

AI-Based Forecasting of Groundwater Storage Depletion Using GRACE-GLDAS Data

Mohammed S. Al Nadabi, Mohammed El-Diasty, Talal Etri and Mohammad Reza Nikoo

Department of Civil and Architectural Engineering
Sultan Qaboos University
Muscat, Oman

s48708@student.squ.edu.om, m.eldiasty@squ.edu.om, t.etri1@squ.edu.om, m.reza@squ.edu.om

Abstract

The Najd region of the Sultanate of Oman experiences critical depletion in its groundwater resources due to the hyper-arid climate, rapid population growth, and high demand for agricultural water. In this research, six deep learning and machine learning models are used to estimate Groundwater Storage Anomaly (GWSA) in the Najd region based on monthly Terrestrial Water Storage Anomaly (TWSA) data obtained from the Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow-On (GRACE-FO) satellite missions and hydroclimatic variables of the Global Land Data Assimilation System (GLDAS) from April 2002 to February 2026. The models are CNN, ERNN, GRU, LSTM, MLP, and TCN. The mass balance approach is adopted to isolate GWSA by removing soil moisture anomalies from TWSA. Linear interpolation is used to fill 35 missing GRACE months, resulting in a complete 287-month continuous time series. The analysis is limited to the 123 GLDAS cell centers located inside the Najd boundary. These cells represent 14 independent CSR mascon-signal groups, divided into 10 training groups (87 cells), 2 validation groups (18 cells), and 2 testing groups (18 cells), with no group overlap. Eight GLDAS variables are retained, in addition to latitude, longitude, normalized time, and cyclical month. The basin analysis indicates a statistically significant declining GWSA trend of -0.171 cm/year. Among all models, CNN achieves the best independent spatial-validation performance with MAE = 0.902 cm, RMSE = 1.131 cm, $R^2 = 0.798$, and KGE = 0.865. The findings provide useful information for groundwater monitoring in the Najd region and similar hyper-arid environments.

Keywords

Artificial Intelligence, Deep Learning, GRACE, GLDAS, Groundwater Storage

1. Introduction

Groundwater is a vital natural resource that is used by many communities for drinking, agricultural, and livestock purposes (Pulla et al., 2023). However, rapid population growth coupled with accelerating climate change is pushing fresh groundwater systems to the limit in most arid and hyper-arid regions where water resources are already scarce (Missik et al., 2021). The Najd area is located in the Dhofar Governorate of the Sultanate of Oman. This region is one of the driest regions in the world where groundwater occurs in four separate aquifers, A to D, within the Hadhramaut Group mostly consisting of formations of limestone and dolomite. The approximate area of the Najd is about 90,000 km². It is situated to the north of the Dhofar Mountains extending deep into Rub Al-Khali desert and is bordered by Yemen on the west, Saudi Arabia on the north, and the Al-Wusta Governorate on the east (Al-Mashaikhi et al., 2012).

The recent launch of the GRACE (Gravity Recovery and Climate Experiment) satellite mission has made it possible to monitor terrestrial water storage changes at large scales remotely, using a good proxy for variation in groundwater storage when combined with land surface models such as the Global Land Data Assimilation System (GLDAS) (Rzepecka and Birylo, 2020). The use of ML methodologies in the analysis of hydrological time series has gained much popularity recently. Over the years, methods like Support Vector Machines (SVM), Artificial Neural Networks

(ANN), Autoregressive Integrated Moving Average (ARIMA), and Random Forests (RF) have been widely used in groundwater level forecasting for different hydroclimatic regimes (Feng et al., 2024). More recently, different deep learning architectures such as Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRU), Elman Recurrent Neural Networks (ERNN), Multilayer Perceptron (MLP), Convolutional Neural Networks (CNN), and Temporal Convolutional Networks (TCN) have shown better performance in capturing complicated, nonlinear, temporal patterns within groundwater storage time series compared to traditional statistical methods (Khan et al., 2023). While groundwater decline in the Najd basin is well-recognized, an integrated framework combining satellite-based GRACE/GLDAS data with advanced techniques like deep learning to predict storage depletion has not yet been extensively explored. In the context of Najd being the main source of water supply to its surrounding domestic and agricultural population, there is a high necessity to assess the performance of various deep learning models in predicting groundwater storage anomalies (GWSA) from satellite-based datasets. This research attempts to fill this gap by providing a forecast of groundwater storage change in the Najd region, Oman, using six different deep learning models (MLP, ERNN, LSTM, GRU, CNN, and TCN) for comparative performance evaluation and recommending the best model configuration in terms of accuracy and computational efficiency for groundwater prediction in hyper-arid and data-scarce conditions.

1.1 Objectives

This study aims to fill this gap by using GRACE and GLDAS satellite data to develop a model for the estimation of groundwater storage changes within the Najd boundary and determine the optimal deep learning model for this hyper-arid and data-scarce region. The specific objectives of the study are:

1. Extract TWSA data for the Najd region from the GRACE satellite and surface water component data from GLDAS. To develop GWSA for the Najd region using a mass balance approach.
2. Estimate groundwater storage changes within the Najd boundary with multiple deep learning models: MLP, ERNN, LSTM, GRU, CNN, and TCN.
3. Evaluate and compare the performance of these deep learning models using standard error metrics: MSE, RMSE, MAE, and R^2 .

2. Literature Review

The use of ML and DL techniques in GWL forecasting has greatly increased in the past few years, mainly due to the increasing availability of satellite-based hydrological datasets and the greater need for low-cost monitoring solutions in regions with little data. The most popular methods used to forecast groundwater dynamics in different hydroclimatic regions include SVM, ANN, ARIMA, ANFIS, and RF. Takafuji (2019) worked in the Bauru Aquifer System in São Paulo State, Brazil, and made a groundwater availability forecast by extreme dry years of the 2013/2014 hydrological cycle affecting groundwater table levels using time series autoregressive integrated moving average (ARIMA) and geostatistical Sequential Gaussian Simulation (SGS). Sattar and Khalid (2021) estimated variations of groundwater storage in the Indus River Basin using GRACE satellite observations along with hydrological data from GLDAS. Though the study was more estimation-centric than forecasting, the authors developed a methodology to separate the variations of groundwater storage from changes in total water storage. The results were able to capture seasonal fluctuations in groundwater and also identify a long-term trend of depletion of groundwater between 2005 and 2015.

In a study conducted by Kanito et al. (2025), different machine learning models were evaluated for the prediction of groundwater levels in Ruataniwha Basin, New Zealand. These models include Random Forest (RF) and Support Vector Machine (SVM). Input data used in the study was based on groundwater observations, seasonal components, and hydrogeological characteristics to predict the dynamic behavior of groundwater levels at six monitoring wells. Results showed that the machine learning models were able to capture groundwater fluctuations, hence RF and SVM can be considered as good predictors for groundwater management applications. Deep learning architecture has recently been able to prove better capturing abilities of complex, nonlinear temporal dependence present in the time series of groundwater as compared to statistical conventional methods.

Though the number of research is growing, such an application of satellite-based GRACE and GLDAS data with deep learning techniques has remained largely unexplored in hyper-arid, data-scarce basins similar to the Najd region of Oman. Previous hydrogeological studies in Najd mainly focused on hydrochemical and isotopic characterization of the aquifer system (Al-Mashaikhi et al., 2012). Groundwater level declines of about 0.73 meters per year between 1992 and 2010 were reported, attributing these declines to desert reclamation activities (Saleh et al., 2025). No study, however, has ever used deep learning-based forecasting models with remote sensing data to predict groundwater

storage depletion within this basin. The paper tries to fill this gap by assessing six deep learning architectures for forecasting groundwater storage anomalies in the Najd region. The methodological approach builds on the one previously validated in comparable arid and semi-arid regions around the world.

3. Methods and Data Collection

The methodology consists of five main stages: (1) Delineation of the Najd study-area boundary, (2) Extraction of GRACE TWSA and gap filling by linear interpolation, (3) Extraction of GLDAS variables for cell centers inside the Najd boundary and unit conversion, (4) Computation of GWSA by mass balance, and (5) Spatially independent training and evaluation of the deep learning models. All satellite data cover the period April 2002 to February 2026.

3.1 Study Area

The Najd region is located in the Dhofar Governorate of the Sultanate of Oman, covering approximately 90,000 km² and bounded by Lat 16.69°N - 20.00°N and Lon 52.00°E - 55.88°E. The region is hyper-arid with average annual precipitation below 50 mm and mean annual air temperature of 27.6°C. Groundwater occurs in four distinct aquifers within the Hadhramaut Group, consisting predominantly of limestone and dolomite formations. The study area boundary (Figure 1) was delineated using ArcGIS.

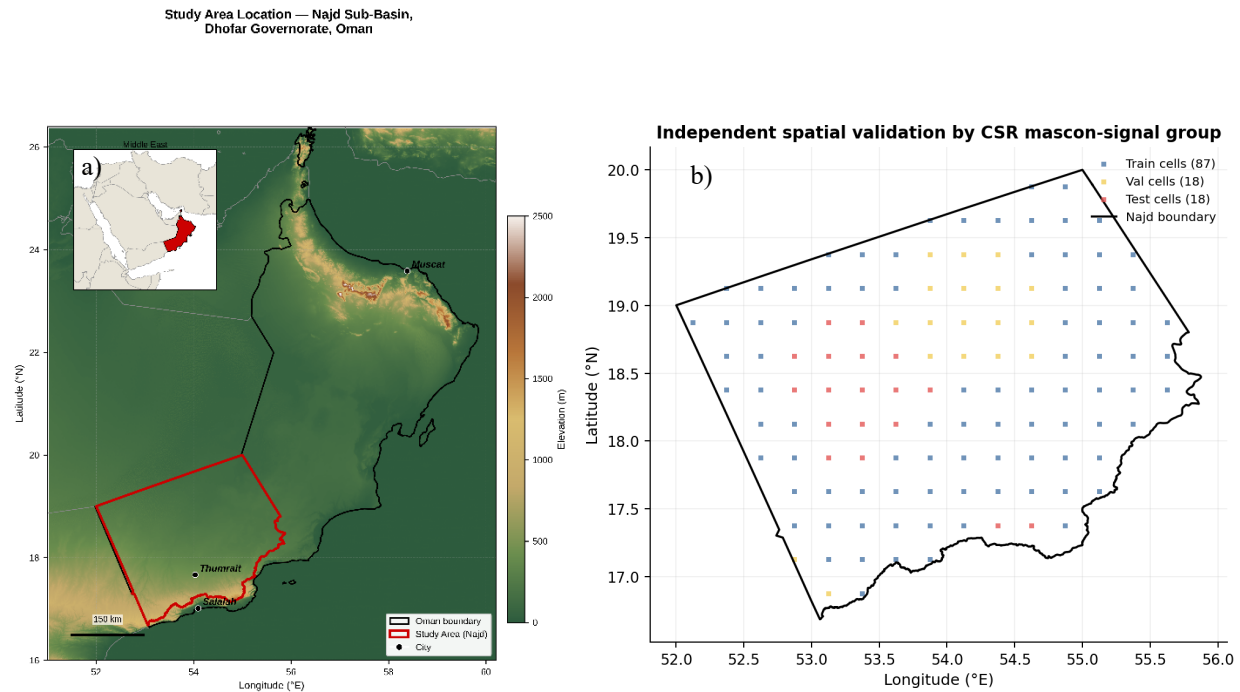


Figure 1: a) Location of the study area; b) GLDAS cells inside the Najd boundary and the independent CSR mascon-signal group split.

3.2 GRACE/GRACE-FO and GLDAS Datasets

Monthly liquid water equivalent thickness (lwe_thickness) in centimeters was obtained from the CSR GRACE and GRACE-FO RL06.3 mascon product (Watkins et al., 2015). The dataset was filtered to April 2002 to February 2026 and sampled only at GLDAS cell centers within the Najd boundary. The period contains 252 observed GRACE months out of 287 expected months. TWSA represents anomalies relative to the 2004-2009 baseline mean. The GRACE/GRACE-FO dataset contains 35 missing months, including scattered gaps and the transition between the GRACE and GRACE-FO missions. These months were identified by comparing the available timestamps with the complete monthly calendar and were filled by linear interpolation, producing a continuous 287-month series.

The GLDAS Noah Land Surface Model monthly product (GLDAS_NOAH025_M V2.1) was downloaded at $0.25^\circ \times 0.25^\circ$ resolution for January 2002 to February 2026. The following 12 variables were extracted for the 123 cell centers

located inside the Najd boundary: four soil moisture layers (0-10, 10-40, 40-100, and 100-200 cm), evapotranspiration, total precipitation, storm surface runoff, baseflow-groundwater runoff, air temperature, surface skin temperature, potential evapotranspiration, and root zone soil moisture.

Najd basin storage anomalies from CSR GRACE and GLDAS Noah V2.1

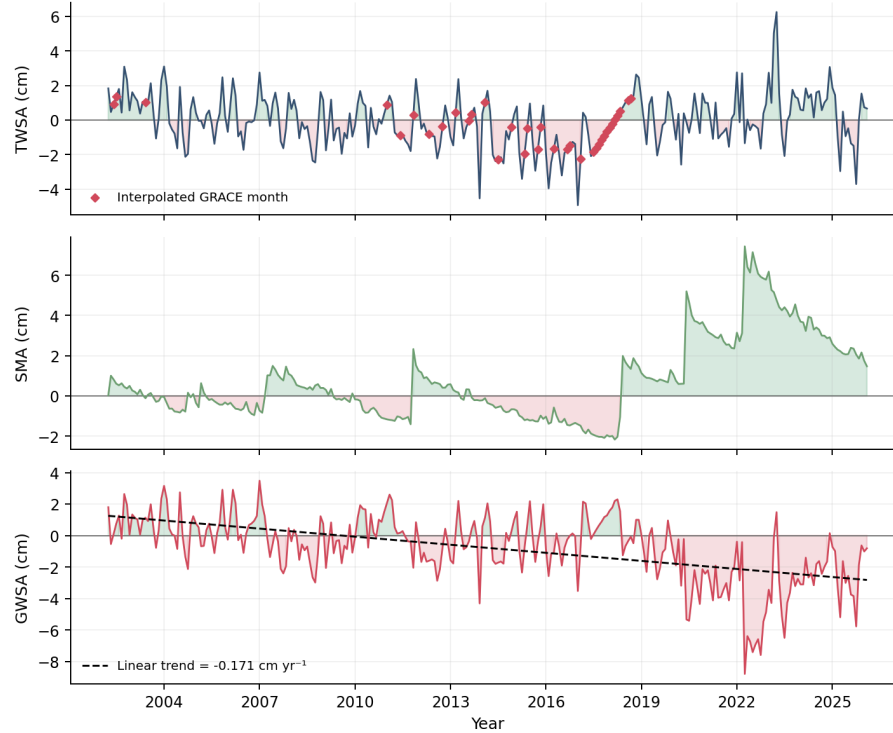


Figure 2: Najd basin-average TWSA, SMA, and GWSA from April 2002 to February 2026. Red diamonds identify interpolated GRACE months; the dashed line shows the GWSA trend.

3.3 GWSA Computation

GWSA was obtained by expressing the mass balance of GRACE TWSA and GLDAS total soil moisture anomaly (SMA), as in Equation (1):

$$\text{GWSA} = \text{TWSA} - \text{SMA} \quad (1)$$

where SMA is the deviation of total column soil moisture (sum of four GLDAS layers in cm) from its long-term mean. Snow water equivalent and canopy interception are negligible in the hyper-arid Najd region. Both datasets were time-synchronized on a common monthly time axis for April 2002 to February 2026.

Six data-driven models were trained using the same eight GLDAS variables: evapotranspiration (Evap), rainfall (Rainf), subsurface runoff (Qsb), root zone soil moisture (RootMoist), average surface temperature (AvgSurfT), air temperature (Tair), potential evapotranspiration (PotEvap), and surface runoff (Qs). The minimal adjustment added only latitude, longitude, normalized time, month sine, and month cosine. For CNN, ERNN, GRU, LSTM, and TCN, the features were organized into genuine six-month rolling sequences, while MLP retained the current-month flat representation. All 123 cells were inside the Najd boundary and belonged to 14 independent CSR mascon-signal groups. A group-level split assigned 10 groups (87 cells) to training, 2 groups (18 cells) to validation, and 2 groups (18 cells) to testing, with zero group overlap. Predictors were standardized from the training partition only, and the models were trained for up to 45 epochs using validation-based model selection.

Water-flux and storage variables were converted to centimeters and temperature variables to degrees Celsius. The test groups were excluded completely from model fitting and validation, while their full available temporal record was retained. This design evaluates transfer to unseen locations within Najd rather than using a chronological forecast split.

4. Results

4.1 Najd-Only Spatial Validation and GWSA Trend

Figure 1 shows the 123 GLDAS cells retained inside the Najd boundary and their independent group allocation. The spatial split contains 87 training, 18 validation, and 18 testing cells drawn from separate CSR mascon-signal groups. The updated Najd GWSA series has a mean of -0.78 cm and ranges from -8.78 to +3.48 cm. A statistically significant declining trend of -0.171 cm/year indicates progressive groundwater depletion under the hyper-arid climate and continued abstraction pressure.

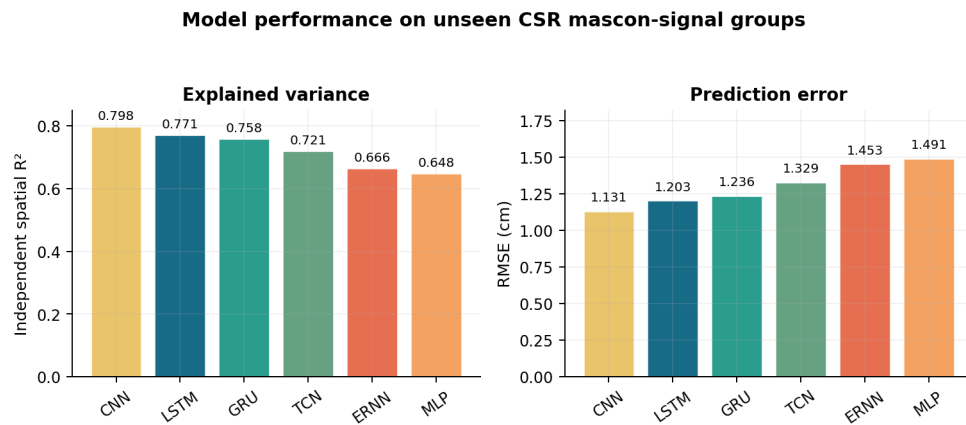


Figure 3: Independent spatial-validation R^2 and RMSE for the six models using Najd-boundary cells only.

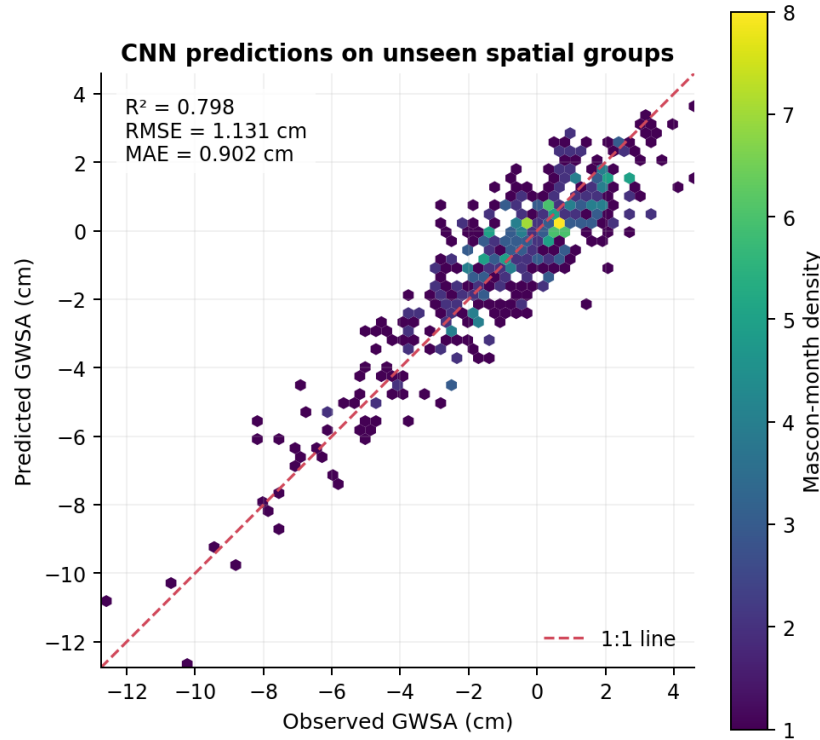


Figure 4: Observed versus CNN-predicted GWSA for the independently withheld Najd CSR signal groups.

4.2 TWSA, SMA, and GWSA Time Series Analysis and Model Performance

Figure 2 shows the complete Najd time series of TWSA, soil moisture anomaly (SMA), and derived GWSA. The GWSA exhibits a statistically significant decreasing trend of -0.171 cm/year ($p < 0.001$), equivalent to an estimated decline of about 4.07 cm over the study period. It ranges from -8.78 to $+3.48 \text{ cm}$, with a mean of -0.78 cm and a standard deviation of 2.09 cm . The strongest negative anomalies occur after 2020, while short positive anomalies reflect temporary storage recovery that does not reverse the long-term decline.

Table 1 presents the performance of all six models on independently withheld CSR mascon-signal groups located entirely inside the Najd boundary. The primary evaluation contains 498 independent mascon-month observations. Figures 3 and 4 summarize the model comparison and the best-model agreement with observed GWSA, while Figures 5 and 6 show the full-period all-model comparison and the minimal-adjustment ablation.

Table 1: Performance metrics of the six models for independent spatial validation within the Najd boundary.

Model	MAE (cm)	MSE (cm ²)	RMSE (cm)	R ²	Rank
CNN	0.902	1.279	1.131	0.798	1st ★
ERNN	1.169	2.112	1.453	0.666	5th
GRU	1.007	1.528	1.236	0.758	3rd
LSTM	0.974	1.447	1.203	0.771	2nd
MLP	1.175	2.224	1.491	0.648	6th
TCN	1.061	1.766	1.329	0.721	4th

The CNN model records the best independent spatial-validation performance with the lowest MAE (0.902 cm), MSE (1.279 cm²), RMSE (1.131 cm), and highest R² (0.798). LSTM ranks second (R² = 0.771), followed by GRU (R² = 0.758), TCN (R² = 0.721), ERNN (R² = 0.666), and MLP (R² = 0.648). Figure 3 compares R² and RMSE for all six models, while Figure 4 shows that CNN predictions follow the observed GWSA pattern across the two completely withheld CSR signal groups.

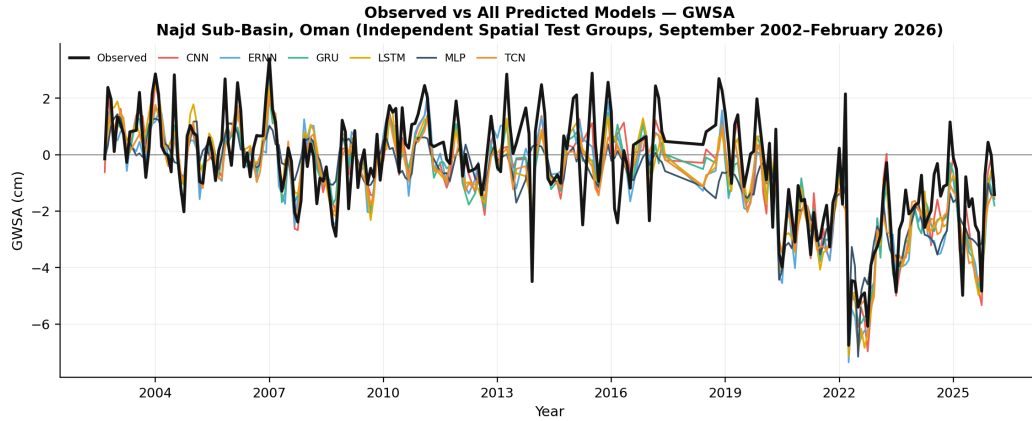


Figure 5: Observed and predicted GWSA for all six models across the independently withheld Najd spatial groups from September 2002 to February 2026.

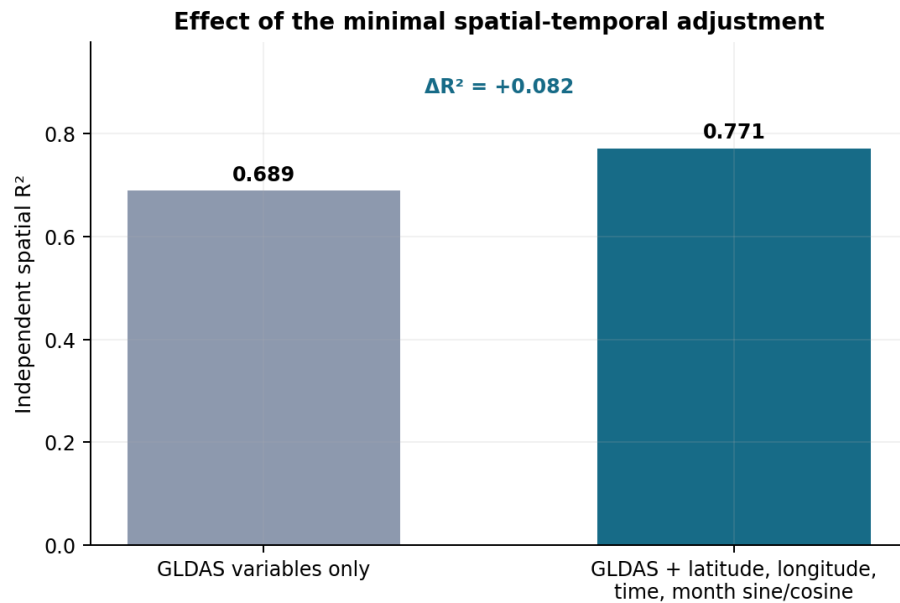


Figure 6: LSTM ablation for Najd-only independent spatial validation. The minimal adjustment increases R^2 from 0.689 to 0.771.

5. Discussion

The combination of GRACE/GLDAS satellite data with deep learning proved useful for estimating groundwater storage inside the Najd boundary. All sequence-based models received genuine six-month temporal inputs. The CNN model performed best ($R^2 = 0.798$; $RMSE = 1.131$ cm), followed by LSTM and GRU. An important finding is that retaining the same eight GLDAS variables and adding only latitude, longitude, normalized time, and cyclical month identifiers improved LSTM independent spatial R^2 from 0.689 to 0.771. Latitude and longitude represent stable spatial differences among the Najd cells, while normalized time and the cyclical month terms represent long-term change and annual seasonality. Because the test set contains only two of the 14 CSR signal groups, the result should be interpreted as independent spatial validation within Najd, not as forecasting beyond February 2026. Future work should integrate in-situ groundwater levels, extraction records, and additional independent spatial observations.

6. Conclusion

Six deep learning models were evaluated for GWSA estimation inside the Najd boundary using GRACE/GRACE-FO and GLDAS satellite data from April 2002 to February 2026. The analysis used 123 GLDAS cells belonging to 14 independent CSR mascon-signal groups, divided into 10 training, 2 validation, and 2 testing groups. Thirty-five missing GRACE months were linearly interpolated to produce a complete 287-month series. The GRACE-based analysis indicates a statistically significant declining GWSA trend of -0.171 cm/year, equivalent to about 4.07 cm over the study period. Among the six models, CNN achieved the best independent spatial-validation results (MAE = 0.902 cm, RMSE = 1.131 cm, $R^2 = 0.798$, and KGE = 0.865). LSTM ranked second ($R^2 = 0.771$), followed by GRU ($R^2 = 0.758$).

The study confirms the value of satellite-based GRACE/GLDAS monitoring with deep learning for groundwater assessment in data-scarce arid regions. The minimal feature adjustment improved LSTM R^2 from 0.689 to 0.771 without changing the original eight GLDAS variables. Subsequent studies should integrate in-situ groundwater levels, groundwater extraction data, land-use information, and climate projections to improve model reliability and support long-term water-resources planning in Najd and other hyper-arid environments of the Arabian Peninsula.

References

- Al-Mashaikhi, K. et al., Evaluation of groundwater dynamics and quality in the Najd aquifers located in the Sultanate of Oman, *Environmental Earth Sciences*, vol. 66, no. 4, pp. 1195-1211, 2012.
- Band, S. S. et al., Groundwater level prediction in arid areas using wavelet analysis and Gaussian process regression, *Engineering Applications of Computational Fluid Mechanics*, vol. 15, no. 1, pp. 1147-1158, 2021.
- Feng, F., Ghorbani, H. and Radwan, A. E., Predicting groundwater level using traditional and deep machine learning algorithms, *Frontiers in Environmental Science*, vol. 12, 2024.
- Kanito, D., Benaafi, M. and Baalousha, H. M., Machine learning models for groundwater level prediction and uncertainty analysis in Ruataniwha Basin, New Zealand, *Hydrology*, vol. 12, no. 11, p. 282, 2025.
- Khan, J. et al., A comprehensive review of conventional, machine learning, and deep learning models for groundwater level (GWL) forecasting, *Applied Sciences*, vol. 13, no. 4, p. 2743, 2023.
- Missik, J. E. C. et al., Groundwater regulates interannual variations in evapotranspiration in a riparian semiarid ecosystem, *Journal of Geophysical Research: Atmospheres*, vol. 126, no. 7, p. e2020JD033078, 2021.
- Pham, Q. et al., Groundwater level prediction using machine learning algorithms in a drought-prone area, *Neural Computing and Applications*, vol. 34, pp. 1-23, 2022.
- Pulla, S. T., Yasarer, H. and Yarbrough, L. D., GRACE Downscaler: a framework to develop and evaluate downscaling models for GRACE, *Remote Sensing*, vol. 15, no. 9, p. 2247, 2023.
- Rzepecka, Z. and Birylo, M., Groundwater storage changes derived from GRACE and GLDAS on smaller river basins—a case study in Poland, *Geosciences*, vol. 10, no. 4, p. 124, 2020.
- Saleh, H. et al., Intensifying tropical cyclones in the Arabian Sea replenish depleting aquifers, *Communications Earth & Environment*, vol. 6, no. 1, p. 536, 2025.
- Sattar, Y. and Khalid, Z., Estimation of groundwater storage variations in Indus River Basin using GRACE data, *Proceedings of the 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4540-4544, 2021.
- Takafuji, E. H. de M., Rocha, M. M. da and Manzione, R. L., Groundwater level prediction/forecasting and assessment of uncertainty using SGS and ARIMA models: a case study in the Bauru Aquifer System (Brazil), *Natural Resources Research*, vol. 28, no. 2, pp. 487-503, 2019.
- Watkins, M. M. et al., Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons, *Journal of Geophysical Research: Solid Earth*, vol. 120, no. 4, pp. 2648-2671, 2015.

Biographies

Mohammed S. Al Nadabi is a Ph.D. candidate in the Department of Civil and Architectural Engineering at Sultan Qaboos University, Muscat, Oman. He is a remote sensing and geospatial professional with 17 years of experience spanning GIS analysis, UAV surveying, satellite image interpretation, environmental monitoring, and water resource applications. Concurrently, he serves as Head of the Surveys and Remote Sensing Section at the Ministry of Agriculture, Fisheries and Water Resources of the Sultanate of Oman, where he leads national geospatial and remote sensing activities related to water resource monitoring and land management. He is proficient in Google Earth Engine, ArcGIS Pro, MATLAB, and machine learning techniques for geospatial decision support. His research interests include remote sensing applications in hydrology, Geographic Information Systems (GIS), satellite-based

groundwater monitoring, drought assessment, environmental change analysis, and water resources management in arid and hyper-arid environments.

Dr. Mohammed El-Diasty is currently an Associate Professor of Geomatics Engineering in the Department of Civil and Architectural Engineering at Sultan Qaboos University, Oman. He obtained a bachelor's degree in civil engineering from Mansoura University, Egypt, in 1997, an M.Sc. degree in Civil Engineering (Geomatics Engineering) from Ryerson University, Canada, in 2003, and a Ph.D. in Geomatics Engineering from York University, Canada, in 2010. He was previously at Mansoura University, Egypt for twelve years and at King Abdulaziz University, Saudi Arabia for nine years. His research interests include positioning and navigation, mobile mapping, GNSS/INS integration, hydrography, UAV-based mapping, digital twin, urban modeling, and digital heritage documentation.

Dr. Talal Etri is an Assistant Professor in the Department of Civil and Architectural Engineering at Sultan Qaboos University, Oman, specializing in Hydraulic and Coastal Engineering. He holds a B.Sc. in Civil Engineering from the University of Tripoli (1995), an M.Sc. in Computational Engineering from Ruhr-Universität Bochum, Germany (2002), and a Ph.D. in Coastal Geosciences and Engineering from Christian-Albrechts-Universität zu Kiel, Germany (2007). His research interests include coastal hydrodynamic measurements and numerical simulations covering tides, waves and currents; wave boundary layer measurement and modelling; sediment transport in wadi channels and the coastal zone; and short, medium, and long-term numerical modelling of coastal morphodynamics. He has been involved in coastal research projects in Germany, Brazil, Saudi Arabia, and Indonesia.

Dr. Mohammad Reza Nikoo is an Associate Professor in the Department of Civil and Architectural Engineering at Sultan Qaboos University, Oman, where he has served since 2021. He was awarded the prestigious Oman National Research Award 2024 and was recognized as one of the world's top 2% scientists in 2024 and 2025, based on Stanford's Top 2% Scientists Ranking. In 2026, he received the Best Researcher Award at Sultan Qaboos University and the Best Research Paper Award from the Environmental and Water Resources Institute (EWRI). He has served as an Honorary Visiting Fellow at the University of Technology Sydney, Australia, for four consecutive years (2022–2025), and was previously an Associate Professor at Shiraz University, Iran (2017–2021), where he was selected as the Top Young Researcher for three consecutive years (2018–2020) and Iran's Eminent Young Scientist in 2020. His research interests encompass water resources management, water quality management, and environmental engineering, with over 461 publications receiving more than 8,600 citations and an H-index of 49. He has supervised 14 Ph.D. and 50 M.Sc. theses.