

Machine learning using the XGBoost algorithm for the prediction of the groundwater level in the Tafilalet aquifer (Morocco)

aicha ousrhire

aicha.ousrhire@gmail.com

Research Article

Keywords: Groundwater management, Groundwater sustainability, Groundwater Level forecasting, Machine learning, XGBoost, Tafilalet Aquifer, Morocco

Posted Date: September 30th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10113648/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Machine learning using the XGBoost algorithm for the prediction of the groundwater level in the Tafilalet aquifer (Morocco)

Abstract

The tremendous expansion of technological advances in recent decades has been accompanied by several significant advances in computer modelling. One of these advances is machine learning techniques. Machine learning and data-driven approaches have revolutionized the world and are becoming increasingly applied because of their effectiveness in solving many complex problems, such as groundwater-related issues. Anticipating fluctuations of groundwater level is important for groundwater management and sustainability, since it provides forward-looking insights into the impacts of groundwater development and abstraction practices, while also delivering critical information on the temporal behavior and underlying dynamics of the aquifer system. Such insights enhance the understanding of how aquifers respond to both natural variability and anthropogenic pressures, thereby supporting more informed decision-making, improved resource planning, and the implementation of sustainable groundwater management strategies. In this study, a machine learning algorithm named XGBoost was used for the prediction of groundwater level based on historic groundwater level data. The model resulted in one-year forecasting, which is a very important outcome since it will enable water resource planners to anticipate and assess potential risks of overexploitation, and design appropriate abstraction and recharge strategies in a proactive manner which will contribute to the sustainable management of groundwater resources and help to balance water demand with aquifer resilience and long term availability.

Keywords: Groundwater management; Groundwater sustainability; Groundwater Level forecasting; Machine learning; XGBoost; Tafilalet Aquifer; Morocco.

1. Introduction

Around the world, Groundwater plays a vital role in providing drinking water for households, supplying water for agricultural activities, and ensuring the sustainability of ecosystems that depend on it. In general, groundwater represents one of the most reliable sources of freshwater that is very important for socio-economic development, especially in countries where arid to semi-arid climates prevail, as it is in most cases their only source of water supply. As a result, groundwater resources are under enormous pressure, combined with unreasonable use and poor management, which has led to their degradation in terms of quality and quantity. In arid to semi-arid countries, this situation is much more worrying because reduced rainfall and high temperatures do not facilitate the response to the high demand for water, which has caused the depletion and sometimes the drying up of some wells, so enormous efforts must be made to improve the resilience of aquifer systems to such pressure, in this situation ensuring a balance between natural recharge and human withdrawals is important to achieve and guarantee groundwater sustainability. Improving groundwater resources management and sustainability requires the mobilization of all stakeholders, but also the mobilization of the scientific community whose efforts were, are, and will always be fruitful. Numerous tools, mechanisms, procedures, methods, and scientific approaches have been developed to study groundwater in order to ensure its proper management. Monitoring the water level in an aquifer is among the solutions adopted

since the dawn of time to quantify availability and to ensure the sustainability and efficient management of groundwater resources. Having an accurate knowledge about the current groundwater level situation is important, but having an idea about what will happen to these levels in the future is more important. Will they decrease or rather increase? in order to know what actions and measures to take. In fact, many techniques for predicting trends in groundwater levels have been developed, which can be divided into two categories: physical models and data-driven models. Physically-based models or process-based models are commonly used to simulate groundwater level changes in aquifers; many of them have been widely used to get groundwater level information, such as GMS, FEFLOW, MODFLOW, CABLE, CWatM, PCR-GLOBWB, etc. Certainly, they are powerful and effective tools however they necessitate a tremendous effort to create and calibrate model parameters (Chen et al., 2020; Yin et al., 2021), as well as need significant amount of data and take a lot of time to build especially for a complex systems like groundwater system (Ibrahim Ahmed Osman et al., 2021; Rajaei et al., 2019) besides, they require a deep knowledge about system dynamics and relevant synthesis of aquifer geological and hydrogeological properties in order to describe the sub surface's spatial variability which is a difficult to come by information and often not readily available (Kenda et al., 2018; Taormina et al., 2012), Moreover, To a large extent, the model's accuracy is determined by the correctness and accuracy of the model inputs (Sun et al., 2015).

Understanding the underlying mechanisms is less critical than making accurate predictions (Rajaei et al., 2019). On this basis, data-driven models were built. Data-driven models are another approach for predicting groundwater levels and an alternative to physically-based models that have begun to gain interest in the scientific community over the past two decades to address many complex and challenging issues in environmental science. Without explicit understanding of the physical behaviour, a data driven model is based on data analysis about investigated system, in particular determining relationships between system state variables (input, internal, and output variables) (Solomatine et al., 2008), because data are now collected in abundance in every experiment at a low cost, unlike in the past when data for research and science were scarce, limited, and difficult to obtain. Data-driven models are a way to implement artificial intelligence, and they represent today a paradigm shift in the way many problems are approached. Because of the inherent difficulty in deriving rules, laws, and equations to explain particular phenomena, various fields of science have adopted and used data-driven models in recent years (Montáns et al., 2019).

Many data-driven models have been developed and applied for groundwater level prediction, especially machine learning (ML) based models. Machine learning (ML) models are very useful because they have brought about Improved and practical solutions for better understanding very complex natural systems since they don't require physical information on the investigated system nor knowledge of the physical system's underlying processes as physically-based models do (Chen et al., 2020; Ibrahim Ahmed Osman et al., 2021). In Fact, Machine Learning is a branch of science, and more specifically, a subfield of artificial intelligence (Zhang, 2020), but it also overlaps with a variety of other scientific fields, including statistics, cognitive science, and information theory (Ławrynowicz and Tresp, 2014). Machine learning is the process of creating computer algorithms that can mimic human intelligence (El Naqa and Murphy, 2015). Machine learning, therefore, provides the mathematical tools to concretise the objective of artificial intelligence, which is to imitate human behaviour.

Since the dawn of humanity, tools and machines have been invented to ease labour and minimize the effort required to execute a variety of tasks. In fact, the starting point of machine learning started in the seventeenth century with the invention of computational applications that mimic the human ability to do addition and subtraction operations (Ifrah, 2001). In the 1950s, Arthur Samuel created a program for playing checkers (Samuel, 1959), which is considered the first realization of machine learning and proof

that computers can be programmed to play checkers and certainly other things. Arthur Samuel then became the pioneer of machine learning, and to him the credit goes for popularizing and inventing the term 'machine learning'. After that, many other inventions followed.

Machine Learning is a data analysis technique that consists of letting algorithms autonomously discover patterns and learn from data, which allows them to be trained. Once the training phase is done, the algorithm will be able to find patterns in new data. Besides, machine learning has experienced exponential expansion in its application in various domains, and this is due to the ever-increasing rise in the development of data storage and information processing technologies, the diversified and increasing number of data, the emergence of supercomputers that perform operations in a faster manner, and important advances in computer modelling. Machine Learning algorithms are varied and numerous, but some are used much more than others. Common categories of machine learning systems are: supervised learning, semi-supervised learning, reinforcement learning, transduction, unsupervised learning, and learning to learn.

Machine learning algorithms have been widely used and applied for the prediction of groundwater level fluctuations during the last two decades, the most common ones are the artificial neural network, which is used in many studies (Ardana et al., 2021; Chitsazan et al., 2015; Guezgouz et al., 2021; Nayak et al., 2006; Pandey and Malik, 2022; Rakhshandehroo et al., 2012; Sun et al., 2015; Wunsch et al., 2021), besides, other machine algorithms are also used for the prediction of groundwater fluctuations like Support vector machine, multivariate linear regression, random forest, multilayer perceptron, support vector regression, etc. In this study, Extreme gradient boosting, known as XGBoost with python language, is used to predict groundwater variations in the Tafilalet aquifer, which is located in the Tafilalet plain. This plain is part of the Drâa-Tafilalet region located in the southeast of the Kingdom of Morocco, as shown in Figure 1, and is characterized by an arid climate, low rainfall, high evapotranspiration rates, almost no vegetation cover, scarcity of surface water, and therefore overexploitation of existing groundwater resources. In such circumstances, knowledge of long-term water level trends is necessary to anticipate appropriate measures. In fact, groundwater level forecasting plays a central role, as it enables water managers and policymakers to evaluate trends, identify early warning signs of overexploitation, and assess the potential impacts of different management and climate scenarios. In this work, the prediction of groundwater level variations is done based on temporal data inputs of available historic groundwater level data from 2012 to 2017. Extreme Gradient Boosting (XGBoost) was chosen to be used due to its relatively limited application in groundwater-related research compared to other widely used machine learning algorithms, such as artificial neural networks, support vector machines, and random forests. While these algorithms have been extensively explored in the literature for groundwater level modeling and prediction, XGBoost has only been mentioned infrequently in this context. This relative lack of application highlights a clear research gap and provides an opportunity to explore and assess the suitability and added value of XGBoost for groundwater level forecasting. By investigating the performance of XGBoost in this underexplored domain, this study not only broadens the methodological landscape of groundwater modeling but also contributes novel insights into the use of advanced machine learning techniques for sustainable groundwater management. The following sections will be dedicated to presenting the Extreme Gradient Boosting algorithm and explaining the approach used. After that, presentation and discussion of the results will follow, and the last part will be devoted to the conclusion of the entire study.

2. Materials and Methods

The XGBoost algorithm was used for groundwater level forecasting regarding the Tafilalet aquifer, notably the XGBoost library in Python language. Extreme gradient boosting, abbreviated as XGBoost, is a machine learning algorithm developed by (Chen and Guestrin, 2016) belonging to the category of supervised learning algorithms and regarded as among the most effective algorithms of this category (Ibrahim Ahmed Osman et al., 2021), XGBoost is a high speed, high performance algorithm, it is based on the gradient boosting concept, and is an implementation of gradient boosted decision trees. XGBoost is available in almost all useful Machine Learning programming languages, mainly Python and R.



Figure 1. Tafilalet plain location.

The purpose of the XGBoost approach is to generate a K regression tree in order to get the tree group's predicted value as close to the true value as feasible while achieving the best generalization ability. More explicitly, the XGBoost algorithm is a lifting technique that uses residual fitting to build many weak learners and to get a strong learner; the generated weak learners are accumulated (Liu et al., 2021). More precisely, XGBoost's approach is to continually add and train new trees to fit residual errors from the previous iteration (Dong et al., 2020). Mathematically, XGBoost uses an addition expression to predict the final result and can be explained as follows (Bhati et al., 2021; Dong et al., 2020; Ibrahim Ahmed Osman et al., 2021):

suppose a data set DS of m features and n number of examples:

$$DS = \{(x_i, y_i): i = 1, 2, \dots, n, x_i \in \mathbb{R}^m, y_i \in \mathbb{R}\}$$

Note $\hat{y}_i$ as the predicted output and expressed as:

$$\hat{y}_i = \sum_{k=1}^k f_k(x_i) \quad (1)$$

where k denotes the number of trees in the model, f_k represents the result of input x_i for the k th tree,

To find a solution for the aforementioned equation, the optimal set of functions must be found by minimizing the loss and regularization objective. XGBoost's objective function is divided into two parts: training error and regularization, which are stated as:

$$\text{obj}(\theta) = \sum L(\theta) + \sum \Omega(\theta) \quad (2)$$

where L is the loss function that measures the difference between the predicted and actual output values. Ω is a regularization function that measures the training model's complexity in order to avoid overfitting, and it is calculated using the following function:

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|W\|^2 \quad (3)$$

where T denotes the number of leaves on the tree, and W denotes the weight of each leaf. γ and λ represent controlling parameters used to avoid overfitting the model. In decision trees, boosting is employed in the training phase of the model to minimize the objective function, which works by adding a new function f as the model is trained. As a result, a new function (tree) is introduced on the t -th iteration so, the objective function is rewritten as:

$$\text{obj}^{(t)} = \sum_{i=1}^n L(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (4)$$

where $\hat{y}_i^{(t-1)} + f_t(x_i)$ is the predicted score for the t -th tree, The inclusion of f increases model accuracy, allowing the objective function to be optimized in a fast way, and second-order Taylor approximation is required to make use of Euclidean space optimization techniques. Thus, the objective function can be expressed as:

$$\text{obj}^{(t)} = \sum_{i=1}^n [L(\hat{y}_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \quad (5)$$

Where:

$$g_i = \partial_{\hat{y}_i^{(t-1)}} L(\hat{y}_i, \hat{y}_i^{(t-1)})$$

$$h_i = \partial_{\hat{y}_i^{(t-1)}}^2 L(\hat{y}_i, \hat{y}_i^{(t-1)})$$

and they are respectively the first-order derivative and the second-order derivative.

Since the residual errors (y) of preceding $(t - 1)$ trees have little impact on the objective function, the constant functions can be removed, so, the objective function can be approximated as:

$$\text{obj}^{(t)} = \sum_{i=1}^n [g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \quad (6)$$

After getting the predictions, the performance of the model should be evaluated. To do that, there are performance indexes that must be calculated:

RMSE: Root mean squared error, this index determines the error between the model's output ($\hat{y}$) and the target values (y). Moreover, the lower value of RMSE implies higher accuracy of the model. Mathematically RMSE is expressed as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

R^2 : Coefficient of determination, is pronounced "R squared" and is symbolized by the letter (R^2), it is a statistical measure that calculates the proportion of the variation in the dependent variable that is predictable from the independent variable in other terms it indicates the goodness of regression fitting by measuring the deviation of a group of data (Dong et al., 2020) and is expressed as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (8)$$

Where:

($\hat{y}$) is the predicted value

(y) is the target values

($\bar{y}_i$) is the average of the measurements

The study was structured following the methodology presented in Figure 2. Firstly, data provided in Figure 3 of historic groundwater levels of seven years from 2012 to 2017 were prepared. The data were provided by the Hydraulic basin agency of Guir Ziz Rheris, and they describe monthly monitoring of the water level in a well located in the plain of Tafilalet and fed by the Tafilalet aquifer, as illustrated in

Figure 4. The well is represented by its coordinates ($x=606797$, $y=80384$) in the projection system Lambert zone Morocco (Zone I).

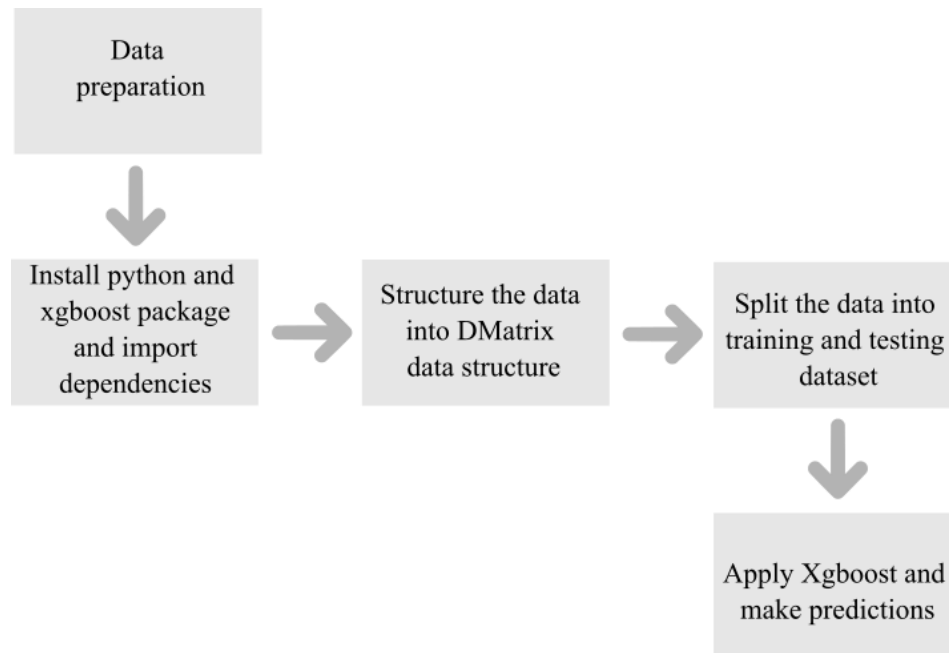


Figure 2. Study methodology.

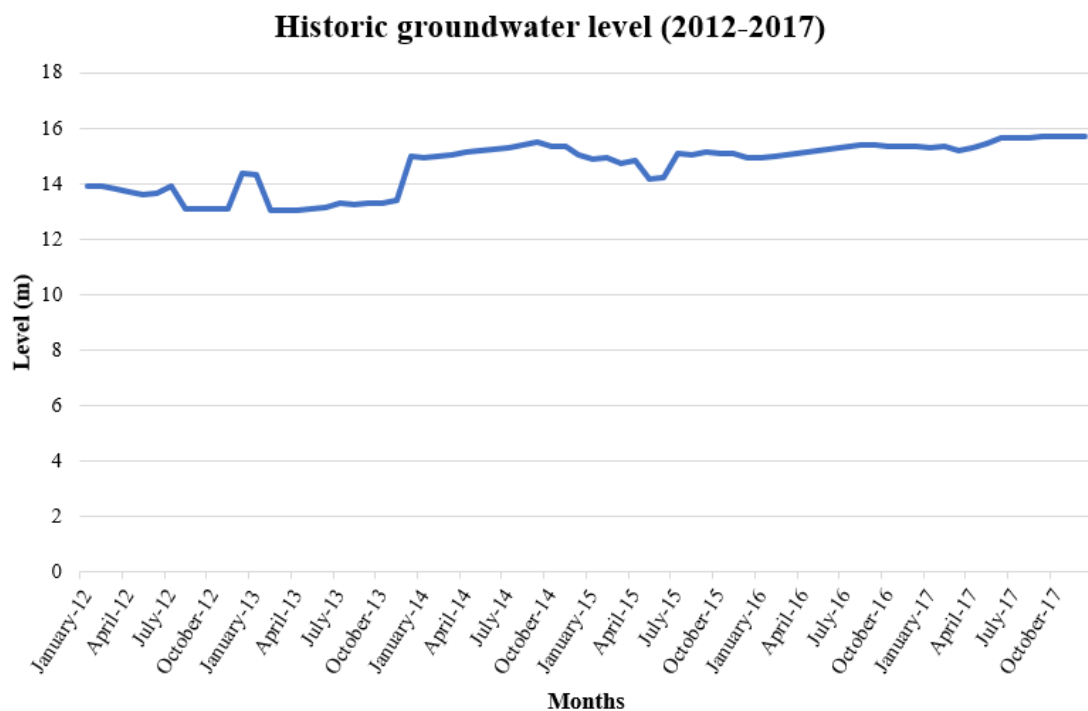


Figure 3. Historic groundwater level of the well ($x=606797$, $y=80384$).

Secondly, preparation of the environment was followed, mainly Python installation, packages of XGBoost were also installed, as well as relevant dependencies. After that, the dataset was loaded and converted to DMatrix format. It is an efficient data structure that the inventors of XGBoost created, and it is responsible for the package's improved performance and efficiency. Furthermore, the dataset was split into a training and a testing dataset. The training set will comprise a known output, from which the model will learn, and the testing set will be utilized to validate and test the model's predictions. Thirdly, the XGBoost algorithm was launched to learn from the training set and applied to make predictions on the testing data, then RMSE and R^2 were measured. Lastly, the XGBoost algorithm was applied to make level predictions of the year 2018 after it had been learned and tested from groundwater level data of the period 2012-2017, so the groundwater levels for the year 2018 were predicted.

3. Results

The training and test phase of the algorithm has resulted in acceptable values of RMSE=0.312626 and $R^2=0.936817$, because lower RMSE indicates higher model accuracy, and the closer R^2 is to 1, the better the model is at predicting groundwater levels. Figure 5 shows the comparison between predicted and original data of groundwater level regarding the test set; it also illustrates good accuracy of the model, since for most of the points, the difference is not noticeable. After building an accurate model, it was used to predict groundwater level values of the year 2018, and the result is shown in Figure 6. Based on that graph (Figure 6), the groundwater level trend is increasing, which means that the Tafilalet aquifer is in continuous depletion.

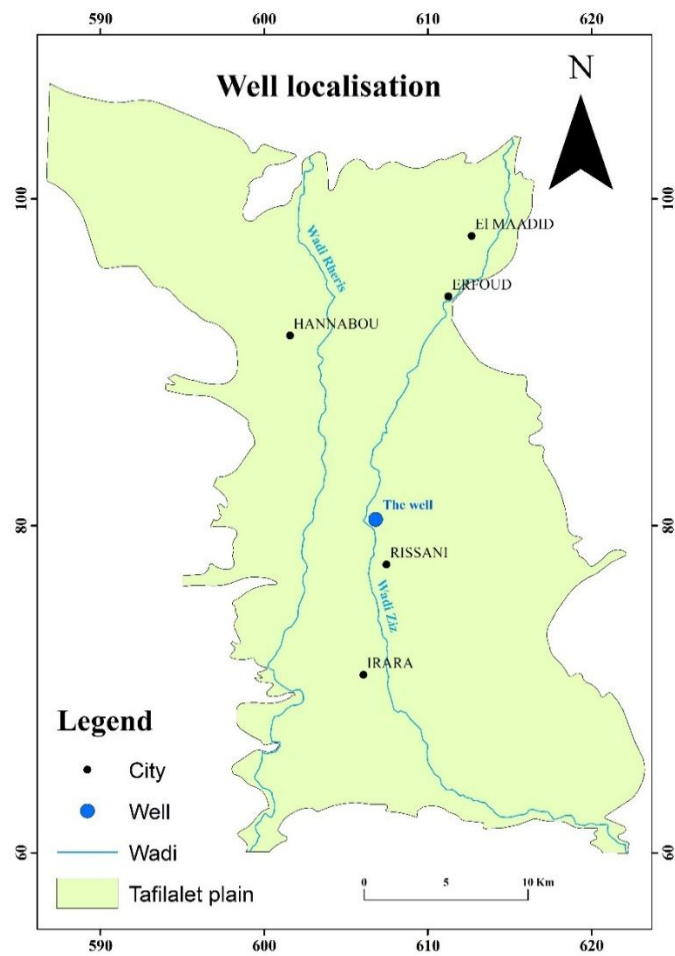


Figure 4. Localization of the well.

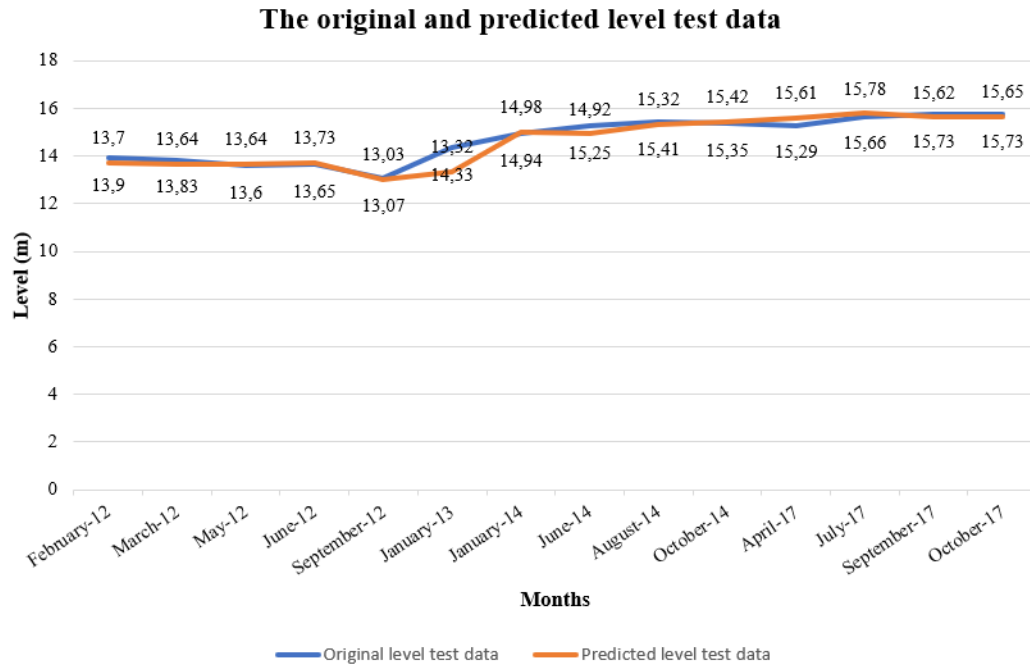


Figure 5. Comparison of the original and predicted level test data.

4. Discussion

After using the model to generate groundwater levels for one year, it has also been used to predict groundwater levels for more than one year, but after trying, it has been noticed that beyond one year, the same predictions are repeated, and the same graph in Figure 6 is repeated. This is quite normal since XGBoost requires a large amount of data in the learning phase to be able to predict over a large period of time. So, extending the prediction horizon led to repetitive prediction patterns. This limitation highlights the challenge of obtaining long-term groundwater level monitoring data from relevant institutions and administrative bodies, which remains a constraint of this work.

Moreover, several inputs can be used in addition to the historical water level data to have a better performance and predictions much closer to reality. Future research will therefore focus on enriching the input dataset by integrating meteorological data and increasing the temporal extent of the training data. This is expected to improve the model's ability to simulate groundwater variability over longer periods and to better support sustainable groundwater management strategies in the Tafilalet aquifer.

The predictions obtained are very valuable for groundwater management as they will help to get an idea of future trends of groundwater level, which will guide water managers and stakeholders to take appropriate actions and measures. Groundwater level forecasts are beneficial for operational decision-making in regions like the Tafilalet plain, where water is a scarce resource. As an example, these predictions can be of great help in evaluating the impact of artificial aquifer recharge on groundwater level in the Tafilalet aquifer, as there is an ongoing project of artificial recharge of the Tafilalet aquifer. Consequently, the integration of machine learning-based forecasting into groundwater management frameworks represents a promising pathway toward maintaining aquifer health, ensuring long-term water security, and promoting sustainable development.

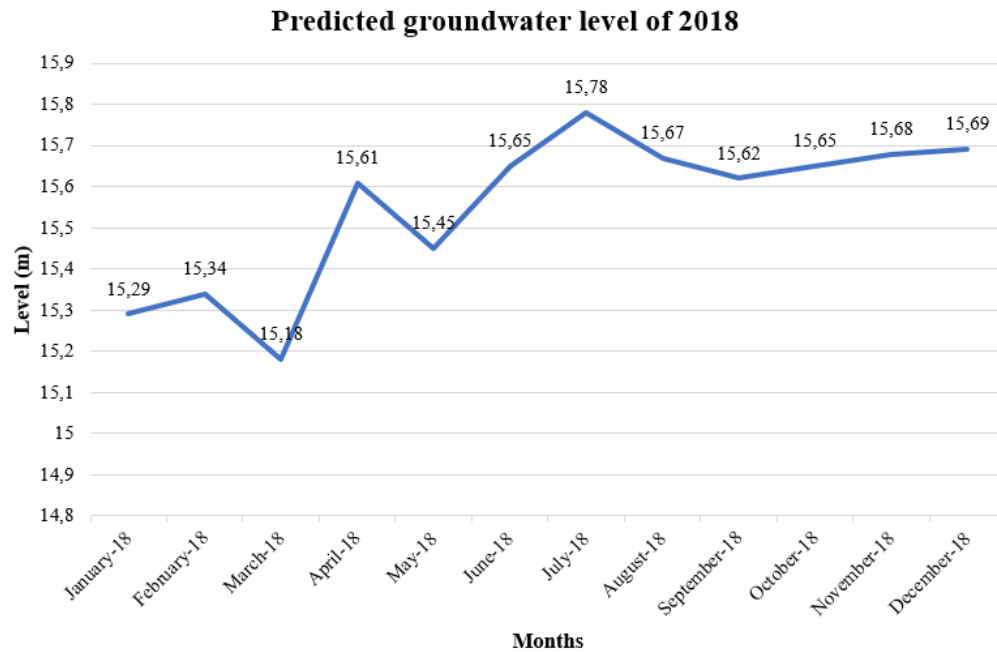


Figure 6. Predicted groundwater level of the year 2018.

5. Conclusions

The machine learning algorithm XGBoost was applied to make predictions of the water level of the aquifer of Tafilalet for a period of one year (2018). The results of this study are limited to one year, because beyond one year, the same result is repeated, which is due to several factors, among others, the amount of data used for the learning phase was not large enough. So, to improve and refine the predictions, other inputs can be used as meteorological data, which is our aspiration in the future. In an upcoming study, this algorithm will be filled with weather data (precipitation and temperature) in order to boost the result and to see what is the impact of adding weather data on the predictions.

Despite the above-mentioned limitation, the outcome of this research is important, as accurate groundwater level forecasting is a key component of sustainable aquifer management. By enabling early detection of declining trends and supporting informed decision-making, the proposed approach can assist water managers in regulating abstraction, improving resource allocation, and mitigating the long-term risks of overexploitation, which will contribute to the sustainable development of groundwater resources and ensure their availability for future generations while supporting socio-economic stability in groundwater-dependent areas.

References

- Ardana, P.D.H., Redana, I.W., Yekti, M.I., Simpen, I.N., 2021. Prediction of Groundwater Level in The Shallow Aquifer Using Artificial Neural Network Approach. IOP Conf. Ser.: Earth Environ. Sci. 837, 012010. <https://doi.org/10.1088/1755-1315/837/1/012010>
- Bhati, B.S., Chugh, G., Al-Turjman, F., Bhati, N.S., 2021. An improved ensemble based intrusion detection technique using XGBoost. Trans Emerging Tel Tech 32. <https://doi.org/10.1002/ett.4076>

- Chen, C., He, W., Zhou, H., Xue, Y., Zhu, M., 2020. A comparative study among machine learning and numerical models for simulating groundwater dynamics in the Heihe River Basin, northwestern China. *Sci Rep* 10, 3904. <https://doi.org/10.1038/s41598-020-60698-9>
- Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System, in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Presented at the KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, San Francisco California USA, pp. 785–794. <https://doi.org/10.1145/2939672.2939785>
- Chitsazan, M., Rahmani, G., Neyamadpour, A., 2015. Forecasting groundwater level by artificial neural networks as an alternative approach to groundwater modeling. *J Geol Soc India* 85, 98–106. <https://doi.org/10.1007/s12594-015-0197-4>
- Dong, W., Huang, Y., Lehane, B., Ma, G., 2020. XGBoost algorithm-based prediction of concrete electrical resistivity for structural health monitoring. *Automation in Construction* 114, 103155. <https://doi.org/10.1016/j.autcon.2020.103155>
- El Naqa, I., Murphy, M.J., 2015. What Is Machine Learning?, in: El Naqa, I., Li, R., Murphy, M.J. (Eds.), *Machine Learning in Radiation Oncology*. Springer International Publishing, Cham, pp. 3–11. https://doi.org/10.1007/978-3-319-18305-3_1
- Guezgouz, N., Boutoutaou, D., Hani, A., 2021. Prediction of groundwater flow in shallow aquifers using artificial neural networks in the northern basins of Algeria. *Journal of Water and Climate Change* 12, 1220–1228. <https://doi.org/10.2166/wcc.2020.067>
- Ibrahim Ahmed Osman, A., Najah Ahmed, A., Chow, M.F., Feng Huang, Y., El-Shafie, A., 2021. Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. *Ain Shams Engineering Journal* 12, 1545–1556. <https://doi.org/10.1016/j.asej.2020.11.011>
- Ifrah, G., 2001. *The universal history of computing: from the abacus to the quantum computer*. Wiley, New York Weinheim.
- Kenda, K., Čerin, M., Bogataj, M., Senožetnik, M., Klemen, K., Pergar, P., Laspidou, C., Mladenčić, D., 2018. Groundwater Modeling with Machine Learning Techniques: Ljubljana polje Aquifer. *Proceedings* 2, 697. <https://doi.org/10.3390/proceedings2110697>
- Ławrynowicz, A., Tresp, V., 2014. Introducing Machine Learning, in: *Perspectives on Ontology Learning*. pp. 35–50.
- Liu, Yang, Wang, H., Fei, Y., Liu, Ying, Shen, L., Zhuang, Z., Zhang, X., 2021. Research on the Prediction of Green Plum Acidity Based on Improved XGBoost. *Sensors* 21, 930. <https://doi.org/10.3390/s21030930>
- Montáns, F.J., Chinesta, F., Gómez-Bombarelli, R., Kutz, J.N., 2019. Data-driven modeling and learning in science and engineering. *Comptes Rendus Mécanique* 347, 845–855. <https://doi.org/10.1016/j.crme.2019.11.009>
- Nayak, P.C., Rao, Y.R.S., Sudheer, K.P., 2006. Groundwater Level Forecasting in a Shallow Aquifer Using Artificial Neural Network Approach. *Water Resour Manage* 20, 77–90. <https://doi.org/10.1007/s11269-006-4007-z>
- Pandey, K., Malik, A., 2022. Seasonal Groundwater Table Depth Prediction Using Fuzzy Logic and Artificial Neural Network in Gangetic Plain, India, in: Rao, C.M., Patra, K.C., Jhaharia, D., Kumari, S. (Eds.), *Advanced Modelling and Innovations in Water Resources Engineering, Lecture Notes in Civil Engineering*. Springer Singapore, Singapore, pp. 549–564. https://doi.org/10.1007/978-981-16-4629-4_37
- Rajae, T., Ebrahimi, H., Nourani, V., 2019. A review of the artificial intelligence methods in groundwater level modeling. *Journal of Hydrology* 572, 336–351. <https://doi.org/10.1016/j.jhydrol.2018.12.037>
- Rakhshandehroo, G.R., Vaghefi, M., Aghbolaghi, M.A., 2012. Forecasting Groundwater Level in Shiraz Plain Using Artificial Neural Networks. *Arab J Sci Eng* 37, 1871–1883. <https://doi.org/10.1007/s13369-012-0291-5>

- Samuel, A.L., 1959. Some Studies in Machine Learning Using the Game of Checkers. *IBM J. Res. & Dev.* 3, 210–229. <https://doi.org/10.1147/rd.33.0210>
- Solomatine, D., See, L.M., Abrahart, R.J., 2008. Data-Driven Modelling: Concepts, Approaches and Experiences, in: Abrahart, Robert J., See, Linda M., Solomatine, D.P. (Eds.), *Practical Hydroinformatics, Water Science and Technology Library*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 17–30. https://doi.org/10.1007/978-3-540-79881-1_2
- Sun, Y., Wendi, D., Kim, D.E., Liong, S.-Y., 2015. Technical Note: Application of artificial neural networks in groundwater table forecasting – a case study in Singapore swamp forest (preprint). *Groundwater hydrology/Stochastic approaches*. <https://doi.org/10.5194/hessd-12-9317-2015>
- Taormina, R., Chau, K., Sethi, R., 2012. Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon. *Engineering Applications of Artificial Intelligence* 25, 1670–1676. <https://doi.org/10.1016/j.engappai.2012.02.009>
- Wunsch, A., Liesch, T., Broda, S., 2021. Groundwater level forecasting with artificial neural networks: a comparison of long short-term memory (LSTM), convolutional neural networks (CNNs), and non-linear autoregressive networks with exogenous input (NARX). *Hydrol. Earth Syst. Sci.* 25, 1671–1687. <https://doi.org/10.5194/hess-25-1671-2021>
- Yin, W., Fan, Z., Tangdamrongsub, N., Hu, L., Zhang, M., 2021. Comparison of physical and data-driven models to forecast groundwater level changes with the inclusion of GRACE – A case study over the state of Victoria, Australia. *Journal of Hydrology* 602, 126735. <https://doi.org/10.1016/j.jhydrol.2021.126735>
- Zhang, X.-D., 2020. Machine Learning, in: *A Matrix Algebra Approach to Artificial Intelligence*. Springer Singapore, Singapore, pp. 223–440. https://doi.org/10.1007/978-981-15-2770-8_6