

¹ Spatio-Temporal Graph Neural Networks for Regional Groundwater ² Level Forecasting: A Case Study of the Haouz Aquifer, Morocco

³ Lahcen Hicham^a, Lhoussaine ELMEZOUARY^b

⁴

⁵ ARTICLE INFO

⁶

⁷ **Keywords:**

⁸ Graph convolutional network

⁹ Long short-term memory

¹⁰ Groundwater forecasting

¹¹ Spatiotemporal graph neural networks

¹²

¹³

¹⁴

¹⁵

¹⁶

¹⁷

¹⁸

¹⁹

²⁰

ABSTRACT

Groundwater plays a critical role in sustaining ecosystems, agriculture, and human water demand. Accurate forecasting of groundwater levels is essential for sustainable water resource management, particularly in regions experiencing climate change and increasing demand. Traditional time-series and statistical models often struggle to capture the nonlinear dependencies and spatial interactions inherent in groundwater systems. This paper explores the application of Spatio-Temporal Graph Neural Networks (STGNNs) for groundwater level forecasting. By modeling monitoring wells as nodes, and node features consist of groundwater level observations combined with relevant climatic variables. Spatial relationships between wells are encoded through edges defined by geographic proximity. STGNNs effectively capture both spatial dependencies and temporal dynamics. The findings highlight the potential of graph-based deep learning methods as a valuable tool for groundwater monitoring and management.

²¹ CRediT authorship contribution statement

²² Lahcen Hicham: contribution . Lhoussaine ELMEZOUARY: contribution .

²³ 1. Introduction

²⁴ Groundwater is an indispensable natural resource that sustains ecosystems, supports intensive agriculture, and
²⁵ fulfills domestic and industrial demands worldwide. As one of the primary freshwater reservoirs, it plays a critical role
²⁶ in mitigating the impacts of droughts and maintaining water security, particularly in semi-arid and arid regions where
²⁷ surface water availability is increasingly erratic [Scanlon et al. \(2023\)](#). However, mounting anthropogenic pressures and
²⁸ the accelerating effects of climate change manifested through irregular precipitation patterns, rising temperatures, and
²⁹ land-use transformations pose unprecedented challenges to groundwater sustainability [Taylor et al. \(2013\)](#); [Famiglietti](#)
³⁰ [\(2014\)](#). Over-extraction for irrigation and inadequate recharge have contributed to alarming declines in groundwater
³¹ levels (GWLS) across many major aquifers, leading to long-term ecological degradation and severe socio-economic
³² consequences [Wada et al. \(2010\)](#); [Mukherjee et al. \(2024\)](#).

³³ 1.1. Challenges in Groundwater Modeling

³⁴ In this context of scarcity and stress, accurate and timely forecasting of GWLs is a prerequisite for sustainable water
³⁵ resource management, efficient irrigation planning, and drought mitigation [Sun et al. \(2022\)](#). Yet, modeling ground-
³⁶ water dynamics remains a formidable challenge due to the nonlinear interactions between meteorological drivers, hy-
³⁷ drological processes, geological heterogeneity, and anthropogenic interventions such as pumping [Sophocleous \(2002\)](#).

ORCID(s): 0009-0007-8193-5359 (L. Hicham)

38 Historically, hydrogeologists have relied on physically-based models (e.g., MODFLOW). While these models are
 39 grounded in physical laws and provide robust understanding of flow dynamics [Anderson et al. \(2015\)](#), they require
 40 extensive site-specific hydrogeological data—often unavailable in data-scarce regions—and involve computationally
 41 intensive parameterization and calibration processes [Refsgaard \(1997\)](#). Alternatively, geostatistical approaches, while
 42 valuable for spatial interpolation, typically assume stationarity and linear relationships, limiting their capacity to ex-
 43 trapolate under the non-stationary conditions induced by climate change.

44 **1.2. The Evolution from Temporal ML to Spatio-Temporal Deep Learning**

45 To overcome the limitations of physical and statistical models, data-driven Machine Learning (ML) and Deep
 46 Learning (DL) approaches have gained prominence in recent decades. Early applications utilizing Artificial Neural
 47 Networks (ANNs), Support Vector Machines (SVMs), and Random Forests (RF) demonstrated superior performance
 48 in capturing nonlinear relationships compared to traditional multiple linear regression [Nayak et al. \(2006\)](#); [Sahoo and](#)
 49 [Jha \(2013\)](#). More recently, Recurrent Neural Networks (RNNs), and specifically Long Short-Term Memory (LSTM)
 50 networks, have become the state of the art for hydrological time series forecasting due to their ability to learn long
 51 term temporal dependencies [Lin et al. \(2022\)](#).

52 However, a critical methodological gap persists: standard deep learning models like LSTM typically treat monitor-
 53 ing wells as isolated entities. They rely exclusively on temporal sequences, ignoring the *spatial dependencies* inherent
 54 in an aquifer system [Li et al. \(2023\)](#). Groundwater levels in a monitoring network are not independent; they are physi-
 55 cally interconnected through hydraulic gradients, where fluctuations in one well are influenced by pumping, recharge,
 56 and geological conditions at neighboring locations [Chang et al. \(2025\)](#). Neglecting this spatial interconnectivity lim-
 57 its the predictive accuracy and physical interpretability of forecasting models, particularly in complex, over-exploited
 58 aquifers.

59 **1.3. The Emergence of Spatio-Temporal Graph Neural Networks (STGNNs)**

60 To address the dual challenge of spatial complexity and temporal dynamism, Graph Neural Networks (GNNs)
 61 have emerged as a transformative framework. By representing monitoring wells as nodes and their hydrogeological
 62 relationships as edges in a graph structure, GNNs explicitly encode spatial dependencies [Scarselli et al. \(2009\)](#); [Kipf](#)
 63 and [Welling \(2017\)](#). Recent advancements have extended this paradigm to Spatio-Temporal Graph Neural Networks
 64 (STGNNs), which integrate graph convolutions (to capture spatial features) with sequence learning modules (to capture
 65 temporal dynamics) [Yu et al. \(2018\)](#); [Sahili and Awad \(2023\)](#).

66 The application of STGNNs to groundwater forecasting represents the cutting edge of hydro-informatics. Recent
 67 studies have demonstrated that these architectures significantly outperform traditional ML and temporal-only DL mod-
 68 els. For instance [Bai and Tahmasebi \(2023\)](#) utilized a GNN with a self-adaptive adjacency matrix to forecast GWLs,

69 proving the model could learn spatial dependencies even when physical connectivity data was incomplete. Similarly
 70 Taccari et al. (2024) applied STGNNs to the Overbetuwe area in the Netherlands, effectively integrating auxiliary vari-
 71 ables like precipitation and evaporation to handle missing data robustly. Furthermore Liang et al. (2025) proposed a
 72 GCN-LSTM framework to serve as a computationally efficient surrogate for numerical models in Quebec, highlighting
 73 the scalability of the approach. Recent work by Wu et al. (2025) further characterized multiple spatial dependencies
 74 such as hydraulic gradients and sub-basin delineations demonstrating that capturing these complex interactions is vital
 75 for regional forecasting.

76 **1.4. Contextualizing the Study: The Haouz Region, Morocco**

77 Despite these global advancements, the application of STGNNs in the specific context of North African semi-arid
 78 aquifers remains unexplored. The Haouz region in Morocco exemplifies the "data-scarce" and "high-stress" environ-
 79 ments where such advanced modeling is most needed Borzì (2025). Similar to the situation in the Rabat-Salé-Kénitra
 80 region Elmotawakkil and Enneya (2024), the Haouz aquifer faces severe depletion due to intensive irrigation and re-
 81 current drought. While recent local studies have employed dimensionality reduction and neural networks to analyze
 82 these trends Bouramtane et al. (2025), they have largely relied on temporal correlations, leaving the spatial network
 83 dynamics unmodeled.

84 The complex subsurface geometries and hydrogeological challenges observed in regions such as the Al Haouz Mej-
 85 jate basin characterized by intensive groundwater abstraction and limited natural recharge require advanced modeling
 86 approaches to reliably estimate aquifer substrate topography and predict groundwater dynamics El Mezouary et al.
 87 (2024). Recent studies have demonstrated that nonlinear machine learning techniques, including Gaussian Process
 88 Regression and deep neural network architectures, can effectively infer substrate depths by integrating sparse bore-
 89 hole information with regional geospatial datasets, achieving coefficient of determination (R^2) values exceeding 0.8
 90 when validated against independent borehole measurements El Mezouary et al. (2024). These findings highlight the
 91 growing potential of data-driven computational frameworks for characterizing key hydraulic parameters of ground-
 92 water reservoirs, parameters that are traditionally costly, time-consuming, and often uncertain when estimated solely
 93 through field-based investigations.

94 **1.5. Research Objectives and Contribution**

95 This study aims to bridge this research gap by introducing a Spatio-Temporal Graph Neural Network (STGNN)
 96 framework specifically tailored for the Haouz aquifer. By conceptualizing the monitoring network as a dynamic graph,
 97 we move beyond isolated time-series analysis to explicitly model the hydraulic connectivity between wells.

98 The main contributions of this study are as follows:

99 **1. Methodological Innovation:** We develop a unified STGNN forecasting framework that integrates hydrological

and climatic dependencies, addressing the limitations of standard LSTM models in capturing spatial correlations Wang et al. (2024); Chen et al. (2025).

- 102 **2. Regional Application:** We provide the first application of STGNN for GWL forecasting in the Haouz region, offering a high-accuracy tool for managing groundwater resources under conditions of data scarcity and climate stress Talib et al. (2024).
- 103 **3. Benchmarking and Validation:** We comprehensively benchmark the proposed approach against traditional deep learning (LSTM, GRU) baselines, demonstrating superior predictive performance and stability.
- 104 **4. Decision Support:** We highlight the potential of the graph-based approach to serve as a decision-support tool, facilitating interpretable and scalable management strategies for the Haouz basin Mohammad et al. (2023).

109 2. Methodology

110 2.1. Data Collection

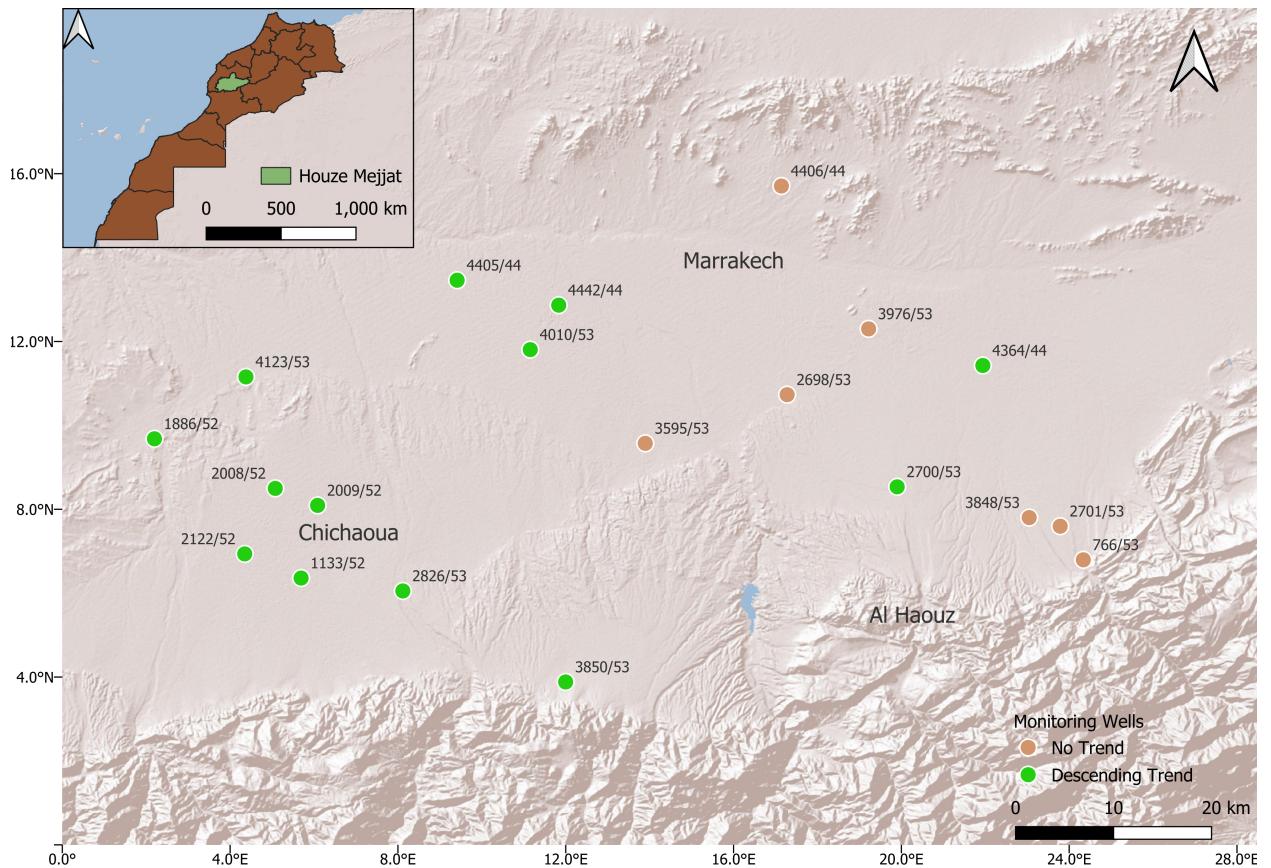


Figure 1: Location of the study area and monitoring wells, the brown dots indicates wells where the ground water level has an obvious trend, the green dots indicates wells with no trend.

111 Groundwater level data are typically obtained from monitoring well networks from the Tensift Hydraulic Basin

112 Agency (Agence des Bassins Hydraulique du Tensift, ABHT <https://abht.ma/>). Supplementary variables such as pre-
 113 precipitation, evapotranspiration, soil moisture, and land surface temperature can improve forecasting accuracy Sun et al.
 114 (2022).

115 The primary dataset used in this study consists of groundwater level (GWL) observations collected from a network
 116 of monitoring wells located within the study area. These measurements are available at a monthly temporal resolution
 117 and represent the target variable for the forecasting task. Each well provides a continuous time series of groundwater
 118 levels, allowing for the characterization of seasonal and interannual variations in groundwater storage.

119 To improve the predictive capacity of the forecasting models, complementary hydro-meteorological and land-
 120 surface variables were integrated. These explanatory variables were selected based on their relevance to groundwater
 121 recharge and depletion processes, and were obtained from a combination of remote sensing products and reanalysis
 122 datasets:

123 • **Precipitation:** Obtained from the CHIRPS dataset Funk et al. (2015), which provides quasi-global rainfall
 124 estimates at high spatial resolution.

125 • **Evapotranspiration (ET):** Extracted from the FAO WaPOR database Food and Agriculture Organization of the
 126 United Nations (FAO) (2020), offering spatially explicit data on actual evapotranspiration.

127 • **Land Surface Temperature (LST):** Retrieved from MODIS MOD11C3/MYD11C3 products Wan et al. (2021).

128 • **Normalized Difference Vegetation Index (NDVI):** Acquired from the MODIS MOD13Q1 vegetation index
 129 dataset Didan (2021).

130 • **Soil Moisture and Soil Temperature:** Taken from the ERA5 reanalysis dataset Rodell et al. (2004).

131 All variables were collected at a monthly temporal scale and spatially aligned with the locations of the monitoring
 132 wells. When necessary, gridded datasets were resampled to match the geographic coordinates of the wells, ensuring
 133 consistency across time series inputs. The resulting multi-source dataset thus combines in situ observations with
 134 satellite-derived indicators of hydrological processes.

135 To illustrate the temporal variability of the datasets, exploratory plots were generated. Figure 2 presents the monthly
 136 evolution of groundwater levels alongside selected climatic and land-surface variables for representative wells. These
 137 visualizations highlight the seasonal cycles and potential lagged relationships between groundwater response and cli-
 138 matic drivers.

139 Figure 3 illustrates the temporal evolution of groundwater levels for a selection of monitoring wells within the
 140 study area. Several important observations can be made from these plots. First, clear declining trends are visible in a
 141 number of wells, suggesting sustained groundwater depletion likely associated with long-term pumping for irrigation

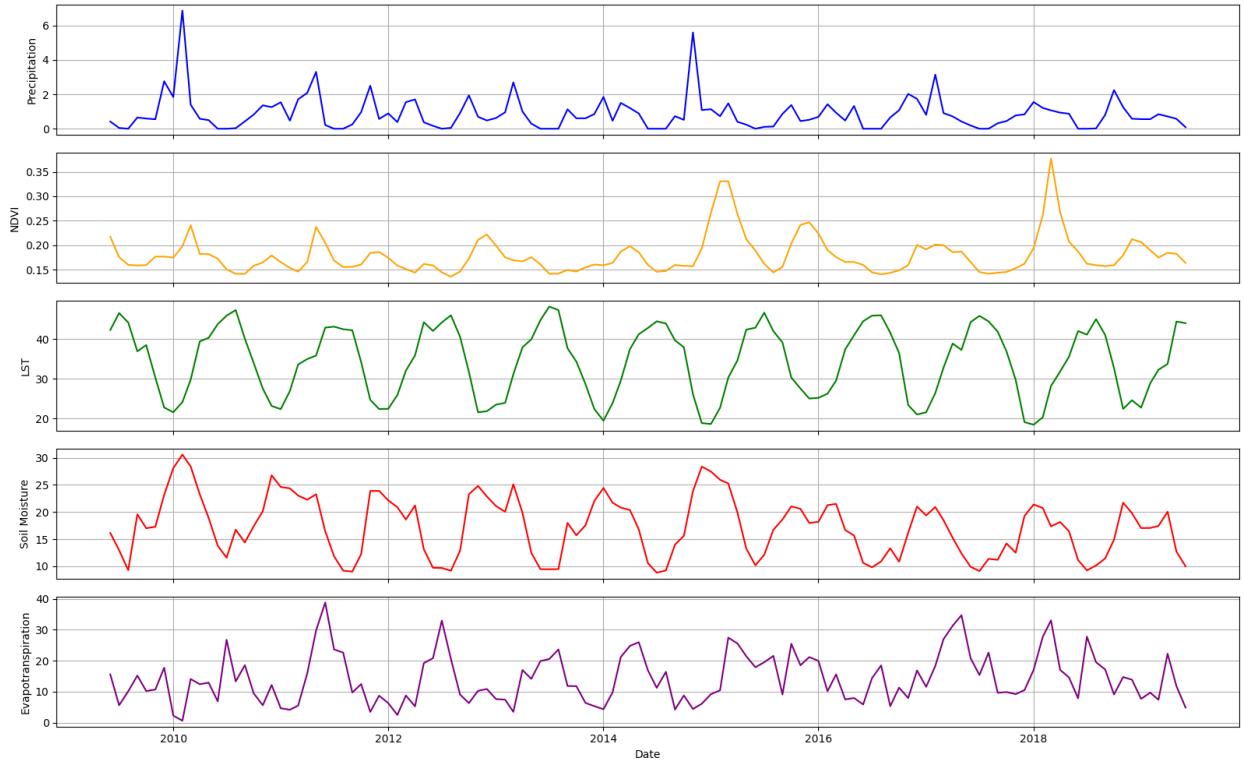


Figure 2: Example of temporal evolution of a selected explanatory variables for a representative well.

142 and domestic use. In contrast, other wells exhibit more stable or fluctuating dynamics, indicating that groundwater
143 responses are not uniform across the aquifer system.

144 The variability observed between wells highlights the combined influence of multiple controlling factors. Anthro-
145 pogenic drivers, such as groundwater abstraction rates and land-use practices, exert a strong influence in certain areas.
146 At the same time, natural processes such as snow accumulation and melt, precipitation variability, soil moisture con-
147 ditions, and evapotranspiration patterns also contribute to the temporal evolution of groundwater levels. The interplay
148 of these factors introduces significant spatial heterogeneity, making groundwater forecasting a challenging task. These
149 observations emphasize the need to integrate complementary hydro-meteorological variables alongside groundwater
150 measurements when developing predictive models.

151 **2.2. Data Preprocessing**

152 The raw datasets required preprocessing to ensure consistency across the time series and to filter out irrelevant
153 predictors.

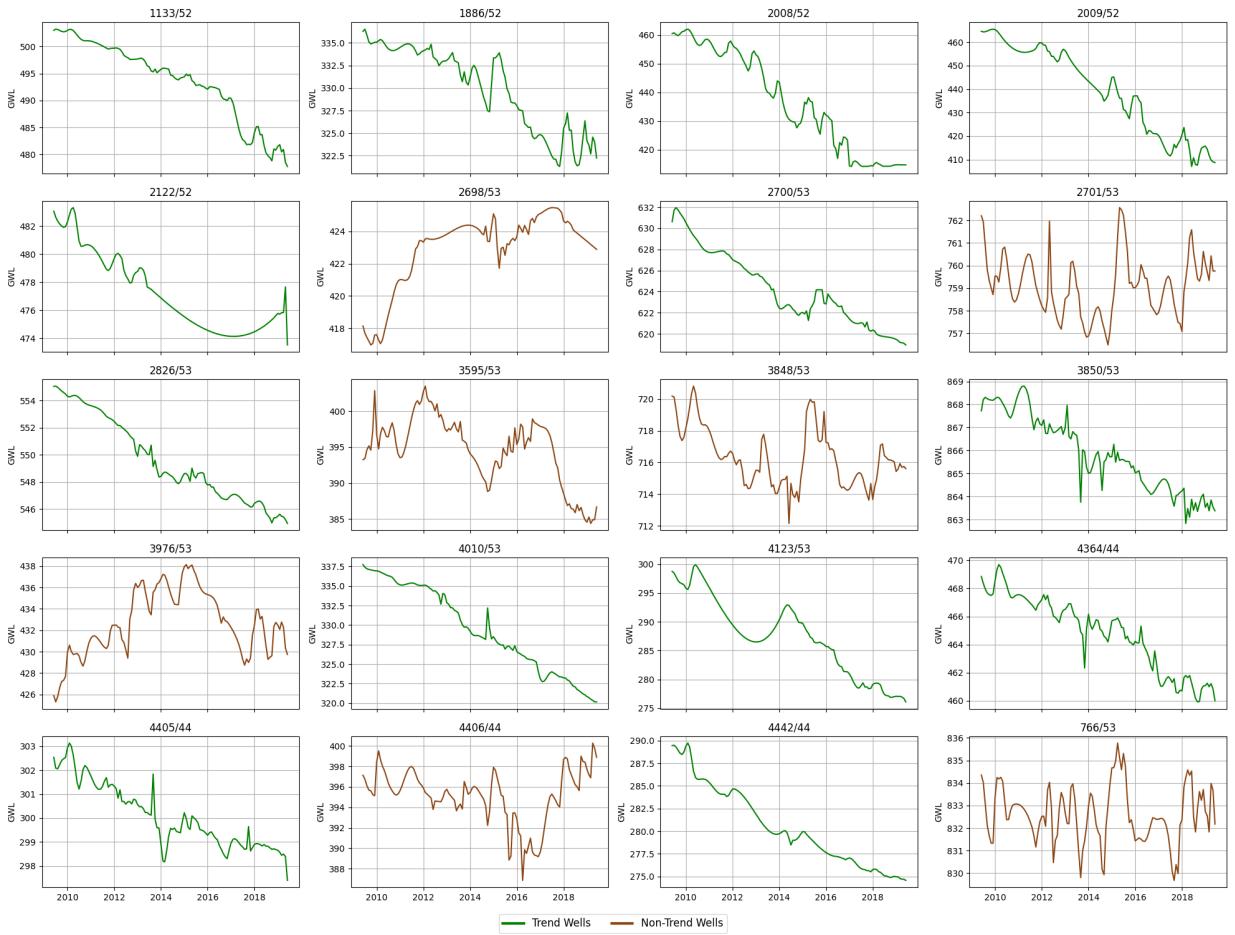


Figure 3: Temporal evolution of groundwater levels for all the wells used in the dataset. The plots highlight both declining trends in some wells and heterogeneous dynamics across the study area.

154 **Groundwater Level Standardization**

155 The groundwater level dataset was originally available at a monthly temporal resolution. However, some wells con-
 156 tained missing monthly observations, leading to discontinuities in the time series. Wells exhibiting large temporal gaps
 157 were excluded from the analysis to ensure data reliability and temporal consistency. For the remaining wells, missing
 158 monthly values were imputed using polynomial interpolation to obtain continuous time series. This interpolation was
 159 applied only to short gaps and was used solely to reconstruct missing months, without altering the original temporal
 160 resolution or introducing additional aggregation. This preprocessing step ensured uniform monthly sequences across
 161 wells while preserving the underlying groundwater level dynamics.

162 **Feature Selection**

163 Initially, a broad set of potential explanatory variables was considered. To mitigate the risk of overfitting and ensure
 164 model parsimony, a correlation-based feature selection procedure was implemented. We calculated the Pearson cor-

relation coefficient between each candidate predictor (e.g., various temperature indices, raw precipitation, vegetation metrics) and the target groundwater levels. Variables exhibiting weak correlations ($|r| < 0.1$) or high multicollinearity (redundant features) were excluded from the final input set. This process ensured that only the most relevant drivers—specifically Precipitation, Adjusted ET, LST, and NDVI—were retained for model training.

169 Variable Transformation and Evapotranspiration Partitioning

170 All predictor variables were resampled to a monthly temporal resolution and temporally aligned with the groundwater
 171 level observations to ensure consistency between remotely sensed data and in situ measurements. Standard scaling
 172 was applied where appropriate prior to model training.

173 Evapotranspiration (ET) data were obtained from the FAO WaPOR database and represent total actual evapo-
 174 transpiration, including both precipitation-driven and groundwater-supported components. To derive a variable more
 175 directly linked to groundwater availability, the precipitation-driven component of evapotranspiration was estimated
 176 and removed.

177 Monthly potential evapotranspiration (PET) was computed using the Hargreaves–Samani method [George H. Harg-
 178 reaves and Zohrab A. Samani \(1985\)](#), a temperature- and radiation-based formulation suitable for data-scarce regions.
 179 Daily minimum and maximum air temperatures were approximated from land surface temperature (LST) by applying
 180 a ± 5 °C offset. Daily PET was estimated as:

$$\text{PET}_d = 0.0023 R_a (T_{\text{mean}} + 17.8) \sqrt{T_{\text{max}} - T_{\text{min}}}, \quad (1)$$

181 where T_{mean} is the mean daily air temperature, T_{max} and T_{min} are the daily maximum and minimum temperatures,
 182 respectively, and R_a is the extraterrestrial radiation ($\text{MJ m}^{-2} \text{ day}^{-1}$), computed as a function of latitude and day of year
 183 following standard formulations. Daily PET estimates were aggregated to monthly totals.

184 The precipitation-driven component of evapotranspiration was estimated using a parametric Budyko framework
 185 [Budyko \(1974\)](#) based on the Fu formulation [Li et al. \(2022\)](#), which relates evapotranspiration to the ratio between
 186 potential evapotranspiration and precipitation. The Budyko relationship is expressed as:

$$\frac{ET_{\text{rain}}}{P} = 1 + \frac{PET}{P} - \left(1 + \left(\frac{PET}{P} \right)^{\omega} \right)^{1/\omega}, \quad (2)$$

187 where P denotes monthly precipitation, PET is monthly potential evapotranspiration, and ω is a dimensionless
 188 shape parameter controlling the partitioning behavior. A value of $\omega = 2.6$ was adopted to represent typical semi-arid

189 climatic conditions. The resulting ET_{rain} represents precipitation-driven (rainfed) evapotranspiration.
 190 Groundwater-dependent evapotranspiration was then estimated as the residual between total actual evapotranspi-
 191 ration and the precipitation-driven component:

$$ET_{\text{gw}} = ET_{\text{total}} - ET_{\text{rain}}. \quad (3)$$

192 To avoid non-physical estimates, negative values of ET_{gw} were constrained to zero. While this approach represents
 193 a simplified treatment of land–atmosphere and subsurface interactions, it provides a physically motivated proxy for
 194 groundwater-related evapotranspiration dynamics in the vicinity of the monitored wells.

195 2.3. Graph Construction

196 The groundwater monitoring network was modeled as a graph, where nodes represent wells and edges represent
 197 relationships between them, following approaches commonly used in recent spatio-temporal graph neural network
 198 (STGNN) frameworks for groundwater prediction Taccari et al. (2024). Several strategies can be adopted to define
 199 edges in such a graph. A straightforward approach is to use spatial proximity, under the assumption that wells lo-
 200 cated near each other are likely to share similar hydrogeological conditions; distance-based adjacency using Haversine
 201 distances and k -nearest neighbors (k -NN) is widely adopted in GNN-based groundwater studies Bai and Tahmasebi
 202 (2023). In this study, pairwise distances between wells were calculated using the Haversine formula, and edges were
 203 assigned following a k -NN strategy. This ensures that each well is connected to its closest neighbors, capturing local
 204 spatial dependencies; similar graph construction strategies are standard in machine learning and spatial modeling Chen
 205 et al. (2009).

206 Alternative definitions of connectivity are also possible. Hydrogeological similarity can be used when detailed
 207 subsurface data (e.g., aquifer structure or soil properties) are available, and multi-form spatial dependency models
 208 combining distance, hydrogeologic attributes, and functional similarity have been proposed Wu et al. (2025). Another
 209 option is correlation-based connectivity, where wells exhibiting similar groundwater dynamics are connected regard-
 210 less of spatial distance; such functional graphs have been applied in data-driven hydrogeological studies Wu et al.
 211 (2025). While these approaches may capture teleconnection patterns more directly, they can also introduce spurious or
 212 non-physical dependencies and risk temporal data leakage if not restricted to the training window Taccari et al. (2024).

213 For the purposes of this study, a distance-based adjacency was adopted as a robust and physically interpretable
 214 baseline. This choice balances simplicity, hydrogeologic plausibility, and reproducibility in the absence of detailed
 215 subsurface information, and aligns with existing STGNN groundwater forecasting literature Taccari et al. (2024); Bai
 216 and Tahmasebi (2023).

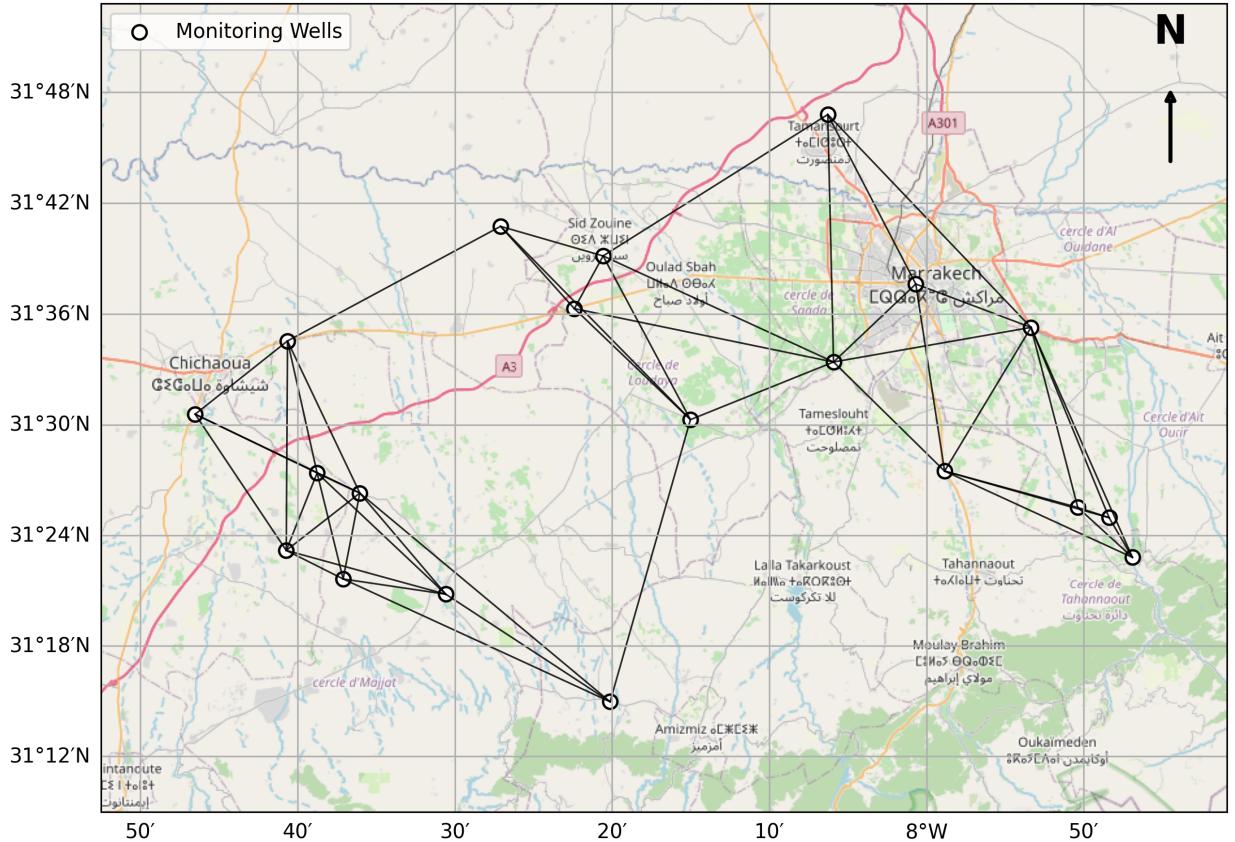


Figure 4: Graph representation of the groundwater monitoring network. Nodes correspond to wells, and edges are defined based on spatial proximity using a k -nearest neighbors (k -NN) approach with pairwise Haversine distances. This distance-based adjacency captures local spatial dependencies among wells while providing a physically interpretable and reproducible baseline for spatio-temporal graph neural network modeling.

2.4. Model Architecture

2.4.1. Graph Neural Networks (GNNs)

Graph Neural Networks (GNNs) extend deep learning methods to graph-structured data, where the relationships between nodes are as important as the attributes of the nodes themselves. The core principle of GNNs is message passing, formalized in early works on neural networks for graphs and later generalized in the Message Passing Neural Network (MPNN) framework Kipf and Welling (2017); Gilmer et al. (2017). At each layer, every node aggregates features from its neighbors and updates its own representation.

Formally, let a graph be defined as $G = (V, E)$, where V is the set of nodes (wells) and E the set of edges. For a node $v \in V$, the update rule in a generic GNN layer can be written as:

$$h_v^{(l+1)} = \sigma\left(W^{(l)} \cdot \text{AGG}\left(\{h_u^{(l)} : u \in \mathcal{N}(v)\} \cup \{h_v^{(l)}\}\right)\right), \quad (4)$$

226 where $h_v^{(l)}$ is the representation of node v at layer l , $\mathcal{N}(v)$ denotes the neighborhood of v , $W^{(l)}$ is a learnable weight
 227 matrix, and σ is a non-linear activation. The function AGG is a permutation-invariant aggregator (mean, sum, or max),
 228 as established in the MPNN formulation Gilmer et al. (2017).

229 2.4.2. Graph Convolutional Networks (GCNs)

230 The Graph Convolutional Network (GCN) is a specific type of GNN that simplifies the message passing scheme
 231 using a normalized adjacency matrix. The spectral GCN formulation used in most applications was popularized by
 232 Kipf and Welling Kipf and Welling (2017). Given feature matrix $X \in \mathbb{R}^{N \times d}$ and adjacency matrix $A \in \mathbb{R}^{N \times N}$, the
 233 forward propagation of a GCN layer is:

$$H^{(l+1)} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)}), \quad (5)$$

234 where $\tilde{A} = A + I$ is the adjacency matrix with self-loops, \tilde{D} is the diagonal degree matrix of \tilde{A} , $H^{(0)} = X$ is the input
 235 feature matrix, $W^{(l)}$ are learnable weights, and σ is a non-linear activation. This formulation ensures that each node
 236 updates its representation as a weighted average of its neighbors' features, normalized to avoid scale issues Kipf and
 237 Welling (2017).

238 2.4.3. Spatio-Temporal Graph Convolutional Networks (STGCNs)

239 Groundwater levels exhibit both temporal dependencies (fluctuations over time) and spatial dependencies (inter-
 240 actions among wells influenced by hydrogeological conditions). The STGCN integrates temporal and spatial modules
 241 into a unified framework, introduced originally for traffic forecasting by Yu et al. (2018) and later applied to environ-
 242 mental and hydrological systems Taccari et al. (2024).

243 Each spatio-temporal block consists of:

244 1. **Temporal convolution:** For an input sequence $X \in \mathbb{R}^{N \times F \times T}$, a 1D convolution is applied along the temporal
 245 axis:

$$Z = \text{Conv1D}_{\text{time}}(X), \quad (6)$$

246 as proposed in temporal convolutional architectures such as STGCN Yu et al. (2018).

247 2. **Graph convolution:** At each time step, spatial dependencies are captured using the GCN operation:

$$H_t^{(l+1)} = \sigma(\tilde{D}_t^{-\frac{1}{2}} \tilde{A}_t \tilde{D}_t^{-\frac{1}{2}} H_t^{(l)} W^{(l)}), \quad \forall t \in [1, T], \quad (7)$$

248 which follows the spectral GCN formulation Kipf and Welling (2017).

249 **3. Residual connections and normalization:** These stabilize learning and allow deeper stacking, as adopted in
 250 STGCN and related spatio-temporal architectures Yu et al. (2018).

251 Stacking multiple spatio-temporal blocks results in a deep encoder-like structure, followed by fully connected layers
 252 to predict future groundwater levels, consistent with recent applications in hydrology Taccari et al. (2024).

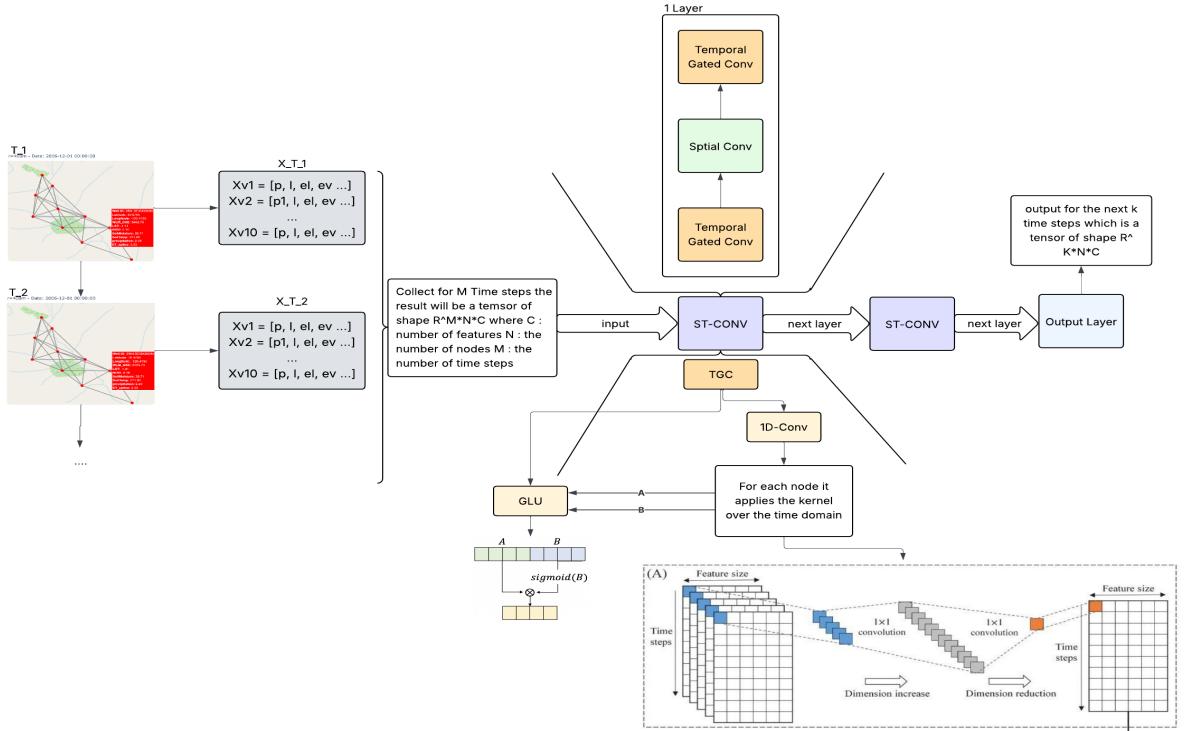


Figure 5: Architecture of the Spatio-Temporal Graph Convolutional Network (STGCN) used in this study. The model stacks multiple spatio-temporal blocks to capture spatial and temporal dependencies in groundwater data, followed by fully connected layers for forecasting future groundwater levels, consistent with recent hydrological applications Taccari et al. (2024).

253 2.5. Training and Evaluation

254 The model was trained to learn the complex spatio-temporal dependencies governing groundwater level (GWL)
 255 dynamics across the wells. Each training sample consists of a sequence of past groundwater levels and corresponding
 256 features (e.g., precipitation, evapotranspiration, temperature), and the target is the GWL value at the next time step.

257 Training Procedure

258 The dataset was divided into training, validation, and testing subsets. The model was trained for a fixed number of
 259 epochs using the Adam optimizer, with a learning rate scheduler to improve convergence stability.

260 The forward pass produces predicted groundwater levels \hat{y}_t for each node and time step, and the parameters of the
 261 model are updated by minimizing the loss function between the predicted and observed values y_t .

262

The loss function used is the **Mean Squared Error (MSE)**, defined as:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (y_{i,t} - \hat{y}_{i,t})^2 \quad (8)$$

263 where: N is the number of wells (graph nodes), T is the number of time steps, $y_{i,t}$ is the observed groundwater
264 level for well i at time t , $\hat{y}_{i,t}$ is the model's prediction.

265 This loss penalizes larger errors more heavily, encouraging the model to focus on capturing extreme variations in
266 groundwater levels.

267 The training and validation losses across epochs are shown in Figure 6. The gradual convergence of both losses
268 indicates that the model successfully generalizes without overfitting.

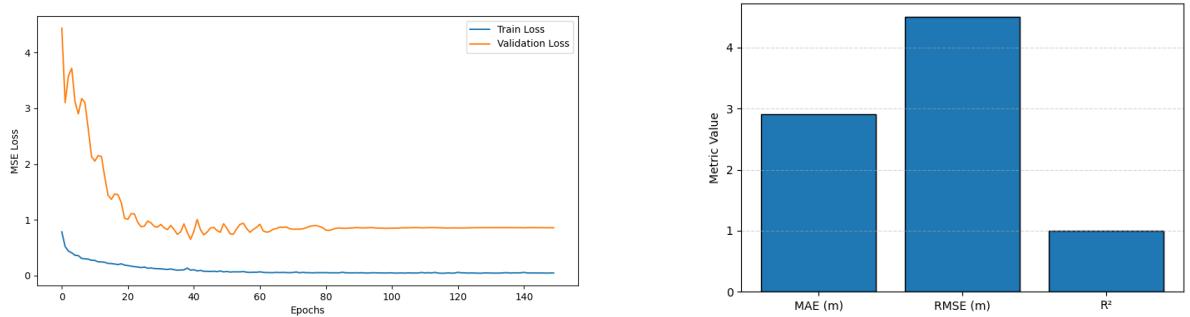


Figure 6: Overview of model performance. (a) Training and validation loss across epochs, showing gradual convergence and indicating effective generalization without overfitting. (b) Evaluation metrics on the test set, illustrating the model's accuracy in capturing groundwater level dynamics and its sensitivity to extreme variations.

269 2.5.1. Evaluation Metrics

270 The model performance was evaluated on the test set using three complementary metrics to capture different aspects
271 of prediction quality:

- 272 • **Mean Absolute Error (MAE)** measures the average magnitude of errors:

$$\text{MAE} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T |y_{i,t} - \hat{y}_{i,t}| \quad (9)$$

- 273 • **Root Mean Square Error (RMSE)** emphasizes larger errors, providing insight into extreme deviations:

$$\text{RMSE} = \sqrt{\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (y_{i,t} - \hat{y}_{i,t})^2} \quad (10)$$

- 274 • **Coefficient of Determination (R^2)** measures how well the predictions explain the variance in the observed data:

$$R^2 = 1 - \frac{\sum_{i,t} (y_{i,t} - \hat{y}_{i,t})^2}{\sum_{i,t} (y_{i,t} - \bar{y})^2} \quad (11)$$

276 where \bar{y} is the mean of the observed values.

277 Higher R^2 values (close to 1) and lower MAE/RMSE indicate better forecasting performance.

278 Overall, these results provide quantitative evidence that the STGNN model effectively captures both temporal
279 patterns and spatial dependencies in groundwater level dynamics.

280 2.6. Baseline Models

281 For the baseline models, LSTM and GRU architectures were used due to their proven effectiveness in time-series
282 prediction tasks. The comparison results show that the our STGNN generally provides more accurate predictions over
283 time. However, this improvement is not consistent across all wells, which can be attributed to the heterogeneity and
284 complexity of the data. For wells with more regular temporal patterns, the baseline models perform slightly better
285 than the STGNN. Nevertheless, our focus is on the challenging, irregular cases where the STGNN demonstrates its
286 advantage.

Model	MAE (m)	RMSE (m)	R^2
STGNN	2.9022	4.4968	0.9994
LSTM	2.9929	5.2329	0.9991
GRU	3.7084	6.2149	0.9988

Table 1

Comparison of model performance in groundwater level prediction. Metrics include mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination (R^2). The STGNN generally outperforms LSTM and GRU models, particularly for wells with irregular temporal patterns, highlighting its ability to capture complex spatio-temporal dependencies.

287 **Analysis of Aggregate Performance :** Figure 7 highlights the aggregate performance gap between the proposed
288 STGNN and the recurrent baselines (LSTM and GRU). The STGNN (blue bars) achieves the lowest error rates across
289 the board, reducing the RMSE by approximately 14% compared to the LSTM and 27% compared to the GRU. The R^2
290 metric further corroborates this, with the graph-based model maintaining a near-perfect global fit (> 0.99), whereas the
291 GRU struggles to capture the variance in the dataset. This suggests that while recurrent units are capable of modeling
292 temporal sequences, they fail to leverage the spatial information that stabilizes predictions in a regional aquifer system.

293 **Well-Specific Performance Heterogeneity :** Table 2 provides a granular breakdown of performance, revealing an
294 important distinction in model behavior. While the STGNN outperforms baselines in the majority of wells, there are
295 specific instances (e.g., Well 766/53 and Well 3595/53) where the LSTM achieves marginally higher R^2 scores. These

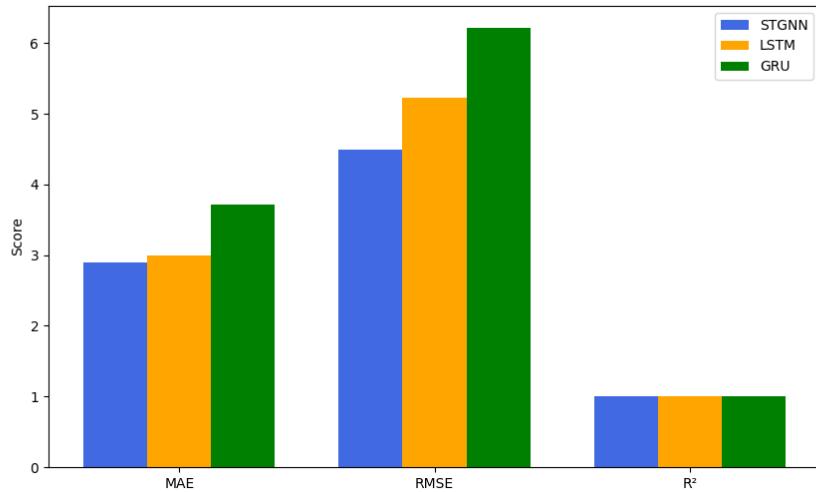


Figure 7: Metrics comparison of the two baseline models (LSTM and GRU) with the STGNN model.

296 exceptions generally occur in wells exhibiting highly localized behavior or those located at the aquifer's periphery,
 297 where spatial neighbor connections may be less informative. However, in critical wells with complex dynamics and
 298 high variability (such as Well 2009/52 and Well 4123/53), the STGNN demonstrates superior robustness, significantly
 299 mitigating the extreme negative R^2 values observed in the baseline models. This indicates that the graph structure
 300 effectively acts as a regularizer, preventing the massive prediction errors that isolated time-series models are prone to
 301 during volatile periods.

Well	STGNN			LSTM			GRU		
	MAE	RMSE	R ²	MAE	RMSE	R ²	MAE	RMSE	R ²
766/53	0.8665	1.1158	0.4055	0.7510	0.9139	0.6011	0.9550	1.0998	0.4223
4442/44	1.8359	2.0590	-21.3732	0.8750	1.0113	-4.3980	1.6207	1.8058	-16.2102
4406/44	1.5358	1.9704	-0.2581	0.7454	1.2240	0.5145	0.7290	1.2146	0.5220
4405/44	0.4545	0.6071	-1.9435	0.2223	0.3789	-0.1466	0.4186	0.5395	-1.3246
4364/44	2.2890	2.4086	-16.6905	2.1384	2.2168	-13.9858	2.6081	2.6380	-20.2208
4010/53	2.9373	3.1758	-5.1524	3.1319	3.3357	-5.7877	4.0897	4.2436	-9.9855
3976/53	2.5702	2.7680	-1.7261	0.7270	0.9496	0.6791	0.9775	1.1862	0.4993
3850/53	0.8655	0.9802	-4.7492	0.7026	0.8378	-3.1998	0.9689	1.1178	-6.4766
3848/53	0.9689	1.4465	-1.4433	0.5632	0.6950	0.4360	0.4095	0.5556	0.6396
3595/53	3.5369	4.0585	-0.3420	2.7094	3.0312	0.2514	3.9927	4.2385	-0.4637
2826/53	1.2117	1.6496	-7.4196	1.1774	1.2174	-3.5857	1.5095	1.5915	-6.8373
2122/52	0.9623	1.3955	-1.8567	0.7014	1.0147	-0.5105	1.1232	1.3621	-1.7217
2701/53	0.9972	1.4371	-0.6562	0.5161	0.6364	0.6752	0.5061	0.6429	0.6686
2700/53	1.4206	1.7508	-6.7185	0.8646	0.9750	-1.3935	1.2645	1.4619	-4.3811
2698/53	0.7377	0.8827	-0.0226	0.2892	0.3606	0.8294	0.3709	0.4216	0.7667
2009/52	12.5478	13.2876	-9.1205	16.6293	17.1444	-15.8483	18.9665	19.3803	-20.5292
2008/52	7.7916	8.5157	-569.1771	9.7309	9.8571	-762.9545	13.3363	13.4321	-1417.5862
1886/52	2.8002	3.1453	-2.4510	4.5317	4.8088	-7.0668	5.0457	5.2322	-8.5496
4123/53	3.8988	4.1425	-15.4744	4.0058	4.1366	-15.4270	5.5304	5.5917	-29.0167
1133/52	7.8165	8.0567	-16.9498	8.8452	9.0294	-21.5454	9.7452	9.9290	-26.2616

Table 2

Per-well performance comparison for STGNN, LSTM, and GRU models. Best values per metric (MAE, RMSE, R²) are highlighted in bold.

3. Results and Discussion

The proposed Spatio-Temporal Graph Convolutional Network (STGNN) demonstrated strong predictive capabilities for groundwater level (GWL) forecasting across the study area. Figure 8 shows a visual comparison between the observed and predicted groundwater levels during both training and testing periods. The close alignment between the two indicates that the model effectively captures the long-term trends and short-term fluctuations of GWL dynamics.

Temporal Dynamics and Generalization : Figure 8 presents the hydrographs for three representative monitoring wells, delineating the training, validation, and testing phases. A visual inspection reveals the model's ability to track disparate hydrological regimes:

• **Trend Capture:** In Well 4442/44 (Top Panel), the STGCN accurately reproduces the seasonal cyclicity while adhering to the gradual recovery trend observed in the validation phase.

• **Response to Abrupt Changes:** Well 4406/44 (Middle Panel) exhibits a sharp drawdown event during the testing phase (right of the purple vertical line). Unlike standard regression models that often suffer from lag or smoothing, the STGCN anticipates this drop, attributed to its ability to aggregate information from neighboring wells that may have experienced the stress earlier.

• **Handling Noise:** Well 4405/44 (Bottom Panel) represents a high-frequency fluctuation scenario. The model

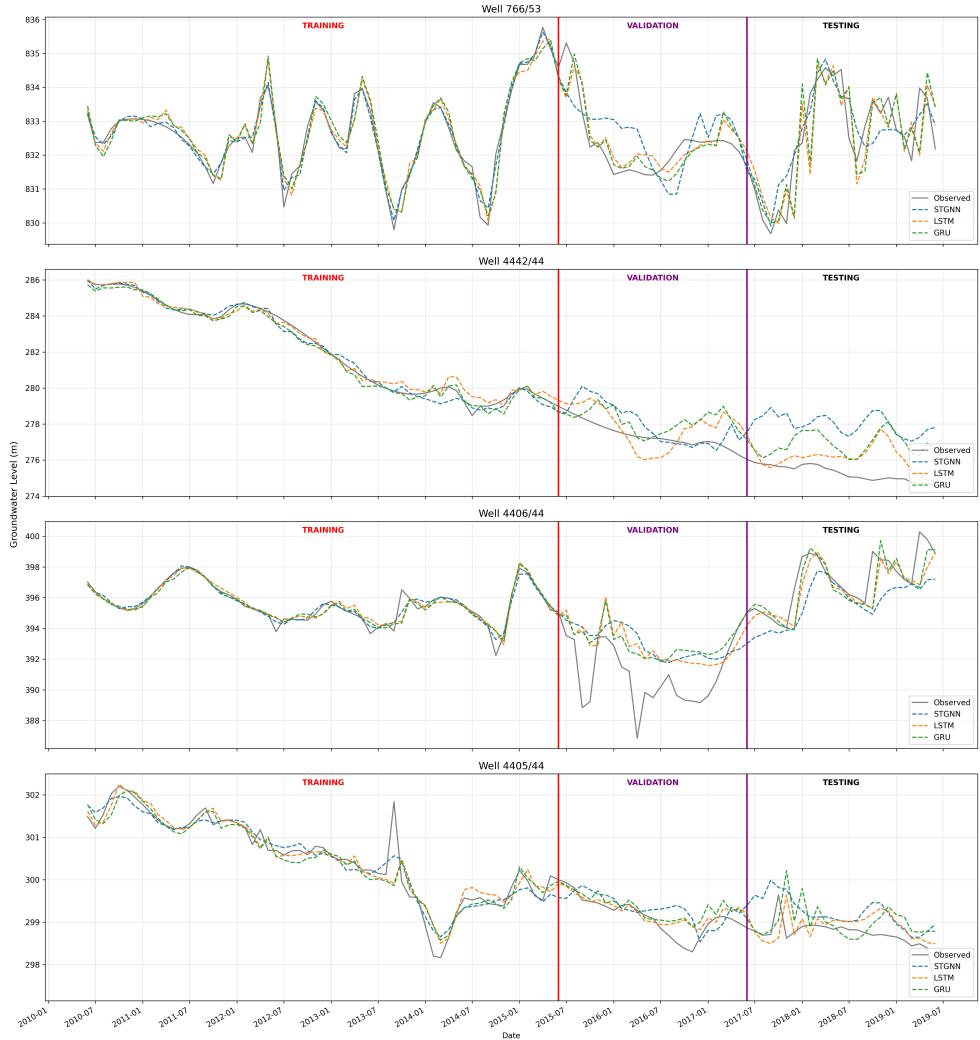
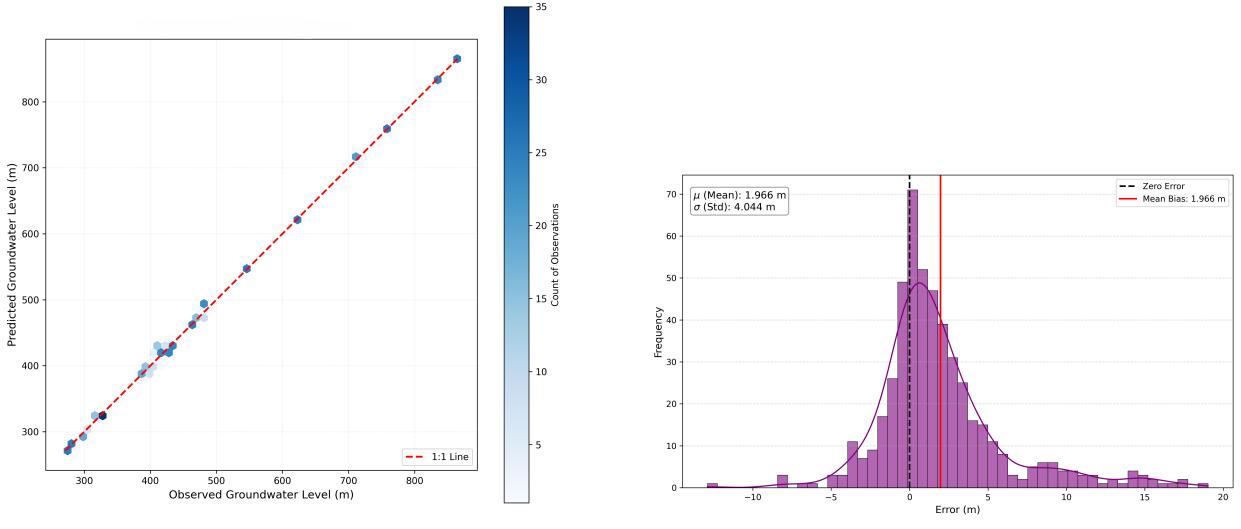


Figure 8: Comparison of observed and predicted groundwater levels for the training and testing periods. The close alignment illustrates the STGNN's ability to accurately capture both long-term trends and short-term fluctuations in groundwater dynamics.

317 predictions (orange dashed line) tightly hug the observed data (grey line), indicating that the model successfully
 318 disentangles signal from noise without overfitting to the training data.

319 The close alignment in the testing phase confirms that the model does not merely memorize historical sequences but
 320 generalizes well to unseen climatic and hydrological conditions.

321 **Error Distribution and Goodness-of-Fit :** To further assess the reliability of the forecasts, we analyzed the statistical
 322 properties of the residuals. Figure 9a displays the scatter plot of observed versus predicted groundwater levels. The
 323 data points cluster tightly around the 1:1 diagonal line (red dashed), indicating a lack of systematic bias across the
 324 range of groundwater depths. There is no significant deviation at the tails, suggesting the model performs equally well



(a) Observed vs predicted groundwater levels

(b) STGNN: Distribution of Residuals.

Figure 9: Evaluation of STGNN prediction performance. (a) Scatter plot of observed versus predicted groundwater levels, demonstrating strong agreement and model generalization to unseen conditions. (b) Distribution of residuals across all wells, highlighting the model's error characteristics and overall prediction reliability.

325 for both shallow and deep water tables.

326 Figure 9b illustrates the distribution of prediction residuals. The error histogram approximates a Gaussian (normal)
 327 distribution centered near zero ($\mu \approx 1.96$ m), with a controlled standard deviation. The symmetry of the bell curve
 328 indicates that the model is not biased toward overestimation or underestimation. The slight positive mean bias suggests
 329 a very marginal tendency to under-predict drawdown in extreme cases, likely due to the inherent smoothing effect of
 330 the graph convolution operator, yet the majority of errors fall within an acceptable range for regional management
 331 planning.

332 3.1. Model Performance and Interpretation

333 The evaluation metrics (MAE, RMSE, and R^2) revealed that the STGNN achieved high accuracy and stable gener-
 334 alization across the different wells. The model outperformed baseline approaches such as classical LSTMs, particularly
 335 during periods of high variability (e.g., seasonal transitions). This improvement is consistent with findings from related
 336 studies, where STGNN-based methods reduced forecasting errors by approximately 15–20% compared to sequence-
 337 only models.

338 **Drivers of Groundwater Dynamics :** To understand the physical drivers influencing the STGNN's predictions, we
 339 conducted a permutation feature importance analysis (Figure 10a). The results identify evapotranspiration (ad-
 340 justed via the Budyko framework) as the single most influential predictor. This aligns with the hydrogeological reality
 341 of the Haouz region, where groundwater depletion is primarily driven by irrigation demands rather than natural fluc-

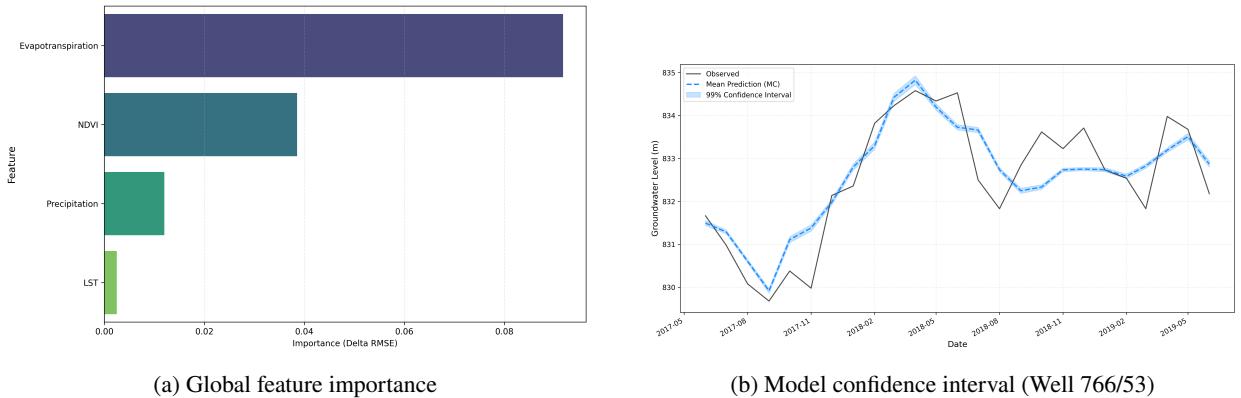


Figure 10: Model interpretability and reliability. (a) Global feature importance across all wells, showing which inputs most strongly influence STGNN predictions. (b) Prediction confidence intervals for well 766/53, illustrating the model's uncertainty and robustness during periods of high groundwater variability.

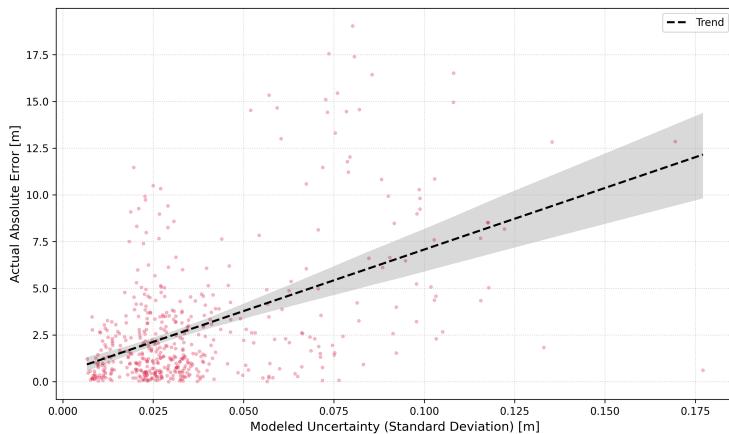


Figure 11: Relationship between predicted uncertainty and actual prediction error. The scatter plot shows that higher modeled uncertainty (standard deviation of predictions) corresponds to larger absolute errors, indicating that the STGNN's uncertainty estimates are well-calibrated and can serve as a reliable proxy for risk in groundwater management decisions.

342 situations.

343 Interestingly, Precipitation exhibits a lower importance score relative to NDVI. This suggests that the aquifer's
 344 response to rainfall is highly non-linear and lagged (due to the infiltration process through the vadose zone), whereas
 345 variables like LST and NDVI serve as immediate proxies for evaporative stress and agricultural water withdrawal.

346 **Uncertainty Calibration :** Figure 11 validates the reliability of the model's uncertainty estimates. The scatter plot
 347 illustrates the relationship between the modeled uncertainty (standard deviation of the prediction distribution) and the
 348 Actual Absolute Error. The distinct positive trend, highlighted by the regression line, demonstrates that the model is
 349 well-calibrated: as the predicted uncertainty increases, the likelihood of a larger error also increases.

350 For decision-makers, this linear relationship is valuable. It implies that the uncertainty metric output by the STGNN
 351 can be used as a trustworthy proxy for risk. If the model predicts a groundwater level with a high uncertainty variance,

352 water managers can prioritize those specific wells for manual verification or additional sensor deployment, thereby
 353 optimizing monitoring resources.

354 3.2. Effect of Spatial Relationships

355 Using spatial proximity as the basis for the graph structure proved effective in capturing coherent spatial patterns
 356 in groundwater dynamics. However, wells influenced by anthropogenic factors (e.g., pumping, irrigation) or localized
 357 hydrogeological differences exhibited deviations from purely distance-based similarity. Future work could benefit
 358 from incorporating more physically meaningful relationships, such as hydraulic conductivity, lithology, or correlation
 359 in groundwater dynamics, to refine the graph construction.

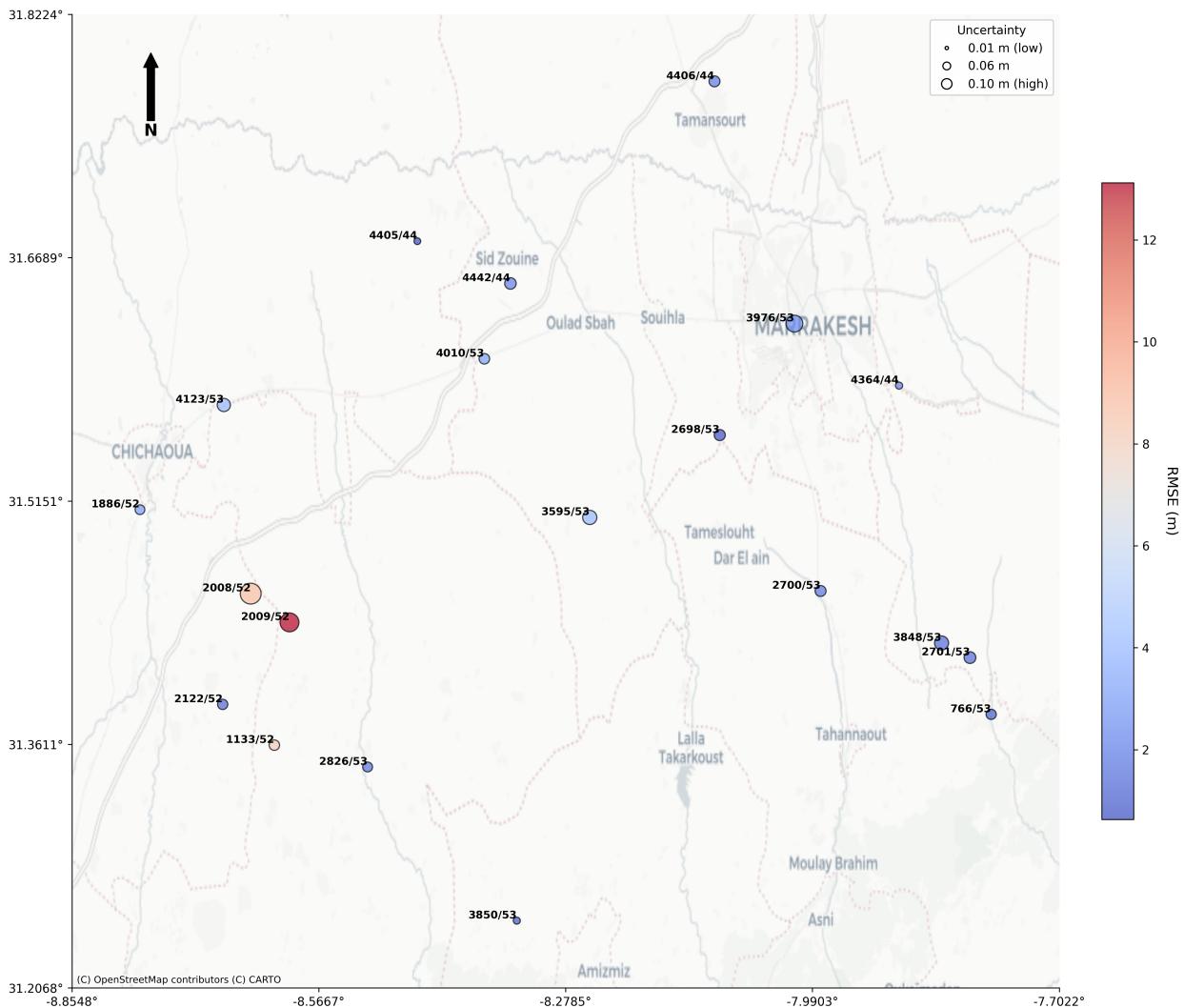


Figure 12: Spatial diagnostics of STGCN performance across the monitoring network. Node color represents prediction accuracy, while node size indicates model confidence. This visualization highlights wells where anthropogenic activities or local hydrogeological heterogeneity lead to deviations from the distance-based graph assumptions, providing insights for targeted model refinement and monitoring prioritization.

360 **Spatial Distribution of Error and Uncertainty :** To assess the reliability of the STGNN across the heterogeneous
 361 landscape of the Haouz aquifer, we visualized the spatial distribution of model performance in Figure 12. In this
 362 diagnostics map, the color scale represents the prediction error (RMSE), while the size of the markers is proportional
 363 to the model's predictive uncertainty (confidence interval width).

364 The map reveals a dominant prevalence of blue markers, indicating low RMSE values across the majority of the
 365 monitoring network. This confirms the model's capability to generalize well over spatially disjoint locations. How-
 366 ever, distinct clusters of larger, red-hued nodes are observable, particularly in zones known for intensive agricultural
 367 activity. The correlation between marker size and color intensity is notable: wells where the model exhibits high
 368 prediction error often coincide with high uncertainty estimates. This "self-awareness" of the model is a critical safety
 369 feature; it indicates that the STGNN can flag its own limitations in areas with complex, non-stationary dynamics driven
 370 by unmeasured anthropogenic factors (e.g., unreported illegal pumping), rather than making confident but incorrect
 371 predictions.

372 3.3. Temporal Dynamics and Seasonal Variability

373 The STGNN successfully reproduced seasonal oscillations associated with precipitation and evapotranspiration
 374 cycles, as well as gradual long-term declines in wells affected by persistent overextraction. However, during abrupt
 375 hydrological events (e.g., extreme drought or heavy rainfall), the model exhibited slightly higher prediction errors.
 376 This suggests that while STGCNs can capture regular spatio-temporal dependencies effectively, additional external
 377 drivers (e.g., surface water–groundwater interactions or land use changes) could further enhance forecasting accuracy.

378 4. Conclusion

379 This study presented a novel application of Spatio-Temporal Graph Neural Networks (STGNNs) for regional
 380 groundwater level forecasting in the semi-arid Haouz Aquifer, Morocco. By conceptualizing the monitoring network as
 381 a dynamic graph, we successfully integrated spatial dependencies—representing hydraulic connectivity—with tempo-
 382 ral hydrological sequences. This approach addressed a critical methodological gap in standard deep learning models
 383 (such as LSTM and GRU), which treat monitoring wells as isolated entities and often fail to capture the systemic
 384 response of the aquifer to anthropogenic and climatic stressors.

385 The empirical results demonstrate that the proposed STGNN framework significantly outperforms baseline tem-
 386 poral models, reducing the Root Mean Square Error (RMSE) by approximately 14% compared to the LSTM and 27%
 387 compared to the GRU. The model exhibited remarkable robustness, maintaining high predictive accuracy ($R^2 > 0.99$)
 388 even in wells characterized by irregular fluctuations and sharp seasonal drawdowns. Furthermore, the inclusion of
 389 uncertainty quantification revealed a strong correlation between predicted uncertainty and actual error, establishing

390 the model as a reliable decision-support tool for identifying high-risk areas in data-scarce regions. Ultimately, this
 391 research confirms that explicit modeling of spatial interactions is a prerequisite for accurate groundwater forecasting
 392 in complex, over-exploited aquifer systems.

393 5. Future Work

394 While the proposed STGNN framework offers a robust data-driven solution, several avenues remain for further
 395 enhancement, particularly regarding the physical consistency of predictions and long-term scenario planning.

396 5.1. Integration of Physics-Informed Neural Networks (PINNs)

397 A primary direction for future research is to bridge the gap between data-driven efficiency and physical realism
 398 by incorporating Physics-Informed Neural Networks (PINNs). While STGNNs effectively learn statistical spatio-
 399 temporal patterns, they do not explicitly enforce fundamental hydrogeological laws, such as the conservation of mass
 400 or Darcy's Law. Consequently, purely data-driven models may occasionally produce physically inconsistent predictions
 401 in unmonitored locations.

402 Future iterations of this work aim to embed these governing partial differential equations (PDEs) directly into
 403 the network's loss function. A hybrid architecture combining STGNNs with PINNs would ensure that predictions
 404 remain hydrogeologically plausible, effectively acting as a regularization mechanism. This approach is expected to
 405 significantly improve performance in data-scarce zones and enhance the model's ability to generalize during extreme
 406 climatic events where historical training data is insufficient.

407 5.2. Dynamic and Functional Graph Construction

408 The current study utilized distance-based adjacency to define the graph structure. However, hydraulic connectivity
 409 is not solely a function of geometric distance but is also influenced by subsurface heterogeneity, geological faults, and
 410 transmissivity. Future work could explore learning the graph structure dynamically from data (adaptive adjacency
 411 matrices) or constructing graphs based on functional similarity and geological surveys to better represent the physical
 412 flow paths within the aquifer.

413 5.3. Climate Change and Scenario Analysis

414 Finally, to support long-term sustainable management, the framework should be extended to simulate future ground-
 415 water trajectories under various climate change scenarios and socioeconomic abstraction pathways. Integrating these
 416 long-term projections will transform the model from a short-term forecasting tool into a comprehensive platform for
 417 strategic water resource planning in the Tensift basin.

418 **6. Acknowledgments**

419 **Code availability section**

420 Name of the code/library

421 Contact: and phone number

422 Hardware requirements: ...

423 Program language: ...

424 Software required: ...

425 Program size: ...

426 The source codes are available for downloading at the link: [https://github.com/](https://github.com/)427 **References**

- 428 Anderson, M.P., Hunt, R.J., Woessner, W.W., 2015. Applied groundwater modeling: simulation of flow and advective transport. Second edition
429 ed., Academic Press, London, UK.
- 430 Bai, T., Tahmasebi, P., 2023. Graph neural network for groundwater level forecasting. Journal of Hydrology 616, 128792. URL: <https://www.sciencedirect.com/science/article/pii/S0022169422013622>, doi:10.1016/j.jhydrol.2022.128792.
- 432 Borzì, I., 2025. Modeling Groundwater Resources in Data-Scarce Regions for Sustainable Management: Methodologies and Limits. Hydrology
433 12, 11. URL: <https://www.mdpi.com/2306-5338/12/1/11>, doi:10.3390/hydrology12010011. publisher: Multidisciplinary Digital
434 Publishing Institute.
- 435 Bouramtane, T., Mohsine, I., Karmouda, N., Leblanc, M., Estève, Y., Kacimi, I., Hilali, M., Mdhaffar, S., Tweed, S., Tahir, M., Kassou, N., El Bilali,
436 A., Chafki, O., 2025. Dimensionality reduction for groundwater forecasting under drought and intensive irrigation with neural networks. Journal
437 of Hydrology: Regional Studies 60, 102477. URL: <https://www.sciencedirect.com/science/article/pii/S2214581825003027>,
438 doi:10.1016/j.ejrh.2025.102477.
- 439 Budyko, M.I., 1974. Climate and life [by] M. I. Budyko. Edited by David H. Miller. International geophysics series ; v. 18., Academic Press, New
440 York.
- 441 Chang, Y.W., Sun, W., Kow, P.Y., Lee, M.H., Chang, L.C., Chang, F.J., 2025. Advanced groundwater level forecasting with hybrid deep learning
442 model: Tackling water challenges in Taiwan's largest alluvial fan. Journal of Hydrology 655, 132887. URL: <https://www.sciencedirect.com/science/article/pii/S0022169425002252>, doi:10.1016/j.jhydrol.2025.132887.
- 444 Chen, J., Fang, H.r., Saad, Y., 2009. Fast Approximate kNN Graph Construction for High Dimensional Data via Recursive Lanczos Bisection. J.
445 Mach. Learn. Res. 10, 1989–2012.
- 446 Chen, L., Zhang, D., Xu, J., Zhou, Z., Jin, J., Luan, J., Wulamu, A., 2025. Enhancing the accuracy of groundwater level prediction at dif-
447 ferent scales using spatio-temporal graph convolutional model. Earth Science Informatics 18, 250. URL: <https://doi.org/10.1007/s12145-025-01741-z>, doi:10.1007/s12145-025-01741-z.
- 449 Didan, K., 2021. Modis/terra vegetation indices 16-day l3 global 1km sin grid v061. NASA Land Processes Distributed Active Archive Center.
450 URL: <https://doi.org/10.5067/MODIS/MOD13A2.061>, doi:10.5067/MODIS/MOD13A2.061. data set. Accessed: 2025-12-11.
- 451 El Mezouary, L., Hadri, A., Kharrou, M.H., Fakir, Y., Elfarchouni, A., Bouchaou, L., Chehbouni, A., 2024. Contribution to advancing aquifer
452 geometric mapping using machine learning and deep learning techniques: a case study of the AL Haouz-Mejjate aquifer, Marrakech, Morocco.
453 Applied Water Science 14, 102. URL: <https://doi.org/10.1007/s13201-024-02162-x>, doi:10.1007/s13201-024-02162-x.

- 454 Elmotawakkil, A., Enneya, N., 2024. Advanced machine learning for predicting groundwater decline and drought in the Rabat–Salé–Kénitra region,
455 Morocco. *Journal of Hydroinformatics* 26, 2980–3007. URL: <https://doi.org/10.2166/hydro.2024.328>, doi:10.2166/hydro.2024.328.
- 456
- 457 Famiglietti, J.S., 2014. The global groundwater crisis. *Nature Climate Change* 4, 945–948. URL: <https://www.nature.com/articles/nclimate2425>, doi:10.1038/nclimate2425. publisher: Nature Publishing Group.
- 458
- 459 Food and Agriculture Organization of the United Nations (FAO), 2020. WaPOR V2 Database Methodology. Technical Report. FAO. Rome. URL:
460 <https://www.fao.org/3/ca9894en/CA9894EN.pdf>.
- 461 Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The
462 climate hazards infrared precipitation with stations (chirps): A new environmental record for monitoring extremes. *Scientific Data* 2, 150066.
463 doi:10.1038/sdata.2015.66.
- 464 George H. Hargreaves, Zohrab A. Samani, 1985. Reference Crop Evapotranspiration from Temperature. *Applied Engineering in Agriculture*
465 1, 96–99. URL: <http://elibrary.asabe.org/abstract.asp?JID=3&AID=26773&CID=aeaj1985&v=1&i=2&T=1>, doi:10.13031/
466 2013.26773.
- 467 Gilmer, J., Schoenholz, S.S., Riley, P.F., Vinyals, O., Dahl, G.E., 2017. Neural message passing for Quantum chemistry, in: Proceedings of the 34th
468 International Conference on Machine Learning - Volume 70, JMLR.org, Sydney, NSW, Australia. pp. 1263–1272.
- 469 Kipf, T.N., Welling, M., 2017. Semi-Supervised Classification with Graph Convolutional Networks. URL: <https://openreview.net/forum?id=SJU4ayYg1>.
- 470
- 471 Li, S., Du, T., Gippel, C.J., 2022. A Modified Fu (1981) Equation with a Time-varying Parameter that Improves Estimates of Inter-annual Variability
472 in Catchment Water Balance. *Water Resources Management* 36, 1645–1659. URL: <https://doi.org/10.1007/s11269-021-03057-1>,
473 doi:10.1007/s11269-021-03057-1.
- 474 Li, Y., Yu, D., Liu, Z., Zhang, M., Gong, X., Zhao, L., 2023. Graph Neural Network for spatiotemporal data: methods and applications URL:
475 <https://arxiv.org/abs/2306.00012>, doi:10.48550/ARXIV.2306.00012. publisher: arXiv Version Number: 1.
- 476 Liang, X.X., Gloaguen, E., Claprood, M., Paradis, D., Lauzon, D., 2025. Graph Neural Network Framework for Spatiotemporal Groundwater
477 Level Forecasting. *Mathematical Geosciences* 57, 1071–1093. URL: <https://doi.org/10.1007/s11004-025-10194-5>, doi:10.1007/
478 s11004-025-10194-5.
- 479 Lin, H., Gharehbaghi, A., Zhang, Q., Band, S.S., Pai, H.T., Chau, K.W., Mosavi, A., 2022. Time series-based groundwater level forecasting
480 using gated recurrent unit deep neural networks. *Engineering Applications of Computational Fluid Mechanics* 16, 1655–1672. URL:
481 <https://doi.org/10.1080/19942060.2022.2104928>, doi:10.1080/19942060.2022.2104928. publisher: Taylor & Francis _eprint:
482 <https://doi.org/10.1080/19942060.2022.2104928>.
- 483 Mohammad, S., Islam, A., Shit, P.K., Towfiqul Islam, A.R.M., Alam, E., 2023. Groundwater level dynamics in a subtropical fan delta region and its
484 future prediction using machine learning tools: Sustainable groundwater restoration. *Journal of Hydrology: Regional Studies* 47, 101385. URL:
485 <https://www.sciencedirect.com/science/article/pii/S2214581823000721>, doi:10.1016/j.ejrh.2023.101385.
- 486 Mukherjee, A., Coomar, P., Sarkar, S., Johannesson, K.H., Fryar, A.E., Schreiber, M.E., Ahmed, K.M., Alam, M.A., Bhattacharya, P., Bundschuh,
487 J., Burgess, W., Chakraborty, M., Coyte, R., Farooqi, A., Guo, H., Ijumulana, J., Jeelani, G., Mondal, D., Nordstrom, D.K., Podgorski, J.,
488 Polya, D.A., Scanlon, B.R., Shamsudduha, M., Tapia, J., Vengosh, A., 2024. Arsenic and other geogenic contaminants in global groundwater.
489 *Nature Reviews Earth & Environment* 5, 312–328. URL: <https://www.nature.com/articles/s43017-024-00519-z>, doi:10.1038/
490 s43017-024-00519-z. publisher: Nature Publishing Group.
- 491 Nayak, P.C., Satyaji Rao, Y., Sudheer, K., 2006. Groundwater Level Forecasting in A Shallow Aquifer Using Artificial Neural Network Approach.

- 492 Water Resources Management 20, 77–90. doi:[10.1007/s11269-006-4007-z](https://doi.org/10.1007/s11269-006-4007-z).
- 493 Refsgaard, J.C., 1997. Parameterisation, calibration and validation of distributed hydrological models. Journal of Hydrology 198, 69–97. URL:
<https://www.sciencedirect.com/science/article/pii/S002216949603329X>, doi:[10.1016/S0022-1694\(96\)03329-X](https://doi.org/10.1016/S0022-1694(96)03329-X).
- 495 Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin,
496 J.K., Walker, J.P., Lohmann, D., Toll, D., 2004. The global land data assimilation system. Bulletin of the American Meteorological Society 85,
497 381–394.
- 498 Sahili, Z.A., Awad, M., 2023. Spatio-Temporal Graph Neural Networks: A Survey. URL: <http://arxiv.org/abs/2301.10569>, doi:[10.48550/arXiv.2301.10569](https://arxiv.org/abs/2301.10569). arXiv:2301.10569 [cs].
- 500 Sahoo, S., Jha, M.K., 2013. Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a com-
501 parative assessment. Hydrogeology Journal 21, 1865–1887. URL: <https://doi.org/10.1007/s10040-013-1029-5>, doi:[10.1007/s10040-013-1029-5](https://doi.org/10.1007/s10040-013-1029-5).
- 503 Scanlon, B.R., Fakhreddine, S., Rateb, A., de Graaf, I., Famiglietti, J., Gleeson, T., Grafton, R.Q., Jobbagy, E., Kebede, S., Kulu, S.R., Konikow,
504 L.F., Long, D., Mekonnen, M., Schmied, H.M., Mukherjee, A., MacDonald, A., Reedy, R.C., Shamsuddoha, M., Simmons, C.T., Sun, A.,
505 Taylor, R.G., Villholth, K.G., Vörösmarty, C.J., Zheng, C., 2023. Global water resources and the role of groundwater in a resilient water
506 future. Nature Reviews Earth & Environment 4, 87–101. URL: <https://www.nature.com/articles/s43017-022-00378-6>, doi:[10.1038/s43017-022-00378-6](https://doi.org/10.1038/s43017-022-00378-6). publisher: Nature Publishing Group.
- 508 Scarselli, F., Gori, M., Ah Chung Tsoi, Hagenbuchner, M., Monfardini, G., 2009. The Graph Neural Network Model. IEEE Transactions on Neural
509 Networks 20, 61–80. URL: <http://ieeexplore.ieee.org/document/4700287/>, doi:[10.1109/TNN.2008.2005605](https://doi.org/10.1109/TNN.2008.2005605).
- 510 Sophocleous, M., 2002. Interactions between groundwater and surface water: the state of the science. Hydrogeology Journal 10, 52–67. URL:
<https://doi.org/10.1007/s10040-001-0170-8>, doi:[10.1007/s10040-001-0170-8](https://doi.org/10.1007/s10040-001-0170-8).
- 512 Sun, J., Hu, L., Li, D., Sun, K., Yang, Z., 2022. Data-driven models for accurate groundwater level prediction and their practical significance
513 in groundwater management. Journal of Hydrology 608, 127630. URL: <https://www.sciencedirect.com/science/article/pii/S0022169422002050>, doi:[10.1016/j.jhydrol.2022.127630](https://doi.org/10.1016/j.jhydrol.2022.127630).
- 515 Taccari, M.L., Wang, H., Nuttall, J., Chen, X., Jimack, P.K., 2024. Spatial-temporal graph neural networks for groundwater data. Scientific Reports
516 14, 24564. URL: <https://www.nature.com/articles/s41598-024-75385-2>, doi:[10.1038/s41598-024-75385-2](https://doi.org/10.1038/s41598-024-75385-2). publisher: Nature
517 Publishing Group.
- 518 Talib, A., Desai, A.R., Huang, J., 2024. Spatial and temporal forecasting of groundwater anomalies in complex aquifer undergoing climate and land
519 use change. Journal of Hydrology 639, 131525. URL: <https://www.sciencedirect.com/science/article/pii/S0022169424009211>,
520 doi:[10.1016/j.jhydrol.2024.131525](https://doi.org/10.1016/j.jhydrol.2024.131525).
- 521 Taylor, R.G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J.S., Edmunds, M., Konikow,
522 L., Green, T.R., Chen, J., Taniguchi, M., Bierkens, M.F.P., MacDonald, A., Fan, Y., Maxwell, R.M., Yechieli, Y., Gurdak, J.J., Allen, D.M.,
523 Shamsuddoha, M., Hiscock, K., Yeh, P.J.F., Holman, I., Treidel, H., 2013. Ground water and climate change. Nature Climate Change 3, 322–
524 329. URL: <https://www.nature.com/articles/nclimate1744>, doi:[10.1038/nclimate1744](https://doi.org/10.1038/nclimate1744). publisher: Nature Publishing Group.
- 525 Wada, Y., van Beek, L.P.H., van Kempen, C.M., Reckman, J.W.T.M., Vasak, S., Bierkens, M.F.P., 2010. Global depletion of ground-
526 water resources. Geophysical Research Letters 37. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1029/2010GL044571>,
527 doi:[10.1029/2010GL044571](https://doi.org/10.1029/2010GL044571). _eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2010GL044571>.
- 528 Wan, Z., Hook, S., Hulley, G., 2021. Modis/terra land surface temperature/emissivity 8-day l3 global 1km sin grid v061. NASA Land Processes
529 Distributed Active Archive Center. URL: <https://doi.org/10.5067/MODIS/MOD11A2.061>, doi:[10.5067/MODIS/MOD11A2.061](https://doi.org/10.5067/MODIS/MOD11A2.061).

Short title

- 530 Wang, L., Jiang, Z., Song, L., Yu, X., Yuan, S., Zhang, B., 2024. A groundwater level spatiotemporal prediction model based on graph convolutional
531 networks with a long short-term memory. *Journal of Hydroinformatics* 26, 2962–2979. URL: <https://doi.org/10.2166/hydro.2024.226>,
532 doi:10.2166/hydro.2024.226.
- 533 Wu, Y., Mei, G., Shao, K., Xu, N., Peng, J., 2025. Forecasting Groundwater Level by Characterizing Multiple Spatial Dependencies
534 of Environmental Factors Using Graph-Based Deep Learning. *Journal of Geophysical Research: Machine Learning and Computation*
535 2, e2024JH000520. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1029/2024JH000520>, doi:10.1029/2024JH000520.
536 _eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024JH000520>.
- 537 Yu, B., Yin, H., Zhu, Z., 2018. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting, in: Proceed-
538 ings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, pp. 3634–3640. URL: [http://arxiv.org/abs/1709.
539 04875](http://arxiv.org/abs/1709.04875), doi:10.24963/ijcai.2018/505. arXiv:1709.04875 [cs].