Impact of using additional precipitation data from the uppermost region on improving the performance of AI models in predicting groundwater levels

제주특별자치도개발공사
Jeju Special Self-Governing Province Development Corp.

Mun-Ju Shin[1], Jeong-Hun Kim[1], Su-Yeon Kang[1], Su-Hyeon Moon[1], Jeong-Wook Kim[1], Hyuk-Joon Koh[1], Soo-Hyoung Moon[1]

[1]Jeju Special Self-Governing Province Development Corp. (Mun-Ju Shin: munjushin@jpdc.co.kr)

European Geosciences Union

2025.4.27. – 5.2.

Sharing is encouraged

ABSTRACT

In the case of Jeju Island, located in southern South Korea, groundwater is an indispensable water resource that accounts for 82% of the total water supply. Therefore, scientific prediction and management of groundwater levels are very important for the sustainable use of groundwater by citizens. This study additionally used precipitation data from the Baekrokdam Climate Change Observatory located on the summit of Jeju Island in artificial intelligence (AI) models (ANN and LSTM) to accurately predict one-month-ahead future groundwater levels for the mid-mountainous areas of Jeju Island, where groundwater levels are highly variable. When additional Baekrokdam precipitation data were used, the two AI models showed improved groundwater level prediction performance by having NSE values of 0.907 or higher. This means that the additional use of precipitation data located in the uppermost region provides more information to help interpret groundwater levels, allowing AI models to better interpret the characteristics of groundwater level fluctuations. In addition, the use of Baekrokdam precipitation data was more helpful in improving groundwater level prediction for the monitoring well, which has highly variable groundwater levels that are difficult to predict, and the ANN model with relatively low groundwater level prediction performance.

INTRODUCTION

Groundwater, along with surface water, is an important water resource that can be used for agriculture, industry, and daily life. In the case of Jeju Island, where groundwater accounts for 82% of total water resources, prediction and management of groundwater levels using useful data are very important for sustainable use of groundwater. When physical data on hydrogeology is limited and the goal is to derive accurate groundwater level prediction results rather than a physical understanding of hydrological processes, artificial intelligence models are more suitable than numerical models (Adamowski and Chan, 2011). In particular, Artificial Neural Network (ANN) and Long Short-Term Memory (LSTM) have been successfully used in various hydrological and water resource studies (Sit et al., 2020). The groundwater level fluctuation characteristics of the mid-mountainous region of Jeju Island are very different and complex due to the complex underground geology formed by volcanic activity. Although there have been studies on groundwater level simulations using artificial intelligence models in island areas (Mohanty et al., 2010; Payne et al., 2022; Kim et al., 2023), there is insufficient research on the use of additional data to accurately predict groundwater levels in island areas. The purpose of this study is to compare and analyze the effect of artificial intelligence models on improving groundwater level prediction performance by additionally using precipitation data from the Baekrokdam Climate Change Observatory located in the uppermost region to accurately predict monthly groundwater levels at observation wells located in the mid-mountainous region of Jeju Island.

STUDY AREA

The study area focuses on two groundwater level observation wells located in the mid-mountainous region of the southeastern part of Jeju Island, South Korea (Fig. 1). Baekrokdam Climate Change Observatory, located on the summit of Mt. Halla, receives more precipitation than the other two meteorological observatories in the downstream area due to the mountainous effect of the region. The groundwater levels of observation wells fluctuate at different rates depending on the complex geological characteristics caused by multiple volcanic activities (Fig. 2).

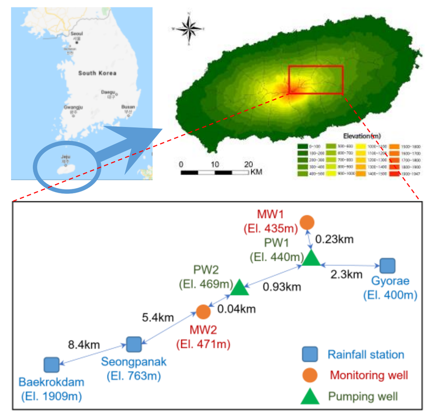

Fig. 1. Mid-mountainous region of Jeju Island in South Korea

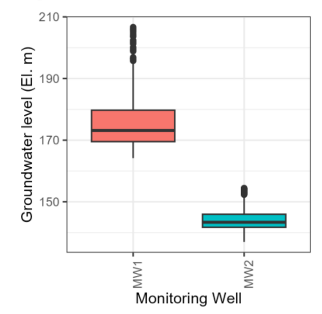

Fig. 2. Groundwater levels of observation wells

MODEL DESCRIPTION

1. Long Short-Term Memory (LSTM)

This study utilized Long Short-Term Memory (LSTM), a recurrent neural network model developed by Hochreiter and Schmidhuber (1997)(Fig. 3). This model was developed to solve the vanishing gradients problem (Bengio et al., 1994), which hinders long-term dependencies of data information during learning in artificial intelligence. The hyper-parameter values of LSTM were estimated using a trial-and-error method (Table 1).

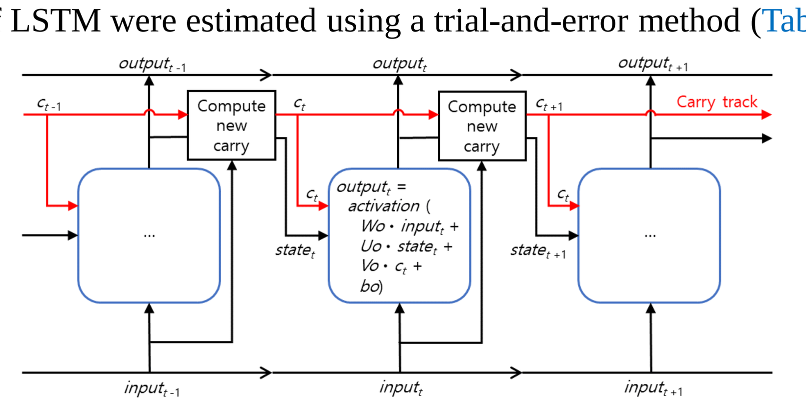

Fig. 3. Schematic of LSTM (Shin et al., 2020)

Table 1. Hyper-parameter values of LSTM model

Hyper-parameter	Range	Value	Description
n_units	-	100	Number of hidden units in hidden layer
batch_size	-	6	Number of samples fed to LSTM in one sub-simulation
dropout	0 - 1	0.5	Fraction of the units to drop for the linear transformation of the inputs
learning_rate	float >= 1	0.001	Learning rate of Adam optimizer
n_epochs	-	50	Number of iterations
patience	-	10	Number of epochs for early termination of training when simulation values for validation do not improve

2. Artificial Neural Network (ANN)

The Artificial Neural Network (ANN) model using the feed-forward method simulates nonlinear phenomena by using a network in which multiple input variables related to a target variable are connected by numerous neurons (nodes) (Adamowski and Chan, 2011). Rectified Linear Unit (ReLU) (Hahnloser et al., 2000) was used as the activation function of the hidden neurons.

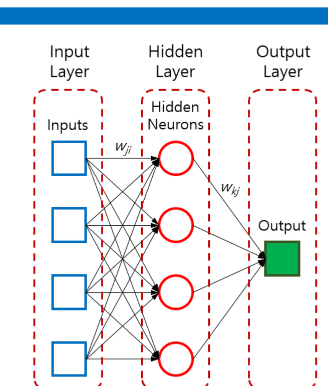

Fig. 4. Schematic of ANN (Shin et al., 2021)

METHODS

We analyzed the impact of additional use of Baekrokdam precipitation data in the AI models on improving one-month-ahead future groundwater level predictions. The data period for precipitation, groundwater withdrawal, and groundwater level used is from 2016 to 2022. Two hidden layers were applied to the AI models. Adam (Kingma and Ba, 2014) was applied as an optimization technique, and the mean absolute error was used as the objective function. Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970) and Root Mean Square Error (RMSE) were used as evaluation indices of the simulation results.

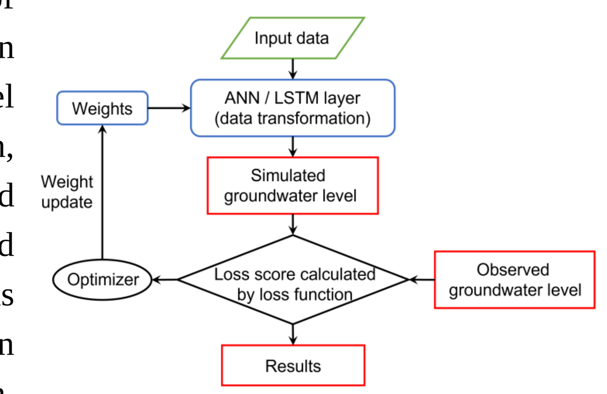

Fig. 5. Learning procedure of AI models for groundwater level prediction

RESULTS

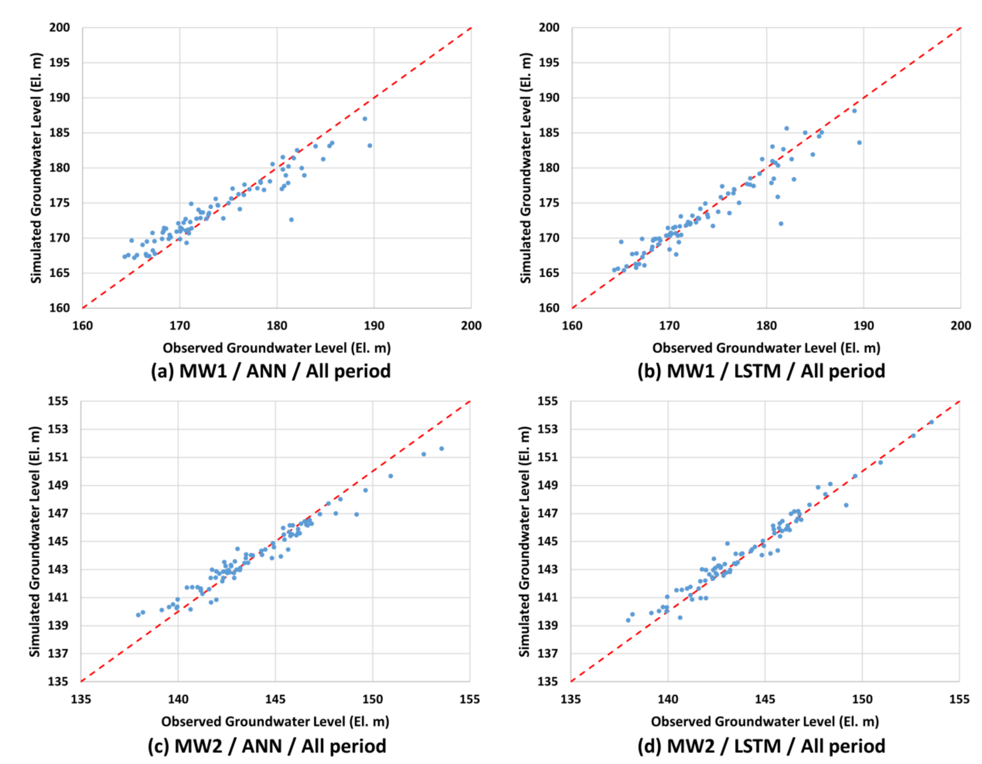

Fig. 6. Groundwater level prediction results from the AI models excluding Baekrokdam precipitation

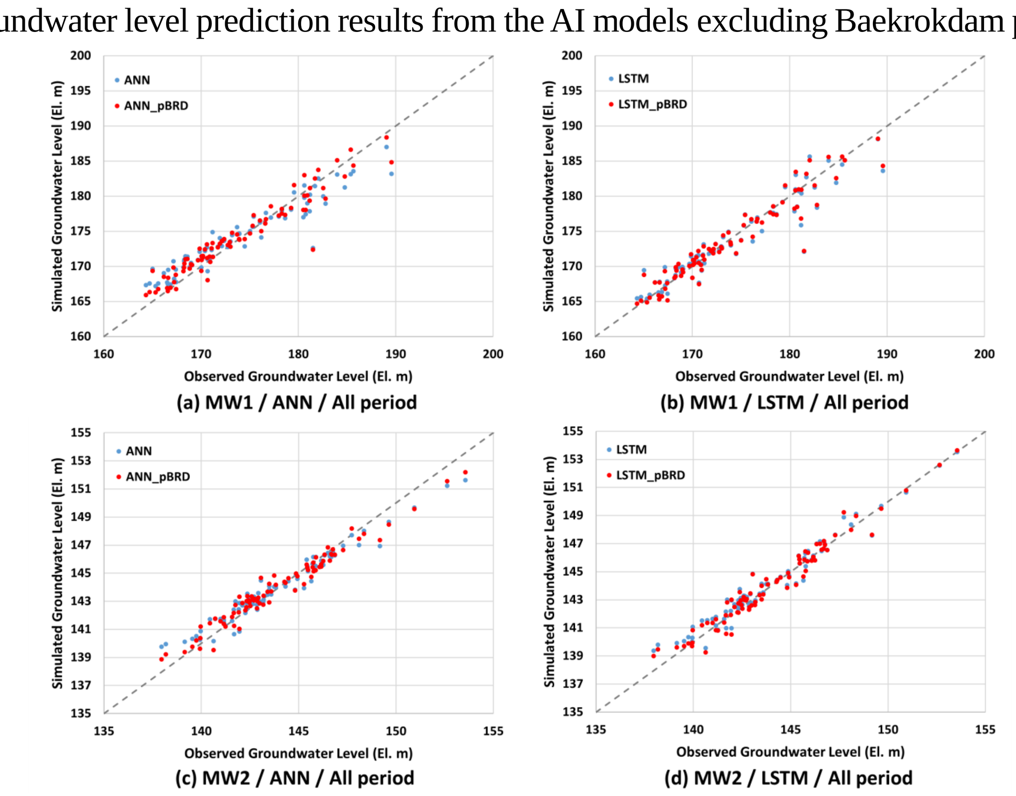

Fig. 7. One-month-ahead future groundwater level prediction results from AI models depending on whether or not Baekrokdam precipitation is utilized

Table 2. Statistics on one-month-ahead future groundwater level prediction performance of AI models depending on whether or not Baekrokdam precipitation is utilized

Monitoring Well	Performance statistics	ANN	LSTM	ANN_pBRD	LSTM_pBRD	ANN_pBRD-ANN	LSTM_pBRD-LSTM
MW1	NSE	0.871	0.897	0.907	0.908	0.036	0.011
	RMSE	2.199	1.966	1.868	1.855	-0.331	-0.111
MW2	NSE	0.936	0.952	0.948	0.956	0.012	0.004
	RMSE	0.761	0.658	0.688	0.633	-0.073	-0.025

DISCUSSION

1. When the Baekrokdam precipitation data were not used, the two AI models showed excellent groundwater level prediction performance with NSE values of 0.871 or higher (Fig. 6, Table 2). The LSTM model showed relatively higher prediction performance for high and low groundwater levels than the ANN model (Fig. 6). This means that the LSTM model adequately incorporates the seasonal effects of wet and dry periods into groundwater level simulations.
2. The more volatile the observed groundwater level, the more difficult it is for the AI models to interpret the characteristics of groundwater level fluctuations, and the lower the performance of predicting future groundwater levels (Fig. 2, Table 2).
3. When additional Baekrokdam precipitation data were used, the two AI models showed improved groundwater level prediction performance by having NSE values of 0.907 or higher (Fig. 7, Table 2). This means that the additional use of precipitation data located in the uppermost region provides more information to help interpret groundwater levels, allowing AI models to better interpret the characteristics of groundwater level fluctuations.
4. In addition, the use of Baekrokdam precipitation data was more helpful in improving groundwater level prediction for the monitoring well, which has highly variable groundwater levels that are difficult to predict, and the ANN model with relatively low groundwater level prediction performance (Table 2).
5. When additional Baekrokdam precipitation data was used for a specific monitoring well, the groundwater level prediction performance of the ANN model was improved to a level comparable to that of the LSTM model, which is a deep learning AI, even with a relatively simple ANN model structure (MW1 in Table 2). This is an example of how important it is to use additional useful data in research using AI models.

CONCLUSIONS

In this study, we compared and analyzed the improvement in groundwater level prediction performance of LSTM and ANN models by additionally using precipitation data from the Baekrokdam Climate Change Observatory located in the uppermost region to accurately predict one-month-ahead future groundwater levels. As a result, when additional Baekrokdam precipitation data was used, the ANN model and LSTM model showed even higher and improved prediction performance. It means that the AI models can more appropriately interpret the fluctuation characteristics of groundwater levels when additional precipitation data from the upstream region of the observation well are used. What is surprising is that the larger the groundwater level fluctuation range of the observation well and the lower the groundwater level prediction performance of the AI model, the more helpful the additional use of Baekrokdam precipitation data was in improving groundwater level prediction. The methods and results developed in this study can be usefully used for groundwater level prediction and management using AI models.

REFERENCES

- Adamowski, J., and Chan, H.F. (2011). “A wavelet neural network conjunction model for groundwater level forecasting.” Journal of Hydrology, Vol. 407, No. 1-4, pp. 28-40.
- Bengio, Y., Simard, P., and Frasconi, P. (1994). “Learning long-term dependencies with gradient descent is difficult.” IEEE Transactions on Neural Networks, Vol. 5, No. 2, pp. 157-166.
- Hahnloser, R.H., Sarpeshkar, R., Mahowald, M.A., Douglas, R.J., and Seung, H.S. (2000). “Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit.” Nature, Vol. 405, No. 6789, pp. 947-951.
- Hochreiter, S., and Schmidhuber, J. (1997). “Long short-term memory.” Neural Computation, Vol. 9, No. 8, pp. 1735-1780.
- Kim, D., Jang, C., Choi, J., and Kwak, J. (2023). “A case study: Groundwater level forecasting of the gyorae area in actual practice on jeju island using deep-learning technique.” Water, Vol. 15, No. 5, p. 972.
- Kingma, D.P., and Ba, J. (2014). “Adam: A method for stochastic optimization.” arXiv preprint arXiv:1412.6980.
- Mohanty, S., Jha, M.K., Kumar, A., and Sudheer, K.P. (2010). “Artificial neural network modeling for groundwater level forecasting in a river island of eastern India.” Water Resources Management, Vol. 24, pp. 1845-1865.
- Nash, J.E., and Sutcliffe, J.V. (1970). “River flow forecasting through conceptual models part I—A discussion of principles.” Journal of Hydrology, Vol. 10, No. 3, pp. 282-290.
- Payne, K., Chami, P., Odle, I., Yawson, D.O., Paul, J., Maharaj-Jagdip, A., and Cashman, A. (2022). “Machine learning for surrogate groundwater modelling of a small carbonate island.” Hydrology, Vol. 10, No. 1, p. 2.
- Shin, M.J., Kim, J.W., Moon, D.C., Lee, J.H., and Kang, K.G. (2021). “Comparative analysis of activation functions of artificial neural network for prediction of optimal groundwater level in the middle mountainous area of Pyoseon watershed in Jeju Island.” Journal of Korea Water Resources Association, Vol. 54, No. 12, pp. 1143-1154.
- Shin, M.J., Moon, S.H., Kang, K.G., Moon, D.C., and Koh, H.J. (2020). “Analysis of Groundwater Level Variations Caused by the Changes in Groundwater Withdrawals Using Long Short-Term Memory Network.” Hydrology, Vol. 7, No. 3, p. 64.
- Sit, M., Demiray, B.Z., Xiang, Z., Ewing, G.J., Sermet, Y., and Demir, I. (2020). “A comprehensive review of deep learning applications in hydrology and water resources.” Water Science and Technology, Vol. 82, No. 12, pp. 2635-2670.

ACKNOWLEDGEMENTS

This study used meteorological data from the Baekrokdam Climate Change Observatory provided by the Jeju Regional Meteorological Administration. We would like to thank them for providing the data.