Applied Computing and Geosciences 28 (2025) 100304

Contents lists available at ScienceDirect

Applied Computing and Geosciences

journal homepage: www.sciencedirect.com/journal/applied-computing-and-geosciences

Groundwater level forecasting using data-driven models and vadose zone: A comparative analysis of ARIMA, SARIMAX, Prophet, and NeuralProphet

Alessandro Galdelli [a,1,*], Davide Fronzi [b,1], Gagan Narang [a], Adriano Mancini [a], Alberto Tazioli [b]

[a] VRAI - Vision Robotics and Artificial Intelligence Lab, Department of Information Engineering (DII), Università Politecnica delle Marche, Italy
[b] Department of Science and Matter Engineering, Environment and Urban Planning (SIMAU), Università Politecnica delle Marche, Italy

ARTICLE INFO

Dataset link: https://github.com/vrai-group/Groundwater-level-forecasting

Keywords:
Groundwater
Time series forecasting
Hydrogeology
RAPS
Artificial intelligence

ABSTRACT

Forecasting water availability is becoming increasingly vital due to rising human demands and changing climatic pressures have caused declines in groundwater levels across many regions. While numerous studies employ data-driven approaches to predict groundwater fluctuations using meteorological data and groundwater level observations, few incorporate measurements from the vadose zone into predictive models. This study proposes a novel method leveraging an advanced hydrogeological monitoring system with high spatio-temporal resolution to forecast groundwater levels in a shallow alluvial aquifer used for drinking purposes. The monitoring system comprises a thermo-pluviometric station and three probes that measure soil water content, electrical conductivity, and temperature at depths of 0.6, 0.9, and 1.7 meters, in addition to a piezometer with a permanent water-level sensor. Data was collected at 15 min intervals over two hydrological years and integrated as exogenous inputs to enhance model predictive performance. Statistical, machine learning and deep learning architectures were tested through ARIMA, SARIMAX, Prophet and NeuralProphet providing a comprehensive evaluation of different approaches. For a robust evaluation, a rolling K-fold cross-validation strategy was implemented and coupled with a grid search to fine-tune all the models. Evaluation metrics and correlation coefficients are employed to assess the predictive capabilities of each model. Our findings indicate that prediction accuracy improves across all models with increasing depth in the vadose zone, with machine learning and deep learning models showing the most significant improvements. Specifically, at 1.7 m depth, Prophet achieves a MAPE of 4.5%, and NeuralProphet achieves a MAPE of 4.1% compared to statistical models. This study has successfully highlighted the enhancement of AI-based models for estimating levels of groundwater incorporating subsurface information from the vadose zone at different depths and phreatic zones, alongside climatic variables.

1. Introduction

Water plays an indispensable role in supporting human life, agricultural productivity, industrial development, and ecological balance (Mishra, 2023; De Fraiture et al., 2010; Connor, 2015). While episodes of intense rainfall can lead to flooding, declining reserves of groundwater for public use are increasingly apparent in vulnerable regions such as the Mediterranean basin (Cos et al., 2022; Tuel and Eltahir, 2020). Under these evolving conditions, the growing irregularity and intensity of precipitation events complicate the task of forecasting subsurface water availability. This dynamic environment demands that predictive models continually adapt to reflect ongoing hydrological shifts (Davamani et al., 2024; Milly et al., 2008; Goderniaux et al., 2011). To address these complexities, emerging research advocates for redirecting observational priorities from groundwater storage to the unsaturated (vadose) zone, aiming to gain deeper insight into the mechanisms driving aquifer recharge (Turkeltaub et al., 2015). The vadose zone, stretching from the land surface to the phreatic boundary, plays a crucial role in controlling recharge–discharge balance and offers valuable information about infiltration dynamics during rainfall events (Hopmans and Van Genuchten, 2005; Hendrickx et al., 2001). Current advancements in environmental sensing and hydrological monitoring technologies have further strengthened efforts to model groundwater level (GWL) variability, supporting sustainable water management. Sensor advancements and new data horizons are at the core of emerging groundwater monitoring strategies. Using

* Corresponding author.
E-mail address: a.galdelli@univpm.it (A. Galdelli).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.acags.2025.100304
Received 5 May 2025; Received in revised form 15 October 2025; Accepted 27 October 2025
Available online 30 October 2025

hydrological monitoring data, the levels are forecasted and then combined with geophysical surveys to track water levels and subsurface conditions (Dangar et al., 2021). The common management strategy involves integrating hydrological models with observational data to improve predictive accuracy and decision-making. Machine learning adoption in geosciences recently had a new found traction (Maity et al., 2024), demonstrating strong ability to model complex, nonlinear relationships across various applications (Lary et al., 2016; Mammoliti et al., 2022; Maniar et al., 2018; Karpatne et al., 2018; Dramsch, 2020; Mancini et al., 2023). Leveraging machine learning algorithms is well established to capture underlying patterns in groundwater fluctuations, outperforming conventional numerical methodologies in adaptability and performance (Galdelli et al., 2023).

Recently, there has been a growing shift toward including Artificial Intelligence (AI)-based techniques in GWL modeling (Rajaee et al., 2019; Vadiati et al., 2022; Sharafati et al., 2020). The trend is primarily motivated by the superior capability of AI methods to capture complex, non-linear, and dynamic patterns in time series data, which often challenge conventional modeling techniques (Djimadoumngar, 2023). Popular AI-based forecasting approaches include machine learning models such as Random Forest (RF) and Support Vector Machine (SVM), and deep learning models like Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), and Sequence-to-Sequence (Seq2Seq) models (Afrifa et al., 2022; Najafabadipour et al., 2022; Yan et al., 2023; Dadhich et al., 2021; Khozani et al., 2022; Chen et al., 2023). Numerical methods are still studied and widely applied (Shirmohammadi et al., 2013; Sarma and Singh, 2022; Amiri et al., 2023), where the most common statistical models like AutoRegressive Integrated Moving Average (ARIMA) and its seasonal counterpart (SARIMA) have presented interpretable results (Ren et al., 2022). Khan et al. (2023) in the comparative benchmark study indicated that machine learning and deep learning models were able to capture complex, nonlinear relationships within GWL datasets. While RF and SVM are popular, there is a lack of consideration for mechanisms for capturing temporal trends, especially in ANN-based models, particularly LSTM and Seq2Seq, which rely heavily on careful architecture design and hyperparameter tuning. The race to improve simplicity is relentless, with new models and methods continually emerging.

As research advances, newer approaches are beginning to incorporate the forecaster in the loop, offering a shift toward more interpretable forecasting techniques. Prophet, an open-source tool, is specifically designed for tasks with strong seasonality, trends, and abrupt effects (Taylor and Letham, 2018). Though originally designed for understanding social media traffic for internal use of the social media service Facebook, its source code was made public, and since then, it has found cross-domain applications. Tasks such as product demand (Negre et al., 2024), gas consumption (Tian et al., 2024), and many more investigated and lauded the robust and decomposable modeling approach. Triebe et al. (2021) extended the capabilities of Prophet by incorporating neural network components into the underlying architecture. AutoRegressive component in NeuralProphet adds a feedforward neural network with fully connected layers, which is trained alongside the decomposable components of Prophet, i.e., trend, seasonality, and holiday. Zarinmehr et al. (2022) tested Prophet for GWL estimation and compared it with other models such as ARIMA, Multivariate Adaptive Regression Splines, and Exponential Smoothing using average water level from satellite data, and analyzed to forecast water availability in the Northwester regions of Urmia Lake basin in Iran. The study demonstrated that Prophet outperformed comparison models in terms of accuracy, and further, its residuals were normally distributed, indicating its reliability for predictive purposes. Aguilera et al. (2019) obtained similar results that confirmed the prediction capabilities of Prophet on a study in the wetlands of the Spanish region of Doñana on GWL time series. The results of the new data-driven approaches are useful but may not robustly represent the actual groundwater recharge mechanism. There is a pressing need to employ vadose zone data in AI-based predictive models where only a few studies have contributed. Instead, traditional approaches often rely on precipitation and air temperature data to estimate water availability, either through empirical or physical methods. A clear demand exists for flexible and reliable forecasting methods that can accommodate seasonal variations in time series data that are common in GWL fluctuations and effectively model complex, non-linear dynamics (Zhang et al., 2023; Fronzi et al., 2024).

This research addresses the identified gap by utilizing high-resolution vadose zone observations from multiple depths to estimate GWL in an alluvial aquifer located in the Ancona province of the Marche region, central Italy. The study site employs a system, including a thermo-pluviometric station, soil water content probes, electrical conductivity sensors, soil temperature monitors, and a piezometer. The data collected using the system from different depths within the vadose zone are incorporated as exogenous variables in the modeling process to assess their contribution to GWL prediction.

In this context, the specific contributions of the present study are: (i) attempts to forecast GWL using multi-depth vadose-zone observations as exogenous drivers rather than relying solely on meteorological inputs; (ii) a systematic benchmarking of statistical, and advanced time-series models (ARIMA, SARIMAX, Prophet, NeuralProphet) using cross-validation; (iii) a comprehensive trend–correlation– and residual analysis across depths, to link model behavior to hydrogeological dynamics. Since the study represents one of the preliminary efforts in modeling the forecastability of GWL by monitoring the water dynamics in the vadose zone, a detailed correlation and trend analysis is additionally conducted by exploiting the Rescaled Adjusted Partial Sums (RAPS) method and computing the correlation coefficient between variables. The models were used to capture the GWL patterns and evaluate their response to key hydro-meteorological variables, offering insights into the vadose zone's role in groundwater recharge. This study, therefore, advances the understanding of vadose zone dynamics at different depths and their integration into forecasting models, contributing to improved hydrogeological modeling and management.

2. Material and methods

2.1. Study site and hydrological monitoring strategy

The study area lies in the Mediterranean basin in central Italy, focusing on a catchment (~5 km^2) within the Conero Mt. Regional Park (Fig. 1a). The area's climate is classified as temperate and sublittoral according to the Köppen classification (Fratianni and Acquaotta, 2017). The average annual temperature is approximately 14 °C at the peak (572 m above sea level, a.s.l.) and 16.6 °C along the coast, with historical extremes ranging from −6.5 °C to +35 °C. Annual rainfall averages around 900 mm, with the lowest precipitation occurring in July and the highest in September (Chemeri et al., 2025; Busico et al., 2024, 2020). Actual evapotranspiration accounts for about 60% of total precipitation annually, resulting in an estimated effective rainfall of 300 mm/y for the period 1995–2015 (Mussi et al., 2017). In the catchment, two main aquifers are identified (Chemeri et al., 2025; Fronzi et al., 2024): (1) the phreatic alluvial aquifer hosted in the alluvial deposits, and (2) the semi-confined Scaglia Calcarea aquifer hosted the Scaglia Rossa and Scaglia Bianca geological formations (Fms.). The alluvial aquifer mainly consists of silty-sands, sandy-silts, and gravel with sand layers, sustains the aquatic ecosystems through groundwater–surface water interactions (Fronzi et al., 2022). Recharge for both aquifers is from meteoric input. These aquifers are crucial for drinking water supplies, as tapped by the local water management company. However, reduced recharge in 2019–2021 caused significant groundwater depletion and prolonged drying of watercourses (Fronzi et al., 2022). The recent meteoric and hydrological dry conditions required an in-depth analysis of future groundwater availability, emphasizing infiltration dynamics

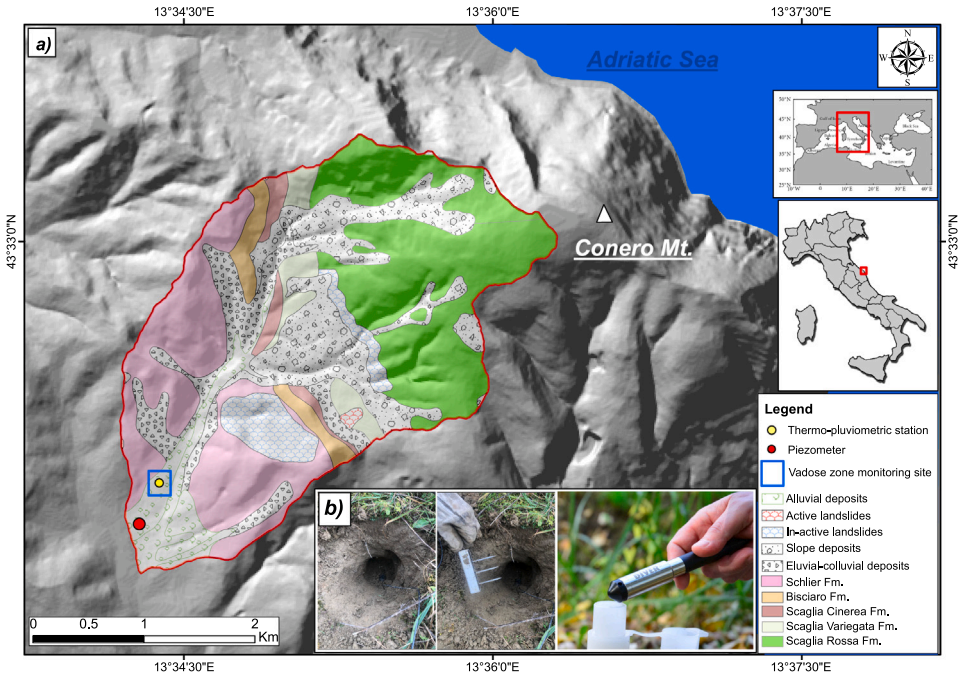


Fig. 1. Study site location, with geological information (a) and sensors (TEROS 12 and TD-Diver) used for hydrological data collection (b).

and vadose zone water flow toward the alluvial aquifer to support local water utilities and public agencies.

Monitoring began on 1 January 2022, with hydrological parameters recorded every 15 min over more than two hydrological years until 30 April 2023, except for GWL, which was logged daily at noon. Internet of Things (IoT) technology was utilized for robust and efficient automatic data collection (Galdelli et al., 2019). The natural site setting was simplified for modeling purposes as illustrated in Fig. 2. Layer 1 represents the Atmosphere and related variables: Precipitation and Air temperature (i.e., P and T_{air}), Layer 2 is referred to the vadose zone and associated variables measured at different depths: Volumetric Water Content, Soil Temperature, and Bulk Electric Conductivity (i.e., VWC, T_{soil}, EC), and Layer 3 relates to the saturated zone and its corresponding variable: Groundwater Level (i.e. GWL). The implementation of the forecasting models consists of four distinct scenarios designed to thoroughly evaluate the model's predictive capabilities. Initially, atmospheric variables from Layer 1 were utilized to predict GWL in Layer 3. Subsequently, three additional tests were performed using variables from Layer 2 at depths of 0.6 m, 0.9 m, and 1.7 m, each independently applied to forecast GWL in Layer 3. These scenarios provide key insights into how hydrological variables impact groundwater level, improving the forecasting model for different system conditions. The resulting high-resolution spatial and temporal dataset was pre-processed at a daily scale and analyzed through trend and correlation analysis and, eventually, employed for developing the GWL prediction models.

2.2. Collected data

The collected data were processed at a daily scale by computing the cumulative daily precipitation, and the mean daily values for T_{air}, T_{soil}, EC, and VWC at all the monitored depths. The time series processed at a daily scale are reported in (Fig. 3). The graph of GWL shows seasonal fluctuations, indicating dynamic groundwater conditions within the aquifer during the hydrologic year. The recharge starts in November until April and a depletion phase can be observed starting from the late spring until October. Atmospheric temperature demonstrates clear seasonal patterns, reflecting expected meteoric conditions at the study site with maximum values during summer and minimum values during the winter season. Rainfall data is intermittent with significant peaks, highlighting periods of intense rainfall (e.g., in May and September) alternated with prolonged dry periods (e.g., July and August). Soil temperatures in the vadose zone at varying depths (0.6 m, 0.9 m, and 1.7 m) generally follow a consistent trend, mimicking the T_{air} across the hydrologic year but show variations influenced by depth. The VWC and the EC display a similar pattern if considering single depths. The VWC at shallower depths is characterized by a sharp increase after rainfall events during wet periods. The response to rainfall events in the soil is smoothed going deep. The EC generally mimics the VWC. It is important to note that at this stage, the primary objective of the research is to enhance and compare data-driven models for GWL forecasting using vadose zone data. At the same time, the characterization of aquifer recharge processes in the study area remains beyond the scope of this work. To ensure scientific rigor, a more detailed specification of the dynamics between precipitation events, infiltration patterns, and aquifer recharge is provided in Chemeri et al. (2025).

2.3. Assessment of data uncertainty

The high-resolution hydrogeological monitoring system was established in January 2022. P and T_{air} are monitored using a thermo-pluviometric station. A piezometer installed in the alluvial aquifer is instrumented with a TD-Diver hydrometric pressure transducer (Eijkelkamp, accuracy ±0.5 cmH2O and resolution 0.2 cmH2O) for continuous GWL monitoring (Fig. 1b), compensated for atmospheric pressure. Three advanced TEROS 12 sensors (Meter Group Inc., Pullman, WA, USA) for measuring T_{soil}, VWC, and EC were installed at 1.7, 0.9 and, 0.6 m depth into the alluvial aquifer soil, approximately 200 m upstream from the equipped piezometer. The TEROS 12 sensors (Fig. 1b) operate based on the TDR principle, allowing precise monitoring of hydro-physical properties in the vadose zone. Soil moisture is measured as VWC in m^3/m^3, with a resolution of 0.001 m^3/m^3 and an accuracy of ±0.03 m^3/m^3. T_{soil} is measured over a range of −10 to +60 °C, with a resolution of 0.1 °C and an accuracy of ±0.5 °C from −10 to 0 °C and ±0.3 °C from 0 to 60 °C. EC is measured within a range of 0–20 mS/cm, with a resolution of 0.001 mS/cm and an accuracy of ±5%. A comprehensive quality check was performed on all collected hydrological time-series data to detect missing values, verify data consistency, and validate the reliability of the field observations. The assessment

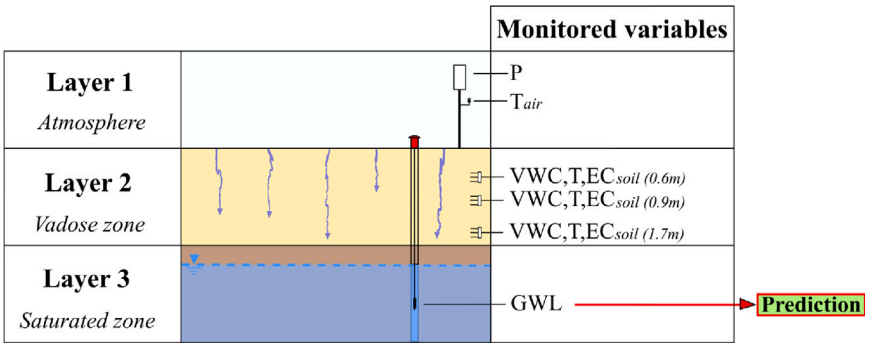


Fig. 2. Illustration of monitoring scenarios and associated variables used during forecasting modeling settings.

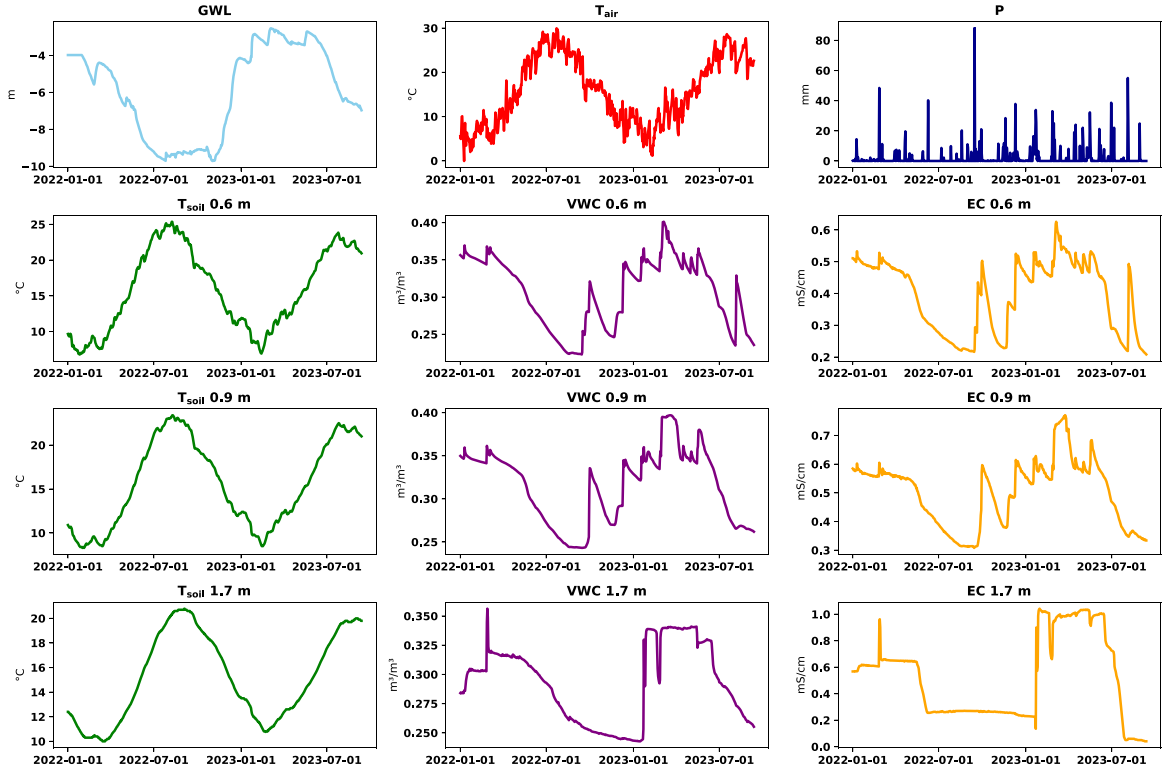


Fig. 3. Time series data of all the monitored variables.

confirmed that the dataset is suitable for further data processing and analysis. The hydrological sensors' uncertainties are further influenced by data processing steps in the experimental setup. When used as inputs for forecasting models, such measurement uncertainties propagate through the modeling process. In data-driven frameworks, small variations or noise in the input variables can affect model parameters, leading to fluctuations in the predicted groundwater levels. The extent of this propagation depends on the model structure, sensitivity to input features, and the amount of available training data. While a detailed quantitative uncertainty analysis is beyond the scope of this work, its qualitative assessment is important to interpret the results and evaluate the reliability of the forecasts. In the following sections, the influence of uncertainty propagation is indirectly examined through the comparison of model performance under different input configurations. To better capture the impact of this uncertainty, the residual analysis presented in the following sections serves as an indirect quantitative measure of its propagation. Residuals, defined as the difference between observed and predicted groundwater levels, reflect both model-related error and input-data uncertainty. Therefore, the variance and dispersion of residuals provide valuable insight into the stability and reliability of the forecasting models under real measurement conditions. Moreover, recent forecasting frameworks, such as RDIT, explicitly use residual statistics to calibrate prediction intervals and better reflect evolving uncertainty (Lai et al., 2025). In hydrologic contexts, residual variance has also been used to inform uncertainty bounds under varying flow regimes (Chen et al., 2025), and in precise engineering systems residual modeling (via GARCH) has proven effective for time-varying uncertainty quantification (Talebpour and Ilbeigi, 2025).

2.4. Trend and correlation

The time-series data of exogenous variables shows visible trends and correlations, requiring empirical investigation for comprehensive analysis. The RAPS method is a popular technique employed in hydrological studies (Šrajbek et al., 2023). The method can effectively compare trends, periodicities, and fluctuations in the observed GWL time series and compare it against the other measured hydrological variables. This approach has been employed in past studies to compare such tendencies, cycles, and variations on time series (Garbrecht and Fernandez,

1994) and, helping visualize and understand the oscillations within collected data, make it an essential tool for groundwater studies (Fiorillo et al., 2015; Fronzi et al., 2024). The method works by standardizing the scale between different time series of target and exogenous variables. Eventually, RAPS for each variable is calculated daily by aggregating the deviations from the mean value of each variable and adjusting against the standard deviation, as reported in Eq. (1).

$$\mathrm{RAPS}_N = \sum_{i=1}^{N} \frac{Y_i - \bar{Y}}{S_Y} \tag{1}$$

where, N, is the amount of data in the time series; Y_i is the value of an individual sample; and the value ranges from $i = 1, 2, \ldots, N$. $\bar{Y}$ is the average value of the observed sample; and S_Y is the standard deviation of the time series.

When multiple hydrological variables are employed as exogenous regressor in predictive modeling methods, it is crucial to assess their interrelationships and their association with the target time series (Dalton and Bekker, 2022; Rocke et al., 2024). This assessment was achieved through a multivariate approach, enabling the joint analysis of multiple variables to understand their collective influence on the dependent variable over time, performing a correlation analysis between hydrological variables at a daily scale. The Pearson correlation coefficient (denoted here as r) was computed using the following equation:

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

where, n is the number of data points, x and y are the individual data points of the two variables being compared.

The correlation coefficient statistically measures the linear relationship between the variables. It ranges from −1 to 1, where 1 indicates a perfect positive linear relationship, −1 indicates a perfect negative linear relationship, and 0 indicates no correlation between explored variables.

2.5. Forecasting models

The time series processed daily was utilized to model and retrain the forecasting models. The internal modeling process is illustrated below:

2.5.1. ARIMA models

ARIMA is a widely used time series forecasting method combining auto regressive (AR) and moving average (MA) components with differencing to address non-stationary time series data. The model is represented by three parameters: 'p', 'd', and 'q', i.e., ARIMA (p, d, q) (Korstanje, 2021). The combined terms facilitate forecasting future values based on historical data as shown in the following equations:

$$Y_t = c + \epsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \delta t \tag{2}$$

$$ARIMA(p, d, q) = AR(p) + I(d) + MA(q) \tag{3}$$

where, p is the autoregressive term, indicating the relationship between an observation and its past values; d is the differencing term, representing the difference between the current and past values; q is the moving average term, relating observations to residual errors from past values; φ and θ are coefficients associated with the AR and MA components, respectively; ϵ_t is the error term. ARIMA is the representative statistical model capable of capturing linear temporal dependencies in groundwater level fluctuations. Its strength lies in its interpretability and suitability for stationary processes that exhibit stable autoregressive patterns. However, ARIMA cannot account for nonlinear dependencies or abrupt changes caused by complex vadose-zone behavior or external climatic variations. Most of these limitations were addressed by the seasonal version of ARIMA.

2.5.2. SARIMAX models

SARIMAX extends ARIMA by incorporating exogenous atmospheric and soil-related variables. This enables the model to capture external climatic influences and delayed responses in groundwater behavior. Its main strength is its ability to represent seasonality and multivariate interactions. The model is formulated as:

$$\phi_p(L)\phi_P(L^s)\Delta^d \Delta_s^d u_t = A(t) + \theta_q(L)\theta_Q(L^s)\zeta_t \tag{4}$$

In this expression, $\phi_p(L)$ denotes the non-seasonal autoregressive operator of order p; $\phi_P(L^s)$ accounts for seasonal autoregression with lag s and order P; Δ^d and Δ_s^D represent non-seasonal and seasonal differencing of orders d and D, respectively; $A(t)$ reflects any deterministic trends (e.g., seasonal cycles); $\theta_q(L)$ is the non-seasonal moving average operator of order q; $\theta_Q(L^s)$ models the seasonal moving average effects with lag s and order Q; ζ_t denotes the stochastic error term; and s is the seasonal periodicity. SARIMAX incorporates the full set of seasonal ARIMA terms, seasonal autoregression (P), seasonal differencing (D), and seasonal moving average (Q), alongside the non-seasonal ARIMA parameters (p, d, q). The seasonal AR term captures dependencies across seasonal lags, while the seasonal MA component adjusts for residual structures within those intervals. Differencing terms address trends and periodic variations to ensure the time series remains stationary.

2.5.3. Prophet

Prophet models time series through a decomposable summation of three components: trend, seasonality, and holidays, expressed mathematically as follows:

$$y(t) = g(t) + s(t) + h(t) + \epsilon(t) \tag{5}$$

where, $g(t)$ denotes the trend component, accounting for long-term, non-cyclic variations; $s(t)$ captures seasonal effects using a Fourier series representation; $h(t)$ models the influence of holiday-related anomalies; and $\epsilon(t)$ reflects random noise or unmodeled variability. Prophet approaches time series modeling by breaking it down into these additive elements, using piecewise linear or logistic growth for the trend and Fourier-based functions for seasonality. Prophet was used to capture nonlinear and quasi-periodic patterns that are typical of hydrological time series. Its automated changepoint detection and multi-seasonality handling make it well-suited for systems with irregular recharge cycles. The model is flexible and interpretable, though its additive decomposition may simplify certain interdependent subsurface processes. It also supports incorporating abrupt changes and holiday effects, although the latter was not utilized in this analysis.

2.5.4. NeuralProphet

NeuralProphet enhances Prophet's flexibility by incorporating autoregressive components and neural networks within a unified structure. This enables it to learn nonlinear interactions between atmospheric and subsurface variables while maintaining partial interpretability. Due to its architecture, the model expands on the complexity, adds a computational overhead, and requires careful tuning to achieve stable results. The decomposable composition is as shown in the following equation:

$$\tilde{y}_t = T(t) + S(t) + E(t) + F(t) + A(t) + L(t) \tag{6}$$

where, $\tilde{y}_t$ denotes the forecasted output; $T(t)$ captures the underlying trend at time t; $S(t)$ represents the seasonal component; $E(t)$ accounts for events and holidays; $F(t)$ models the influence of known future exogenous variables; and $A(t)$ captures auto-regressive patterns derived from historical values at time t.

$L(t)$ is the regression effects for lagged observations of exogenous variables at time t. NeuralProphet adds features like automatic hyperparameter selection during training and the ability to incorporate an auto-regressive component managed by the AR-Net neural network model. Further hidden layers or more output nodes can be added for multi-step forecasting.

3. Experiments

3.1. Problem formulation

The aim is to forecast GWL fluctuations based on a Mediterranean Basin alluvial aquifer by leveraging vadose zone data monitored in high resolution using select statistical, machine learning, and deep learning techniques. The forecasting problem is formulated as follows:

Inputs:

1. $X = x_t : t \in T$ a time series of GWL values over a specified time horizon $T = \{t_1, t_2, \ldots, t_n\}$
   - each $x_t \in \Re$ is the GWL value at time t;
   - the data is collected through an in-situ piezometer with a continuous monitoring sensor.
2. External features:
   - hydro-meteorological variables: $H_t = \{P, T_{\text{air}}, \text{VMC}, T_{\text{soil}}, \text{EC}\}$ from different layers, including precipitation, air temperature, soil moisture, soil temperature and bulk electrical conductivity;

Output:

1. $\hat{Y} = \hat{y}_{t+h} : h \in H$ the forecasted GWL values for a future horizon $H = \{t_{n+1}, t_{n+2}, \ldots, t_{n+h}\}$.
   - The forecast must accurately capture: (i) short-term fluctuations in GWL, and (ii) long-term trends in recharge–discharge cycles to make an accurate estimate of the future availability of the GWL.

Objective function: minimizes the error between the observed GWL values (y_t) and ($\hat{y}_t$) over the prediction horizon H. To ensure robustness and generalizability, we employ a K-fold cross-validation strategy. This involves splitting the entire dataset into K subsets $D_1, D_2, \ldots, D_K$, where each sample is used for both training and validation.
For each fold $k \in \{1, 2, \ldots, K\}$, training is performed on the subset $D^{(k)}_{\text{train}} = D \setminus D_k$, while evaluation is carried out on the corresponding validation fold $D^{(k)}_{\text{val}} = D_k$. The overall cross-validation error is calculated as the mean loss over all K folds:

$$\min_{\Theta} \mathcal{L}_{\text{cv}} = \frac{1}{K} \sum_{k=1}^{K} \mathcal{L}(D^{(k)}_{\text{val}}, f^{(k)}_{\Theta})$$

where $\mathcal{L}$ represents the selected loss function. In this study, we use Mean Absolute Percentage Error (MAPE) as the primary loss metric:

$$\mathcal{L}_{MAPE} = \frac{1}{H} \sum_{h=1}^{H} \left| \frac{y_{t+h} - \hat{y}_{t+h}}{y_{t+h}} \right| \times 100$$

Θ represents the model parameters (e.g., trend coefficients, seasonal components, hyperparameters etc.).
Constraints:

1. Seasonal and cyclical patterns inherent in GWL data:

   $$\hat{y}_t = g(t) + s(t) + h(t) + \epsilon_t$$

   - $g(t)$: long-term trend;
   - $s(t)$: seasonal variations (e.g., annual recharge cycles);
   - $h(t)$: effects from external events (e.g., extreme rainfall events);
   - ϵ_t: residual noise.
2. Handling missing data or irregular sampling intervals using pre-processing techniques (for our case, interpolation and imputation).

Model Selection: We evaluate the formulated problem statement on our selected forecasting models based on their ability to adapt to the characteristics of GWL time-series data. The models are compared in a fair manner with a consistent tuning procedure applied to all models as described in Fig. 4. The models Prophet and NeuralProphet come with built-in functions specifically designed for hyperparameter optimization via cross-validation; the other models required the development of *ad hoc* solutions to implement a similar optimization process.

3.2. Experimental settings

The experiments are performed using a GPU NVIDIA® Quadro RTX 6000 with 48 GB of RAM, and the script is implemented in Python v3.8. The implementation is made open source and made available on GitHub[2] to experiment and extend. The portion of the test dataset is made publicly available for evaluating on other kinds of models. All model selection and tuning are performed on the in-sample data (January 1, 2022–July 12, 2023) using a rolling-origin, expanding-window procedure. The model is first trained on an initial training window covering 90% of the in-sample period, a 12-day forecast horizon is produced, the training window is then advanced by 10 days, the model is retrained, and forecasting is repeated. Errors from each origin are averaged to obtain the cross-validation score used for grid-search hyperparameter selection. Finally, the chosen model is retrained on the complete in-sample data and evaluated once on the held-out test period.

Forecasting accuracy is assessed using metrics such as MAPE, Mean Error (ME), Mean Absolute Error (MAE), Mean Percentage Error (MPE), Root Mean Square Error (RMSE), Correlation (Corr), and Nash–Sutcliffe Efficiency (NSE), which evaluates how well a model predicts relative to the variance in observed data, with values closer to one indicating stronger performance (Jain and Sudheer, 2008). These metrics collectively capture various aspects of model performance, including error magnitude, directional bias, relative size, sensitivity to outliers, strength of linear association, and overall predictive reliability.

3.3. Hyperparameter tuning

The hyperparameter tuning is performed on the validation set. The validation loss, measured using $\mathcal{L}_{MAPE}$, is computed for each hyperparameter configuration as described by the grid shown in Table 1. The process iterates over all candidate hyperparameters in a grid, selecting the configuration that minimizes the average error loss across all folds (as shown in Fig. 4). Once the optimal hyperparameters were determined, they are applied to a test set that the model had never seen before. The model then attempts to forecast the test set, and its predictions are used to evaluate performance. The tuned parameters for each model are discussed, and for ARIMA, the optimal value of autoregressive term p moving average term, and the differencing term was determined to be (5, 1, 5) on cross-validation results. For SARIMAX, the order parameter for non-seasonal components was tuned similarly to ARIMA, with the optimal values being (1, 1, 1). For the seasonal_order parameter, specifying (P, D, Q, M), the optimal seasonal order is $(3, 1, 3, 12)$, which captured the recurring patterns in the dataset without overfitting.

Prophet was configured with an additive seasonality mode and tuned for a balance between flexibility and robustness. Optimal results were achieved using *changepoint_prior_scale* = 0.7 to capture key trend

[2] https://github.com/vrai-group/Groundwater-level-forecasting.

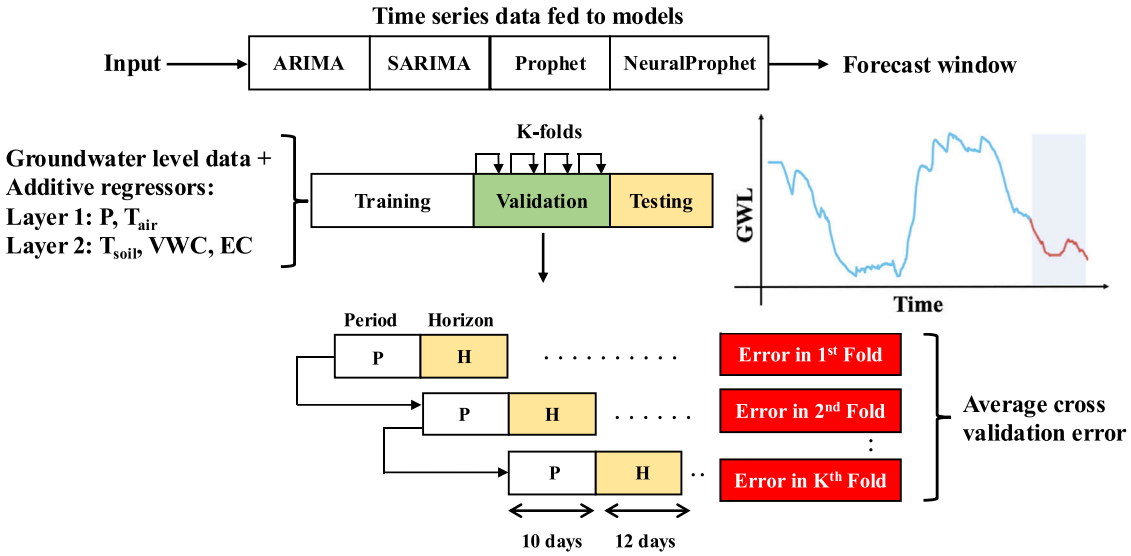


Fig. 4. Cross-validation is used for hyperparameter tuning.

Table 1
Hyperparameter grids used for model tuning.

Model	Hyperparameter	Tuning grid
ARIMA	p	{0, 5, 10, 15}
	d	{0, 1, 2}
	q	{0, 5, 10, 15}
SARIMAX	p	{0, 1, 2}
	d	{0, 1}
	q	{0, 1, 2}
	P	{3, 4, 5}
	D	{0, 1}
	Q	{3, 4, 5}
	M	{3, 6, 12}
Prophet	changepoint_prior_scale	{0.01, 0.3, 0.7, 1, 10}
	seasonality_prior_scale	{0.001, 0.01, 0.1, 0.5}
	changepoint_range	{0.9, 1.0}
	seasonality_mode	{additive, multiplicative}
NeuralProphet	n_changepoints	{5, 30, 50, 100}
	changepoints_range	{0.1, 0.2, 0.3, 0.5, 0.75, 0.9}
	seasonality_reg	{0.001, 10, 150}
	trend_reg	{0.001, 10, 100, 150, 200}
	ar_reg	{0.001, 10, 100}
	seasonality_mode	{additive, multiplicative}
	optimizer	{SGD, Adam}
	impute_missing	{true}
	growth	{linear}

shifts without overfitting, *seasonality_prior_scale* = 0.001 to limit seasonal noise, and *changepoint_range* = 0.9 to constrain changepoints to earlier data segments. NeuralProphet, which enhances Prophet with neural and autoregressive components, was tuned for optimal performance. Key settings included *n_changepoints* of 50 and a *changepoints_range* of 0.75 for balanced flexibility, "additive" seasonality mode based on data behavior, and regularization parameters: *seasonality_reg* = 10, *trend_reg* = 150, and *ar_reg* = 10 to avoid overfitting. The "SGD" optimizer proved most effective for convergence.

3.4. Trend and correlation analysis

Fig. 5 presents the trend analysis using the RAPS method (y-axis) to visualize temporal patterns of GWL alongside atmospheric and vadose zone data at varying levels. Subfigure (a) shows that T_{air} follows a clear seasonal cycle, while P naturally remains relatively irregular and uncorrelated with GWL trends. In contrast, subfigures (b), (c), and (d) display a closer alignment between GWL and vadose zone variables, particularly VWC and EC at 0.6 m, 0.9 m, and 1.7 m depths. The strongest agreement in temporal trends is obtained at 1.7 m, as well, where both VWC and EC exhibit seasonal dynamics that closely mirror the measured GWL, suggesting a more representative relationship between Layer 2 and Layer 3. T_{soil}, showing a marked seasonal trend, consistently diverges from GWL pattern, with a stationary shift of about six months at all depths. In this context, it is worth noting that the extreme rainfall event that affected the study area in September 2022 (Morelli et al., 2023) caused visible changes in the RAPS trend of both EC and VWC at depths of 0.6 m and 0.9 m. Contrarily, at 1.7 m, the impact of the precipitation was dampened, with no significant changes observed in EC, VWC, as well as in GWL. This evidence supports the idea that variables measured deeper in Layer 2 can effectively filter out the noise associated with shallower events, which may not correspond to actual recharge processes and therefore GWL variations.

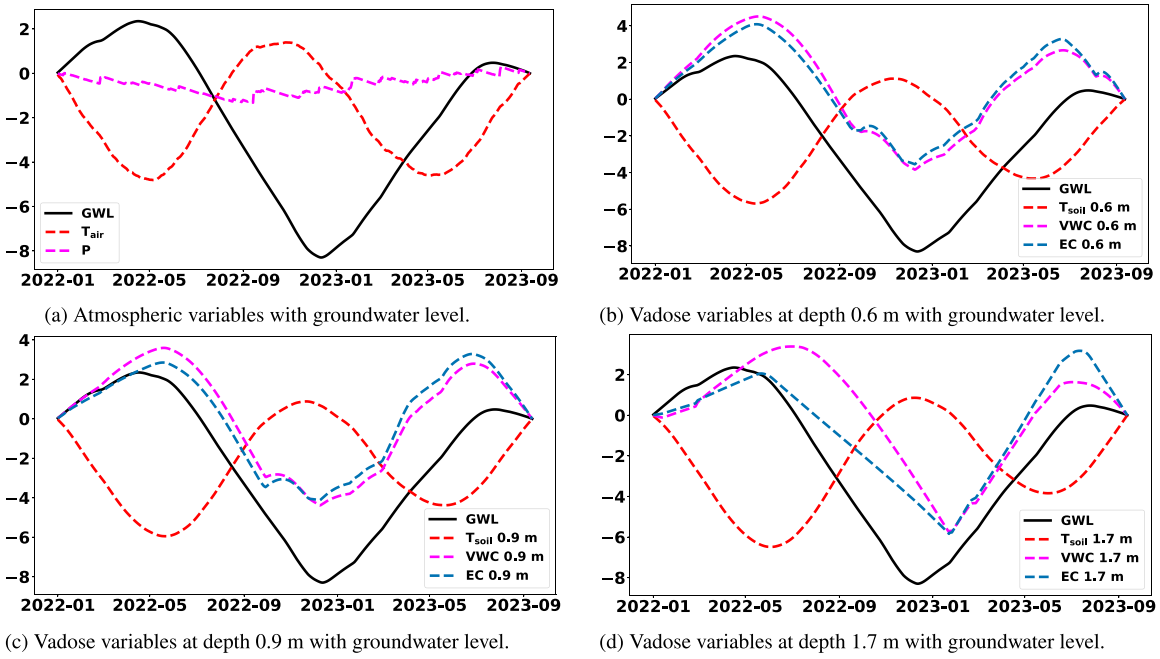
Fig. 6 presents the correlation analysis using the Pearson correlation coefficient between GWL and meteorological and soil-related parameters. The strongest correlations with GWL were found in the vadose zone at 1.7 m, particularly with EC (r = 0.79) and VWC (r = 0.71), indicating that deeper soil measurements are more representative of actual groundwater recharge processes, as expected. Conversely, P showed a very weak correlation with GWL (r = 0.033), while T_{air} had a moderate negative correlation (r = −0.39), likely due to seasonal influences. T_{soil} at all depths also showed negative correlations, especially at 0.6 m (r = −0.5). The values depicted from the correlation matrix were used to decide which exogenous variables are most strongly associated with the target variable (GWL), helping to refine the selection of predictors and improve the accuracy of the model. The trend analysis (i.e., RAPS), together with the insights obtained by the correlation matrix, supports the integration of deeper vadose zone variables, particularly at 1.7 m, in forecasting models for the GWL dynamic.

3.5. Residual computation and uncertainty metrics

Residual analysis was conducted to assess model accuracy and quantify predictive uncertainty across the different groundwater depths. For each forecasting model, residuals were defined as the difference between observed and predicted groundwater levels:

$$r_t = y_t - \hat{y}_t \tag{7}$$

where y_t denotes the observed groundwater level at time t, and $\hat{y}_t$ represents the corresponding model prediction. Residuals provide a



(a) Atmospheric variables with groundwater level.

(b) Vadose variables at depth 0.6 m with groundwater level.

(c) Vadose variables at depth 0.9 m with groundwater level.

(d) Vadose variables at depth 1.7 m with groundwater level.

Fig. 5. Trend analysis using RAPS (y-axis) for the hydrologic variables collected in the different layers.

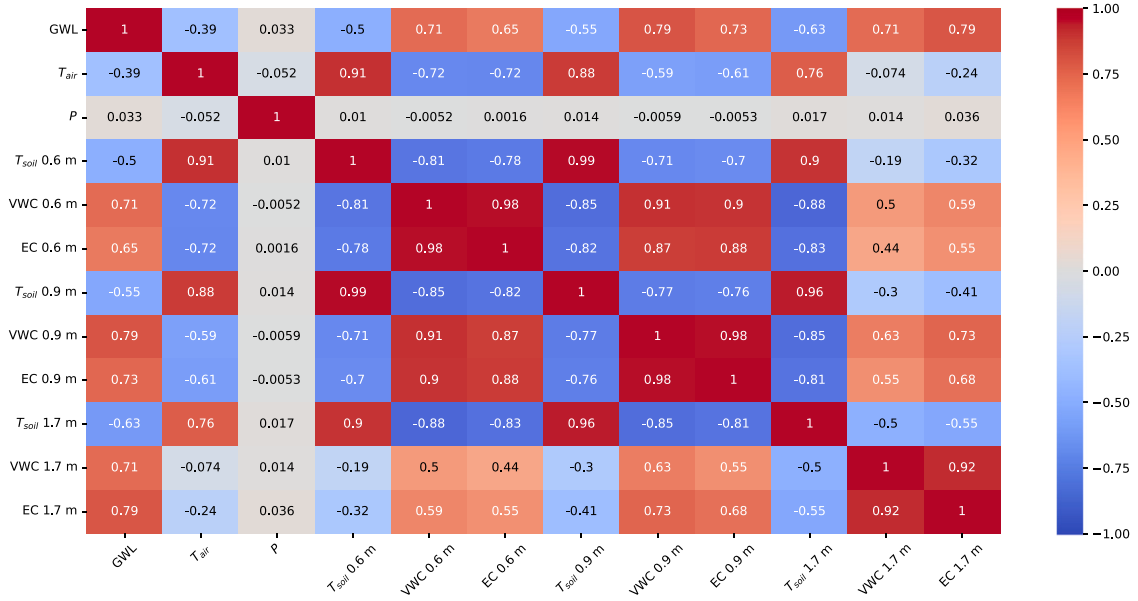


Fig. 6. Correlation matrix for the collected time series.

direct measure of the deviation between simulated and observed values, allowing for the characterization of model bias and dispersion. From the residual distribution, the mean residual $\bar{r}$, variance $\mathrm{Var}(r)$, and standard deviation $\mathrm{SD}(r)$ were computed as follows:

$$\mathrm{Var}(r) = \frac{1}{N-1}\sum_{t=1}^{N}(r_t - \bar{r})^2, \quad \mathrm{SD}(r) = \sqrt{\mathrm{Var}(r)} \tag{8}$$

where N is the total number of samples. These statistics provide insights into the spread and stability of the predictions, with lower variance and standard deviation indicating more consistent model behavior. To complement these measures, the Mean Absolute Percentage Error (MAPE) was computed to express the average relative deviation between observed and predicted groundwater levels:

$$\mathrm{MAPE} = \frac{100}{N}\sum_{t=1}^{N}\left|\frac{r_t}{y_t}\right| \tag{9}$$

Residual statistics were evaluated for each model and for each monitoring depth within the vadose zone, enabling the assessment of uncertainty propagation across soil layers. Lower residual variance and standard deviation were interpreted as indicators of improved robustness and reduced sensitivity to input variability, providing a quantitative measure of model reliability under changing hydrogeological conditions.

Table 2
Performance comparison of groundwater forecasting models across different depths. Best values for each metric and depth are shown in bold.

Sample	Depth (m)	Model	MAPE	ME	MAE	MPE	RMSE	Corr	NSE
Layer 1	–	ARIMA	0.247	−1.631	1.631	0.247	2.071	0.914	−16.660
		SARIMAX	0.084	0.530	0.530	**−0.084**	0.563	**0.921**	−0.308
		Prophet	**0.076**	**−0.175**	0.485	0.023	0.585	0.916	**−0.117**
		NeuralProphet	**0.076**	0.228	**0.469**	−0.031	**0.524**	0.896	−0.130
Layer 2	0.6	ARIMA	0.115	−0.749	0.761	0.112	1.073	0.910	−3.744
		SARIMAX	0.141	0.916	0.916	−0.141	1.020	0.465	−3.281
		Prophet	**0.055**	**−0.259**	**0.364**	**0.038**	**0.526**	**0.921**	−1.276
		NeuralProphet	0.090	0.250	0.561	**−0.038**	0.696	0.515	**−0.993**
Layer 2	0.9	ARIMA	0.111	−0.733	0.733	0.111	0.984	0.926	−2.987
		SARIMAX	0.135	0.876	0.876	−0.135	0.972	0.720	−2.889
		Prophet	**0.037**	**−0.073**	**0.235**	**0.009**	**0.287**	**0.950**	0.010
		NeuralProphet	0.052	−0.180	0.294	0.020	0.355	0.945	**0.380**
Laye 2	1.7	ARIMA	0.148	0.955	0.955	−0.148	1.025	0.952	−3.326
		SARIMAX	0.119	0.768	0.768	−0.119	0.815	**0.958**	−1.737
		Prophet	0.045	0.174	0.276	−0.030	**0.310**	0.953	**0.605**
		NeuralProphet	**0.041**	**−0.194**	**0.268**	**0.029**	0.340	**0.958**	0.523

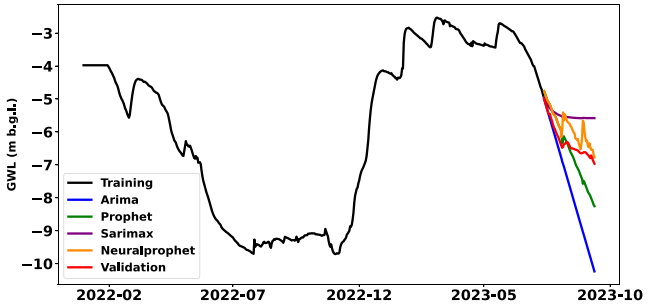
Table 3
Residual statistics for groundwater level forecasting models across different Layers and depths. Best values for each metric and depth are shown in bold.

Sample	Depth	Model	Mean residual	Variance	Std. Dev.	MAPE (%)
Layer 1	–	ARIMA	1.369	1.207	1.099	20.84
		SARIMAX	−0.764	0.139	0.373	11.86
		Prophet	**0.278**	0.347	0.589	7.94
		NeuralProphet	−0.402	**0.052**	**0.229**	**6.37**
Layer 2	0.6 m	ARIMA	0.782	0.580	0.761	11.84
		SARIMAX	−0.882	0.226	0.475	13.64
		Prophet	**0.293**	**0.205**	**0.453**	**5.38**
		NeuralProphet	−0.622	0.465	0.682	9.67
Layer 2	0.9 m	ARIMA	0.766	0.419	0.647	11.68
		SARIMAX	−0.843	0.199	0.445	13.06
		Prophet	**0.106**	0.072	0.267	**3.35**
		NeuralProphet	0.410	**0.050**	**0.223**	6.89
Layer 2	1.7 m	ARIMA	−0.922	0.156	0.395	14.35
		SARIMAX	−0.734	0.089	0.297	11.48
		Prophet	**−0.141**	**0.062**	**0.248**	**4.11**
		NeuralProphet	0.189	0.080	0.282	4.42

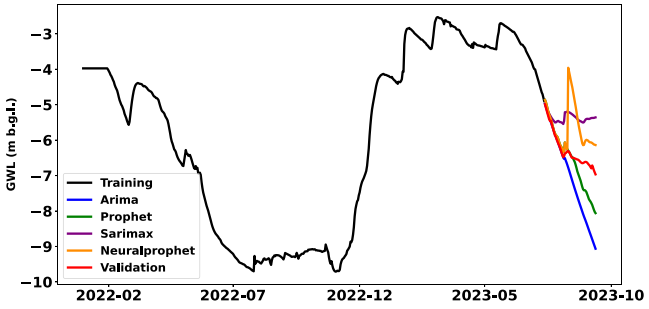
4. Results

The model performance was assessed across varying depths to evaluate predictive accuracy and reliability under different experimental conditions. Results of the model predictions associated with errors between observed and predicted GWL are reported in Table 2 and visually presented in Fig. 7. By using the variables collected in Layer 1, ARIMA exhibited the highest MAPE (0.247) and the lowest NSE (−16.660), indicating poor predictive accuracy and model reliability. SARIMAX, which incorporates seasonal adjustments, partially mitigates this issue, but still struggles with capturing external variability beyond its predefined seasonal windows. This explains its moderate improvement in correlation (0.921) but persistent error accumulation (RMSE: 0.563). SARIMAX and Prophet performed better, achieving MAPE values of 0.084 and 0.076, respectively. Looking at the graphical results (Fig. 7(a)), the graph indicates that shallow statistical models lack adaptive mechanisms, in comparison to Prophet and NeuralProphet, both relying on Fourier series-based seasonality extraction. These models, due to their capacity to isolate periodic GWL fluctuations without assuming a static seasonal pattern, are adaptable to highly non-linear patterns. Prophet demonstrated a slightly higher RMSE (0.585) compared to SARIMAX (0.563), though both models maintained strong correlations above 0.91, noting that SARIMAX does not visually capture much of the fluctuations from the original time series. NeuralProphet exhibited the lowest RMSE (0.524) and the lowest MAE (0.469), suggesting strong predictive accuracy. However, its correlation (0.896) is lower than the other models, indicating weaker capture of temporal dependencies. This is in agreement with the fact that the deep learning approaches are more robust on large sizes of datasets to get better temporal accuracy. Exploring the predictive models using Layer 2 with variables collected at 0.6 m depth, presented in Fig. 7(b), ARIMA exhibited slight performance improvements, reducing its MAPE to 0.115 while maintaining a good correlation (0.910). This improvement is expected, as shallow subsurface water fluctuations exhibit strong autoregressive behavior, making past values more informative at this depth. However, SARIMAX underperformed (MAPE: 0.141, Correlation: 0.465), indicating that its predefined seasonal adjustment fails to align with the dynamic, short-period oscillations.

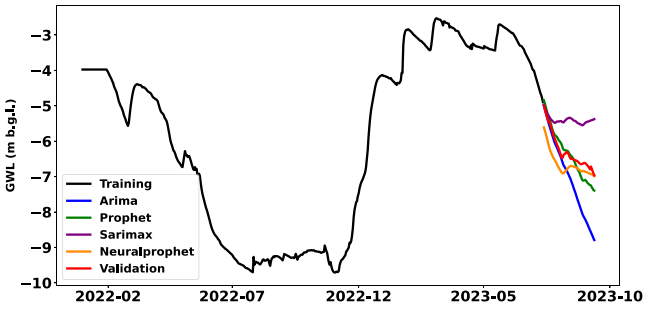
Even at this depth, it is still the case that visually, the statistical models still are not flexible to the nature of the time series. Prophet delivered a stable performance, achieving the lowest MAPE (0.055) and a high correlation (0.921). NeuralProphet, with MAPE 0.090 and correlation of 0.515, exhibited resilience but showed a distinctive sensitivity to high-frequency disturbances. Unlike Prophet, which decomposes trends and seasonality separately, NeuralProphet integrates autoregression within its neural architecture, making it more reactive to abrupt fluctuations but less effective in capturing the long-term smoothing effects of the regressors. Using the variables collected at 0.9 m (Fig. 7(c)), ARIMA and SARIMAX followed similar trends, with MAPE values of 0.111 and 0.135, respectively. Both models exhibited improved correlations (0.926 and 0.720), though their RMSE values remained relatively high. Prophet provided the best predictive performance, with



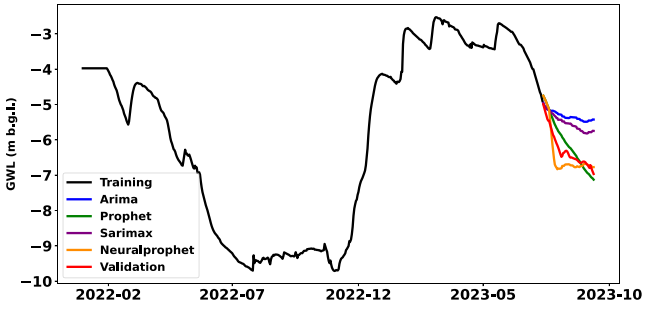
(a) *Layer 1 to predict Layer 3.*



(b) *Layer 2 at depth 0.6 m to predict Layer 3.*



(c) *Layer 2 at depth 0.9 m to predict Layer 3.*



(d) *Layer 2 at depth 1.7 m to predict Layer 3.*

Fig. 7. Performance comparison of forecasting Layer 3 (Saturated zone) from monitored variables in preceding layers.

the lowest MAPE (0.037), RMSE (0.287), and the highest correlation (0.950), confirming its robustness in predictions. NeuralProphet also exhibited strong performance with a MAPE of 0.052 and a correlation of 0.945, indicating its adaptability at this depth. At this depth, where the groundwater signal is moderately stable, NeuralProphet's AR component introduced minor overfitting to short-term fluctuations, limiting its ability to fully smooth out seasonal variations. In the last scenario (using variables of layer 2 collected at 1.7 m) Fig. 7(d), ARIMA's MAPE increased to 0.148, while SARIMAX and NeuralProphet displayed more stable performance, achieving MAPE values of 0.119 and 0.041, respectively. Prophet continued to perform well, with a MAPE of 0.045 and RMSE of 0.310. NeuralProphet outperformed all models at this depth, achieving the lowest MAPE (0.041), a high correlation (0.958), and an NSE of 0.523, demonstrating strong generalization capabilities and reliable predictions across varying depths. ARIMA relies purely on past GWL values without external inputs, making it more susceptible to residual fluctuations. SARIMAX, which incorporates external predictors, remained stable.

4.1. Residual analysis across vadose zone depths

Table 3 reports the residual statistics computed for all forecasting models across the Layers and the monitoring depths. The results provide a quantitative assessment of prediction stability and uncertainty propagation through the vadose zone. In general, residual variability decreases progressively with increasing depth, suggesting that measurements closer to the groundwater body are less affected by short-term meteorological fluctuations and exhibit more predictable dynamics. At the Layer 1, NeuralProphet achieved the lowest residual variance (0.052) and standard deviation (0.229), together with the lowest MAPE (6.37%), confirming its ability to capture short-term fluctuations while maintaining overall stability. In the Layer 2 at 0.6 and 0.9 m depths, Prophet consistently performed well compared to the other models, achieving near-zero mean residuals and the smallest MAPE values (5.38% and 3.35%, respectively). These results suggest that Prophet is particularly effective in modeling the mid-depth dynamics of the vadose zone, where both infiltration and retention processes interact. Layer 2 at 1.7 m depth, Prophet again yielded comparatively strong performance, with a mean residual of −0.141, variance of 0.062, standard deviation of 0.248, and a MAPE of 4.11%, highlighting its robustness in capturing the long-term and smoother groundwater responses. By contrast, traditional statistical models (ARIMA and SARIMAX) exhibited larger residual variances and higher MAPE values, at shallower depths, where short-term rainfall and evapotranspiration effects dominate. While SARIMAX reduced bias through the inclusion of exogenous variables, its overall variance remained higher than that of Prophet-based architectures. These findings highlight the advantage of hybrid and neural forecasting approaches over purely statistical ones in representing complex hydrogeological dynamics. Overall, the residual analysis supports the robustness of data-driven models that integrate seasonality, nonlinearity, and external forcing, confirming the advantage of Prophet and NeuralProphet architectures in capturing complex hydrogeological processes with higher accuracy and lower uncertainty. Prophet demonstrated greater adaptability across depths, while NeuralProphet exhibited superior stability and minimal residual variance, making both models promising candidates for operational groundwater level forecasting (see Fig. 8).

5. Discussion

The correlation and trend analysis conducted before developing the forecasting models provided a standardized visual framework for assessing variability in the input features. This analysis confirmed that external regressors in the vadose zone exhibit associations with GWL, with noise decreasing in Layer 2 as depth increases. These statistical findings suggest that incorporating these regressors into the forecasting model setup is beneficial, potentially enhancing predictive accuracy. In this context, it is important to emphasize that the variable (P) monitored in Layer 1, which corresponds to precipitation, is the least correlated regressor with fluctuations in the GWL. Since rainfall is a discrete variable, it does not directly correlate with groundwater level oscillations, negatively impacting forecasting in the developed models. This result aligns with several authors who have developed predictive models based on observed groundwater levels and external atmospheric variables (e.g., precipitation) (Sun et al., 2022; Yin et al., 2021), highlighting the limitations of this approach. The correlations between GWL and external regressors used in the models increase when

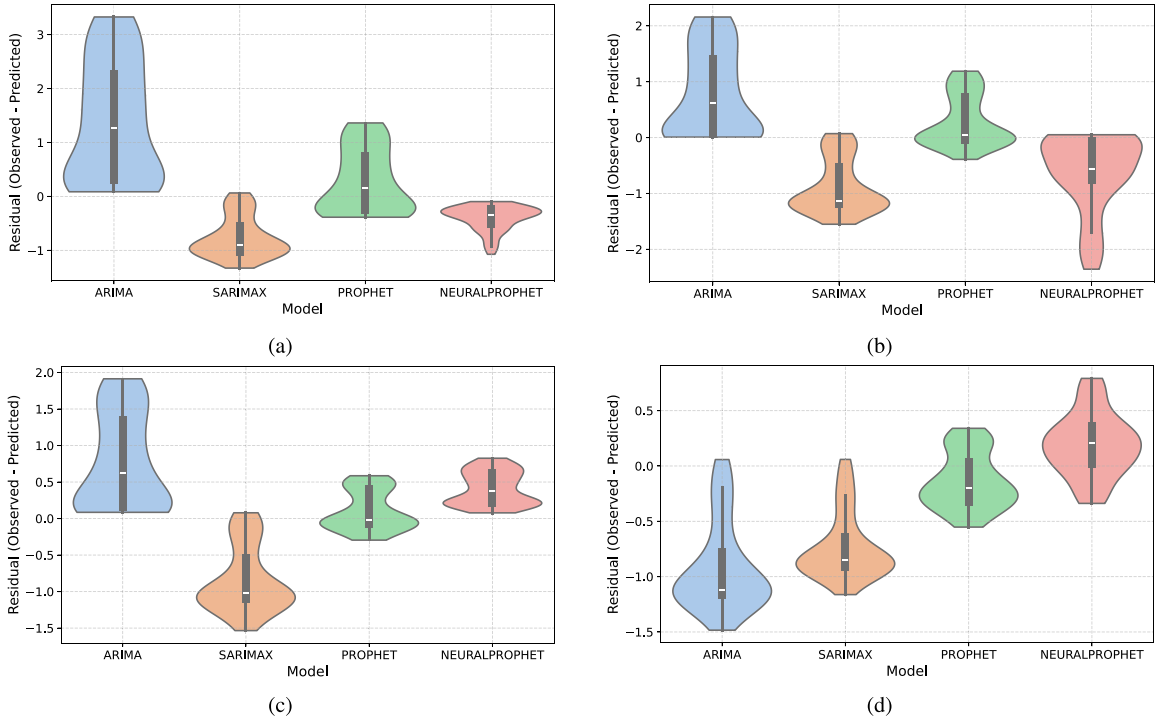


Fig. 8. Distribution of residuals across models (ARIMA, SARIMAX, Prophet, and NeuralProphet) at different monitoring depths: (a) Layer 1, Layer 2 at depth (b) 0.6 m, (c) 0.9 m, and (d) 1.7 m. Each violin plot summarizes the residual spread and central tendency, highlighting the variance and bias characteristics of each model.

shifting to variables from Layer 2. Indeed, our approach, based on a spatio-temporal high-resolution hydrological monitoring system in the vadose zone, overcomes the limitations of using discrete variables such as rainfall and air temperature to predict groundwater availability, providing time series that mimic GWL, to be used as exogenous variables. Across varying depths in Layer 2, the general performance of forecasting models is found to be influenced by both their underlying architectures and the stability of key regressors. If using hydrogeological variables collected at shallower depths to predict GWL, the modeled GWL time series displays high-frequency fluctuations deeply influenced by external factors like sharp temperature variation, or sudden response of VWC and EC to rainfall events. The moderate to strong positive correlation between GWL, VWC, and EC (Fig. 6) provides critical insight into the dynamics of infiltration and recharge within the vadose zone. The results obtained from the high-resolution hydrogeological monitoring system reveal that variations in these parameters at different depths reflect the temporal and spatial progression of infiltrating water through unsaturated soil layers toward the saturated zone. The interplay among these variables governs the hydrological response of the system, linking surface meteorological inputs with subsurface groundwater fluctuations (Dahan, 2020; Singh et al., 2018). Infiltration begins when precipitation exceeds the infiltration capacity of the soil, allowing water to percolate vertically through the vadose zone (Seiler and Gat, 2007; Harter et al., 2004). The upper layers (0.6 m depth) are characterized by rapid changes in VWC immediately following rainfall events, indicating prompt but transient wetting responses. These fluctuations are largely controlled by atmospheric conditions, evaporation, and plant transpiration, producing short-term increases in moisture content and EC due to ion dissolution during wetting phases. However, this layer is strongly influenced by evapotranspiration, which reduces its contribution to effective recharge. The intermediate depth (0.9 m) exhibits a delayed VWC and EC response compared to the surface, reflecting the attenuation of short-term fluctuations and the beginning of downward infiltration continuity. Soil temperature and EC variations at this depth indicate ongoing percolation, although evapotranspiration effects are still present. At greater depth (1.7 m), the behavior of VWC and EC becomes markedly smoother and more stable. Here, the influence of atmospheric variability decreases, and the observed changes correspond more closely with GWL oscillations. The lagged increase in VWC and EC at this depth coincides with subsequent rises in GWL, signifying that infiltration fronts have reached the lower vadose zone and contributed to aquifer recharge. This pattern indicates the dominance of effective percolation over surface processes and the transition from transient storage to recharge flow. EC in the soil is a powerful integrative indicator of water movement and solute transport in the vadose zone (Cassiani et al., 2006; Poulain et al., 2018). Its variations result from both changes in soil moisture (VWC) and the ionic concentration of pore water. Because EC reflects the combined effects of hydrological and geochemical processes, it often exhibits a measurable correlation with groundwater level (GWL) oscillations, particularly when infiltration fronts propagate through the unsaturated zone and influence aquifer recharge (Singh et al., 2018; Dahan, 2020). When rainfall infiltrates, the advancing wetting front increases the degree of saturation, causing an immediate rise in EC due to enhanced electrical connectivity between soil particles and the dissolution of surface-bound ions (Lin, 2010). Therefore, fluctuations in EC can be interpreted as a leading indicator of groundwater level changes, especially at monitoring depths where infiltration signals are not obscured by rapid surface processes (Harter et al., 2004). Fronzi et al. (2024) highlight that the strongest correlation between vadose zone variables and groundwater level ($r = 0.923$) occurs at 1.7 m depth, confirming that deeper soil layers act as integrative indicators of infiltration reaching the saturated zone. In contrast, shallow sensors primarily capture short-lived hydrodynamic and thermal effects, introducing noise related to surface energy exchanges. These results emphasize the importance of depth-dependent monitoring of VWC and EC to distinguish between superficial soil moisture variations and true recharge signals influencing groundwater dynamics. Overall, the observed relationships among GWL, VWC, and EC delineate the natural infiltration process from the atmosphere through the vadose zone to

the aquifer. The sequential increase in VWC from upper to deeper layers, accompanied by a corresponding stabilization of EC and delayed GWL rise, reflects the downward migration of wetting fronts and the temporal coupling between unsaturated and saturated zone processes. Understanding these interactions provides a basis for refining predictive models of groundwater level fluctuations and supports more accurate assessments of aquifer recharge under variable climatic conditions. This integrated approach confirms that continuous multi-depth monitoring in the vadose zone is essential for characterizing infiltration pathways, quantifying recharge efficiency, and improving the management of groundwater resources in hydrologically sensitive environments.

Some previous research on GWL prediction explored meteorological and hydrological inputs. However, despite their success, these models primarily relied on surface meteorological variables and mathematical decomposition techniques to represent subsurface dynamics, which limits the ability to capture the true physical processes linking recharge and groundwater response. While modeling based solutions increasingly remain the area of focus in many current literature, as shown by Najafabadipour et al. (2022) and Chen et al. (2023), however, research still is progressively optimistic about vadose zone for modeling (Stewart et al., 2025). Our work improves on the existing body of research by introducing vadose-zone monitoring as a new predictive pathway and by integrating it with modern decomposable and neural forecasting models. Previous studies, that used soft-computing and heuristic modeling approaches using meteorological and hydrological inputs such as Shiri and Kişi (2011), compared Genetic Programming (GP) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) for short-term water table prediction, showing that heuristic models can effectively capture nonlinear relationships but require careful lag selection. Similarly, Kisi and Shiri (2012) enhanced ANFIS through wavelet decomposition, demonstrating that separating high- and low-frequency components improves prediction stability. In contrast, in our study Prophet and NeuralProphet models achieve report similar accuracy without heuristic calibration. Their decomposable and neural autoregressive architectures inherently model trend, seasonality, and nonlinear dependencies while maintaining interpretability. Moreover, by incorporating vadose-zone regressors (VWC and EC), our framework achieves a level of physical transparency by capturing the physical infiltration process through the data compared to mathematical methods such as GP, directly linking predictive signals to subsurface hydrological dynamics. Compared with the soft-computing techniques assessed by Shiri et al. (2013), our models rely on direct vadose-zone observations rather than statistically optimized lags. The observed improvement on the predictive performance, with increasing depth reflects the role of deeper vadose layers as natural low-pass filters, reducing short-term meteorological noise and emphasizing recharge-driven signals. Finally, consistent with the findings of Shiri et al. (2022), who showed that adding tidal information as an external regressor improved coastal aquifer predictions, our results confirm that including physically meaningful exogenous variables such as vadose-zone parameters improves the forecasting performance. Overall, Prophet and NeuralProphet, supported by vadose-zone data, provide a robust, interpretable, and physically grounded alternative to traditional statistical and heuristic groundwater models. Collectively, by integrating physical sensing with advanced data-driven modeling, the present study bridges the gap between heuristic based computing and hydrological process understanding, offering a robust and scalable approach for short-term groundwater forecasting.

6. Conclusions

The results demonstrate that model performance is significantly influenced by hydrogeological aspects such as depth-dependent noise levels and time-dependent hydrodynamic processes in terms of precipitation-infiltration-recharge dynamics. Our approach, leveraging a high-resolution spatio-temporal hydrological monitoring system in the vadose zone, effectively overcomes the limitations of discrete variables like rainfall in predicting groundwater dynamics. By providing continuous time series that closely mimic GWL patterns, our method enhances the accuracy and reliability of groundwater level predictions, offering a more robust framework for hydrological modeling. Traditional statistical models such as ARIMA and SARIMAX exhibited limitations in capturing complex hydrological dependencies where external regressors like precipitation and air temperature are less stable. Prophet consistently performed well across all depths due to its structured trend-seasonality decomposition, while NeuralProphet outperformed all models at higher depths in the vadose zone where reduced high-frequency fluctuations allowed its neural autoregressive component to leverage long-term dependencies effectively. Our study features hyperparameter tuning via K-fold rolling cross-validation, which is essential for optimizing model performance and ensuring robustness under limited data scenarios. Future research should explore hybrid approaches integrating machine learning with physically-based hydrological models to improve the interpretability of the data-driven approaches, especially for critical groundwater management applications.

CRediT authorship contribution statement

Alessandro Galdelli: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Formal analysis, Conceptualization. **Davide Fronzi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Gagan Narang:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis. **Adriano Mancini:** Writing – review & editing, Supervision, Resources, Project administration. **Alberto Tazioli:** Writing – review & editing, Supervision, Project administration.

Computer code availability

Name of the code/library: Groundwater-level-forecasting
Contact e-mail: a.galdelli@univpm.it, g.narang@pm.univpm.it
Date first available: March 13, 2025
Software required: The program requires Python 3.8 and the necessary libraries listed in requirements.txt, including Pandas, NumPy, Scikit-learn, Matplotlib, Statsmodels, Prophet, and NeuralProphet and containerized in a Docker environment.
Hardware requirements: A multi-core CPU, at least 8 GB RAM, and 10 GB of free disk space are required to run the forecasting models. A GPU is optional, but it may benefit NeuralProphet. These specifications ensure efficient data processing, model training, and evaluation.
Program language: Python
Program size: For standard use, at least 10 GB of free disk space is recommended.
The source codes are available for downloading at the link: https://github.com/vrai-group/Groundwater-level-forecasting
Documentation: Detailed documentation for application installation, testing, and deployment can be found at https://github.com/vrai-group/Groundwater-level-forecasting/blob/main/README.md
Data availability: The time series data used for model training collected from the thermo-pluviometer stations and the three probes, are not publicly available. The test data can be accessed here: https://github.com/vrai-group/Groundwater-level-forecasting/tree/main/data_test, while the forecasted data by all models can be accessed here: https://github.com/vrai-group/Groundwater-level-forecasting/tree/main/forecast_data

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to acknowledge Viva Servizi S.p.A., Consorzio Gorgovivo Azienda Speciale, and Stefano Palpacelli.

Data availability

The source codes and test data are available for downloading at the link: https://github.com/vrai-group/Groundwater-level-forecasting.

References

Afrifa, S., Zhang, T., Appiahene, P., Varadarajan, V., 2022. Mathematical and machine learning models for groundwater level changes: A systematic review and bibliographic analysis. Futur. Internet 14 (9).

Aguilera, H., Guardiola-Albert, C., Naranjo-Fernández, N., Kohfahl, C., 2019. Towards flexible groundwater-level prediction for adaptive water management: using Facebook's prophet forecasting approach. Hydrol. Sci. J. 64 (12).

Amiri, S., Rajabi, A., Shabanlou, S., Yosefvand, F., Izadbakhsh, M.A., 2023. Prediction of groundwater level variations using deep learning methods and GMS numerical model. Earth Sci. Inform. 1–15.

Busico, G., Colombani, N., Fronzi, D., Pellegrini, M., Tazioli, A., Mastrocicco, M., 2020. Evaluating SWAT model performance, considering different soils data input, to quantify actual and future runoff susceptibility in a highly urbanized basin. J. Environ. Manag. 266, 110625.

Busico, G., Fronzi, D., Colombani, N., Mastrocicco, M., Tazioli, A., 2024. Identification and quantification of nutrients sources in the Aspio watershed (Italy). Insight from geogenic mineralization and anthropogenic pressure. Catena 236, 107759.

Cassiani, G., Binley, A., Ferré, T.P., 2006. Unsaturated zone processes. In: Vereecken, H., Binley, A., Cassiani, G., Revil, A., Titov, K. (Eds.), Applied Hydrogeophysics. Springer Netherlands, Dordrecht, pp. 75–116.

Chemeri, L., Taussi, M., Fronzi, D., Cabassi, J., Mazzoli, S., Tazioli, A., Renzulli, A., Vaselli, O., 2025. Groundwater hydrogeochemical changes predating and following the november 9, 2022 Mw 5.5 adriatic offshore earthquake (central Italy). J. Hydrol. 132792.

Chen, C., Li, B., Zhang, H., Zhao, M., Liang, Z., Li, K., An, X., 2025. Uncertainty-informed multi-reservoir flood control optimization: A probabilistic forecasting and stochastic decision-making framework. J. Hydrol. 663, 134285.

Chen, H.-Y., Vojinovic, Z., Lo, W., Lee, J.-W., 2023. Groundwater level prediction with deep learning methods. Water 15 (17), 3118.

Connor, R., 2015. The United Nations World Water Development Report 2015: Water for a Sustainable World, vol. 1, UNESCO publishing.

Cos, J., Doblas-Reyes, F., Jury, M., Marcos, R., Bretonnière, P.-A., Samsó, M., 2022. The mediterranean climate change hotspot in the CMIP5 and CMIP6 projections. Earth Syst. Dyn. 13 (1), 321–340.

Dadhich, A.P., Goyal, R., Dadhich, P.N., 2021. Assessment and prediction of groundwater using geospatial and ANN modeling. Water Resour. Manag. 35.

Dahan, O., 2020. Vadose zone monitoring as a key to groundwater protection. Front. Water Volume 2 - 2020.

Dalton, A., Bekker, B., 2022. Exogenous atmospheric variables as wind speed predictors in machine learning. Appl. Energy 319, 119257.

Dangar, S., Asoka, A., Mishra, V., 2021. Causes and implications of groundwater depletion in India: A review. J. Hydrol. 596, 126103.

Davamani, V., John, J.E., Poornachandhra, C., Gopalakrishnan, B., Arulmani, S., Parameswari, E., Santhosh, A., Srinivasulu, A., Lal, A., Naidu, R., 2024. A critical review of climate change impacts on groundwater resources: A focus on the current status, future possibilities, and role of simulation models. Atmosphere 15 (1), 122.

De Fraiture, C., Molden, D., Wichelns, D., 2010. Investing in water for food, ecosystems, and livelihoods: An overview of the comprehensive assessment of water management in agriculture. Agricult. Water. Manag. 97 (4), 495–501.

Djimadoumngar, K.-N., 2023. Parallel investigations of remote sensing and ground-truth lake Chad's level data using statistical and machine learning methods. Appl. Comput. Geosci. 20, 100135.

Dramsch, J.S., 2020. 70 years of machine learning in geoscience in review. Adv. Geophys. 61, 1–55.

Fiorillo, F., Petitta, M., Preziosi, E., Rusi, S., Esposito, L., Tallini, M., 2015. Long-term trend and fluctuations of karst spring discharge in a Mediterranean area (central-southern Italy). Environ. Earth Sci. 74, 153–172.

Fratianni, S., Acquaotta, F., 2017. The climate of Italy. In: Landscapes Landforms Italy. pp. 29–38.

Fronzi, D., Gaiolini, M., Mammoliti, E., Colombani, N., Palpacelli, S., Marcellini, M., Tazioli, A., 2022. Groundwater-surface water interaction revealed by meteorological trends and groundwater fluctuations on stream water level. Acque Sotterranee-Italian J. Groundw. 11 (2), 19–28.

Fronzi, D., Narang, G., Galdelli, A., Pepi, A., Mancini, A., Tazioli, A., 2024. Towards groundwater-level prediction using prophet forecasting method by exploiting a high-resolution hydrogeological monitoring system. Water 16 (1), 152.

Galdelli, A., Mancini, A., Tassetti, A., Vega, C.F., Armelloni, E., Scarcella, G., Fabi, G., Zingaretti, P., 2019. A cloud computing architecture to map trawling activities using positioning data 15th IEEE/ASME int. In: Conf. Mechatron. Embed. Syst. Appl. Vol. 9.

Galdelli, A., Narang, G., Migliorelli, L., Izzo, A.D., Mancini, A., Zingaretti, P., 2023. An AI-driven prototype for groundwater level prediction: Exploring the gorgovivo spring case study. In: International Conference on Image Analysis and Processing. Springer, pp. 418–429.

Garbrecht, J., Fernandez, G.P., 1994. Visualization of trends and fluctuations in climatic records 1. JAWRA J. Am. Water Resour. Assoc. 30 (2), 297–306.

Goderniaux, P., Brouyere, S., Blenkinsop, S., Burton, A., Fowler, H.J., Orban, P., Dassargues, A., 2011. Modeling climate change impacts on groundwater resources using transient stochastic climatic scenarios. Water Resour. Res. 47 (12).

Harter, T., Hopmans, J.W., Feddes, R., et al., 2004. Role of Vadose Zone Flow Processes in Regional Scale Hydrology: Review, Opportunities and Challenges. Kluwer.

Hendrickx, J.M., Flury, M., et al., 2001. Uniform and preferential flow mechanisms in the vadose zone. Concept. Model. Flow Transp. Fract. Vadose Zone 149–187.

Hopmans, J., Van Genuchten, M., 2005. Vadose zone: Hydrological process. Encycl. Soils Environ. 209–216.

Jain, S.K., Sudheer, K., 2008. Fitting of hydrologic models: a close look at the Nash–Sutcliffe index. J. Hydrol. Eng. 13 (10), 981–986.

Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H.A., Kumar, V., 2018. Machine learning for the geosciences: Challenges and opportunities. IEEE Trans. Knowl. Data Eng. 31 (8), 1544–1554.

Khan, J., Lee, E., Balobaid, A.S., Kim, K., 2023. A comprehensive review of conventional, machine leaning, and deep learning models for groundwater level (GWL) forecasting. Appl. Sci. 13 (4), 2743.

Khozani, Z.S., Banadkooki, F.B., Ehteram, M., Ahmed, A.N., El-Shafie, A., 2022. Combining autoregressive integrated moving average with long short-term memory neural network and optimisation algorithms for predicting ground water level. J. Clean. Prod. 348.

Kisi, O., Shiri, J., 2012. Wavelet and neuro-fuzzy conjunction model for predicting water table depth fluctuations. Hydrol. Res. 43 (3), 286–300.

Korstanje, J., 2021. Advanced forecasting with Python. Springer.

Lai, C.-Y., Ning, Y.-C., Boning, D.S., 2025. RDIT: Residual-based diffusion implicit models for probabilistic time series forecasting. (Accessed 8 October 2025), arXiv preprint: arXiv:2509.02341.

Lary, D.J., Alavi, A.H., Gandomi, A.H., Walker, A.L., 2016. Machine learning in geosciences and remote sensing. Geosci. Front. 7 (1), 3–10.

Lin, H., 2010. Earth's critical zone and hydropedology: concepts, characteristics, and advances. Hydrol. Earth Syst. Sci. 14 (1), 25–45.

Maity, R., Srivastava, A., Sarkar, S., Khan, M.I., 2024. Revolutionizing the future of hydrological science: Impact of machine learning and deep learning amidst emerging explainable AI and transfer learning. Appl. Comput. Geosci. 24, 100206.

Mammoliti, E., Di Stefano, F., Fronzi, D., Mancini, A., Malinverni, E.S., Tazioli, A., 2022. A machine learning approach to extract rock mass discontinuity orientation and spacing, from laser scanner point clouds. Remote. Sens. 14 (10), 2365.

Mancini, A., Cosoli, G., Galdelli, A., Violini, L., Pandarese, G., Mobili, A., Blasi, E., Tittarelli, F., Revel, G.M., 2023. A monitoring platform for the built environment: towards the development of an early warning system in a seismic context. In: 2023 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv). IEEE, pp. 102–106.

Maniar, H., Ryali, S., Kulkarni, M.S., Abubakar, A., 2018. Machine-learning methods in geoscience. In: SEG International Exposition and Annual Meeting. SEG, pp. SEG–2018.

Milly, P.C., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., 2008. Stationarity is dead: Whither water management? Science 319 (5863), 573–574.

Mishra, R.K., 2023. Fresh water availability and its global challenge. Br. J. Multidiscip. Adv. Stud. 4 (3), 1–78.

Morelli, S., Bonì, R., Guidi, E., De Donatis, M., Pappafico, G., Francioni, M., et al., 2023. L'alluvione delle marche del 15 settembre 2022, cause e conseguenze. Cult. Territ. Linguaggi 24, 136–147.

Mussi, M., Nanni, T., Tazioli, A., Vivalda, P.M., 2017. The Mt Conero limestone ridge: The contribution of stable isotopes to the identification of the recharge area of aquifers. Ital. J. Geosci. 136 (2), 186–197.

Najafabadipour, A., Kamali, G., Nezamabadi-pour, H., 2022. The innovative combination of time series analysis methods for the forecasting of groundwater fluctuations. Water Resour. 49 (2), 283–291.

Negre, P., Alonso, R.S., Prieto, J., García, Ó., de-la Fuente-Valentín, L., 2024. Prediction of footwear demand using prophet and SARIMA. Expert Syst. Appl. 124512.

Poulain, A., Watlet, A., Kaufmann, O., Van Camp, M., Jourde, H., Mazzilli, N., Rochez, G., Deleu, R., Quinif, Y., Hallet, V., 2018. Assessment of groundwater recharge processes through karst vadose zone by cave percolation monitoring. Hydrol. Process. 32 (13), 2069–2083.

Rajaee, T., Ebrahimi, H., Nourani, V., 2019. A review of the artificial intelligence methods in groundwater level modeling. J. Hydrol. 572, 336–351.

Ren, H., Cromwell, E., Kravitz, B., Chen, X., 2022. Using long short-term memory models to fill data gaps in hydrological monitoring networks. Hydrol. Earth Syst. Sci. 26 (7), 1727–1743.

Rocke, R.J.A., Zhang, L., BEng, A., Jackson, I., Sáenz, M.J., Sheffi, Y., 2024. Impact of exogenous variables on AI/ML predictive algorithm prophet, in the FMCG industry. Signature.

Sarma, R., Singh, S., 2022. A comparative study of data-driven models for groundwater level forecasting. Water Resour. Manag. 36 (8), 2741–2756.

Seiler, K.-P., Gat, J.R., 2007. Groundwater Recharge from Run-Off, Infiltration and Percolation, vol. 55, Springer Science & Business Media.

Sharafati, A., Asadollah, S.B.H.S., Neshat, A., 2020. A new artificial intelligence strategy for predicting the groundwater level over the Rafsanjan aquifer in Iran. J. Hydrol. 591, 125468.

Shiri, J., Kişi, Ö., 2011. Comparison of genetic programming with neuro-fuzzy systems for predicting short-term water table depth fluctuations. Comput. Geosci. 37 (10), 1692–1701.

Shiri, J., Kisi, O., Yoon, H., Kazemi, M.H., Shiri, N., Poorrajabali, M., Karimi, S., 2022. Prediction of groundwater level variations in coastal aquifers with tide and rainfall effects using heuristic data driven models. ISH J. Hydraul. Eng. 28 (sup1), 188–198.

Shiri, J., Kisi, O., Yoon, H., Lee, K.-K., Hossein Nazemi, A., 2013. Predicting groundwater level fluctuations with meteorological effect implications—A comparative study among soft computing techniques. Comput. Geosci. 56, 32–44.

Shirmohammadi, B., Vafakhah, M., Moosavi, V., Moghaddamnia, A., 2013. Application of several data-driven techniques for predicting groundwater level. Water Resour. Manag. 27, 419–432.

Singh, G., Kaur, G., Williard, K., Schoonover, J., Kang, J., 2018. Monitoring of water and solute transport in the Vadose zone: A review. Vadose Zone J. 17 (1), 160058.

Šrajbek, M., Đurin, B., Sušilović, P., Singh, S.K., 2023. Application of the RAPS method for determining the dependence of nitrate concentration in groundwater on the amount of precipitation. Earth 4 (2), 266–277.

Stewart, R.D., Flury, M., Ajami, H., Anderson, R.G., Green, T.R., Jin, Y., Patrignani, A., Shillito, R., Zhang, W., Najm, M.R.A., Babaeian, E., Berli, M., Brookshire, E.N.J., Daigh, A.L.M., Franklin, S., Giovando, J., Heinse, R., Heitman, J., Huang, J., Kelleners, T., Naseri, M., Ochsner, T.E., Radolinski, J., Sadeghi, M., Sasidharan, S., Shukla, M.K., Tuller, M., Wendroth, O., Wu, J., Wyatt, B.M., Yang, Y., Yu, Y., Zhang, Z.F., 2025. Emerging issues and research opportunities in vadose zone processes. Vadose Zone J. 24 (4), e70030.

Sun, J., Hu, L., Li, D., Sun, K., Yang, Z., 2022. Data-driven models for accurate groundwater level prediction and their practical significance in groundwater management. J. Hydrol. 608, 127630.

Talebpour, N., Ilbeigi, M., 2025. Modeling and predicting uncertainty in tidal turbine power output: A data-driven time-series approach. IEEE Access 13, 82986–82994.

Taylor, S.J., Letham, B., 2018. Forecasting at scale. The American Statistician 72 (1), 37–45.

Tian, N., Shao, B., Bian, G., Zeng, H., Li, X., Zhao, W., 2024. Application of forecasting strategies and techniques to natural gas consumption: A comprehensive review and comparative study. Eng. Appl. Artif. Intell. 129, 107644.

Triebe, O., Hewamalage, H., Pilyugina, P., Laptev, N., Bergmeir, C., Rajagopal, R., 2021. Neuralprophet: Explainable forecasting at scale. (Accessed 8 October 2025), arXiv preprint arXiv:2111.15397.

Tuel, A., Eltahir, E.A., 2020. Why is the Mediterranean a climate change hot spot? J. Clim. 33 (14), 5829–5843.

Turkeltaub, T., Kurtzman, D., Bel, G., Dahan, O., 2015. Examination of groundwater recharge with a calibrated/validated flow model of the deep vadose zone. J. Hydrol. 522, 618–627.

Vadiati, M., Rajabi Yami, Z., Eskandari, E., Nakhaei, M., Kisi, O., 2022. Application of artificial intelligence models for prediction of groundwater level fluctuations: Case study (Tehran-Karaj alluvial aquifer). Environ. Monit. Assess. 194 (9), 619.

Yan, Z., Lu, X., Wu, L., 2023. Exploring the effect of meteorological factors on predicting hourly water levels based on CEEMDAN and LSTM. Water 15 (18), 3190.

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. J. Hydrol. 602, 126735.

Zarinmehr, H., Tizro, A.T., Fryar, A.E., Pour, M.K., Fasihi, R., 2022. Prediction of groundwater level variations based on gravity recovery and climate experiment (GRACE) satellite data and a time-series analysis: a case study in the Lake Urmia basin, Iran. Environ. Earth Sci. 81 (6), 180.

Zhang, Z., Bao, N., Yan, X., Zhu, A., Li, C., Liu, M., 2023. Implementing a new fully stepwise decomposition-based sampling technique for the hybrid water level forecasting model in real-world application. (Accessed 8 October 2025), arXiv preprint arXiv:2309.10658.