*"**Importance of spatial and depth-dependent drivers in groundwater level modeling through machine learning**" by Pragnaditya Malakar, Abhijit Mukherjee, Soumendra N. Bhanja, Dipankar Saha, Ranjan Kumar Ray, Sudeshna Sarkar, Anwar Zahid*

Reviewer #1: This study investigated the spatial patterns of the performance of machine learning method for the groundwater level modeling using a large number of observations. the topic of this manuscript is interesting, considering the importance of groundwater resources management in the south Asia. I think this manuscript can be accepted after major revisions. The comments are given below:

Reply: We would like to thank the reviewer for his/her appreciation. We have addressed the reviewer's comments and done a complete major revision of the manuscript. Doing so, we have added details on the method, results, and discussion, rewritten several sections, and added a few new figures, which we believe have greatly improved the manuscript.

Highlights of the revision:

We have

a) Explained the potential rationale behind SVM's better performance than ANN
b) Described the hydrogeological conditions and aquifer characteristics and two maps added in this regard
c) Added two figures and a table showing data availability of the observation wells and also included analyses to show the influence of missing data on model performances
d) Including a detailed discussion of the previous studies, differences with our study, and originality of the present study
e) Moved the flowchart to the main text

Rev 1. Major Comment 1: The SVM performs better than ANN model, the reason behind this result should be explained.

Reply: We would like to thank the reviewer for the comment. In this study, our findings suggest that SVM performs better than ANN, which is in agreement with other studies *(Yoon et al., 2011; Yoon et al., 2016; Mukherjee et al., 2018, Bhanja et al., 2019)* on groundwater level prediction or simulation. The exact reason for the improved model performance is difficult to explain since the models have different structures. One of the potential reasons could be SVM's better robustness than ANN for its application on groundwater level time-series data. Following the reviewer's comment, we have added a paragraph in the Discussion section.

We added,

"In this context, we used similar input data in ANN and SVM to compare the performance of these computational methods. Our results are in general agreement with the previous studies (Yoon et al., 2011; Yoon et al., 2016; Mukherjee et al., 2018, Bhanja et al., 2019) on SVM performing better than ANN in groundwater level time series simulation and prediction. The exact rationale behind the SVM model performance improvement is difficult to explain since the models have different structures. A potential reason could be linked to SVM's ability of converging on global minimum and allowing a better tolerance to the noise (based on the inherent pattern associated with the datasets). Thus SVM may have certain benefits over ANN regarding the robustness and convergence (Burges et al., 1998; Karamizadeh et al., 2014). Thus, the results further highlight the importance of developing multiple methods for groundwater level modeling using machine learning. The comparison may indicate the best way forward, which is one of the motivations of this study."

References

Burges, C. J. C.: A tutorial on support vector machines for pattern recognition, Data Min. Knowl. Discov., 2(2), 121–167, doi:10.1023/A:1009715923555, 1998.

Karamizadeh, S., Abdullah, S. M., Halimi, M., Shayan, J. and Rajabi, M. J.: Advantage and drawback of support vector machine functionality, I4CT 2014 - 1st Int. Conf. Comput. Commun. Control Technol. Proc., 63–65, doi:10.1109/I4CT.2014.6914146, 2014.

Mukherjee, A. and Ramachandran, P.: Prediction of GWL with the help of GRACE TWS for unevenly spaced time series data in India : Analysis of comparative performances of SVR, ANN and LRM, J. Hydrol., 558(October 2008), 647–658, doi:10.1016/j.jhydrol.2018.02.005, 2018.

Yoon, H., Jun, S. C., Hyun, Y., Bae, G. O. and Lee, K. K.: A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer, J. Hydrol., 396(1–2), 128–138, doi:10.1016/j.jhydrol.2010.11.002, 2011.

Yoon, H., Hyun, Y., Ha, K., Lee, K. K. and Kim, G. B.: A method to improve the stability and accuracy of ANN- and SVM-based time series models for long-term groundwater level predictions, Comput. Geosci., 90, 144–155, doi:10.1016/j.cageo.2016.03.002, 2016.

Rev 1. Major Comment 2: Figure 4. For Bhahmaputra (DP), it seems that all the models show large errors in the testing period. The model may loss stability in this region for deep depth modeling. This needs explanations.

Reply: We thank the reviewer for the comment. Figure 4 demonstrates the comparative time series of the observed and simulated median groundwater levels for all the basin, sub-basin, and depth categories using ANN and SVM. As pointed by the reviewer, it is evident that Brahmaputra (DP) shows a large error. Please note that there are only three observation wells in the Brahmaputra (DP) category. Thus from a statistical perspective, the median observed and simulated median groundwater levels are less reliable.

We have mentioned this in the text.

"*However, the DP observations for the Brahmaputra basin is less reliable for the use of a comparatively lower number of deep wells (Table S4) in the basin.*"

Rev 1. Major Comment 3: Section 2.1. The aquifer is heterogeneous. The spatial variations in permeability and other hydrogeology conditions, such as the character of the rocks, the depth of the aquifer, etc, need to be described here.

Reply: We thank the reviewer for his/her comment. The IGBM basin exhibits a wide range of permeability, transmissibility, hydraulic conductivity, and aquifer depth. The diverse depositional settings and environment of Pleistocene to Holocene sediments resulted in variable aquifer properties across the basin. Following the reviewer's suggestions, we have added a brief description of the hydrogeological conditions and aquifer characteristics of the IGBM.

Furthermore, we also added two figures showing the aquifer type, horizontal hydraulic conductivity, transmissivity, and specific yield of India and Bangladesh.

We added,

"*The sediment (both recent Plio-Pleistocene to Holocene alluvium and older Miocene rocks) thickness of IGBM is up to 2 km (Singh et al., 1996). However, the effective thickness of the aquifer in most of the IGBM is generally the top 200 m. Notably, in the Bengal basin area in the eastern part of the Ganges basin and the Indus basin area, the effective aquifer thickness could be greater than 300 m (Mukherjee et al., 2007; Macdonald et al., 2015; Bonsor et al., 2017). The diverse depositional setting and environment of Pleistocene to Holocene sediments resulted in variable aquifer properties across the basin (Bonsor et al., 2017). A distinct systematic reduction in permeability is found away from the mountain and towards the coast in most of the IGBM; however, the distribution is more complex for the Ganges basin (Macdonald et al., 2015). The transmissivity within the upper and middle Ganges basin and most of the Brahmaputra basin ranges from several 100 $m^2 day^{-1}$ to more than 5000 $m^2 day^{-1}$ (Bonsor et al., 2017), which is representative of permeability values of 5 – 100 m/d (CGWB 2010). However, in the Indus basin, the permeability values of <10 $m \cdot day^{-1}$ to >60m $m^2 day^{-1}$ is reported. Figure S1 show the aquifer type, horizontal hydraulic conductivity, and transmissivity of India and Bangladesh (Bhanja et al., 2017a, 2019a). The specific yield in the unconsolidated sedimentary (high hydraulic conductivity) aquifer part of the IGBM ranges from 0.06 to 0.20 (mean 0.013). However, the specific yield values up to 0.08 are reported in the consolidated sedimentary (medium hydraulic conductivity) part of the basin (Bhanja et al., 2016). A specific yield map for India and Bangladesh is shown in Figure S2.*"

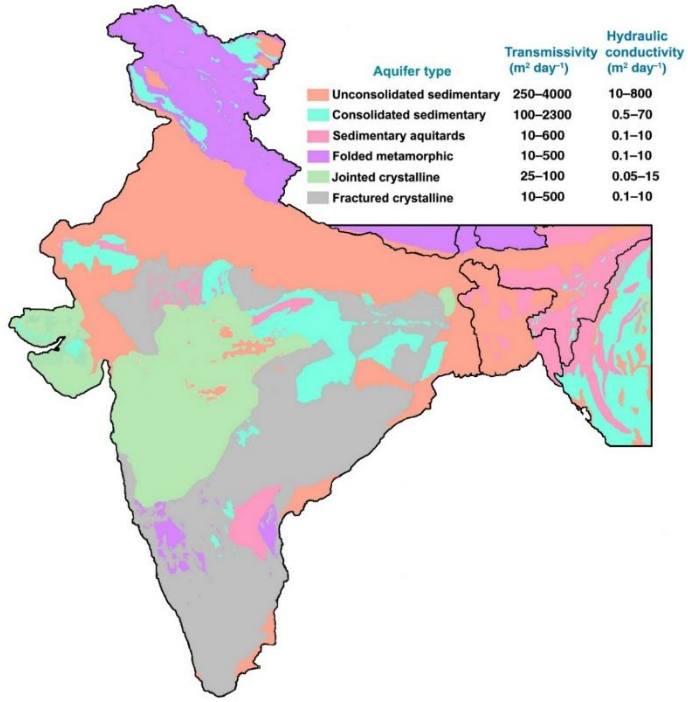


Figure S1. Different aquifer types, horizontal hydraulic conductivity (mday^{-1}) and transmissivity (m^{2}day^{-1}) for India and Bangladesh (modified from Bhanja et al., 2019a).

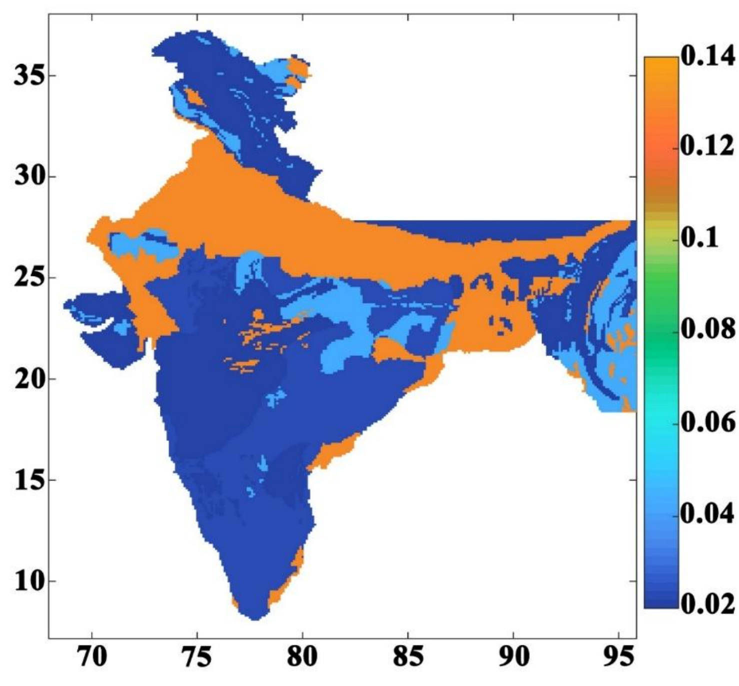


Figure. S2. Specific yield map for India and Bangladesh (modified from Bhanja et al., 2016).

Reference

Bonsor, H. C., MacDonald, A. M., Ahmed, K. M., Burgess, W. G., Basharat, M., Calow, R. C., Dixit, A., Foster, S. S. D., Gopal, K., Lapworth, D. J., Moench, M., Mukherjee, A., Rao, M. S., Shamsudduha, M., Smith, L., Taylor, R. G., Tucker, J., van Steenbergen, F., Yadav, S. K. and Zahid, A.: Hydrogeological typologies of the Indo-Gangetic basin alluvial aquifer, South Asia, Hydrogeol. J., 25(5), 1377–1406, doi:10.1007/s10040-017-1550-z, 2017.

Bhanja, S. N., Mukherjee, A., Saha, D., Velicogna, I. and Famiglietti, J. S.: Validation of GRACE based groundwater storage anomaly using in-situ groundwater level measurements in India, J. Hydrol., 543, 729–738, doi:10.1016/j.jhydrol.2016.10.042, 2016.

Bhanja, S. N., Mukherjee, A., Rodell, M., Wada, Y., Chattopadhyay, S., Velicogna, I., Pangaluru, K. and Famiglietti, J. S.: Groundwater rejuvenation in parts of India influenced by water-policy change implementation, Sci. Rep., 7(1), 7453, doi:10.1038/s41598-017-07058-2, 2017a.

Bhanja, S. N., Mukherjee, A., Rangarajan, R., Scanlon, B. R., Malakar, P. and Verma, S.: Long-term groundwater recharge rates across India by in situ measurements, Hydrol. Earth Syst. Sci., 23(2), 711–722, doi:10.5194/hess-23-711-2019, 2019a.

Central Ground Water Board: Groundwater quality in shallow aquifers of India, Govt. of India, Ministry of Water Resources, Faridabad, 76 pp., 2010

MacDonald, A.M., Bonsor, H. C., Taylor, R., Shamsudduha, M., Burgess, W. G., Ahmed, K. M., Mukherjee, A., Zahid, A., Lapworth, D., Gopal, K., Rao, M. S., Moench, M., Bricker, S. H., Yadav, S. K., Satyal, Y., Smith, L., Dixit, A., Bell, R., van Steenbergen, F., Basharat, M., Gohar, M. S., Tucker, J., Calow, R. . C. and Maurice, L.: Groundwater resources in the Indo-Gangetic Basin: resilience to climate change and abstraction, 2015.

Mukherjee, A., Fryar, A. E. and Howell, P. D.: Regional hydrostratigraphy and groundwater flow modeling in the arsenic-affected areas of the western Bengal basin, West Bengal, India, Hydrogeol. J., 15(7), 1397–1418, doi:10.1007/s10040-007-0208-7, 2007.

Singh, I.B.: Geological evolution of Ganga Plain: an overview. J Palaeontol Soc India 41:99–137, 1996

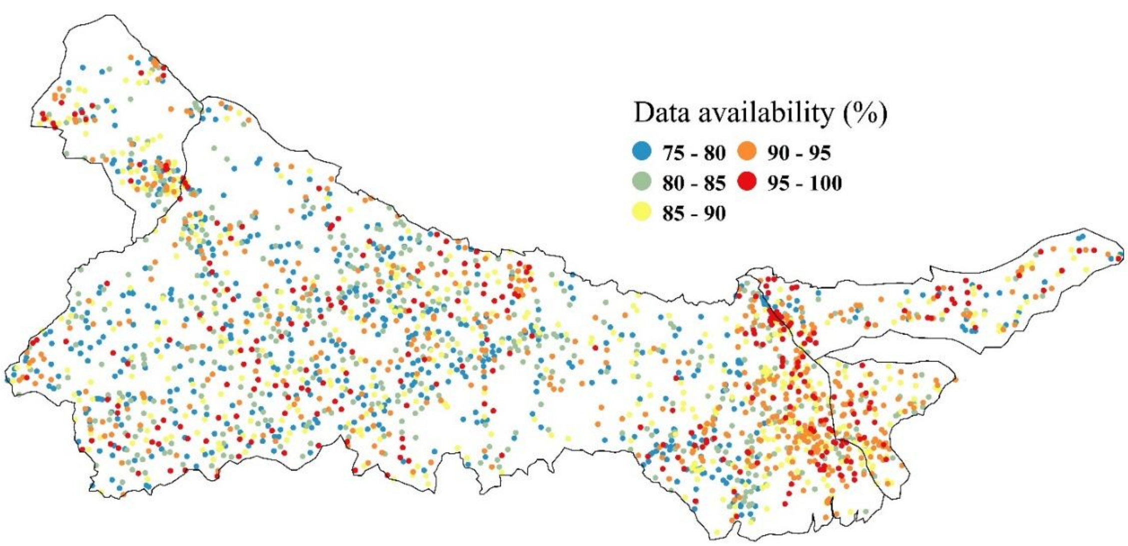
Rev 1. Major Comment 4: Line 145 “The missing values in the GWL time series data were filled using Multiple imputation”. It needs to be clarified that how many wells have missing data. Did filling missing data influence the result? A comparison of modeling results at wells with missing data and other wells without missing data is needed.

Reply: We thank the reviewer for his/her comment.

Following the reviewer's comment, we have added a spatial figure showing the data availability of each observation wells. We have also added a histogram showing the distribution of the number of observation wells (x axis) with groundwater level data records (y axis).

We added in the text,

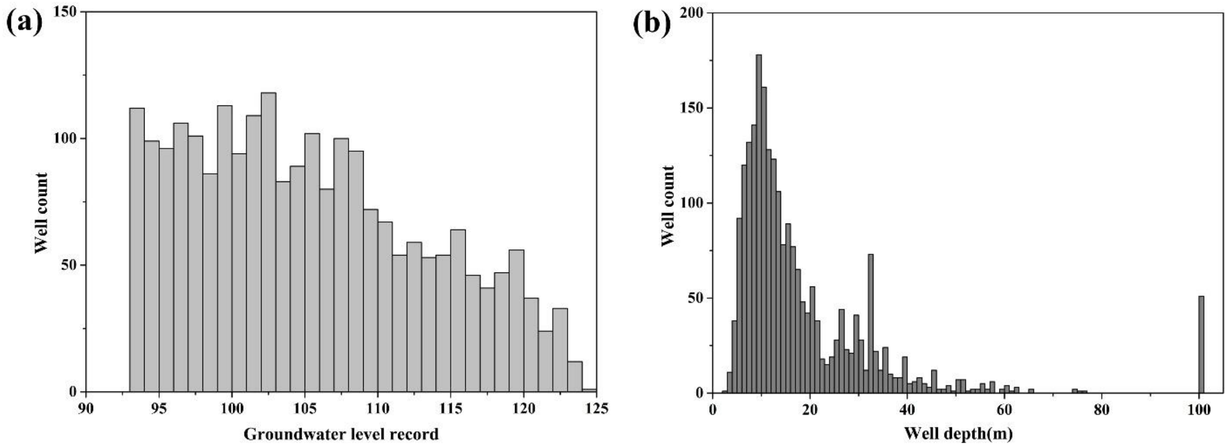
“The location wise GWL data availability is shown in Figure S4.”



“Figure S4. Location-wise groundwater level data availability (%).”

We also modified,

“The usable number of observation wells (Figure S5) was reduced significantly (from n=13465 to usable n=2303), following the application of these filters and data processing.”



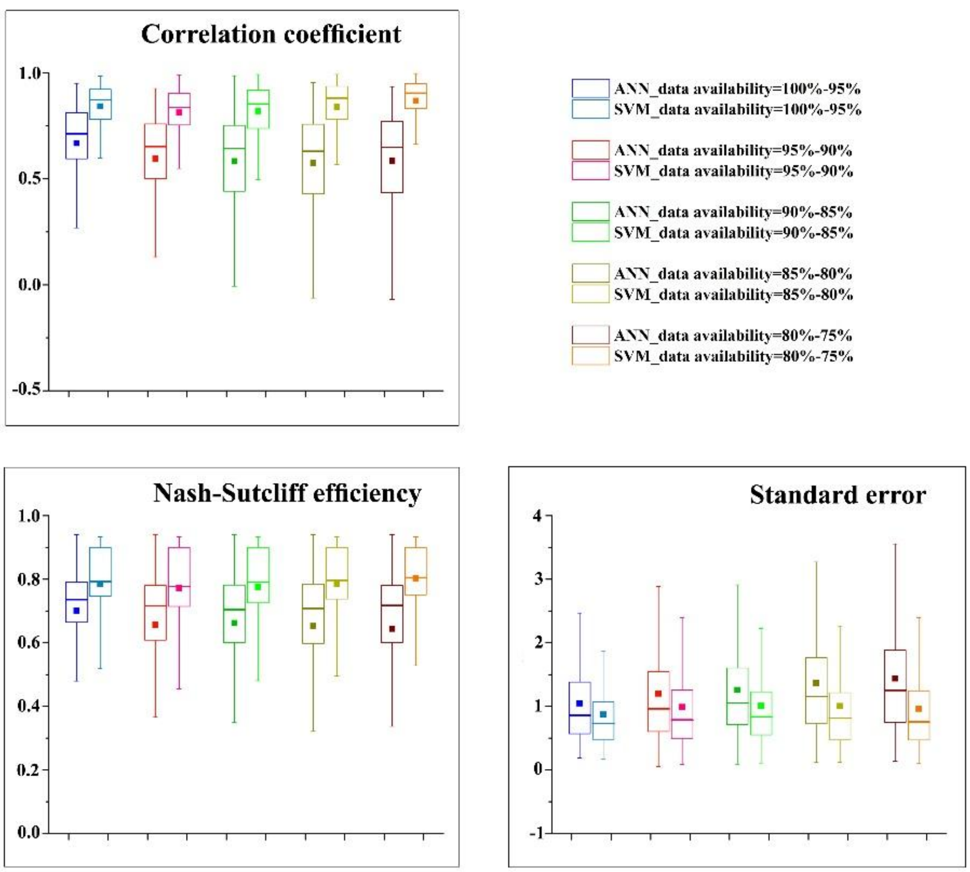
"Figure S5. Summary of observation well data included in the study."

To show the influence of missing data, we divided the observation wells into five categories (i.e., 75-80%, 80-85%, 85-90%, 90-95%, and 95-100% data availability) based on the data availability. We also computed the model performance for these five categories of data. We have included these results and discussed the finding in the revised manuscript. In general, the result suggests slightly better performance in the observation wells with higher data availability. However, no major increase or decrease in correlation coefficient, Nash-Sutcliff efficiency, and standard error are observed.

"Groundwater level dataset contains few observation wells with no missing data (i.e., 100% data availability). Thus, to show the influence of missing data on groundwater level simulation results, we divided the observation wells according to highest to lowest data availability (Figure S5; Table S5). Similar to the sensitivity analysis, we used the best performing model (Model B) for the shallow wells (as the spatial variability is better for shallow wells with a good number of measurement location availability) in the IGBM basin. In general, the result (Figure S18) suggests slightly better performance for the observation wells with higher data availability. However, no major increase or decrease in correlation coefficient, Nash-Sutcliff efficiency, and standard error are observed (Figure S18)."

Table S5. *Classes of observation wells adapted to observe the effect of missing data on model performance.*

Data availability (%)	*ANN*	*SVM*	*Number of wells*
100 - 95	*ANN_data availability=100%-95%*	*SVM_data availability=100%-95%*	*275*
95 - 90	*ANN_data availability=95%-90%*	*SVM_data availability=95%-90%*	*430*
90 - 85	*ANN_data availability=90%-85%*	*SVM_data availability=90%-85%*	*423*
85 - 80	*ANN_data availability=85%-80%*	*SVM_data availability=85%-80%*	*442*
80 - 75	*ANN_data availability=80%-75%*	*SVM_data availability=80%-75%*	*510*



"Figure S18. Comparison of model performances with data availability."

Rev 1. Major Comment 5: I suggest the author to add a discussion with previous similar studies to illustrate the differences between this study and other studies.

Reply: We thank the reviewer for the comment. In recent times, machine learning-based methods have been widely used to simulate and predict groundwater levels across the world. Most of these studies used methods like autoregressive integrated moving average (ARIMA), Artificial Neural Network (ANN), hybrid-ANN, Adaptive neuro-fuzzy inference system (ANFIS), genetic programming (GP), Support Vector Machine (SVM) and nonlinear auto-regressive exogenous model-based (NARX), long-short term memory (LSTM) model few others using a wide range of frequency and temporal data on past GWLs, satellite observations derived groundwater storage (GWS), Normalized difference vegetation index (NDVI)), meteorological variables, river discharge, variables of groundwater use, few dummy variables to simulate and/or predict GWLs. However, most studies (including studies on India and Bangladesh) are mainly small-scale studies, and due to the small number of observation wells, they are unable to characterize the spatial variability in model performances extensively. Moreover, the temporal extent of the studies on India and Bangladesh is often short. Hence the predictions are based on the short-term trends of dependent variables and do not consider the long-term variability. Furthermore, to our knowledge, none of the studies have considered the spatial and depth-wise performance variability of machine learning models in predicting GWL. The originality of this study lies in addressing some critical aspects which were not included in the previous studies. Firstly, to understand the spatial variability in machine learning-based model performances, we have considered a large network of monitoring wells (n = 2303) from 1985 to 2015 to simulate GWLs in the IGBM. Secondly, considering the variable depth-wise patterns of groundwater abstraction, we showed the significance of well depth (intake depth of the observation wells) information in GWL modeling using machine learning. Thirdly, we used meteorological variables exclusively to simulate in-situ GWLs. Fourthly, based on dominance analysis and outputs from the machine learning models, we investigated the most influential basin specific predictor(s) (both natural and human-induced) in GWL modeling. Following the reviewer's suggestion, we have modified the text, including the findings of the previous studies, and added new text to show the differences and originality of the present study.

"Over the years, the simplistic approach and acceptable results of the machine learning (ML) methods are preferred when the underlying physical system is not well understood. Sometimes, decoding the physical system becomes much more complex due to interactions and feedbacks between multiple processes. GWL modeling based on ML has the unique ability to find the likely relationships between GWL and controlling hydro-climatic-anthropogenic variables without constructing knowledge-driven conceptual or physically-based models. Therefore, researchers have studied the performance of ML methods for GWL modeling in India and Bangladesh (Nayak et al., 2006; Nury et al., 2017; Malakar et al., 2018; Mukherjee and Ramachandran, 2018; Bhanja et al., 2019b; Sun et al., 2019; Yadav et al., 2020; and the references therein) and other parts of the world (Coulibaly et al., 2001; Feng et al., 2008; Sun, 2013; Nourani and Mousavi, 2016; Sun et al., 2016; Yoon et al., 2016; Barzegar et al., 2017; Ebrahimi and Rajaee, 2017; Wunsch et al., 2018; Zhang et al., 2018; Chen et al., 2019; Lee et al., 2019; Jeong et al., 2019 and the references therein). Most of these studies used methods like autoregressive integrated moving average (ARIMA), Artificial Neural Network (ANN), hybrid-ANN, Adaptive neuro-fuzzy inference system (ANFIS), genetic programming (GP), Support Vector Machine (SVM) and nonlinear auto-regressive exogenous model-based (NARX), long-short term memory (LSTM) model few others using a wide range of frequency and temporal data on past GWLs, satellite observations derived groundwater storage (GWS), Normalized difference vegetation index (NDVI)), meteorological variables, river discharge, variables of groundwater use, few dummy variables to simulate and/or predict GWLs. In a study by Yoon et al. (2011), ANN and SVM models were used using tide level, precipitation, and past GWLs as inputs to predict GWL fluctuations at a South Korean coastal aquifer. They reported that precipitation and tide levels are the most important input variables, and SVM performs better than ANN. Furthermore, the ability of GP, ANFIS, ANN, SVM, and ARIMA methods was evaluated by Shiri et al. (2013) in predicting GWL in Korea. The results suggest that the performance of GP is better than others. Using hybrid ANN with preprocessing approach Sahoo et al. (2017) predicted GWL change in some of the agriculture alluvial aquifers of the USA. Another recent study (Jeong et al., 2019) reported that NARX and LSTM methods provide good accuracy in predicting water level of two observation wells in the Korean peninsula using preprocessed climatic variables (temperature, precipitation, humidity, sunshine hours, and atmospheric pressure) that potentially affect GWL through changing the evapotranspiration and recharge. Zhang et al. (2018) identified stressed aquifer conditions by comparing the observed

and estimated GWL with LSTM using precipitation, temperature, water diversion, and evaporation as input. A recent study by Mukherjee and Ramachandran (2018) simulated GWLs for a small number (n = 5) of in-situ observation wells in India using Linear Regression Model (LRM), Artificial Neural Network (ANN), and Support Vector Regression (SVR) using Gravity Recovery and Climate Experiment (GRACE) derived terrestrial water storage (TWS) change and meteorological variables. However, the above-mentioned studies (including studies on India and Bangladesh) are mainly small-scale studies, and due to the small number of observation wells, they are unable to characterize the spatial variability in model performances extensively. Furthermore, the temporal extent of the studies on India and Bangladesh is often short (e.g., Mukherjee and Ramachandran (2018) considered the time period from 2005 to 2018). Hence the predictions are based on the short-term trends of dependent variables and do not consider the long-term variability. Moreover, using a combination of physically-based modeling and deep convolutional neural network (CNN), Sun et al. (2019) matched the GRACE based and simulated (by a land surface model as inputs) terrestrial water storage anomalies (TWSA). They further compared the calculated in-situ GWS (using specific yields and in-situ GWLs) with the variation between the observed and simulated model values and found a good correlation. However, this study does not use in-situ GWLs as model input and mainly based on the satellite observations and land surface model outputs. Moreover, a recent study by Yadav et al. (2020) used ANN and SVM on preprocessed data on GWL, precipitation, Southern Oscillation Index, Northern Oscillation Index, Niño3, and population as input to predict GWL in the urban areas of Bengaluru, India. They also discussed the significant impact of population growth in GWL estimation and prediction in urban areas in India (Yadav et al., 2020).”

We further added on the originality of our study,

“The previous studies, as well as the studies on Bangladesh and India, are mostly based on a small spatial and a short temporal extent. Furthermore, to our knowledge, none of the studies have considered the spatial and depth-wise performance variability of machine learning models in predicting GWL. The originality of this study lies in addressing some critical aspects which were not included in the previous studies. Firstly, to understand the spatial variability in machine learning-based model performances, we have considered a large network of monitoring wells (n = 2303) from 1985 to 2015 to simulate GWLs in the IGBM. Secondly, considering the variable

patterns of groundwater abstraction, we showed the significance of well depth (intake depth of the observation wells) information in GWL modeling using machine learning. Thirdly, we used meteorological variables exclusively to simulate in-situ GWL. Fourthly, based on dominance analysis and outputs from the machine learning models, we investigated the most influential basin specific predictor(s) (both natural and human-induced) in GWL modeling."

Reference

Barzegar, R., Fijani, E., Asghari Moghaddam, A. and Tziritis, E.: Forecasting of groundwater level fluctuations using ensemble hybrid multi-wavelet neural network-based models, Sci. Total Environ., 599–600, 20–31, doi:10.1016/j.scitotenv.2017.04.189, 2017.

Bhanja, S. N., Malakar, P., Mukherjee, A., Rodell, M., Mitra, P. and Sarkar, S.: Using Satellite-Based Vegetation Cover as Indicator of Groundwater Storage in Natural Vegetation Areas, Geophys. Res. Lett., 46(14), 8082–8092, doi:10.1029/2019GL083015, 2019b.

Chen, H., Zhang, W., Nie, N. and Guo, Y.: Long-term groundwater storage variations estimated in the Songhua River Basin by using GRACE products, land surface models, and in-situ observations, Sci. Total Environ., 649, 372–387, doi:10.1016/j.scitotenv.2018.08.352, 2019.

Coulibaly, P., Anctil, F., Aravena, R. and Bobée, B.: Artificial neural network modeling of water table depth fluctuations, Water Resour. Res., 37(4), 885–896, doi:10.1029/2000WR900368, 2001.

Ebrahimi, H. and Rajaee, T.: Simulation of groundwater level variations using wavelet combined with neural network, linear regression and support vector machine, Glob. Planet. Change, 148, 181–191, doi:10.1016/j.gloplacha.2016.11.014, 2017.

Feng, S., Kang, S., Huo, Z., Chen, S. and Mao, X.: Neural networks to simulate regional ground water levels affected by human activities, Ground Water, 46(1), 80–90, doi:10.1111/j.1745-6584.2007.00366.x, 2008.

Jeong, J. and Park, E.: Comparative applications of data-driven models representing water table fluctuations, J. Hydrol., 572(March), 261–273, doi:10.1016/j.jhydrol.2019.02.051, 2019.

Lee, S., Lee, K. K. and Yoon, H.: Using artificial neural network models for groundwater level forecasting and assessment of the relative impacts of influencing factors, Hydrogeol. J., 27(2), 567–579, doi:10.1007/s10040-018-1866-3, 2019.

Malakar, P., Mukherjee, A., Sarkar, S.: Potential Application of Advanced Computational Techniques in Prediction of Groundwater Resource of India, in: Mukherjee, A. (Ed.), Groundwater of South Asia. Springer Singapore, Singapore, pp. 643–655, 2018.

Mukherjee, A. and Ramachandran, P.: Prediction of GWL with the help of GRACE TWS for unevenly spaced time series data in India : Analysis of comparative performances of SVR, ANN and LRM, J. Hydrol., 558(October 2008), 647–658, doi:10.1016/j.jhydrol.2018.02.005, 2018.

Nayak, P. C., Satyaji Rao, Y. R. and Sudheer, K. P.: Groundwater level forecasting in a shallow aquifer using artificial neural network approach, Water Resour. Manag., 20(1), 77–90, doi:10.1007/s11269-006-4007-z, 2006.

Nourani, V. and Mousavi, S.: Spatiotemporal groundwater level modeling using hybrid artificial intelligence-meshless method, J. Hydrol., 536, 10–25, doi:10.1016/j.jhydrol.2016.02.030, 2016.

Nury, A. H., Hasan, K. and Alam, M. J. Bin: Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh, J. King Saud Univ. - Sci., 29(1), 47–61, doi:10.1016/j.jksus.2015.12.002, 2017.

Sahoo, S., Russo, T. A., Elliott, J. and Foster, I.: Machine learning algorithms for modeling groundwater level changes in agricultural regions of the U.S., Water Resour. Res., 53(5), 3878–3895, doi:10.1002/2016WR019933, 2017.

Shiri, J., Kisi, O., Yoon, H., Lee, K. K. and Hossein Nazemi, A.: Predicting groundwater level fluctuations with meteorological effect implications-A comparative study among soft computing techniques, Comput. Geosci., 56, 32–44, doi:10.1016/j.cageo.2013.01.007, 2013.

Sun, A. Y.: Predicting groundwater level changes using GRACE data, Water Resour. Res., 49(9), 5900–5912, doi:10.1002/wrcr.20421, 2013.

Sun, A. Y., Scanlon, B. R., Zhang, Z., Walling, D., Bhanja, S. N., Mukherjee, A. and Zhong, Z.: Combining Physically Based Modeling and Deep Learning for Fusing GRACE Satellite Data: Can We Learn From Mismatch?, Water Resour. Res., doi:10.1029/2018WR023333, 2019.

Sun, Y., Wendi, D., Kim, D. E. and Liong, S. Y.: Technical note: Application of artificial neural networks in groundwater table forecasting-a case study in a Singapore swamp forest, Hydrol. Earth Syst. Sci., 20(4), 1405–1412, doi:10.5194/hess-20-1405-2016, 2016.

Wunsch, A., Liesch, T. and Broda, S.: Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX), J. Hydrol., 567, 743–758, doi:10.1016/j.jhydrol.2018.01.045, 2018.

Yadav, B., Gupta, P. K., Patidar, N. and Himanshu, S. K.: Ensemble modelling framework for groundwater level prediction in urban areas of India, Sci. Total Environ., 135539, doi:10.1016/j.scitotenv.2019.135539, 2019.

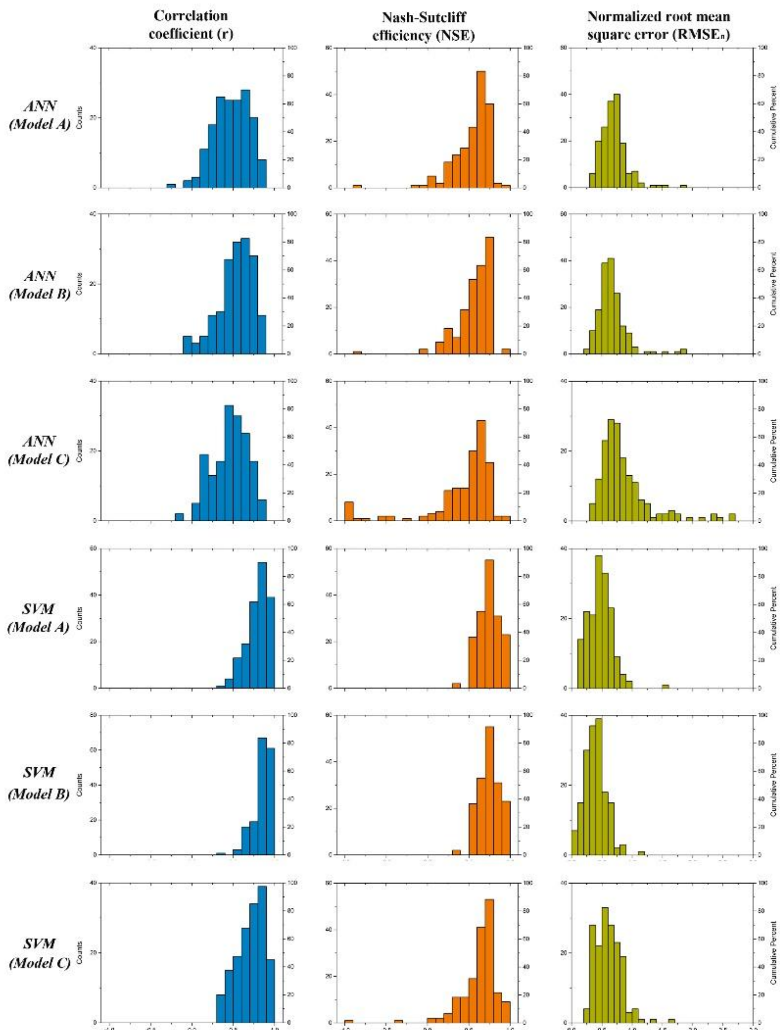
Yoon, H., Jun, S. C., Hyun, Y., Bae, G. O. and Lee, K. K.: A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer, J. Hydrol., 396(1–2), 128–138, doi:10.1016/j.jhydrol.2010.11.002, 2011.

Yoon, H., Hyun, Y., Ha, K., Lee, K. K. and Kim, G. B.: A method to improve the stability and accuracy of ANN- and SVM-based time series models for long-term groundwater level predictions, Comput. Geosci., 90, 144–155, doi:10.1016/j.cageo.2016.03.002, 2016.

Zhang, J., Zhu, Y., Zhang, X., Ye, M. and Yang, J.: Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas, J. Hydrol., 561(April), 918–929, doi:10.1016/j.jhydrol.2018.04.065, 2018.

Rev 1. Minor Comment 1: Figure S14. There are errors for the label of y-axis."SVM A" should not be followed by "ANN B" and "ANN C"

Reply: We thank the reviewer for noticing the typo. Following the reviewer's concern, we have corrected the typo in Figure S14.



"Figure S14. Distribution of observation well counts with r, NSE RMSE$_n$ for the Brahmaputra basin."

Rev 1. Minor Comment 2: Figure S3. I suggest the flowchart can be moved to the main text.

Reply: We thank the reviewer for the suggestions. Following the reviewer's comment, we have moved the flowchart to the main text as Figure 2.

Rev 1. Minor Comment 3: The manuscript analyzed the influence of population and groundwater withdrawal. These two variables may be related, this need to be clarified, and what are the purpose for groundwater abstraction?

Reply: We would like to thank the reviewer for this concern. We agree with the reviewer that population and groundwater withdrawal are interlinked parameters in some aspects. For example, assuming per capita groundwater withdrawal for domestic purposes is nearly equal, a net rise in population is directly proportional to the rise in groundwater withdrawal for domestic purposes. However, domestic withdrawal is limited to only ~4-8% of the total groundwater withdrawal in the basin, while irrigation-linked groundwater withdrawal contributes more than 90%. Irrigation strategies are shifting from flood irrigation to drip and sprinkler based irrigation systems, and this would continue in the near future. Thus, the water withdrawal for irrigation purposes (being the highest consumer of groundwater) is not directly linked to the population increase of the study area. This is the reason we have considered two separate parameters for designing this study.

"*We considered both the population and groundwater withdrawal as input parameters in the dominance analysis. Although these two parameters seem to be interlinked, however, in reality, they are not directly related in the IGBM basin. For example, assuming per capita groundwater withdrawal for domestic purposes is not changing over the years, a net rise in population is directly proportional to the rise in groundwater withdrawal for domestic purposes. However, domestic withdrawal is limited to only ~4-8% of the total groundwater withdrawal in the basin (Sharma et al., 2008; CGWB, 2019). Irrigation-linked groundwater withdrawal contributes more than 90% throughout the basin (Sharma et al., 2008; CGWB, 2019). Irrigation strategies are shifting from flood irrigation to drip and sprinkler based irrigation systems, and this would continue in the near future. Thus, the water withdrawal for irrigation purposes (being the highest consumer of groundwater) is not directly linked to population increase; rather, it is dependent upon the irrigation strategies used (Bhanja et al., 2017a).*"

References

Sharma, B. R., Amarasinghe, U. A., & Sikka, A.: Indo-Gangetic river basins: Summary situation analysis. New Delhi, India: International Water Management Institute. Pp14, 2008

Central Ground Water Board: Dynamic groundwater resources, India, 2017, Govt. of India, Ministry of Water Resources, Faridabad, 306 pp., 2019

Bhanja, S. N., Mukherjee, A., Rodell, M., Wada, Y., Chattopadhyay, S., Velicogna, I., Pangaluru, K. and Famiglietti, J. S.: Groundwater rejuvenation in parts of India influenced by water-policy change implementation, Sci. Rep., 7(1), 7453, doi:10.1038/s41598-017-07058-2, 2017a.