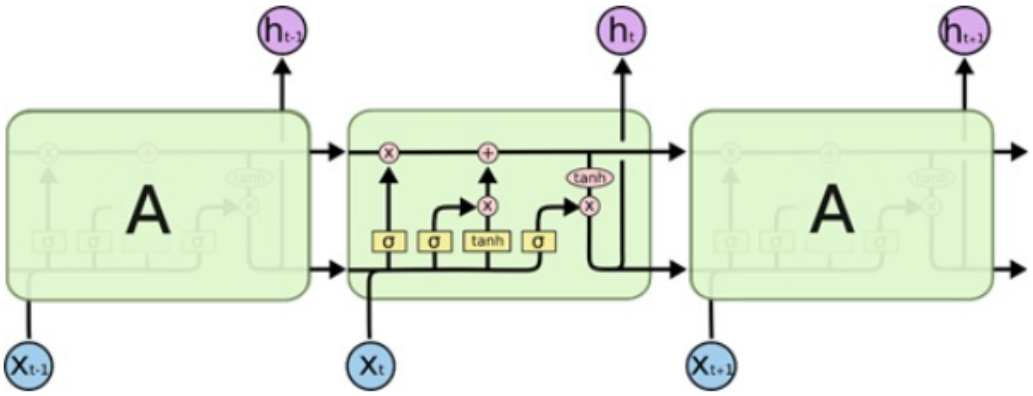
CHALMERS
UNIVERSITY OF TECHNOLOGY



Evaluating deep learning models for estimating groundwater level perturbation sources in areas with non-natural groundwater level regimes

Master's thesis in Infrastructure and Environmental Engineering

Shreyas Raja Ramesh

Department of Architecture & Civil Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2021
www.chalmers.se

Master's thesis 2021

Evaluating deep learning models for estimating groundwater level perturbation sources in areas with non-natural groundwater level regimes

Shreyas Raja Ramesh

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Architecture and Civil Engineering
Division of Geology and Geotechnics
Chalmers University of Technology
Gothenburg, Sweden 2021

Evaluating deep learning models for estimating groundwater level perturbation sources in areas with non-natural groundwater level regimes.
Shreyas Raja Ramesh

Supervisors: Ezra Haaf, Division of Geology and Geotechnics
Jonas Sundell, Trafikverket

Examiner: Lars Rosen, professor in the Division of Geology and Geotechnics

Master's Thesis 2021
Department of Architecture and Civil Engineering
Division of Geology and Geotechnics
Engineering geology
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: A block of LSTM used for groundwater time series modelling

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2021

Evaluating deep learning models for estimating groundwater level perturbation sources in areas with non-natural groundwater level regimes.
Shreyas Raja Ramesh
Department of Architecture and Civil Engineering
Chalmers University of Technology

Abstract

The study uses deep learning models such as artificial neural networks (ANN) to tackle applications within the scope of hydrogeology. Time series modelling is one of many such applications used in analysing historical groundwater heads over an area of interest. A perturbed source, i.e. the Haga service tunnel of the West Link commuter rail project is the case study used to examine the model. Reduced groundwater heads are observed from monitoring wells around the perturbed source when plotted against time around the service tunnel. This reveals a certain impact on the groundwater caused by underground constructions. In this thesis, a branch under ANN called RNN (Recurrent Neural Networks) using LSTM (Long-Short Term Memory) network is investigated for the case study and the model is completely data-driven and statistically evaluated. The standard evaluation error metrics used are Mean squared error, Mean absolute error, Coefficient of determination and Pearson correlation coefficient. The predictions made on the observed GWL are checked with the error metrics and the results evaluated the impacts caused by the perturbed source for different cases. The groundwater heads are forecasted for different time periods based on best case identified from sequence to sequence (seq2seq) predictions. From the forecast, the groundwater heads are evaluated with unseen data using the Pearson correlation coefficient. However, shorter forecasting periods showed promising accuracy compared to longer forecasting accuracy indicating the groundwater movements are complex in real world influenced by natural parameters such as precipitation rates, relative humidity, evapo-transpiration.

Keywords: AI, ANN, LSTM neural network, Python, deep learning , groundwater, groundwater level time series.

Acknowledgements

Throughout my thesis work, I received great support and knowledge from my supervisors. In particular, I would like to thank my supervisor Dr.Ezra Haaf (Post Doc) at Chalmers University of technology for giving me an interesting project topic and being the backbone of this project. A special thanks to Dr.Jonas Sundell (Specialist hydrogeologi) at Trafikverket for providing knowledge on the study area and insights to the problems encountered during my thesis. I would like to thank my examiner Prof. Lars Rosén for being a part of the thesis assessment and providing me fruitful feedback. I would also like to thank Eric McGiveny for giving insights on handling the data in this project. A final thanks to my family and friends who supported me during this pandemic and making it possible to complete my journey.

Shreyas, Raja Ramesh, Gothenburg, June 2021

Contents

List of Figures **ix**

List of Tables **xi**

1 Introduction **1**
1.1 Purpose of the study . . . 2
1.2 Research questions . . . 2

2 Background **3**
2.1 Literature Review . . . 3
2.2 Theory . . . 6
2.2.1 Groundwater flow and impacts . . . 6
2.3 Deep learning Neural Networks . . . 7
2.3.1 Recurrent Neural Networks(RNN) . . . 8
2.3.1.1 LSTM units . . . 8
2.3.1.2 Hyperparameters . . . 10
2.3.1.3 Training the LSTM neural network . . . 10
2.4 WestLink - Commuter Rail Project . . . 11
2.4.1 Linné service tunnel . . . 12
2.4.1.1 Geology . . . 13
2.4.1.2 Hydrogeology . . . 14
2.4.1.3 Groundwater lowering . . . 16

3 Method **17**
3.1 Working Strategy . . . 17
3.2 Conceptual model . . . 19
3.3 Data Collection & Preparation . . . 20
3.3.1 Data preprocessing . . . 21
3.3.1.1 Data Resampling . . . 21
3.3.1.2 Data Scaling . . . 21
3.4 Groundwater Modelling using LSTM . . . 22
3.4.1 Training and testing the data . . . 22
3.4.2 Neural Network Architecture & Computational Aspects . . . 23

4 Results & Discussions **25**
4.1 Response time estimation using precipitation . . . 26
4.2 Groundwater predictions . . . 27

4.2.1 Sequential modelling using meteorological parameters 27
4.2.2 Sequential modelling with Meteorological, Infiltration & Leakage data 29
4.2.3 Sequential modelling with GWL, Meteorological, Infiltration & Leakage data 32
4.2.4 Error Metrics on groundwater predictions for different scenarios 34
4.2.5 Limitations of sequence modelling 35
4.3 Groundwater Forecasting 36
4.3.1 Sequential groundwater forecasting for different training periods 36
4.3.1.1 Forecasting HK4004U groundwater levels 37
4.3.1.2 Forecasting HK4256U groundwater levels 38
4.3.2 Forecasting correlations 40

5 Conclusion **43**

Bibliography **45**

A Appendix 1 **I**

List of Figures

2.1 Neural Network structure (Lelli 2021) 7
2.2 RNN architecture (Daliakopoulos, Coulibaly, and Tsanis 2005) 8
2.3 LSTM block with input, output and forget gates at different. + and X sign indicates pointwise addition & multipication(Sagheer and Kotb 2019) 9
2.4 Top View of the WestLink tunnel(Ulf. Sundkvist and Thomas. Wallroth 2016) 12
2.5 Stratification of soil layers from Olive-Annedal from west to east of the service tunnel region(Ulf. Sundkvist and Thomas. Wallroth 2016) 13
2.6 Weakness Zones around the service tunnel shown by red and blue dashed lines(Ulf. Sundkvist and Thomas. Wallroth 2016) 14
2.7 Surface water being divided based on topography around the West Link Commuter rail project(Ulf. Sundkvist and Thomas. Wallroth 2016) 15
2.8 Representation of Groundwater lowering in tunnels 16

3.1 Location of selected wells with respect to the Haga service tunnel . . 17
3.2 Flowchart depicting the working approach 18
3.3 GWL time series of HK4004U and HK4256U around the Haga service tunnel 19
3.4 Training and testing data splitting arrangement 23

4.1 Groundwater level time series of selected wells along with meteorological and hydrological time series 25
4.2 Correlation coefficients for different precipitation rolling sum periods for HK4004U and HK4256U 26
4.3 Correlations for 15 days rolling sum between HK4004U and precipitation 26
4.4 Correlations for 15 days rolling sum between HK4256U and precipitation 26
4.5 Predictions on HK4004U groundwater well with meteorological data . 28
4.6 Predictions on HK4256U groundwater well with meteorological data . 28
4.7 Impacting average gradients for HK4004U 29
4.8 Impacting average gradients for HK4256U 29
4.9 Correlations between observed & predicted GWL for HK4004U 29
4.10 Correlations between observed & predicted GWL for HK4256U 29

4.11 Predictions on HK4004U groundwater heads with meteorological, infiltration and leakage data 30
4.12 Predictions on HK4256U groundwater heads with meteorological, infiltration and leakage data 31
4.13 Correlation between observed & predicted GWL for HK4004U 31
4.14 Correlation between observed & predicted GWL for HK4256U 31
4.15 Impacting average gradients for HK4004U 32
4.16 Impacting average gradients for HK4256U 32
4.17 Predictions on HK4004U groundwater heads with historical GWL's, meteorological, infiltration and leakage data 32
4.18 Predictions on HK4256U groundwater heads with historical GWL's, meteorological, infiltration and leakage data 33
4.19 Correlation between observed & predicted GWL for HK4004U 33
4.20 Correlation between observed & predicted GWL for HK4256U 33
4.21 Impacting average gradients for HK4004U 34
4.22 Impacted average gradient for HK4256U 34
4.23 Splitting arrangement for forecasting different training periods of time series data used for modelling 36
4.24 Forecasted series of days for 12 months vs unseen GWL for HK4004U 37
4.25 Forecasted series of days for 18 months vs unseen GWL for HK4004U 37
4.26 Forecasted series of days for 27 months vs unseen GWL for HK4004U 38
4.27 Forecasted series of days for 12 months vs unseen GWL for HK4256U 39
4.28 Forecasted series of days for 18 months vs unseen GWL for HK4256U 39
4.29 Forecasted series of days for 27 months vs unseen GWL for HK4256U 40
4.30 Correlations of HK4004U for different training periods and forecasted days 41
4.31 Correlations of HK4256U for different training periods and forecasted days 41

List of Tables

2.1 List of Articles reviewing groundwater predictions using AI and Machine learning techniques with influencing input parameters. * (P) - precipitation, * (rH) - relative humidity, * (T) - temperature, *(ET) - evapotranspiration, *(GWL) - groundwater level 4

3.1 List of wells selected for modelling with parametric information . . . 19
3.2 Input data units of measurement 20
3.3 List of input data, range of input data and time steps used in the the LSTM model 22

4.1 Hyperparameters used for groundwater modelling 27
4.2 Error metric values for different features based on the predictions. (MAE) - Mean absolute error, (MSE) - Mean squared error, (RMSE) - Root mean squared error, (R^2 score) - Coefficient of determination . 35

1
Introduction

Groundwater is an essential source for drinking water, domestic usage and agricultural activities in many parts of the world (Wunsch, Liesch, and Broda 2020). Groundwater resources are affected by densely populated areas by the impacts from urbanisation and development. Subsequently large underground infrastructures such as tunnels, subways, basement car parks and deep foundations for high rise buildings are commissioned for this purpose keeping in mind the impacts on groundwater (De Caro, Crosta, and Previati 2020). The natural groundwater flow pattern is obstructed during the construction phase in underground infrastructures leading to increased risks of damages over time. For example, lower level flooding (in basements) and in-specific tunnel leakages leading to drawdowns of the groundwater table which in turn cause settlements. In this context, hydrogeological modelling helps to evaluate the impacts on groundwater levels in the vicinity caused by construction activities. However, conceptual and physical models are the main tools used for determining the hydrological parameters such as groundwater levels observed in the system. But in some cases, lack of data availability leads to poor results from evaluating physical models and requires calibration for improvement in performance (Daliakopoulos, Coulibaly, and Tsanis 2005). Supervised learning techniques have influenced hydrogeological model evaluation in the recent past adding to the conservative methods.

Deep data-driven models such as artificial neural networks (ANN) applied in hydrological time series have variety of applications from event based to real-time modelling of hydrological investigations. Real time modelling helps predicting and forecasting targets with relevant influencing factors from the dataset through neural networks computations. Studies based on ANN's in estimating groundwater level time series have shown promising response in predicting required targets (Rajaee, Ebrahimi, and Nourani 2019). Moreover, there is lack of studies that show the applications of ANN's in urban settings to tackle hydrological time-series problems caused by underground infrastructural activities. Fortunately, literature surveys relevant to predicting groundwater level time series in rural and urban areas motivated in selecting an appropriate ANN to examine in this thesis project. A brief summary of the literature survey is presented in section 2.1 and listed in table 2.1. However, it should be noted that there are very few or almost no studies reviewing impacts on groundwater from underground constructions using ANN's inturn making this project research oriented.

1.1 Purpose of the study

The study is based on assessing the change in groundwater heads using ANN's models to analyse the impacts caused by the construction of the Haga service tunnel of the West Link project handled by Trafikverket. The main aim is to predict the groundwater level time series using meteorological, infiltration and inleakage data and evaluate it against the observed groundwater level to analyse impacts from the perturbed source. A detailed comparison between parameters selected for modelling is presented to evaluate the credibility of the method adopted. This aims to investigate the source of groundwater impact and uncertainties in modelling using ANN's which open gates for spatial analysis and development of strategies in the future.

1.2 Research questions

1. To what degree are ANN's used for groundwater time series predictions supported by existing research studies?
2. What input data can be selected and prepared for groundwater time series modelling using the selected neural network for the case study ?
3. What is the influence on the selected input features used for groundwater predictions?
4. What is the best suited forecasting range from sequence to sequence modelling the groundwater heads with true future data?

2
Background

2.1 Literature Review

Deep learning, Artificial Intelligence and Machine Learning methods have seen growth in the field of hydrogeology to tackle quantitative groundwater problems. Hydrological time series are investigated using these methods with regards to the study area . A few relevant studies on AI used in applications of GWL time series data driven models are listed in table 2.1 and discussed further in this section.
The problem statement is motivated with studies using neural networks, deep learning & machine learning in groundwater forecasting using time series and other relevant information. Rajaee et al 2019 (Rajaee, Ebrahimi, and Nourani 2019), surveyed studies concerning AI and machine learning applications used in GWL time series prediction. The review paper shows 67 peer reviewed articles modelling GWL time series using AI and machine learning techniques. These data-driven approaches model GWL time series with uncertainties and showcase the highest number of studies were conducted using ANN's. The uncertainties in modeling were time step selection, input data consideration, data splitting arrangement, data set size, type of aquifer in the area, and finally, the software program used (Rajaee, Ebrahimi, and Nourani 2019). Discussions on the correct use of the uncertainty parameters are reviewed to develop a consistent model in this article (Rajaee, Ebrahimi, and Nourani 2019). Also, the paper suggests that it is impossible to use one AI method to predict groundwater heads and relevant comparison is required in supervised learning models. However, hybrid models such as wavelet-AI, SOM-AI, GP-AI models perform better than regular models such as ARIMA (Auto regression integrated moving averages), SVM (support vector machines) and MLR(Multiple linear regression).
Recent studies using different AI models with relevant input data and time steps for GWL time-series predictions are depicted in table 2.1. The studies show groundwater predictions for different aquifer locations and scenarios using deep learning models, AI, and other machine learning models. However, they are oriented with predictions of GWL's in rural and urban environments concerning natural factors influencing GWL's.

Title of Article	Journal Published	Used AI models	Input Variables	Time	Range of input data
Groundwater Level Forecasting using ANN (Wunsch, Liesch, and Broda 2020)	Hydrology and Earth Systems	LSTM, CNN & NARX	(P), (rH),(GWL) & (T)	Weekly	2012-2016
Reconstruction of missing GWL data using LSTM deep neural network (Vu et al. 2021)	Journal of Hydrology	LSTM, SVR, FNN	(P), (rH), (ET) & (GWL)	Weekly	1970-2020
Developing a LSTM model for predicting water table depth in agricultural areas (Zhang et al. 2018)	Journal of Hydrology	LSTM	Water Diversion, Evap. Vol. , (T), (GWL)	Monthly	2000-2014
Ground Water Level forecasting using ANN (Daliakopoulos, Coulibaly, and Tsanis 2005)	Journal of Hydrology	ANN(FNN, RNN, LM, BR)	(P) & (T)	Monthly	1981-2000
A review of AI methods in groundwater modeling (Rajaee, Ebrahimi, and Nourani 2019)	Journal of Hydrology	ANN, ANFIS, GP & SVM	-	-	-
A novel deep learning algorithm for GWL prediction based on spatiotemporal deep neural network (Chen et al. 2020)	-	LSTM, FNN & SVR	(P), (ET) & Surface Water	Monthly	-
A wavelet NN conjunction model for GWL forecasting	Journal of Hydrology	ANN, ARIMA & (WA-ANN)	(P) total, (T) avg & (GWL)	Monthly	2002-2009
Comparison study of AI methods for short term GWL prediction in the northeast Garchsaran unconfined aquifer(Khedri, Kalantari, and Vadiati 2020)	-	ANN,FL,ANFIS, GMDH & LSSVM	(GWL), (P),(T) & (ET)	Monthly	-
Groundwater Modelling with machine learning techniques (Kenda et al. 2018)	-	LR, Decision treea, Random Forest, Gradient Boosting	(T),(P), Sun duration, cloud cover and snow blanket	Daily	2010

Table 2.1: List of Articles reviewing groundwater predictions using AI and Machine learning techniques with influencing input parameters. * (P) - precipitation, * (rH) - relative humidity, * (T) - temperature, *(ET) - evapotranspiration, *(GWL) - groundwater level

GWL time series using deep learning models, particularly RNN's are studied in detail based on their ability to learn non-linear data over long periods (Daliakopoulos, Coulibaly, and Tsanis 2005). The study is examined on different aquifers in Messara valley, Greece to predict missing GWL's and is successfully implemented. A comparative study between three ANN models such as NARX, LSTM, and CNN is implemented to forecast the performance of GWL's. The study was conducted on the largest groundwater resource for drinking water in the Upper Rhine Graben in central Europe (Wunsch, Liesch, and Broda 2020). It is seen that NARX performs well for short and mid-term weekly data while LSTM's and CNN's struggle. However, their performance is said to be improved for longer training periods such as 10 years or more. In agricultural areas, the GWL's are predicted using LSTM's to analyze the water table depth primarily to develop irrigation strategies. In China, a study was evaluated on the Heihe river basin groundwater conditions using deep learning models such as LSTM, FNN, and SVR (Chen et al. 2020). The goal was to develop water management strategies caused by the exploitation of groundwater resources by domestic, industrial and agricultural use. Furthermore, a spatiotemporal model is developed for groundwater predictions by Chen (Chen et al. 2020) and is not examined due to its heavy complexity in this thesis project. Similarly, groundwater condition on the Garschan aquifer in Iran is predicted for short periods of 1, 2, and 3 months respectively using ANN's and other machine learning methods (Khedri, Kalantari, and Vadiati 2020). The main goal of this study was to examine the decline in GWL caused by increasing water demands exploited by deep drilling of wells.

It should be noted that all the studies involving time series using ANN and machine learning methods are evaluated using statistical error metrics such as Mean absolute error(MAE), Coefficient of determination (R^2), Mean squared error (MSE), and Root mean squared error (RMSE). Furthermore, the importance of these error metrics is to analyse the model performance accuracy on the predictions with the original dataframe. Overall, machine learning and deep learning models are successfully used in applications of hydrological studies to estimate GWL's using influencing factors. However, most studies showed comparing different methods is better to reach a conclusion using black-box models involving uncertainties (Wunsch, Liesch, and Broda 2020)(Rajaee, Ebrahimi, and Nourani 2019)(Chen et al. 2020). Based on literature studies, LSTM neural networks have the applicability to be used in hydrological time series modelling and hence, it is adopted for the case study in this project.

2.2 Theory

2.2.1 Groundwater flow and impacts

Groundwater flows primarily following the law of physics and thermodynamics (C.W.Fetter 2014). Groundwater patterns are influenced by surrounding ecosystems and climate change with respect to time (Haaf 2020). Identifying the natural parameters impacting the groundwater recharge is challenging considering historical geological formations in the area. Dealing with complex variability in groundwater movements through process based, numerical models and conceptual models have proved to be efficient for successful water management over the past years (Haaf 2020). However, this is a continuous problem to handle for water managers which varies with the local hydrogeological conditions.

Furthermore, urbanisation is another reason for groundwater lowering caused by on-ground and underground infrastructures (Sundell et al. 2019). Groundwater lowering is caused by impermeable constructions at surface levels, for eg roads, pavements, tunnels, deep foundations etc. Subsequently, stormwater drainage's complement these surface level constructions by diverting surface runoff into the ground. Additionally, increased groundwater recharge occurs through leakages from drinking water systems. Deep foundations and underground constructions, for e.g tunnels are complex projects with regards to tackling groundwater impacts. The factors influencing the impacts of groundwater in underground constructions include permeability, hydraulic properties, groundwater flow direction, waterproofing ability, nature of aquifer and other relevant properties (De Caro, Crosta, and Previati 2020).

Another approach to examine impacts on groundwater is GWL time series analysis, i.e. data driven. This helps visualizing the groundwater head in wells with respect to time. The groundwater level in different aquifers is achieved visually using observed GWL data relative to time by plotting the relationship in 2D. Also, the variations in heads of the groundwater is influenced by stresses on aquifers caused by rainfall, evapotranspiration, pumping, infiltration rates and surface water levels (Bakker and Schaars 2019). Time series analysis helps solving problems related to groundwater heads around desired locations. For example, the drawdown conditions in a well due to pumping, recovery of heads in wells after droughts, level of heads after rainfall and the impacts of climate change (Bakker and Schaars 2019).

2.3 Deep learning Neural Networks

Artificial Neural networks(ANN's) are complex systems resembling the functions of how neurons work in human brains for various perceptions observed in a biological context (Lelli 2021). Similarly, data is perceived as complex neurons for problem specific data-driven modelling. The neural network consists of layers, named as input layers, hidden layers and output layers, see fig 2.1. Input data fed to the neural network is captured by the input neurons and the outputs are fed to the neurons in the hidden layers defined in the architecture and subsequently neurons from these hidden layers are fed to the output neurons producing the desired final output.

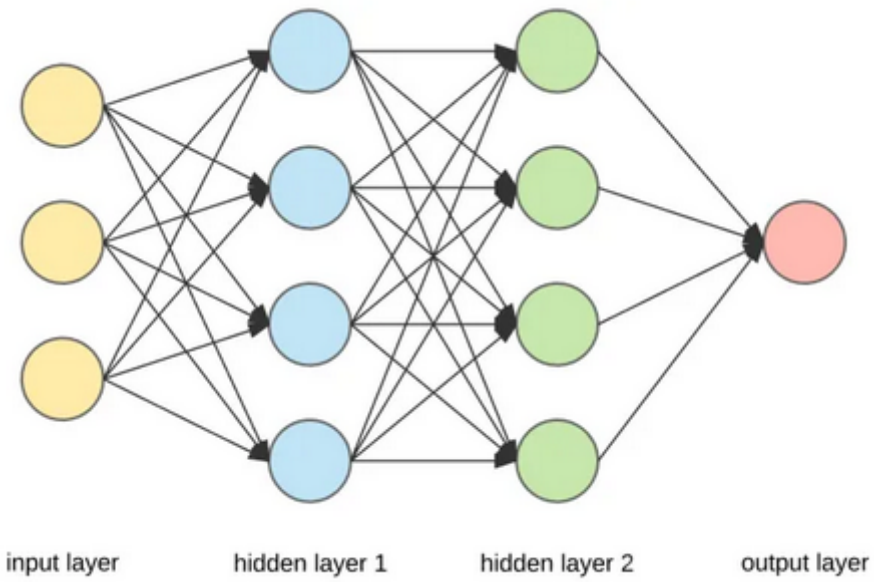


Figure 2.1: Neural Network structure (Lelli 2021)

Each neuron in the network is computed by an activation function and the information always flows in one direction as seen in fig 2.1. The function associated with each neuron is the input feature of the neuron connected with a weight to another neuron and a bias factor is included carrying a certain value to that particular neuron in the hidden layers. Also, an activation function is applied to the weighted sum of inputs. The mathematical representation of each neuron in the neural network architecture is seen in eqn 2.1 where 'b' is bias, 'x' is the input to neuron, 'w' is the weights attached and 'f' is the activation layer.

$$f\left(b + \sum_{i=1}^{n} x_i w_i\right) \tag{2.1}$$

This activation function can be selected based on research and there are common ones like Rectified Linear Units(ReLU), hyperbolic tangent (tanh) of the function and a sigmoid function. This is the basic functioning of neurons in the neural network. There are many ways this can be tuned based on different applications

using deep learning approaches. In this study, the margin for selecting a special type of ANN is narrowed based on the literature survey corresponding to the problem statement. The special type of ANN is termed as Recurrent neural Networks (RNN) which is discussed in the next subsection.

2.3.1 Recurrent Neural Networks(RNN)

These networks are special types of ANN's which have the ability to remember and reconsider the output of each neuron back to itself for adaptive learning of the problem. Fig 2.2 represents the network outline of a typical RNN.

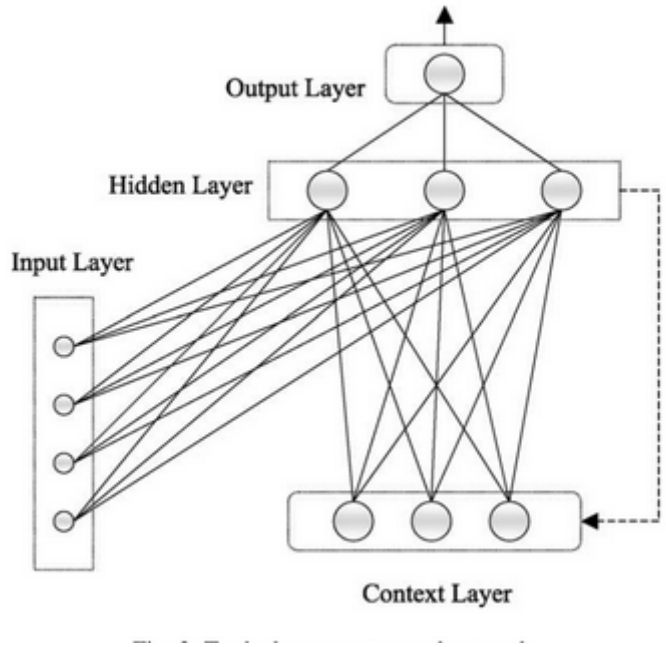


Figure 2.2: RNN architecture (Daliakopoulos, Coulibaly, and Tsanis 2005)

Another speciality of RNN's is that a memory cell is created in the network for the model output to depend on the current state of the network and the current input across time steps. Unlike Feedforward neural networks where the weights and biases of the neurons are fixed, RNN's use recurrent layers where the inputs have the ability to modify the internal state of the network . Each neuron should have the ability to store information and modify it when the neural network is trained. The commonly used units for this purpose are LSTM and GRU(Gated Recuurent Units).

2.3.1.1 LSTM units

LSTM model was first intorduced by Horchreiter and Schmidhuber in the late 1990's (Sagheer and Kotb 2019). LSTM's are special memory units in neural networks that adapt to learn long-term dependencies that can be used in analysis of time series data (Sagheer and Kotb 2019). LSTM blocks are comprised of an input gate, a forget gate and an output gate, see fig 2.3.

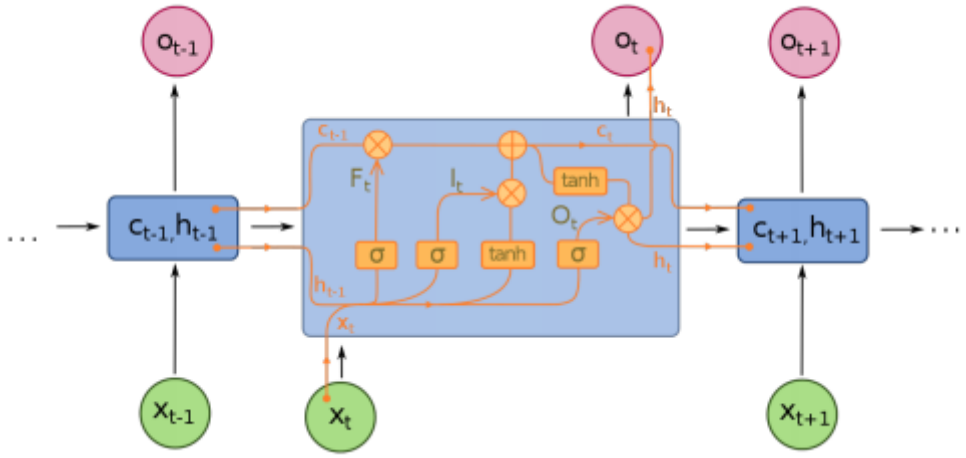


Figure 2.3: LSTM block with input, output and forget gates at different. + and X sign indicates pointwise addition & multipication(Sagheer and Kotb 2019)

LSTM's process information by either adding or removing it through the whole neural network chain that are regulated by different gates. Each gate uses a point-wise operation to estimate the outcomes of the hidden state as seen in fig 2.3. Computation of the hidden state 'St' at time step 't', the input is 'Xt' and the hidden state from the previous state is 'St-1'and a step wise operation of the gates are explained below. In the initial phase, the forget gate decides what information is going to be discarded based on prior information and is represented as 'ft', see eqn 2.2

$$f_t = \sigma\left(X_t U^f + S_{t-1} W^f + b_f\right) \tag{2.2}$$

In the next phase, which new information is kept in the cell state is decided using the input gate and the cell state values. This operation happens using an input gate 'it' which updates the new values and a tanh layer that uses vectorization to create new cell state values 'Ct'.These two equations are represented in eqn 2.3 and eqn 2.4

$$\mathrm{i_t} = \sigma\left(X_t U^i + S_{t-1} W^i + b_i\right) \tag{2.3}$$

$$f_t = \sigma\left(X_t U^f + S_{t-1} W^f + b_f\right) \tag{2.4}$$

Now, the old cell state 'Ct-1' is updated into the new cell state 'Ct' given by eqn 2.5

$$C_t = C_{t-1} \otimes f_t \oplus i_t \otimes C_t \tag{2.5}$$

In the final step, the output gate is computed using the cell state contributing to the output, see eqn 2.6. The cell state is passed through a tanh layer which is multiplied with the output to obtain the desired hidden state, see eqn 2.7.

$$0_t = \sigma\left(X_t U^0 + S_{t-1} W^0 + b_o\right) \tag{2.6}$$

$$S_t = o_t \otimes \tanh\left(C_t\right) \tag{2.7}$$

The notations used in the LSTM are as follows:
Input Weights: $\mathrm{U}', U^i, U^o, U^c$
Recurrent weights: $\mathrm{W}^f, W^i, W^o, W^c$
Bias: $\boldsymbol{b}_f, \boldsymbol{b}_i, \boldsymbol{b}_0, \boldsymbol{b}_c$

2.3.1.2 Hyperparameters

In deep learning, hyperparameters form the backbone in modelling and learning data. The training dataset is influenced by hyperparameters optimization to achieve good performance. Implementing these parameters in the LSTM network architecture is an exploratory task which helps define how many fully connected layers to consider, number of units in the network, what type of activation function to use, the dropout layers to use in the network, batch size of data etc. The selection of these parameters can be done either by grid seach where all hyperparamters are searched in combinations or random search where the hyperparameters are sampled randomly or by other relevant techniques such as trial and error by assessing the fit of the learning curve(Wikipedia 2021).

2.3.1.3 Training the LSTM neural network

The performance of neural networks is determined by training the networks with the right hyperparameters. Evaluation of the neural network is done using a loss function, i.e. Mean Squared Error(MSE) used for regression tasks. Moreover, training data is usually passed through the neural network using backpropagation(BP) method (Rumelhart, Hinton, and Williams 1986).

2.4 WestLink - Commuter Rail Project

The Swedish transport administration, Trafikverket has commissoned the WestLink rail project which is about 6.6km tunnel covering major hubs of Gothenburg city. The project is ongoing and expected to be completed in 2026 (Ulf. Sundkvist and Thomas. Wallroth 2016). WestLink carries the responsibility to facilitate mobility for an estimated higher population in the future keeping in mind the sustainable development.The tunnel will have three underground stations at Haga, Centralen and Korsvagen. The tunnel is mainly driven through clayey soil and crystalline bedrock (Ulf. Sundkvist and Thomas. Wallroth 2016). Construction techniques vary depending on the stratification of the site, for example a concrete tunnel section would suffice in the soil and drilling/blasting with grouting would suffice in case of a rock tunnel section. The tunnel encounters groundwater inleakage tackled by measures such as draining water from openshafts, pumping of water through pipes at the mouth of the tunnels and protective water systems during construction.(T. Wallroth et al. 2016). A detailed outline of the WestLink can be seen in fig 2.4 . In this thesis the scope is narrowed by understanding the geological and hydrogeological features around the Linné service tunnel which is marked in fig 2.4.

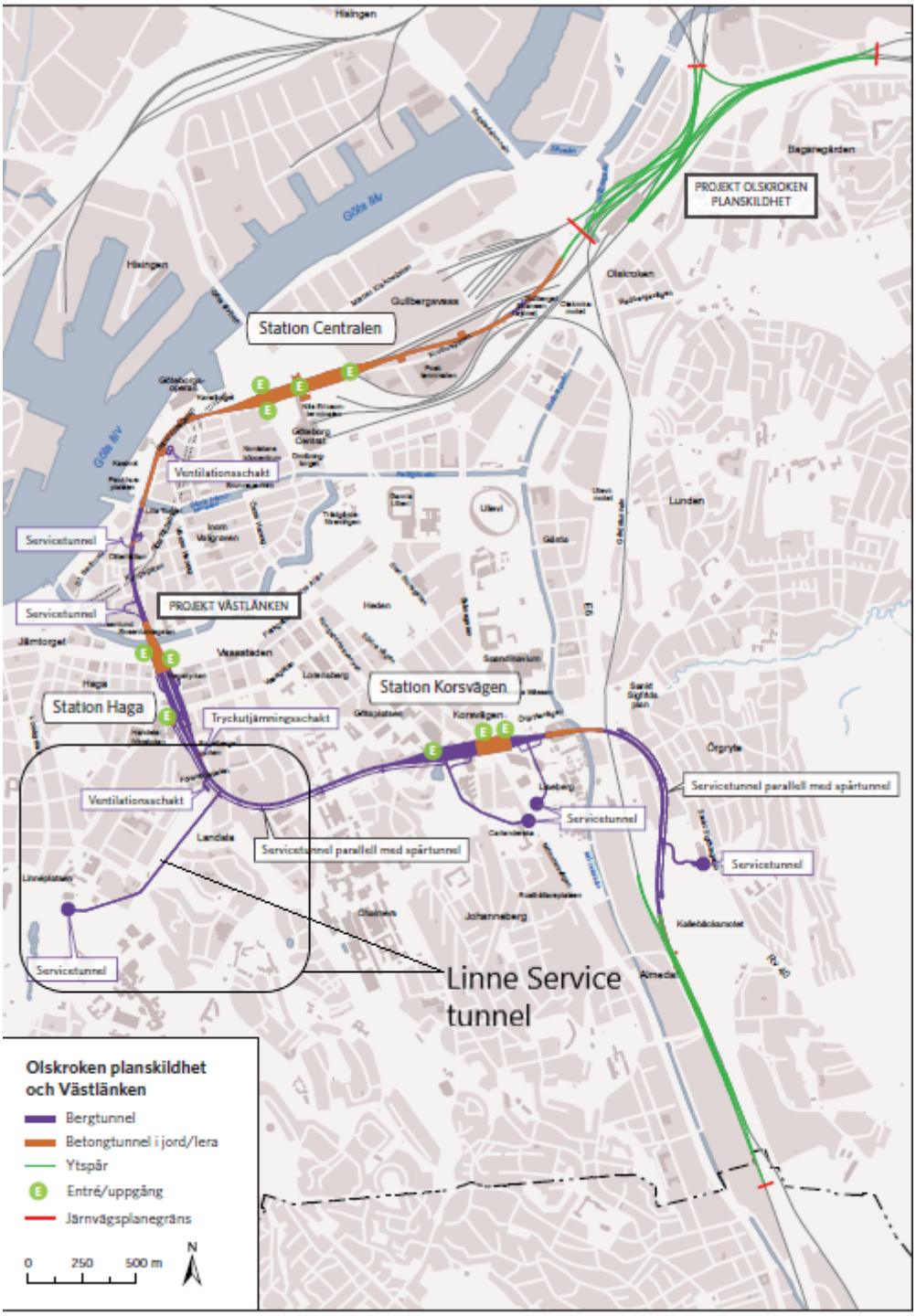


Figure 2.4: Top View of the WestLink tunnel(Ulf. Sundkvist and Thomas. Wallroth 2016)

2.4.1 Linné service tunnel

This section of the WestLink is about 920m long and acts as a service provider to the main tunnel (at Haga) once the construction is completed. The Haga/Linné service tunnel stretches from Linnéplatsen through a crystalline rocky mountain

which connects to the main tunnel passing through Landala (Lithén and Wadsten 2016). The tunnel is driven through crystalline bedrock along all of its path and will be used as a working tunnel during the main tunnel construction. The tunnel mouth begins at Linnéplatsen +20 m.a.s.l and connects to the main tunnel at Landala in the north-east direction with -20 m.a.s.l.(Ramboll 2014).

2.4.1.1 Geology

The geological setting around the service tunnel mainly comprises of bedrock below the ground surface and is surrounded by clay in the north-west regions towards Haga. The thickness of the clay layers vary from about 40-100m around this region with a minimal layer of glacial till below followed by the bedrock (Lithén and Wadsten 2016). The geological stratification of the area between Annedal and Olivedal is seen in fig 2.5.

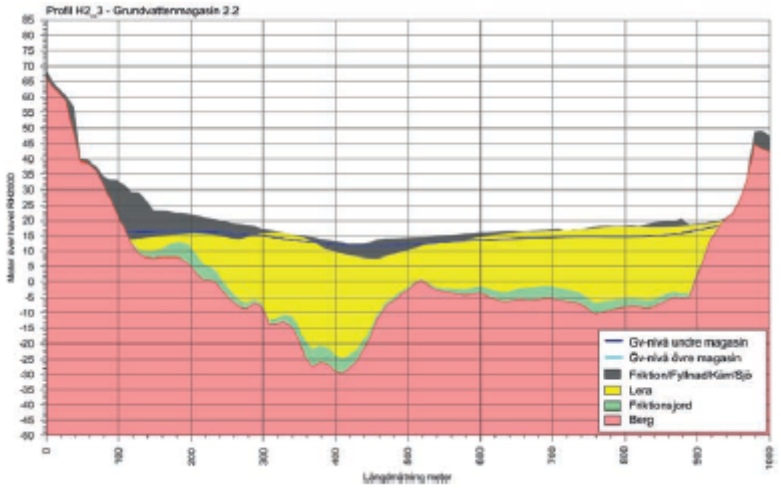


Figure 2.5: Stratification of soil layers from Olive-Annedal from west to east of the service tunnel region(Ulf. Sundkvist and Thomas. Wallroth 2016)

The weakness zones of the bedrocks are seen around the Haga Service tunnel in fig 2.6 obtained from SGU archive maps.

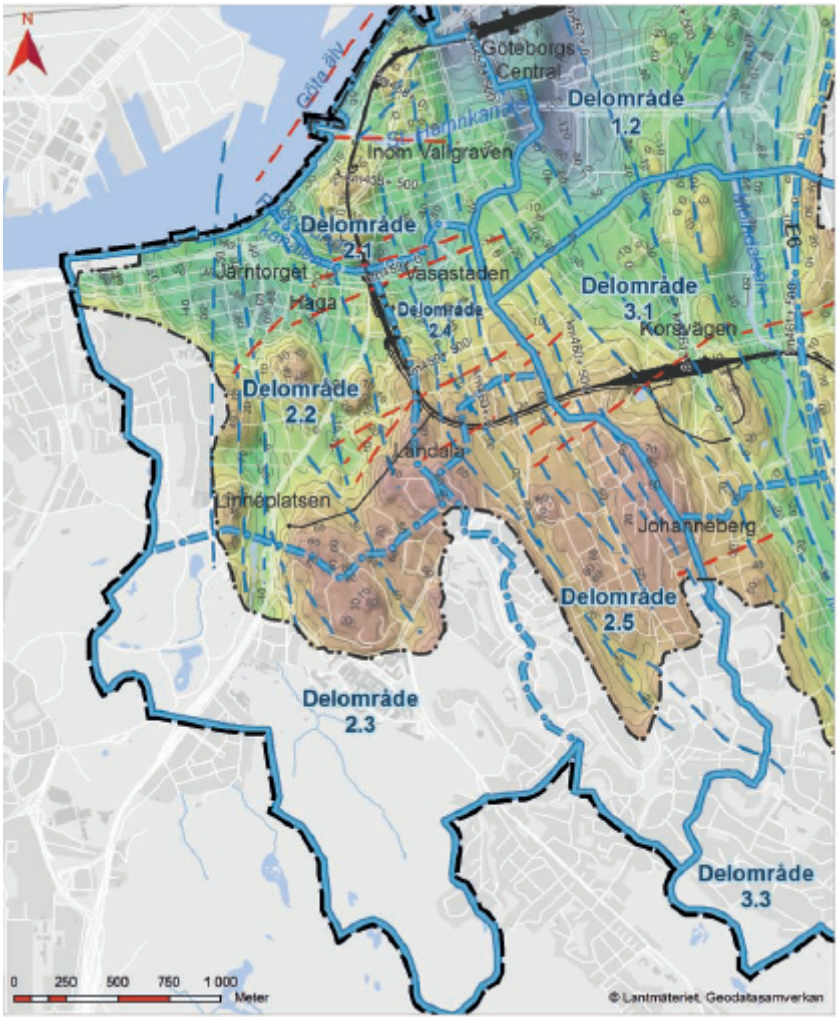


Figure 2.6: Weakness Zones around the service tunnel shown by red and blue dashed lines(Ulf. Sundkvist and Thomas. Wallroth 2016)

These zones are marked at approximate locations at the rock surface in the area of influence.The weakness zones can have significant amount of permeability in the rocks.

2.4.1.2 Hydrogeology

Groundwater conditions in the bedrock and soil layers around the Haga service tunnel region gives a brief understanding of the hydrogeological features further used in conceptual modelling. Majority of the aquifers are confined by a layer of clay and other filling materials in the study area based on historical geological formation. Moreover, shallow aquifers are present in the filling materials which is contact with the mountainous part of the area as seen in fig 2.5. Surface runoff is also a primary factor affecting groundwater conditions by penetrating into the earth stratum. An extensive amount of surface water runoff around the service tunnel flows to relevant groundwater aquifers divided according to their topography and finally flows into the Göta älv river(Ulf. Sundkvist and Thomas. Wallroth 2016). Also natural groundwater recharge to the confined aquifer occurs in transition zones

in areas where the clay layers are thin. Also from fig 2.7, we can see the topography with surface water dividers and surface water courses.

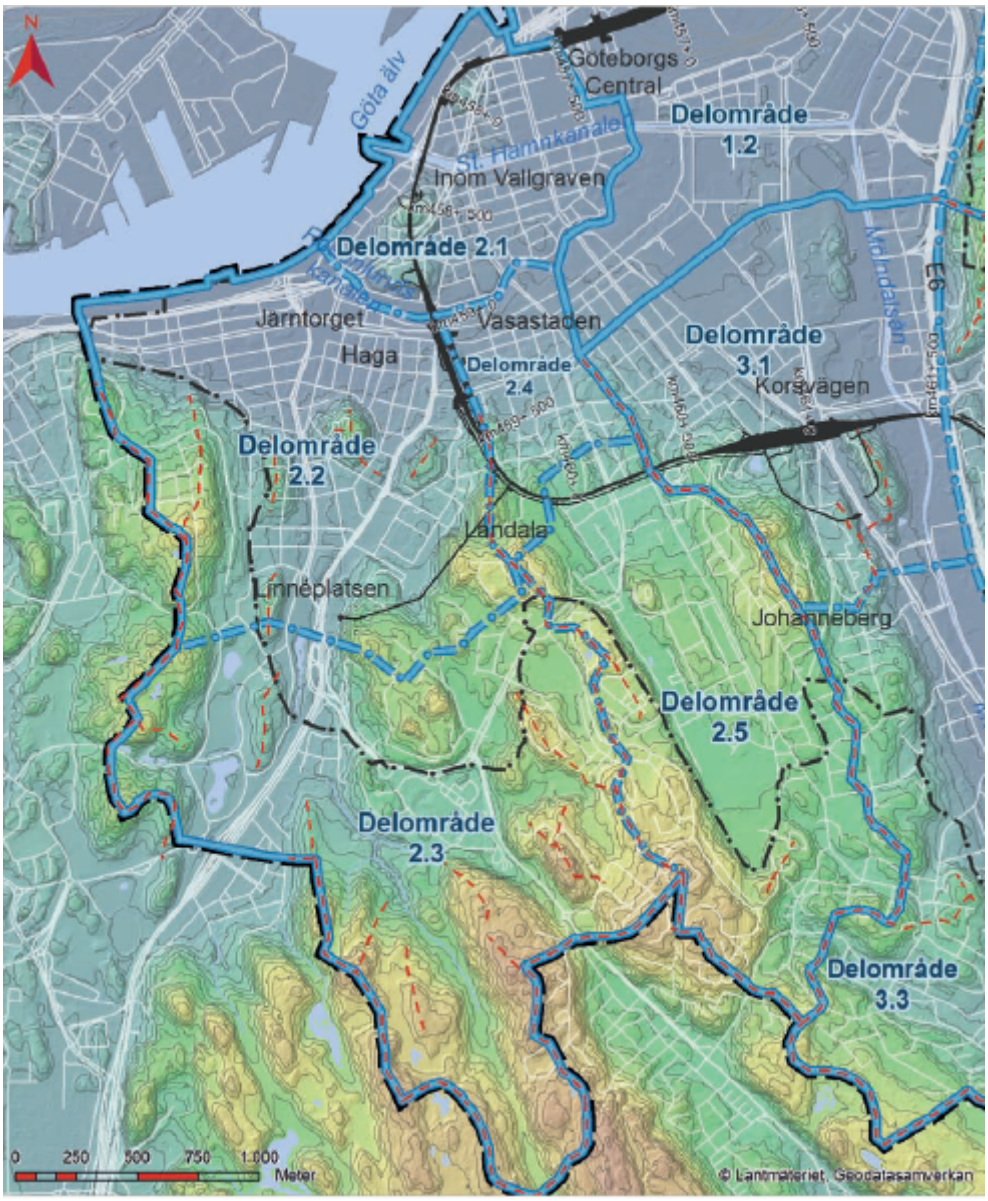


Figure 2.7: Surface water being divided based on topography around the West Link Commuter rail project(Ulf. Sundkvist and Thomas. Wallroth 2016)

Groundwater levels in rocks are complex to visualize with its irregularity in flow patterns caused by fractures and cracks.(T. Wallroth et al. 2016)

The hydraulic properties of the soil around the service tunnel is seen to be around $1 * 10^{-}7$ to $1 * 10^{-}4$ m/s in soils (Ulf. Sundkvist and Thomas. Wallroth 2016). The lower values indicate dense soils such as silty - muddy silty soil around the area based on prior soil tests conducted on early constructions in the area. The higher values indicate the exposure of soil stratum with fractured rock surfaces, typically seen with sandy-moraine and friction soils. Hydraulic conductivity in rocks sections upto the tunnel depth is estimated to be around $4 * 10^{-}8$ to $2 * 10^{-}7$ m/s and in

the bedrock along with zones of weakness it is found to be $2 * 10^{-}7$ to $5 * 10^{-}6$ m/s (Ulf. Sundkvist and Thomas. Wallroth 2016).

2.4.1.3 Groundwater lowering

The groundwater levels vary based on the topography and it is seen that the southern parts of the Haga-Annedal region have highest water levels at about +50 in Guldheden and progressively reduces towards Sahlgrenska Hospital with +30 to the west and Chalmers University at +36 to the east. Lower groundwater heads are observed at Linné sqaure at +13 due north and further reduces at Vasastaden and subsequently at the Göta älv river at its lowest(Ulf Sundkvist 2016). This gives an insight on the natural groundwater lowering conditions around the Haga service tunnel. Groundwater lowering occurs in the confined aquifer of the region based on the topography of the construction. Leakage of groundwater into a rock tunnel occurs generally if the groundwater level is above the tunnel and this subsequently causes groundwater lowering given the impact from the tunnel construction(Lithén and Wadsten 2016). The magnitude of groundwater lowering has an influence from the geological materials and its hydraulic properties. See fig 2.8 showing the leakage in tunnels causing groundwater lowering.

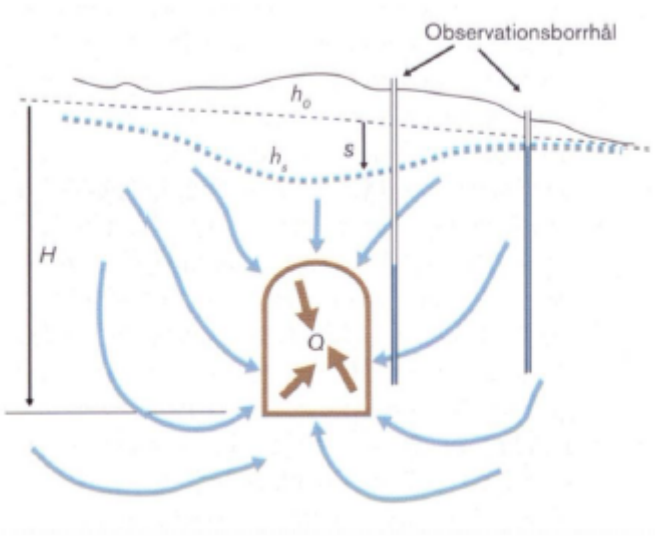


Figure 2.8: Representation of Groundwater lowering in tunnels

3
Method

3.1 Working Strategy

Groundwater Levels(GWL) predictions using deep learning methods from relevant historical time series data and meteorological data can be challenging in adopting the modelling approach. Groundwater wells around the service tunnel are selected and mapped from coordinates provided by Trafikverket, see fig 3.1.

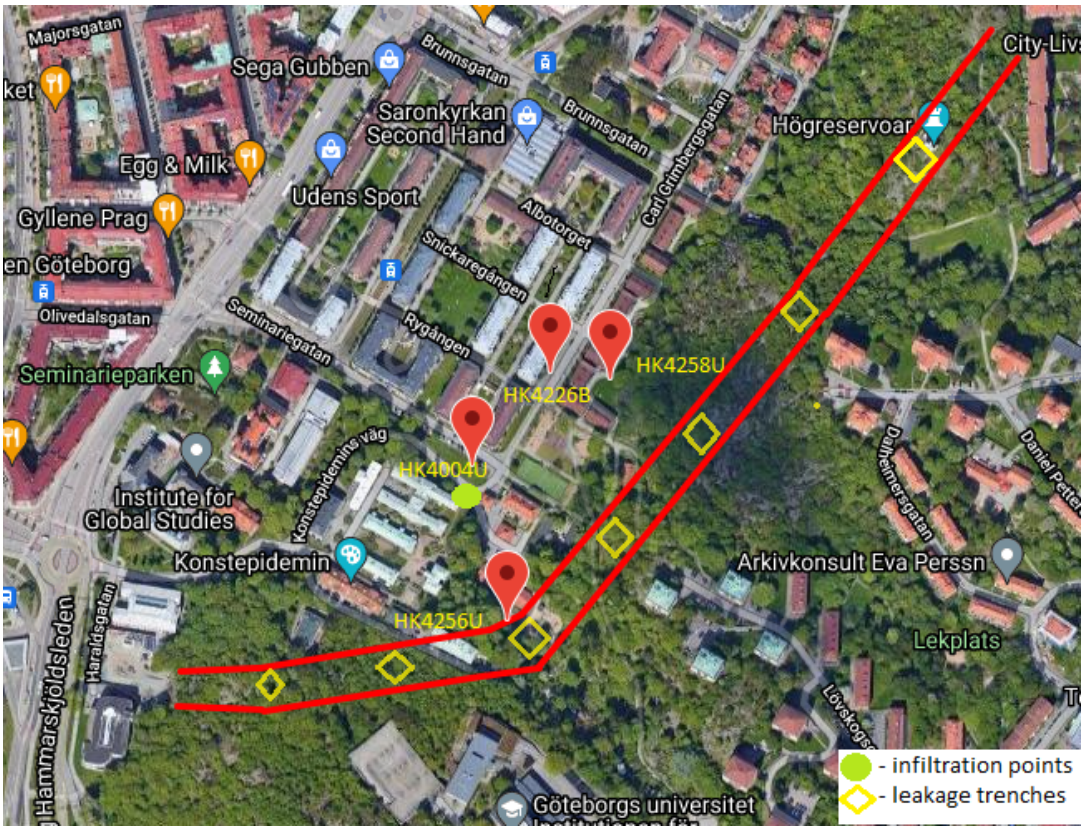


Figure 3.1: Location of selected wells with respect to the Haga service tunnel

The meteorological factors affecting groundwater conditions naturally around the perturbed source include precipitation, relative humidity and evapotranspiration. The wells chosen for modelling require thorough data collection and preparation with pre-processing techniques such as interpolation, negating missing values, resampling etc. In this thesis, python is used as an open source interface to perform statistical data modelling using open source packages such as pandas, numpy, matplots and

other deep learning libraries like tensorflow, Keras and SciKit-Learn. Applications of these deep learning libraries is explained later in this chapter. In fig 3.2, the working method is represented by a flowchart to achieve desirable results.

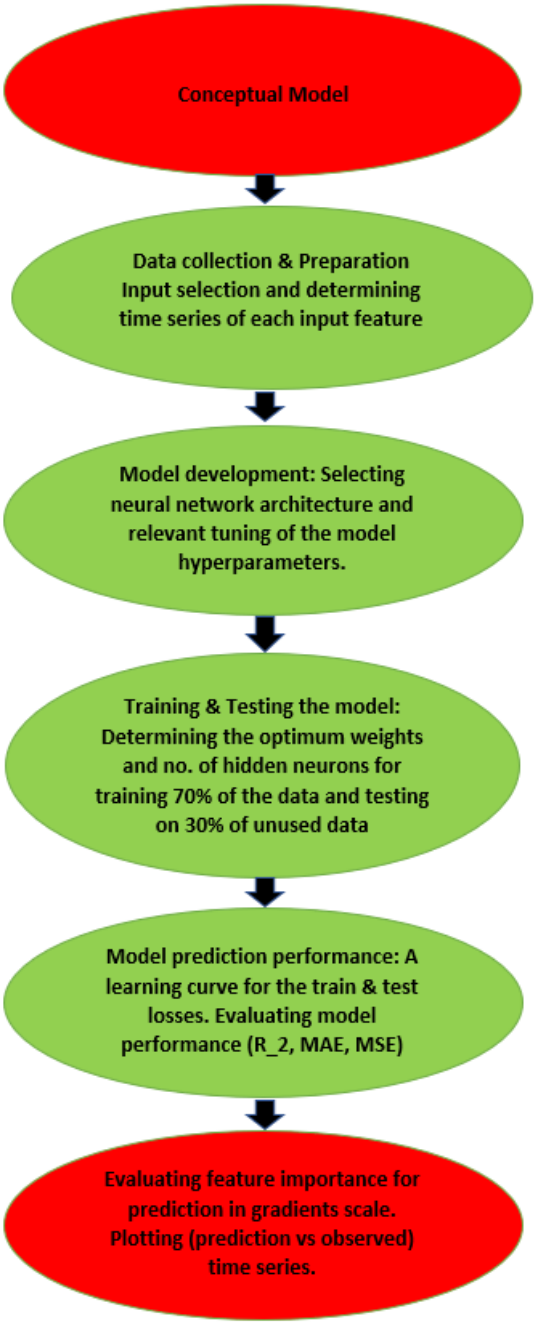


Figure 3.2: Flowchart depicting the working approach

A conceptual model around the wells location allows interpretation of the hydrogeological features of the area used for evaluating results from data modelling. The wells are located around the Haga-Korsvagen strip of the Vastlanken project, specifically at Annedal for this study.

3.2 Conceptual model

Conceptualization of the tunnel and well location surrounding the service tunnel reveals the hydrogeological features of the study area. The conceptual parameters include hydraulic conductivity, groundwater head, groundwater recharge and storage coefficient(Stejmar Eklund 2002).However, all these parameters are not considered in modelling the groundwater heads based on information available of the study area. Additional information on the wells are seen in table 3.1.

Well ID	Well status	Reference GWL level (m.a.s.l)	Groundwater resovoir
HK4004U	Active	20.85	Confined
HK4256U	Active	27.53	Confined

Table 3.1: List of wells selected for modelling with parametric information

Time series plot of the wells are plotted as seen in fig 3.3. This reveals the groundwater pattern in the wells and the trend in the pattern shows seasonality changes. The plot also features drawdown caused by infrastructural activities around the wells. From fig 3.3, we see the wells selected with code HK (Haga-Korsvagen) at Annedal have reduced groundwater levels since the commission of tunnel construction. Meteorological factors show seasonal variation which influences the groundwater patterns.

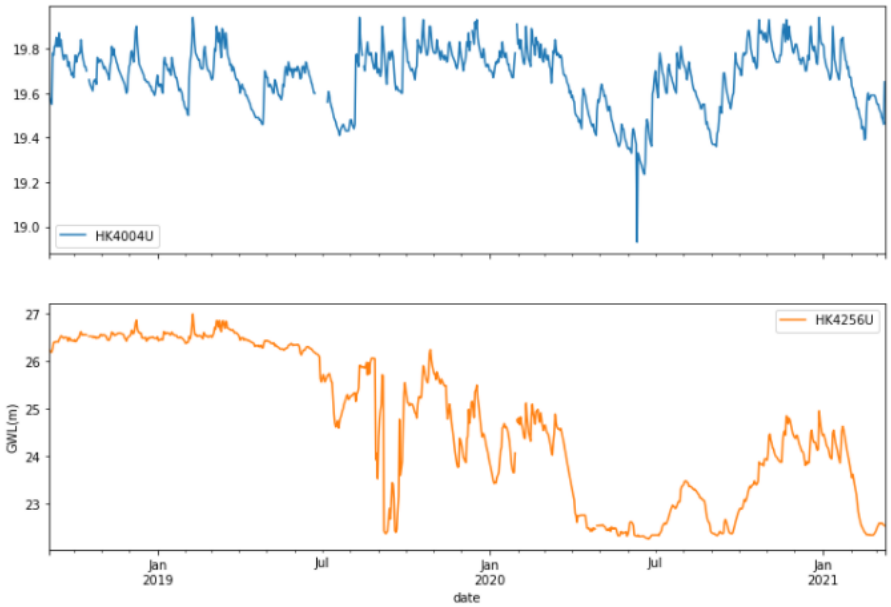


Figure 3.3: GWL time series of HK4004U and HK4256U around the Haga service tunnel

Furthermore, the movement of water occurs from the higher elevated areas in the south to the lower elevated regions in the north and finally flowing into the Gota Alv river with regards to the study area (Ulf. Sundkvist and Thomas. Wallroth 2016). From fig 2.5, the east side of Guldheden is mountainous sloping downwards

to Linnéplatsen and rising again to the west of Olivedal. This shows a valley like formation, where the rocky depressions are filled with clay, friction soil and filling materials (Ulf. Sundkvist and Thomas. Wallroth 2016). However, the groundwater reservoir is confined by a layer of clay indicating the aquifers are primarily confined in this region. Also, there are traces of unconfined aquifers found in the region but the stresses are mainly applied on the confined aquifer for water extraction. Detailed explanation on the geological and hydrogeological features of the project site is seen in section 2.4.1.

3.3 Data Collection & Preparation

The data used for modelling is collected from two organisations, one being Trafikverket (Swedish Transportation Administration) and the other one SMHI (Swedish Meteorological and Hydrological Institute)(SMHI 2021). Trafikverket provided groundwater data for different wells around the Linne service tunnel with their respective hydrogeological features.
The input data selected for statistical modelling using neural networks are groundwater levels, infiltration rate, leakage rate and meteorological parameters like precipitation, relative humidity, evapo-transpiration. The meteorlogical parameters are used to check for natural impacts on the final predictions in the study. Infiltration rate is the amount of water pumped into the ground to achieve groundwater recharge. Leakage is measured at trenches along the tunnel to quanitfy the leakage rate of groundwater around the tunnel construction. Input data comprised of 19 wells around the service tunnel region with daily time periods from years July 2014 - March 2021. Most of the wells had missing groundwater head information and required thorough data pre processing. Finally, 2 groundwater wells were selected with the longest daily data from August 2018 - March 2021 for the modelling process. Precipitation and relative humidity data is retrieved from Gothenburg station A using SMHI web (SMHI 2021). Evapo-transpiration is calculated using a python package called Eto (Kittridge 2021) which is estimated using temperature, relative humidity, latitude and longitude as variables retrieved from SMHI. The unit of measurement of all the input parameters used in this project are listed in table 3.2.

Input data	**Unit of measurement**
Groundwater level	Meters(m)
Relative humidity	%
Precipitation	Milli-meters(mm)
Evapotranspiration	Milli-meters/day(mm/d)
Infiltration	Litres/min(l/min)
Leakage	Litres/min(l/min)

Table 3.2: Input data units of measurement

3.3.1 Data preprocessing

With data gathered, the next step involves pre-processing data for modelling using python(Pandas 2021). The time series data is made stationary to have a constant mean, variance and auto-correlation thoughout the modelling process for a dataset. This is usually implemented to render statistical modelling on the dataset with less complexity. It is then used to checked for null values, outliers in the dataset and missing timestamps to check data density. This strategy accounts for the quality and quantity of available data in the dataset. Groundwater wells with maximum groundwater head information is selected for modelling. In the dataset, out of 19 different wells only 2 wells had longest information with least missing groundwater heads and are used in the model. Imputing missing values in the dataset can be done various ways, for example by forward/backward filling, interpolation etc. Very few missing values for the wells selected from this dataset are imputed using the linear interpolation function with respect to the time using Pandas (Pandas 2021). The interpolated values to some extent influence the modelling performance. There are many ways to interpret these values using ANN's instead of simple interpolations. In this thesis, advance methods to interpret missing values are not discussed.

3.3.1.1 Data Resampling

In time series data, resampling is widely used to reduce or add time steps from the dataset for statistical modelling. In this study, relative humidity is deduced from hourly timestep to daily time step using the sampling function in Excel before importing the file into python. The observed groundwater data is in daily time steps except in some cases, hourly measurement recorded is tackled similarly. Precipitation is available in daily time steps.

3.3.1.2 Data Scaling

Normalizing and standardizing data is beneficial for neural network models to train and learn faster. Normalization is a scaling method where the input features are scaled between range 0 and 1 using eqn 3.1.

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{3.1}$$

Similarly, standardization scales the data where the values are estimated from the mean and the standard unit deviation using eqn 3.2, where μ is the mean and σ is the standard deviation in the dataset.

$$X' = \frac{X - \mu}{\sigma} \tag{3.2}$$

The idea behind scaling is to numerically set the columns in the dataset to a common scale and to ignore the measurement unit of each input feature. There is no thumb rule to use any particular scaling technnique, it is usually better to try both normalization and standardization with the dataset and check for the learning ability of the neural network. For this project, standardized data is used as an input

for the LSTM model. Another important reason for scaling data is for the gradients in the model to converge faster for better learning rate of the training and testing curves. The implementation is done using a powerful machine learning library called Scikit-Learn using python and is referred to as StandardScaler (ScikitLearn 2021).

Another useful function in pandas is the rolling sum technique. This allows in checking the average response time between two independent variables in a dataset, especially useful to examine correlations for time series data. For example, initially two variables do not correlate but when using the rolling sum method the values add up cumulatively for a given time period showing different correlations for different time periods. This motivates the use of an appropriate window length in the model.

On data collection and preparation, the range of input data used for modelling is listed in table 3.3

DataSet	First Date	Last Date	No. of observations	Time Steps
HK4004U (Well ID)	2018 - 09 - 04	2021 - 03 - 10	919	Daily
HK4256U (Well ID)	2018 - 09 - 04	2021 - 03 - 10	919	Daily
Precipitation	2018 - 09 - 04	2021 - 03 - 10	919	Daily
Relative Humidity	2018 - 09 - 04	2021 - 03 - 10	919	Daily
Evapotranspiration	2018 - 09 - 04	2021 - 03 - 10	919	Daily
Infiltration	2018 - 09 - 04	2021 - 03 - 10	919	Daily
Leakage	2018 - 09 - 04	2021 - 03 - 10	919	Daily

Table 3.3: List of input data, range of input data and time steps used in the the LSTM model

3.4 Groundwater Modelling using LSTM

Implementing the LSTM neural network for the processed and scaled input data for groundwater predictions is explained step by step in this section. Time series differencing is the initial step implemented to stabilize the mean and eliminate the seasonality trend in the dataset(Hyndman and Athanasopoulos 2018). Differencing is basically the difference between consecutive data points in the dataframe and this makes the time series stationary. To ensure the data does not shuffle in the modelling process as it is independent of time after it is being scaled, a time series generator function is implemented from Keras (Keras 2021). Groundwater level is the target for prediction and the features used are the observed groundwater levels, infiltration level, leakage and all the meteorological parameters respectively. Using historic observed groundwater as a feature to predict the groundwater levels improves model performance. Especially considering GWL time series always have significant auto-correlations for many lags (Wunsch, Liesch, and Broda 2020).

3.4.1 Training and testing the data

Input data is split as training and testing before fitting it into the neural network model. The idea is to train the data and test it on the test set for assessing the

performance of the model. Generally, a standard rule is to split the data as 70% training and 30% for testing which is followed in this project, see fig 3.4.

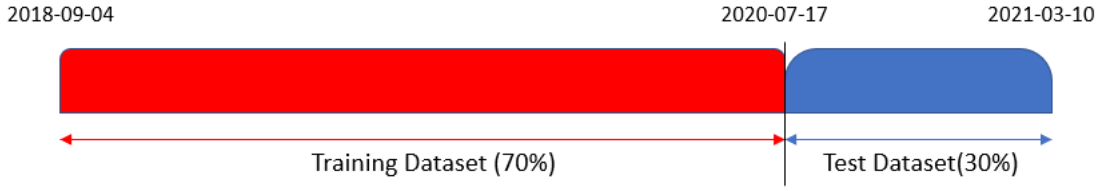


Figure 3.4: Training and testing data splitting arrangement

3.4.2 Neural Network Architecture & Computational Aspects

To fit the data in the network architecture, we need to define the number of days to predict in future based on the window length and select a certain batch size for training the values. Window length is defined as the number of days in the past used for predicting one day in future per batch size . The network is defined using tensorflow/Keras (Keras 2021), which is an open source deep learning library for ANN's. Keras works as an interface to tensorflow and is used for sequential modelling of data that enables constructing the model. Each LSTM layer is defined using keras and tensorflow with special hyperparameters like activation functions and dropout layers attached which enhances the training ability of the network. An activation function produces the final value computed by a neuron in the network and helps understand the complex pattern in the data. A leakyReLU (Rectified Linerar Unit) activation function is used in this project to handle gradients with null values, i.e. it has the ability to return a very small linear value into the model preventing redundancy in neuron behaviour. Finally, the model is constructed using one input layer along with two hidden layers and one output layer forming an embedded LSTM network.

The model is complied using tensorflow and keras, that accounts for metric losses, i.e Mean squared error, mean absolute error coupled with an optimizer for training the model accurately. The optimizer used is called Adam which is widely used in deep learning models. Subsequently, this is an alternative deep learning algorithm for stochastic gradient descent optimizer. The weights attached to each neuron in the network is continuously updated using the Adam optimizer to reach a global minimum during the training process. Additionally, this optimizer learns and adapts to noisy patterns observed in time series data during computation. The model is then fit to a train generator and is validated against the test generator to determine the error metrics and to evaluate the learning curve. Since the learning process is iterative using the gradient descent algorithm, the efficiency of training is defined using another hyperparameter called epochs. One epoch basically passes the entire dataset thorugh a neural network once which is not enough for the gradients to converge faster and learn quickly. Therefore, higher epochs improves the training process however, there is no real proof of using a certain number of epochs and is

defined by trial and error (Sharma 2021). A learning curve is plotted between the training and testing losses once the model is compiled. Based on the fit from the learning curve, hyperparameter tuning by trial and error could improve the performance of the model. For example, adding or removing hidden layers, increasing or decreasing the hidden units, changing the dropout parameters, tuning the epochs and tuning the activation functions. However, in this project the hyperparamters were evaluated based on the fit of the learning curve. To maintain consistency in supervised learning problems, the hyperparameters are made constant for different scenarios presented in the results.

The final step involves re-scaling the predictions back to its original values using inverse transform functions with pandas and attaching it with the test set of the original dataframe to visualize the fit of the model keeping in mind the date-time index (Pandas 2021). The inference on predictions performance is examined using the tensorflow gradient tape function. This function gives the average impacting input feature contributing to the predictions made one day ahead (Cerliani 2020). Furthermore, an inbuilt correlation function in pandas following the Pearson correlation coefficient (R) principle is used for evaluating the linearity of two variables and this also helps in determining the R^2 score, i.e. the coefficient of determination. This is helpful when evaluating the model performance of GWL's predictions and forecasts with unseen data.

4

Results & Discussions

In deep learning or supervised learning or data driven models, there is scope for continuous improvements of results and is influenced by human trial and error hyperparameter tuning. From the data preparation and processing step , it is concluded to use the wells from table 3.3 for prediction analysis. A groundwater time series is plotted along with the meteorological and hydrological data, see fig 4.1. We can see a clear drawdown in the groundwater heads in both the wells HK4004U and HK456U from July 2020 until Jan 2021, and the idea is to understand the impacting source and also the stresses on the aquifers of the well location.

The results presented are broadly divided into two sections where both use LSTM models to predict the target variable, i.e. the groundwater heads influenced by time. In the first case, the data set is divided into two sets namely the training and testing dataset where the predictions are tested on the test data set. While in the second case, the idea is to forecast the groundwater heads by training the entire data set with a small validation data set to examine the best suited time period for predictions with unseen data.

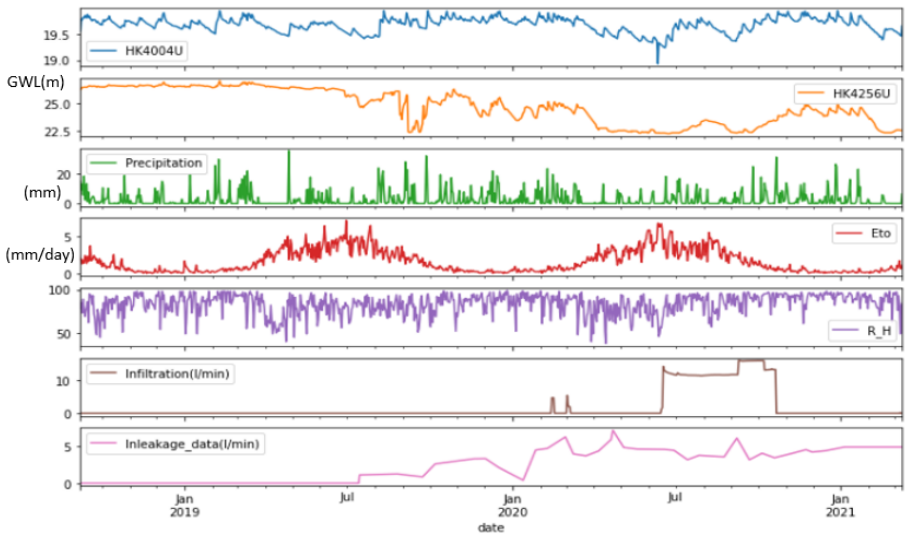


Figure 4.1: Groundwater level time series of selected wells along with meteorological and hydrological time series

4.1 Response time estimation using precipitation

Evaluating the performance with Pearson correlations between 0.5, 1, 2, 3 and 4 months using the rolling sum function for precipitation from the data set. It is seen that 0.5 months shows good response time with promising correlations against groundwater heads for HK4004U and no improvements in response time for HK4256U for more than 15 days rolling sum as seen in fig 4.2 .

	HK4256U	HK4004U	Precipitation	KB_2	15D	1M	2M	3M	4M
HK4256U	1.000000	0.379562	0.031410	0.294204	0.146520	0.142153	0.030727	-0.067152	-0.198699
HK4004U	0.379562	1.000000	0.143238	0.446041	0.665713	0.699348	0.558479	0.454145	0.318675
Precipitation	0.031410	0.143238	1.000000	0.015493	0.291640	0.216241	0.089056	0.072955	0.045062

Figure 4.2: Correlation coefficients for different precipitation rolling sum periods for HK4004U and HK4256U

Correlation plots between precipitation and the groundwater heads for HK4004U and HK4256U are seen in fig 4.3 and fig 4.4. Also, giving insights on selecting an appropriate window length for predictions and forecasting GWL's using the LSTM neural network.

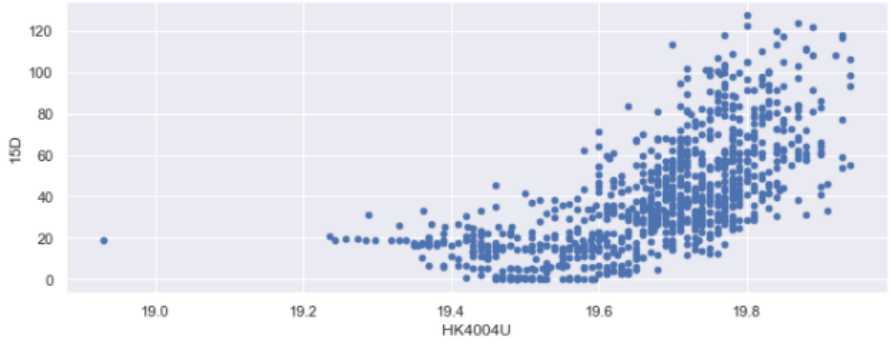


Figure 4.3: Correlations for 15 days rolling sum between HK4004U and precipitation

Figure 4.4: Correlations for 15 days rolling sum between HK4256U and precipitation

It is evident that the response time for HK4256U does not follow a linear trend with highly varying GWL's in meters compared to HK4004U which varies in a few decimeters in GWL's. This indicates that precipitation is independent of groundwater heads for HK4256U and failing to respond well using the rolling sum function. It can be due to the location of HK4256U which is on the service tunnel section as seen in fig 3.1.

4.2 Groundwater predictions

What is the influence on the selected input features used for groundwater predictions?

4.2.1 Sequential modelling using meteorological parameters

The network is modelled using all the meteorological parameters as input features and the groundwater head as the target variable. A window length of 14 days is used for training the model parameters to predict 1 day in future with a batch size of 10. The entire network is built using LSTM layers and compiled using the loss functions for error metric evaluations to analyse predictions. The predictions are made on the test set which comprises about the last 30% of the data which is trained on 70% of the initial data following a time sequence. The hyperparameters used for modelling is listed in table 4.1.

Hyperparameters	Value
Number of LSTM layers	3
Number of LSTM units	64
Activation rate	0.2
Batch size	10
Window length	14
Dropout rate	0.3
Epochs	50

Table 4.1: Hyperparameters used for groundwater modelling

Once the model is compiled, the resulting predictions are attached to the test set and plotted as seen in fig 4.5 and fig 4.6. There is a good fit of the predictions on the observed groundwater levels for HK4004U although it does not account for sudden drawdowns. Also, an overestimation in predictions indicates a possible impact on the groundwater levels. Since the variation in groundwater head is a few decimeters, the stresses observed on the groundwater aquifer are minimal in the area. The predictions made on the HK4256U groundwater well have difference of 2-3 decimeters as the model underestimates the high fluctuations observed. This indicates poor performance of the model lacking the ability to account for impacts which develops critical GWL's. Another reason for poor performance of HK4256U is the location of the well which lies adjacent to the line of the tunnel construction. This also implies

stresses on the aquifer region occur on seasonal patterns and also from non-natural influence in the area.

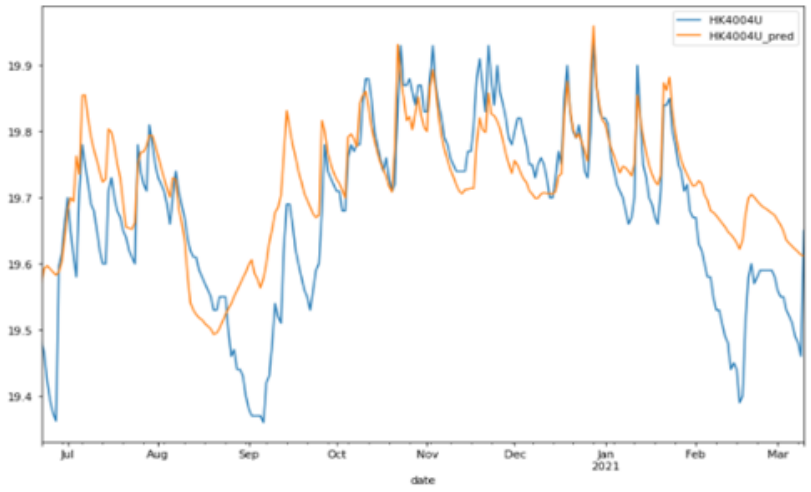


Figure 4.5: Predictions on HK4004U groundwater well with meteorological data

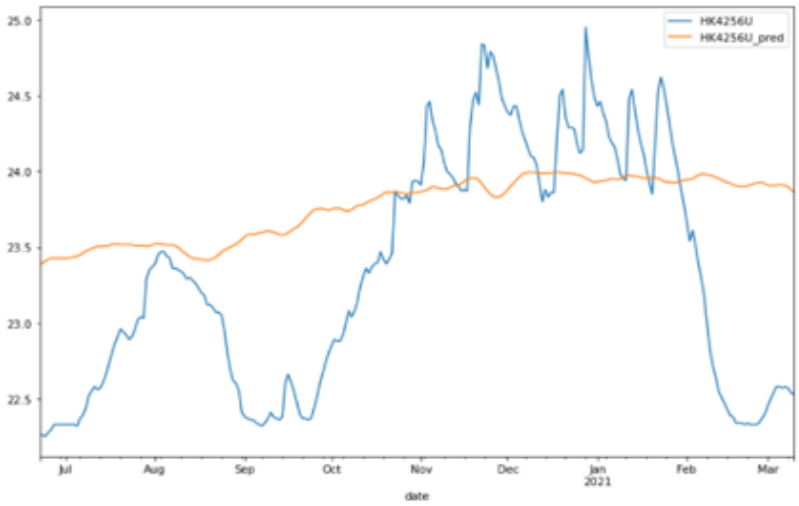


Figure 4.6: Predictions on HK4256U groundwater well with meteorological data

Moreover, in neural network models the impacts on predictions are inferred in terms of average gradients. From fig 4.7 we can see that precipitation has the highest average gradient to predict the groundwater head for one day ahead followed by relative humidity and evapotranspiration for HK4004U. The inference on predictions for the well HK4256U from fig 4.8 indicates both precipitation and evapotranspiration have negative average gradients and do not contribute to the predictions made one day ahead while relative humidity contributes to some extent.

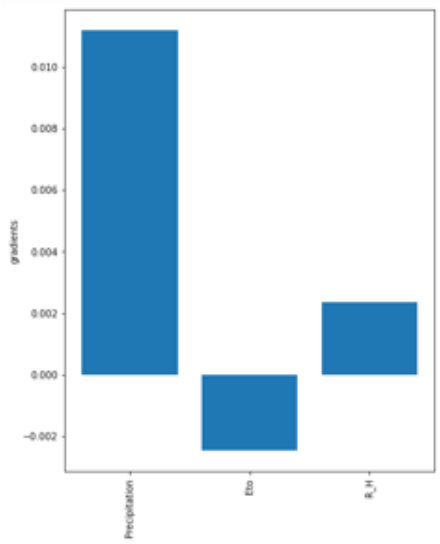


Figure 4.7: Impacting average gradients for HK4004U

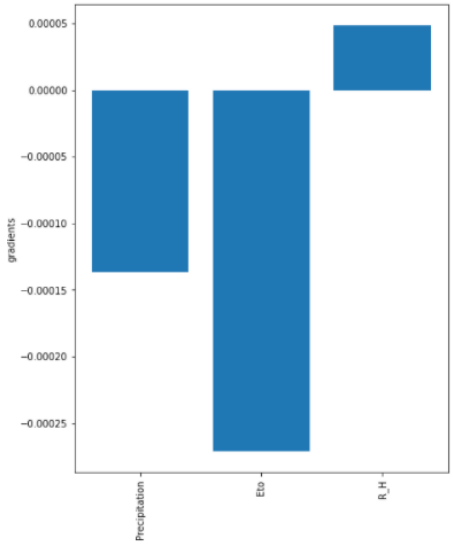


Figure 4.8: Impacting average gradients for HK4256U

Furthermore, to assess the fit, a correlation between the observed and the predicted GWL is plotted. From fig 4.9 and fig 4.10 we can see a linear correlation between the parameters for HK4004U and HK4256U. However, the linearity in HK4256U is hardly observed as the points do not follow a clear direction and are grouped to a certain extent indicating poor performance of the model. From the error metrics, an R^2 score of 0.68 for HK4004U and 0.36 for HK4256U is obtained. Moreover, a higher MAE and RMSE is seen for HK4256U when compared to HK4004U indicating poor performance of the model on evaluating high fluctuations in GWL's.

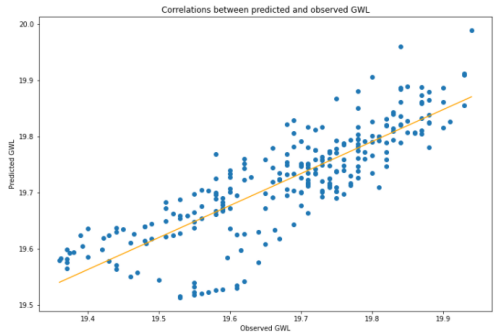


Figure 4.9: Correlations between observed & predicted GWL for HK4004U

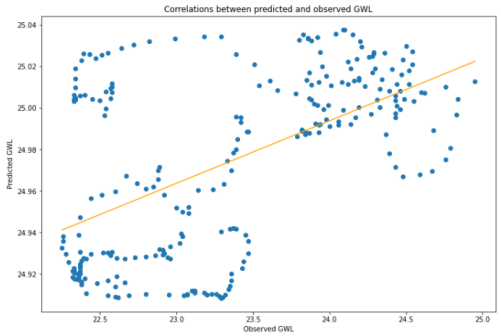


Figure 4.10: Correlations between observed & predicted GWL for HK4256U

4.2.2 Sequential modelling with Meteorological, Infiltration & Leakage data

Assessing the model performance using influencing hydrological parameters along with meteorological parameters gives a better understanding on the GWL's impacted

in the area. The model is compiled using the same hyperparameters listed in table 4.1. Groundwater predictions on the test set for HK4004U and HK4256U are seen in fig 4.11 & 4.12. The predictions for HK4004U has a slower learning rate influenced by additional infiltration and inleakage data. However, there is an overestimation seen in HK4256U indicating the learning rate is slower in observing the groundwater heads with high fluctuations. Also, the model understands the pattern much better when additional hydrological data is added compared to only meteorological data used as input features for HK4256U.

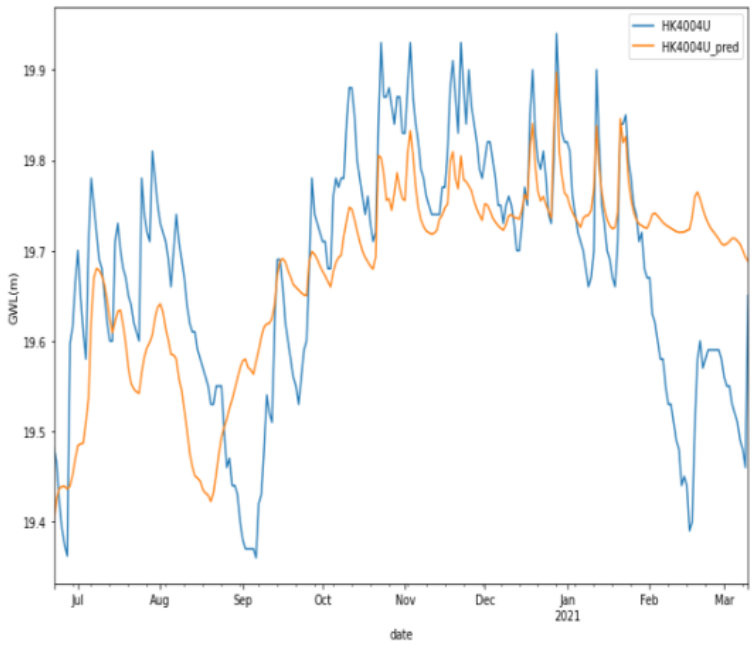


Figure 4.11: Predictions on HK4004U groundwater heads with meteorological, infiltration and leakage data

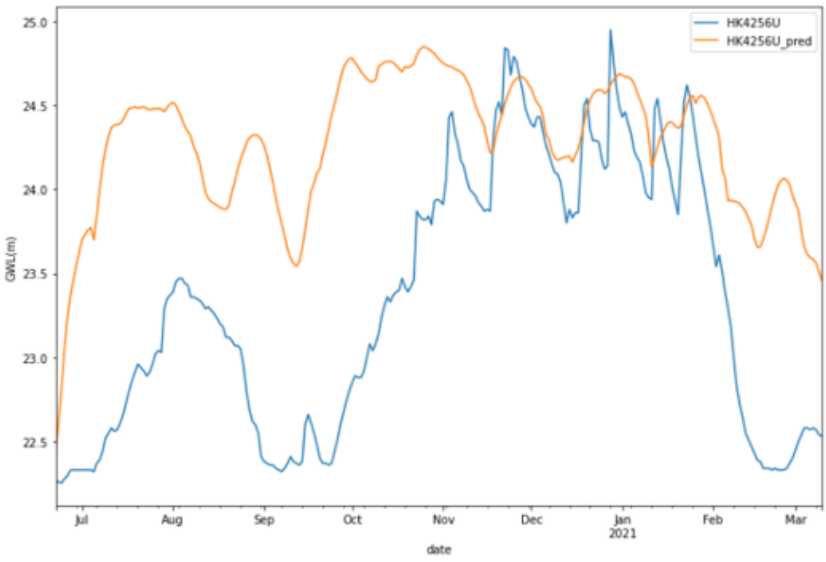


Figure 4.12: Predictions on HK4256U groundwater heads with meteorological, infiltration and leakage data

Similarly, the correlations between the predicted and observed groundwater heads are plotted for HK4004U & HK4256U seen in fig 4.13 and 4.14. The linearity between the points are slightly away from each other indicating an influence from hydrological parameters like infiltration and leakage. From the error metrics,an R^2 score of 0.48 for HK4004U and 0.37 for HK4256U is obtained. This shows the model performance decreases on adding infiltration and leakage data with meteorological data to predict GWL's. Adding on, the MAE and RMSE for both the wells, i.e. HK4004U and HK4256U are higher when compared to the performance with only meteorological data as seen in table 4.2.

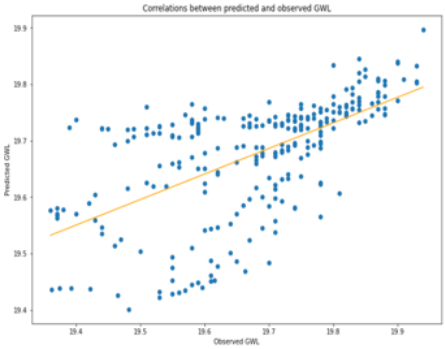


Figure 4.13: Correlation between observed & predicted GWL for HK4004U

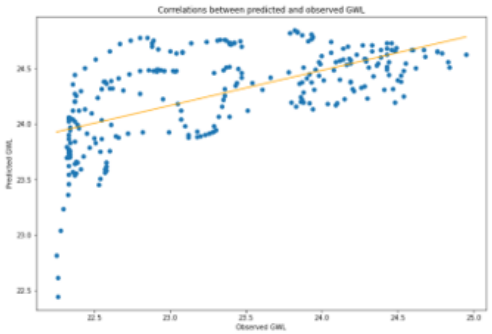


Figure 4.14: Correlation between observed & predicted GWL for HK4256U

Furthermore, inference on the input features average gradients are plotted for HK4004U and HK4256U using the tensorflow gradient tape function as seen in fig 4.15 & 4.16. We see that precipitation has the highest impact on the predictions while the evapotranspiration does not contribute in the learning process, meaning it does not improve in the training process during back propagation.

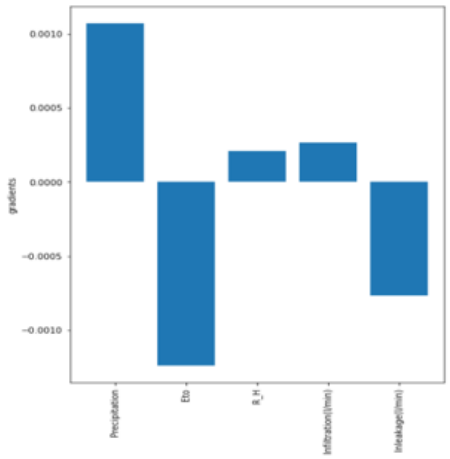


Figure 4.15: Impacting average gradients for HK4004U

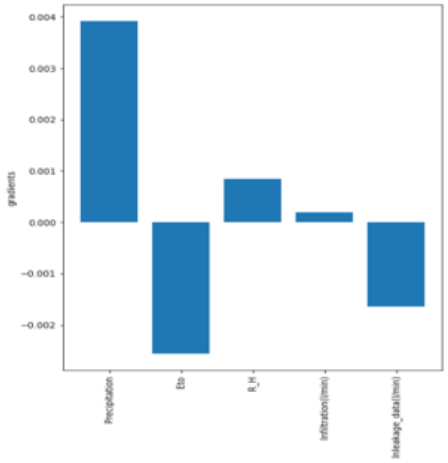


Figure 4.16: Impacting average gradients for HK4256U

4.2.3 Sequential modelling with GWL, Meteorological, Infiltration & Leakage data

Additional GWL and hydrological parameters like infiltration and inleakage(in the tunnel) data is compiled with the same hyperparameters listed in table 4.1. The groundwater head predictions on wells HK4004U and HK4256U are seen in fig 4.17 and fig 4.18. The predictions shows an improvement in the model performance as it understands the seasonality patterns and trends without over estimation for well HK4004U. But looking at well HK4256U, which is located on the tunnel pathway undergoes severe drawdowns due to heavy inleakage into the tunnel. However, the network model over-estimates the groundwater head indicating it is slow in understanding the impacted GWL's.

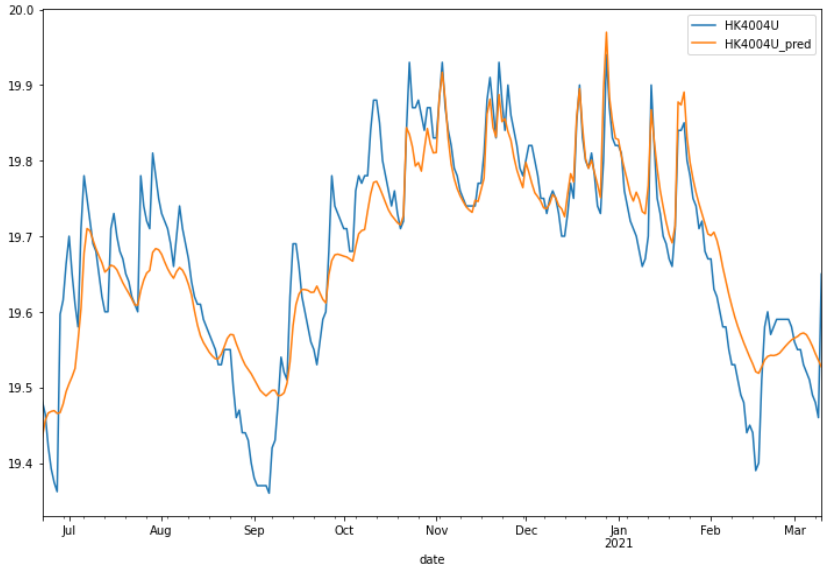


Figure 4.17: Predictions on HK4004U groundwater heads with historical GWL's, meteorological, infiltration and leakage data

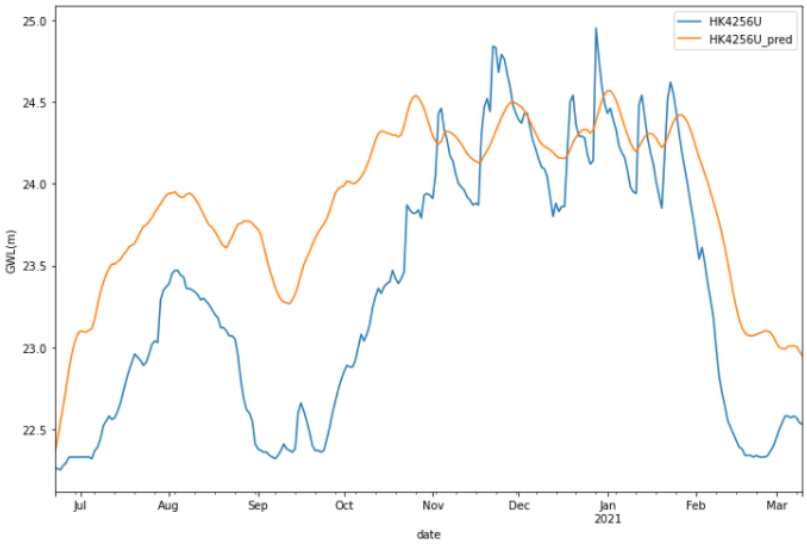


Figure 4.18: Predictions on HK4256U groundwater heads with historical GWL's, meteorological, infiltration and leakage data

The same evaluation technique is adopted to assess the fit, i.e. the correlations depicted between the observed and predicted groundwater head for HK4004U and HK4256U are seen in fig 4.19 and 4.20. From the correlations seen below, it is clear that there is a linear relationship indicating the model performs well with historic GWL, infiltration and leakage data as additional inputs. The R^2 score using all the modelling parameters for both wells significantly increase when compared to using only meteorological factors and, adding infiltration along with leakage data. R^2 scores of upto 0.81 for HK4004U and 0.73 for HK4256U are achieved. Furthermore, errors such as MAE and RMSE also decreased showing representative model performance for both the wells as seen in table 4.2. However, this is expected on adding historic GWL as an input feature.

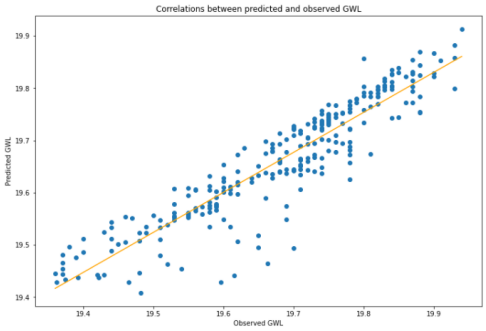


Figure 4.19: Correlation between observed & predicted GWL for HK4004U

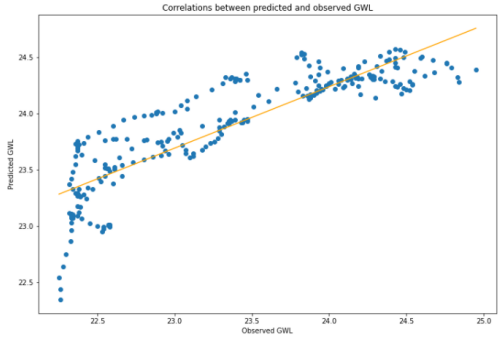


Figure 4.20: Correlation between observed & predicted GWL for HK4256U

Likewise an assessment on the input features average gradients for the two wells are seen in fig 4.21 and fig 4.22. We can see the impacted gradients contributing the most to the prediction of groundwater level one day ahead are the historic groundwater

levels, precipitation and, followed by relative humidity and evapotranspiration. The hydrological parameters such as infiltration and leakage data do not contribute to the predictions as expected, with most of the data being null during the training period.

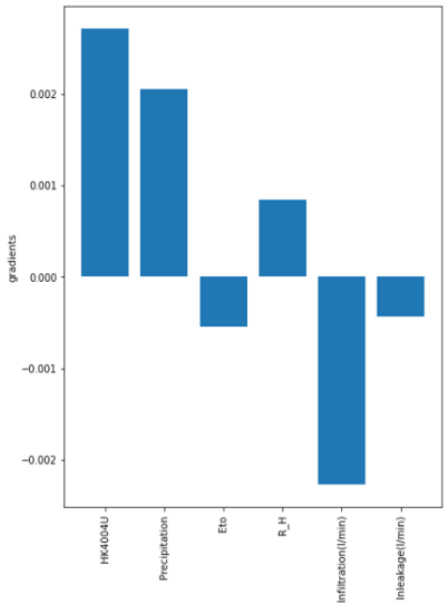


Figure 4.21: Impacting average gradients for HK4004U

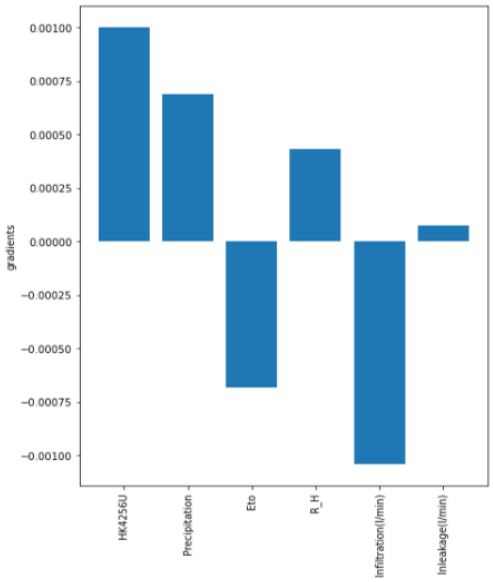


Figure 4.22: Impacted average gradient for HK4256U

4.2.4 Error Metrics on groundwater predictions for different scenarios

Quantifying the groundwater predictions on the observed groundwater heads statistically, opens for discussions on the accuracy of the model performance. The error metrics on the predictions are listed in table 4.2. Investigating the model, we see that the mean absolute error is less than the root mean squared error for all cases which indicates the model is correct in terms of evaluating the errors for all the cases listed. However, it should be noted that these scores are negatively oriented ranging from 0 to infinity. This means that values closer to 0 for MSE and MAE are preferred and indicate better model performance.

Features used for GWL predictions	Well_ID	MSE	RMSE	MAE	R^2 score
Only meteorological input	HK4004U	0.436	0.66	0.510	0.68
	HK4256U	1.40	1.18	1.08	0.36
Meteorological data + Hydrological data(infiltration & leakage)	HK4004U	0.65	0.806	0.668	0.48
	HK4256U	0.534	0.730	0.62	0.37
Historic GWL + Meteorological data + Hydrological data(infiltration & leakage)	HK4004U	0.256	0.50	0.381	0.82
	HK4256U	0.273	0.522	0.342	0.74

Table 4.2: Error metric values for different features based on the predictions. (MAE) - Mean absolute error, (MSE) - Mean squared error, (RMSE) - Root mean squared error, (R^2 score) - Coefficient of determination

Moreover, the results improved considering any error metric value when historic groundwater levels and hydrological parameters are added to the model. On comparing the three cases as mentioned in table 4.2 in terms of their metric performance we see an overall lower error values in the third case. This implies that the average error magnitudes are reduced when historic GWL and hydrological input are additionally fed to the model. This shows that the network understands the target in the model much better and elevates the confidence in prediction values Also, an R^2 score of 0.82 in HK4004U shows promising predictions on the GWL signifying that more than 80% of the variation in predictions are explained by the input features. While a comparitively lower R^2 score of 0.74 is seen in HK4256U implying more than 70% variations in predictions are explained by the input features.

4.2.5 Limitations of sequence modelling

The results obtained for GWL time series using LSTM neural networks have a series of limitations when implementing the model. Beginning with data handling, many sequences of missing values in daily time steps were discarded. However, short missing time periods were interpolated using linear interpolation to maintain a standard time sequence. It is known that neural networks train better with longer datasets wunsch2020groundwater. In this study, with data availability is limited to 2.5 years, the model could perform much better with longer datasets. Trivial correlations between the input features and the targets are not representative in terms of linearity. However, a rolling sum function is implemented to examine the response times between two variables, i.e. GWL heads and precipitation. There is no improvement in the correlation results after a certain rolling sum time period implying a need for improvement.

Deep learning networks are highly influenced by trial and error, and the model hyperparameters have to be made consistent at a certain point to show comparable results in this thesis zhang2018developing. As otherwise, this is a continuous learning process that leads to several set of results. A comparison study with other deep learning models is not in the scope of this thesis due to time constrain.

4.3 Groundwater Forecasting

Forecasting is basically making predictions with lack of knowledge on the outcomes in the future with regards to time. The idea behind forecasting is to have the ability to make strategical decisions on relevant prior information for the problem tackled. In this section, we forecast the groundwater level to check if there are any critical impacts using LSTM neural networks following simple linear regression principle.

4.3.1 Sequential groundwater forecasting for different training periods

The two wells selected from table 3.3 are used for forecasting different training periods and are plotted against unseen GWL data once the model is compiled. Considering the results from the error metrics evaluation of groundwater predictions, the network uses all the inputs, i.e. GWL heads, meteorological factors, infiltration and leakage data for forecasting GWL's. The same computational conditions are used for forecasting, see table 4.1 to maintain a fixed model for the project dataset. The splitting arrangement for different training periods used are seen in fig 4.23, i.e, 12 months, 18 months and 27 months of data are evaluated against unseen GWL data.

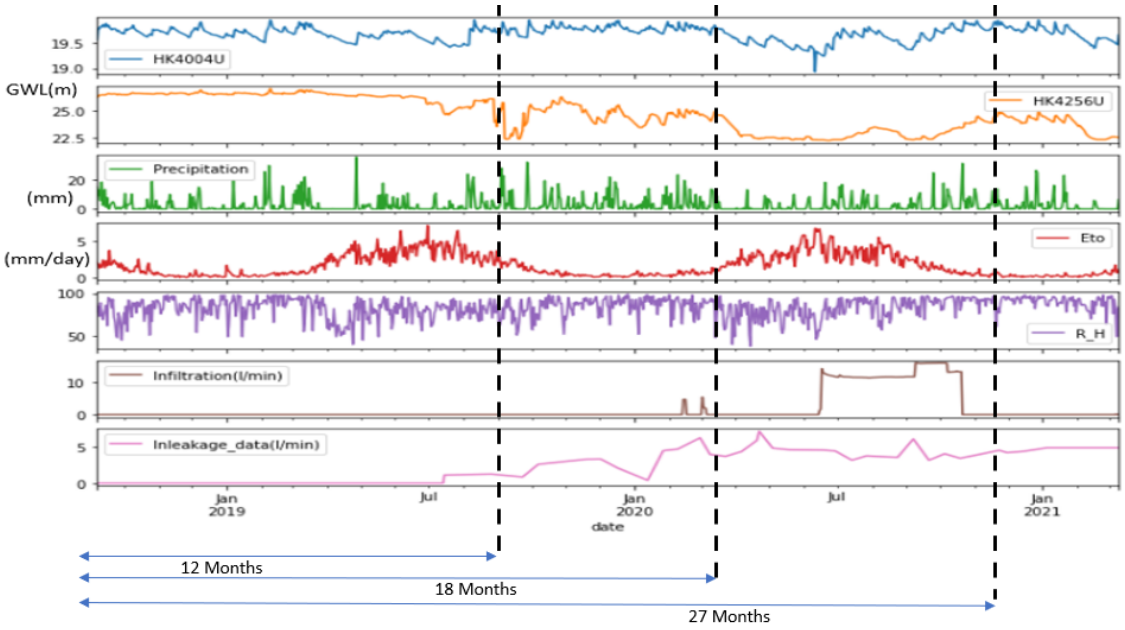


Figure 4.23: Splitting arrangement for forecasting different training periods of time series data used for modelling

To achieve reliable results, it is necessary to validate the neural network (LSTM) model with the training dataset. The model is validated against the last 10% of the dataset before forecasting.

4.3.1.1 Forecasting HK4004U groundwater levels

For all the training periods, forecasts of 7, 15, 30 and 60 days are plotted against unseen data .

1. Training period of 12 months

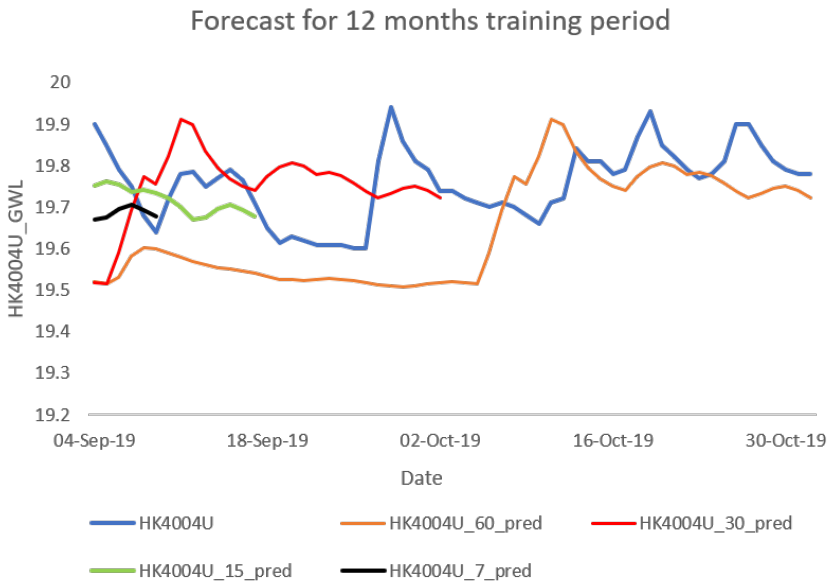


Figure 4.24: Forecasted series of days for 12 months vs unseen GWL for HK4004U

2. Training period of 18 months

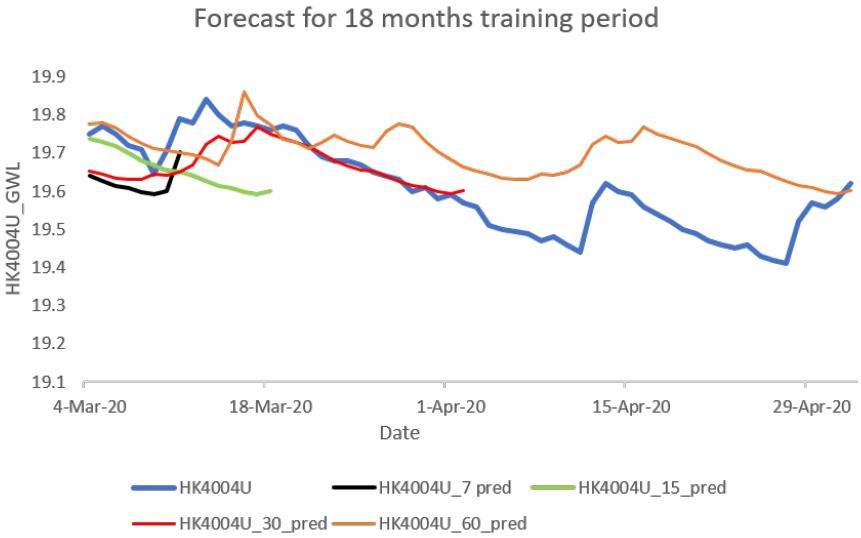


Figure 4.25: Forecasted series of days for 18 months vs unseen GWL for HK4004U

3. Training period of 27 months

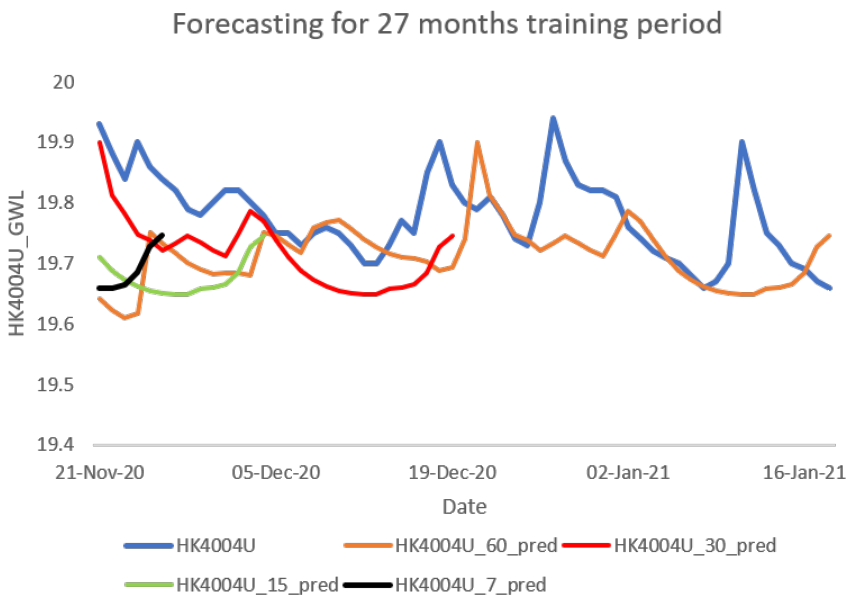


Figure 4.26: Forecasted series of days for 27 months vs unseen GWL for HK4004U

The forecasted results are plotted against unseen groundwater heads as seen above in fig 4.24, fig 4.25 and fig 4.26 Also, we see no real critical drop in the groundwater levels as the forecasted groundwater values vary in few decimeters with unseen data. There is no significant improvement in the different training lengths opted for forecasting. However, a longer training period should show higher accuracy in forecasting and it is seen that 30 days of forecast show good forecasting accuracy for 18 and 27 months of training the data.

4.3.1.2 Forecasting HK4256U groundwater levels

To maintain uniformity for comparison, the same set of days are used for forecasting the GWL's and are plotted against unseen GWL's data as seen in fig 4.27, fig 4.28 and fig 4.29.

1. Training period of 12 months

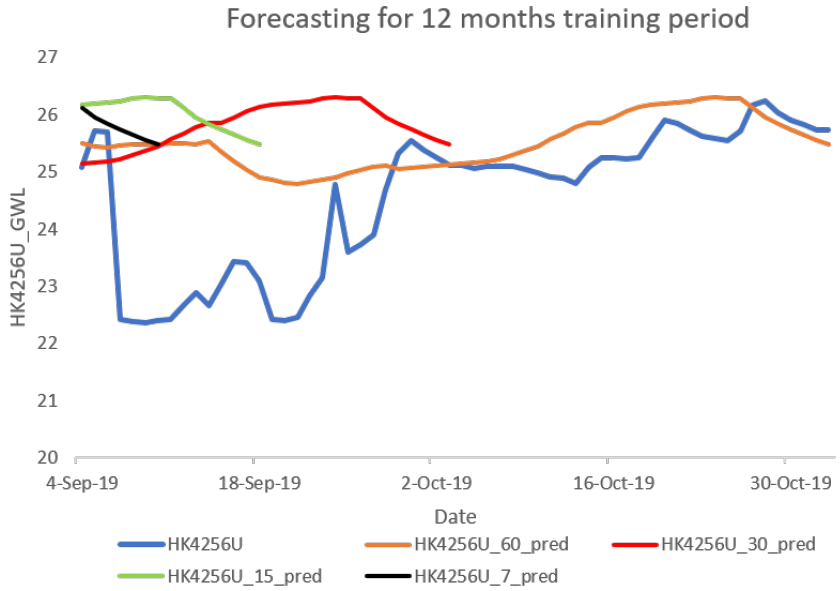


Figure 4.27: Forecasted series of days for 12 months vs unseen GWL for HK4256U

2. Training period of 18 months

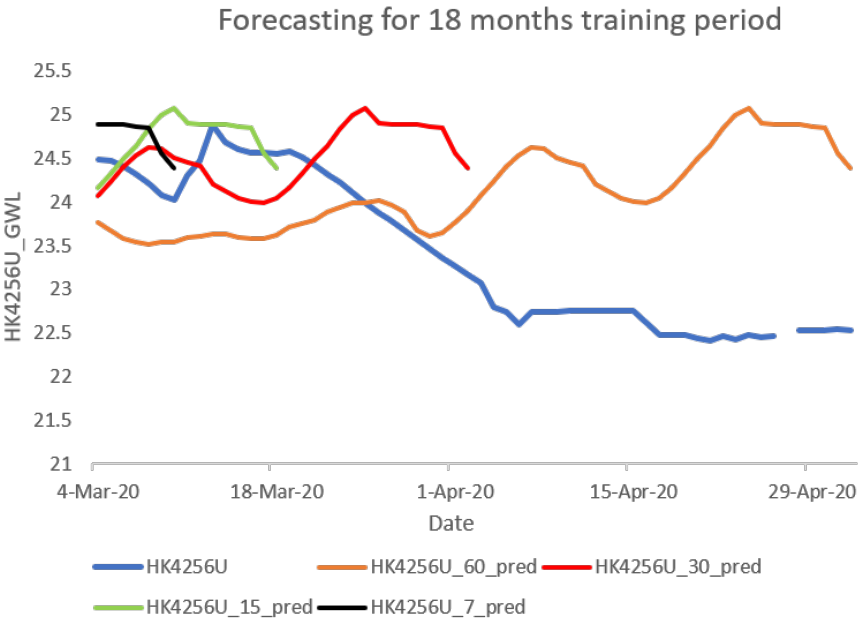


Figure 4.28: Forecasted series of days for 18 months vs unseen GWL for HK4256U

3. Training period of 27 months

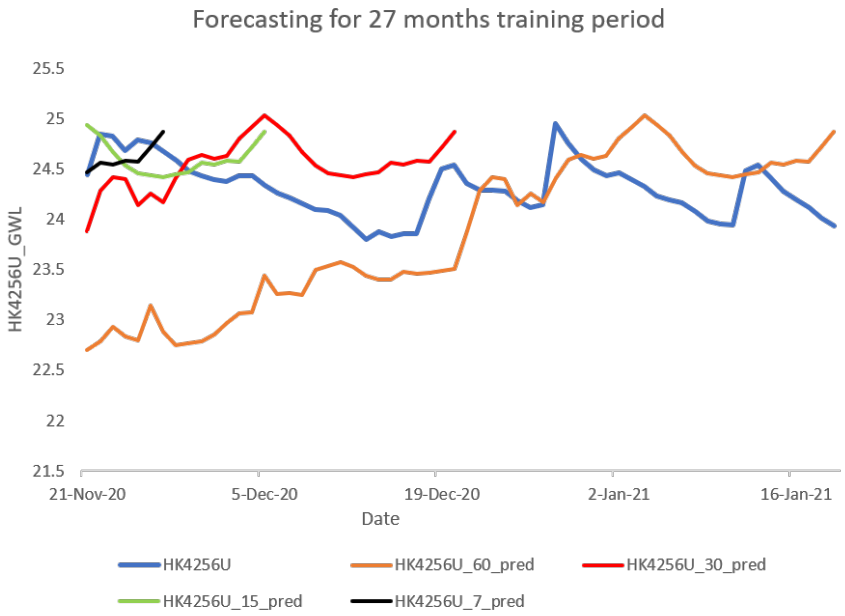


Figure 4.29: Forecasted series of days for 27 months vs unseen GWL for HK4256U

Considering this well being in the critical location, we see huge differences in GWL's over 12 months training period forecasting, see fig 4.27. This clearly shows an impact in the groundwater around the construction of the tunnel and the model performs poor in estimating the patterns observed in unseen data. However, a training period of 18 months reveals promising results for short forecasting duration while it does not account for a drawdown observed in unseen data after the first 30 days, see fig 4.28. Finally, after 27 months of data trained , we see a fitting pattern in the forecast for short periods of 7 days and 15 days varying a few decimeters in GWL's indicating promising results from fig 4.29.

4.3.2 Forecasting correlations

Using correlations to assess the forecasted GWL values with unseen GWL values from the model gives the accuracy of the model performance. Moreover, an interpretation of the time series plots gives the forecasted trend of the GWL while considering this approach is more visual oriented. To counteract this, we use the Pearson correlation coefficient to examine the linearity of relationship between the forecasted and observed GWL for different training periods and forecasted ranges. However, the Spearman correlation coefficient is not used in evaluating the correlations due to its monotonic relationship between variables and also it does not consider the normality of raw data. A bar chart of the correlations for different training periods and forecasted days ahead are plotted for HK4004U and HK4256U, see fig 4.30 and fig 4.31.

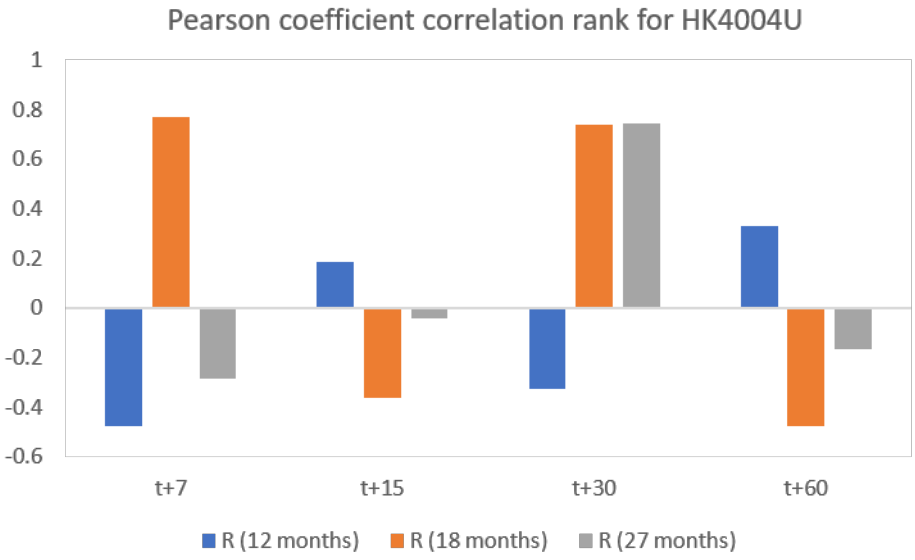


Figure 4.30: Correlations of HK4004U for different training periods and forecasted days

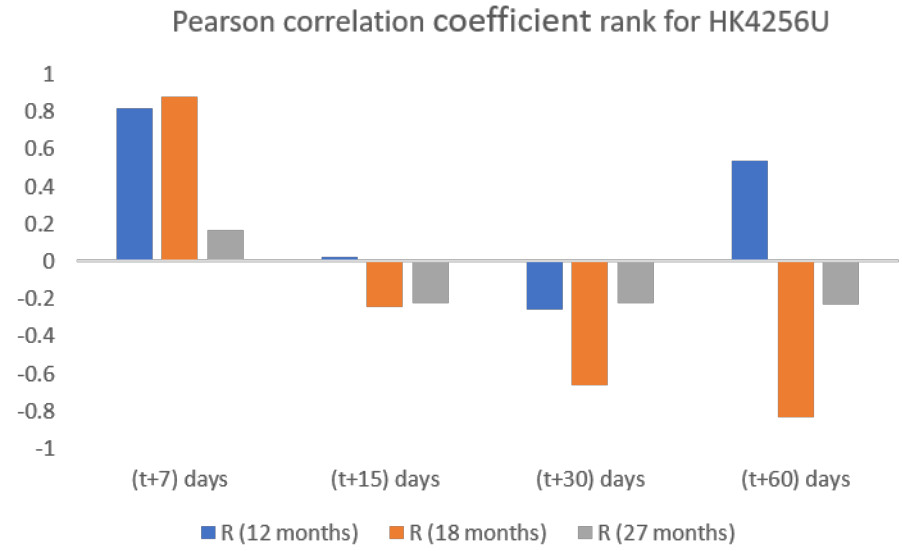


Figure 4.31: Correlations of HK4256U for different training periods and forecasted days

The correlations observed from HK4004U and HK4256U groundwater heads show interesting results in terms of linearity between the forecasted and unseen data (observed data) plotted in Excel. The idea is to find the optimum range of forecasting days influenced by different training periods. From fig 4.30, there is a strong correlation for the 30 days forecast for training periods 18 months and 27 months and a medium correlation for 12 months of training. Moreover, the correlations observed for different days of forecast do not indicate strong impacts on groundwater levels

due to minimal variations in corresponding GWL heads upto 1-3 decimeters. A negative correlation does not indicate weak performance instead it gives the overall linearity of the values compared. However, this can vary with the range of magnitude of values correlated for any specific problem. Additionally, a certain positive impact could be achieved from the location of the infiltration well with regards to the location of HK4004U to keep the GWL's non-critical.

The correlations observed for HK4256U indicate the best forecasting range for a well in a critical location is 7 days ahead. Moreover, it should be noted that for 7 days of forecasting, a higher correlation for 27 months of training should have been achieved in HK2456U instead of a lower correlation. Therefore, this gives insights on the model not being able to capture critical drop in GWL's in meters for such long training periods. As the forecasting days increase for different training periods, the correlations observed between forecasted and unseen data tend to become negative. However, this can be due to the fact that infiltration and inleakage data influences longer training periods complemented by non-natural seasonality trends observed in HK4256U. However, it should be kept in mind that inleakage and infiltartion data is not available on daily basis during most of the tunnel construction.

5
Conclusion

In conclusion, the LSTM neural network model is used to assess the impacts on groundwater caused by pertubed sources and proved to be one of many methods to tackle the problem based on literature studies. However, in such models there is no physical relation between the parameters and it is completely data-driven executed using many deep learning libraries like tensorflow, keras and scikit-learn (Wunsch, Liesch, and Broda 2020). Supervised learning models to evaluate the impacts for the case study presented have a high degree of uncertainty in terms of data preprocessing, hyperparameter tuning, network architecture and the computational aspect. The influence of human trial and error is applicable in this study signifying LSTM neural networks is a continuous learning process and has to be constrained at a certain point based on different evaluation techniques (Zhang et al. 2018).
The results from predicting GWL by only adding meterological data to assess if the impact is non natural seems to result in an overestimation, meaning in some cases the model struggles to understand complex pattern for short training periods. However, this shows that the impact is from a pertubed source which is not capturing complex variations in GWL.
It can be further concluded that the addition of historic GWL along with meterological, infiltration and leakage data inputs showed good fit to the observed GWL with high accuracy and low error metrics upon evaluation. This gives insights on the GWL being impacted non-naturally by looking at the goodness of the fit. Response time estimations for 15 days showed highest correlation when just using precipitation with GWL and the forecasted GWL's correlates to the observed GWL's on these terms based on the results obtained. Moreover to test if there is a critical impact on the GWL, forecasting the GWL for shorter periods and evaluating it with unseen data showed no impacts from the natural parameters considered for modelling. In addition, it should be noted that the model performance can be approached more effectively if additional data is available, testing for different training times and thorough optimization of hyperparameters i.e., assessed by the learning curve of the model. However, using all the available resources effectively it is concluded that variations of a few decimeters in the GWL is seen in HK4004U indicating there could be a possible impact from a non-natural source. Moreover, considering the performance of HK4256U it is seen that the impact is non-natural especially due to the location of the well being in the line of tunnel construction and it is difficult in achieving accurate forecasts of GWL's for longer time periods.

Bibliography

Bakker, Mark and Frans Schaars (2019). "Solving groundwater flow problems with time series analysis: you may not even need another model". In: *Groundwater* 57.6, pp. 826–833.

C.W.Fetter (2014). *Applied hydrogeology*. Pearson custom library. Harlow, Essex : Pearson. ISBN: 9781292022901.

Cerliani, M (2020). *Feature importance with time series and recurrent neural network*. `https://towardsdatascience.com/feature-importance-with-time-series-and-recurrent-neural-network-27346d500b9c`. accessed: 12.05.2021.

Chen, Chong et al. (2020). *A Novel Deep Learning Algorithm for Groundwater Level Prediction based on Spatiotemporal Attention Mechanism. Doi : 10.21203/rs.3.rs-59191/v1*.

Daliakopoulos, Ioannis N, Paulin Coulibaly, and Ioannis K Tsanis (2005). "Groundwater level forecasting using artificial neural networks". In: *Journal of hydrology* 309.1-4, pp. 229–240.

De Caro, Mattia, Giovanni B Crosta, and Alberto Previati (2020). "Modelling the interference of underground structures with groundwater flow and remedial solutions in Milan". In: *Engineering Geology* 272, p. 105652.

Haaf, Ezra (2020). *Towards prediction in ungauged aquifers–methods for comparative regional analysis*. `http://hdl.handle.net/2077/63694`.

Hyndman, Rob J and George Athanasopoulos (2018). *Forecasting: principles and practice*. OTexts.

Kenda, Klemen et al. (2018). "Groundwater modeling with machine learning techniques: Ljubljana polje aquifer". In: *Multidisciplinary Digital Publishing Institute Proceedings*. Vol. 2. 11, p. 697.

Keras (2021). *The Python deep learning API*. `https://keras.io/`.

Khedri, Akbar, Nasrollah Kalantari, and Meysam Vadiati (2020). "Comparison study of artificial intelligence method for short term groundwater level prediction in the northeast Gachsaran unconfined aquifer". In: *Water Supply* 20.3, pp. 909–921.

Kittridge, Mike (2021). *The Python deep learning API*. `https://pypi.org/project/ETo/`. accessed: 15.05.2021.

Lelli, Francesco (2021). *Neural Networks: The Basics and a Collection of YouTube Videos*. `https://francescolelli.info/tutorial/neural-networks-a-collection-of-youtube-videos-for-learning-the-basics/`. accessed: 10.04.2021.

Lithén, Johanna. and Maria. Wadsten (2016). *PM Hydrogeologi berg Underlagsdokumenttill PM Hydrogeologi, ansökan om tillstånd till vattenverksamhet*. Tech. rep. TRV 2016/3151. Trafikverket.

Pandas (2021). *Python Data Analysis Library*. `https://pandas.pydata.org/`. accessed: 03.03.2021.

Rajaee, Taher, Hadi Ebrahimi, and Vahid Nourani (2019). "A review of the artificial intelligence methods in groundwater level modeling". In: *Journal of hydrology* 572, pp. 336–351.

Ramboll (2014). *VÄSTLÄNKEN SERVICETUNNEL-PROFIL*.

Rumelhart, David E, Geoffrey E Hinton, and Ronald J Williams (1986). "Learning representations by back-propagating errors". In: *nature* 323.6088, pp. 533–536.

Sagheer, Alaa and Mostafa Kotb (2019). "Time series forecasting of petroleum production using deep LSTM recurrent networks". In: *Neurocomputing* 323, pp. 203–213.

ScikitLearn (2021). *Standard Scaler - ScikitLearn*. `https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html`. accessed: 15.03.2021.

Sharma, S (2021). *Epoch vs Batch Size vs Iterations*. `https://towardsdatascience.com/epoch-vs-iterations-vs-batch-size-4dfb9c7ce9c9`. accessed: 03.05.2021.

SMHI (2021). *Ladda ner meteorologiska observationer*. `https://www.smhi.se/data/meteorologi/ladda-ner-meteorologiska-observationer#param=airtemperatureInstant,stations=all`. accessed: 03.03.2021.

Stejmar Eklund, H (2002). *Hydrogeologiska typmiljöer: Verktyg för bedömning av grundvattenkvalitet, identifiering av grundvattenförekomster samt underlag för riskhantering längs vägar.(Hydrogeological type settings: A tool to assess groundwater quality, to identify groundwater recourses and for risk-decision analysis along roads.) Licentiate thesis Publ A 101. Department of Geology, Chalmers University of Technology, Gothenburg*.

Sundell, Jonas et al. (2019). "Risk mapping of groundwater-drawdown-induced land subsidence in heterogeneous soils on large areas". In: *Risk Analysis* 39.1, pp. 105–124.

Sundkvist, Ulf (2016). *PM Hydrogeologiska beräkningar*. Tech. rep. TRV 2016/3 (44),1–44. Trafikverket.

Sundkvist, Ulf. and Thomas. Wallroth (2016). *Ansökan om tillstånd enligt miljöbalkenför anläggandet av Västlänken och Olskroken planskildhet Göteborgs Stad,Mölndals stad, Västra Götalands län. PM Hydrogeologi*. Tech. rep. TRV 2016/3151. Trafikverket.

Vu, MT et al. (2021). "Reconstruction of missing groundwater level data by using Long Short-Term Memory (LSTM) deep neural network". In: *Journal of Hydrology* 597, p. 125776.

Wallroth, T. et al. (2016). *Ansökan om tillstånd enligt miljöbalken för anläggandet avVästlänken och Olskroken planskildhet, Teknisk Beskrivning*. Tech. rep. (2016/3151). Trafikverket.

Wikipedia (2021). *Hyperparameter optimization — Wikipedia, The Free Encyclopedia*. accessed: 10.05.2021.

Wunsch, Andreas, Tanja Liesch, and Stefan Broda (2020). "Groundwater Level Forecasting with Artificial Neural Networks: A Comparison of LSTM, CNN and NARX". In: *Hydrology and Earth System Sciences Discussions* 2020, pp. 1–23.

Zhang, Jianfeng et al. (2018). "Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas". In: *Journal of hydrology* 561, pp. 918–929.

A

Appendix 1

The following link is attached to my GitHub page concerning the codes implemented to obtain results for this thesis work.
`https://github.com/raja-ui/Lstm-Gwl-Model`

DEPARTMENT OF SOME SUBJECT OR TECHNOLOGY
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY