[8] also did a comparative study between artificial neural networks 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 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.

Several other studies used FFNNs ([9], [10], [11], [12], [13]) and RNNs ([14], [15], [16]) to predict groundwater levels while other studies compared the performance of FFNNs and RNNs in predicting time dependant patters ([17] and [18]). In most studies FFNNs were able to accurately predict future values for a short period while RNNs were able to predict future values over an indefinite lead time. While FFNN 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 [19]. On the other hand, RNN have feedback connections that make them inherently dynamic in nature. In several studies mentioned earlier, RNNs were able to cope with the seasonality trend and time-varying behaviour of the semi-arid aquifer systems better than FFNNs. Hence, most studies concluded by recommending RNNs as useful tools for predicting hydrological systems.

These recent studies show how machine learning models are able to recognize patterns hidden in historical data and then apply those patterns to make future predictions without deep knowledge of the study area.

Due to our data being time-dependant we focused on the use of RNNs in this study particularly the LSTM model.

III. METHODOLOGY

In this paper we attempt focus on three main objectives. Firstly, we quantify the influence of groundwater extraction, precipitation, temperature and discharge on groundwater levels in the aquifer using mutual information, secondly, we attempt to predict groundwater levels for four boreholes under different climate and groundwater extraction conditions by training and testing an LSTM model and finally comparing results of each case scenario with the base case to give recommendations.

This section is structured as follows: Section III-A describes the case study area; Section III-B discusses data acquisition and pre-processing; Section III-D outlines the feature selection method used; and Section III-E will provide a brief descriptions of model setup and evaluation metrics employed in this study.

A. Study Area

The dolomite aquifers located in the northern part of South Africa are one of the most important aquifers in the country [3]. The Grootfontein dolomite aquifer specifically is a highly productive aquifer that contains good quality water used for industrial, agricultural and domestic water needs of people in the region [20]. Several towns and settlements such as Mahikeng, Lichtenburg and Itsoseng largely rely on groundwater as their main source of water. Average rainfall in the region is between 300 to 700mm annually while average temperatures in the region range between 2 to 16 degrees in winter (May to July) and 22 to 34 degrees in summer (August to March) [21]. Further details of the Grootfontein aquifer can be found in [3]. Figure 1 shows the map of the area under study.

B. Data Collection and Pre-processing

This study made use of data obtained from the Department of Water and Sanitation (DWS), Water Authorization and Registration Management System (WARMS) and South African Weather Service (SAWS). Monthly groundwater levels and discharge rates where obtained from DWS while monthly rainfall and temperature datasets was obtained from SAWS and abstraction dataset was obtained from WARMS. Four bore-holes were selected of those obtained from DWS. Selection criterion of these four was based on boreholes that had the longest continuous time period and least missing data points. The selected boreholes therefore ran for over 30 years and had less that 15% missing data points.

For climate data 3 weather stations were present within close proximity of the compartment as evidenced in Figure 1. One was within Marico (North of the compartment), one within the Mafikeng region (North West of the compartment) and another in Litchenburg (South of the compartment). A time series of monthly temperature and precipitation was compiled from the three stations. The three stations were selected due to their positions relative to the compartment boundaries providing an indication of the temperature and rainfall variability across the compartment. The stations were also selected due to the availability of longer historical data.

Discharge data used was from station D4H014 (location shown in Figure 1). The spring discharge at Molopo is considered hydraulically connected to groundwater across the Molopo/Grootfontein, and is therefore used as a representation of groundwater discharge from the compartment.

It should be noted that the groundwater abstraction dataset obtained for this study was licensed water use and not actual water use. Therefore, the registrations obtained may be over-estimates in an attempt to secure a supply or under-estimates for the purposes of securing a license. Furthermore, several individuals may not be registered (unlicensed abstraction). However, this is the dataset that was readily available. Therefore, total groundwater abstraction was calculated using the registrations obtained.

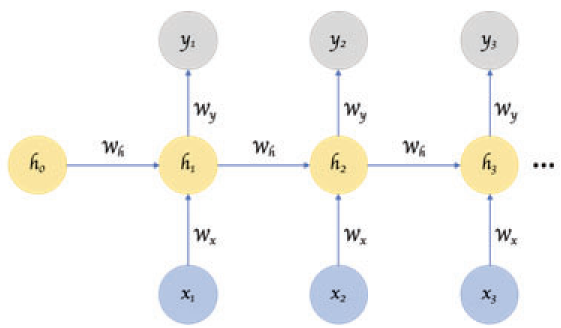

Fig. 2. A Recurrent Neural Network, with a hidden state that is meant to carry pertinent information from one input item in the series to others

C. Scenario tests

The LSTM model was run under different cases to explore the future trends of groundwater behaviour under different scenarios. The scenarios in table I were created to mimic possible changes in groundwater abstraction and climate conditions. Amplifying and decreasing rainfall peaks are case scenarios seen in climate variability and doubling abstraction amounts were also tested to verify whether the the neural network can replicate the impact on groundwater levels which is essential for groundwater planning. 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. Table I defines each scenario case tested.

D. Feature Analysis

Generally, input parameters are not equally informative in predicting the target variable. This is because some parameters may correlated, noisy or have an insignificant relationship with the target variable.

To quantify the influence of each feature (groundwater abstraction, temperature, precipitation and discharge) against the target variable (groundwater levels for the four boreholes) feature analysis was performed. Using mutual information (MI), we were able to quantify the relationship between each feature and each target variable. MI is a statistical method that is used to measure the degree of relatedness between variables [22]. MI is greater than zero when X and Y exhibit mutual dependence regardless of how nonlinear that dependence is because MI is able to model both linear and nonlinear rela-tionships. Therefore, the stronger the relatedness the higher the MI value.

In the next section we will discuss the machine learning model employed in this paper.

E. Regression and Evaluation

RNNs take two sources of input. To determine how it will respond to new data the 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 [23]. Therefore an event downstream in time depends upon, and is a function of one or more events that came before. Figure 2 shows an example of a RNN where Wx represents the weight vector for the hidden layer, W_y represents the weight vector for the output layer, Wh 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.

Table II Mutual Information results summary

BH Site	Temperature	Precipitation	Discharge	Abstraction
D4N0037	0.01	0.36	0.45	0.61
D4N0127	0.01	0.32	0.35	0.69
D4N0110	0.05	0.39	0.43	0.58
D4N0142	0.03	0.35	0.32	0.62
Average	0.03	0.36	0.39	0.63

F. Evaluation Metrics

The LSTM model was evaluated using three metrics, namely: RMSE, MAE and R2. RMSE is used as an indicator of how much much error the predicted results contain by comparing actual vs predicted results [24]. MAE calculates the averages of the absolute differences between prediction and actual observation where all individual differences have equal weight [24]. MAE is a good measure of error that can serve as a loss function to minimize. R2 is used to measure how close the data are to the fitted regression line. It explains how much variability of one factor can be caused by its relationship to another factor (co-linearity between the observed and predicted data) [25].

G. Ethics Clearance

No ethics considerations were required for this study. All data used in this study was obtained freely.

In the next section we will discuss the results obtained from the study.

IV. RESULTS AND DISCUSSION

This section presents the results obtained for each case scenario and the MI results. In section IV-A we discuss the results obtained from feature analysis and section IV-B presents the regression results.

A. Feature ranking

There were 4 features used in this paper. Using MI we deduced the contribution of each feature to classify the features as a value from the target variables. Table II shows the results of applying MI to the feature set for each target variable.

Table II illustrates the scores of the contribution of each feature to each target variable. The first column indicates the borehole site ID (BH site), the remaining 4 columns indicate feature names and the corresponding scores to each target variable.

Table II shows that on average regional abstraction correlates with water levels the strongest in each borehole with an average entropy of 0.63. Correlation does not imply causality, but these results essentially show that the method detects a correlation between the curve of increasing groundwater use over time with the groundwater levels at each borehole. It should be noted that the results obtained are quite surprising given that the groundwater level dataset is significantly noisier than that of groundwater abstraction which shows an increasing curve over time. However, the results also support the assertion that the compartment is highly impacted by regional abstraction as recorded in previous studies.

Table III Performance of LSTM model on the test data set for 4 boreholes in the Grootfontein aquifer.

BH Site	R2	RMSE	MAE
D4N0037	0.87	0.09	0.06
D4N0127	0.81	0.57	0.36
D4N0110	0.77	0.13	0.10
D4N0142	0.85	0.18	0.13

The second most correlating feature with groundwater levels is discharge with an entropy average of 0.39. A strong relationship is expected, given that the hydraulic gradient between the surrounding aquifer and the spring is what drives discharge hence changes in groundwater level will directly translate to changes in discharge rate.

The rainfall dataset is the feature with 3rd strongest correlation with average entropy value of 0.36. A high correlation was expected as groundwater levels respond to rainfall events. Temperature has the lowest correlation (and therefore least influence in predicting water levels) with an average entropy average of 0.03.

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, but it does not tell us what drives the groundwater levels. However, MI gain provides a useful framework for feature selection by indicating the contribution of each feature relative to all the other features.

B. Scenario testing results

Table III shows the results obtained for the base case by training and testing the LSTM model with actual input values. As can be seen by the results obtained the LSTM model performed quite well in the prediction of all four monitoring stations using precipitation discharge and groundwater abstraction as model inputs. On this basis the predictive capabilities of the LSTM model was considered adequate for the main objectives of this study.

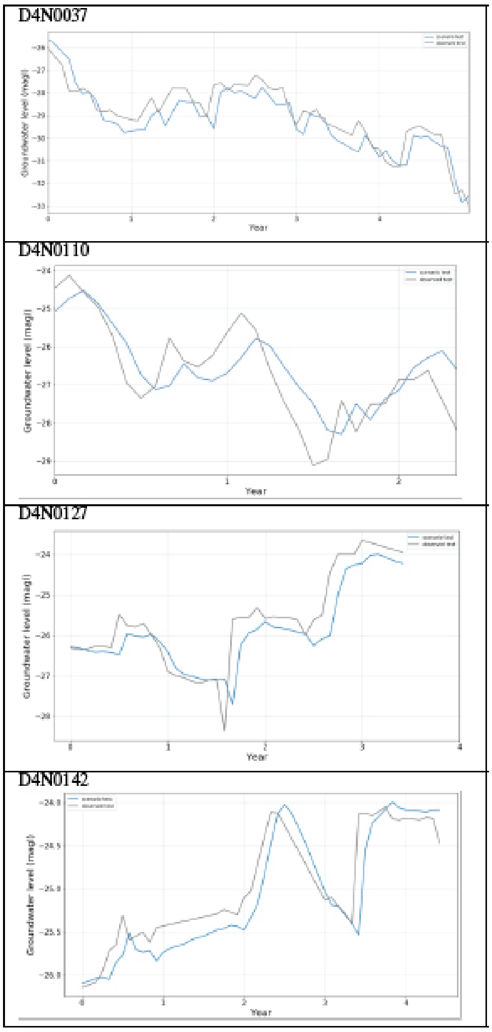

Table IV. Model prediction for scenario 1 - decrease rainfall peaks

In the next section we discuss results obtained through each scenario test.

1) Scenario 1: Subtracting 100mm from rainfall events that had peaks higher than 100mm resulted in a drop of groundwater levels at the specific month were precipitation was reduced. It is important to note that precipitation peaks above the 100mm threshold were not abundant in the testing phase but instead only a few high rainfall events occurred during 2000 – 2004 (duration of test data for most boreholes) that passed the threshold. For this reason, the drop in groundwater levels is only seen at the particular months were recharge was reduced. For borehole D4N0142, the decrease in water levels between year 0 to year 2 is clearly 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 the base case the model had 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 the base case. For borehole D4N0110 a decline in water levels (0.2 to roughly 0.4m) can be observed from year 0.

2) Scenario 2: Doubling the regional abstraction rates in the compartment resulted in a the highest decrease of groundwater levels for all boreholes (ranging between 0.5 to 1.7m) in the compartment when compared to the base case. Although the doubled abstraction rates seemed to have a higher effect on boreholes D4N0110 and D4N0142, the magnitude for boreholes D4N0127 and D4N0037 was much smaller.

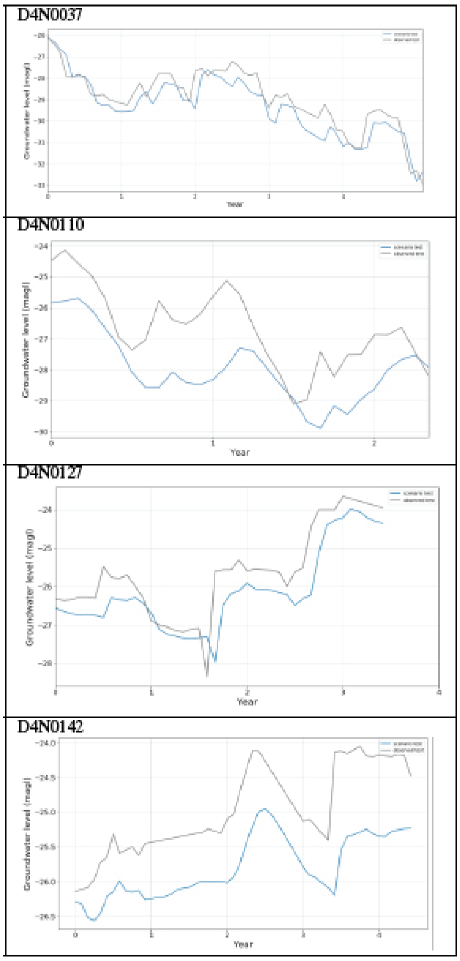

Table V. Model prediction for scenario 2 - double abstraction

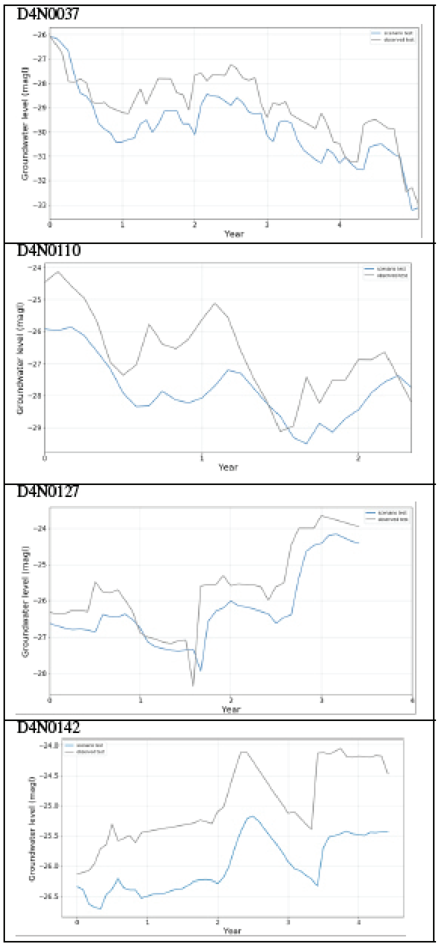

Table VI. Model prediction for scenario 3a - double abstraction + decrease rainfall peaks

3) Scenario 3a: The highest negative change was observed from the third scenario were regional abstraction was doubled and recharge was reduced by 100mm for high precipitation events. Groundwater levels for borehole D4N0142 dropped by almost 1m. For boreholes D4N0037, D4N0127 and D4N0110 groundwater levels dropped by at least 0.5m. All selected boreholes showed a decline in groundwater levels. This indicates that the models are able to generate a change in groundwater level that would be expected if abstraction rates increase. This could be useful for assessing the impact of future changes in groundwater abstraction.

4) Scenario 3b: By halving all precipitation values in the test set, groundwater levels declined considerably when compared to the base case. The decline in groundwater levels is consistent throughout the years since all precipitation values were halved unlike in scenario 1a where specific events decreased (providing it was above 100m threshold). 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.

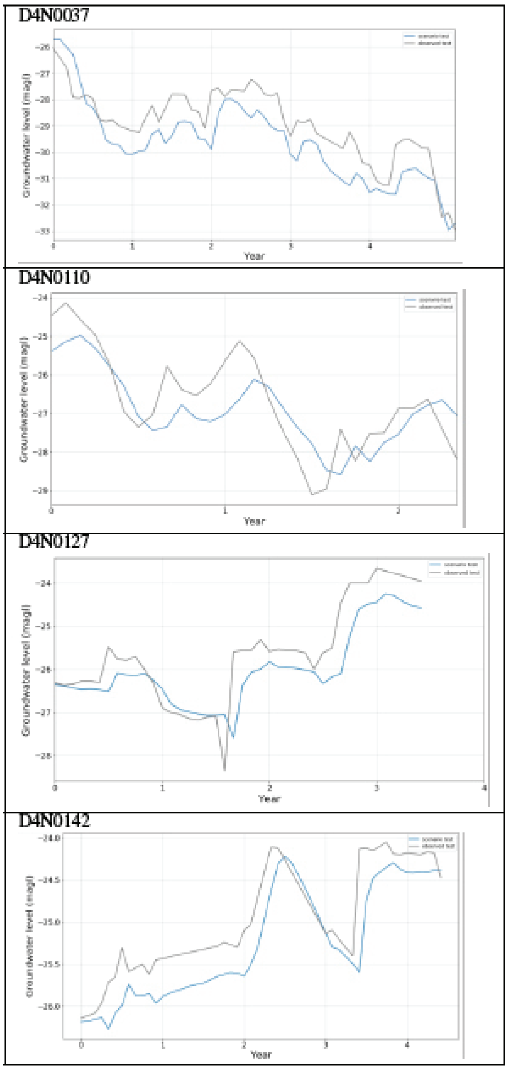

Table VII. Model prediction for scenario 3b - half all rainfall values

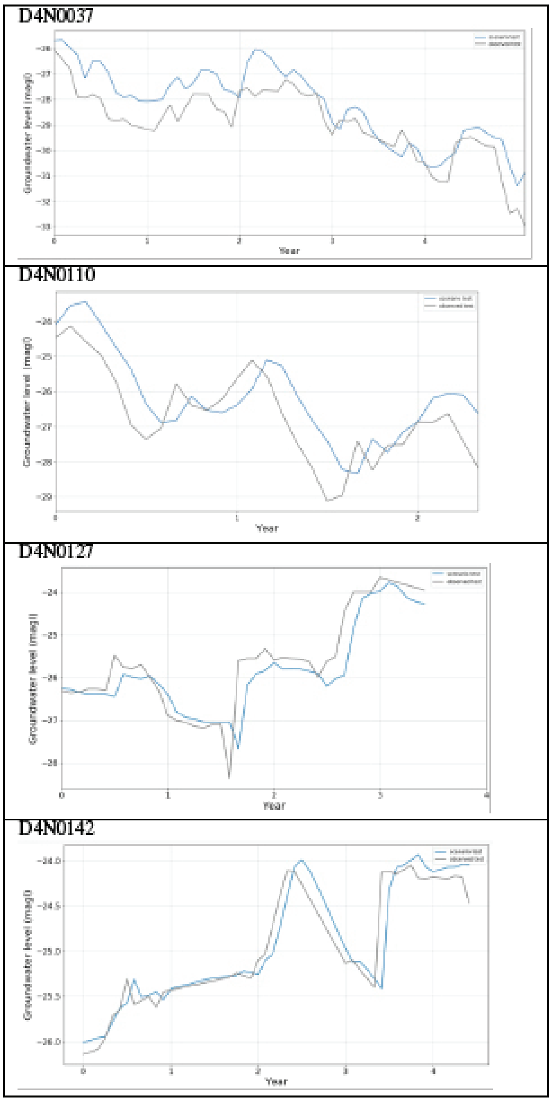

Table VIII. Model prediction for scenario 3c - increase rainfall peaks

5) Scenario 3c: Similar to scenario 3a 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 increased groundwater levels for boreholes D4N0037 and D4N0110. In each of these boreholes an increase in recharge increased groundwater levels a lot higher than compared to borehole D4N0142. For borehole D4N0127 an increase in water levels was seen 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.

6) Scenario 4: For the final scenario, precipitation, abstraction and discharge data were formulated by averaging the 30 years of data for each variable and predicting 15 years into the future for borehole D4N0037. The graph showed a reasonable prediction of groundwater levels by simply feeding the model averaged values of precipitation, 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 15 years given sufficient training data.

V. CONCLUSION

The purpose of this study was to simulate the effects of adjusted groundwater abstraction and precipitation on ground-water levels in the Grootfontein aquifer. The aim was to test the applicability of recurrent neural network models in the estimation of groundwater levels under adjusted parameters. 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 water service authorities including those in the North West. Groundwater is an essential resource in the North West providing water to a significant part of the population. 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.

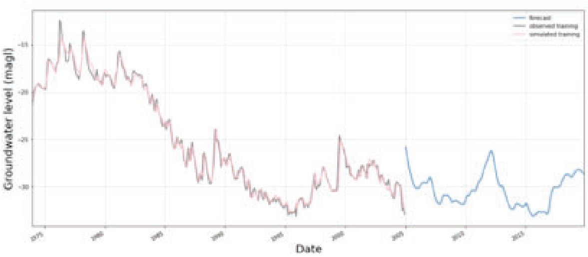

Fig. 3. Model prediction for scenario 4 - long term prediction for borehole D4N0037

Although the RNN model was unable to predict the sporadic peaks and troughs which could be due to lurking variables such as daily pumping records from high-capacity irrigation wells or measurement errors, the model still performed quite well. The LSTM model was first tested to predict groundwater levels using actual input values obtained (considered as the base case). The model performed quite well with the modelled water levels matching the observed reasonably well. On this basis the predictive capabilities of the LSTM model was considered adequate for the main objectives of this study.

Scenario testing results 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 the groundwater levels in the compartment decreased by almost 2 m. These results indicated the importance the chosen variables had on groundwater levels in the region further strengthening our entropy results. The framework used in this study 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.

This study only considered a handful of wells at a regional scale thereby potentially restricting large scale applicability of the model and results. Further investigation is therefore needed at a larger-scale. This study only tested a single model, further investigation could test applicability of different models in the prediction of groundwater levels. Lastly, testing with actual groundwater abstraction could further strengthen results.

ACKNOWLEDGEMENTS

The authors would like to thank the Water Research Commission in Pretoria, and the wider consortium, for supporting and encouraging this work, and providing a forum for useful discussions. This work forms part of a 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. This work is also based on the research supported in part by the National Research foundation of South Africa (Grant number: 121835). A special thanks to Kirsty Gibson for creating the map used in the study.

REFERENCES

[1] R. C. O. Modisha, “Investigation of the Ramotswa Transboundary Aquifer Area, groundwater flow and pollution,” Ph.D. dissertation, Witswatersrand, 2017.

[2] K. G. Villholth, C. Tøttrup, M. Stendel, and A. Maherry, “Integrated mapping of groundwater drought risk in the Southern African Development Community (SADC) region,” *Hydrogeology Journal,* 2013.

[3] J. E. Cobbing, “An updated water balance for the Grootfontein aquifer near Mahikeng,” *Water SA,* vol. 44, no. 1, pp. 54–64, jan 2018.

[4] X. Chen, Y. Zhang, Y. Zhou, and Z. Zhang, “Analysis of hydrogeological parameters and numerical modeling groundwater in a karst watershed, southwest China,” *Carbonates and Evaporites,* 2013.

[5] M. Steyn and W. H. Brink, “Short-term stream flow forecasting and downstream gap infilling using machine learning techniques,” Ph.D. dissertation, Stellenbosch University, 2018.

[6] P. Coulibaly and C. Baldwin, “Dynamic neural networks for non-stationary hydrological time series modeling,” *Journal of Hydrology,* vol. 67, pp. 71–75, 2008.

[7] C. Chen, W. He, H. Zhou, Y. Xue, and M. Zhu, “A comparative study among machine learning and numerical models for simulating groundwater dynamics in the Heihe River Basin, northwestern China,” *Scientific Reports,* vol. 10, no. 1, pp. 1–13, dec 2020.

[8] M. Chitsazan, G. Rahmani, and A. Neyamadpour, “Forecasting ground-water level by artificial neural networks as an alternative approach to groundwater modeling,” *Journal of the Geological Society of India,* 2015.

[9] H. Herawati, Suripin, and Suharyanto, “River flow modeling using artificial neural networks in Kapuas river, West Kalimantan, Indonesia,” in *AIP Conference Proceedings,* vol. 1903. American Institute of Physics Inc., nov 2017.

[10] Y. Kanyama, R. Ajoodha, H. Seyler, and H. 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,* Melborne, Australia, 2020.

[11] P. D. Sreekanth, N. Geethanjali, P. D. Sreedevi, S. Ahmed, N. Ravi Kumar, and P. D. Kamala Jayanthi, “Forecasting groundwater level using artificial neural networks,” *Current Science,* vol. 96, no. 7, pp. 933–939, 2009.

[12] P. C. Nayak, Y. R. Satyaji Rao, and K. P. Sudheer, “Groundwater level forecasting in a shallow aquifer using artificial neural network approach,” *Water Resources Management,* vol. 20, pp. 77–90, 2006.

[13] C. M. Zealand, D. H. Burn, and S. P. Simonovic, “Short term streamflow forecasting using artificial neural networks,” *Journal of Hydrology,* vol. 214, no. 1-4, pp. 32–48, 1999.

[14] B. Maheshwara Babu, G. Srinivasa Reddy, U. Satishkumar, and P. Kulkarni, “Simulation of Groundwater Level Using Recurrent Neural Network (RNN) in Raichur District, Karnataka, India,” *Int.J.Curr.Microbiol.App.Sci,* vol. 7, no. 12, pp. 3358–3367, 2018. [Online]. Available: https://doi.org/10.20546/ijcmas.2018.712.386

[15] X. H. Le, H. V. Ho, G. Lee, and S. Jung, “Application of Long Short-Term Memory (LSTM) neural network for flood forecasting,” *Water (Switzerland),* vol. 11, no. 7, 2019.

[16] P. Coulibaly, F. Anctil, R. Aravena, and B. Bobe´e, “Artificial neural network modeling of water table depth fluctuations,” *Water Resources Research,* vol. 37, no. 4, pp. 885–896, apr 2001. [Online]. Available: http://doi.wiley.com/10.1029/2000WR900368

[17] B. S. Supreetha, N. Shenoy, and P. Nayak, “Lion Algorithm-Optimized Long Short-Term Memory Network for Groundwater Level Forecasting in Udupi District, India,” *Applied Computational Intelligence and Soft Computing,* 2020.

[18] D. Nagesh Kumar, K. Srinivasa Raju, and T. Sathish, “River flow forecasting using recurrent neural networks,” *Water Resources Management,* vol. 18, pp. 143–161, 2004.

[19] D. Brezak, Bacek Tomislav, Majetic Dubravko, Kasac Josip, and B. Novakovic, “A Comparison of Feed-forward and Recurrent Neural Networks in Time Series Forecasting,” Tech. Rep., 2011.

[20] WRC, “Understanding groundwater: A stakeholders guide to the North West Dolomite Aquifer,” Tech. Rep., 2003.

[21] SA-Venues, “North West Province Climate and Weather, South Africa,” p. 1, 2019. [Online]. Available: https://www.sa-venues.com/weather/ northwest.php

[22] A. Sudharsan, “Why, How and When to apply Feature Selection,” jan 2018. [Online]. Available: https://towardsdatascience. com/why-how-and-when-to-apply-feature-selection-e9c69adfabf2

[23] 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,* vol. 60, no. 7, pp. 1242–1265, 2015.

[24] A. P. Engelbrecht, Wiley: *Computational Intelligence: An Introduction, 2nd Edition - Andries P. Engelbrecht.* Wiley, 2007. [Online]. Available: http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0470035617.html

[25] P. Krause, D. P. Boyle, and F. Base, “Comparison of different efficiency criteria for hydrological model assessment,” *Advances in Geosciences,* vol. 5, pp. 89–97, 2005.