Int. J. Aquat. Biol. (2024) 12(6): 581-589
ISSN: 2322-5270; P-ISSN: 2383-0956
Journal homepage: www.ij-aquaticbiology.com

Original Article

Application of machine learning in predicting sources of water pollution in the Euphrates and Tigris rivers in Iraq

Mohammed Kareem Rashid[1], Israa Ramadhan Salman[2], Abbas Luaibi Obaid[3], Saif Al-Deen H. Hassan[*4], Mohammed Raoof Al-musawi[4], Moumal Al-Saady[5]

[1]Electronic Computing Center, University of Misan, Maysan, Iraq.
[2]College of Pharmacy, University of Misan, Maysan, Iraq.
[3]College of Agriculture, University of Misan, Maysan, Iraq.
[4]Department Business Administrator, College of Administration and Economics, University of Misan, Maysan, Iraq.
[5]Monash University, Clayton 3168, Melbourne, Victoria, Australia.

Abstract: New evaluation and control methods are required to address the ecological, economic, and public health concerns raised by the contamination of the rivers Tigris and Euphrates. To minimize negative effects on ecosystems, our research built and implemented a machine learning framework to track down and foresee potential water contamination hotspots. To examine the causes of pollution and its consequences on aquatic ecosystems, researchers combined data from multiple sources, such as aerial photographs, field surveys, and official government documents. Predictive models encompass significant attributes such as pesticides, mineral composition, suspended particulates, diversity of macroinvertebrates, and habitat quality. Feature selection techniques, including LASSO regression and recursive feature elimination, ensured dependable model construction. Four machine learning algorithms of MCP, K-nearest neighbors, decision tree, and multi-layer perceptron were employed for pollution source recognition and impact prediction. The models correctly identified significant pollution sources, including untreated sewage, agricultural runoff, and industrial discharges. The concentration and distribution patterns of pollutants were elucidated by clustering and regression techniques. The results indicated reduced biodiversity, habitat degradation, and toxic algal blooms, as well as identified significant pollution areas. This research shows that machine learning can transform environmental monitoring and water resource management. The study's practical findings, which integrate ecological and computational methodologies, can assist policymakers and water resource managers.

Article history:
Received 17 October 2024
Accepted 22 December 2024
Available online 25 December 2024

Keywords:
Aquatic Ecology
Predictive modeling
Classification algorithms
Pollutant sources

Introduction

Water pollution in the Euphrates and Tigris Rivers in Iraq threatens the integrity of these vital ecosystems, presenting a significant and escalating issue. The Euphrates and Tigris rivers support diverse freshwater ecosystems, including lotic and lentic habitats, riparian zones, and wetlands (Evans, 1994; Al-Ansari, 2021a; Fadhil et al., 2024). These ecosystems are crucial in regional biodiversity and support over 52% of the Iraqi population. However, water quality has deteriorated due to industrial discharges, untreated sewage, and agricultural runoff, negatively impacting the complex biological communities inhabiting these rivers (Al-Ansari, 2021a). Drought conditions and runoff from irrigation exacerbate pollution in these rivers. Migration due to conflict increases reliance on river water for sanitation, further degrading water quality and impacting ecological function. Analysis of surface water shows significant contamination levels with negative ecological effects. In addition, for economies, businesses, public health systems, and ecosystems to be viable over the long term, the freshwater resources of the Tigris and Euphrates rivers are essential (FAO, 2016). Algal turbidity and agricultural pesticides are common pollutants that damage aquatic ecosystems and biodiversity (Anderson et al., 2002; World Health Organization, 2017; Al-Abboodi and Kareem, 2023). UNEP (2023)

*Correspondence: Saif Al-Deen H. Hassan
E-mail: saif_aldeen@uomisan.edu.iq

DOI: https://doi.org/10.22034/ijab.v12i6.2421

and Fadhil et al. (2024) indicate that hazardous substances substantially affect water quality and the complex biological processes, local economies, and human societies dependent on these resources.

A promising new avenue for environmental restoration has arisen from exploring innovative wastewater treatment methods, including diatomaceous earth and other natural resources, which are sustainable and ecologically advantageous (Benouis et al., 2022). The bacterial breakdown of oil-contaminated soil exemplifies how biological remediation of pollutants can contribute to sustainability and global development objectives (Aziz et al., 2024). It is critical to locate and eliminate the causes of water contamination to purify water efficiently and ensure everyone can access clean water (UNESCO, 2018).

Due to time, cost, and scope constraints, earlier research in the area had to be limited (Rahi and Halihan, 2021; Mueller et al., 2021). Consequently, there is an immediate need for better methods. An option proposed by Jordan and Mitchell (2015) and Hassan et al. (2023) is machine learning, which allows for rapidly examining large and complex datasets. The spatial distribution of water quality indicators can be quantified efficiently in Iraq and similar developing nations using machine learning (Nazari et al., 2021). The Iraq River Studies Centre possesses in-house expertise in applying the Random Forest machine learning model to address such challenges (Rash et al., 2023; Al-Abadi et al., 2024). However, to ensure relevance to aquatic ecology, it is essential to integrate ecological concepts, data, and validation strategies into the modeling. Traditional water safety assessments based on physical and chemical analyses often fail to capture the complex ecological dynamics in this basin. Machine learning has the potential to streamline the process, though its application in environmental studies is still relatively novel (Mahmud, 2021; Ahram, 2021; Lorenz and Erickson, 2023). Traditional techniques for measuring pollution are expensive, time-consuming, and labor-intensive. There is a lack of research on the ecological effects of pollution, as machine learning can identify patterns in large datasets (Proshad et al., 2021).

This study employs machine learning to examine stakeholder viewpoints, predicted sources, and ecological implications to assess environmental and water quality needs. This will help us understand pollutant dynamics, water management, and ecosystem preservation. The main goal is to create machine learning algorithms to identify water contamination sources and environmental implications. To fully understand, we will compare stakeholder perceptions with expected pollution sources and ecosystem effects. Therefore, this study aims (1) to use machine learning to forecast ecological indicators and water quality and (2) to clarify these models as pollution indicators and ecological effects using stakeholder comments. This method will contain important water users, sociodemographic data, sample methods, ecological indicators, and water sufficiency assessments.

Materials and Methods

The study employs machine learning to identify causes of water contamination, emphasizing environmental implications using ecological health, water quality, and environmental variables. Macroinvertebrate community structure, fish diversity and abundance, algal composition, suspended particles, pesticide concentrations, mineral content, and habitat quality were considered. To ensure accuracy, data was preprocessed and feature-selected. CART decision trees, support vector machines, backpropagation multi-layer perceptrons, and K-nearest neighbors with an RBF kernel were examined. For cross-validation, 30% of the dataset was used for model testing (Sarker, 2021). This approach integrates ecological data with machine learning to understand water contamination and its effects on aquatic ecosystems.

Data collection: The ecological health and pollution levels of the Tigris and Euphrates rivers were evaluated using data from multiple sources. Government databases, field surveys, and satellite imagery were some data sources.

Satellite imagery: Landsat 8 and Sentinel-2 imagery captured spatial and temporal variations of water quality parameters and surrounding vegetation cover. Data from 2021 to 2023 were analysed for turbidity, chlorophyll levels, vegetation indices (NDWI, NDVI), and land use changes (Jensen, 2015). The spatial resolution was 30 m and 10-20 m, respectively, with temporal resolutions of 16 and 5 days.

Field surveys: Field surveys were conducted at 15 sampling locations along the rivers. Water samples were analysed for pH, temperature, dissolved oxygen (DO), electrical conductivity (EC), total dissolved solids (TDS), nitrate, phosphate, heavy metals (lead, cadmium, mercury), and pesticide concentrations. Biological surveys included the collection of macroinvertebrates, fish surveys, and algal samples (Reynolds, 1996; Barbour et al., 1999; Chorus and Bartram, 1999). Habitat assessments included the physical habitat, water flow, and riparian vegetation structure using metrics of the Index of Biotic Integrity (IBI) (Karr, 1981). The questionnaires focused on water use, perceived pollution sources, and associated health and ecological impacts. The interviews covered management strategies and perceptions of pollution on water quality and ecosystem health.

Government databases: Data from municipality and household census reports, industrial activity reports, and the Ministry of Water Resources databases were utilized. This data included population density, industrial discharge locations, wastewater treatment plant outputs, and available ecological data (Iraq Ministry of Planning, 2020).

Preprocessing: Data were cleaned and preprocessed prior to analysis. Missing values were imputed using multiple imputations by chained equations. Normalisation was performed using min-max scaling to ensure all variables ranged from 0 to 1.

Feature selection: Feature selection aims to reduce overfitting and simplify the model. The LASSO regression method removes unnecessary variables by prioritizing ecological significance. Recursive feature elimination (RFE) using a random forest model validated the selected features (Guyon et al., 2002). To improve prediction models, we included ecological indicators and water quality metrics in feature selection.

Machine learning algorithms: Machine learning uses clustering, classification, and classification-regression. Most models classified pollution as natural or industrial using a binary method. Regression models predicted contaminant and ecological indicator amounts. Clustering revealed pollution and environmental trends.

Decision tree (CART): The classification and regression tree (CART) algorithm developed interpretable models for identifying pollution sources and making rule-based predictions about ecological health.

K-nearest neighbors (KNN): The ease and efficacy of employing Euclidean distance in implementing KNN for classification tasks led to its selection. Cover and Hart's (1967) method used grid search to optimize the k-value, which tested values ranging from 3 to 15.

Support vector machines (SVM): The ecological impacts and sources of pollution were classified using a support vector machine (SVM) trained with an RBF kernel. According to Cortes and Vapnik (1995), the hyperparameters C and gamma were optimized using grid search to guarantee optimal model performance.

Multi-layer perceptron (MLP): For regression and classification, a backpropagation neural network (MLP) was employed. The design had a single hidden layer that used the ReLU activation function. To ensure the hyperparameters (learning rate, momentum, and neurons) were tuned to suit the data properly, we used grid search with cross-validation (Rumelhart et al., 1986).

Precision, recall, F1-score, accuracy, and the area under the ROC curve were typical classification measures used to assess the model's performance (Fawcett, 2006). According to Moriasi et al. (2007), the following metrics were used to evaluate the performance of the models: RRSE, MAE, R, R^2, and NSE.

Results and Discussions

The Euphrates and Tigris rivers originate from Syria and Turkey, and they supply Iraqi water, which decreases over time due to the construction of many dams. The Tigris and Euphrates Rivers are subjected to extensive pollution from untreated sewage, agricultural runoff (notably soil erosion), and elevated levels of suspended solids, all of which significantly affect their ecology (Oleiwi and Al-Dabbas, 2022; Madhi and Azeez, 2024; Olson and Speidel, 2024). These rivers pass metropolitan areas that release untreated garbage, resulting in health issues and adversely affecting aquatic ecosystems (WHO, 2019). The increase of opportunistic organisms resulting from sewage discharge further disturbs natural ecosystems (Smith, 2003).

Primary pollutants that directly impact aquatic flora and fauna are pesticides, fertilizers, food industry waste, and detrimental livestock practices. Studies indicate that contamination adversely affects invertebrate biodiversity, primary production patterns, and food web dynamics (Al-Maliki et al., 2021; Al-Yasiri, 2021; Hamza et al., 2024). Resilient aquatic species proliferate, but fragile ones decline in population. Bryan and Langston (1992) and Gilliom et al. (2006) indicate that bioaccumulation of heavy metals and pesticides can lead to detrimental impacts, alterations in species diversity, and compromised fish health. Turbidity diminishes light penetration, impacting primary production and visual predation (Kirk, 1994). Given these data, it is evident that further extensive research into the sources and effects of pollution is necessary, besides the public discourse around pollution control measures (Hassan et al., 2024).

Due to water shortage and demand, water quality and pollution testing are critically needed (UNESCO, 2018). Many studies have examined river pollution effects on aquatic ecosystems and humans (Al-Ansari, 2021b; Al-Saadi and Sadkhan, 2021), showing that wastewater discharge and heavy metal contamination harm biodiversity and ecological integrity (USEPA, 2012). The main objectives of pollution mitigation initiatives are to locate and remove the sources of pollutants and to increase knowledge of how the environment responds to these modifications. Conventional techniques, including chemical analysis, flow manipulation, and fish and benthic creature evaluations, are time-consuming and have limitations in complex systems (APHA, 2017). Although theoretical frameworks exist to assess the effects of hazardous metals, traditional methodologies fail to consider the wider environmental repercussions of pollution. Contemporary approaches, such as content-weighted mixture models, provide economical methods for identifying pollution sources, although they lack comprehensive ecological understanding (Yadav and Sharma, 2023). Hence, this work addresses this gap by utilizing machine learning to link pollution sources to ecological effects. Innovative methodologies, including machine learning, must be employed to evaluate the effects on surface water quality and ecological health (Elbeltagi et al., 2021).

Conventional methods for identifying pollution sources, although employing digital elevation models, remote sensing, and water quality monitoring, continue to depend on visual inspections, geographical analysis, and verification of receiving water quality (USGS, 2023). Minor catchments utilize techniques such as water sampling, field research, and yard tests (Agibayeva et al., 2022; Pongpiachan et al., 2024; Chen et al., 2024; Mullen et al., 2024). While precise, these methods are arduous and do not consistently consider all the subtleties of water contamination's impact on the environment. Ecological monitoring approaches for assessing ecosystem health encompass macroinvertebrate sampling, fish surveys, algal sampling, and habitat structure evaluation (Karr, 1981; Plafkin et al., 1989; Barbour et al., 1999; Chorus and Bartram, 1999). Environmental data can be linked with machine learning to enhance the identification of pollution sources and assess ecological impacts.

Freshwater contamination was a major concern in seventeen nations in 2018, including Iraq (Al-Ansari et al., 2021b). Altahaan and Dobslaw (2024) and Todd (2024) state that the Tigris and Euphrates rivers are significant for the water supply in Iraq. Groundwater recharge, biodiversity, and ecosystem function are all greatly aided by their presence. Here, we look at the contaminants, hydrochemical parameters, and

Table 1. Water quality parameters and ecological indicators at sampling stations.

Sampling location	Date	Time	pH	DO (mg/L)	EC (µS/cm)	TDS (mg/L)	Nitrate (mg/L)	Phosphate (mg/L)	Lead (µg/L)	Cadmium (µg/L)	Mercury (µg/L)	EPT Index	Fish species richness	Chlorophyll-a (µg/L)	Habitat score
Location 1	2023-07-15	10:00	7.8	6.5	550	300	1.2	0.3	2	0.5	0.1	10	8	12	75
Location 2	2023-07-15	11:30	8.1	5.8	700	400	2.5	0.7	5	1.2	0.3	5	5	25	60
Location 3	2023-07-15	13:00	7.5	7.2	480	280	0.8	0.2	1	0.3	0.05	12	10	8	80
Location 4	2023-07-16	9:00	8.2	4.5	850	500	3.0	1.0	10	2.0	0.5	2	3	35	40
Location 5	2023-07-16	10:30	7.9	6.0	620	350	1.8	0.5	3	0.8	0.2	7	7	18	70
Location 6	2023-07-16	12:00	7.7	5.5	900	550	3.2	1.2	12	2.5	0.7	1	2	40	30
Location 7	2023-07-17	10:00	8.0	7.0	500	250	0.5	0.1	0.5	0.1	0.02	15	12	5	85
Location 8	2023-07-17	11:30	8.3	4.0	950	600	3.5	1.5	15	3.0	0.8	0	1	50	20
Location 9	2023-07-17	13:00	7.6	7.5	520	270	0.7	0.2	0.8	0.2	0.04	13	11	7	82
Location 10	2023-07-18	9:00	7.4	5.0	780	450	2.0	0.8	6	1.5	0.3	4	4	30	50
Location 11	2023-07-18	10:30	7.9	6.8	600	320	1.5	0.4	2.5	0.6	0.1	9	7	15	72
Location 12	2023-07-18	12:00	7.6	4.2	920	580	3.3	1.3	11	2.3	0.6	2	3	38	35
Location 13	2023-07-19	10:00	8.0	7.1	490	240	0.6	0.15	0.6	0.15	0.03	14	13	6	88
Location 14	2023-07-19	11:30	8.3	4.3	940	610	3.6	1.4	14	2.8	0.75	0	2	45	25
Location 15	2023-07-19	13:00	7.8	7.3	510	260	0.75	0.25	0.9	0.25	0.045	12	9	9	85

ecological indicators associated with pollution sources and their effects on the ecosystem using the data obtained. Incorporating the necessary complicated datasets has been a common omission in earlier efforts to handle these issues thoroughly (Suzuki et al., 2022). Our research will use ecological data, machine learning techniques, and local experience to improve environmental monitoring and management.

Performance evaluation of models: The efficacy of machine learning models was assessed using a range of statistical metrics. According to Moriasi et al. (2007), the algorithms used to evaluate regression jobs included RRSE, MAE, R, R^2, and NSE, which stood for root relative squared error and mean absolute error, respectively. Classification task performance was evaluated using F1-score, recall, accuracy, precision, and the area under the receiver operating characteristic (ROC) curve (Fawcett, 2006). Furthermore, the specific impact on ecological indicator forecasting and the percentage enhancement in accuracy were calculated. Various metrics demonstrated that the models had strong performance.

Identification of pollution sources and ecological impacts: Time series data from 15 monitoring stations along the Euphrates River was analysed. This included parameters such as EC, temperature, pH, ammonia, nitrite, nitrate, phosphate, sulphate, bicarbonate, macroinvertebrate community composition (quantified by the EPT Index), fish diversity (measured as species richness), algal abundance (using Chlorophyll-a concentration) and a calculated habitat quality score. Classification models (SVM, Random Forest, and Neural Networks) were used to identify whether the sources of pollution were primarily natural or industrial and what type of impact they have on the biological community. The random forest model and Boruta feature selection were used to identify critical factors affecting water quality and ecological health. The model's predictions were compared with stakeholder inputs gathered through questionnaires and interviews to validate its reliability. The final model can classify the sources of pollution, predict ecological impacts, and show an associated accuracy level. This will be followed by an ecological interpretation of these results.

Fifteen sampling stations are shown in Table 1. The data shows variations in water quality parameters and ecological health indicators across locations, revealing the differing pollution levels and their impact on the ecosystem. For example, station 8 shows high levels of nitrates and phosphates and low levels of DO, along with a low EPT index and low fish species richness, indicative of anthropogenic impacts. Conversely, station 13 shows a high EPT index and fish species richness, low levels of contaminants, and high DO, indicating a healthier ecosystem. These findings are complemented by stakeholder perceptions (Table 2). Many stakeholders from station 4, for instance,

Table 2. Stakeholder perception data.

Stakeholder ID	Stakeholder Group	stations	Perceived Source of Pollution	Perceived Ecological Impact
S101	Farmer	4	Agricultural Runoff	Decline in fish abundance
S102	Local Community	2	Industrial discharge	Increased algal blooms
S103	Industry Rep	6	Untreated effluent	Reduced water clarity
S104	Farmer	8	Fertilizer runoff	Reduction in plant diversity
S105	Local Community	1	Domestic sewage	More insects
S106	Industry Rep	7	Industrial discharge	Reduced habitat quality
S107	Farmer	11	Pesticide use	Decrease in bird population
S108	Local Community	9	Agricultural and domestic	Reduced fish health
S109	Industry Rep	12	Industrial discharges	Reduced water clarity
S110	Farmer	15	Lack of proper management	Increased invertebrate abundance
S111	Local Community	13	Sewage runoff	Increased algal blooms
S112	Industry Rep	3	Industrial discharge	Decrease in oxygen level
S113	Farmer	10	Runoff	Reduction in bird abundance
S114	Local Community	5	Domestic and agricultural	Decline in fish abundance
S115	Industry Rep	14	Untreated effluent	Increase in opportunistic species

Table 3. Feature importance from machine learning model (example random forest).

Feature	Importance Score
Nitrate (mg/L)	0.18
Phosphate (mg/L)	0.15
Lead (µg/L)	0.12
EPT Index	0.10
Habitat Score	0.09
Fish Species Richness	0.08
DO (mg/L)	0.07
Cadmium (µg/L)	0.06
EC (µS/cm)	0.05
Chlorophyll-a (µg/L)	0.04
Mercury (µg/L)	0.03
TDS (mg/L)	0.02
pH	0.01

identified agricultural runoff as a significant issue, correlating well with the high levels of nitrates and phosphates measured at that location. Similarly, stakeholders in stations 2 and 6 described industrial discharge as a major cause of pollution, corresponding with the low EPT index, low fish species richness, high metal contamination, and low habitat quality at those locations. However, the stakeholders at stations 13 and 15 largely reported minimal environmental concerns, which aligns with the low pollutant concentrations and high ecological health observed.

The random forest model and Boruta feature selection identified key predictors influencing ecological health (Table 3). Nitrate (mg/L), phosphate (mg/L), and lead (µg/L) emerged as the most influential factors, suggesting that agricultural runoff and heavy metal contamination are major drivers of ecological degradation in the region. The importance of the EPT index, fish species richness, and habitat score highlights that biological parameters are vital in assessing ecological health.

The classification performance of the different models is presented in Table 4. The Multi-Layer Perceptron (MLP) model yielded the highest performance across all metrics, indicating its potential for accurately identifying pollution sources and their impacts. The relatively high performance of the KNN model, compared to other models, may indicate that the ecological parameters and locations of concern may be better classified using this model. Further analysis will clarify the nuances of these findings.

The data and the models suggest that the sources of pollution are diverse and have variable impacts. Further ecological interpretation of the results will allow us to more accurately describe the specific biological communities most affected by various sources of pollution. This analysis demonstrates the value of combining machine learning with ecological

Table 4. Classification performance of machine learning models.

Model	Accuracy	Precision	Recall	F1-Score	ROC AUC
Decision Tree	0.75	0.70	0.72	0.71	0.74
KNN	0.82	0.78	0.80	0.79	0.81
SVM	0.85	0.83	0.84	0.83	0.86
MLP	0.88	0.86	0.87	0.86	0.89

data for understanding pollution sources and their consequences in complex freshwater ecosystems. The integration of stakeholder perceptions provides real-world validation of the models. This approach enhances our understanding of the complex interactions between water quality, ecological health, and local community perceptions.

Conclusion

This research presented a machine learning model for identifying water pollution sources in the Euphrates and Tigris rivers, specifically focusing on their ecological impacts. The model effectively classified water samples based on the pollution source and predicted impacts on water quality and the local ecosystem. The research provides recommendations for policymakers regarding the main sources of pollution impacting the rivers and the negative effects on their unique ecosystems. This work expands on previous research and provides a model that effectively combines complex datasets to provide real-world implications for water management. Machine learning techniques provided a suitable approach for environmental monitoring, particularly for identifying sources of pollution and their impact on ecological health. Classification algorithms (KNN, decision tree, random forest, logistic regression, and Naive Bayes) were applied to identify sources of pollution in the Euphrates and Tigris rivers. The models predicted pollution sources, including hospitals, domestic areas, upstream sources, agricultural areas, and industrial zones, each with an associated predicted pollution density. The models produced accurate results based on metrics including accuracy, recall, F1-score, and the relevant ecological metrics. The study delivered practical methods for identifying pollution sources, their impacts, and areas of concern. These findings aid the introduction of purification techniques and improvements in ecosystem management. Furthermore, a logistic regression model was also successfully applied to identify pollution sources.

References

Agibayeva A., Khalikhan R., Guney M., Karaca F., Torezhan A., Avcu E. (2022). An air quality modeling and disability-adjusted life years (DALY) risk assessment case study: Comparing statistical and machine learning approaches for PM2. 5 Forecasting. Sustainability, 14(24): 16641.

Ahram A.I. (2021). Human ecology of the marshes. In Southern Iraq's Marshes: Their Environment and Conservation. Springer International Publishing. pp: 55-76.

Al-Abadi A.M., Hassan A.A., Al-Moosawi N.M., Handhal A.M., Alzahrani H., Jabbar F.K., Anderson N.L. (2024). Drought susceptibility mapping in Iraq using GRACE/GRACE-FO, GLDAS, and machine learning algorithms. Physics and Chemistry of the Earth, Parts A/B/C, 134: 103583.

Al-Ansari N. (2021a). Water resources of Iraq. Journal of Earth Sciences and Geotechnical Engineering, 11(2): 15-34.

Al-Ansari N., Saleh S., Abdullah T., Abed S.A. (2021b). Quality of surface water and groundwater in Iraq. Earth Sciences and Geotechnical Engineering, 11(2): 161-199.

Al-Maliki L.A., Farhan S.L., Jasim I.A., Al-Mamoori S.K., Al-Ansari N. (2021). Perceptions about water pollution among university students: A case study from Iraq. Cogent Engineering, 8(1): 1895473.Al-Saadi N., Sadkhan K.A.M. (2021). Water resources decrease of river Euphrates and its impacts on the environment in Iraq. International Journal of Ecology and Environmental Sciences, 33-55.

Altahaan Z., Dobslaw D. (2024). Assessment of the impact of war on concentrations of pollutants and heavy metals and their seasonal variations in water and sediments of the Tigris River in Mosul/Iraq. Environments, 11(6): 66.

Al-Yasiri A.J. (2021). Iraq's post-2003 environmental pollution: Issues and solutions. Akkad Journal of Contemporary Economic Studies, 1(1): 1-16.

Anderson D.M., Glibert P.M., Burkholder J.M. (2002). Harmful algal blooms and eutrophication: nutrient sources, composition, and consequences. Estuaries, 25(4): 704-726.

APHA (American Public Health Association), (2017). Standard methods for the examination of water and wastewater. American Public Health Association.

Aziz Z.S., Jazaa S.H., Dageem H.N., Banoon S.R., Balboul B.A., Abdelzaher M.A. (2024). Bacterial biodegradation of oil-contaminated soil for pollutant abatement contributing to achieve sustainable development goals: A comprehensive review. Results in Engineering, 102083.

Barbour M.T., Gerritsen J., Snyder B.D., Stribling J.B. (1999). Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish. US Environmental Protection Agency.

Benouis K., Khane Y., Ahmed T., Albukhaty S., Banoon S.R. (2022). Valorization of diatomaceous earth as a sustainable eco-coagulant for wastewater treatment: optimization by response surface methodology. Egyptian Journal of Chemistry, 65(9): 777-788.

Bryan G.W., Langston W.J. (1992). Bioavailability, accumulation and effects of metals in aquatic systems. In: Metal Ecotoxicology: Concepts and Applications. Lewis Publishers, Boca Raton, FL. pp: 35-70.

Chen T., Wang T., Xue L., Brasseur G. (2024). Heatwave exacerbates air pollution in China through intertwined climate–energy–environment interactions. Science Bulletin.

Chorus I., Bartram J. (1999). Toxic cyanobacteria in water: a guide to their public health consequences, monitoring and management. World Health Organization.

Cortes C., Vapnik V. (1995). Support-vector networks. Machine learning, 20(3): 273-297.

Cover T., Hart P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1): 21-27.

Elbeltagi A., Mohamed E.S., Masoud M., Islam A.R.M.T. (2021). Groundwater quality assessment using machine learning algorithms and GIS in the north Sinai, Egypt. Journal of Hydrology, 597: 126190.

Evans M.I. (1994). Important bird areas in the Middle East. BirdLife International. 410 p.

Fadhil A.A., Swaid S.Y., Mohammed S.J., Al-Abboodi A. (2024). Impact of salinity on tomato seedling development: A comparative study of germination and growth dynamics in different cultivars. Journal of Chemical Health Risks, 14(1): 183-190.

FAO (Food and Agriculture Organization of the United Nations). (2016). Water for Sustainable Food and Agriculture. A report by the FAO.

Fawcett T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8): 861-874.

Gilliom R.J., Barbash J.E., Crawford C.G., Hamilton P.A., Martin J.D., Nakagaki N., Wolock D.M. (2006). The quality of our nation's waters: pesticides in the nation's streams and ground water. US Geological Survey.

Guyon I., Weston J., Barnhill S., Vapnik V. (2002). Gene selection for cancer classification using support vector machines. Machine Learning, 46: 389-422.

Hamza A.R., Ali H.A.A., Jabbar H.W. (2024). Environmental pollution and its impact on economic growth in Iraq (an applied study on the reality of the Iraqi economy). International Journal of Business and Management Sciences, 4(05): 91-106.

Hassan S.A.H., Al-Furiji H., Rashid M.K., Hussein Z.A., Ambudkar B. (2024). Trending Algorithm on Twitter through 2023. Data and Metadata, 3: 384-384.

Hassan S.A., Almaliki M.N., Hussein Z.A., Albehadili H.M., Banoon S.R., Al-Abboodi A., Al-Saady M. (2023). Development of nanotechnology by artificial intelligence: a comprehensive review. Journal of Nanostructures, 13(4): 915-932.

Iraq Ministry of Planning. (2020). Annual statistical abstract.

Jensen J.R. (2015). Remote sensing of the environment: an earth resource perspective. Pearson Education, Inc.

Jordan M.I., Mitchell T.M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245): 255-260.

Karr J.R. (1981). Assessment of biotic integrity using fish communities. Fisheries, 6(6): 21-27.

Kirk J.T.O. (1994). Light and photosynthesis in aquatic ecosystems. Cambridge University Press. 528 p.

Lorenz F., Erickson E.J. (2023). Strategic water: Iraq and security planning in the Euphrates-Tigris Basin. *dtic.mil.*

Madhi A.K., Azeez N.M. (2024). A review of water quality and pollution assessment of the Euphrates River. Science.

Mahmud S.F. (2021). Opportunities and challenges of sustainable agricultural development in Iraq. International Journal of Social Relevance and Concern, 9(1): 12-24

Moriasi D.N., Arnold J.G., Van Liew M.W., Bingner R.L., Harmel R.D., Veith T.L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE, 50(3): 885-900.

Mueller A., Detges A., Pohl B., Reuter M.H., Rochowski L., Volkholz J., Woertz E. (2021). Climate change, water and future cooperation and development in the Euphrates-Tigris basin. Research Gate / Geoscience/ Report, 38-53.

Mullen C.J., Grineski S.E., Collins T.W., Flores A.B. (2024). Air quality sensors and distributional environmental justice: a case study of Salt Lake County, Utah. Environmental Sociology, 10(2): 179-191.

Nazari A., Karbasi M., Abdollahi H. (2021). A review on application of machine learning in water resources engineering. International Journal of Environmental Science and Technology, 18(5): 1191-1206.

Oleiwi A.S., Al-Dabbas M. (2022). Assessment of contamination along the Tigris River from Tharthar-Tigris canal to Azizziyah, middle of Iraq. Water, 14(23): 3862.

Olson K.R., Speidel D.R. (2024). Tigris, Euphrates, and Shatt Al-Arab River System: Historic and Modern Attempts to Manage and Restore Iraq's Lifeline. Open Journal of Soil Science, 14(1): 1.

Plafkin J.L., Barbour M.T., Porter K.D., Gross S.K., Hughes R.M. (1989). Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. US Environmental Protection Agency.

Pongpiachan S., Hirisajja J., Aekakkararungroj A., Gupta P., Rungsiyanon S., Choochuay C., Poshyachinda S. (2024). Impacts of Meteorological Parameters on COVID-19 Transmission Trends in the Central Part of Thailand. Aerosol Science and Engineering, 1-14.

Proshad R., Islam S., Tusher T.R., Zhang D., Khadka, S., Gao J., Kundu S. (2021). Appraisal of heavy metal toxicity in surface water with human health risk by a novel approach: a study on an urban river in vicinity to industrial areas of Bangladesh. Toxin Reviews, 40(4): 803-819.

Rahi K.A., Halihan T. (2021). Surface water salinity of the Euphrates, Tigris, and Shatt al-Arab Rivers. In: Tigris and Euphrates Rivers: Their Environment from Headwaters to Mouth. Springer, Cham. pp: 309-336.

Rash A., Mustafa Y., Hamad R. (2023). Quantitative assessment of land use/land cover changes in a developing region using machine learning algorithms: A case study in the Kurdistan Region, Iraq. Heliyon, 9(11).

Reynolds J.B. (1996). Electrofishing. American Fisheries Society.

Rumelhart D.E., Hinton G.E., Williams R.J. (1986). Learning representations by back-propagating errors. Nature, 323(6088): 533-536.