Kirkuk University Journal for Agricultural Sciences
ISSN:2958-6585
https://kujas.uokirkuk.edu.iq
RESEARCH ARTICLE https://doi.org. 10.58928/ku25.16430

Wetland Areas Trend And Examining Effective Factors With Machine Learning

Srwsht Mohammed jaza[1] **Tarza Jabar Baqi[2]**
Hiwa Farajpanah[3]
[1]*Surveying Department,Darbandikhan technical institute,sulaimani polytechnic university,sulaimani,KGR,IRAQ.*
[2]*Surveying Department,Darbandikhan technical institute,sulaimani polytechnic university,sulaimani,KGR,IRAQ.*
[3]*Civil Engineering and Architecture Faculty, Shahid Chamran University of Ahvaz, Ahvaz – IRAN.*
*Corresponding Author: srwsht.m.abdulla@spu.edu.iq.

Received:01/10/2025 **Revised: 20/11/2025** **Accepted: 01/12/2025** **Published: 22/12/2025**

ABSTRACT
Wetlands are special ecosystems providing crucial hydrological, ecological, and socio-economic services. The study investigates the long-term development of the Hammar Marsh in Iraq from 2000 to 2025, focusing on water level trends and the driving environmental forces of the changes. Remote sensing imagery is analyzed using Google Earth Engine to obtain monthly water surface areas and other key climatic and ecological variables. Mann–Kendall test and Sen's slope estimator were applied to detect significant trends in water level, and there was an overall increase, with summer and autumn being particularly so, while winter and early spring had slower changes. Stepwise Variance Inflation Factor (VIF) analysis was performed to reduce multicollinearity among predictors so that all remaining variables had VIF values below 10. A Random Forest model was then executed to infer the relative importance of environmental drivers. The model exhibited test set R^2 of 0.690 and RMSE of 0.154, indicating good predictability. Calculation of the variable importance indicated that the Palmer Drought Severity Index (PDSI) and soil moisture were the dominant controlling factors of water level change, followed by vegetation cover (NDVI) and land surface temperature (LST), with the other variables of precipitation, vapor pressure, wind speed, runoff, and aerosol optical depth having secondary effects.
The results highlight the synergistic effects of climatic and hydrological drivers on wetland dynamics and demonstrate the effectiveness of integrating remote sensing, trend analysis, and machine learning for wetland monitoring. The outcomes of this study have significant implications for the sustainable management and conservation of Hammar Marsh and other similar wetland ecosystems in the face of changing environmental conditions.

***Keywords*:** Wetlands, Remote Sensing, Hydro-climatic Drivers, machine learning, NDVI..

INTRODUCTION

Wetlands are among the most vital ecosystems on earth, providing a range of ecological, hydrological, and socio-economic benefits [1,2,3,4]. They play an important role in climatic regulation [5], conservation of biodiversity [6,7], ecosystem balance maintenance [8], water purification [9], carbon fixation [10,11], medicinal resource supply, and tourism. Though they cover just a fraction of the Earth's surface, wetlands are of unequal importance in supporting ecological equilibrium and human well-being [12]. They alone cover millions of square kilometers in Asia, and are among the most widespread and diverse ecosystems in the region [13].

Wetlands are increasingly facing threats from human activities and natural processes. Population growth, rapid urban development, and climate change impacts have triggered massive degradation and continuous loss of wetland cover [14,15,16,17].

Remote sensing, as such, has proven to be one of the most effective and efficient tools for monitoring and evaluating wetlands [18]. In addition to remote sensing, machine learning approaches have recently shown remarkable capabilities in analyzing complex hydrological and environmental systems. Several studies have employed hybrid and data-driven models to predict water level fluctuations, drought indices, and ecosystem responses with high accuracy [19,20,21]. For instance,[22] demonstrated that combining Random Forest and LSTM models with PCA and stepwise regression provided highly accurate predictions of river water quality parameters in arid regions, highlighting the robustness of hybrid ML approaches for hydrological and environmental modeling. Due to the fact that it provides consistent, historical, and large-scale information without spatial limits, remote sensing enables the accurate monitoring of significant wetland resources, like water, vegetation, and soil [23,24]. This technological approach, besides supporting improved understanding of wetland dynamics, also assists in supporting decision-making for the conservation, restoration, and sustainable management of these valuable ecosystems [25,26].

Wetlands have significantly evolved over the last decades as a result of climatic variability, terrestrial characteristics, and increasing human-induced stress [27]. Therefore, long-term and repetitive monitoring is key to establishing whether these ecosystems are augmenting, stable, or degrading, and to guiding conservation and restoration efforts. However, due to the complex interlinkage of wetlands with their surrounding environment drivers, the identification of the most important drivers that influence their dynamics becomes particularly essential for significantly sensitive or very degraded systems [28,29,30]. Many factors control the expansion or shrinkage of wetlands, including climatic conditions (rainfall, temperature, evapotranspiration, drought) [31], hydrological drivers (river discharge, groundwater level, catchment inflow) [32], and terrestrial drivers (soil moisture, vegetation cover, topography) [23]. In addition, anthropogenic drivers such as land use and land cover change, agricultural intensification, urbanization, and industrialization further aggravate wetland loss and fragmentation [33,34,35]. Quantifying and ascertaining the relative contribution of these natural and anthropogenic parameters are essential to rational conservation planning, sustainable management, and fostering adaptive capacity of wetlands in the face of persisting environmental change.

In research on trends in a wetland, various factors are considered. Some research focuses on ecological degradation, while others research water level trends, and others look at soil cover and vegetation cover. [36] estimated spatial and temporal wetland degradation trends in Jiangsu Province, China, from 1980 to 2020 using remote sensing data and a landscape directional succession model. Their research showed that 3,020.67 km² (42.74% of the overall coastal wetland coverage) of the wetlands deteriorated, with overall degradation being characterized as mild in character. Degradation was higher in Yancheng City, mostly covering Sheyang County, Dafeng District, Dongtai City, and Rudong County. The trend was cumulative, and the overall degradation score rose from 0.45 in 1985 to 1.67 in 2020, with a maximum in 2000. The major forms of degradation were conversion to construction land, fish farming, arable land, and invasion by exotics. The study focuses on imposing stricter controls over wetland development to ensure sustainable management and long-term conservation.[37] summarized recent advances in satellite remote sensing of wetland ecosystems in Sub-Saharan Africa, highlighting the fundamental role of wetlands as highly productive ecosystems that accommodate numerous plants and animals. While they are of significant ecohydrological importance, wetlands here are gravely threatened by global environmental change and anthropogenic pressures, particularly poor management practices leading to overexploitation and underutilization. The lack of regular monitoring and up-to-date spatial information has limited data on the wetland loss rates and effective management. The study brings out the fact that remote sensing is an effective means of accurate mapping, monitoring, and documentation of past and present wetland state. Likewise, other researchers have highlighted that integrating artificial intelligence and hybrid modeling techniques enhances the understanding of wetland degradation patterns and hydrological dynamics, offering complementary insights to remote sensing analyses [38,39]. Satellite observation delivers spatially explicit and temporal data, which enables better understanding of ecohydrological processes, wetland dynamics, and state of the environment as well as focusing on challenges and limitations involved with such an approach.

[40] analyzed wetland and watershed degradation in the Tabunio Watershed using multi-temporal remote sensing and a high-resolution land degradation index. The study monitored spatial and temporal land degradation from 2005 to 2020 based on a mixture of land use/land cover, vegetation coverage, soil erosion, and soil moisture content. The projected index was considerably precise (kappa > 0.73, overall accuracy > 86%) and provided a better assessment than isolated indices. Results showed an overall pattern towards higher degradation of land, with 2010 being the worst year and most of the other years having moderate degradation. Key drivers for degradation were reduced water uptake, flooding during the rainy season, dry season droughts, and impacts from conventional gold mining operations. The study emphasizes that advanced monitoring methods like the integrated land degradation index have important roles in sustainable management, early detection, and protection of valuable watershed areas with freshwater resources. [41] sought to characterize and monitor wetlands in Pakistan using satellite remote sensing and emphasizing important parameters like wetness, greenness, turbidity, temperature, and landscape changes. Despite wetlands receiving international recognition and value through initiatives like the Ramsar Convention, a majority of wetlands in less developed nations remain under-valued. Supervised classification and TCW index were employed by this research to classify wetlands, and change detection indices, TCG, and NDTI were utilized to detect water quality, ecological, and climate-related effects. 2016-2019 Sentinel-2 imagery, along with ASTER DEM and MODIS LST data, was applied for spatial and hydrological analyses and complemented by rainfall data from ANN databases. Small but notable

alterations in water fractions of large lakes such as Borith, Phander, Upper Kachura, Satpara, and Rama Lake were depicted in outcomes, indicating ongoing ecological transformations. The study emphasizes the strong need for proactive conservation measures to conserve wetlands and enhance ecosystem dynamics against environmental stress.[42] discussed multiscenario degradation in the Maidika Alpine Wetland Nature Reserve, Qinghai–Tibet Plateau, using Landsat time series data. The study developed an AW-CCD, a continuous degradation detection and classification algorithm that integrates spectral–temporal characterization, classification, and degradation detection to map alpine wetland dynamics. The method detected water-related landscape change, like snowmelt, lake and river contraction, and swampy meadow to alpine meadow transition with decreased soil wetness. Using spectral–temporal indices and seasonally varying soil wetness, AW-CCD enabled yearly wetland mapping and multiscenario degradation assessment. Results indicated better mapping precision (94.9% in 2022) and demonstrated spatial–temporal patterns of degradation across two decades: snow and river regions lost 5.04% and 16.74%, respectively, and 3.23% of swamp meadows were transformed into alpine meadows. Degradation was strongest prior to 2009, followed by relative stability up to 2015, and subsequent degradation. The study confirms the applicability of AW-CCD in the valuation of the complex responses of alpine wetlands to climate changes in high-mountain ecosystems.

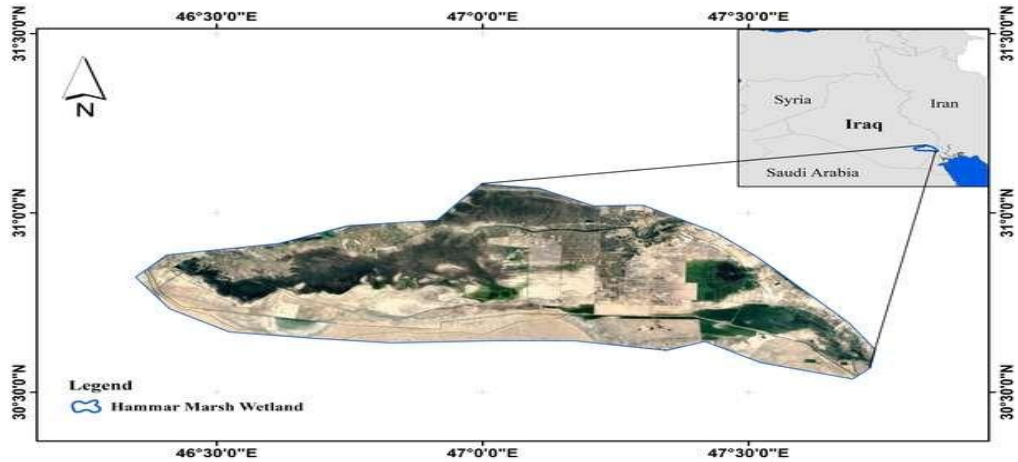
Investigation in this field has also been further conducted by other scholars, including: [43,44].

Based on the existing scientific literature, numerous studies have examined wetland degradation and dynamics from perspectives of hydrological change, vegetation cover change, and impacts of human activities. The primary goal of this present study, however, is to examine the trends of the water level of a selected wetland in Iraq, whether its levels rise or fall with time, and identify the leading environmental drivers of these trends. This study is run on a suite of climatic parameters in conjunction with advanced machine learning techniques in the form of the Random Forest algorithm to enable a firm and data-driven assessment of the drivers of wetland dynamics over time. This research endeavors to have a better understanding of processes that shape wetland activity as well as provide insight into sustainable management and conservation strategies.

Materials And Methods

2-1-Study area

The research location in this study is the Hammar Marsh, which is one of the three great marshes originally making up the extensive and diverse Mesopotamian marshes of south Iraq, along with the Central Marshes [45] and Hawizeh Marsh [46] (Figure 1). The three marshes are designated Ramsar Sites and are internationally conserved. Historically, Hammar Marsh covered up to 4,500 km² at seasonally elevated water. Enormous drainage in the early 1990s, together with the construction of roads, canals, and embankments, largely transformed the natural condition. Re-flooding and restoration of works in the 2000s partially recovered water levels and biodiversity. Flood control, water storage and filtration, and climate regulation services are provided by the marsh ecosystem. It is home to globally threatened species such as the endangered Basra reed warbler (Acrocephalus griseldis) and the marbled duck (Marmaronetta angustirostris), besides the vulnerable Mesopotamichthys sharpeyi and Euphrates softshell turtle (Rafetus euphraticus). Redbelly tilapia (Tilapia zillii), a non-native and invasive fish, is the most prominent fish species. Hammar Marsh is also of great historical and cultural significance, as the resources of the Marsh Arab indigenous community are being used for building special floating houses and perpetuation of their traditional way of life. The main economic activities are fishing, hunting, and agriculture, while in recent decades increasing oil exploration has been taking place. The ecological, cultural, and economic value of Hammar Marsh, as well as its Ramsar designation, highlights the significance of continued monitoring and sustainable use of this critical wetland



system.

Figure 1: Study area of the Hammar Marsh wetland

2-2-Data and Process

Within this research, all analysis and calculations were conducted on the Google Earth Engine (GEE) platform, offering access to an extensive array of satellite products alongside robust cloud-based geospatial processing features [47]. The employment of GEE enabled us to process large datasets of multi-temporal data with ease and maintain data consistency in handling and reproducibility of the findings. In order to examine the Hammar Marsh dynamics, the water surface area was approximated from MODIS MOD09A1 product with the help of Normalized Difference Water Index (NDWI) (eq1).

$$NDWI = \frac{Green - NIR}{Green + NIR} \qquad (Eq.1)$$

NIR and Green in this context denote near-infrared and green respectively [48]. Water bodies were also identified through positive NDWI, and their covered areas were approximated every month for each wetland.

Monthly water surface area values for the 2000 to 2025 time period were extracted, providing time-series perspective on hydrologic wetland modification. Concomitant with surface water monitoring, additional climatic and environmental variables were integrated to enhance knowledge of their contribution to forcing wetland fluctuations. Vegetation change was monitored in particular through use of the Normalized Difference Vegetation Index (NDVI, MOD13Q1), and thermal regimes were approximated from Land Surface Temperature (LST, MOD11A2). Climate-related drivers were precipitation (CHIRPS), aerosol optical depth (AOD, MCD19A2), and certain variables of the TERRACLIMATE dataset, namely actual evapotranspiration (AET), reference evapotranspiration (PET), wind speed (WS), vapor pressure (VAP), soil moisture, Palmer Drought Severity Index (PDSI), and runoff (R). These datasets as a whole constitute a rich basis for understanding hydrological and ecological variability of the study region on season as well as interannual timescales. The specifications in terms of detail of the datasets, including their sources, spatial and temporal resolution, and study duration, are listed in Table 1.

Table 1: Specifications of the sensors and datasets used

Variable	Sensor	Units	S R
NDWI(WA)	MOD09A1	----	500m
NDVI	MOD13Q1	----	250m
LST	MOD11A2	Kelvin	1000m
Precipitation (P)	CHIRPS	mm	5566m
Aerosol Optical Depth (AOD)	MCD19A2	----	1000m
Actual evapotranspiration (AET)	TERRACLIMATE	mm	4638m
Reference evapotranspiration (PET)	TERRACLIMATE	mm	4638m

Wind Speed (WS)	TERRACLIMATE	m/s	4638m
Vapor pressure (VAP)	TERRACLIMATE	kPa	4638m
Soil moisture	TERRACLIMATE	mm	4638m
PDSI	TERRACLIMATE	----	4638m
Runoff (R)	TERRACLIMATE	mm	4638m

2-3-Method

The method of research employed in this study consists of five principal stages (figure 2).

1- Preprocessing and temporal harmonization: All data were cleaned and resampled to monthly temporal frequency since each of the data sets was initially provided in several alternative temporal resolutions.

2- Data standardization: To ensure comparability between variables and reduce the impact of scale differences, all data sets were standardized prior to analysis.

3- Trend Analysis using the Mann–Kendall Test: The MK test was employed to detect monotonic trends in the time series of wetland water level. As a non-parametric test, it is not affected by non-normality and outliers and hence can be used for long-term hydrological data. MK test was utilized in this study work for monthly water level data (2000–2025) to determine the direction and statistical significance of change with time.

4- Multicollinearity test (VIF test): The Variance Inflation Factor (VIF) was applied to identify and remove highly correlated variables to ensure model stability and interpretability.

5- Machine learning modeling (Random Forest): The Random Forest algorithm was employed to simulate wetland water level behavior, evaluate model performance, and estimate the relative importance of every explanatory variable.

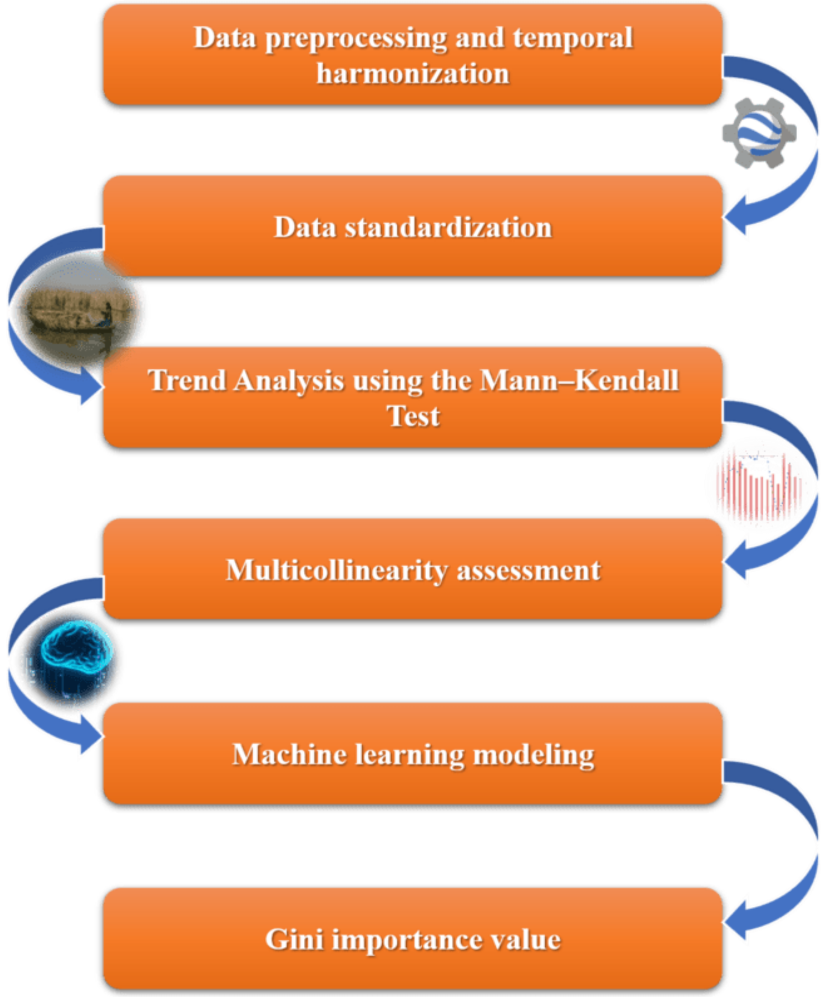


Figure 2: Flowchart of the methodology used in this study

2-3-1- Data preprocessing and temporal harmonization

All the preprocessing was performed within the Google Earth Engine (GEE) environment. Since the data sets applied in this study had different temporal resolution, they were resampled and reinitialized into their actual temporal resolutions in order to maintain the subsequent monthly analysis consistent.

2-3-2- Data standardization

Standardization is a significant step in the preprocessing process, especially when working with mixed variables that differ in scale and unit. This operation minimizes differential measurement range bias and ensures all variables contribute equally to the analysis [49]. Additionally, it facilitates improved collinearity tests and enhances the robustness of machine learning models. All variables in this study were standardized through Equation (2) to enable a fair comparison and integration of the datasets.

$$Z_i = \frac{X_{i-}\mu}{\sigma} \qquad (Eq.2)$$

In the formula above, Z_i is the standard score for data X_i, μ is the mean and σ is the standard deviation of the data. By doing this, the Z_i's will have a mean of 0 and a variance of 1.

2-3-3- Trend analysis

There are many statistical procedures which can be employed to analyze time series, but non-parametric methods are particularly effective to use when dealing with hydrological and meteorological data [50]. These methods have a number of benefits since they do not rely on the statistical distribution of the dataset and are effective for those series that exhibit skewness or unpredictable fluctuations [51]. Trend analysis in this case aims at determining whether a dataset has a consistent rise or decline trend over time. The Mann–Kendall test is one of the most common non-parametric tests for this purpose [52]. It is based on two hypotheses: under the null hypothesis, the data series is trend-free and random, and under the alternative hypothesis, the data series has a monotonic trend [53]. The methodology lies in the calculation of the S statistic, which tests the difference between all observations and all subsequent observations, as presented in Equation (3).

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \mathrm{sgn}(x_j - x_k) \qquad (Eq.3)$$

In this formula, n is the number of observations in the time series, and xj and xk are the j-th and k-th data points of the series, respectively. Then, the variance of S is calculated and the standardized Z statistics are calculated using equations (2) and (3):

$$\mathrm{VAR(S)} = \frac{1}{18}[n(n-1)(2n+5)] \qquad (Eq.4)$$

$$Z = \begin{cases} \dfrac{S-1}{\sqrt{\mathrm{VAR(S)}}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{\mathrm{VAR(S)}}} & \text{if} S < 0 \end{cases} \qquad (Eq.5)$$

In a two-sided trend analysis, the null hypothesis (H_0), which assumes that the data series has no trend, is retained if the condition $|Z| < Z\alpha/2$ is satisfied at a 95% confidence level. If this condition is not met, the alternative hypothesis (H_1), indicating the presence of a trend, is accepted. Positive Z values represent an increasing trend, while negative Z values indicate a decreasing trend in the dataset.

To estimate the trend rate in a time series, the Sen's Slope estimator is commonly used. This non-parametric method, originally proposed by Theil and later refined by Sen, calculates the slope based on the differences between all pairs of observations in the series. It is particularly suitable for detecting linear trends, where the trend value at time t can be expressed as:

$$f(t) = Qt + B \qquad (Eq.6)$$

Where Q is the slope of the trend line and B is the constant value.

2-3-4- Multicollinearity Analysis (VIF)

Collinearity is when an independent variable in a regression equation has a high correlation with one or more other independent variables and essentially becomes a composite of them. When there are interdependent predictors, it is called multicollinearity [54]. Multicollinearity reduces the reliability of the findings of a regression, as it is difficult to determine individual effect of each variable on the dependent variable. This usually causes excessive variances of estimates of coefficients and can lead to unstable or biased predictions, where slight variations in the data cause gigantic changes in the coefficients. In order to test and determine such relationships among variables, the Variance

Inflation Factor (VIF) is often employed, offering a measurement of multicollinearity, as seen in Equation (7) (Salmerón-Gómez et al., 2025).

$$VIF = \frac{1}{TC} = \frac{1}{(1-R_i^2)} \quad \text{(Eq.7)}$$

In this context, Ri^2 represents the unadjusted coefficient of determination obtained when the independent variable is regressed on all other independent variables. The tolerance coefficient (TC) is calculated as the inverse of the VIF. A small TC value (less than 0.2) indicates a strong correlation between independent variables, while values above 0.2 suggest that multicollinearity is not a significant concern. To address multicollinearity, a common approach is to exclude variables that show high correlation with others, which is the procedure applied in this study.

2-3-4- Random Forest (RF)

Random Forest (RF) is one of the well-known machine learning techniques for the estimation of variable importance and feature selection. RF performance is extremely sensitive to the number of trees constructed in the model [55]. Feature importance is typically estimated through the Gini Index, which operates well to capture the predictors with the highest contribution towards the model. A random subset of variables is selected at each decision node as potential candidates for the split [43]. The decrease in heterogeneity for each candidate variable is then computed, which measures how much the split improves node purity. Decreases in heterogeneity for all splits are accumulated over all nodes and averaged over all trees to receive the overall Gini index [56]. This process facilitates the robust ranking of the most influential variables such that the primary drivers are always identified in the model. Similar applications of the Random Forest algorithm and hybrid learning frameworks have been successfully used for hydrological prediction, groundwater level estimation, and water quality monitoring in previous environmental studies [21,39].

$$Gini\ index = \frac{1}{n}\sum_{z}[d(x,z) \cdot I(x,z)] \qquad (Eq.8)$$

Here, I(h,z) is a function that equals 1 if variable x-th variable is used for splitting at the node z, and equals 0 if it is not used.

Result

3-1- Results of Parameter Variations

In Figure 3, the monthly patterns of the parameters discussed in the methodology are displayed. By eye, the water surface appears to have an increasing trend, but in the following section, a more formal test based on the Mann–Kendall test will be employed to assess the trends formally.

3-2-Maan-Kendall

The Mann–Kendall test was employed to assess monthly and seasonal trends in the water surface of the wetland being studied. Results include Kendall's Tau, p-values, and Sen's slope estimates (in hectares) for each month and season. The overall analysis reveals a trend that is generally positive for most of the year, although the magnitude and significance of the trend vary across months and seasons. For monthly trends, the highest positive Kendall's Tau values were those of September (0.46, $p = 0.00098$), June (0.42, $p = 0.0028$), and July (0.393, $p = 0.0054$), indicating a statistically significant increase in water surface expansion in mid to late summer. Moderate positive trends were also observed in March (0.34, $p = 0.017$) and April (0.353, $p = 0.013$), suggesting that early spring months are also seeing water level rises (figure 4). December (Tau = 0.16, $p = 0.275$) and January (Tau = 0.213, $p = 0.142$) had weaker, non-significant trends, suggesting comparatively stable winter water levels. Sen's slope estimates quantify these changes, with the steepest monthly increases occurring in March (1,698 ha), April (1,694 ha), and February (1,556 ha), and the lowest increases during winter months of December and January (774 ha and 1,234 ha, respectively). Seasonally, the strongest positive trends were in Fall (Tau = 0.413, $p = 0.0034$) and Summer (Tau = 0.373, $p = 0.0085$), while Spring (Tau = 0.24, $p = 0.0975$) and Winter (Tau = 0.267, $p = 0.0646$) exhibited weaker and less significant increases. Sen's slopes also support this trend, as Fall and Summer present slopes of 924 ha and 1,222 ha, respectively, whereas Spring and Winter present slopes of 1,458 ha and 893 ha. It can be inferred that the wetland experiences the fastest water surface increases during the warmer seasons, whereas winter and early spring exhibit slower or more stable tendencies. In summary, the Mann–Kendall test and Sen's slope analysis combined detect a significant positive trend in the water surface of the wetland over the period of study with seasonal variation in more pronounced summer and fall increments. The findings constitute the foundation for investigation of the climatic and environmental causes of these trends and guide the next phase machine learning-based assessment of principal controlling factors.

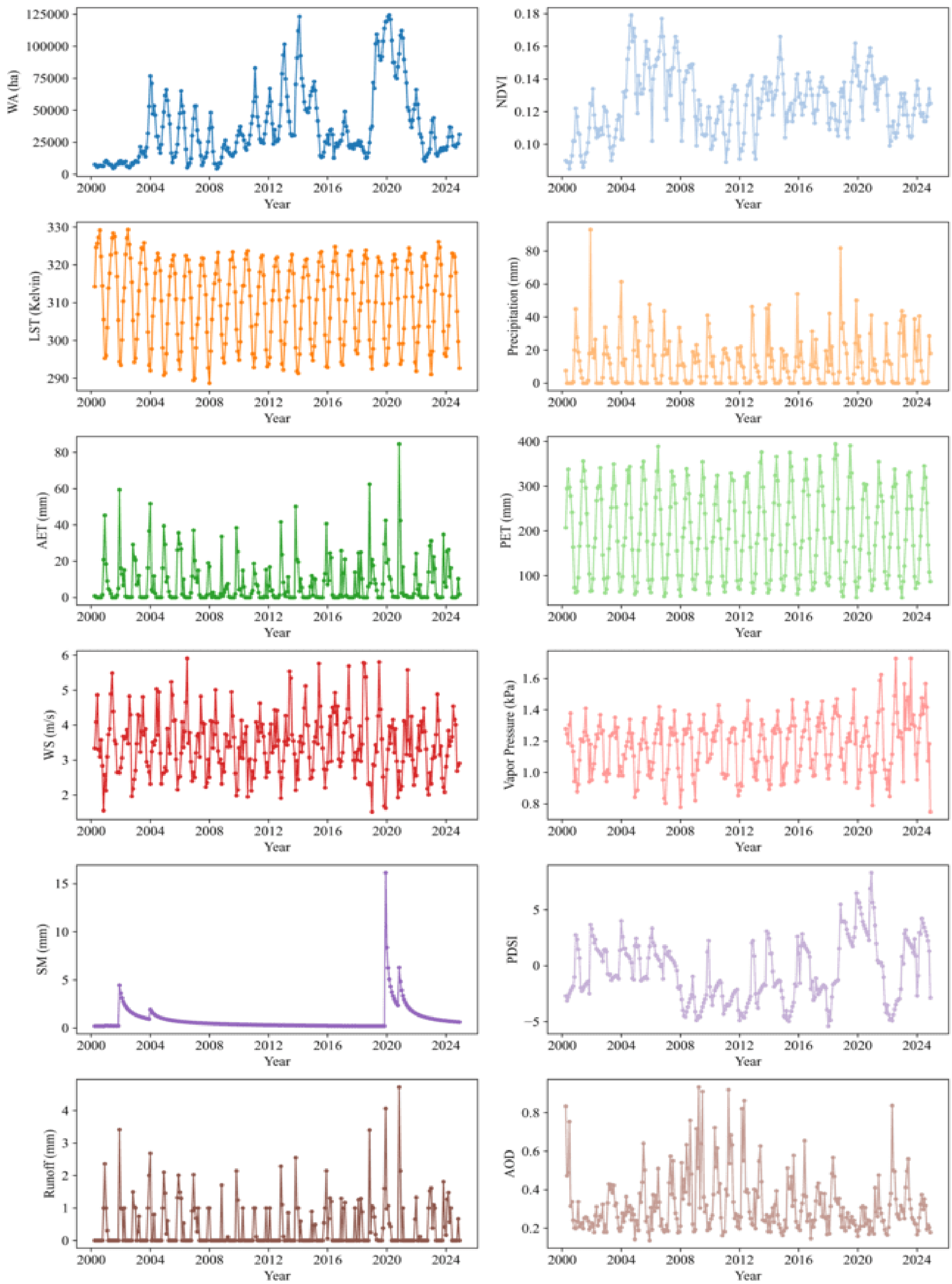


Figure 3: Monthly variations of the different variables

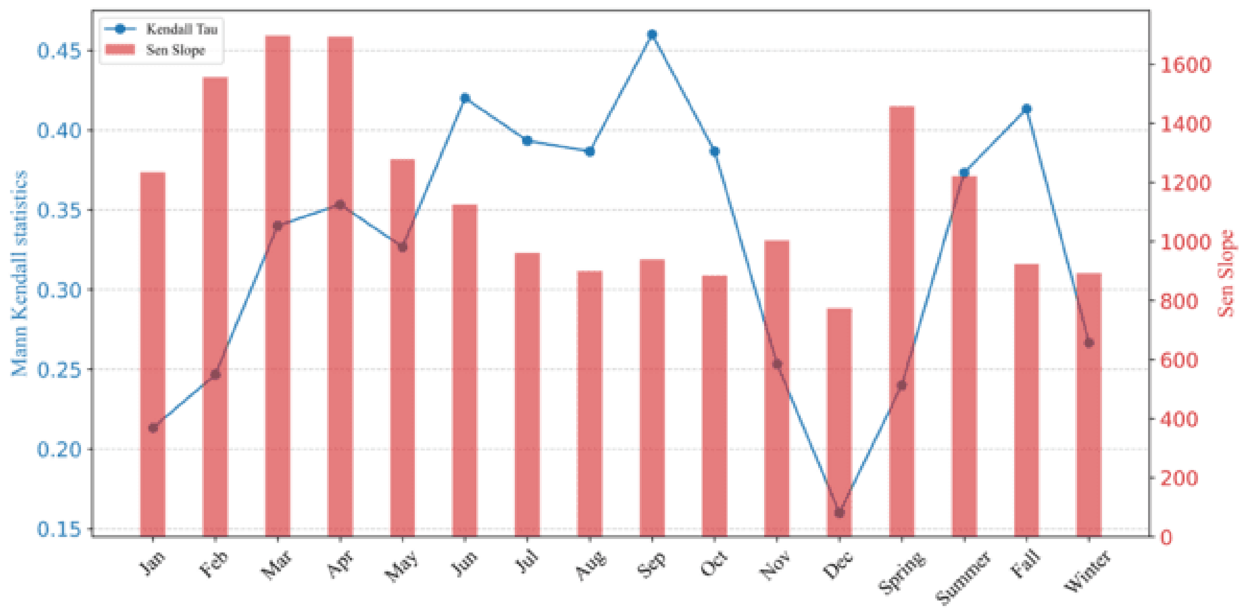


Figure 4: Results of the Mann–Kendall test and analysis of trend variations

3-3-VIF and Random Forest

Before implementing the Random Forest model, a stepwise Variance Inflation Factor (VIF) analysis was conducted to address multicollinearity among the predictors. In the first step, the highest VIF value corresponded to Actual Evapotranspiration (AET), which was removed from the dataset. In the subsequent step, Reference Evapotranspiration (PET) exhibited the next highest VIF and was also excluded. After these changes, all the remaining variables had VIF values below 10, indicating an acceptable multicollinearity. VIF values for all the variables after this are presented in Table 2.

Table 2: Results of the Variance Inflation Factor (VIF) Test

Step_1		Step_2		Step_3	
NDVI	1.37	NDVI	1.36	NDVI	1.19
LST	20.34	LST	20.34	LST	6.51
P	6.94	P	6.54	P	6.48
AET	34.95	PET	30.09	WS	1.8
PET	30.11	WS	5.33	VAP	3.14
WS	5.33	VAP	3.16	SM	1.43
VAP	3.22	SM	1.43	PDSI	1.71
SM	1.69	PDSI	1.71	RO	4.43
PDSI	1.88	RO	4.49	AOD	1.43
RO	30.87	AOD	1.6		
AOD	1.63				

The Random Forest (RF) model was trained on 80% of the data, while the remaining 20% was utilized for testing. The model was very predictive, as shown by the R^2 of 0.690 for the test set, meaning that the selected environmental variables were able to account for approximately 69% of water surface variation. The RMSE of 0.154 shows that the model predictions are very close to the observations, revealing the success of the RF approach in simulating the dynamics of the wetland.

Variable importance plots revealed that the Palmer Drought Severity Index (PDSI) was the strongest driver (importance = 0.302), suggesting that long-term drought and water availability are fundamental drivers controlling water level variation. Soil moisture (SM) was the second strongest predictor (0.250), which further underlines the impact of hydrological conditions on wetland dynamics. Vegetation health, as indicated by NDVI, was in the third position (0.110), showing the interaction between vegetation cover and water storage. The land surface temperature (LST) also played an important role (0.102), which indicates the influence of temperature-induced evaporation on wetland water levels (figure 5).

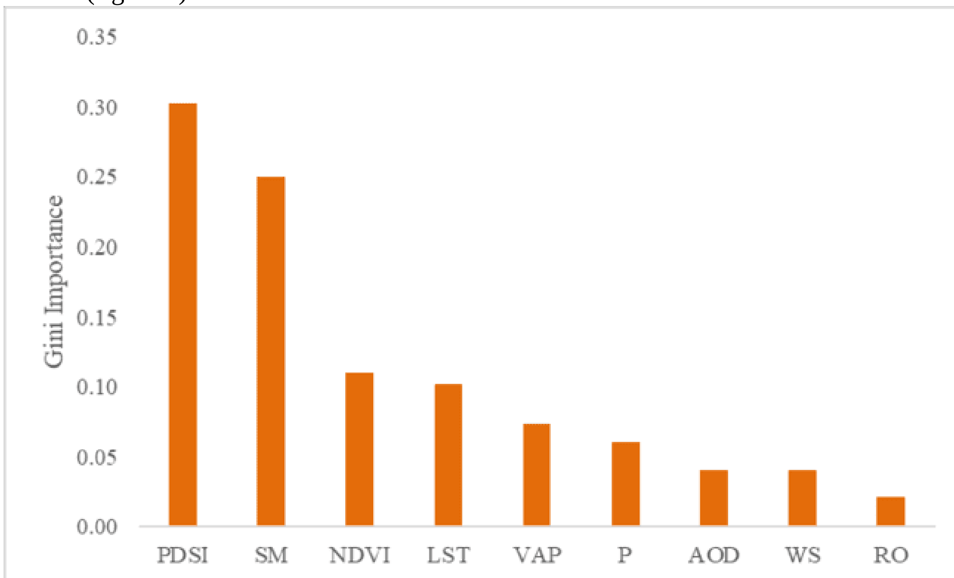


Figure 5: Results of the Gini Importance Coefficients

Other variables, including vapor pressure (VAP, 0.074), precipitation (P, 0.061), aerosol optical depth (AOD, 0.041), wind speed (WS, 0.040), and runoff (RO, 0.021), had lower but still measurable contributions. While their individual impacts were smaller, collectively they influence wetland dynamics by modulating local climate, hydrology, and sediment or pollutant transport.

Conclusions

This study analyzed long-term trends in the water levels of Hammar Marsh in Iraq from 2000 to 2025 using remote sensing techniques, non-parametric statistical methods, and machine learning models. The aim was to identify the main factors influencing wetland hydrological variability. Mann–Kendall trend analysis combined with Sen's slope estimator revealed an overall increasing trend in water levels, with the most substantial rise occurring during summer and fall, whereas winter and early spring exhibited weaker or insignificant changes. These results highlight strong seasonal fluctuations and the significant climatic influence on water availability in the marsh.

To ensure reliability in the predictive modeling, Stepwise Variance Inflation Factor (VIF) testing was used to remove highly collinear predictors such as Actual Evapotranspiration (AET) and Potential Evapotranspiration (PET). After refinement, all remaining variables had VIF values below 10, indicating minimal multicollinearity. The Random Forest model demonstrated strong predictive performance with $R^2 = 0.690$ and RMSE = 0.154 for the test dataset, suggesting that the selected environmental factors explained most water-level variance.

Variable importance analysis identified the Palmer Drought Severity Index (PDSI) and soil moisture as the most influential drivers, confirming the dominant role of drought and hydrological factors in marsh regulation. NDVI and land surface temperature (LST) were also important, reflecting vegetation–temperature interactions. Other climatic elements including precipitation, wind speed, vapor pressure, aerosol optical depth, and runoff showed secondary but meaningful influence.

Overall, the study highlights the effectiveness of integrating multi-temporal remote sensing and machine learning for wetland assessment and supports improved management strategies focusing on drought mitigation and soil moisture monitoring in Hammar Marsh.

References

[1].Sharma S, Phartiyal M, Madhav S, Singh P. Global wetlands: categorization, distribution and global scenario. Wetlands conservation: current challenges and future strategies. 2021 Nov 8:1-6.

[2].Nayak A, Bhushan B. Wetland ecosystems and their relevance to the environment: importance of wetlands. InHandbook of research on monitoring and evaluating the ecological health of wetlands 2022 (pp. 1-16). IGI

Global Scientific Publishing.
[3].Isaac GO, Hamzat AB, Daramola OO. Ecological Importance and Functions of Wetlands. InWetland Ecosystems: Conservation Strategies, Policy Management and Applications 2025 Jul 2 (pp. 23-42). Cham: Springer Nature Switzerland.
[4].Nema D, Vishwakarma A, Chauhan SS, Singh R. HYDROLOGICAL AND ECOLOGICAL INTERCONNECTIONS OF URBAN WETLANDS OF PRAYAGRAJ.
[5].Daba MH, Dejene SW. The role of biodiversity and ecosystem services in carbon sequestration and its implication for climate change mitigation. Environmental Sciences and Natural Resources. 2018;11(2):1-0.
[6].Zecheng W, Xinsheng Z, Lijuan C, Yinru L, Ziliang G, Jinzhi W, Jing L, Xiajie Z, Rumiao W, Wei L. Identification and optimization of urban wetland ecological networks in highly urbanized areas: A case study of Haidian District, Beijing. Ecological Indicators. 2025 Jan 1;170: 113028.
[7].Luo J, Zhu L, Fu H. Construction of wetland ecological network based on MSPA-Conefor-MCR: A case study of Haikou City. Ecological Indicators. 2024 Sep 1;166: 112329.
[8]. Li D, Tian P, Luo Y, Dong B, Cui Y, Khan S. Importance of stopping groundwater irrigation for balancing agriculture and wetland ecosystem. Ecological Indicators. 2021 Aug 1;127: 107747.
[9].Shravani M, Banerjee R, Pallavi N. Natural & Constructed Wetland: A Review on Water Purification and Ecosystem Services. J. Mater. Environ. Sci., 16 (7), 1245. 2025;1269.
[10]. Hemes KS, Chamberlain SD, Eichelmann E, Knox SH, Baldocchi DD. A biogeochemical compromise: The high methane cost of sequestering carbon in restored wetlands. Geophysical Research Letters. 2018 Jun 28;45(12):6081-91.
[11]. Wang L, Xie Y, Wang W, Li Y, Hou N, Yin R, Song Z, Sardans J, Ge M, Liao Y, Lai DY. Vegetation-driven differences in soil CO2 emissions and carbon-sequestering microbiomes of estuarine salt marsh and mangrove wetlands. Environmental Research. 2025 Oct 1;282: 122053.
[12]. Maltby E. The wetlands paradigm shift in response to changing societal priorities: A reflective review. Land. 2022 Sep 9;11(9):1526.
[13]. Corlett R. The ecology of tropical East Asia. Oxford University Press; 2019.
[14]. Newton A, Icely J, Cristina S, Perillo GM, Turner RE, Ashan D, Cragg S, Luo Y, Tu C, Li Y, Zhang H. Anthropogenic, direct pressures on coastal wetlands. Frontiers in Ecology and Evolution. 2020 Jul 7;8: 144.
[15]. Xiong Y, Mo S, Wu H, Qu X, Liu Y, Zhou L. Influence of human activities and climate change on wetland landscape pattern—A review. Science of the Total Environment. 2023 Jun 25;879: 163112.
[16]. Chakraborty SK, Sanyal P, Ray R. Pollution, environmental perturbation and consequent loss of wetlands. InWetlands Ecology: Eco-biological uniqueness of a Ramsar site (East Kolkata Wetlands, India) 2023 Mar 12 (pp. 521-582). Cham: Springer International Publishing.
[17]. Livinus MU, Bala SZ, Abdulsalam M, Ojeba Innocent M, Hassan M, Yusuf K, Akinola JO, Ali GM. Threats to Wetland Environments: Human Activities, Pollution and Climate Change. InWetland Ecosystems: Conservation Strategies, Policy Management and Applications 2025 Jul 2 (pp. 107-129). Cham: Springer Nature Switzerland.
[18]. Yuan S, Liang X, Lin T, Chen S, Liu R, Wang J, Zhang H, Gong P. A comprehensive review of remote sensing in wetland classification and mapping. arXiv preprint arXiv:2504.10842. 2025 Apr 15.
[19]. Farajpanah H, Lotfirad M, Adib A, Esmaeili-Gisavandani H, Kisi Ö, Riyahi MM, Salehpoor J. Ranking of hybrid wavelet-AI models by TOPSIS method for estimation of daily flow discharge. Water Supply. 2020 Dec 1;20(8):3156-71.
[20]. Esmaeili-Gisavandani H, Farajpanah H, Adib A, Kisi O, Riyahi MM, Lotfirad M, Salehpoor J. Evaluating ability of three types of discrete wavelet transforms for improving performance of different ML models in estimation of daily-suspended sediment load. Arabian Journal of Geosciences. 2022 Jan;15(1):29.
[21]. Adib A, Farajpanah H, Shoushtari MM, Lotfirad M, Saeedpanah I, Sasani H. Selection of the best machine learning method for estimation of concentration of different water quality parameters. Sustainable Water Resources Management. 2022 Dec;8(6):172.
[22]. Daloye AM, Ali FH, Azeez SN, Pana HF. Machine Learning-Based Prediction of River Water Quality Using LSTM and RF Models with PCA and Stepwise Regression for Dimensionality Reduction: A Case Study of the Maroon River Basin. Journal of Kirkuk University for Agricultural Sciences. 2025 Sep 1;16(3).
[23]. Shinkarenko SS, Bartalev SA. Application of remote sensing data in large-scale monitoring of wetlands. Cosmic Research. 2024 Dec;62(Suppl 1): S100-14.
[24]. Abdelmajeed AY, Juszczak R. Challenges and limitations of remote sensing applications in northern peatlands: present and future prospects. Remote Sensing. 2024 Feb 4;16(3):591.
[25]. Barman A, Rajak F, Jha R. Integrating Traditional Knowledge Systems for Wetland Conservation and

Management: A Critical Analysis. Nature Environment & Pollution Technology. 2025 Mar 1;24(1).
[26]. Eftekhari H, Karamidehkordi E, Mojarradi G, Naderi L, Owfi F. Coastal wetland participatory management strategies: A systematic literature review. Caspian Journal of Environmental Sciences. 2025 Jul 15:1-6.
[27]. Babaniyi GG, Akor UJ, Olagoke OV, Daramola OE. Climate Change Impacts on Wetlands: Vulnerabilities, Adaptation, and Mitigation Strategies. InWetland Ecosystems: Conservation Strategies, Policy Management and Applications 2025 Jul 2 (pp. 165-195). Cham: Springer Nature Switzerland.
[28]. Ballut-Dajud GA, Sandoval Herazo LC, Fernández-Lambert G, Marín-Muñiz JL, López Méndez MC, Betanzo-Torres EA. Factors affecting wetland loss: A review. Land. 2022 Mar 17;11(3):434.
[29]. Dube T, Dube T, Marambanyika T. A review of wetland vulnerability assessment and monitoring in semi-arid environments of sub-Saharan Africa. Physics and Chemistry of the Earth, Parts A/B/C. 2023 Dec 1;132:103473.
[30]. Ebrahimi-Khusfi Z, Samadi-Todar SA, Okati N, Kaskaoutis D. A 25-Year Assessment of Aerosol Dynamics and Environmental Drivers in Iran's Lakes and Wetlands.
[31]. Xi Y, Peng S, Ciais P, Chen Y. Future impacts of climate change on inland Ramsar wetlands. Nature Climate Change. 2021 Jan;11(1):45-51.
[32]. Chen K, Cong P, Qu L, Liang S, Sun Z, Han J. Identifying changes in the hydrological connectivity and their drivers in the Liaohe Delta wetland. Ocean & Coastal Management. 2023 Aug 1;242:106733.
[33]. Ekumah B, Armah FA, Afrifa EK, Aheto DW, Odoi JO, Afitiri AR. Assessing land use and land cover change in coastal urban wetlands of international importance in Ghana using Intensity Analysis. Wetlands Ecology and Management. 2020 Apr;28(2):271-84.
[34]. Jamali AA, Naeeni MA, Zarei G. Assessing the expansion of saline lands through vegetation and wetland loss using remote sensing and GIS. Remote Sensing Applications: Society and Environment. 2020 Nov 1;20: 100428.
[35]. Ji P, Su R, Wu G, Xue L, Zhang Z, Fang H, Gao R, Zhang W, Zhang D. Projecting Future Wetland Dynamics Under Climate Change and Land Use Pressure: A Machine Learning Approach Using Remote Sensing and Markov Chain Modeling. Remote Sensing. 2025 Mar 20;17(6):1089.
[36]. Cui L, Li G, Liao H, Ouyang N, Li X, Liu D. Remote sensing of coastal wetland degradation using the landscape directional succession model. Remote Sensing. 2022 Oct 21;14(20):5273.
[37]. He K, Zhang Y, Li W, Sun G, McNulty S. Detecting coastal wetland degradation by combining remote sensing and hydrologic modeling. Forests. 2022 Mar 3;13(3):411.
[38]. Adib A, Farajpanah H, Mahmoudian Shoushtari M, Ahmadeanfar I. Estimation of Water Quality Parameters in the Sepidrood River by ANFIS, GEP and LS-SVM Models. Journal of Water and Wastewater. 2020 Nov 21;31(5):1-0.
[39]. Farajpanah H, Adib A, Lotfirad M, Esmaeili-Gisavandani H, Riyahi MM, Zaerpour A. A novel application of waveform matching algorithm for improving monthly runoff forecasting using wavelet–ML models. Journal of Hydroinformatics. 2024 Jul 1;26(7):1771-89.
[40]. Ridwan I, Kadir S, Nurlina N. Wetland degradation monitoring using multi-temporal remote sensing data and watershed land degradation index. Global Journal of Environmental Science and Management. 2024 Jan 1;10(1):83-96.
[41]. Aslam RW, Shu H, Javid K, Pervaiz S, Mustafa F, Raza D, Ahmed B, Quddoos A, Al-Ahmadi S, Hatamleh WA. Wetland identification through remote sensing: Insights into wetness, greenness, turbidity, temperature, and changing landscapes. Big Data Research. 2024 Feb 28;35: 100416.
[42]. Chen Y, Ci R, Zhong D, Liu L, Yu J, Zhang D, Tong Y, Fu Y. Characterizing and Detecting Multiscenario Degradation of the Maidika Alpine Wetland Nature Reserve in the Qinghai–Tibet Plateau Using Landsat Time Series. Journal of Remote Sensing. 2025 Jan 2;5: 0380.
[43]. Li Q, Li M, Zhao J, Mishu RA, Hu Z, Hasan M, Kamruzzaman M, Bhattacharjee A, Rahman M, Kabir Z, Rahman MM. Future spatiotemporal inequality of precipitation extreme in monsoon-driven country using Gini coefficient and random forest model. Theoretical and Applied Climatology. 2025 Jul;156(7):382.
[44]. Tholo HM, Masuku P, Mwase W, Kachamba D, Njala A, Tchongwe W, Mvula RL, Tchuwa I, Mzumara TI, Chisenga C, Kadewa W. Integrating Remote Sensing and Machine Learning for High-Resolution Vulnerability Mapping of Seasonal Wetland Shrinkage under Climate Change in Southern Malawi.
[45]. Ramsar Sites Information Service. (2014, April 7). Hammar Marsh (Site No. 2242). Ramsar Convention Secretariat. Retrieved August 21, 2025, from https://rsis.ramsar.org/ris/2242
[46]. Ramsar Sites Information Service. (2012, April 7). Hammar Marsh (Site No. 2242). Ramsar Convention Secretariat. Retrieved August 21, 2025, from https://rsis.ramsar.org/ris/1718
[47]. Amani M, Ghorbanian A, Ahmadi SA, Kakooei M, Moghimi A, Mirmazloumi SM, Moghaddam SH, Mahdavi S, Ghahremanloo M, Parsian S, Wu Q. Google earth engine cloud computing platform for remote sensing big data applications: A comprehensive review. IEEE Journal of Selected Topics in Applied Earth Observations and

Remote Sensing. 2020 Sep 1;13: 5326-50.
[48]. Gao BC. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote sensing of environment. 1996 Dec 1;58(3):257-66.
[49]. Gal MS, Rubinfeld DL. Data standardization. NYUL Rev..2019;94:737.
[50]. Galeati G. A comparison of parametric and non-parametric methods for runoff forecasting. Hydrological sciences journal. 1990 Feb 1;35(1):79-94.
[51]. Mircioiu C, Atkinson J. A comparison of parametric and non-parametric methods applied to a Likert scale. Pharmacy. 2017 May 10;5(2):26.
[52]. Libiseller C, Grimvall A. Performance of partial Mann–Kendall tests for trend detection in the presence of covariates. Environmetrics: The official journal of the International Environmetrics Society. 2002 Feb;13(1):71-84.
[53]. Coccia M, Roshani S. Evolution of topics and trends in emerging research fields: multiple analyses with entity linking, Mann–Kendall test and burst methods in cloud computing. Scientometrics. 2024 Sep;129(9):5347-71.
[54]. Dutta P, Deka S. A novel approach to flood risk assessment: Synergizing with geospatial based MCDM-AHP model, multicollinearity, and sensitivity analysis in the Lower Brahmaputra Floodplain, Assam. Journal of Cleaner Production. 2024 Aug 15;467: 142985.
[55]. Bargam B, Boudhar A, Kinnard C, Bouamri H, Nifa K, Chehbouni A. Evaluation of the support vector regression (SVR) and the random forest (RF) models accuracy for streamflow prediction under a data-scarce basin in Morocco. Discover Applied Sciences. 2024 Jun 3;6(6):306.
[56]. Mustafa OM, Ahmed OM, Saeed VA. Comparative analysis of decision tree algorithms using gini and entropy criteria on the forest covertypes dataset. InThe International Conference on Innovations in Computing Research 2024 Aug 1 (pp. 185-193). Cham: Springer Nature Switzerland.

مساحة الأراضي الرطبة واختبار العوامل المؤثرة عليها باستخدام التعلم الآلي

سروشت محمد جزا[1] **تەرزە جبار باقي[2]**

هيوا فرج پناه[3]

جامعة السليمانية التقنية، المعهد دربنديخان التقني، قسم المساحة، السليمانية، إقليم كردستان، العراق
جامعة السليمانية التقنية، المعهد دربنديخان التقني، قسم المساحة، السليمانية، إقليم كردستان، العراق
جامعة الشهيد شمران، كلية الهندسة المدنية والعمارة، الأهواز - ايران

الخلاصة

الأراضي الرطبة أنظمة بيئية خاصة تُقدم خدمات هيدرولوجية وبيئية واجتماعية واقتصادية بالغة الأهمية. تبحث الدراسة في التطور طويل الأمد لهور الحمار في العراق من عام 2000 إلى عام 2025، مع التركيز على اتجاهات منسوب المياه والعوامل البيئية الدافعة للتغيرات. حُلّلت صور الاستشعار عن بُعد باستخدام محرك جوجل إيرث للحصول على بيانات شهرية عن مساحات المياه السطحية وغيرها من المتغيرات المناخية والبيئية الرئيسية. طُبّق اختبار مان-كيندال ومُقدّر منحدر سين للكشف عن اتجاهات مهمة في منسوب المياه، وقد لوحظت زيادة بشكل عامة، لا سيما في الصيف والخريف، بينما كانت التغيرات أبطأ في الشتاء وأوائل الربيع.
أُجري تحليل معامل تضخم التباين التدريجي (***VIF***) لتقليل التعدد الخطي بين المتنبئات، بحيث تكون قيم معامل تضخم التباين لجميع المتغيرات المتبقية أقل من 10. ثم طُبّق نموذج الغابة العشوائية لاستنتاج الأهمية النسبية للعوامل البيئية الدافعة. أظهر النموذج قيمة ***R2*** لمجموعة الاختبار بلغت 0.690 و***RMSE*** بلغت 0.154، مما يدل على قدرة جيدة على التنبؤ. وأشار حساب أهمية المتغير إلى أن مؤشر بالمر لشدة الجفاف (***PDSI***) ورطوبة التربة هما العاملان المسيطران على تغير منسوب المياه، يليهما الغطاء النباتي (***NDVI***) ودرجة حرارة سطح الأرض (***LST***). وتُبرز النتائج التأثيرات التآزرية للعوامل المناخية والهيدرولوجية على ديناميكيات الأراضي الرطبة، وتُظهر فعالية دمج الاستشعار عن بُعد وتحليل الاتجاهات والتعلم الآلي في رصد الأراضي الرطبة. ولنتائج هذه الدراسة آثار هامة على الإدارة المستدامة والحفاظ على مستنقع الحمار وغيره من النظم البيئية المماثلة للأراضي الرطبة في ظل الظروف البيئية المتغيرة.

الكلمات المفتاحية: الأراضي الرطبة، الاستشعار عن بعد، العوامل الهيدرولوجية المناخية، التعلم الآلي، مؤشر الغطاء النباتي الوطني.