Balaban Desalination Publications

مؤتمر الخليج الخامس عشر للمياه
The 15th Gulf Water Conference

(2025) 81–83
doi: 10.5004/700105

WSTA 15th Gulf Water Conference
Water in the GCC, The Role of Technology in Effective Water Management
28–30 April 2024, Doha, Qatar

Artificial intelligence application in hydrogeology and groundwater management

Khalid ElHaj[1,2], Dalal Alshamsi[1,2]

[1]Department of Geosciences, United Arab Emirates University, Al Ain, P.O. Box 15551, UAE
[2]National Water and Energy Center, United Arab Emirates University, Al Ain, P.O. Box 15551, UAE
emails: 200734761@uaeu.ac.ae (K. ElHaj); dalal.shamsi@uaeu.ac.ae (D. Alshamsi)

EXTENDED ABSTRACT

Groundwater management involves overseeing groundwater resources to ensure their ideal utilization and long-term sustainability. More than two billions of the world population depend on groundwater resources as their main water source (Famiglietti, 2014), as result, managing such an important resources would be of paramount importance to governments. This would be even more important for countries like UAE where groundwater provides about 51% of the total water budget (Al Rashed et al., 2023; Ministry of Environment & Water, 2014). Managing groundwater resources involves many aspects, including observing its usage and predicting its future utilization, evaluating the scenarios affecting its utilization and any kind of adverse effects it can have on public health. In addition to aquifer delineation, pollution indicators, and any kind of relation the different aquifers may maintain with each other.

Our work involved exploring many of the key aspects of groundwater management. We utilized state-of-the-art Artificial Neural Networks to predict the fluctuations in Al Ain city groundwater levels. The algorithm utilized state-of-the-art Artificial Neural Networks that used the location, time and meteorologic conditions to predict the groundwater level in any location within Al Ain city. It achieved an accuracy of 0.952 in Coefficient of Determination (R^2). The trained model was used in conjunction with a geomodelling framework that utilized GIS to create variability maps to visually highlight the changes with time in a spatial form (ElHaj et al., 2023b)

We also investigated the critical issue of rising groundwater salinity using machine learning techniques. In a recent study submitted to COP28, we developed predictive models trained on historical total dissolved solids (TDS) data from multiple wells across Al Ain. The models were then used to forecast potential future changes in groundwater salinity under business-as-usual conditions. Additionally, we leveraged the trained models to simulate salinity responses under various climate scenarios, including substantial increases or decreases in precipitation and temperature. This provided actionable insights into how groundwater quality could evolve under different climate futures, enabling proactive management strategies to mitigate salinization risks. Overall, our integrated modeling approach generated data-driven projections to support evidence-based policies aimed at sustaining fresh groundwater resources in the face of environmental change.

We developed an integrated machine learning workflow to elucidate relationships between aquifers using their unique hydrogeochemical fingerprints. The approach involved first leveraging supervised learning

*Corresponding authors

ISBN 0-86689-078-5

techniques to train models on the available hydrochemical data. These models were then used to impute any missing values in the dataset — a crucial step enabling robust subsequent analysis. We next utilized unsupervised clustering algorithms to objectively group wells based solely on similarities in their hydrochemical makeup. This delineated distinct groundwater bodies in a completely data-driven manner. The clustered wells were then visualized spatially using our custom GeoZ library for visualizing clustering outputs, revealing meaningful spatial patterns aligning with the hydrostratigraphy (ElHaj et al., 2023a). By integrating supervised learning for imputation with unsupervised clustering for classification, the workflow provided an automated means of elucidating connections between aquifers based on their hydrogeochemical signatures. The approach extracted new hydrogeological insights, including groundwater mixing, flow paths, recharge sources, and more. Overall, it demonstrated the power of artificial intelligence techniques to uncover meaningful patterns in complex and imperfect hydrochemical data.

Accurately delineating aquifer boundaries provides numerous critical benefits that support sustainable groundwater management and utilization. Well-mapped aquifers empower decision-makers to effectively monitor, regulate, and protect groundwater resources. Aquifer delineation further aids pollution prevention efforts, disaster preparedness planning, informed land use development, efficient infrastructure sitting, and environmental conservation. Additionally, clearly defined aquifer boundaries enable cooperative legal and transboundary management of aquifers spanning multiple jurisdictions. Aquifer mapping also facilitates economic development planning that accounts for available groundwater resources. It can even assist geothermal energy exploration by elucidating subsurface hydrogeological structures. Overall, precise aquifer delineation is an essential foundation that facilitates informed, integrated groundwater management from local to regional scales.

To address the need for accurate yet efficient aquifer delineation, we developed a novel data-driven methodology leveraging hierarchical clustering analysis of groundwater hydrographs. The approach depends solely on leveraging existing hydrograph data from monitoring wells to delineate aquifer boundaries, without requiring expensive and time-consuming geophysical surveys. A key novelty lies in the custom distance metric we formulated to specifically address unique aspects of hydrogeological time-series data. This custom function compares the hydrographs from different wells based on similarities in groundwater level fluctuations over time. Wells with similar hydrographs are grouped into clusters using agglomerative hierarchical clustering. We selected this algorithm for two main reasons; first, agglomerative hierarchical clustering had the most flexibility in tuning among all the clustering methods tested during the project first phase, this allowed us to customize it to work more efficiently with hydrogeological datasets. The second reason we chose agglomerative hierarchical clustering was due to its hierarchical nature when choosing the number of clusters. We utilized this characteristic to address the issue of aquifer heterogeneity.

The algorithm is embedded within an overarching optimization framework enabling systematic tuning of parameters for optimal performance. This allowed further flexibility in tuning the preprocessing aspect of the data pipeline in addition to manipulating the clustering algorithm parameters. Creating the method in this way allowed us to test numerus scenarios, different methods, and use different metrics to assess the model performance. Not to mention that this approach also made the model future proof, as introducing any new methods within the framework would be done with minimal effort and without any major change to the main codebase. The output of the clustering algorithm would be general clusters containing the most similar wells. We developed GeoZ to map these clusters into geographic maps and used these maps to delineate the aquifer boundaries based on the available data.

The last step of the project was to enclose the developed framework within an objective function and run it in an optimization algorithm to find the best parameters to achieve the highest accuracy in addition to find the limits of the requirement to operate the model on any area of interest. The optimization algorithm achieved 90% accuracy in aquifer delineation using a monthly average of synthetic observation data that spanned a minimum of 400 months (33 years and 4 months). Using the same observation period limit on real groundwater observation data obtained from Texas state aquifers resulted in 73% accuracy. Longer hydrograph records improve accuracy but reduce the number of usable wells. Striking an appropriate balance is key to maximizing delineation while utilizing the most data.

The method can be optimized in many ways, including new imputation methods, outlier detection, similarity calculations among others, and the framework approach would make implementing them much easier. Also, the method was tested on monthly average regional scale datasets. With the advancement of groundwater observations methods and the new continuous telemetric observation equipment introduced in the Gulf Cooperation Council (GCC) Countries, this approach might provide new insights with more granular data with local and regional scale aquifers in the middle east. The efficiency of the method has not yet been tested on local aquifers with more granular data such as local dataset with daily or hourly readings, so it could be a very good research opportunity to explore the feasibility of the method in this aspect.

Overall, the methodology provides an inexpensive and efficient means of extracting value from existing hydrograph data to map aquifer boundaries. Precisely delineated aquifers empower communities and governments to enact informed, sustainable groundwater management practices. The work highlights the power of machine learning to advance large-scale hydrogeological characterization from imperfect data sources.

Keywords: Groundwater management; Machine learning; Aquifer delineation; Groundwater salinity; Artificial neural networks; Hydrogeochemistry; Hydrochemical fingerprints; Hierarchical clustering; Geomodelling

References

Al Rashed, M., Sefelnasr, A., Sherif, M., Murad, A., Alshamsi, D., Aliewi, A., Ebraheem, A.A., Novel concept for water scarcity quantification considering nonconventional and virtual water resources in arid countries: Application in Gulf Cooperation Council countries. Sci. Total Environ. 882 (2023) 163473. https://doi.org/10.1016/j.scitotenv.2023.163473

ElHaj, K., Alshamsi, D., Aldahan, A., . GeoZ: a region-based visualization of clustering algorithms. J. Geovisualization Spat. Anal. 7 (2023a) 15. https://doi.org/10.1007/s41651-023-00146-0

ElHaj, K., Issa, S., Alshamsi, D., Cherkose, B.A., Modeling and prediction of groundwater level fluctuations using geoinformatics and artificial neural networks in Al Ain City, UAE. in Water Resources Management and Sustainability, 2023b, pp. 205–217. https://doi.org/10.1007/978-3-031-245060_13

Famiglietti, J.S., The global groundwater crisis. Nat. Clim. Chang. 4 (2014) 945–948. https://doi.org/10.1038/nclimate2425

Ministry of Environment & Water, Hydroatlas United Arab Emirate, 2014.