# DEVELOPMENT OF LONG SHORT-TERM MEMORY MODEL FOR PREDICTION OF WATER TABLE DEPTH IN UNITED ARAB EMIRATES

Pavitra Kumar[1], Ali Najah Ahmed[2], Mohsen Sherif [3,4], Ahmed Sefelnasr[4], Ahmed Elshafie[1,4,*]

[1]Department of Civil Engineering, Faculty of Engineering, Universiti Malaya (UM), 50603 Kuala Lumpur, Malaysia; pavitrakumar27@gmail.com,
[2]Institute of Energy Infrastructure (IEI), Department of Civil Engineering, College of Engineering, Universiti Tenaga Nasional (UNITEN), 43000 Selangor, Malaysia; mahfoodh@uniten.edu.my
[3]Civil and Environmental Eng. Dept., College of Engineering, United Arab Emirates University, Al Ain P.O. Box 15551, United Arab Emirates; MSherif@uaeu.ac.ae
[4]National Water and Energy Center, United Arab Emirates University, Al Ain P.O. Box 15551, United Arab Emirates; ahmed.sefelnasr@uaeu.ac.ae, elshafie@um.edu.my
[*] Corresponding Author: elshafie@um.edu.my

**KEYWORDS**
Declining Groundwater, Less Precipitation, Monthly Data, LSTM Model, Time-step lag

**ABSTRACT**
Water table depth is declining in most part of the world, especially in those countries which have high temperature almost throughout the year and receive very less precipitation throughout the year. Due to increasing population, intensive agricultural and industrial practices, the demand of freshwater is increasing and is predicted to increase in upcoming years. The countries which receive less rainfall throughout the year have limited groundwater recharge, resulting in declining of water table. United Arab Emirates belong to this category of countries where there is high temperature almost throughout the year and receives very less rainfall (less than 200 mm annually). Modeling groundwater in such an arid climate is of serious concern. This paper proposes LSTM models for prediction of water table depth at six different wells in different parts of United Arab Emirates. Data obtained for this study comprises of times series monthly water table depth data in meters from ground level from six different wells. Analysis of the data showed the drastic decline of water table depth between 1977 and 2011. These data were used to generate the input and target variables by adding three time-step lags in the given data. The time-step lag data was used as input to predict the current water table depth. In other words, the water table depth data of current target month was predicted using the previous three months water table depth data as input. Training of LSTM models was carried forward using TensorFlow libraries in python programming language. The trained models provided good accuracy in testing dataset. The training $R^2$ values of all the six models were more than 0.96 and the testing $R^2$ values of all the six models were more than 0.91.

## 1. INTRODUCTION
Ground water is a very crucial source of water supply for humankind (Hanoon et al., 2021). It fulfills one-third of world's water demand (Kombo et al., 2020; Sapitang et al., 2021). Apart from drinking, it is used for lot many purposes, such as, household, agriculture, industries etc. Ground water is being exploited and overused since many years. The demand of groundwater is predicted to rise in upcoming years. However, the recent researches report the decline of ground water level in various part of the world (Kombo et al., 2020), thus requiring awareness for management of groundwater level along with its quality (Afan et al., 2021; Hanoon et al., 2022). The decline of this level is either because of overuse or exploitation of the groundwater or because of the limited groundwater recharge. Overpopulation and intensive use of ground water in agriculture, household and industries have resulted in its

overuse (Abuelgasim and Elkamali, 2019; Osman et al., 2022). Reduced precipitation and climate change are considered as the main cause for limited ground water recharge (de Graaf et al., 2017; Haas and Birk, 2017; Kombo et al., 2020), thus altering the groundwater balance which results in decline of groundwater level.

United Arab Emirates (UAE) belongs to the category of those countries where there is less precipitation, elevated temperature resulting in high evaporation rate. Water management is a top priority concern for UAE government (Rizk and Alsharhan, 2003). Apart from the groundwater source of water supply, UAE has an alternate source of water supply i.e. desalinated water supply (Al-Ruzouq et al., 2019). Water from the groundwater source is mainly used for agricultural activities and desalinated water is used for non-agricultural activities (UKEssays, 2018). Most of the groundwater in entire country is saline, which is mostly used for irrigating date palms due to their salt-tolerance property. Hence, for drinking purpose and other non-agricultural purposes desalinated water is the only option. Artificial recharge of ground water using desalinated water has begun on large scale in some part of the country since 2008. Artificial recharge is an attempt to bring back the declining water table level and also to reduce the dependency on desalinated water, which may not be available in case of any emergency in the country or in case of oil spill in the sea. However, bringing back the declining water level is a tough task for UAE, as most part of the country has seen great decline in water table in recent years (Abuelgasim and Elkamali, 2019). According to Sherif et al. (2021), groundwater aquifers have seen a drastic decline in water storage. Quaternary aquifer was recorded to have fresh groundwater of 238 km$^3$ in 1969 and was again recorded to have just 10 km$^3$ in 2015. Hence, a great step towards management and restoration of groundwater is urgently needed for UAE. Lot of research work is needed to understand the pattern of the water table level in order to help in its recovery. This paper puts forward an attempt to understand the pattern of water table level and predict the future water table level.

The main objective of this study is to investigate the potential of the deep learning modeling approach in developing a model that could be able to predict the groundwater table. This paper proposes a Long Short-Term Memory (LSTM) model for prediction of water table depth one month in advance at six different wells located in different parts of UAE.

## 2. MATERIALS AND METHODOLOGY

UAE is situated in the southeastern part of Arabian Peninsula (figure 1) between the latitude 22˚50′- 26˚4′ N and longitudes 51˚5′ - 56˚25′ E with total mainland surface area of 83,600 km$^2$. It has more than 75% of its area covered with desert (Sherif et al., 2014). The temperature of UAE varies from 10˚C in winter season to 48˚C in summer season. UAE is considered as one of the driest places on earth with only less than 200 mm average annual rainfall (Abuelgasim and Elkamali, 2019). Such low precipitation throughout the year in entire country is primary reason for limited groundwater recharge in UAE.

For LSTM model development, data from six different water table measuring wells were collected, which are situated at different stations in UAE. Those stations include Madam, 2 wells at Dhaid (Al Naseem, Meliha), Jahili, Idhen, and Hamdah (as presented in table 1 and figure 1 (plotted on MATLAB)). Data consisted of monthly water table depth, in meters below ground level, with different time period, as presented in table 1. The obtained data was pre-processed to remove the outliers to enable smooth learning of the models. Using these data, six different LSTM models were developed for prediction of water table depth at each selected station. Analyzing the data, the drastic fall of water table was observed in all the six wells. For well1 (station: Madam), water table has declined from 10.85 m in 1977 to 45.68 m in 2011. The decline of water table in other wells can be analyzed from the given opening and closing water table depth data in table 1, where opening data is the depth data at the beginning of the data time period and closing data is the depth data at the end of the data time period. Such drastic decline of the water table depth in such a short time frame in all wells is a very serious concern for the nation.

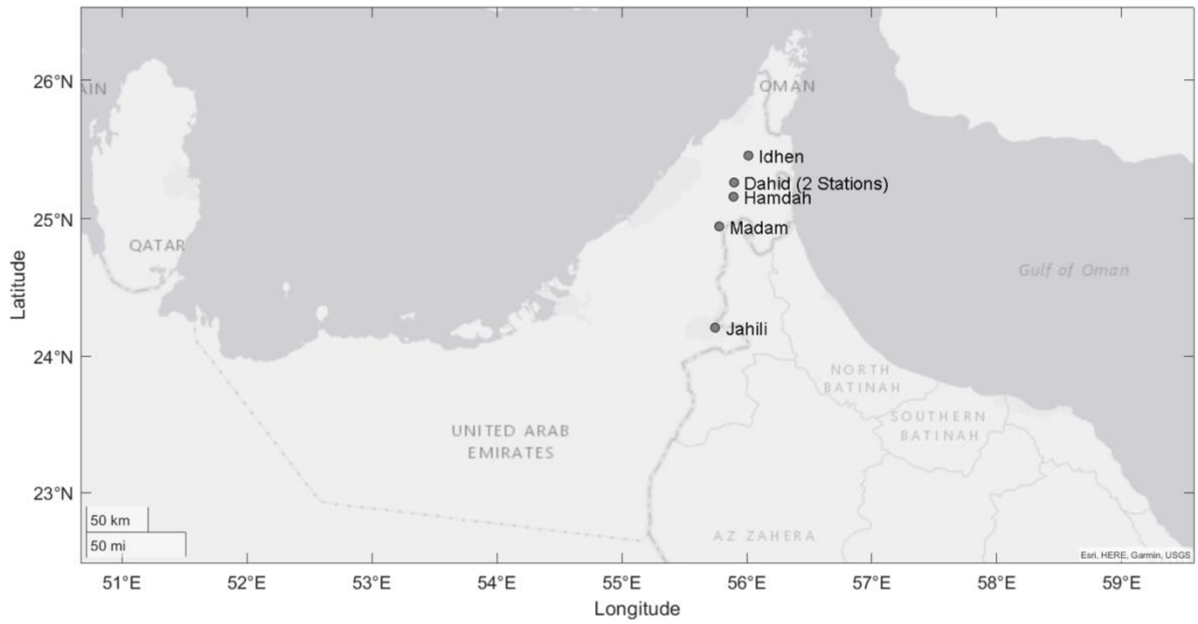


**Figure 1.** Location of all wells on UAE map

**Table 1.** Analysis of obtained data

| *Well* | *Station Name* | *Data time period* | *Opening water table depth (m)* | *Closing water table depth (m)* |
|---|---|---|---|---|
| ***Well1*** | Madam | 1977-2011 | 10.85 | 45.68 |
| ***Well2*** | Dhaid (Al Naseem) | 1983-2011 | 16.74 | 59.42 |
| ***Well3*** | Jahili | 1985-2011 | 31.67 | 45.26 |
| ***Well4*** | Idhen | 1979-2007 | 9.72 | 58.9 |
| ***Well5*** | Hamdah | 1983-2011 | 19.5 | 30.78 |
| ***Well6*** | Dhaid (Meliha) | 1983-2011 | 21.9 | 27.41 |

Using the data obtained for all the six wells, six different LSTM models were developed. LSTM model is a special variant of recurrent neural network model. It was designed to overcome the limitation of recurrent neural network, where it faces difficulty in learning long term dependency between the input and target variables (Chung and Shin, 2018; Zhang et al., 2017), due to vanishing and exploding gradient problem (Karim et al., 2018; Supreetha et al., 2020). LSTM network contain certain unique gates which is used to learn and memorize the long-term dependence between input and output presented in long sequence of data, which helps in achieving high-precision prediction (Chen, 2020). LSTM network is designed to remove or add certain piece of information from the sequence data through these gates (Le et al., 2019; Tian and Pan, 2015). Thus, enabling the LSTM network to learn useful piece of information or delete the obsolete piece of information through these gates, thus, improving the learning capability. Figure 2 represents the basic structure of LSTM model containing, input ($x_t$), output ($o_t$), input from previous time step ($h_{t-1}$, $c_{t-1}$) and output to next time stop ($h_t$, $c_t$). These inputs are processed using the equations (1-6) which uses weight matrices (W, U) and biases b to generate outputs to all output gates.

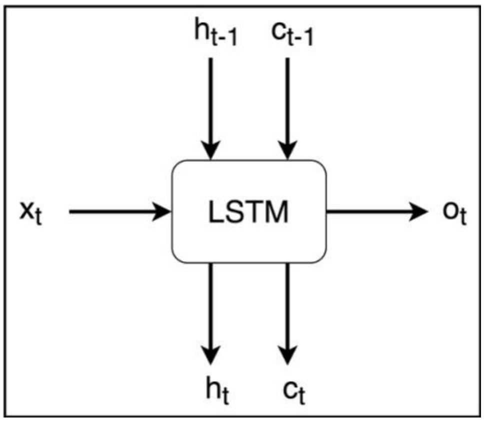


**Figure 2.** Basic structure of LSTM

$$f_t = sigmoid\left(W_f \times x_t + U_f \times h_{t-1} + b_f\right) \quad (1)$$
$$i_t = sigmoid(W_i \times x_t + U_i \times h_{t-1} + b_i) \quad (2)$$
$$o_t = sigmoid(W_o \times x_t + U_o \times h_{t-1} + b_o) \quad (3)$$
$$c'_t = tanh(W_c \times x_t + U_c \times h_{t-1} + b_c) \quad (4)$$
$$c_t = f_t \cdot c_{t-1} + i_t \cdot c'_t \quad (5)$$
$$h_t = o_t \cdot tanh(c_t) \quad (6)$$

Where: $f_t$ is forget gate; $i_t$ is input gate; $o_t$ is output gate; $c_t$ is cell state; $h_t$ is hidden state

Observing the unique characteristic of LSTM model, it was chosen for modeling the water table depth at different wells. Data obtained for modeling was only the monthly time series water table depth data. Three time-step lag data was generated using the monthly time series data. The previous three time-step data were used as the input data and the current data was used as the target data for the model. In other words, the target water table depth data of a particular month is predicted using the input of the previous three months water table depth data, as presented in table 2. A sample of 10 input and target records of water table depth data of well 1 is presented in table 2. This procedure was followed to generate input and target data for all the six wells. The generated input and target data were then fed to the LSTM model for training. Modeling process was carried forward using TensorFlow libraries in python programming language.

**Table 2.** Sample of input and target for well 1

| Input | | | Target |
|---|---|---|---|
| Depth(t-3) | Depth(t-2) | Depth(t-1) | Depth(t) |
| 10.81 | 10.78 | 10.68 | 10.85 |
| 10.78 | 10.68 | 10.85 | 10.84 |
| 10.68 | 10.85 | 10.84 | 10.75 |
| 10.85 | 10.84 | 10.75 | 10.73 |
| 10.84 | 10.75 | 10.73 | 10.7 |
| 10.75 | 10.73 | 10.7 | 10.8 |
| 10.73 | 10.7 | 10.8 | 11.08 |
| 10.7 | 10.8 | 11.08 | 11.19 |
| 10.8 | 11.08 | 11.19 | 11.26 |
| 11.08 | 11.19 | 11.26 | 11 |

Model performance was analyzed using several performance criteria, such as: training and testing $R^2$, training and testing mean square error (MSE) and maximum percentage error.

$R^2$ is computed as:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (7)$$

MSE us computed as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x-y)^2 \quad (8)$$

Maximum percentage error is computed as:

$$Max\ \%\ Error = max\left(\frac{|x-y|}{x} * 100\right) \quad (9)$$

where: x is target value, y is predicted value, n is number of data samples.

## 3. RESULTS AND DISCUSSION

Modeling process resulted in six different LSTM models trained on the generated input and target data for all the six wells. The architecture of the trained model consists of two LSTM layers with dropout layer after each LSTM layer and a dense layer at the end of sequential layers. Trained models provided good prediction accuracy based on above-mentioned performance criteria, as presented in table 3. The training $R^2$ value of all the models are greater than 0.96 and the testing $R^2$ value of all the models are greater than 0.91. in addition, it is obvious that the proposed LSTM model could achieved a prediction accuracy all wells with acceptable level of accuracy except for Well4. With careful investigation for Table 3, it could be depicted that the maximum error is ranged between 10% and 28% for all wells except Well4 the maximum error was 96%. It should be noticed here that only one model structure has been applied for all wells, and hence, it is expected the model performance accuracy might be relatively low as in Well4. However, this particular drawback in the accuracy achieved for Well4 could be improved by adapting the model internal parameters for this particular well to enhance the accuracy.

**Table 3.** Performance analysis of all the six models

| *Well* | *Well1* | *Well2* | *Well3* | *Well4* | *Well5* | *Well6* |
|---|---|---|---|---|---|---|
| ***Training $R^2$*** | 0.9881 | 0.9858 | 0.9775 | 0.9679 | 0.9833 | 0.9893 |
| ***Testing $R^2$*** | 0.9566 | 0.9905 | 0.9192 | 0.9881 | 0.9557 | 0.9787 |
| ***Training MSE*** | 2.2711 | 1.5339 | 0.8185 | 18.4891 | 0.1596 | 0.1564 |
| ***Testing MSE*** | 2.3679 | 2.0507 | 0.4358 | 14.7226 | 0.1964 | 0.5542 |
| ***Testing Max % Error*** | 28.47 | 25.16 | 11.27 | 95.96 | 12.39 | 10.34 |

Figure 3 (a-f) present the plot of all the performance criteria for wells (1 - 6), which include the regression plot for testing and training data superimposed in one plot, the percentage error plot with marked maximum percentage error and plot of target and predicted data for both training and testing sets. The plot of target and predicted testing data, presented in figure 3 (a-f), represents the prediction capability of the model. It represents that model has learned the pattern of variation of water table depth and is capable of predicting accurately the future water depth data.

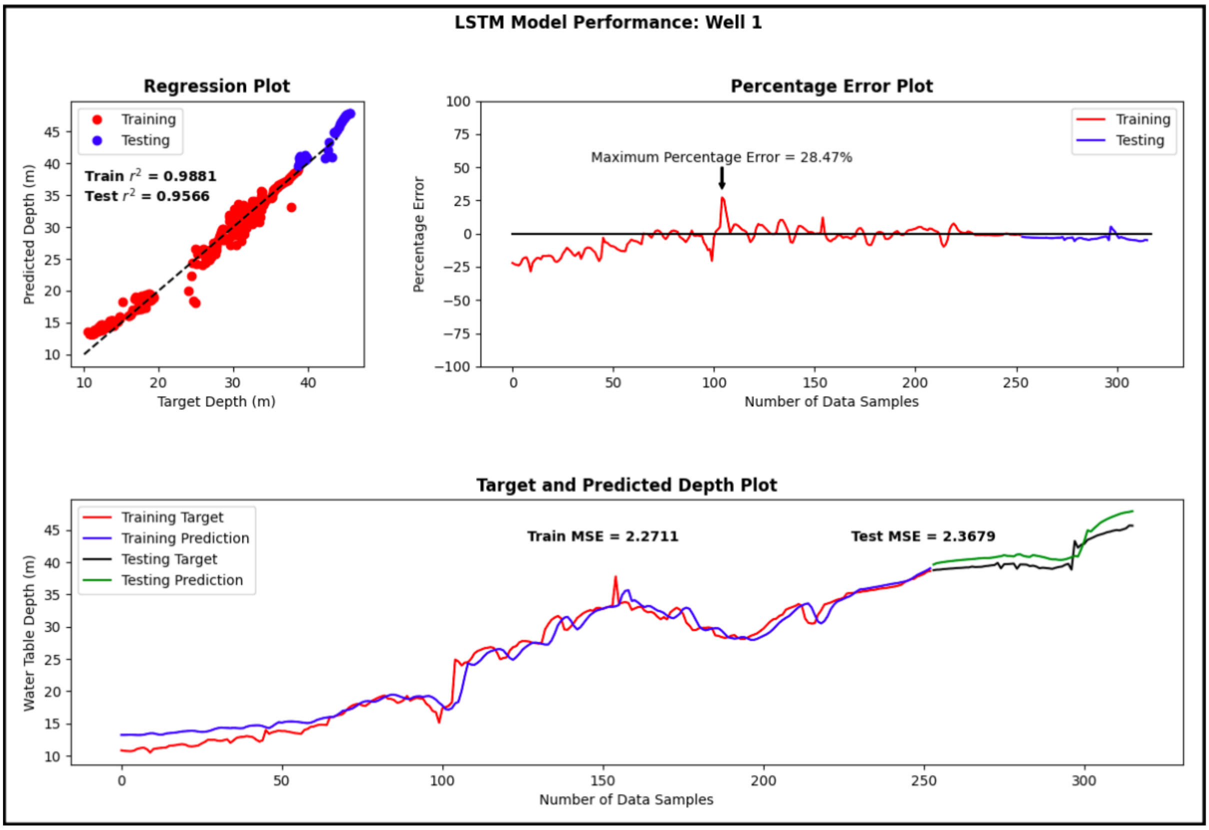


**Figure 3(a).** LSTM model performance plot for well 1

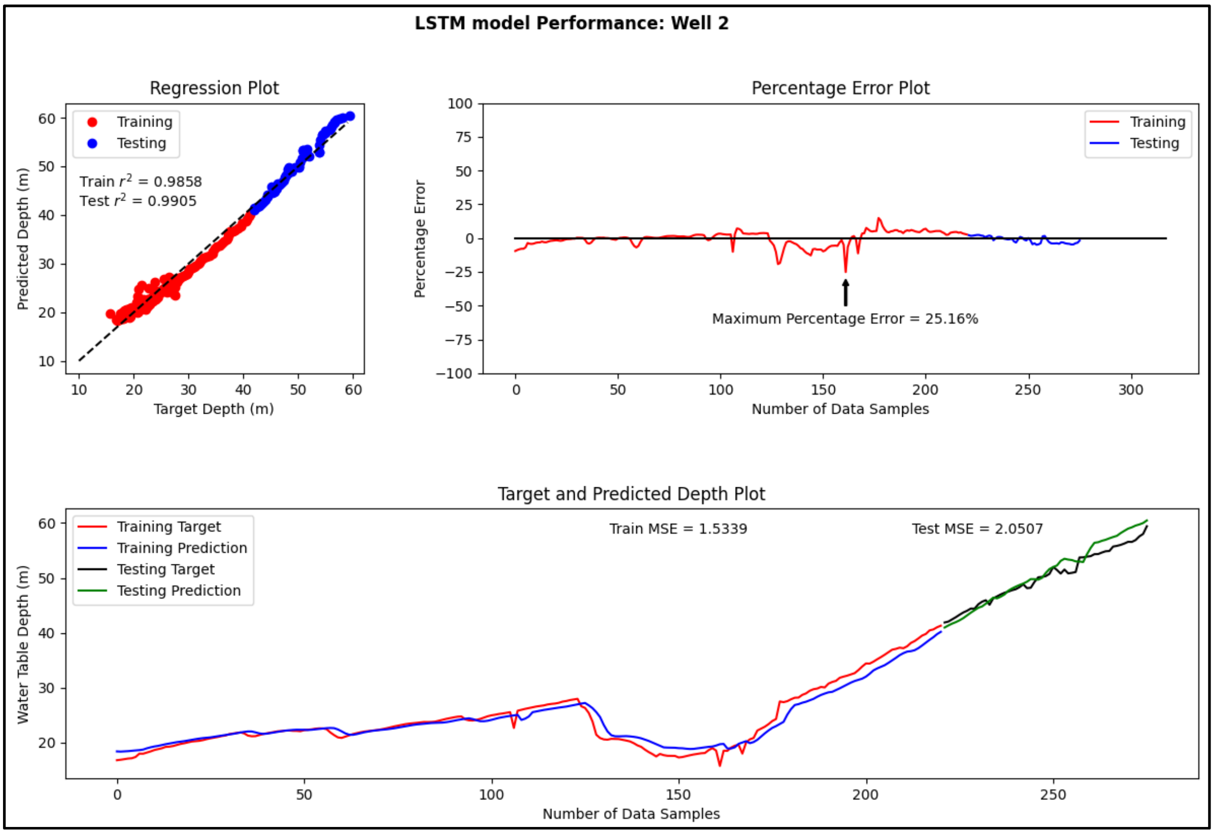


**Figure 3(b).** LSTM model performance plot for well 2

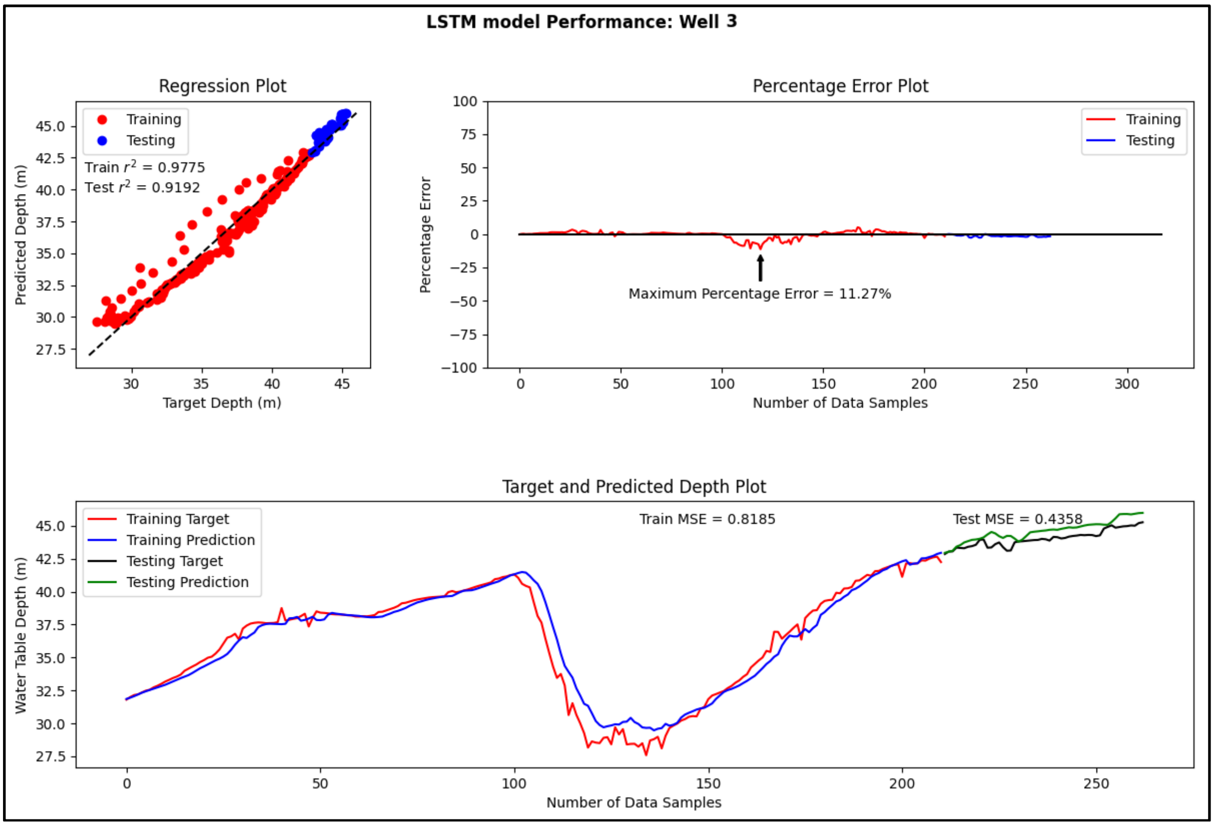


**Figure 3(c).** LSTM model performance plot for well 3

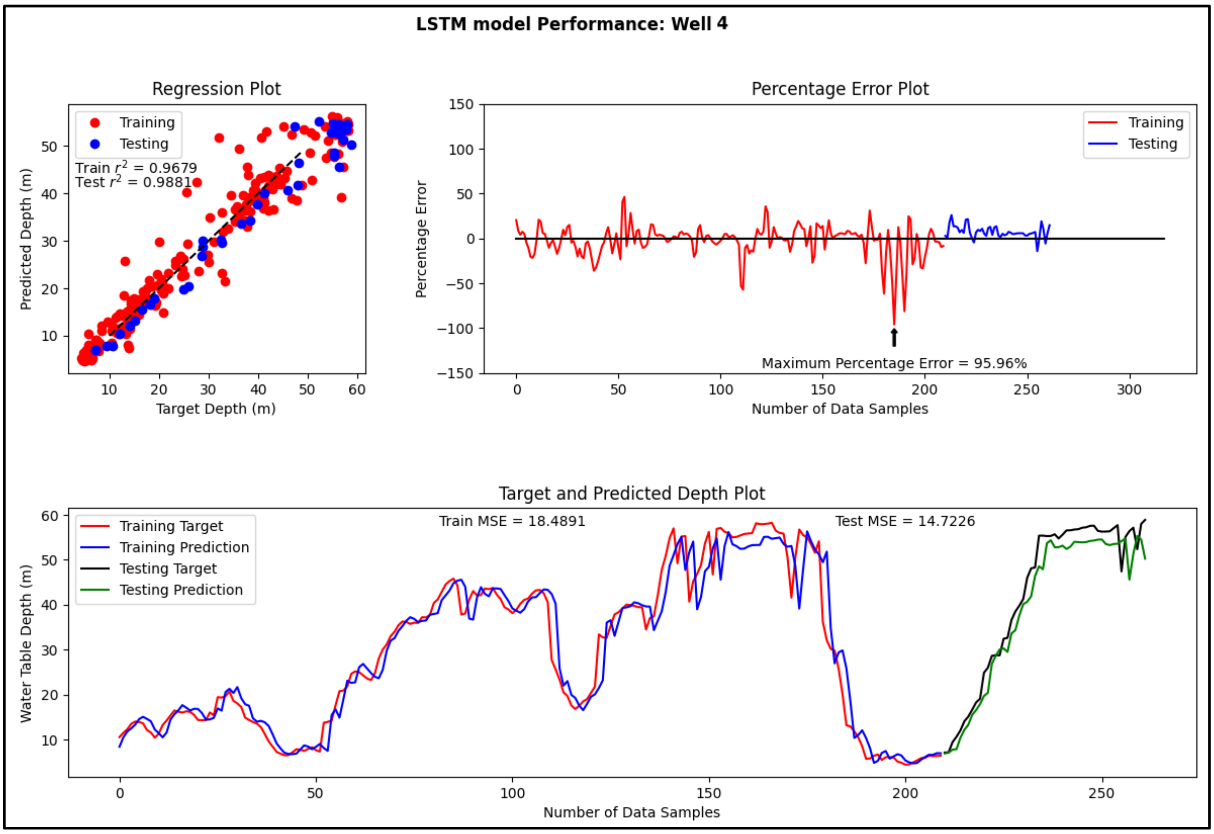


**Figure 3(d).** LSTM model performance plot for well 4

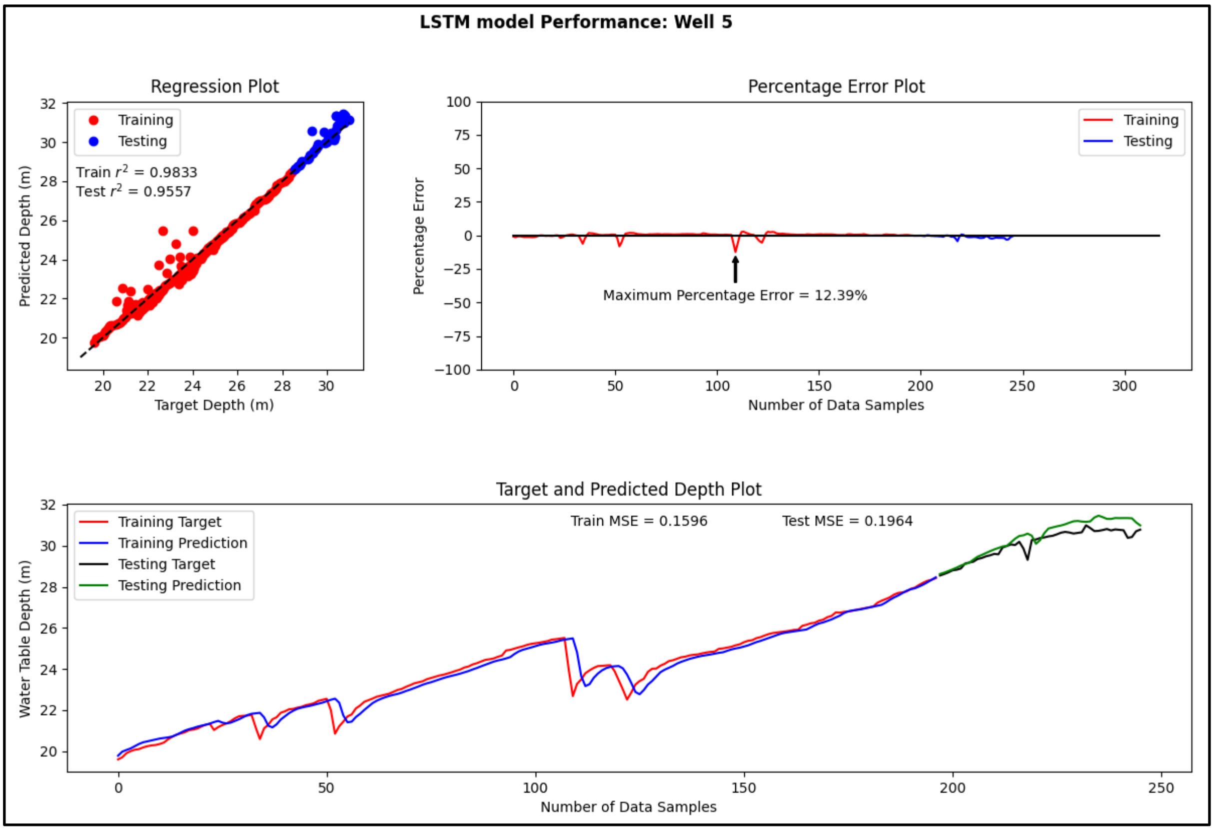


**Figure 3(e).** LSTM model performance plot for well 5

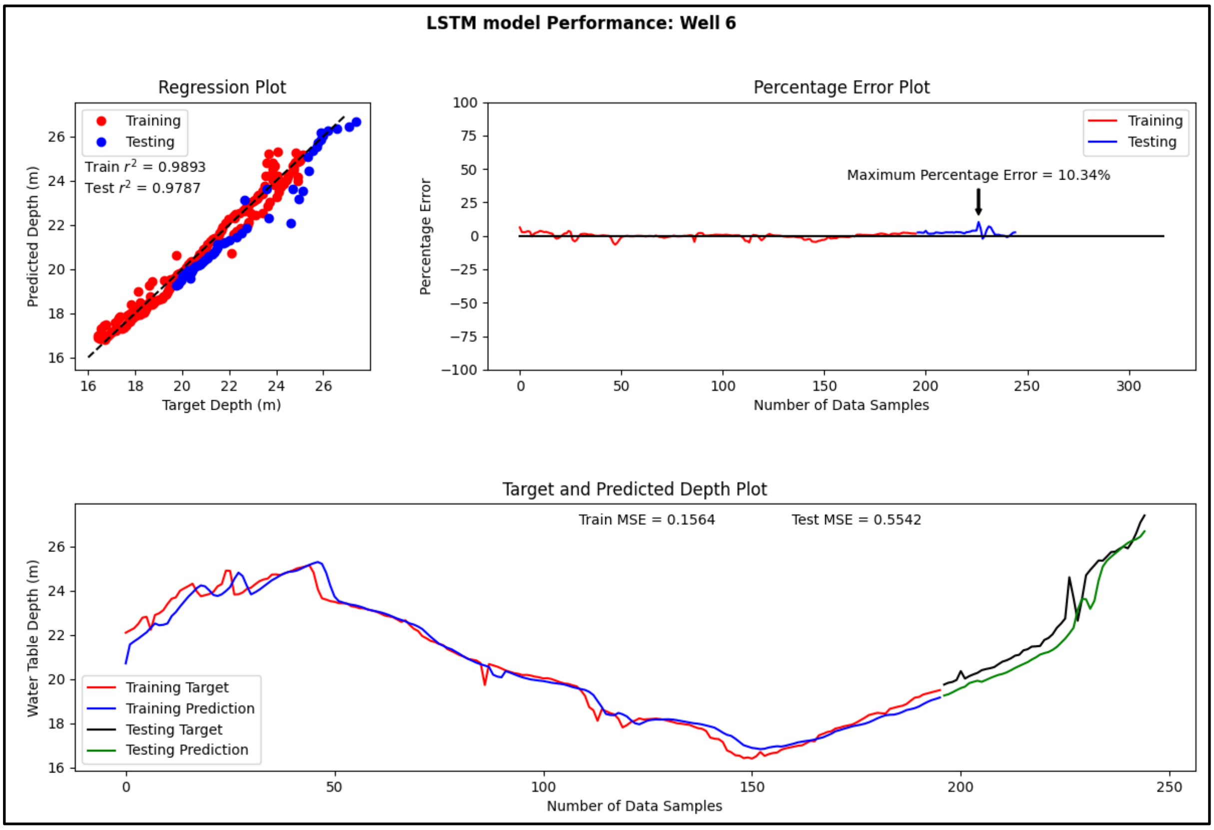


**Figure 3(f).** LSTM model performance plot for well 6

Separate LSTM model has been trained for all the six wells which predicts the water table depth one month in-advance using the previous three-month water table depth data as input. These models can help in prediction of advanced water table depth at different wells, thus helping a step forward in understanding and managing the groundwater resource in such a situation of drastic downtrend of water table depth. However, accuracy of all the models could be further enhanced to bring down the maximum percentage error and improve the $R^2$ values using hybrid modeling. In fact, rather than only using the previous groundwater level as input for the model, further enhancement could be developed by introducing other variables that influence the groundwater level in the model using different scenarios. In addition, the model performance could be improved by integrating it using advanced optimization algorithms that might improve the convergence accuracy and accelerate the time-consuming to achieve the optimal value of the model internal parameters. Furthermore, the model could be examined against other groundwater dataset in different aquifers or different climatic zones.

## 4. CONCLUSION

The groundwater table prediction model has been developed using LSTM models using TensorFlow libraries in python programming language. The model has been applied using data of six wells at different locations in UAE. These models predict the water table depth one month in-advance using previous three months water table depth data as input. Analysis of the data stated that groundwater has declined drastically between 1977 to 2011. The performance of the models was analyzed through testing and training $R^2$ values, testing and training MSE values and maximum percentage error. Models trained on this data provide good accuracy in all performance criteria. The training $R^2$ values for all the six models were greater than 0.96 and the testing $R^2$ values of all the six models were greater than 0.91. Predicted one-month in advance water table depth can help in management and restoration of drastically depleting groundwater. However, further modification could be proposed to enhance the accuracy of the results of LSTM models using hybrid models. New direction for modeling structure could be proposed by introducing different input variables that affect the groundwater level and integrate the model with meta-heuristic optimization algorithms to improve the convergence process.

## ACKNOWLEDGMENT

The project was funded from UAE University within the initiatives of Asian Universities Alliance collaboration.

## REFERENCES

Abuelgasim, A., Elkamali, M., 2019. TIME SERIES ANALYSIS OF GROUNDWATER LEVELS USING GEOGRAPHIC INFORMATION SYSTEMS IN AL AIN REGION, IN THE UNITED ARAB EMIRATES.

Afan, H.A. et al., 2021. Modeling the fluctuations of groundwater level by employing ensemble deep learning techniques. Engineering Applications of Computational Fluid Mechanics, 15(1): 1420-1439. DOI:10.1080/19942060.2021.1974093

Al-Ruzouq, R. et al., 2019. Potential groundwater zone mapping based on geo-hydrological considerations and multi-criteria spatial analysis: North UAE. CATENA, 173: 511-524. DOI:https://doi.org/10.1016/j.catena.2018.10.037

Chen, K., 2020. APSO-LSTM: An Improved LSTM Neural Network Model Based on APSO Algorithm. Journal of Physics: Conference Series, 1651: 012151. DOI:10.1088/1742-6596/1651/1/012151

Chung, H., Shin, K.-s., 2018. Genetic Algorithm-Optimized Long Short-Term Memory Network for Stock Market Prediction. Sustainability, 10(10): 3765.

de Graaf, I.E.M. et al., 2017. A global-scale two-layer transient groundwater model: Development and application to groundwater depletion. Advances in Water Resources, 102: 53-67. DOI:https://doi.org/10.1016/j.advwatres.2017.01.011

Haas, J.C., Birk, S., 2017. Characterizing the spatiotemporal variability of groundwater levels of alluvial aquifers in different settings using drought indices. Hydrol. Earth Syst. Sci., 21(5): 2421-2448. DOI:10.5194/hess-21-2421-2017

Hanoon, M.S. et al., 2021. Application of Artificial Intelligence Models for modeling Water Quality in Groundwater: Comprehensive Review, Evaluation and Future Trends. Water, Air, & Soil Pollution, 232(10): 411. DOI:10.1007/s11270-021-05311-z

Hanoon, M.S. et al., 2022. Application of Soft Computing in Predicting Groundwater Quality Parameters. Frontiers in Environmental Science, 10. DOI:10.3389/fenvs.2022.828251

Karim, F., Majumdar, S., Darabi, H., Chen, S., 2018. LSTM Fully Convolutional Networks for Time Series Classification. IEEE Access, 6: 1662-1669. DOI:10.1109/ACCESS.2017.2779939

Kombo, O.H., Kumaran, S., Sheikh, Y.H., Bovim, A., Jayavel, K., 2020. Long-Term Groundwater Level Prediction Model Based on Hybrid KNN-RF Technique. Hydrology, 7(3): 59.

Le, X.-H., Ho, H.V., Lee, G., Jung, S., 2019. Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting, Water. DOI:10.3390/w11071387

Osman, A.I.A. et al., 2022. Past, Present and Perspective Methodology for Groundwater Modeling-Based Machine Learning Approaches. Archives of Computational Methods in Engineering, 29(6): 3843-3859. DOI:10.1007/s11831-022-09715-w

Rizk, Z., Alsharhan, A., 2003. Water resources in the United Arab Emirates. Developments in Water Science, 50: 245-264. DOI:10.1016/S0167-5648(03)80022-9

Sapitang, M., Ridwan, W.M., Ahmed, A.N., Fai, C.M., El-Shafie, A., 2021. Groundwater level as an input to monthly predicting of water level using various machine learning algorithms. Earth Science Informatics, 14(3): 1269-1283. DOI:10.1007/s12145-021-00654-x

Sherif, M., Almulla, M., Shetty, A., Chowdhury, R.K., 2014. Analysis of rainfall, PMP and drought in the United Arab Emirates. International Journal of Climatology, 34(4): 1318-1328. DOI:https://doi.org/10.1002/joc.3768

Sherif, M. et al., 2021. Spatial and Temporal Changes of Groundwater Storage in the Quaternary Aquifer, UAE. Water, 13(6): 864.

Supreetha, B.S., Shenoy, N., Nayak, P., 2020. Lion Algorithm-Optimized Long Short-Term Memory Network for Groundwater Level Forecasting in Udupi District, India. Applied Computational Intelligence and Soft Computing, 2020: 8685724. DOI:10.1155/2020/8685724

Tian, Y., Pan, L., 2015. Predicting Short-Term Traffic Flow by Long Short-Term Memory Recurrent Neural Network, 2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity), pp. 153-158. DOI:10.1109/SmartCity.2015.63

UKEssays, 2018. Developing Groundwater Sources in the UAE. https://www.ukessays.com/essays/biology/developing-groundwater-sources-uae-3714.php?vref=1.

Zhang, Q., Wang, H., Dong, J., Zhong, G., Sun, X., 2017. Prediction of Sea Surface Temperature Using Long Short-Term Memory. IEEE Geoscience and Remote Sensing Letters, 14(10): 1745-1749. DOI:10.1109/LGRS.2017.2733548