Journal of Hydrology: Regional Studies 58 (2025) 102221

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh

Remote sensing assessment of dam impact on arid basins in Southern Saudi Arabia: A machine learning and space-for-time approach

Raid Almalki [a,b,*], Mehdi Khaki [c], Patricia M. Saco [c,d], Jose F. Rodriguez [c]

[a] *School of Environmental and Life Science, University of Newcastle, Callaghan, NSW 2308, Australia*
[b] *Department of Geography, Umm Al-Qura University, Makkah 21955, Saudi Arabia*
[c] *School of Engineering, University of Newcastle, Callaghan, NSW 2308, Australia*
[d] *School of Civil and Environmental Engineering University of Technology Sydney, Ultimo, NSW 2007, Australia*

ARTICLE INFO

Keywords:
Remote sensing
Space-for-time
Machine learning
Dam impacts
Water
Arid regions

ABSTRACT

Study region: This study focuses on four basins in southern Saudi Arabia: Hali, Baish, Yiba, and Reem. These regions are characterized by arid conditions and are significantly impacted by dam construction.
Study focus: The research investigates the environmental impacts of dam construction using a space-for-time substitution approach, remote sensing, and machine learning techniques. A key focus is analyzing non-linear environmental impacts, particularly in data-limited, arid regions where traditional methodologies fall short. The study introduces a novel framework that combines space-for-time substitution and Dynamic Time Warping (DTW) to assess temporal and spatiotemporal changes in key environmental factors such as NDVI, soil salinity, groundwater, and runoff.
New hydrological insights: The results reveal significant changes post-dam construction. In the Yiba-Hali basins, DTW values increased across several parameters: NDVI (0.08–0.25), soil salinity (0.09–0.25), and runoff (0.45–0.90), indicating reduced similarity between pre- and post-dam conditions. In the Reem-Baish basins, the Baish dam caused notable increases in DTW values for NDVI (0.16–0.31), soil salinity (0.15–0.30), groundwater (0.52–1.19), and runoff (0.53–1.33), with the most significant changes observed in groundwater and runoff. Additionally, regression models showed a decrease in predictive accuracy from 2010 to 2020, as evidenced by lower R^2 values for NDVI (0.82–0.37), soil salinity (0.77–0.38), groundwater (0.98–0.34), and soil moisture (0.96–0.24).

1. Introduction

Understanding the environmental impacts of dams is critical for sustainable water resource management and informed policy-making (Beck et al., 2012). Insights gained from such evaluations provide valuable information on the trade-offs associated with dam construction and guide future projects to balance development and environmental preservation. In dry climates, where rainfall is scarce and evaporation rates are high, water becomes a precious and limited resource (Ji et al., 2006), and basins are crucial for

* Corresponding author at: School of Environmental and Life Science, University of Newcastle, Callaghan, NSW 2308, Australia.
E-mail address: C3348452@uon.edu.au (R. Almalki).

https://doi.org/10.1016/j.ejrh.2025.102221
Received 20 October 2024; Received in revised form 27 January 2025; Accepted 30 January 2025
Available online 3 February 2025

sustaining ecosystems by storing and distributing water (Zhou et al., 2020). However, the construction of dams in these basins can significantly impact the downstream environment (Almalki et al., 2023). Changes to natural water flow can disrupt sediment transport, erosion, and groundwater recharge, which in turn affects the broader ecosystem (Almalki et al., 2023). Furthermore, managing water flow can influence downstream water availability, potentially causing water scarcity for ecosystems that depend on these resources (Richter et al., 2010). Therefore, monitoring basin water changes is critical for informed management, particularly in understanding how dams influence water resources and their subsequent impacts on the sustainability of downstream environments in arid regions.

Monitoring water changes in dam-affected areas is challenging, especially where data is scarce and ground-based measurements are limited. This problem becomes larger due to the logistic difficulties and high costs of conducting extensive on-site observations, further perpetuating the data voids (Almalki et al., 2022). The absence of such data for climate, hydrology, and soil characteristics in arid and semiarid areas, including Saudi Arabia, was highlighted by other work (Al-Saeedi, 2022; Alshehri and Mohamed, 2023a, 2023b). Satellite measurements bridge data gaps by providing wide spatial coverage and near-continuous observations, enabling researchers to study areas that are otherwise inaccessible. Remote sensing technology has long been recognized as an essential tool for monitoring vegetation, soil conditions, and hydrological changes, especially in regions with limited ground-based data (Barnes et al., 2003; Thakur et al., 2017). By identifying spatiotemporal patterns, this technology enhances the detection of long-term trends and impacts, making it invaluable for environmental research and management in data-scarce regions (Almalki et al., 2023; Parr et al., 2003).

Nevertheless, detecting changes directly attributable to dams are complex due to many influencing factors, notably fluctuations in water flow compounded by dam operations (Hu et al., 2018; You et al., 2022). These fluctuations can significantly alter downstream water levels, sediment transport dynamics, and vegetation cover, making it very difficult to isolate the environmental changes caused by dams (Petts and Gurnell, 2005). This challenge can be mitigated by using unaffected areas with similar conditions as proxies to benchmark the specific impacts of dam construction (Strobl and Strobl, 2011). Space-for-time substitution is widely used in environmental studies to infer temporal changes by examining spatial patterns (Damgaard, 2019; Pickett, 1989). Through this approach, researchers study different spatial locations with similar hydroclimate conditions that represent various stages of the process of interest (Peuquet, 2001). Space-for-time substitution can assess similarities between datasets, allowing extrapolation across contexts and is useful in environmental succession studies (Burke et al., 1997; Costa et al., 2021; Miyanishi and Johnson, 2021), climate change impacts (Horrocks et al., 2020; Rastetter, 1996), disturbance recovery (Seidl et al., 2018), and land use changes (Rastetter, 1996; Siehoff et al., 2011). This approach offers significant advantages in time efficiency and feasibility. Careful site selection, replication, and the use of control sites are essential for robust findings. Combining space-for-time substitution with long-term monitoring, experiments, and modelling can enhance the validity and comprehensiveness of the conclusions (Yue et al., 2023).

To evaluate environmental variables in arid basins with and without dams, methods included statistical analyses, time series similarity measurements, machine learning, change point detection, clustering, and classification (Almalki et al., 2023; Chowdhury and Al-Zahrani, 2014; El Ghazali et al., 2021; Yaseen et al., 2020). These methods are commonly used in space-for-time analyses to assess environmental changes and patterns. However, DTW effectively measures similarity between time series that vary in timing, aligning out-of-phase sequences (Folgado et al., 2018). For example, Romani et al. (2010) applied a DTW-based method to analyze similarity in sugarcane regions using climate and remote sensing time series. This feature is particularly useful for comparing environmental data across regions or periods, accommodating shifts and distortions. In addition, Dürrenmatt et al. (2013) used DTW on water quality data from upstream and downstream sensors to assess temporal changes. This highlights DTW and ED's effectiveness with remote sensing data in comparing environmental changes between dammed and undammed basins, aiding sustainable water management in arid regions.

Machine learning methods handle large datasets, uncover non-linear patterns, and offer robust predictions and similarity assessments. Methods like DTW, multivariate regression, and machine learning are effective for space-for-time analysis in complex, nonlinear environmental data. For example, Ugbaje et al. (2024) employed machine learning for space-time soil organic carbon stock mapping and its local drivers. Their study highlights the capability of machine learning to predict and quantify soil organic carbon stock and its local drivers. Equere et al. (2021) utilized GIS and artificial neural networks to predict the distribution of Urban Heat Island in Illinois, USA. Their study significantly improved prediction accuracy when topographic factors were incorporated into the model. Several studies have used machine learning, including regression techniques, to analyze groundwater, land use, land cover, and meteorological variables (Azari et al., 2022; Hussein et al., 2020; Rahmati et al., 2019; Stojanova et al., 2010; Wu et al., 2018). These studies aimed to improve predictions and mapping of environmental variables, offering valuable insights for management and planning.

The main goal of this study is to assess environmental changes across multiple basins in arid regions through the space-for-time approach. Using remote sensing data combined with a space-for-time substitution, the research examines the impacts of dam construction on vegetation cover, soil conditions, and hydrological changes. By identifying spatial patterns as proxies for temporal changes, the study provides insights into long-term environmental processes influenced by dams. This assessment enhances our understanding of how dams alter natural systems in arid areas, guiding future water resource management and conservation efforts.

2. Study areas

This study selected four basins—Hali, Baish, Yiba, and Reem—in southern Saudi Arabia. Hali an Baish are adjacent to Yiba and Reem, respectively exhibiting nearly identical environmental conditions. This similarity makes them ideal for a comparative analysis of the impacts of dam construction on environmental variables (see Fig. 1). Two basins have dams, while the others serve as proxies. In this study, Yiba is a proxy for Hali, and Reem is a proxy for Baish due to their proximity and similar environmental conditions. Yiba and

Hali are adjacent basins with comparable climate and soil characteristics, and similarly, Reem and Baish are close to each other and share similar attributes (Şen et al., 2013; Shahin and Shahin, 2007). Using neighbouring basin pairs as proxies ensures relevant environmental conditions, enabling a more accurate comparison of dam construction impacts. The study spans from 2003 to 2020, highlighting long-term trends and dam construction impacts on environmental variables. These basins are the main basins in the southern part with similar geographical characteristics and a combined runoff of around 50–100 × 10^6 m^3 y^{-1} (Shahin and Shahin, 2007).

Fig. 1 shows the Hali Basin near the Red Sea coast in southwestern Saudi Arabia, spanning 160 km and channelling rainfall from nearby hills to the sea. (Alarifi et al., 2022). The basin receives an average of 100 mm of rainfall annually, with headwaters receiving between 300 and 600 mm. A gravity dam, constructed in 2009, is located within the basin and has a catchment area of 5222 km^2 (Hasanean and Almazroui, 2015). The dam has a storage capacity of 249,860,000 m^3 and is surrounded by sparse vegetation, including Prosopis juliflora, halophytes, and mangroves (Avicennia marina) along the Red Sea coast (Alarifi et al., 2022). Yiba basin, located close to Hali, receives an average annual rainfall of 128 mm (Ejaz et al., 2024), with its headwaters experiencing higher rainfall ranging from 300 to 600 mm annually (Arebu et al., 2024). This basin supports vegetation similar to Hali and Baish basins, including drought-resistant species and sparse acacia trees.

The Baish dam, operational since 2009, is in southwestern Saudi Arabia. Standing 106 m tall, it has a catchment area of 4741.07 km^2 (Sallam et al., 2018). This area receives an average of 229 mm of annual rainfall. Maximum temperatures reach 41°C between June and August, while minimum temperatures average 18°C from December to February (Radwan et al., 2017). The vegetation in the Baish downstream area includes grasses, desert shrubs, and scattered acacia trees (Khawfany et al., 2009). In addition, Reem basin is located in southwestern Saudi Arabia and is an important watercourse that plays a crucial role in the ecology and agriculture of the region. The basin receives about 180 mm of annual rainfall and higher precipitation in the headwater areas, enhancing water availability (ElKashouty et al., 2022). This rainfall supports a variety of ecosystems and is vital for maintaining local biodiversity, as well as sustaining agricultural activities that depend on the intermittent water flow. The diverse topography of the Reem basin, including the Red Sea coastal plain (Tihamah), the hills, and the scarp of the Hijaz mountains, further influences the distribution and retention of water, making Reem basin a key resource in its region (Abd El Shafy and Mostafa, 2021).

3. Data and method

To analyze the downstream environmental conditions of each dam, this study examines seven variables: NDVI, soil salinity, soil moisture, groundwater, runoff water, temperature, precipitation and, organic carbon sediment (OCS). These variables help compare proxy basins without dams and impacted basins with dams. These variables were selected for their relevance to vegetation, soil, climate, and water resources, and their common use in assessing the environmental impacts of dams in arid and semi-arid regions (Al-Robai et al., 2018; Al-Sodany et al., 2015; Almalki et al., 2023; Chikodzi et al., 2013; Jafari and Hasheminasab, 2017; Sallam et al., 2018). The study uses monthly satellite and reanalyzed data (see Table 1 in Section 3.1.2.), applying normalization processes to address differences in spatial resolution and acquisition times. All datasets were resampled to a uniform spatial resolution of 1 km to ensure consistent and accurate comparisons, balancing detail with data availability. This standardization allows for uniformity in both temporal and spatial dimensions. The variables were assessed across Hali, Baish, Yiba, and Reem for two periods, 2003–2009 (pre--dam) and 2010–2020 (post-dam), to analyze environmental changes and the impacts of dam construction.

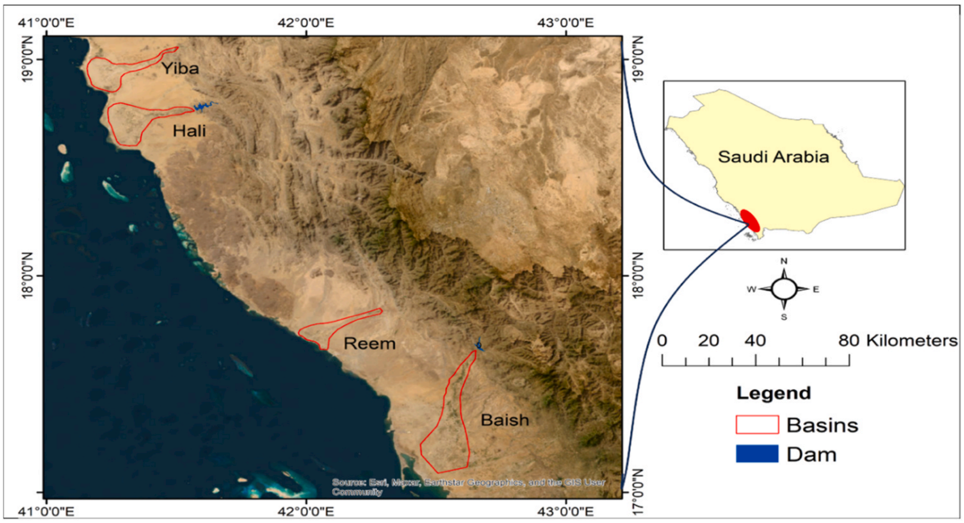


Fig. 1. This study considers four areas in the southwestern side of Saudi Arabia. Yiba and Reem were used as proxy basins for Hali and Baish.

Table 1
Summary of study data (2003–2020) access April-May 2023.

Variable	Type	Resolution	Source
NDVI	MODIS	250 m	https://search.earthdata.nasa.gov
Groundwater	GRACE	0.25 degree	https://giovanni.gsfc.nasa.gov/
Precipitation	GPM	0.1 degree	https://giovanni.gsfc.nasa.gov/
Temperature	AIRS	1 degree	https://giovanni.gsfc.nasa.gov/
Soil moisture	GLDAS	0.25 degree	https://giovanni.gsfc.nasa.gov
Total evapotranspiration	FLDAS	0.1 degree	https://giovanni.gsfc.nasa.gov
OCS	Merra−2	0.5 degree	https://giovanni.gsfc.nasa.gov

3.1. Database

3.1.1. Monthly satellite data

The Moderate Resolution Imaging Spectroradiometer (MODIS) vegetation indices (MOD13Q1) version 6.1 level-3 data were obtained every 16 days with a spatial resolution of 250 m. Additionally, the MODIS Surface Reflectance products (MOD09A1, bands 1–7) were utilized, provided in an 8-day gridded Level-3 format. The Near-Infrared (NIR) and visible red (RED) bands from the MOD09A1 dataset were used to calculate the Normalized Difference Salinity Index (NDSI) The NIR band captured near-infrared reflectance, and the RED band measures visible red reflectance, allowing the NDSI to quantify surface water salinity variations. (Meng et al., 2016). The NDSI is highly effective in arid and semi-arid regions, where fluctuations in soil and water salinity occur due to variations in water availability, evaporation rates, and land use practices (Allbed and Kumar, 2013). The index leverages differences in reflectance properties, with salinized surfaces typically showing higher reflectance in the RED band compared to the NIR band. This allows the NDSI to identify areas with elevated salinity levels, making it ideal for monitoring salinity changes, especially in areas downstream of dams where altered water flows can increase salinity. Previous studies, such as Allbed et al. (2014), have demonstrated its effectiveness in tracking surface salinity variations, highlighting its relevance for assessing the environmental impacts of dams in arid regions. The index is calculated using the following formula:

$$NDSI = \frac{RED - NIR}{RED + NIR} \tag{1}$$

The Global Precipitation Measurement (GPM) provides monthly precipitation data at a spatial resolution of 0.1 degrees. Air temperature data for the daytime were obtained from the Atmospheric Infrared Sounder (AIRS) with a spatial resolution of 1 degree. The Gravity Recovery and Climate Experiment (GRACE) mission provides data on groundwater conditions by measuring changes in Earth's gravity. The GRACEDADM_CLSM025GL_7D v3.0 dataset tracks variations in terrestrial water storage, including groundwater, which is crucial for assessing climate and human impacts on water resources. In addition, OCS data were obtained from the MERRA-2 dataset, which provides global, gridded information on organic carbon deposition in sediments. These data were used to assess the role of sediment quality on soil moisture, vegetation dynamics, and hydroclimatic interactions in the study basins. Normalization was applied using the Nearest resampling method in ArcMap to ensure consistency across datasets with varying spatial resolutions and acquisition times. The data were resampled to a uniform spatial resolution of 1 km, allowing for accurate comparisons and analyses. The 1 km resolution was selected to strike a balance between detail and data availability. The remote sensing datasets used—MODIS, GMP, AIRS, GLDAS, and GRACE—are validated for environmental monitoring in arid regions. MODIS NDVI tracks vegetation, soil salinity, GMP precipitation data aligns with ground measurements, and AIRS is reliable for atmospheric data (Almalki et al., 2025; Alshehri and Mohamed, 2023b). GLDAS and GRACE data effectively monitor soil moisture and groundwater, supported by (Almalki et al., 2025). These validations confirm their applicability in data-limited areas.

3.1.2. Reanalyzed data

Soil moisture data were obtained from the NASA Global Land Data Assimilation System (GLDAS) due to insufficient ground-based data in the study areas. GLDAS combines satellite and ground-based data to generate comprehensive soil moisture simulations (Tavakol et al., 2021). The model provides accurate soil moisture data, filling gaps in ground-based observations, so monthly GLDAS data were used. To estimate downstream runoff, we use a water budget approach that accounts for precipitation, evapotranspiration (including soil and canopy evaporation and plant transpiration), runoff (both surface and subsurface), and water storage (in vegetation canopies, lakes, wetlands, rivers, soil moisture, and groundwater) (Sheffield et al., 2009). Despite dam control, small streams continue to contribute to downstream flow. The water balance is given by:

$$\Delta S = P - E - R \tag{2}$$

The balance of precipitation (P), evapotranspiration (E), and runoff (R) determines the change in water storage (i.e., ΔS) at the Earth's surface. Eq. (2) is used to calculate the runoff for all the case studies. The precipitation and water storage data were acquired by downloading monthly satellite data (i.e., GPM and GRACE, respectively), while evapotranspiration was obtained from reanalysed data (i.e., FLDAS).

3.2. Method

The methodology described in Fig. 2 outlines the process for calculating DTW similarities and applying regression-based machine learning to predict environmental changes in basins with and without dams before and after dam construction. The study utilizes a combination of statistical analyses, time series similarity measurements (ED and DTW), and machine learning techniques, such as multivariate regression, to assess environmental changes. These methods are chosen for their ability to handle large datasets, uncover non-linear patterns, and align environmental data with varying temporal shifts. The process begins with satellite imagery and reanalyzed data acquisition, followed by pre-processing steps like extraction, unit conversion, and spatial resolution adjustments. Detrending is applied to remove trends and focus on fluctuations in environmental variables. DTW aligns time series to assess similarities, comparing monthly environmental variables from 2003 to 2009 and 2010–2020 between proxy basins (Yiba, Reem) and impacted basins (Hali, Baish). DTW is well-suited for comparing time series data that exhibit phase shifts, making it ideal for analyzing remote sensing data from regions with different temporal dynamics. The data is then split for DTW and machine learning analysis, with a 4-fold cross-validation method ensuring robust evaluation. DTW calculations are performed for each fold, generating distance values and plots to assess environmental similarities between the basins. Regression-based models predict environmental variables in dam-impacted basins and compare them to natural changes in basins without dams to assess the impacts of dam construction. The findings are presented separately for each case study to demonstrate the model's effectiveness and the extent of environmental change.

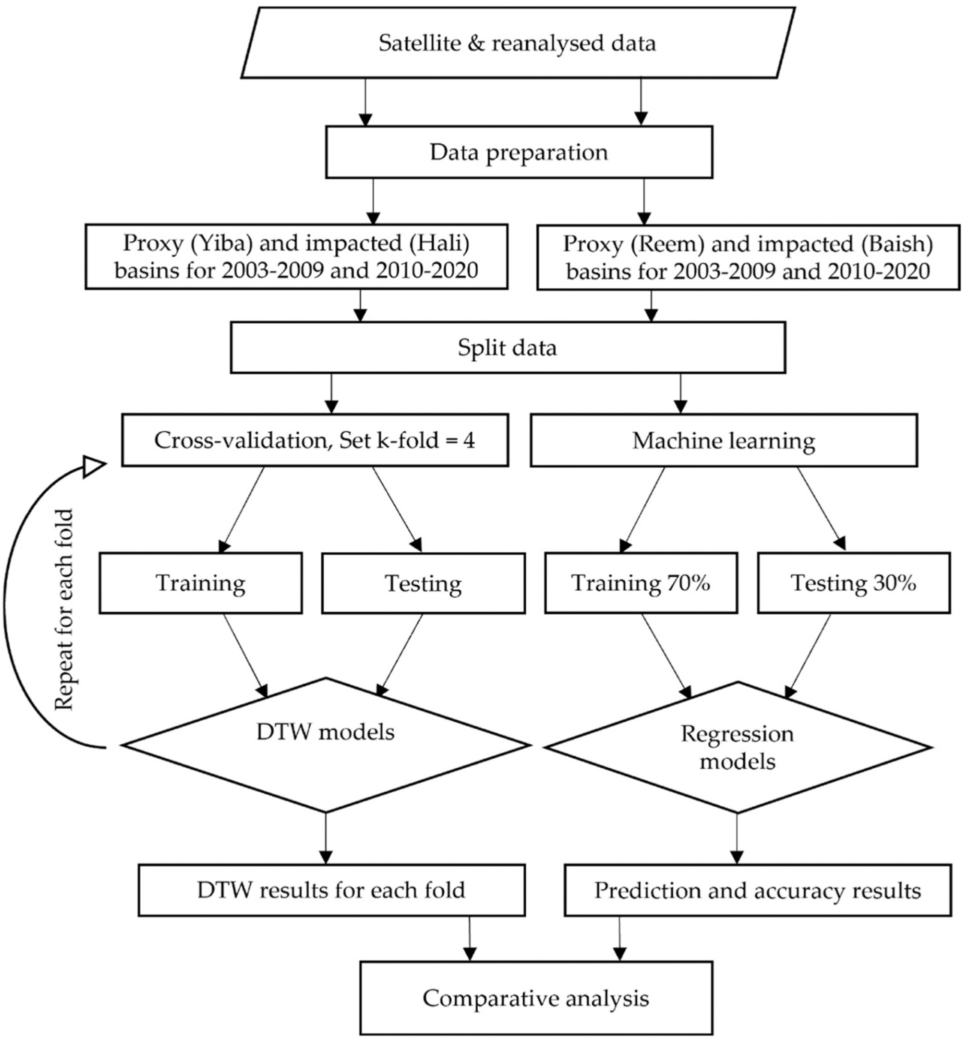


Fig. 2. Flowchart illustrating the methodology for assessing the impact of dam construction on environmental variables using remote sensing and machine learning techniques.

3.2.1. Correlation calculation

Before using DTW and regression-based machine learning to compare the temporal patterns of environmental variables between periods with and without a dam, this study assessed the correlation coefficient between variables in proxy basins and those impacted by dam construction. The correlation coefficient measures the strength and direction of a linear relationship between two datasets. A correlation coefficient for each environmental variable from 2003 to 2009 and 2010–2020 is calculated to assess the consistency of changes caused by the dam. Afterwards, the correlation reduction between the two periods was calculated by subtracting the first period's correlation (e.g., 2003–2009) from the second period's correlation (e.g., 2010–2020).

3.2.2. DTW distance calculation

The DTW distance calculation (Felty, 2024) was performed with k-fold cross-validation in MATLAB to assess the similarity between time series data from dammed and undammed basins, both before and after dam construction. This analysis was focused on key environmental variables, including NDVI, groundwater, soil salinity, soil moisture, runoff water, precipitation, and temperature. DTW optimally aligns the sequences by stretching or compressing the time axis, enabling more accurate comparisons than traditional methods such as Euclidean Distance (ED) (Liu et al., 2024), see Fig. 3. One critical aspect of dam construction is its effect on the timing of water flow. Dams often cause delays in water release, which can alter the natural hydrological cycles and subsequently affect various environmental variables such as soil moisture, vegetation growth, and groundwater levels. DTW is particularly well-suited for this study, which employs space-for-time analysis, as it can align time series data that are out of phase due to delays. Additionally, DTW assesses similarities between datasets, making it a powerful tool for comparing environmental variables across different regions or periods (Oregi et al., 2017).

DTW measures the distance between environmental variables in proxy and impacted basins for each period, reflecting the similarity in their temporal patterns. A lower DTW distance near zero indicates high similarity, whereas a higher distance suggests significant temporal pattern differences (Jeong et al., 2011). This approach was used to quantify the similarity between X (e.g., proxy basin) and Y (e.g., impacted basin) for each environmental variable during the pre-dam and post-dam periods, as shown in Eq. 3:

$$DTW(x,\quad y) = \sqrt{\sum\nolimits_{i=1}^{N}\sum\nolimits_{i=1}^{M} d(i,j)^2} \tag{3}$$

where N and M are the lengths of X and Y respectively, and $d(i,j)$ is the distance between the i-th element of X and the j-th element of Y. A further step was taken with 4-fold cross-validation to enhance DTW performance. By dividing the data into four subsets and iteratively validating the model, the robustness and accuracy of DTW in measuring similarity were improved.

Cross-validation is a key technique for assessing model performance by dividing the dataset into k subsets, known as k-fold cross-validation (A. Ramezan et al., 2019). In this method, the data was split into k equal parts, and for each fold, the model was trained on k-1 folds and tested on the remaining fold, as shown in Fig. 4. This process repeats k times, using each fold as the test set once, which helps evaluate the model's performance, reduce overfitting, and estimate its generalization ability. This study uses 4-fold cross-validation to divide the dataset into four parts, improving evaluation reliability by training on three parts and testing on one, which minimizes overfitting and provides a more accurate performance estimate. Four DTW values were obtained per period by dividing the time series data into four segments, providing a clearer view. The specific number of folds (3, 4, or 5) is less critical (Yoon, 2021). Each cross-validation iteration uses one part as the test set and the rest as the training set, ensuring diverse evaluation and reducing overfitting. (Ghojogh and Crowley, 2019).

3.2.3. Regression-based machine learning

The goal is to use the hydrological and climate variables to investigate the ability of proxy basin variables (e.g., Yiba and Reem) to predict the impacted basin variables (e.g., Hali and Baish), enabling a comparative analysis of changes due to dam construction. Specifically, the study aims to predict the variables in the impacted basins based on models built from the proxy basins to assess how these variables would differ in the absence of dams (2003–2009) compared to their state with dams (2010–2020). By using linear regression as the primary regression-based machine learning model in MATLAB (Bhartendu, 2024), the study aims to predict and quantify the impact of dam construction on key environmental variables, providing insights into the long-term environmental processes influenced by dam infrastructure.

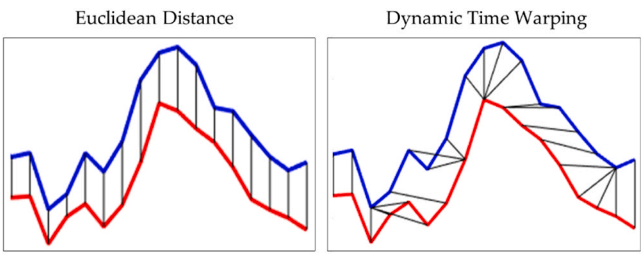


Fig. 3. Differences between Euclidean Distance (left) calculation and Dynamic Time Warping (right).

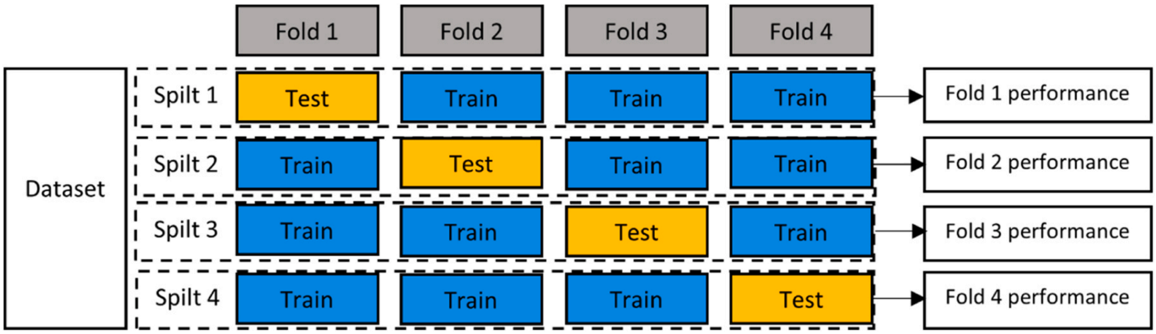


Fig. 4. The dataset underwent 4-fold cross-validation, where it was divided into four subsets. Four subsets were used for training in each fold, while one subset was reserved for validation. Following the evaluation, these four subsets were utilized to train a final model, with the remaining subset used to assess the system's performance.

3.2.3.1. Data preparation. The predictor variables (X) were derived from proxy basins including Yiba and Reem, while the variables from Hali and Baish were used as the targets for prediction. This approach assumes that the environmental conditions in Yiba can estimate the conditions in Hali due to their similar geographic and climatic characteristics and comparable land use patterns. A similar assumption was made for the other two case studies, where conditions in Reem can estimate those in Baish. The dataset was split into training (70 %) and testing (30 %) sets, allowing the model to learn from a substantial portion of the data while reserving enough to evaluate its predictive capabilities. This split ensures the model was trained on diverse data points, helping it generalize well to new, unseen data. As a representative sample, the testing set provides a robust measure of the model's performance, helping to identify overfitting or underfitting issues.

3.2.3.2. Model training and prediction. This section uses the linear regression model to train the data after splitting it. It was chosen for its simplicity, interpretability, and effectiveness in capturing linear relationships between variables (Schielzeth, 2010). Model fitting involves minimizing the difference between observed and predicted values by optimizing coefficients. Specifically, the optimization was carried out by minimizing the Root Mean Squared Error (RMSE), which provides a measure of how closely the model's predictions match the actual values. The linear regression model was expressed by Eq. 4:

$$Y = B_0 + B_1X_1 + B_2X_2 + \ldots + B_nX_n + e \tag{4}$$

Where Y represents the response variable (e.g., NDVI, groundwater, soil salinity, and soil moisture for Hali and Baish), $X_1X_2, \ldots. X_n$ are the predictor variables (e.g., NDVI, soil salinity, groundwater, soil moisture, runoff water, precipitation, temperature, and OCS for Yiba and Reem basins), B_0 is the intercept, and $B_1B_2, \ldots. B_n$ are the coefficients for the predictor variables (e.g., corresponding data from Yiba for Hali and Reem for Baish).

In the Yiba-Hali pair, data from Yiba's 2003–2009 period were used to predict Hali's 2003–2009 and 2010–2020 periods, capturing the baseline relationship, and to assess how the model trained on pre-dam data performs in the post-dam periods. Similarly, for the Reem-Baish pair, data from 2003 to 2009 of Reem were used to predict Baish's 2003 to 2009 and 2010–2020 periods. The testing data for each of these periods consisted of 30 % of the available data from each period, ensuring robust model validation. The training process involves fitting the model to the training data and optimizing the coefficients best to explain the relationship between the predictor and response variables. This process minimizes the difference between the observed and predicted values. Once trained, the linear regression model predicts the environmental variables for the testing set by applying the fitted model to the test data's predictor variables as in Eq. 5:

$$\widehat{Y}_{test} = B_0 + B_1X_{1,test} + B_2X_{2,test} + \ldots + B_nX_{n,test} \tag{5}$$

Where $\widehat{Y}_{test}$ represents the predicted values for the response variables based on the testing data, $X_{i,test}$ are the predictor variables from the testing data, and B_i are the coefficients obtained from the training data. Using the trained linear regression model, the predictive accuracy for environmental variables, both in the presence and absence of a dam, is assessed, enabling comparison across different conditions. This approach allows for the analysis of the relationship between predictor and response variables, providing valuable insights into how dam construction impacts downstream environments.

3.2.3.3. Evaluate the model. Several statistical measures were computed to assess the model's performance. The R-squared (R2) value measures the proportion of variance in the dependent variable predictable from the independent variables and providing an overall measure of the model's goodness of fit (Chicco et al., 2021; Menard, 2001). A higher R^2 value indicates that the model explains a greater portion of the variance in the response variable, suggesting a stronger relationship between the predictors (data from Yiba and Reem) and the response variables (data from Hali and Baish). R^2 is expressed in Eq. 6:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} \tag{6}$$

Where n is the number of observations, Yi is the actual observed value for the ith observations, and $\widehat{Yi}$ is the predicted value for the ith observations.

The Root Mean Squared Error (RMSE) measures a model's accuracy by taking the square root of the average squared errors (Willmott and Matsuura, 2005). It provides an interpretable metric in the same units as the response variable. A lower RMSE signifies better predictive accuracy, showing how closely the model's predictions align with the actual observed values of environmental variables (Chicco et al., 2021). Mathematically, the RMSE expressed in Eq. 7:

$$RMSE = \sqrt{\frac{1}{2} \sum_{i=1}^{n}(Yi - \widehat{Yi})}^{2} \tag{7}$$

where n is the number of observations, Yi is the actual observed value for the ith observations, and $\widehat{Yi}$ is the predicted value for the ith observations.

Mean Absolute Error (MAE) is calculating the average magnitude of errors in predictions (Willmott and Matsuura, 2005). It is the mean of the absolute differences between observed and predicted values. MAE is expressed in Eq. 8:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |Y_i - \widehat{Y}_i| \tag{8}$$

Where n is the number of observations, Yi is the actual observed value for the ith observations, and $\widehat{Yi}$ is the predicted value for the ith observations.

4. Results

Section 4.1 presents the statistical analysis outcomes, focusing on the correlation coefficients and their changes between the study areas for the periods 2003–2009 and 2010–2020 (before and after the dam construction, respectively) for both pairs of basins. Section 4.2 details the DTW analysis, utilizing k-fold cross-validation to compare the Yiba and Hali basins and the Reem and Baish basins during the same periods. Additionally, Section 4.3 discusses the application of regression-based machine learning, which uses the proxy basins dataset (e.g., Yiba and Reem) to predict the environmental impacts on the dam-affected basins (e.g., Hali and Baish) for these periods.

4.1. Correlation coefficients between environmental variables

Table 2 shows the correlation coefficients for environmental factors in 2003–2009 and 2010–2020. Notably, the correlations for NDVI in Yiba-Hali decrease significantly after dam construction (from 0.87 to 0.59), while the change for Reem-Baish (0.71–0.63) is more moderate. For soil salinity, correlations decrease across both pairs, particularly for Yiba-Hali (from 0.81 to 0.52). Similarly, groundwater correlations show a decline in both pairs, with significant drops observed in Yiba-Hali (from 0.92 to 0.74) and Reem-Baish (from 0.91 to 0.57). Soil moisture correlations remain relatively stable, slightly decreasing Yiba-Hali (from 0.94 to 0.85) and Reem-Baish (from 0.96 to 0.92). Therefore, soil moisture variations in the study areas are highly related to the precipitation pattern and follow its trends. In the Yiba-Hali pair, runoff water correlation decreased from 0.88 in the first period to 0.67 in the second, resulting in a correlation reduction of 0.21. Similarly, in the Reem-Baish pair, runoff water correlation dropped from 0.87 to 0.68, with a correlation reduction of 0.19, which indicates a significant impact of dam construction on runoff water dynamic. Temperature correlations exhibit a minimal change, maintaining high consistency in Yiba-Hali (remaining at 0.99) and Reem-Baish (from 0.99 to 0.98). Precipitation correlations also remain almost unchanged for Yiba-Hali (from 0.96 to 0.95) and stable for Reem-Baish (remaining

Table 2
Correlation coefficients and their reduction for environmental variables between proxy and Impacted Basins for the 2003–2009 and 2010–2020. The table presents the correlation coefficients for NDVI, soil salinity, groundwater, soil moisture, runoff, OCS water across the paired basins of Yiba-Hali and Reem-Baish, highlighting the effects of dam construction on these environmental variables.

Correlation coefficient							
	Periods	NDVI	Soil salinity	Groundwater	Soil moisture	Runoff water	OCS
Yiba - Hali	2003–2009	0.87	0.81	0.92	0.94	0.88	0.98
	2010–2020	0.59	0.52	0.74	0.85	0.67	0.92
Reem - Baish	2003–2009	0.71	0.7	0.91	0.96	0.87	0.92
	2010–2020	0.63	0.48	0.57	0.92	0.68	0.89
Correlation change							
		NDVI	Soil salinity	Groundwater	Soil moisture	Runoff water	OCS
Yiba - Hali		−0.28	−0.29	−0.18	−0.09	−0.21	−0.06
Reem - Baish		−0.08	−0.09	−0.34	−0.04	−0.19	−0.03

at 0.92).

The reduction in correlation coefficients between 2003 and 2009 and 2010–2020 underscores notable shifts in relationships between proxy and impacted basins across environmental variables (Table 2). For NDVI, the correlation for Yiba-Hali between the periods decreased by 0.28, indicating a notable impact of dam construction on vegetation patterns due to altered water availability. In comparison, the reduction for Reem-Baish was less pronounced (by 0.08), suggesting considerable changes in vegetation. Soil salinity showed a significant reduction in correlation for Yiba-Hali (0.29), reflecting changes in soil composition due to altered water flow and sediment distribution, while Reem-Baish experienced a moderate reduction of 0.09. Groundwater exhibited the most substantial decrease for Reem-Baish (0.34), indicating significant changes in groundwater dynamics post-dam construction, with Yiba-Hali (0.18) also showing notable impacts. Soil moisture remained relatively stable across both pairs with a minimal reduction of 0.09 for Yiba-Hali and 0.04 for Reem-Baish, suggesting that overall soil moisture content did not vary drastically. Temperature and precipitation patterns were largely unaffected by dam construction, with negligible reductions in correlation values (0–0.01). OCS data showed strong correlations between proxy and dammed basins (Yiba-Hali: 0.98–0.92, Reem-Baish: 0.92–0.89) with minimal post-dam changes (-0.06 and −0.03). These findings show that dam construction significantly affected NDVI, soil salinity, groundwater, and runoff water, increasing variability and reducing similarity between basins.

4.2. Dynamic time warping (DTW) analysis

The DTW method aligns time series data, allowing for detailed comparisons of environmental similarities and differences between basins with and without dams. In the DTW results, similarities for 2003–2009 and 2010–2020 were evaluated using a 4-fold cross-validation approach, rather than just averaging values between proxy and impacted basins. Each time series was divided into 4 portions, and DTW distances were calculated for each fold. In addition, the average DTW is considered to assess the similarities of environmental variables between the case studies. The following DTW sections are Section 4.2.1, which shows the results for the Yiba and Hali basins, and Section 4.2.2, which shows the results for the Reem and Baish basins.

4.2.1. Yiba and Hali basins

In Fig. 5, the DTW values for NDVI, soil salinity, soil moisture, temperature, and precipitation between Yiba and Hali (2003–2009) before the dam construction are close to 0 with average DTW values of 0.08, 0.09, 0.08, 0.01, 0.04 respectively as also shown in Table 3, indicating high similarity in the datasets. This high similarity could be attributed to similar climatic conditions, comparable land use, vegetation cover, and shared environmental variables affecting these parameters, such as the amount and timing of precipitation and temperature variations (Almalki et al., 2025; Şen et al., 2013). However, the groundwater and runoff water DTW values show moderate similarity, with an average fold value of 0.30 for groundwater and 0.45 for runoff water. These differences before the dam construction could be due to variations in groundwater recharge rates and extraction, influenced by factors such as soil permeability, aquifer characteristics, and local geology (Al-Turki, 1995; Luo et al., 2023). Different subsurface geological formations and soil types can affect groundwater storage and flow pattern (de Rooij, 2016). Human activities, such as variations in groundwater extraction and usage for agricultural or other purposes, can also lead to differences in groundwater levels and dynamic (Almalki et al., 2023). In addition, the DTW value for runoff water reflects some alignment in runoff timing and magnitude but also highlights natural variability due to differences in rainfall, topography, and land use. This moderate value captures the complexity of runoff behaviour in arid environments before the dam's influence.

In the second period (2010–2020), after the dam construction, the DTW distance values for the Yiba-Hali analysis increased compared to the first period (2003–2009). This increase is particularly notable for runoff water, which had a high DTW value of 0.90. In comparison, NDVI, soil salinity, groundwater, soil moisture, and OCS had average DTW values of 0.25, 0.25, 0.49, 0.14, and 0.10, respectively, as also shown in Table 3. These higher DTW values indicate moderate similarities in these datasets compared to the first period. In contrast, DTW values for temperature and precipitation remained near 0, with averages of 0.01 and 0.05, similar to the first

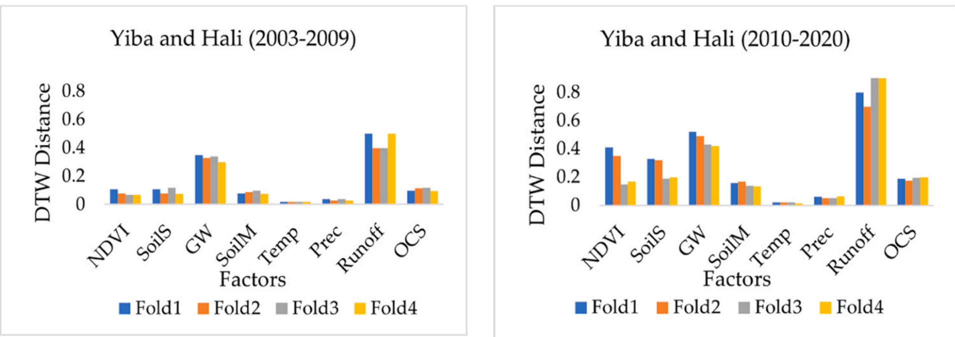


Fig. 5. DTW distances to assess the similarities of the 4-folds between the considered environmental variables of the proxy (e.g., Yiba) and impacted (e.g., Hali) basins for 2003–2009 and 2010–2020. soil salinity (SoilS), groundwater (GW), soil moisture (SoilM), temperature (Temp), precipitation (Prec) and OCS, spatially averaged over each basin. Refer to Fig. 1 for the geographical locations of the basins.

Table 3
The average DTW distances used to assess the similarities between the considered environmental variables of the proxy (e.g., Yiba) and impacted (e.g., Hali) basins for 2003–2009 and 2010–2020.

	Period	NDVI	Soil salinity	Groundwater	Soil moisture	Temperature	Precipitation	Runoff water	OCS
Yiba - Hali	2003–2009	0.08	0.09	0.30	0.09	0.01	0.04	0.45	0.10
	2010–2020	0.25	0.25	0.49	0.14	0.01	0.05	0.90	0.19
Percentage changes		212 %	177 %	63 %	55 %	0 %	25 %	100 %	90 %

period. This indicates high similarity in these variables across proxy and impacted basin periods before and after the dam construction. Table 3 shows that after dam construction, NDVI, soil salinity, and runoff water experienced the largest increases in DTW distances—212 %, 177 %, and 100 %, respectively—indicating significant reductions in similarity. Groundwater, soil moisture, and OCS experienced moderate increases of 63 %, 55 %, and 90 %, respectively. Temperature patterns remained unchanged, showing no impact, and precipitation similarity decreased slightly by 25 %. These results highlight that NDVI and runoff water were most affected by the dam, consistent with correlation analysis.

Fig. 6 shows more distance variation in the second period (2010–2020) than in the first period (2003–2009). The dam construction in Hali during the second period likely explains the changes in DTW values compared to the first period. Additionally, changes in land use, such as increased urbanization or agricultural activities, could contribute to the more significant distance variations. The presence of the dam alters water flow and distribution, impacting various downstream environmental variables such as vegetation, soil salinity, and groundwater, as demonstrated in several studies (Al-Robai et al., 2018; Al-Sodany et al., 2015; Almalki et al., 2023; Sallam et al., 2018). Dam construction significantly affects runoff water, which in turn influences NDVI, soil salinity, and groundwater. Changes in runoff alter vegetation distribution and soil salinity due to shifts in water flow and irrigation. Additionally, altered runoff impacts groundwater recharge and levels, as well as soil moisture distribution and water retention downstream. However, the dam has not significantly impacted temperature and precipitation, as indicated by the similar DTW values between the two periods. This observation is consistent with studies showing that large dams typically have a minimal effect on the local climate in arid areas (Afzal et al., 2023; Degu et al., 2011).

4.2.2. Reem and Baish basins

In the first period for Reem and Baish, DTW values for NDVI, soil salinity, soil moisture, and OCS range from 0.10 to 0.26, with averages of 0.16, 0.15, 0.10, and 0.26, respectively, as shown in Fig. 7 and Table 4. These values, again, suggest high similarity in the datasets between the two basins, indicating comparable vegetation cover and soil salinity patterns during this period. However, the groundwater and runoff water DTW values for all folds are around 0.5, as depicted in Fig. 7, with an average DTW values of 0.52 and 0.53, suggesting a moderate similarity in groundwater conditions between the two basins, with variable recharge rates, different human extraction levels, and geological differences could influence. In contrast, the DTW values for temperature and precipitation are close to 0, similar to those observed between Yiba and Hali. These low DTW values indicate high similarity in temperature and precipitation patterns between Reem and Baish, suggesting similar environmental conditions during the first period.

In the second period, where only Baish has a dam, and Reem does not, the DTW values for all folds of NDVI, soil salinity, soil moisture, and OCS increased, with average DTW values of 0.31, 0.30, 0.18, and 0.41, respectively, as shown in Table 4. These increases indicate changes in the datasets compared to the first period, suggesting that vegetation cover and soil salinity patterns in Baish diverged from those in Reem due to the dam's construction. Additionally, the runoff water DTW value of 1.33 reflects significant changes in water flow patterns between the two basins. The groundwater DTW values for all folds ranged between 1 and 1.4, with an average DTW value of 1.19, indicating very low similarities between the datasets compared to the first period. In Table 4, NDVI, soil salinity, groundwater, and runoff water experienced the largest increases, 93 %, 100 %, 128 % and 150 %, respectively, indicating substantial reductions in similarity. Soil moisture and OCS saw moderate increases of 80 % and 58 %, while temperature and precipitation similarity decreased slightly by 25 % and 22 %. These findings align with the impact observed in correlation analysis, confirming the substantial effects of dam construction.

Fig. 8 shows low variation in the dataset of NDVI, soil salinity, and groundwater in the first period. In the second period, distance values slightly increased for NDVI and soil salinity, while groundwater and runoff water showed high variation. This variation increase corresponds to the dam's presence in Baish during the second period. These results highlight significant differences in groundwater conditions between Baish and Reem in the second period, likely influenced by the dam's impact on Baish's hydrological system. Conversely, the DTW values for temperature and precipitation remained similar to the first period, showing high similarities with DTW values for all folds close to 0, and average values of 0.05 and 0.11, respectively. These findings suggest that the construction of the Baish dam primarily impacted groundwater dynamics, runoff water, NDVI, soil salinity, and soil moisture, while other variables remained relatively consistent between the two periods, as observed in the initial results of the Yiba-Hali analysis. Another study on the environmental consequences of the dam in the Baish area supports the conclusion that the Baish dam had a more significant effect on groundwater and runoff water than on other environmental variables (Sallam et al., 2018).

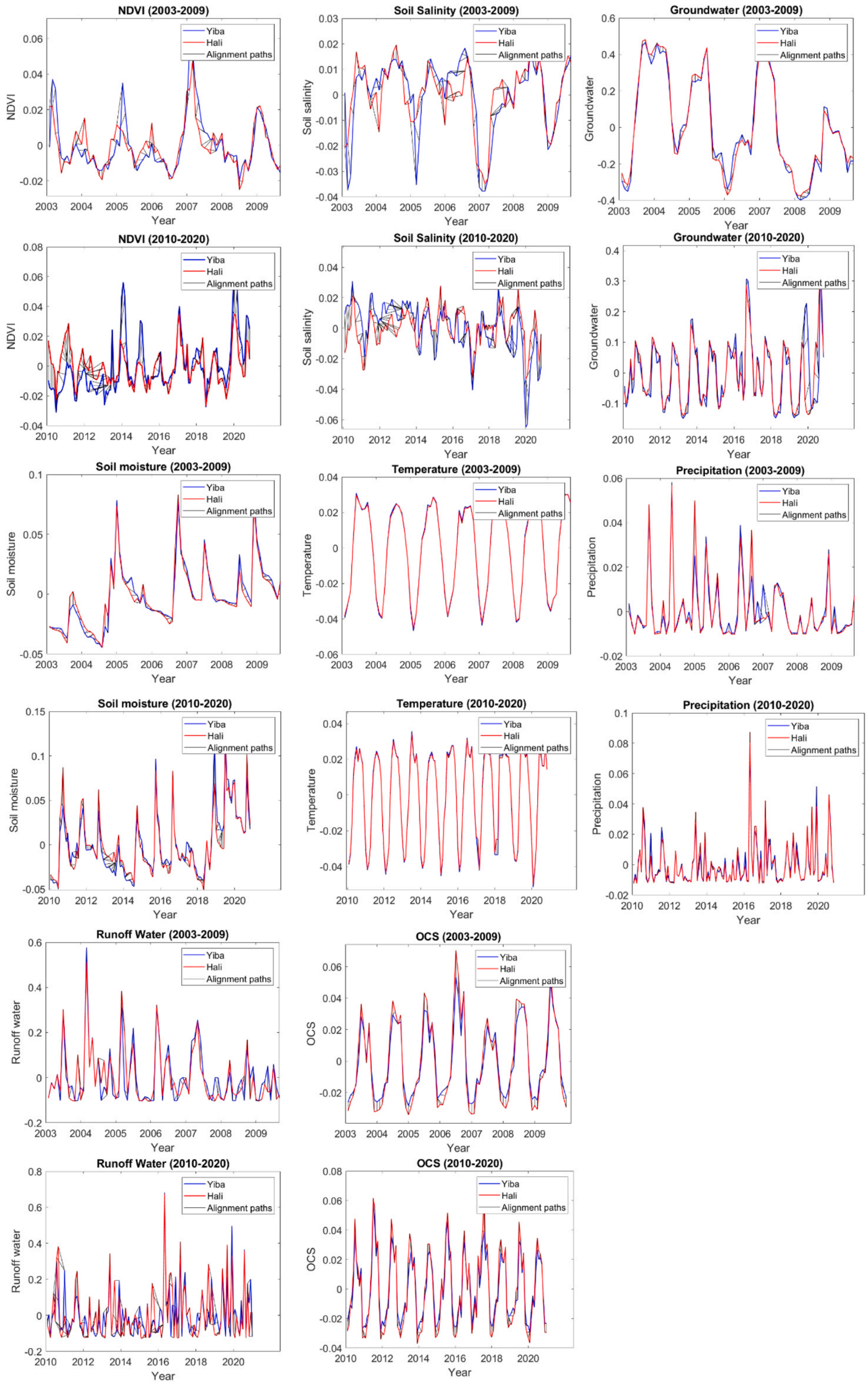
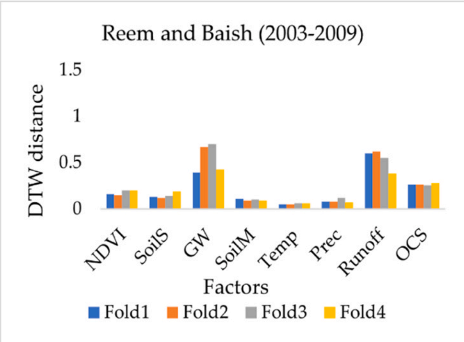


Fig. 6. DTW is used to assess the similarities between the time series of environmental variables from the proxy basin (e.g., Yiba) and the impacted basin (e.g., Hali) for the periods 2003–2009 and 2010–2020. The figure overlays the DTW alignment paths to illustrate how individual time points in one series are aligned with corresponding points in the other series, providing a detailed view of how the time series align with each other over time.



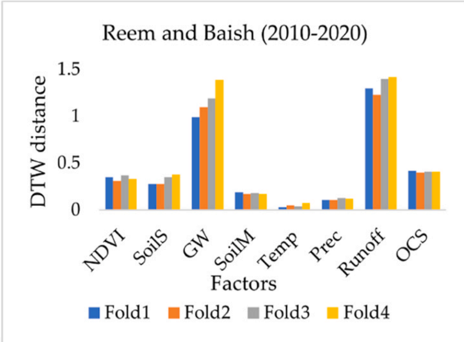


Fig. 7. DTW distances to assess the similarities of the 4-folds between the considered environmental variables of the proxy (e.g., Reem) and impacted (e.g., Baish) basins for 2003–2009 and 2010–2020. Soil Salinity (SoilS), Groundwater (GW), Soil Moisture (SoilM). Temperature (Temp), Precipitation (Prec) and OCS, spatially averaged over each basin. Refer to Fig. 1 for the geographical locations of the basins.

Table 4
The average DTW distances used to assess the similarities between the considered environmental variables of the proxy (e.g., Reem) and impacted (e.g., Baish) basins for 2003–2009 and 2010–2020.

	Period	NDVI	Soil salinity	Groundwater	Soil moisture	Temperature	Precipitation	Runoff water	OCS
Reem-Baish	2003–2009	0.16	0.15	0.52	0.10	0.04	0.09	0.53	0.26
	2010–2020	0.31	0.30	1.19	0.18	0.05	0.11	1.33	0.41
Percentage changes		93 %	100 %	128 %	80 %	25 %	22 %	150 %	58 %

4.3. Regression-based machine learning analysis

Regression-based machine learning models were also used to predict environmental variables in Hali and Baish for 2003–2009 and 2010–2020. Trained on data from Yiba and Reem from the period 2003–2009, these models used 70 % of the dataset for training and 30 % for testing. The primary aim was to predict environmental variables (e.g., NDVI, soil salinity, groundwater, soil moisture) for Hali and Baish during both periods, allowing for a comparison of conditions with and without dams. The results, including R^2 values, correlation between actual and predicted data, and RMSE and RAE values, help highlight these differences and provide insights into the broader implications of dam impacts.

4.3.1. Yiba and Hali basins analysis

In Fig. 9, for 2003–2009, high R^2 values across all environmental variables—NDVI (0.82), soil salinity (0.77), groundwater (0.98), soil moisture (0.96), and OCS (0.96)—indicate strong predictive accuracy. In the period 2010–2020, the R^2 values decreased significantly for NDVI (0.37), soil salinity (0.38), groundwater (0.34), and soil moisture (0.24). This sharp decline in R^2 values across all environmental variables highlights the substantial impact of dam construction, severely reducing the predictive accuracy between Yiba and Hali.

Table 5 shows the correlation values indicating the strength and direction of the linear relationship between actual and predicted values for Yiba and Hali. During 2003–2009, strong correlations were observed for all environmental variables for NDVI (0.84), soil salinity (0.87), groundwater (0.99), soil moisture (0.97), and OCS (0.98). This suggests that the regression-based machine learning model accurately predicted these variables in an undisturbed environment before the dam's construction. However, for 2010–2020, the correlations decreased—NDVI (0.61), soil salinity (0.59), groundwater (0.44), and soil moisture (0.37)—indicating that the dam's construction significantly disrupted the accuracy of the model's predictions for these environmental variables.

Table 5, for 2003–2009, low RMSE values across all variables (NDVI: 0.001, soil salinity: 0.005, groundwater: 0.030, and soil moisture: 0.005) indicate high prediction accuracy. However, in 2010–2020, RMSE values increased for NDVI (0.009), soil salinity (0.010), groundwater (0.081), soil moisture (0.024), and OCS (0.003), indicating decreased prediction accuracy. Additionally, during the 2003–2009 period, the MAE values were low across all variables: 0.006 for NDVI, 0.005 for soil salinity, 0.002 for groundwater, and 0.004 for soil moisture, further demonstrating high prediction accuracy before dam construction. In the 2010–2020 period, the MAE values increased to 0.008 for NDVI, 0.009 for soil salinity, 0.067 for groundwater, and 0.029 for soil moisture, reflecting decreased prediction accuracy post-dam construction. The significant increase in MAE for groundwater (0.067) suggests substantial changes in groundwater dynamics due to altered water flow and storage patterns. Similarly, the increased MAE for soil moisture (0.029) indicates heightened sensitivity to changes in water management and distribution. These results highlight the impact of dam construction on the predictability of environmental variables and underscore the need for more sophisticated modeling approaches. The significant changes in R^2, correlation, RMSE, and MAE values between the two periods can be attributed to the dam construction in Hali during 2010–2020. The dam likely caused substantial alterations in environmental conditions, primarily affecting NDVI, soil salinity, groundwater, and soil moisture. As several studies in arid areas have shown (Al-Robai et al., 2018; Nilsson and Berggren,

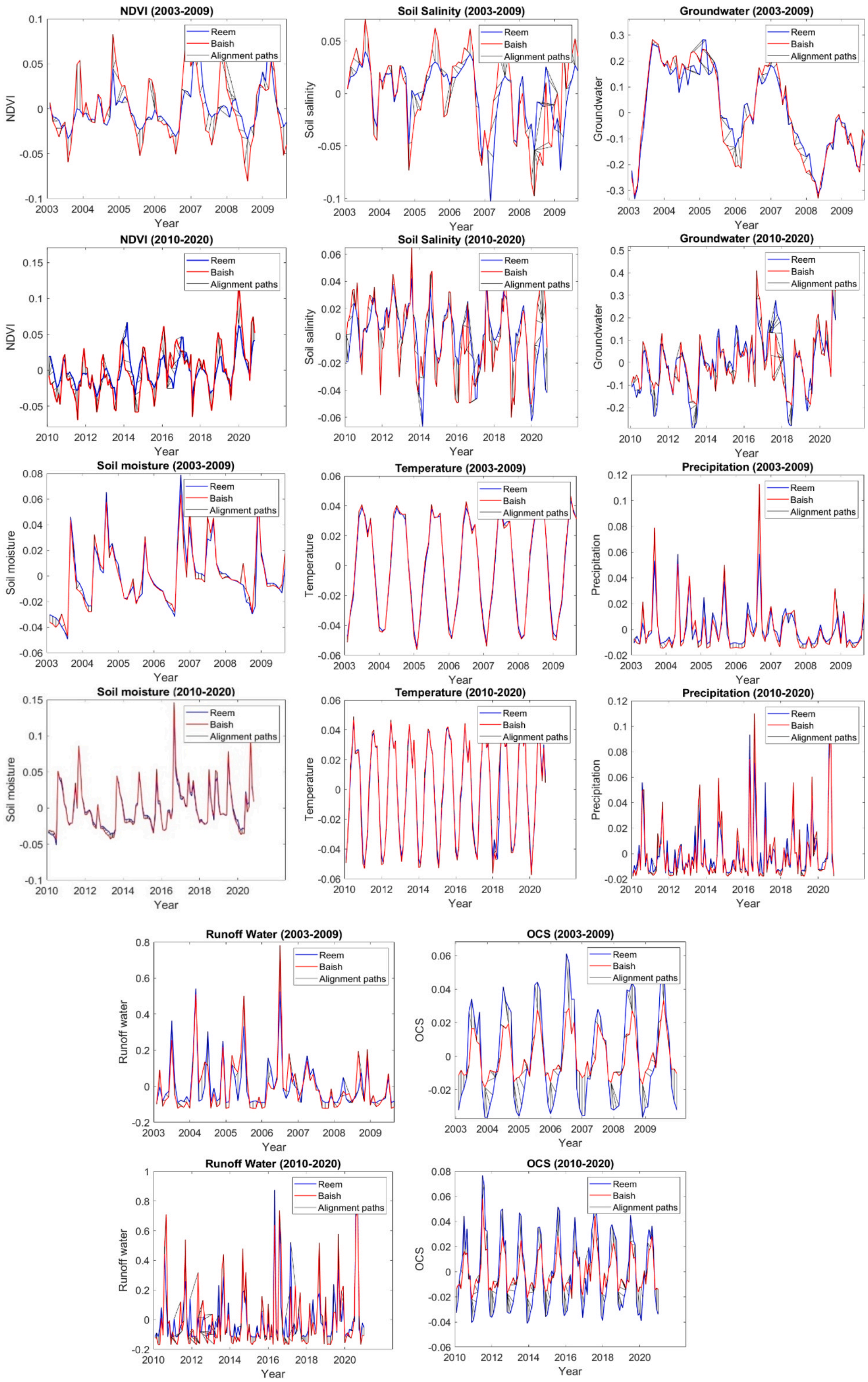


Fig. 8. DTW is used to assess the similarities between the time series of environmental variables from the proxy basin (e.g., Reem) and the impacted basin (e.g., Baish) for the periods 2003–2009 and 2010–2020. The figure overlays the DTW alignment paths to illustrate how individual time points in one series are aligned with corresponding points in the other series, providing a detailed view of how the time series align with each other over time.

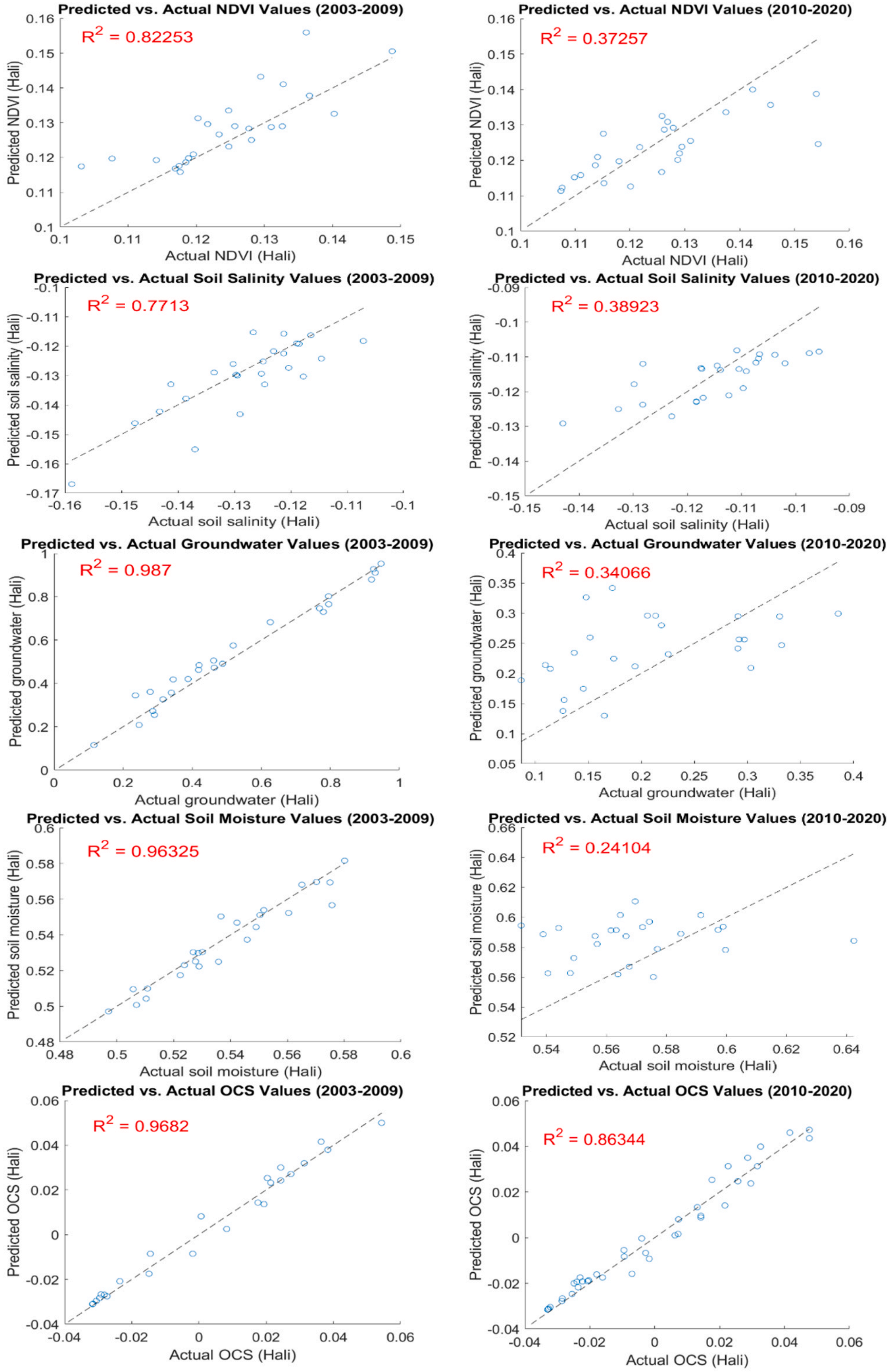


Fig. 9. Regression-based machine learning analysis of environmental variables using proxy basins (e.g., Yiba) to predict impacted basins (e.g., Hali) for 2003–2009 and 2010–2020.

Table 5
Correlation, RMSE, and MAE values Between actual and predicted data for environmental variables in the Yiba-Hali basin for the periods 2003–2009 (before the dam construction) and 2010–2020 (after the dam construction). The table presents the correlation coefficients, RMSE, MAE values for NDVI, soil salinity, groundwater, soil moisture, and OCS highlighting the predictive performance of the machine learning model before and after dam construction.

Correlation between actual and predicted data						
	Period	NDVI	Soil Salinity	Groundwater	Soil moisture	OCS
Yiba - Hali	2003–2009	0.84	0.87	0.99	0.97	0.98
	2010–2020	0.61	0.59	0.44	0.37	0.90
RMSE						
	Period	NDVI	Soil Salinity	Groundwater	Soil moisture	OCS
Yiba - Hali	2003–2009	0.001	0.005	0.030	0.005	0.003
	2010–2020	0.009	0.010	0.081	0.024	0.004
MAE						
	Period	NDVI	Soil Salinity	Groundwater	Soil moisture	OCS
Yiba - Hali	2003–2009	0.006	0.005	0.002	0.004	0.002
	2010–2020	0.008	0.009	0.067	0.029	0.003

2000; Sallam et al., 2018), the dam's construction altered water availability downstream, influencing vegetation patterns and increasing soil salinity levels.

4.3.2. Reem and Baish basins analysis

Fig. 10 illustrates the R^2 values, which indicate the proportion of variance in environmental variables explained by the regression-based machine learning models for the basins studied. For the period 2003–2009, the R^2 values were relatively high: NDVI (0.69), soil salinity (0.71), groundwater (0.94), soil moisture (0.94), and OCS (0.85), reflecting strong predictive power across most variables. However, during 2010–2020, these values decreased significantly: NDVI (0.39), soil salinity (0.40), groundwater (0.33), and soil moisture (0.30). This decline suggests that dam-induced changes in river flow have significantly impacted vegetation and water availability, affecting both groundwater and soil moisture. Reduced downstream recharge rates and increased variability in groundwater levels challenge accurate prediction (Kelly et al., 2013). Additionally, sediment trapping alters soil composition and salinity (van Maren et al., 2013), while reservoirs increase evaporation rates, further affecting local groundwater and soil moisture levels.

In addition, Table 6 shows strong correlations for 2003–2009 are observed for NDVI (0.84), soil salinity (0.79), groundwater (0.97), and soil moisture (0.94), indicating a strong linear relationship between actual and predicted values, reflecting the model's accuracy as shown in Table 6. For 2010–2020, correlations for NDVI (0.70), soil salinity (0.62), groundwater (0.57), and soil moisture (0.62) decreased. Furthermore, low RMSE values for 2003–2009 for NDVI (0.011), soil salinity (0.013), groundwater (0.011), and soil moisture (0.004), indicate high prediction accuracy during this period, as shown in Table 5. In 2010–2020, RMSE values for NDVI (0.024), soil salinity (0.022), groundwater (0.063), and soil moisture (0.026) showed increasing of RMES, suggesting the model's predictive accuracy for these variables experienced significant fluctuations. The increase in RMSE for groundwater likely reflects dam construction's complex and variable impact on groundwater levels, making them harder to predict accurately. From 2003 to 2009, MAE values for the Reem and Baish basins were low for NDVI (0.012), soil salinity (0.011), groundwater (0.009), soil moisture (0.005) and OCS (0.003), indicating high prediction accuracy before dam construction. However, from 2010 to 2020, MAE increased to 0.020 for NDVI, 0.016 for soil salinity, 0.058 for groundwater, and 0.018 for soil moisture, showing decreased prediction accuracy post-dam. The rise in MAE for groundwater suggests major changes in its dynamics, while the increase for soil moisture indicates greater sensitivity to water management, emphasizing the need for more refined modelling.

The construction of the dam in Baish has markedly altered environmental conditions, significantly impacting groundwater dynamics, soil salinity, soil moisture, and NDVI. The increased variability in groundwater and soil moisture, as reflected in the lower R^2 and higher RMSE and MAE values, suggests that the dam disrupted natural recharge and flow patterns, complicating accurate predictions. In addition, the significant impact of the dam on NDVI and soil salinity, as evidenced by the lower R^2 values for these variables, suggests that the model struggled to capture general trends and was challenged by the increased complexity and variability introduced by the dam.

5. Discussion

5.1. Remote sensing and space-for-time substitution approach

Remote sensing data and the space-for-time substitution approach were essential for assessing environmental changes and similarities across basins, particularly regarding dam construction in arid regions. In arid regions, where the lack of ground-based data poses significant challenges, remote sensing overcomes this problem by providing a comprehensive view of environmental conditions, enabling the analysis of large-scale changes over time. The space-for-time substitution approach further enhances this analysis by simulating long-term temporal changes using spatial data, which is particularly valuable when direct temporal data is limited. However, beyond remote sensing and space-for-time, the use of DTW and regression analysis offers a deeper, more nuanced

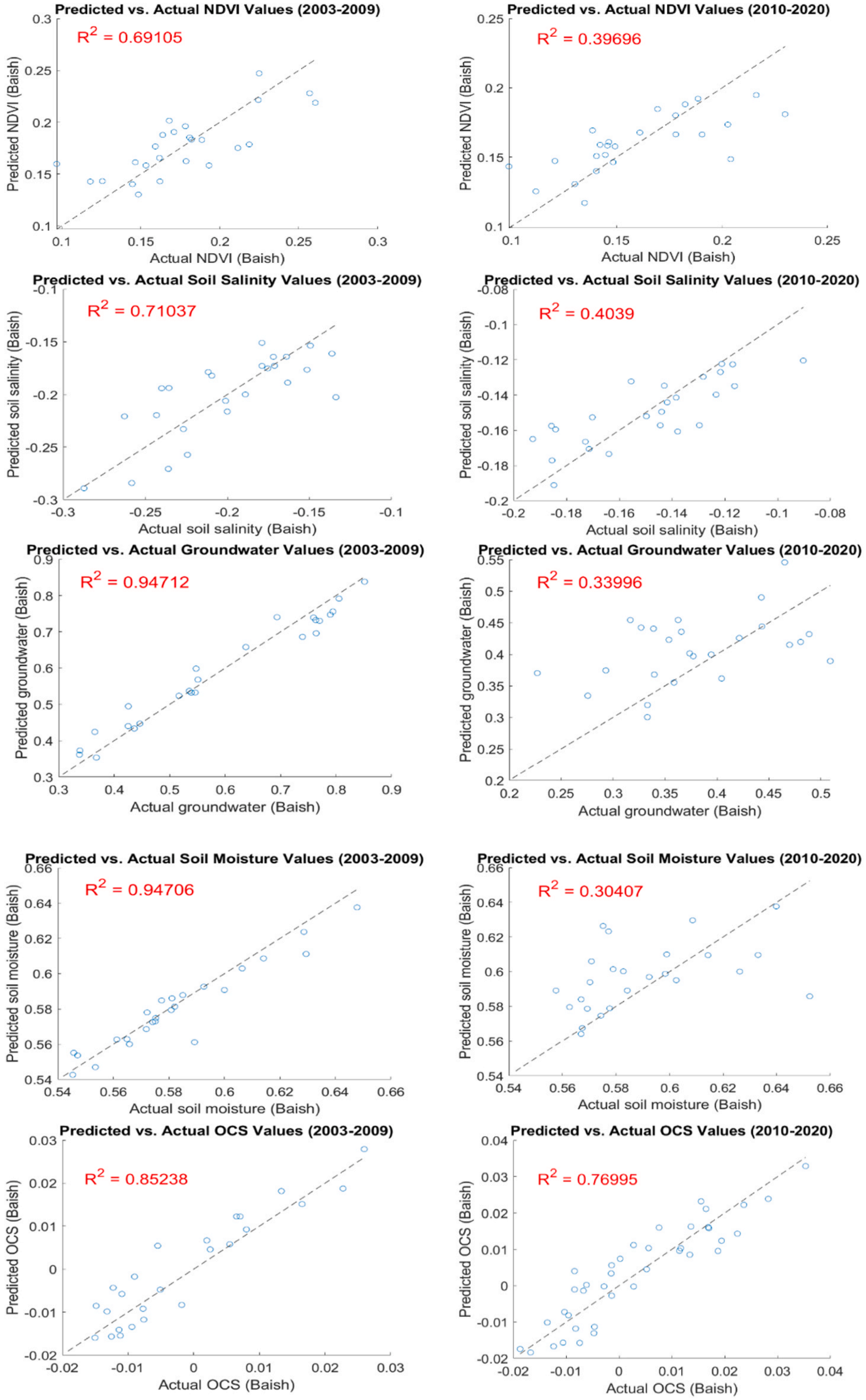


Fig. 10. Regression-based machine learning analysis of environmental variables using proxy basins (e.g., Reem) to predict impacted basins (e.g., Baish) for 2003–2009 and 2010–2020.

Table 6
Correlation, RMSE, and MAE values between actual and predicted data for environmental variables in the Reem-Baish basin for the periods 2003–2009 (before the dam construction) and 2010–2020 (after the dam construction). The table presents the correlation coefficients, RMSE, and MAE values for NDVI, soil salinity, groundwater, soil moisture, and OCS highlighting the predictive performance of the machine learning model before and after dam construction.

Correlation between actual and predicted data						
	Period	NDVI	Soil salinity	Groundwater	Soil moisture	OCS
Reem - Baish	2003–2009	0.84	0.79	0.97	0.94	0.92
	2010–2020	0.70	0.62	0.57	0.62	0.88
RMSE						
	Period	NDVI	Soil salinity	Groundwater	Soil moisture	OCS
Reem - Baish	2003–2009	0.011	0.013	0.011	0.004	0.004
	2010–2020	0.024	0.022	0.063	0.026	0.005
MAE						
	Period	NDVI	Soil salinity	Groundwater	Soil moisture	OCS
Reem - Baish	2003–2009	0.012	0.011	0.009	0.005	0.003
	2010–2020	0.020	0.016	0.058	0.018	0.004

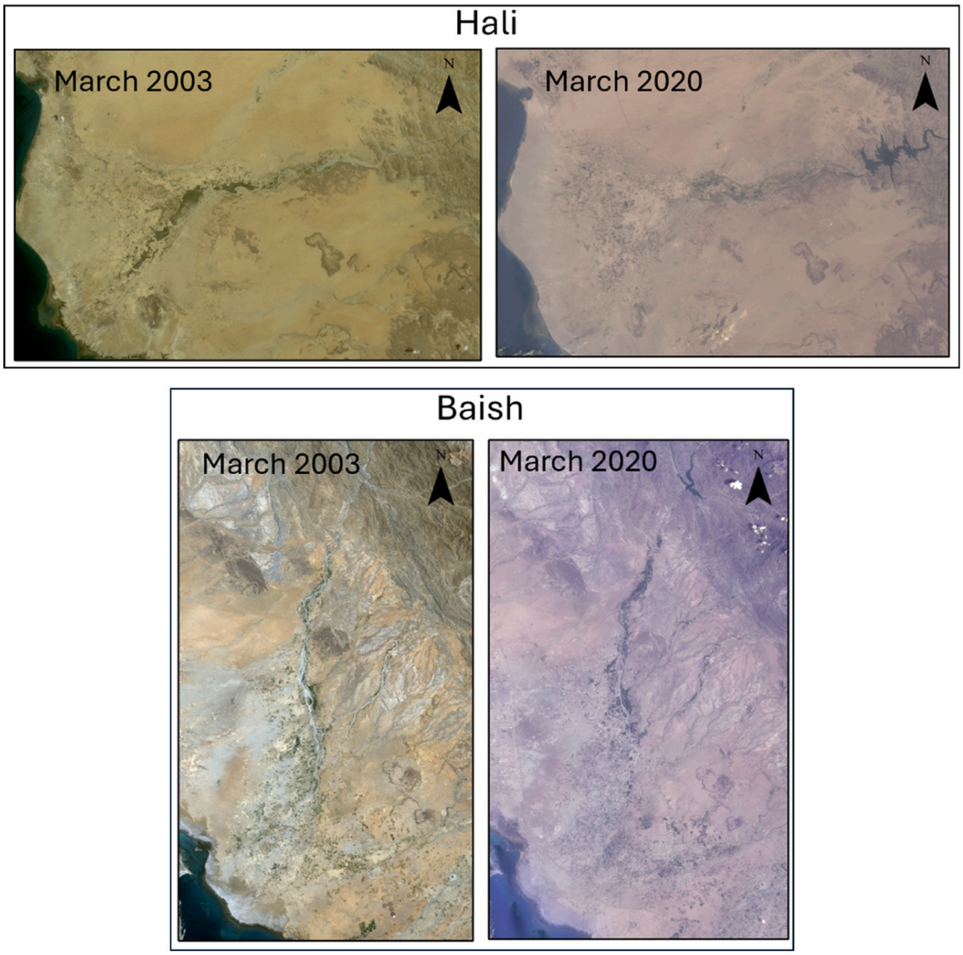


Fig. 11. Landsat images of the Hali and Baish basins before and after dam construction (March 2003 and March 2020, respectively).

understanding of dam impacts. DTW excels in identifying and aligning time-series data that may be out of phase due to delays caused by dam construction. This allows for a more precise analysis of how environmental variables evolve in response to dam interventions. Regression analysis, in turn, provides a robust statistical framework for predicting how these variables would have behaved without dam construction, offering concrete evidence of the extent of these impacts. By utilizing these advanced methodologies, the study fills a crucial research gap by providing statistically supported evidence that links dam construction to environmental changes.

Dam construction altered variables affecting vegetation, with uniform water distribution maintaining higher correlations pre-construction, while uneven distribution post-construction reduced basin similarity. Dams also reduce sediment flow and nutrient distribution, affecting soil composition and decreasing soil salinity correlations (Al-Robai et al., 2018). Water and nutrient availability changes impacted vegetation growth, reducing NDVI correlation, especially in Yiba-Hali. Dams also altered groundwater recharge rates, affecting availability and reducing groundwater correlation, indicating different recharge and usage patterns post-construction (Sherif et al., 2023). For instance, Fig. 11 highlights a clear decline in natural vegetation cover in the Hali basin after dam construction, consistent with findings by Almalki et al. (2023), which attribute this reduction to altered hydrological dynamics and reduced sediment deposition downstream. In contrast, the Baish basin shows an increase in vegetation cover post-dam, likely due to improved water retention and sedimentation processes. These contrasting patterns underscore the basin-specific environmental responses to dam construction and emphasize the complex interplay between hydrology, sedimentation, and vegetation dynamics in arid regions.

5.2. Application of DTW in assessing hydrological and environmental changes

DTW is crucial in this research for comparing spatial data on environmental variables before and after dam construction. One significant finding is the pronounced impact of dam construction across different basins, as revealed by DTW analysis using space-for-time. It highlighted substantial changes in environmental variables post-construction, indicating significant shifts in water flow, vegetation dynamics, soil composition, and groundwater levels, as noted in several studies (Al-Robai et al., 2018; Al-Sodany et al., 2015; Almalki et al., 2023; Sallam et al., 2018). For example, the DTW values for groundwater and runoff water between Reem and Baish show very low similarity, highlighting severe groundwater and runoff water dynamics disruption due to the Baish dam. Similarly, in the Yiba-Hali comparison, the dam also impacted groundwater and runoff water conditions in Hali, with DTW values indicating a disruption in groundwater and runoff water dynamics. For Yiba and Hali, the DTW values for NDVI, soil salinity, and soil moisture increased, suggesting notable changes in these variables, likely due to the dam's impact. However, these values are not as drastically different as those between Reem and Baish. In addition, in arid regions near the sea, local hydrological cycles are influenced by the sea's proximity, which provides additional moisture through sea breezes and evaporation. This influence, combined with sparse vegetation, contributes to precipitation recycling (Dominguez et al., 2006, 2008). Despite limited overall precipitation recycling in arid areas, the sea's steady moisture supply buffers against changes in upstream water availability. Consequently, dam impacts on downstream water flow may not significantly alter local climate conditions, such as precipitation and temperature, due to the mitigating effects of marine moisture (Peixoto and Oort, 1992). The high DTW value for runoff water post-dam indicates significant changes in runoff patterns, with NDVI being most affected, followed by soil salinity and runoff, reflecting the notable post-dam decline in NDVI correlations.

Furthermore, a comparative DTW analysis was conducted between Yiba and Reem, where no dams existed during the two periods (2003–2009 and 2010–2020). In 2003–2009, Reem and Yiba showed DTW values of 0.11 for NDVI and 0.12 for soil salinity, indicating closer vegetation and soil salinity patterns compared to Yiba-Hali (0.08 and 0.09) and Reem-Baish (0.16 and 0.15). Groundwater dynamics had a DTW value of 0.25 for Reem-Yiba and 0.52 for Reem-Baish, suggesting greater initial similarity in groundwater patterns between Reem and Yiba than Reem and Baish. In 2010–2020, DTW values between Yiba and Reem remained relatively unchanged (NDVI: 0.12, soil salinity: 0.13, groundwater: 0.27), reflecting high similarity. In contrast, Yiba-Hali and Reem-Baish exhibited higher DTW values for NDVI and soil salinity (between 0.25 and 0.40), indicating significant environmental changes due to dam construction. Groundwater DTW values differed notably (Yiba-Hali: 0.46, Reem-Baish: 1.19), highlighting the dam's significant impact on Baish's groundwater dynamics. However, these differences also reflect the influence of human activities and groundwater extraction in arid regions (Almalki et al., 2023; Alyami et al., 2022; Khan et al., 2023). In addition, dams not only affect hydrological processes like river runoff and groundwater recharge but also alter landscape dynamics, including soil properties (e.g., salinity) and vegetation cover (NDVI) (Zhao et al., 2012). Changes in water availability influence soil moisture and runoff, while altered landscapes can reduce soil permeability, affecting groundwater recharge (Han et al., 2017). These changes in vegetation and soil salinity further impact ecosystems.

The interannual variability of precipitation and temperature was analyzed across the dammed basins (Hali and Baish) and their corresponding proxy basins (Yiba and Reem) from 2003 to 2020. The average DTW values for precipitation were 0.08 between Yiba and Hali and 0.14 between Reem and Baish, indicating high similarity in precipitation patterns. Similarly, temperature showed even greater consistency, with average DTW values of 0.03 and 0.08 for Yiba-Hali and Reem-Baish, respectively. The trend analysis revealed negligible changes in precipitation, with slopes of 0.0003 (Yiba), 0.0002 (Hali), 0.0004 (Baish), and 0.0005 (Reem), while temperature trends were nearly flat, with slopes of 0.0002 for Yiba and Hali and 0.0000 for both Baish and Reem. These results indicate that climatic factors such as precipitation and temperature remained stable across the study period and were consistent across dammed and proxy basins. Consequently, the changes observed in environmental variables such as NDVI, soil salinity, and groundwater are unlikely to be driven by interannual climatic variability and are more likely a result of dam-induced impacts. This highlights the role of dam construction in altering downstream environmental dynamics, independent of natural climatic trends.

5.3. Linear regression-based machine learning for assessing hydrological and environmental changes

Regression-based machine learning has provided valuable insights into the environmental impacts of dam construction across various basins and periods. Key findings show significant changes in environmental variables, especially vegetation patterns, with NDVI changes indicating altered water availability and distribution. Variations in soil salinity, groundwater, and soil moisture levels also affect irrigation practices and water flow patterns post-dam construction. The R^2 values for NDVI, groundwater, soil moisture, and soil salinity decreased significantly in the Hali and Baish basins post-dam construction, reflecting the increased complexity and variability of these variables due to altered water availability and sediment flow. The increase in RMSE for NDVI and soil salinity in 2010–2020 suggests reduced prediction accuracy, reflecting significant changes in vegetation and soil composition. Similar trends in groundwater and soil moisture support the idea that dam construction disrupts natural groundwater recharge and soil moisture dynamics. These findings align with studies documenting the impact of dams on hydrological and environmental systems (Al-Robai et al., 2018; Nilsson and Berggren, 2000; Sallam et al., 2018). These changes resulted in lower R^2, higher RMSE, and MAE for NDVI and soil salinity, indicating reduced model accuracy (Yang et al., 2019), with more pronounced effects on groundwater and soil moisture, which are directly related to water availability, the main variable affected by the dam. Studies show that dams significantly impact soil moisture and groundwater, especially in arid areas. Due to reduced water flow, Vale et al. (2015) observed a notable decrease in soil moisture three years after the Amador Aguiar dam was built, particularly in the dry season. Al-Munqedhi et al. (2022) also reported significant drying downstream of dams, underscoring the challenges of water shortages. The high R^2 and correlation values, along with moderate RMSE, indicate a significant impact of the dam on water-related variables.

Further analysis was done for Yiba and Reem, where no dams exist and showed that in the first period (2003–2009), Yiba-Hali generally showed higher R^2 values for NDVI (0.73), soil salinity (0.64), and groundwater (0.98) compared to Reem-Yiba (0.67, 0.62, and 0.89, respectively), indicating better predictive accuracy in capturing environmental variations before significant dam impacts. This difference can be attributed to the geographical proximity of the Yiba and Hali basins, which likely share more similar environmental conditions. However, in the second period (2010–2020), Yiba-Hali experienced significant declines in the predictive accuracy of R^2 for NDVI (0.32) and soil salinity (0.23), while Reem-Yiba maintained relatively stable values (0.68 and 0.61), suggesting a more consistent environmental response in the second period in Reem-Yiba. Conversely, Reem-Baish shows consistently lower R^2 values for groundwater (0.33) than Reem-Yiba (0.89), indicating greater variability likely due to the dam's impact on Baish's hydrological systems. These findings highlight the basin-specific effects of dam construction and the importance of machine learning in quantifying these impacts for sustainable development.

5.4. Geological context and sedimentation implications

The geological maps as in Fig. 12 highlight that the Hali and Yiba basins (map A) share the same parent bedrock type, while the Reem and Baish basins (map B) predominantly consist of alluvial and related superficial deposits (Al-Sayari and Zötl, 2012; Al-Washmi et al., 2005). These similarities in geological composition provide a robust basis for selecting Yiba and Reem as proxy basins for Hali and Baish, respectively, ensuring consistent baseline conditions for comparative analysis. Geological characteristics play a critical role in sedimentation dynamics, particularly in arid regions. The shared bedrock type in the Hali-Yiba pair implies that sedimentation patterns, including particle size and mineral composition, are likely comparable. This is important for understanding soil texture, water retention, and vegetation establishment. Similarly, the uniform alluvial deposits in the Reem-Baish pair suggest similar sediment transport and deposition processes. The presence of alluvial deposits often indicates regions where fine sediments settle, influencing soil fertility and water infiltration. Including sedimentation in the analysis strengthens the attribution of environmental changes to dam construction. Dams typically alter sediment transport by trapping sediments upstream, leading to reduced sediment deposition downstream (Kummu and Varis, 2007). In the Baish and Hali basins, this could result in downstream soil degradation, reduced vegetation growth, and altered hydrological behavior. DTW analysis of the OCS revealed increasing divergence in sediment dynamics between the proxy and dam-impacted basins over time. These changes highlight the impact of dams on sediment transport and deposition, with reduced downstream sedimentation likely altering soil fertility, water retention, and vegetation patterns. The sediment trapping effect of dams in the Baish and Hali basins may have resulted in downstream soil degradation and reduced vegetation growth, as reflected in the DTW analysis, while the proxy basins (Yiba and Reem) maintained more stable sediment dynamics.

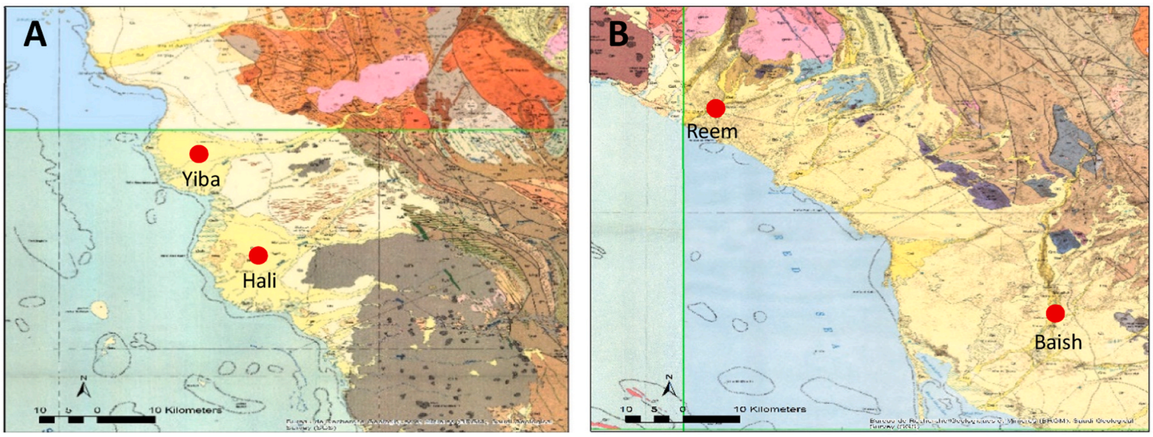


Fig. 12. Geological map A shows that the Hali and Yiba basins share the same parent bedrock type, while map B illustrates that the Reem and Baish basins are primarily composed of alluvial and related superficial deposits. Obtained from National Library of Australia.

Including sedimentation in the analysis strengthens the attribution of environmental changes to dam construction. The observed divergence in OCS dynamics further underscores the role of dams in disrupting natural sediment transport processes. Comparing these effects with their proxy basins provides a clearer understanding of the extent to which dams influence sedimentation, soil properties, and hydrological behavior.

5.5. Future directions and limitations

To mitigate the impacts of dam construction on NDVI, groundwater, and runoff, several strategies can be implemented. Adjusting water release schedules to restore natural flow patterns and improve groundwater recharge is crucial. Additionally, integrating water resource management practices that balance ecological needs and water storage can help. Soil conservation techniques and planting drought-resistant vegetation can reduce soil salinity and improve downstream vegetation. Artificial recharge methods, such as managed aquifer recharge, could also support groundwater restoration. Lastly, adopting efficient irrigation and water conservation practices would help alleviate pressure on water resources. This study offers valuable insights for environmental management and water resource planning, particularly in arid regions. It provides evidence of the long-term impacts of dam construction, helping policymakers balance water development with ecosystem. Using advanced techniques such as remote sensing and DTW improves environmental monitoring by allowing for early detection of degradation. However, this study has some limitations. A limitation of this study is that vegetation distribution, type, and seed dispersal mechanisms were not explicitly analyzed, despite their critical role in shaping vegetation dynamics and responses to hydrological and sedimentation changes. These factors should be considered in future studies to provide a more comprehensive understanding of dam impacts on arid region ecosystems. In addition, it relies on remote sensing data, which may have issues with spatial resolution and accuracy, especially when capturing detailed variations in environmental variables. The study is also limited to specific basins, which might not fully represent the impacts seen in other arid regions with different dam characteristics. Future research should expand to include more basins and ground-truth data for validating remote sensing. Studying long-term dam impacts on ecology and incorporating socio-economic factors could provide a more comprehensive view, while combining ecological, hydrological, and socio-economic data would improve environmental predictions and support better management practices.

6. Conclusion

The study demonstrates that dam construction in arid regions leads to significant environmental transformations, particularly in vegetation, groundwater, soil salinity, and runoff dynamics. Using a combination of space-for-time substitution, remote sensing, and machine learning techniques, the research revealed substantial changes in key environmental variables, with increased non-linearity in post-dam conditions. For example, DTW values for groundwater in Reem-Baish increased substantially from 0.52 to 1.19, and runoff DTW rose from 0.53 to 1.33, indicating notable disruptions in hydrological systems. Similarly, the decline in regression model predictive accuracy was particularly evident, with R^2 for groundwater dropping sharply from 0.98 to 0.34, reflecting the significant disruption of natural water flow and storage processes. The findings highlighted basin-specific disruptions, particularly in groundwater and runoff, underscoring the localized hydrological impacts of dams. This research provides valuable insights for understanding the long-term environmental consequences of dam infrastructure, especially in data-limited regions. Our use of remote sensing allowed us to overcome data limitations, providing a robust framework for assessing long-term environmental trends through the space-for-time substitution technique. This innovative approach is particularly valuable in regions like Yiba, Hali, Reem, and Baish, where ground observations are limited. The study demonstrates the power of combining advanced analytical tools, such as DTW and regression-based models, to capture subtle environmental shifts and assess the extent of dam impacts on local ecosystems. By integrating these methods, the study contributes to a better understanding of dam-induced environmental shifts, with implications for sustainable water resource management and ecosystem conservation.

Funding

This research received no external funding.

CRediT authorship contribution statement

Rodriguez Jose F.: Writing – review & editing, Validation, Supervision, Investigation, Conceptualization. **Saco Patricia M.:** Writing – review & editing, Validation, Supervision, Investigation, Conceptualization. **Khaki Mehdi:** Writing – review & editing, Validation, Supervision, Investigation, Conceptualization. **Almalki Raid:** Writing – original draft, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

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

R.A. would like to express my sincere gratitude to Umm Al-Qura University for its encouragement.

Data availability

The authors do not have permission to share data.

References

Abd El Shafy, M., Mostafa, A., 2021. Flash Flood Modeling Using HEC-RAS (2D) model on Wadi Reem in the western region, Kingdom of Saudi Arabia. J. Egyptian Acad. Soc. Environ. Develop. D., Environ. Stud. 22 (1), 17–32.
Afzal, J., Yihong, Z., Qayum, M., Afzal, U., Aslam, M., 2023. Effects of dam on temperature, humidity and precipitation of surrounding area: a case study of Gomal Zam Dam in Pakistan. Environ. Sci. Pollut. Res. 30 (6), 14592–14603.
Alarifi, S.S., Abdelkareem, M., Abdalla, F., Alotaibi, M., 2022. Flash flood hazard mapping using remote sensing and GIS techniques in Southwestern Saudi Arabia. Sustainability 14 (21), 14145.
Allbed, A., Kumar, L., Aldakheel, Y.Y., 2014. Assessing soil salinity using soil salinity and vegetation indices derived from IKONOS high-spatial resolution imageries: applications in a date palm dominated region. Geoderma 230, 1–8.
Allbed, A., Kumar, L., 2013. Soil salinity mapping and monitoring in arid and semi-arid regions using remote sensing technology: a review. Adv. Remote Sens. 2013.
Almalki, R., Khaki, M., Saco, P.M., Rodriguez, J.F., 2022. Monitoring and mapping vegetation cover changes in arid and semi-arid areas using remote sensing technology: a review. Remote Sens. 14 (20), 5143.
Almalki, R., Khaki, M., Saco, P.M., Rodriguez, J.F., 2023. The impact of dam construction on downstream vegetation area in dry areas using satellite remote sensing: a case study. Remote Sens. 15 (21), 5252.
Almalki, R., Khaki, M., Saco, P.M., Rodriguez, J.F., 2025. Understanding environmental factors influencing vegetation cover downstream of dams. Int. J. Environ. Res. 19 (1), 1–25.
Al-Munqedhi, B.M., El-Sheikh, M.A., Alfarhan, A.H., Alkahtani, A.M., Arif, I.A., Rajagopal, R., Alharthi, S.T., 2022. Climate change and hydrological regime in arid lands: impacts of dams on the plant diversity, vegetation structure and soil in Saudi Arabia. Saudi J. Biol. Sci. 29 (5), 3194–3206.
Al-Robai, Mohamed, H.A., Ahmed, A.A., Alsherif, E.A., 2018. Changes in vegetation structure and soil components at upstream and downstream areas of alaqiq dam, Southwestern Saudi Arabia [Article]. Pak. J. Bot. 50 (3), 1135–1145.
Al-Saeedi, A.H., 2022. Characterizing physical and hydraulic properties of soils in Al-Ahsa, Kingdom of Saudi Arabia. Saudi J. Biol. Sci. 29 (5), 3390–3402.
Al-Sayari, S.S., Zötl, J.G., 2012. Quaternary period in Saudi Arabia: 1: Sedimentological, hydrogeological, hydrochemical, geomorphological, and climatological investigations in central and eastern Saudi Arabia. Springer Science & Business Media.
Alshehri, F., Mohamed, A., 2023b. Investigation of groundwater potential using gravity data in Wadi Fatimah and its surroundings, Western Saudi Arabia. Front. Earth Sci. 11, 1225992.
Alshehri, F., Mohamed, A., 2023a. Analysis of groundwater storage fluctuations using GRACE and remote sensing data in Wadi As-Sirhan, Northern Saudi Arabia. Water 15 (2), 282.
Al-Sodany, Y.M., Fadl, M.A., Farrag, H.F., Saif, T.Y., 2015. Effect of dams on the vegetation of arid regions. J. Environ. Sci. Water Res 4, 71–91.
Al-Turki, S. 1995. Water resources in Saudi Arabia with particular reference to Tihama Asir province Durham University.
Al-Washmi, H., Gheith, A., Nabhan, A., 2005. Geomorphological features, sediment distribution and transport along Ash Shuqayq-Al Huraydah coastal area, southern Red Sea, Saudi Arabia. Mar. Scienes 16 (1).
Alyami, S.H., Alqahtany, A., Ghanim, A.A., Elkhrachy, I., Alrawaf, T.I., Jamil, R., Aldossary, N.A., 2022. Water resources depletion and its consequences on agricultural activities in Najran Valley. Resources 11 (12), 122.
Arebu, B.A., Alamri, N., Elfeki, A., 2024. Evaluation of sediment transport in ephemeral streams: a case study in the Southwestern Saudi Arabia. Arab. J. Sci. Eng. 1–16.
Azari, B., Hassan, K., Pierce, J., Ebrahimi, S., 2022. Evaluation of machine learning methods application in temperature prediction. Environ. Eng. 8 (1), 1–12.
Barnes, E.M., Sudduth, K.A., Hummel, J.W., Lesch, S.M., Corwin, D.L., Yang, C., Daughtry, C.S., Bausch, W.C., 2003. Remote-and ground-based sensor techniques to map soil properties. Photogramm. Eng. Remote Sens. 69 (6), 619–630.
Beck, M.W., Claassen, A.H., Hundt, P.J., 2012. Environmental and livelihood impacts of dams: common lessons across development gradients that challenge sustainability. Int. J. River Basin Manag. 10 (1), 73–92.
Bhartendu. 2024. Linear Regression [Simplest Implementation] https://www.mathworks.com/matlabcentral/fileexchange/64930-linear-regression-simplest-implementation.
Burke, I.C., Lauenroth, W.K., Parton, W.J., 1997. Regional and temporal variation in net primary production and nitrogen mineralization in grasslands. Ecology 78 (5), 1330–1340.
Chicco, D., Warrens, M.J., Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. Peerj Comput. Sci. 7, e623.
Chikodzi, D., Mutowo, G., Makaudze, B., 2013. Impacts of dam construction on tree species diversity in semi-arid regions: the case of Ruti Dam in Zimbabwe. Greener J. Environ. Manag. Public Saf. 2 (1), 16–21.
Chowdhury, S., Al-Zahrani, M., 2014. Water quality change in dam reservoir and shallow aquifer: analysis on trend, seasonal variability and data reduction. Environ. Monit. Assess. 186, 6127–6143.
Costa, A., Salvidio, S., Penner, J., Basile, M., 2021. Time-for-space substitution in N-mixture models for estimating population trends: a simulation-based evaluation. Sci. Rep. 11 (1), 4581.
Damgaard, C., 2019. A critique of the space-for-time substitution practice in community ecology. Trends Ecol. Evol. 34 (5), 416–421.
Degu, A.M., Hossain, F., Niyogi, D., Pielke Sr, R., Shepherd, J.M., Voisin, N., Chronis, T., 2011. The influence of large dams on surrounding climate and precipitation patterns. Geophys. Res. Lett. 38 (4).
Dominguez, F., Kumar, P., Vivoni, E.R., 2008. Precipitation recycling variability and ecoclimatological stability—a study using NARR data. Part II: North American monsoon region. J. Clim. 21 (20), 5187–5203.
Dominguez, F., Kumar, P., Liang, X.-Z., Ting, M., 2006. Impact of atmospheric moisture storage on precipitation recycling. J. Clim. 19 (8), 1513–1530.
Dürrenmatt, D., Del Giudice, D., Rieckermann, J., 2013. Dynamic time warping improves sewer flow monitoring. Water Res. 47 (11), 3803–3816.
Ejaz, N., Khan, A.H., Saleem, M.W., Elfeki, A.M., Rahman, K.U., Hussain, S., Ullah, S., Shang, S., 2024. Multi-criteria decision-making techniques for groundwater potentiality mapping in arid regions: a case study of Wadi Yiba, Kingdom of Saudi Arabia. Groundw. Sustain. Dev., 101223
El Ghazali, F.E., Laftouhi, N.-E., Fekri, A., Randazzo, G., Benkirane, M., 2021. Enhancing the success of new dams implantation under semi-arid climate, based on a multicriteria analysis approach: case of Marrakech region (Central Morocco). Open Geosci. 13 (1), 1494–1508.
ElKashouty, M., Mohy, M., Aziz, A.A.A., 2022. Hydrogeochemical characteristics of the aquifer in southern Assir, southwest Saudi Arabia. Arab. J. Geosci. 15 (1), 73.
Equere, V., Mirzaei, P.A., Riffat, S., Wang, Y., 2021. Integration of topological aspect of city terrains to predict the spatial distribution of urban heat island using GIS and ANN. Sustain. Cities Soc. 69, 102825.
Felty, T. 2024. Dynamic Time Warping. Retrieved May 29, 2024 from 〈https://www.mathworks.com/matlabcentral/fileexchange/6516-dynamic-time-warping〉.

Folgado, D., Barandas, M., Matias, R., Martins, R., Carvalho, M., Gamboa, H., 2018. Time alignment measurement for time series. Pattern Recognit. 81, 268–279.
Ghojogh, B., Crowley, M., 2019. The theory behind overfitting, cross validation, regularization, bagging, and boosting: tutorial. arXiv Prepr. *arXiv:1905.12787*.
Han, D., Currell, M.J., Cao, G., Hall, B., 2017. Alterations to groundwater recharge due to anthropogenic landscape change. J. Hydrol. 554, 545–557.
Hasanean, H., Almazroui, M., 2015. Rainfall: features and variations over Saudi Arabia, a review. Climate 3 (3), 578–626.
Horrocks, C., Newsham, K., Cox, F., Garnett, M., Robinson, C., Dungait, J., 2020. Predicting climate change impacts on maritime Antarctic soils: a space-for-time substitution study. Soil Biol. Biochem. 141, 107682.
Hu, J., Ma, F., Wu, S., 2018. Comprehensive investigation of leakage problems for concrete gravity dams with penetrating cracks based on detection and monitoring data: a case study. Struct. Control Health Monit. 25 (4), e2127.
Hussein, E.A., Thron, C., Ghaziasgar, M., Bagula, A., Vaccari, M., 2020. Groundwater prediction using machine-learning tools. Algorithms 13 (11), 300.
Jafari, R., Hasheminasab, S., 2017. Assessing the effects of dam building on land degradation in central Iran with Landsat LST and LULC time series. Environ. Monit. Assess. 189 (2), 1–15.
Jeong, Y.-S., Jeong, M.K., Omitaomu, O.A., 2011. Weighted dynamic time warping for time series classification. Pattern Recognit. 44 (9), 2231–2240.
Ji, X., Kang, E., Chen, R., Zhao, W., Zhang, Z., Jin, B., 2006. The impact of the development of water resources on environment in arid inland river basins of Hexi region, Northwestern China. Environ. Geol. 50, 793–801.
Kelly, B.F., Timms, W., Andersen, M., McCallum, A., Blakers, R., Smith, R., Rau, G., Badenhop, A., Ludowici, K., Acworth, R., 2013. Aquifer heterogeneity and response time: the challenge for groundwater management. Crop Pasture Sci. 64 (12), 1141–1154.
Khan, M.Y.A., ElKashouty, M., Abdellattif, A., Egbueri, J.C., Taha, A.I., Al Deep, M., Shaaban, F., 2023. Influence of natural and anthropogenic factors on the hydrogeology and hydrogeochemistry of Wadi Itwad Aquifer, Saudi Arabia: assessment using multivariate statistics and PMWIN simulation. Ecol. Indic. 151, 110287.
Khawfany, A., Basaham, A.S., & Gheith, A.M. 2009. Geomorphology, sedimentology, mineralogy and geochemistry of AL-Lith coast, Red Sea, Saudi Arabia. Unpublished M. Sc. thesis submitted to King Abdulaziz University, Jeddah, Saudi Arabia.
Kummu, M., Varis, O., 2007. Sediment-related impacts due to upstream reservoir trapping, the Lower Mekong River. Geomorphology 85 (3-4), 275–293.
Liu, Y., Zhang, Y.-A., Zeng, M., Zhao, J., 2024. A novel distance measure based on dynamic time warping to improve time series classification. Inf. Sci. 656, 119921.
Luo, Y., Zhou, Q., Peng, D., Yan, W., Zhao, M., 2023. Key influence of hydrogeological, geochemical, and geological structure factors on runoff characteristics in karst catchments. J. Hydrol. 623, 129852.
van Maren, D.S., Yang, S.-L., He, Q., 2013. The impact of silt trapping in large reservoirs on downstream morphology: the Yangtze River. Ocean Dyn. 63, 691–707.
Menard, S., 2001. Applied logistic regression analysis. SAGE publications.
Meng, L., Zhou, S., Zhang, H., Bi, X., 2016. Estimating soil salinity in different landscapes of the Yellow River Delta through Landsat OLI/TIRS and ETM+ Data. J. Coast. Conserv. 20, 271–279.
Miyanishi, K., Johnson, E.A., 2021. Coastal dune succession and the reality of dune processes. In Plant disturbance ecology. Elsevier, pp. 253–290.
Nilsson, C., Berggren, K., 2000. Alterations of riparian ecosystems caused by river regulation: dam operations have caused global-scale ecological changes in riparian ecosystems. How to protect river environments and human needs of rivers remains one of the most important questions of our time. BioScience 50 (9), 783–792.
Oregi, I., Pérez, A., Del Ser, J., & Lozano, J.A. 2017. On-line dynamic time warping for streaming time series. Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2017, Skopje, Macedonia, September 18–22, 2017, Proceedings, Part II 10,
Parr, T.W., Sier, A.R., Battarbee, R., Mackay, A., Burgess, J., 2003. Detecting environmental change: science and society—perspectives on long-term research and monitoring in the 21st century. Sci. Total Environ. 310 (1-3), 1–8.
Peixoto, J.P., Oort, A.H., 1992. Phys. Clim.
Petts, G.E., Gurnell, A.M., 2005. Dams and geomorphology: research progress and future directions. Geomorphology 71 (1-2), 27–47. https://doi.org/10.1016/j.geomorph.2004.02.015.
Peuquet, D.J., 2001. Making space for time: issues in space-time data representation. GeoInformatica 5, 11–32.
Pickett, S.T., 1989. Space-for-time substitution as an alternative to long-term studies. In Long-term studies in ecology: approaches and alternatives. Springer, pp. 110–135.
Radwan, F., Alazba, A., Mossad, A., 2017. Watershed morphometric analysis of Wadi Baish Dam catchment area using integrated GIS-based approach. Arab. J. Geosci. 10 (12), 1–11.
Rahmati, O., Choubin, B., Fathabadi, A., Coulon, F., Soltani, E., Shahabi, H., Mollaefar, E., Tiefenbacher, J., Cipullo, S., Ahmad, B.B., 2019. Predicting uncertainty of machine learning models for modelling nitrate pollution of groundwater using quantile regression and UNEEC methods. Sci. Total Environ. 688, 855–866.
Ramezan, A., Warner, T, C., A., Maxwell, A, E., 2019. Evaluation of sampling and cross-validation tuning strategies for regional-scale machine learning classification. Remote Sens. 11 (2), 185.
Rastetter, E.B., 1996. Validating models of ecosystem response to global change. BioScience 46 (3), 190–198.
Richter, B.D., Postel, S., Revenga, C., Scudder, T., Lehner, B., Churchill, A., Chow, M., 2010. Lost in development's shadow: the downstream human consequences of dams. Water Altern. 3 (2), 14.
Romani, L.A., Goncalves, R., Zullo, J., Traina, C., Traina, A.J., 2010. New DTW-based method to similarity search in sugar cane regions represented by climate and remote sensing time series. IEEE Int. Geosci. Remote Sens. Symp. 2010.
de Rooij, G., 2016. Subsurface flow of water in soils and geological formations. Oxf. Res. Encycl. Environ. Sci.
Sallam, A., Bader Alharbi, A., Usman, A.R., Hussain, Q., Ok, Y.S., Alshayaa, M., Al-Wabel, M., 2018. Environmental consequences of dam construction: a case study from Saudi Arabia. Arab. J. Geosci. 11 (3), 1–12.
Schielzeth, H., 2010. Simple means to improve the interpretability of regression coefficients. Methods Ecol. Evol. 1 (2), 103–113.
Seidl, R., Albrich, K., Thom, D., Rammer, W., 2018. Harnessing landscape heterogeneity for managing future disturbance risks in forest ecosystems. J. Environ. Manag. 209, 46–56.
Şen, Z., Khiyami, H.A., Al-Harthy, S.G., Al-Ammawi, F.A., Al-Balkhi, A.B., Al-Zahrani, M.I., Al-Hawsawy, H.M., 2013. Flash flood inundation map preparation for wadis in arid regions. Arab. J. Geosci. 6 (9), 3563–3572.
Shahin, M., Shahin, M., 2007. Wadis and wadi flow. Water Resour. Hydrometeorol. Arab Reg. 279–332.
Sherif, M., Sefelnasr, A., Al Rashed, M., Alshamsi, D., Zaidi, F.K., Alghafli, K., Baig, F., Al-Turbak, A., Alfaifi, H., Loni, O.A., 2023. A review of managed aquifer recharge potential in the Middle East and North Africa Region with examples from the Kingdom of Saudi Arabia and the United Arab Emirates. Water 15 (4), 742.
Siehoff, S., Lennartz, G., Heilburg, I.C., Roß-Nickoll, M., Ratte, H.T., Preuss, T.G., 2011. Process-based modeling of grassland dynamics built on ecological indicator values for land use. Ecol. Model. 222 (23-24), 3854–3868.
Stojanova, D., Panov, P., Gjorgjioski, V., Kobler, A., Džeroski, S., 2010. Estimating vegetation height and canopy cover from remotely sensed data with machine learning. Ecol. Inform. 5 (4), 256–266.
Strobl, E., Strobl, R.O., 2011. The distributional impact of large dams: evidence from cropland productivity in Africa. J. Dev. Econ. 96 (2), 432–450.
Tavakol, A., McDonough, K.R., Rahmani, V., Hutchinson, S.L., Hutchinson, J.S., 2021. The soil moisture data bank: the ground-based, model-based, and satellite-based soil moisture data. Remote Sens. Appl.: Soc. Environ. 24, 100649.
Thakur, J.K., Singh, S.K., Ekanthalu, V.S., 2017. Integrating remote sensing, geographic information systems and global positioning system techniques with hydrological modeling. Appl. Water Sci. 7 (4), 1595–1608.
Ugbaje, S.U., Karunaratne, S., Bishop, T., Gregory, L., Searle, R., Coelli, K., Farrell, M., 2024. Space-time mapping of soil organic carbon stock and its local drivers: potential for use in carbon accounting. Geoderma 441, 116771.
Vale, Schiavini, I., Araújo, G., Gusson, A., Lopes, S., Oliveiral, A., Prado-Júnior, J., Arantes, C., Dia-Neto, O., 2015. Effects of reduced water flow in a riparian forest community: a conservation approach. J. Trop. For. Sci. 13–24.
Willmott, C.J., Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim. Res. 30 (1), 79–82.

Wu, W., Zucca, C., Muhaimeed, A.S., Al-Shafie, W.M., Fadhil Al-Quraishi, A.M., Nangia, V., Zhu, M., Liu, G., 2018. Soil salinity prediction and mapping by machine learning regression in C entral M esopotamia, I raq. Land Degrad. Dev. 29 (11), 4005–4014.
Yang, M., Xu, D., Chen, S., Li, H., Shi, Z., 2019. Evaluation of machine learning approaches to predict soil organic matter and pH using Vis-NIR spectra. Sensors 19 (2), 263.
Yaseen, Z.M., Al-Juboori, A.M., Beyaztas, U., Al-Ansari, N., Chau, K.-W., Qi, C., Ali, M., Salih, S.Q., Shahid, S., 2020. Prediction of evaporation in arid and semi-arid regions: a comparative study using different machine learning models. Eng. Appl. Comput. Fluid Mech. 14 (1), 70–89.
Yoon, H., 2021. Finding unexpected test accuracy by cross validation in machine learning. Int. J. Comput. Sci. Netw. Secur. 21 (12spc), 549–555.
You, Y., Li, Z., Gao, P., Hu, T., 2022. Impacts of dams and land-use changes on hydromorphology of braided channels in the Lhasa River of the Qinghai-Tibet Plateau, China. Int. J. Sediment Res. 37 (2), 214–228.
Yue, C., Kahle, H.-P., Klädtke, J., Kohnle, U., 2023. Forest stand-by-environment interaction invalidates the use of space-for-time substitution for site index modeling under climate change. For. Ecol. Manag. 527, 120621.
Zhao, Q., Liu, S., Deng, L., Dong, S., Yang, Z., Yang, J., 2012. Landscape change and hydrologic alteration associated with dam construction. Int. J. Appl. Earth Obs. Geoinf. 16, 17–26.
Zhou, J., Zhao, Y., Huang, P., Zhao, X., Feng, W., Li, Q., Xue, D., Dou, J., Shi, W., Wei, W., 2020. Impacts of ecological restoration projects on the ecosystem carbon storage of inland river basin in arid area, China. Ecol. Indic. 118, 106803.