

Association for Information Systems

AIS Electronic Library (AISeL)

Proceedings of the 2020 Pre-ICIS SIGDSA
Symposium

Special Interest Group on Decision Support and
Analytics (SIGDSA)

Winter 12-13-2020

Predicting Groundwater Fluctuations in Hard Rock Watersheds – An Application of Data Visualizations and Machine Learning Algorithms

Matthew Stocker

Lakshmi Iyer

Basant Maheshwari

Ramesh Sharda

Follow this and additional works at: <https://aisel.aisnet.org/sigdsa2020>

This material is brought to you by the Special Interest Group on Decision Support and Analytics (SIGDSA) at AIS Electronic Library (AISeL). It has been accepted for inclusion in Proceedings of the 2020 Pre-ICIS SIGDSA Symposium by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Predicting Groundwater Fluctuations in Hard Rock Watersheds – An Application of Data Visualizations and Machine Learning Algorithms

Research-in-Progress – Extended Abstract

Matthew Stocker

Appalachian State University
stockermr@appstate.edu

Basant Maheshwari, Ph.D.

Western Sydney University
B.Maheshwari@westernsydney.edu.au

Lakshmi Iyer, Ph.D.

Appalachian State University
iyerls@appstate.edu

Ramesh Sharda, Ph.D.

Oklahoma State University
ramesh.sharda@okstate.edu

Groundwater is a resource used in a variety of countries with sparse rainfall throughout the year. India, as a whole, uses groundwater for 60% of their irrigation water for crops and 80% of their drinking water. India receives rain almost exclusively during the monsoon season of June through the end of September. The Rajasthan state, in the north-western region of India, receives slightly less rainfall than the southern parts. As a result, its villages often are at a far greater risk of experiencing drought and famine, due to a shortage of water for irrigation and self-sufficiency.

In addressing these challenges, many farmers prefer to drill deeper towards groundwater to allow them to water their crops more generously. Unfortunately, this is a short-term solution as groundwater has a limited amount of water recharge per year, depending on climatic conditions, rainfall, well porosity and a variety of other factors. This has led to growing concern over sustainable water management practices in order to prevent endangering future irrigation and drinking water availability.

The Managing Aquifer Recharge and sustaining Groundwater Use through Village-Level Intervention (MARVI) project aims to educate villagers on their water use, so that they can manage their supply more sustainably (Maheshwari et al. 2014). This project brings cross-disciplinary aspects of the problem together to achieve sustainable groundwater use while realizing improved livelihood outcomes for village communities. The secondary focus is to be a resource for policymakers in regards to groundwater availability. The primary research questions of this MARVI project are:

1. What are the patterns of groundwater changes in the wells over time, across villages? How are these patterns related to parameters such as rainfall?
2. What is the minimum number of wells that should be targeted for data collection to get a sense of an aquifer's behavior? In other words, what is the minimum number of wells that will give farmers a good understanding of the water level fluctuations around their village? How does the water level in their wells change based on rainfall amounts during the monsoon season?
3. How often should well data be collected to get a sense of an aquifer's behavior?

This research in progress addresses the first research question through analysis of data made available. Predicting groundwater fluctuations has grown in prevalence as global concerns rise over water shortages. This study uses rainfall patterns, and geographical patterns as predictors, because these factors play a large role in analyzing shifts in groundwater levels. For this study, we focus on 5 villages in the northwestern part of India which are spread out over roughly a 24 sq. mile area of hard rock area. There were 50 wells monitored weekly per village for a total of 250 wells in the watershed since 2012, but they were reduced to 30 per village with 150 total in 2017. The water readings and rainfall recordings are conducted by villagers trained to monitor the groundwater and rainfall gauges.

This research explores using supervised machine learning (ML) algorithms to predict and better understand groundwater fluctuations in these hard rock aquifers. ML methods allow for non-linear approaches to

modeling these complex interactions within nature, which would otherwise require a large amount of resources to calculate manually. The two methods at the forefront of this field of research are Artificial Neural Networks (ANN), based on neural networks in the brain, and support vector machines (SVM), based on statistical learning theory (Hajji et al. 2012; Yoon et al. 2017). ANNs have been found to perform more efficiently than SVM models in regards to groundwater fluctuations, however, SVMs were found to be less sensitive to the structure of input data, allowing it to better handle more diverse data (Zhou et al. 2017).

While the field is still actively changing, “the most obvious input variables in groundwater level predictions studies are rainfall, evaporation, temperature and pumping patterns” (Yadav et al. 2019). Normally, pumping patterns and evapotranspiration levels would be explicitly recorded, however, due to limited resources, the focus of the study is to test how accurately groundwater level could be predicted using the variables in Figure 1. As a result, these variables are implicitly a part of the date variable. However, due to the self-reported nature of rainfall data collection, the rainfall data was reported more sporadically than expected, given the monsoon season weather.

Variable	Explanation
Date	The date the well was checked on a weekly basis.
Longitude	Longitude of the well
Rainfall	Weekly Total Rainfall as recorded by farmers in millimeters
Elevation	Elevation of the wells, recorded in meters
Latitude	Latitude of the wells
El Niño3 SST	Surface Sea Temperature readings from the Niño3 region that are used to predict El Nino and La Niña and monitor their effects. This data is recorded as a monthly average in °C.
Temperature	Temperature readings for the state of Rajasthan recorded monthly

Table 1. Variables used in ML algorithms

Figure 1 below shows the location of the wells through [interactive visualization](#) created using Google maps. When a user clicks on it, provides information about each well such as owner information, the village located in, elevation and the depth of the well.

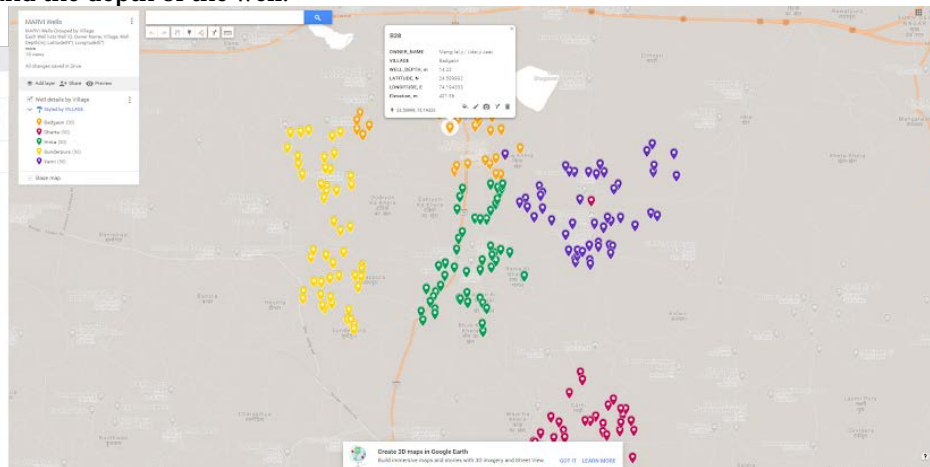


Figure 1. Map providing the wells location and information.

In evaluating model performance, two key metrics were identified, Root Mean Square Error (RMSE), and R². RMSE is the average magnitude of errors, which results in larger errors being given greater weight with a range from 0 to infinity (Shiri et al. 2013). Furthermore, RMSE is the primary performance evaluation metric and functions as a measure of accuracy for prediction models to allow for comparison. The coefficient of determination (R²) is a secondary performance indicator, evaluating the “completeness” of the model.

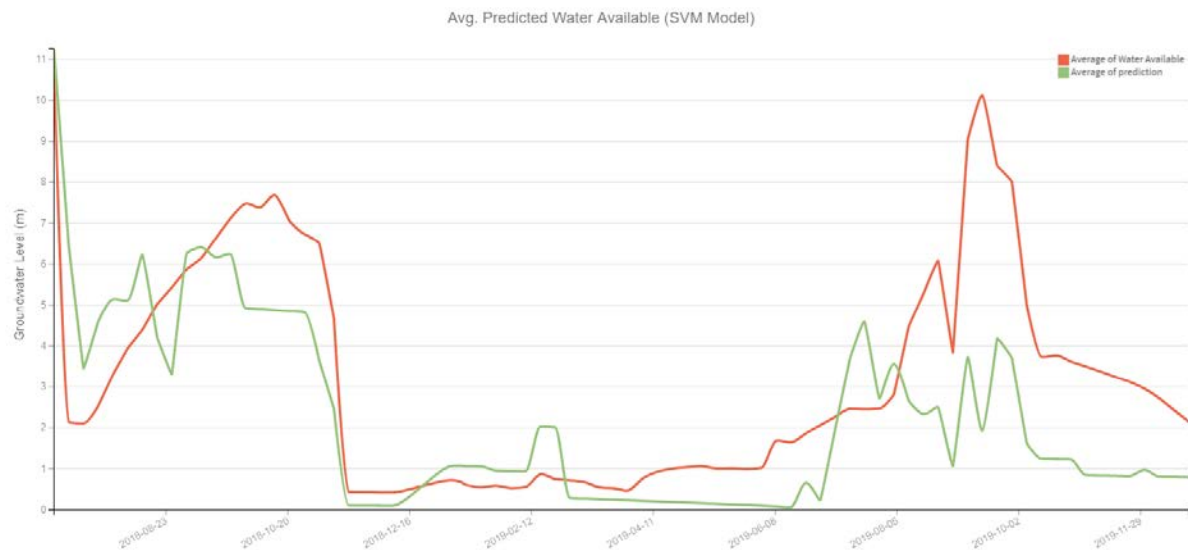


Figure 2. Predicted Groundwater levels compared to the actual groundwater levels using the SVM Model

Initial findings revealed that the SVM model performed the best with a RMSE of 5.4629 and a R2 score of 0.31316. This particular SVM model used a radial basis function (RBF) kernel, with a tolerance of 0.001. The ANN model performed significantly worse with a RMSE score of 6.6080 and R2 score of -0.0049547. Clearly, more data is needed in order to more accurately predict groundwater levels. In Figure 2, the groundwater levels are consistently underpredicted by the SVM Model. As this research is still ongoing, additional weather variables and alternative methods of pre-processing will be introduced in further iterations. The next steps of this project will be an optimization of the well monitoring network to reduce the number of wells monitored while still maintaining an accurate view on the available groundwater and explore how frequently wells should be monitored to maintain this accurate view. These optimizations will reduce the costs and allow easier adoption of this program across the area, and ease the burden on participating villages.

Acknowledgements

This research was conducted under the Managed Aquifer Recharge through Village-level Intervention (MARVI) project, which was funded by the Australian Centre for International Agricultural Research and Australian Water Partnership. Authors thank the MARVI team for the data collection, particularly the farmer volunteers (called *Bhujal Jaankaars*), Dr. Yogita Dashora, Mr. Prahlad Soni, Dr. P.K. Singh, Dr. Yogesh Jadeja and Dr. Peter Dillon.

Selected References

- Hajji, S., Hachicha, W., Bouri, S., & Dhia, H. B. (2012). Spatiotemporal groundwater level forecasting and monitoring using a neural network-based approach in a semi arid zone. *International Journal of Hydrology Science and Technology*, 2(4), 342-361.
- Maheshwari, B., M. Varua, J. Ward, R. Packham, P. Chinnasamy, Y. Dashora, S. Dave, P. Soni, P. Dillon, R. Purohit, Hakimuddin, T. Shah, S. Oza, P. Singh, S. Prathapar, A. Patel, Y. Jadeja, B. Thaker, R. Kookana, H. Grewal, K. Yadav, H. Mittal, M. Chew, P. Rao 2014. The role of transdisciplinary approach and community participation in village scale groundwater management: Insights from Gujarat and Rajasthan, India. *Int Open Access J Water*, 6(6) pp. 3386-3408.
http://www.mdpi.com/journal/water/special_issues/MAR
- Rodell, M., Velicogna, I., & Famiglietti, J. S. (2009). Satellite-based estimates of groundwater depletion in India. *Nature*, 460(7258), 999-1002.
- Yadav, B., Gupta, P. K., Patidar, N., & Himanshu, S. K. (2020). Ensemble modelling framework for groundwater level prediction in urban areas of India. *Science of The Total Environment*, 712, 135539.
- Zhou, T., Wang, F., & Yang, Z. (2017). Comparative analysis of ANN and SVM models combined with wavelet preprocess for groundwater depth prediction. *Water*, 9(10), 781.