Marine Pollution Bulletin xxx (xxxx) 114107

Contents lists available at ScienceDirect

Marine Pollution Bulletin

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

Hydrogeochemical characterization based water resources vulnerability assessment in India's first Ramsar site of Chilka lake

Dipankar Ruidas[a], Subodh Chandra Pal[a,*], Asish Saha[a], Indrajit Chowdhuri[a], Manisa Shit[b]

[a] Department of Geography, The University of Burdwan, Bardhaman, West Bengal 713104, India
[b] Department of Geography, Raiganj University, Raiganj, Uttar Dinajpur, West Bengal 733134, India

ARTICLE INFO

Keywords:
Chilka lake
Ramsar site
Machine learning
Anthropogenic stress
Vulnerability

ABSTRACT

A limnological site is significantly characterized by rich biological, chemical, and physical properties of the environment and is also described as the epitome of a large aquatic ecosystem. During the last few decades, the Chilka lake Ramsar site has experienced substantial degradation of water quality with associated deterioration of aquatic biodiversity. Our study aims to quantify the VWRM of the Chilka lake Ramsar region using the most reliable MLAs, namely ANN and RF, with the help of seventeen hydro-chemical properties of lake water. The produced map is validated through six validating measures (ROC-AUC- 0.89, Sensitivity-0.90, Specificity-0.78, PPV-0.78, NPV-0.88, Taylor diagram (r)-0.94), which depict that ANN is the most reliable ML algorithm in assessing the VWRM of the concerned region followed by RF. The prepared map of our study revealed that the eastern part was remarkably high to very high vulnerable zone covered area with 22.41 % and 7.19 %, respectively.

1. Introduction

Water is one of the crucial renewable natural resources; on which all life forms are based. The water resource is present on earth in three forms (gaseous, liquid, and solid). According to Kumar et al. (2005) and Ruidas et al. (2021), it is a precious gift to all forms of life by nature that is significantly expensive to transport, difficult to de-pollute, and unthinkable to substitute its usage. Generally, liquid water is present on the earth at two important places: at several depressions on the earth's surface as surface water forms of river water, and lake water, wetland, reservoirs, etc., and another one at the underline layer as groundwater. Surface water is not only a source of required water but also provides the habitat of several animal and plant species, making a distinctive ecological diversity in a region. Among all types of surface, water lakes are very significant from an ecological point of view, and their study falls under the umbrella of the field of "limnology" (Bhateria and Jain, 2016a); which may contain salt water as well as fresh water, and may be deep or shallow, and the inland nature of it gives no direct access to the ocean significantly. The ecosystem of lake water substantially depends on several physical, biological and chemical properties of available water and defines the ecological settings prominently. But in the last few decades, the water resource system has significantly come under vulnerable conditions due to many stressors, climatic variability, and increasing demand for it (Anandhi and Kannan, 2018). Noori et al. (2021b) asses how the quality of Sabalan lake water is experiencing rapid deterioration due to several anthropogenic factors by introducing significant pollution load in lake water which negatively impacts the ecological condition of this lake. Similarly, several scholars such as Bhat and Pandit (2014) in Wular lake, Elsayed et al. (2021a), Gad et al. (2021) and Elsayed et al. (2021a, 2021b) assessed the surface water quality of Qaroun lake they showed the severity of the existing lake water. Along with this (Tanaka et al., 2013; Xu et al., 2017) estimated the water condition of Dianchi lake and Taihu lake in China with significant relation to land degradation in these regions. Therefore, it is established that negative anthropogenic activities, including land use (Mack et al., 2019), urbanization (Luo et al., 2020), and pollution load (Tian et al., 2019), have a significant influence on the lake's health in different parts of the world. In history, very little data has been available on global surface water ecosystems; therefore, the United Nations environmental program has employed the Earth Observation technologies and

Abbreviations: VWRM, Vulnerability of Water Resources Map; ROC-AUC, Receiver Operating Characteristics Area Under Curve; VWR, Vulnerability of Water Resources; IDW, Inverse Distance Weighted; MLA, Machine Learning Algorithm; ANN, Artificial Neural Network; PPV, Positive Predictive Values; NPV, Negative Predictive Values; MDA, Mean Decrease Accuracy; BIS, Bureau of Indian Standards; LULC, Land Use Land Cover; VIF, Variance Inflation Factor; TOL, Tolerance; ML, Machine Learning; MC, Multicollinearity; RF, Random Forest; TP, True Positive; TN, True Negative; FP, False Positive; FN, False Negative

* Corresponding author.
E-mail address: geo.subodh@gmail.com (S.C. Pal).

https://doi.org/10.1016/j.marpolbul.2022.114107
Received 5 June 2022; Received in revised form 30 August 2022; Accepted 2 September 2022
0025-326/© 20XX

found significant changes in fresh water ecosystems. Several researchers have studied 75,000 water bodies in 89 countries worldwide and estimated that >40 % of areas were severely polluted (Zhongming et al., 2021). According to Reid (2018), globally, around 785 million people have no access to clean water, specifically one person in ten people on the earth, and statistically, 800 children under five years of age have died every day due to contaminated and unsafe water. Henceforth, the adopted sustainable development goals related to the clean water supply of the United Nations get its prime concern throughout the world. Thus, the water resource vulnerability of surface water becomes urgent for many water bodies by controlling pollution sources and implementing proper protecting measures (Feng et al., 2021). Füssel and Klein (2006) state that the term 'vulnerability' is defined in several contradictory ways, and which Inter-Governmental Panel for Climate Change (IPCC) plays an important role (Hutton et al., 2011). It is nothing but a theoretical idea (Tonmoy et al., 2014) and later, this idea transformed into a methodology and became a decision-making method (Weaver et al., 2013). Fundamentally, vulnerability defines the state of being and the quality of any physical component exposed to the possibility of being susceptible. Henceforth, the key vulnerability assessment and development of mitigation strategies have become primary concerns in recent times (Parry et al., 2007). According to Shabbir and Ahmad (2016), different physical vulnerabilities can be observed, namely coastal zone, water resource, draught area, energy resource, etc., and their assessment is the only essential tool to mitigate them. At present, the vulnerability of water resources (VWR) has become a prime concern of several researchers to control the water quality and conserve the ecological balance in the environment because it is a result of a complex interaction of ecological and social systems (Anandhi and Kannan, 2018; Gain et al., 2013). Worldwide different water deteriorating factors such as the presence of heavy metals (Yilmaz et al., 2021), economic development (Geng and Yi, 2006), land use land cover (LULC) change (Li et al., 2021), the rapid growth of population (Ducey et al., 2018), urbanization (McGrane, 2016), sediment load (Park and Engel, 2016), and rapid increase of water demand create a vulnerable situation for water resources (Al-Kalbani et al., 2014). Various researchers (Carbajal-Hernández et al., 2013; Davis and Hirji, 2003; Hernández-Terrones et al., 2015) have tried to show the deteriorating nature of surface water and their adverse effect on ecosystems throughout the world. Therefore, the Ramsar convention was held in 1971 to conserve the surface water and their existing ecosystem and sustainable development of the entire biodiversity-rich region by considering it a Ramsar site. But in the last few decades maximum Ramsar site was also affected by several anthropogenic and natural stress (Jamal et al., 2022). In India, it has been found that water resources and the existing ecosystem of several important Ramsar sites are under threatening conditions (Mabwoga et al., 2010; Nisari and Sujatha, 2021; Singh et al., 2017, 2021). So, the anthropogenic pressure on the existing lotic ecosystem has caused not only degraded the present water quality (Murray et al., 2010) but also significantly affected the existing aquatic biodiversity that deteriorated the ecological balance (Sala et al., 2000; Sruthy and Ramasamy, 2017). Therefore, Soonthornrangsan and Lowry (2021), Mansour et al. (2021), and Gu et al. (2015) assessed how anthropogenic activities trigger the significant water resource vulnerability in great lakes in the USA and Qiandao lake region in China, respectively. Furthermore, Qin et al. (2020) show the Spatio-temporal changes in surface water resources since 2000 of entire China. Besides, nutrient richness in the surface water is also a prominent reason in making water resource vulnerability (Romshoo et al., 2011) and in alteration of biodiversity (Barica, 1978; Bowen and Valiela, 2001; Smith, 2003). Thus, Pellerin et al. (2016) developed nutrient monitoring sensors as emerging tools in water resource studies; that can play a notable role in water protection. Generally, the surrounding environment plays playing an important role in making almost all issues regarding water resource vulnerability and fragile natural systems (Filik Iscen et al., 2008). Thus, it is necessary to estimate the water resource vulnerability in a region to overcome the aforementioned environmental problems. Therefore, several studies have been done on surface water quality but very few studies have taken place regarding the VWR and its relation to the deteriorating nature of biodiversity in a region due to available nutrient enrichment and immense anthropogenic stress. Although proper water analysis can play a vital role to delineate the severely affected areas; VWRM can be an important assessment tool that can help to correlate derived hydro-chemical properties and anthropogenic pressure on a region.

So far, several researchers have employed various statistical methodological approaches to water resource studies, such as the multivariate statistical technique (Bui et al., 2020; Kazi et al., 2009; Shrestha and Kazama, 2007; Singh et al., 2004), water quality index (Olasoji et al., 2019), fuzzy method (Wang et al., 2014). McBride (2005) states that statistical techniques have a distinctive role in water resource vulnerability studies. But till now, widespread studies have been done in water studies using ML algorithm, which has more efficiency in this field and is emerging as a most capable and promising methodological approach in the recent decade. Most frequently used ML approach in the different area, i.e. support vector regression (Najafzadeh et al., 2021; Ruidas et al., 2021); artificial neural network (ANN) (Chen et al., 2020; May et al., 2008), bagging-boosting (Hong et al., 2018), and random forest (RF) (Chowdhuri et al., 2020) which are also significant applicability in water resource vulnerability assessment due to their reliability and accuracy level. Among them, in this research work, we have employed two crucial machine learning algorithms (MLAs), namely RF and ANN, to develop VWRM of the Chilka Ramsar site, which is a quite first initiative to assess the role of hydro-chemical properties and anthropogenic pressure on water resource vulnerability and existing salt water biodiversity in this region. Thus, our adopted ML algorithms have a substantial contribution to prediction; the output of ANN is associated with discrete values, real values and represented as attribute-value pairs and makes the fast evaluation of target function; but it significantly depended on user ability and experience of trial and error which may consume little more time. Apart from all of this ML algorithm is characteristically robust in nature and minimal error cannot affect the final output. Similarly, RF also solve the classification and regression both problems and it automatically handles the missing data and deals with categorical and continuous data but the complexity and time-taking nature make some drawback. However, its robust capability, stability, and non-linear data handling efficiency make it noteworthy. In last few decades several scholars (Alizadeh and Kavianpour, 2015; Juahir et al., 2004; Palani et al., 2008a; Tiyasha et al., 2021) have been successfully applied ANN modelling approach in their water quality studies and get significant result in their objectives. Similarly (Irandoost et al., 2021; Kouadri et al., 2021) also achieved prominent outcomes in their water quality studies by using RF MLAs. Thus, these adopted modelling approaches are very significant in this research work due to their authenticity, accuracy, and precession level.

Chilka lake Ramsar site (1981) has enormous socio-economic and ecological importance, but this place gradually became vulnerable irrespective of its water resource and biodiversity. This region experienced massive threats due to rapid LULC change that pushed this ecologically diversified region into threatening conditions. Till now, very few studies have taken place to assess the VWR in this region, but among them, most of the works focus only on water quality based on chemical properties, but no work has been done on the development of VWRM. Therefore, our prime objective is to develop a novel tool to correlate the VWR with the degradation of biodiversity due to immense anthropogenic stress and water properties. In our study, we also adopted six validation measures, namely receiver operating characteristics area under the curve (ROC-AUC), sensitivity, specificity, positive predictive values (PPV), negative predictive values (NPV), and Taylor diagram to evaluate the result of the developed prediction.

Our present study has been organized in a sequential structure in which we have tried to display our research outcomes; the first section of this study notably describes the study area, including geographical, hydro-geological, climatic, and ecological conditions of this Ramsar site; the second section deals with the entire methodological framework of this study with adopted models, sampling techniques, and datasets; therefore, the third section is about the derived results and discussion and in the last segment of our study consist with the complete description on the importance of research outcomes in conclusion section that profoundly elaborate the future prospect of this study.

2. Materials and methods

2.1. Materials

2.1.1. Study area

Chilka Lake, the largest sub elliptical saltwater lake in India as well as in Asia, gets recognition as a Ramsar site by Ramsar wetland laws on the Ramsar convention in 1981; this lake is one of the largest tropical lagoons on the east coast of India and falls under category I marine protected areas. The absolute location of this lake is under 19° 28′ to 19° 54′ north latitude and 85° 06′ to 85° 35′ East longitude; the northern and western portion of this Ramsar site is surrounded by the Khordha and Ganjam district of Odisha respectively, a specific part of southern region fringed by Eastern Ghat range (Fig. 1). Basically, this inshore brackish water lake is shallow in nature with an average depth of 2 m situated with parallel to the coastline of the Bay of Bengal with a 65 km length and width of 20.1 km (Panigrahi et al., 2007). The Chilka lake experienced substantial changes in lake water depths in dry (0.3–0.8 m) and rainy seasons (1.8–4.2 m) throughout the year, and bathymetric survey in this site shows that the northern and southern portion of this site is characterized by <1.5 m and 3.9 m depth respectively (Ghosh and Pattnaik, 2005). Generally, this region is characterized by a tropical monsoonal climate with ranges from 14° to 39.9 °C temperatures and 1238 mm average rainfall; thus, Mohanty et al. (1997) state that this shallow water body experienced a very high evaporation rate. The monsoon (July–September) and post-monsoon (October–November) seasons contribute a significant amount of fresh water at the northern end of this lagoon through several large rivers that help to reduce the salinity level of this lake; The salinity gradient fluctuates because of a large amount of fresh water supply from several rivers, poor tidal influx, and weak circulation of lake water, the annual total amount of freshwater supply is 1.76 cubic meters among them 0.87 cubic meters come from direct precipitation on the lake and rest amount from the inland river system. Parallelly, the salinity level also increases due to the intrusion of saline water from the ocean; with the seasonal change, the area of this lake also changed i.e. where the area is 1165 km^2 during monsoon then it dried up in summer with 950 km^2 area (Siddiqui, 1995). The total catchment area is 4406 km^{2,} among them 68 % shared by the western part and 32 % with the Mahanadi delta; in addition, 55 % of the freshwater contributed to three important distributaries (Daya, Bahargavi, Nuna) of Mahanadi, whereas the rest of 45 % contributed by western catchment (Sarkar et al., 2012). The floral and faunal richness enlisted this highly productive Ramsar site as India's tentative world heritage site. The rich fishery resources help sustain the livelihood of many fishermen in this lake area. But in the last few decades, the environment as well as the water quality of the Chilka Ramsar site have been changed and fallen under severe stress due to abnormality in salinity, siltation, emergency in biodiversity, eutrophication etc. (Patnaik, 1998) estimated the sediment load in Chilka lake about 1.8 Million tonnes in 1998 that increase at 2.94 Million tonnes in 2001. The lack of management strategies pushes this Ramsar site into the stage of higher ecological stress.

Geologically, the coastline of western shores, north-eastern regions associated with the Pleistocene era, and the ephemeral characteristics of Chilka lake make this distinctive from surrounding areas. Basically, the entire catchment area of this Ramsar site is made up of rock, sand, and mud substratum. Silt, clay, sand, and gravel, are the significant sedimentary particles in this area, but the major part of this region contains silt; hydro-geologically, the aquifer is unconfined, interconnective, deep, and multi-layered in nature, with a lower recharge rate of 5–10 m3/h (Mishra and Dwibedy, 2015). The lake regions consist of several islands such as Parikud, Phulbari, Berahpura, Nuapara, Nalbana, and Tampara and they all are separated by shallow channels; those channels work as well connectors with the Bay of Bengal. A significant amount of siltation is responsible for closing the mouth of the lake with the sea by increasing the width of the barrier. It also helps in shifting of the mouth of the lake frequently toward the northeast direction.

2.1.2. Sampling techniques

The larger database is the foremost requirement for any novel modelling approach that helps in better performance and gives better prediction result. The ecological characteristics, circulation nature of water, and physiographic characteristics of our study area play a significant role in controlling water quality. Considering all the geohydrological factors, we collected 48 surface water samples from several locations along the shoreline of this lake during the dry season (January–February) of 2022. The sampling locations are selected by using a stratified sampling method based on their shallower depth, and the coastal position of sampling points can define the actual pollution level in lake water because most of the mixing regimes of river water and oceanic water are located in outline area of the Chilka lake Ramsar site. The Garmin GPS was used to make a record of the location of every individual water sample, and a unique ID number was also used for every water sample. In the collection of water samples, two different 500 ml polyethylene bottles were used and preserved by taking desirable sampling precautions and procedures to avoid any biotic activities as well as the exchange of gaseous substances (Bera et al., 2021; Pavlidou et al., 2014). Therefore, all the collected samples were speedily transported to the laboratory, and all samples were tested and analyzed by the laboratory to examine their chemical characteristics. Two separate containers were used to collect water samples, one for testing the chemical properties and another for cross-checking the result and validation. In this study, the Dionex ICS-90 Ion Chromatography and water checker U-10 method have been employed in examining the basic hydro-chemical properties of collected water samples. Several researchers (Islam et al., 2021a; Khan et al., 2013) have applied these two water analyzing techniques in their research work and found very significant results. Afterwards, all samples were examined in different laboratories to ensure the accuracy level. In our study, we have found seventeen hydro-chemical properties through laboratory testing namely, nitrate (NO_3) (mg/l), Fecal coliform (FC) (MPN/100 ml), mercury (Hg) (mg/l), fluoride (F^-) (mg/l), iron (Fe) (mg/l), biological oxygen demand (BOD) (mg/l), cadmium (Cd) (mg/l), pH, total coliform (TC) (MPN/100 ml), dissolve oxygen (DO) (mg/l), sodium absorption ratio (SAR), electrical conductivity (EC) (μS/cm), ammonia (NH_3) (mg/l), solids (S) (total), boron (B) (mg/l), lead (Pb) (mg/l), Turbidity (NTU); Table 1 displays the descriptive statistics of statistical measures of all considered water vulnerability parameters, which significantly describe the actual water quality condition by comparing the existing physiochemical properties values with national and international standards. In our study, the descriptive statistics (Table 1) show that several existing elements (NH_3, B, DO, Turbidity, EC, Pb, Fe) exceed their permissible values; such as NH_3 and B ranges from 0.57 to 1.11 (mg/L) and 0.06–1.74 (mg/L) whereas the permissible limit (BIS, 2012) 0.5 (mg/L). Similarly, whereas, the permissible limit of DO, Turbidity, EC, Pb, and Fe are 5 (mg/L), 1 (NTU), 1500 (μS/cm), 0.01 (mg/L), 0.3 (mg/L) accordingly, the higher values found of DO, Turbidity, EC, Pb, and Fe are 6.97 (mg/L), 4.21 (NTU), 14,950.1 (μS/cm), 0.09 (mg/L), 0.69 (mg/L) respectively which states that they significantly crossed their permissible limit and make an ad-

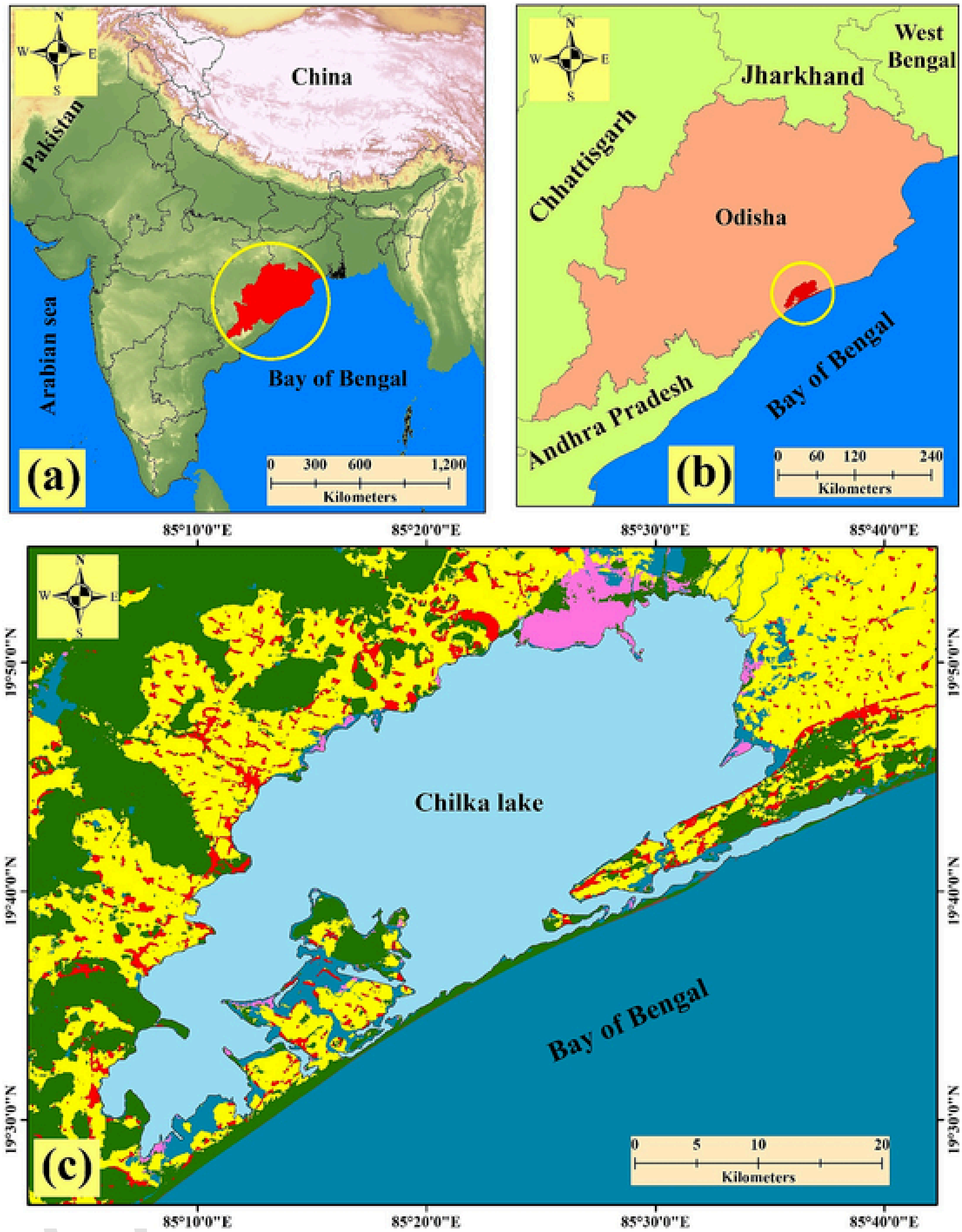


Fig. 1. Location map of study area.

verse effect on water condition. In this study, Fig. 2 shows the entire methodological framework which is adopted in the completion of the research work.

2.1.3. *Inventory data*

In the prediction of VWRM, inventory data making is a foremost concern. Therefore, it is developed in every prediction study that helps identify the possible vulnerable water resource zone by employing past records of vulnerability by combing all causative factors. Field study helps collect the past information regarding water resource vulnerability of our study regions that play an important role in making an inventory map of this region. From the collected 48 collected water datasets by using the stratified random sampling method, 24 sample points were selected for the water resource vulnerable regions, and conversely, 24 water samples regions were chosen as non-vulnerable regions by employing 1 and 0, respectively. Among collected 48 water samples, 34 (70 %) were considered for VWRM, and the rest 14 (30 %) were used for validating the result by using Indian water quality standards (BIS, 2012); that means one section of data was used in making predictions and developing the models and another section for testing the predicted result Chung and Fabbri (2008). Therefore, after examining the chemical properties of water samples, the inverse distance weighted (IDW) method (Bartier and Keller, 1996) has been used to develop the map of 17 causative factors to delineate the VWRM of the study area in the Ar-

Table 1
Descriptive statistics of hydro-chemical parameters of water samples collected from Chilka lake Ramsar site.

Factors	Maximum	Minimum	Average	Standard deviation
NH_3 (mg/l)	1.05	0.57	0.76	0.11
BOD (mg/l)	2.28	1.36	1.79	0.23
B (mg/l)	1.74	0.068	0.25	0.28
Cd (mg/l)	0.0028	0.0008	0.0016	0.0004
DO (mg/l)	6.97	2.41	5.44	0.87
EC (μS/cm)	14,950.1	479.36	3046.12	2430.3
FC (MPN/100 ml)	2141.41	147.95	1089.77	573.19
F (mg/l)	0.31	0.07	0.16	0.06
Turbidity (NTU)	4.21	0.7	1.49	0.58
TC (MPN/100 ml)	1826.41	217.27	707.56	397.56
Solids	62.18	9.87	27.03	14.09
SAR	29.53	1.06	7.35	8.22
pH	8.23	6.98	7.64	0.3
NO_3 (mg/l)	0.26	0.13	0.17	0.03
Hg (mg/l)	0.00065	0.00021	0.00045	0.00011
Pb (mg/l)	0.099	0.003	0.024	0.028
Fe (mg/l)	0.69	0.02	0.32	0.21

cGIS 10.4 platform. Hence, in the present research, two important ML algorithms, namely, ANN and RF were employed to delineate the VWRM with reliable validation techniques.

2.1.4. VWR conditioning factors

The water resource vulnerability is specifically influenced by adverse hydro-chemical properties that can affect the entire aquatic ecosystem and are sufficient to collapse the balance of biodiversity in a region by crossing their permissible limit, specified by the Bureau of Indian Standards (BIS, 2012). Thus, VWRM emerges as a necessary tool to delineate and quantify the most vulnerable zone in our study area. Therefore, in our study, we have considered seventeen water resource vulnerable conditioning factors including NO_3 (mg/l), FC (MPN/100 ml), Hg (mg/l), F^- (mg/l), Fe (mg/l), BOD (mg/l), Cd (mg/l), pH, TC (MPN/100 ml), DO (mg/l), SAR, EC (μS/cm), NH_3 (mg/l), S (total), B (mg/l), Pb (mg/l), Turbidity (NTU) to make VWRM and LULC change map of two time period (2005, 2022) to assess Spatio-temporal change and the anthropogenic stress on Chilka Ramsar site considering extensive literatures (Jaydhar et al., 2022; Pal et al., 2022b, 2022c; Ruidas et al., 2022b). The identification and quantification were taken place through water testing in the laboratory, and the IDW interpolation method was applied to make the spatial distribution map of each parameter (Fig. 3); this method is very helpful in delineating the proper distributional pattern of each conditioning factors (Ruidas and Pal, 2022; Saha et al., 2021b). The map shows, that Chilka lake has a diverse distribution pattern of all identified factors i.e. the significant portion of the central and western region near Sabulia and Rambha are characterized by high NH_3 and ranging from 0.57 to 1.11 (mg/l); about 1.33–2.29 (mg/l) BOD level is present in this region and northwestern (Balugaon, Barkul) and western part (Malakuda, Kumarpur) characterized with high BOD level. The east-central region of this lake dominated by B and EC ranged by 0.06–1.99 (mg/l) and 315.54–17,128.1 (μS/cm), respectively. In addition, high Cd concentration is found in the eastern part near Subhadrapur, Jaganathpur, and Gagadala; DO is distributed more or less in a similar pattern throughout the entire lake. FC has a diverse distributional pattern, and the west-central part is characterized by high concentration, which is quite similar to F^-.

In contrast, the noteworthy portion of the study area was dominated by high Fe ranging from 0.007 to 0.70 (mg/l). In addition, Pb, Hg, SAR, solids, and TC have varied distribution, and the central portion is delineated as a high NO_3, pH and Turbidity region (Fig. 3). Generally, the geographical distribution pattern of VWR conditioning factors helps to understand the influence of water resource vulnerability concerning the degradation of biodiversity.

2.2. Methods

2.2.1. Multi-collinearity test

Multicollinearity (MC) shows the presence of interdependence of two or more independent factors (Saha et al., 2021c; Talukdar et al., 2021), and high linearity has a significant role in reducing of accuracy and precision level of derived results (Ruidas et al., 2022a) which also resulting various regression problems in a modelling approach. The MC test helps identify those factors with significant linear dependency (Saha et al., 2021a). Generally, to survey MC analysis, three popularly used methods such as variance inflation factor (VIF) (Pal et al., 2022b; Saha et al., 2022b), tolerance (TOL), and Pearson's correlation coefficients (r^2 = < _ +0.7) are used. TOL is the ratio between the relative scatter and the scatter of that variable. Several scholars (Chowdhuri et

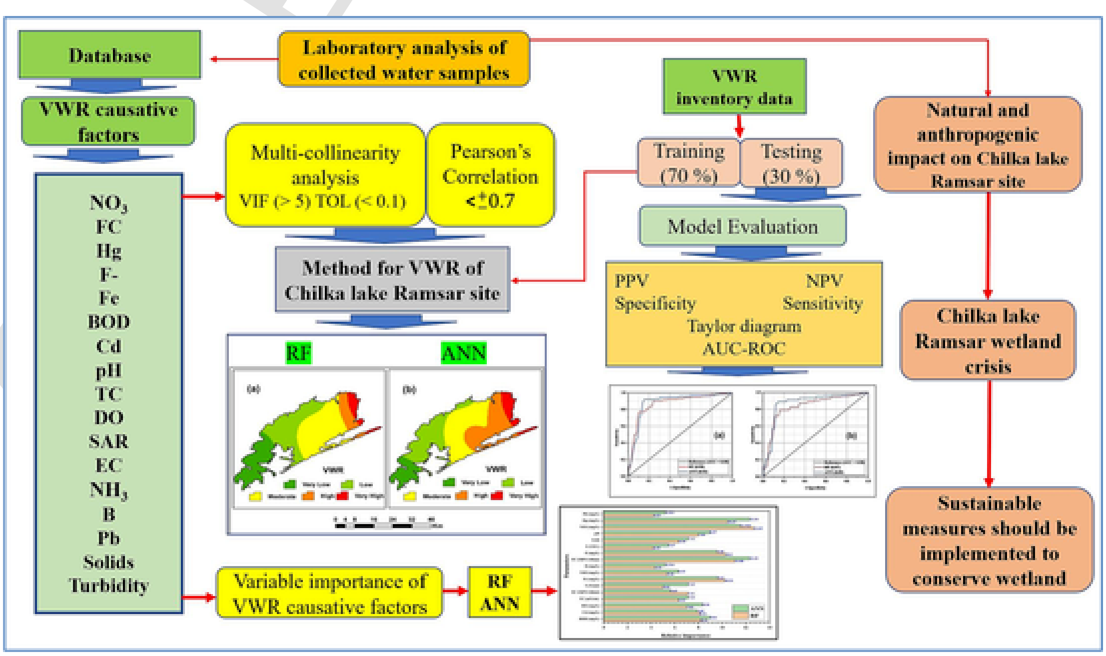
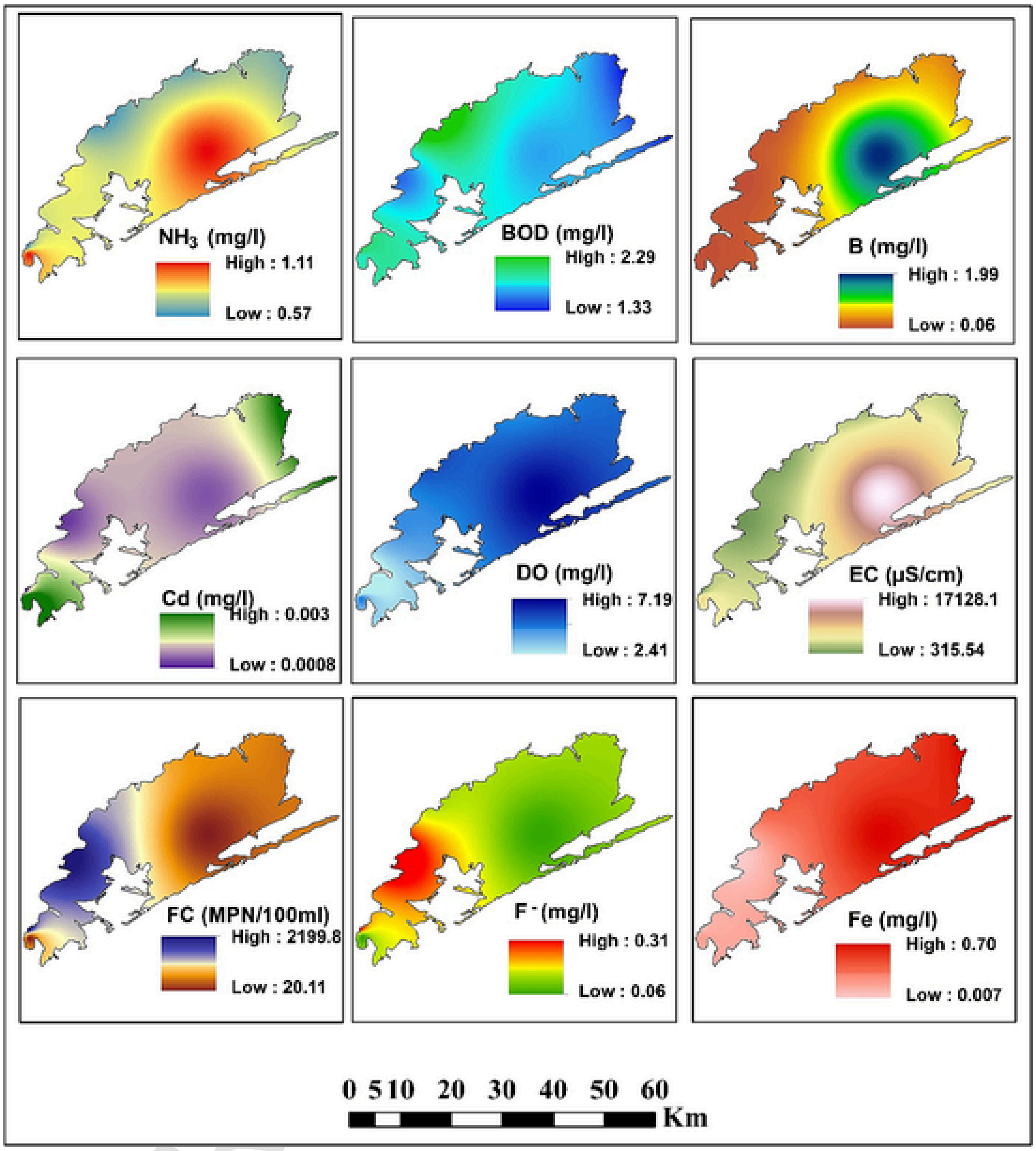


Fig. 2. Methodological framework of the current study.



(*caption on next page*)

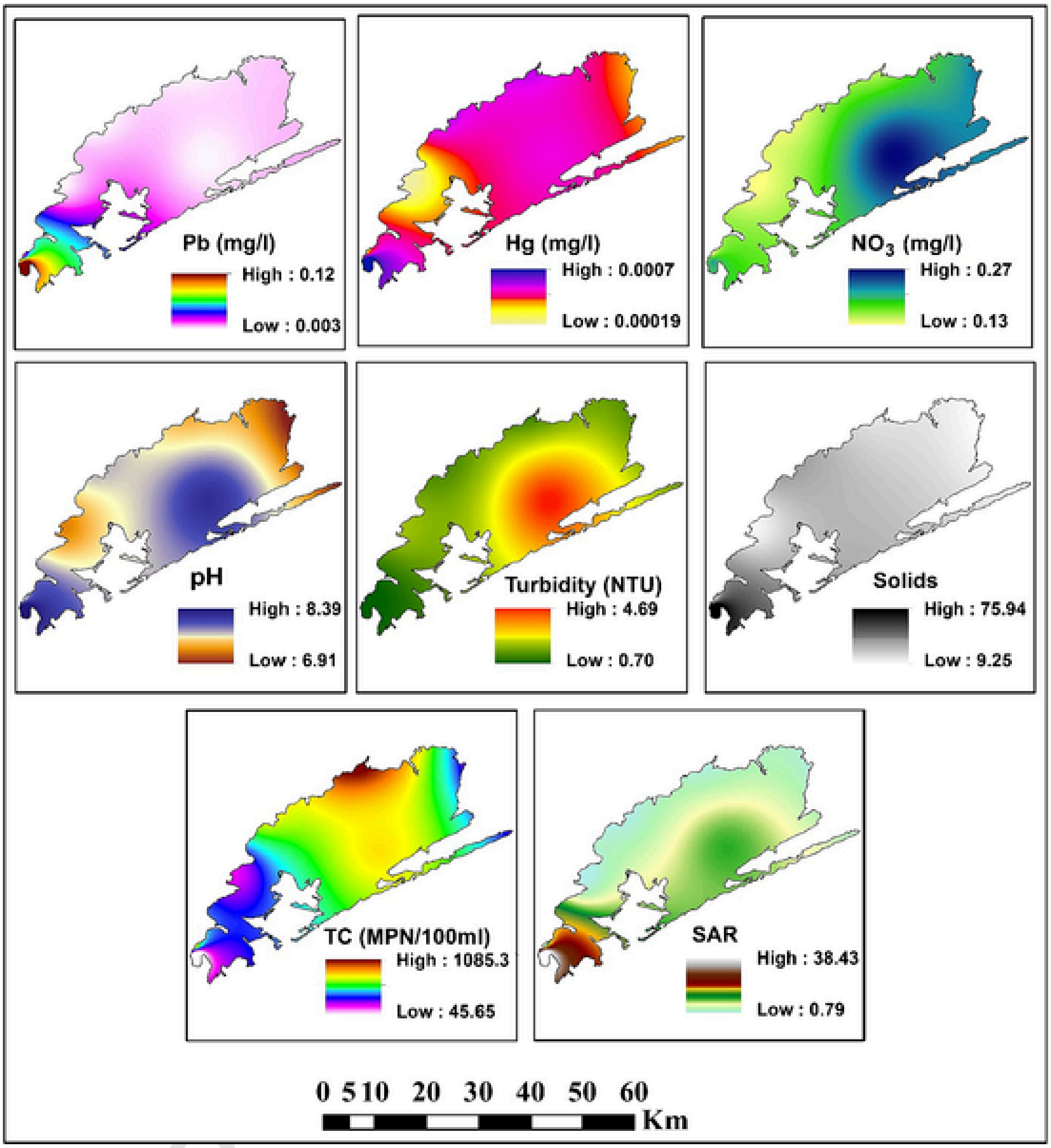


Fig. 3. VWRM conditioning factors.

al., 2021; Pal et al., 2022c; Saha et al., 2022a; Talukdar et al., 2020) established that the presence of factors with VIF value >5 and TOL value <0.10 have significant multicollinearity issues. The acceptable result should be determined by <5 and >0.10 values of VIF and TOL, respectively. The following equations are used in determining MC:

$$\mathrm{TOL} = 1 - r^2 \tag{1}$$

$$\mathrm{VIF} = \frac{1}{\mathrm{TOL}} \tag{2}$$

"Here r^2 represents the coefficient of regression"

2.2.2. *VWRM modelling approach*

Currently, VWR is a significant issue; but timely information about water quality and proper policy can improve this vulnerable condition. Several physicochemical characteristics of collected water samples help to develop the VWRM of this region. In our present study, we have adopted two frequently used predictive ML techniques: ANN and RF algorithms. These two techniques were performed in R programming software with the help of ArcGIS 10.4 software.

2.2.3. *Artificial neural network (ANN)*

ANN is a popular computing system quite similar to a biological neural network containing parallel nonlinear units comprising many interconnected processing elements (Sengorur et al., 2015). According to (Palani et al., 2008b), this began with backpropagation training algorithms (Rumelhart et al., 1986) that are based on the artificial neuron concept (McCulloch and Pitts, 1943). This nonlinear statistical technique solves impossible problems using conventional mathematical and statistical methods. ANN's underlying data generating process helps to a wide application in prediction, classification, forecasting, data processing, and many other purposes (Wu et al., 2014). In the last few years, the use of ANN modelling is increased significantly in the field of hydrological studies due to its significant accuracy and precision level (ASCE, 2000). (Dogan et al., 2009) also shows how the trained particular set of inputs produces a possible and specific set of target outputs; ANN algorithm is characterized by three important characteristics such as nodes, connection strength, and transfer function, which commonly follow three important stages including data collection, processing the data and training of the neural network. Usually, a large amount of historical data is needed to perform ANN algorithms and to get a satisfactory result. The schematic diagram of ANN is presented in Fig. 4. The details about the ANN algorithm can found in the link provided in Supplementary material. The ANN algorithm can be elaborated by the following equations (Hagan et al., 1997).

$$\mathrm{net}_j^l(t) = \sum_{i=o}^{p} \left(y_i^{i-1}(t)\, w_{ji}^l(t) \right) \tag{3}$$

The net input of jth neuron of layer l and I iteration

$$y_j^l(t) = f\left(\mathrm{net}_j^{(l)}(t)\right. \tag{4}$$

$$f(\mathrm{net}) = \frac{1}{1 + e^{(-\mathrm{net})}} \tag{5}$$

$$e_j(t) = c_j(t) - a_j(t) \tag{6}$$

$$\delta_j^l(t) = e_j^l(t)\, a_j(t) \left[1 - a_j x(t)\right] \tag{7}$$

δ factor for neuron jth in the output layer ith

$$\delta_j^l(t) = y_j^l(t) \left[1 - y_j(t)\right] \sum \delta_j^l(t)\, w_{kj}^{(l+1)}(t) \tag{8}$$

δ factor for neuron jth in the hidden layer ith

$$\begin{aligned} w_{ji}^l(t+1) = {} & w_{ji}^l(t) + \alpha \left[w_{ji}^l(t) - w_{ji}^l(t-1) \right] \\ & + n\delta_j^{(l)}(t)\, y_j^{(l-1)}(t) \end{aligned} \tag{9}$$

"where α represents the momentum rate and n is the learning rate".

2.2.4. Random forest (RF)

RF machine learning algorithms are the expression of the arrangement of tree predictors that was introduced by Breiman (2001). Basically, this algorithm separated two types of trees, namely regression and classification trees from the decision trees (Rodriguez-Galiano et al., 2014); that's why this artificial ML algorithm has been employed in unsupervised learning and in regression classification (Liaw and Wiener, 2002). This is also an advanced version of the bagging and ensemble data mining algorithm (Rahman et al., 2019). (Belgiu and Drăguţ, 2016) state that this RF classifier is significantly based on two parameters as the number of variables and decision trees. According to (Chakrabortty et al., 2021; Islam et al., 2021b; Ok et al., 2012; Pal et al., 2022a), this flexible ML algorithm helps in constructing a set of classifiers instead of one classifier to develop similar predictions from similar circumstances by producing new training site through selecting sample site. In the case of regression trees, there has a substantial outfitting problem in the training dataset and therefore, RF have the ability to overcome these shortcomings (Criminisi and Shotton, 2013) and also another such precise method in reducing variance instead of increasing bias (Hastie et al., 2001); it is also a significant method in examining the hierarchical interactions and non-linearities in the large data frame (Olden et al., 2008). The following equations represent the RF model.

$$\log 2\,(M + 1) \tag{10}$$

"where, M is the input number of an algorithm, and generalization i.e., the mean square error is represented by":

$$\varepsilon = \left(v_{\mathrm{observed}} - v_{\mathrm{response}}\right)^2 \tag{11}$$

"where, $v_{observed}$ is the variable from observed data and $v_{response}$ is the variable from the output result in this model. Lastly, RF algorithm can be expressed by":

$$S = \frac{1}{K} \sum K^{\mathrm{th}} v_{\mathrm{response}} \tag{12}$$

"where S represents the future prediction result and K is the individual trees".

2.2.5. Variable importance

Several suitable water resource vulnerable causative factors were used to develop VWRM, but all parameters do not have equal responsibility in water resource vulnerability in a region. Thus, it is compulsory to quantify the relative importance of each adopted causative factor to make a reliable assessment of water resource vulnerability. In our study, we have employed an RF machine learning algorithm to quantify the relative importance of each conditioning factor. Conventional statistical and mathematical techniques are not proficient in handling large datasets. Thus, in this research, we have applied the mean decrease accuracy (MDA) index by using RF (Breiman, 2001) and ANN machine learning algorithm to assess the relative significance of adopted factors. The following equation represents the MDA method.

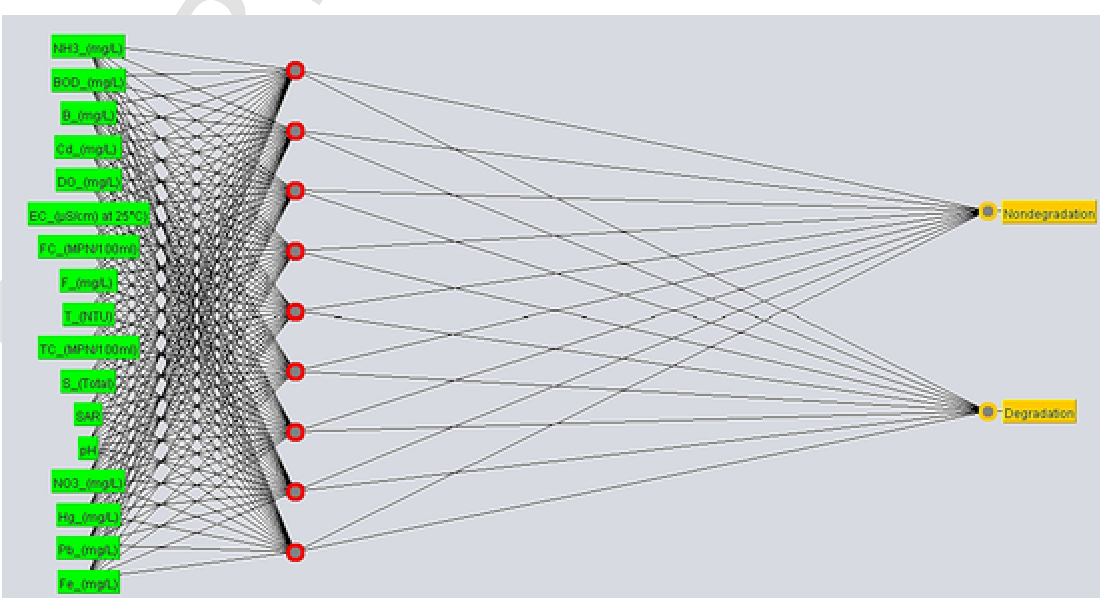


Fig. 4. Schematic diagram of ANN model.

$$\mathrm{VI_j} = \frac{1}{\mathrm{ntree}} \sum_{t=1}^{\mathrm{ntree}} \mathrm{EP_{tj}} - \mathrm{E_{tj}} \tag{13}$$

"where, *VI* represents variables importance, E_{tj} indicates OOB error on tree *t* before permuting the values of X_j and EP_{tj} indicates OOB error on tree *t* after permuting the values of X_j".

2.2.6. Model validation

Model validation is an integral part of research work that can help to quantify the accuracy level and increase the reliability of produced results; it can relate the result to the real world. The derived results have no significance if it is not gone through the model evaluation stage. Therefore, in our present study, the two adopted models (ANN, RF) were evaluated by six validation techniques among them, five are statistical techniques such as sensitivity, specificity, PPV, NPV, receiver operating characteristics curve- area under the curve (ROC-AUC) and one graphical presentation namely Taylor diagram (Taylor, 2001). According to (Chen et al., 2019), in the AUC-ROC technique AUC value varies from 0.5 to 1, and the closer value toward the 1 and 0.5 values implies the higher and lower reliability of the model, respectively. It represents two axes, the x-axis (sensitivity) and y-axis (specificity) and shows its success value. The quality of output results is also quantified by measuring the four indices as true positive (TP), true negative (TN), false positive (FP), and false-negative (FN), which are all used to perform two important validation techniques sensitivity and specificity (Westerhuis et al., 2008). Khosravi et al. (2019) state that the higher values of the aforementioned validation techniques imply the healthier result of the adopted modelling approach and vice versa. The aforementioned techniques were performed in this study by using the following formulas:

$$\mathrm{Specificity} = \frac{\mathrm{TN}}{\mathrm{FP} + \mathrm{TN}} \tag{14}$$

$$\mathrm{Sensitivity} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{15}$$

$$\mathrm{PPV} = \frac{\mathrm{TP}}{\mathrm{FP} + \mathrm{TP}} \tag{16}$$

$$\mathrm{NPV} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FN}} \tag{17}$$

$$\mathrm{AUC} = \sum_{k=1}^{n} \left(\mathrm{X_{k+1}} - \mathrm{X_k}\right)\left(\mathrm{S_{k+1}} - \mathrm{S_k} - \mathrm{S_k}/2\right)\ (18)$$

$$\mathrm{AUC} = \sum_{k=1}^{n} \left(\mathrm{X_{k+1}} - \mathrm{X_k}\right)\left(\mathrm{S_{k+1}} - \mathrm{S_k}/2\right) \tag{18}$$

"Here, specificity and sensitivity are denoted by X_k and S_k respectively".

3. Results

3.1. Multicollinearity test

In our present study, before performing any modelling approach, the MC test was carried out by three aforementioned statistical techniques to quantify the interdependence of each factor (Table 2). Seventeen VWR causative factors were considered after the MC evaluation stage. The result shows that the lowest TOL value and highest VIF values are 0.212 and 4.72, respectively and that all the TOL and VIF values are under the acceptable limits. Thus, it implies no collinearity issues among all 17 VWR conditioning factors. Another adopted statistical technique, Pearson's correlation matrix, shows that several causative factors are slightly correlated (Table 3). The result shows that the r^2 value of EC and Boron, FC and F^-, Boron and Turbidity, Solids and SAR, Solids and pH, Solids and Pb, SAR and Pb were 0.804, 0.906, 0.817, 0.911, 0.925, 0.916, 0.992 respectively but rest of the values are in the considerable limit.

Table 2
Result of multicollinearity test.

SL no.	Factor	Collinearity statistics	
		Tolerance	VIF
1	BOD (mg/L)	0.282	3.551
2	Cd (mg/L)	0.212	4.720
3	DO (mg/L)	0.389	2.568
4	EC (μS/cm)	0.879	1.138
5	TC (MPN/100 ml)	0.634	1.577
6	S (Total)	0.239	4.177
7	Fe (mg/L)	0.277	3.610
8	NH_3 (mg/L)	0.245	4.082
9	B (mg/L)	0.286	3.497
10	FC (MPN/100 ml)	0.257	3.891
11	F (mg/L)	0.433	2.309
12	T (NTU)	0.542	1.845
13	SAR	0.344	2.907
14	pH	0.374	2.674
15	NO_3 (mg/L)	0.486	2.058
16	Hg (mg/L)	0.452	2.212
17	Pb (mg/L)	0.274	3.650

3.2. Vulnerability of water resource map prediction

The available and aforesaid data established that the water quality of the Chilka lake Ramsar site is under significantly vulnerable conditions and makes this an ecologically vulnerable place. Therefore, the water resource vulnerability measurement becomes a foremost concern. Thus, VWRM is an essential tool for evaluating available Chilka lake water quality and vulnerability. In our present study, VWRM was developed using two distinctive ML algorithms, ANN and RF, because proper modelling plays a significant role in quantifying and identifying vulnerable water resource regions. So, the VWRM is predicted at the Chilka lake Ramsar site of Odisha by combining 17 causative factors. The entire Ramsar site is characterized by the diverse distribution pattern of all conditioning factors in the monitoring year of 2022 which is mentioned in Fig. 3. The predicted VWRM was developed by ANN and RF, which classified the entire Ramsar site into five VWR zone based on their magnitude of vulnerability: very low, low, moderate, high, and very high, vulnerable zone. Both the predictions display similar approximate results (Fig. 5), in which the western part of this Ramsar site is characterized by a very low water resource vulnerability zone, surrounded by Arakhakuda, Malakuda, Prayagi, Laxmanpur, Rambha, Sabulia, Samanasi, Barkul region. The Eastern part, surrounded by Karatiasahi, Sitarampur, Jaganathpur, Barikanupada, Basanapur, Chandumal, and Gopinathpur regions, falls under a high to a very high vulnerable zone. The five classes have different areal extensions in each model; if the total Ramsar site is considered as 100 %, then the produced result of the ANN model (Fig. 5a) shows the diverse class-wise spatial distribution that accounts for 120.22 km^2 (10.32 %), 297.30 km^2 (25.52 %), 400.99 km^2 (34.42 %), 261.07 km^2 (22.41 %) and 83.76 km^2 (7.19 %) area are under very low, low, moderate, high and very high category accordingly; In RF (Fig. 5b) 240.57 km^2 (20.65 %), 342.27 km^2 (29.38 %), 423.01 km^2 (36.31 %), 98.20 km^2 (8.43 %) and 60.69 km^2 (5.21 %) area under very low, low, moderate, high and very high category respectively. Both results show that the maximum area of the Chilka Ramsar site is under a moderately vulnerable zone which covers mainly the central and east-central parts of this study area.

3.3. Assessment of anthropogenic stress on Chilka Ramsar site (2005–2022)

In the present decade, the RS-GIS technology has emerged as a significant tool that remarkably imparts a chance to evaluate and interpret surface water quality. Several researchers (Hossain et al., 2014; Milano

Table 3
Pearson's correlation coefficient between pairs of VWRM conditioning factors.

	NH_3 (mg/L)	BOD (mg/L)	B_(mg/L)	Cd (mg/L)	DO (mg/L)	EC (µS/cm)	FC (MPN/100ml)	F (mg/L)	T (NTU)	TC (MPN/100ml)	S (Total)	SAR	pH	NO_3 (mg/L)	Hg (mg/L)	Pb (mg/L)	Fe (mg/L)
NH_3 (mg/L)	1.000	-0.318	0.176	0.366	-0.235	0.612	-0.013	0.145	0.162	-0.611	0.540	0.736	0.528	0.333	0.062	0.683	-0.481
BOD (mg/L)		1.000	-0.340	-0.189	-0.074	-0.273	0.217	-0.194	-0.401	0.057	0.472	0.081	0.590	-0.542	0.550	0.108	-0.127
B (mg/L)			1.000	-0.043	0.485	0.804	-0.577	-0.474	0.817	0.329	-0.343	-0.200	-0.158	0.682	0.067	-0.316	0.631
Cd_(mg/L)				1.000	-0.154	0.400	-0.626	-0.539	-0.420	-0.271	0.557	0.721	0.281	0.657	0.535	0.709	0.133
DO (mg/L)					1.000	0.127	-0.352	-0.397	0.651	0.339	-0.457	-0.529	-0.447	0.206	-0.010	-0.598	0.693
EC (µS/cm)						1.000	-0.639	-0.530	0.470	-0.021	0.233	0.421	0.328	0.797	0.382	0.309	0.303
FC (MPN/100ml)							1.000	0.906	-0.124	-0.460	-0.030	-0.176	0.062	-0.805	-0.605	-0.100	-0.764
F (mg/L)								1.000	-0.019	-0.440	-0.184	-0.157	-0.141	-0.615	-0.792	-0.088	-0.765
T (NTU)									1.000	0.155	-0.568	-0.476	-0.369	0.320	-0.378	-0.571	0.399
TC (MPN/100ml)										1.000	-0.499	-0.528	-0.403	0.000	0.238	-0.555	0.647
S (Total)											1.000	0.911	0.925	-0.004	0.664	0.916	-0.482
SAR												1.000	0.789	0.267	0.538	0.992	-0.462
pH													1.000	-0.090	0.637	0.775	-0.485
NO_3 (mg/L)														1.000	0.230	0.186	0.559
Hg (mg/L)															1.000	0.501	0.248
Pb (mg/L)																1.000	-0.525
Fe (mg/L)																	1.000

	NH_3 (mg/L)	BOD (mg/L)	B_(mg/L)	Cd (mg/L)	DO (mg/L)	EC (µS/cm)	FC (MPN/100ml)	F (mg/L)	T (NTU)	TC (MPN/100ml)	S (Total)	SAR	pH	NO_3 (mg/L)	Hg (mg/L)	Pb (mg/L)	Fe (mg/L)
NH_3 (mg/L)	1.000	-0.318	0.176	0.366	-0.235	0.612	-0.013	0.145	0.162	-0.611	0.540	0.736	0.528	0.333	0.062	0.683	-0.481
BOD (mg/L)		1.000	-0.340	-0.189	-0.074	-0.273	0.217	-0.194	-0.401	0.057	0.472	0.081	0.590	-0.542	0.550	0.108	-0.127
B (mg/L)			1.000	-0.043	0.485	0.804	-0.577	-0.474	0.817	0.329	-0.343	-0.200	-0.158	0.682	0.067	-0.316	0.631
Cd_(mg/L)				1.000	-0.154	0.400	-0.626	-0.539	-0.420	-0.271	0.557	0.721	0.281	0.657	0.535	0.709	0.133
DO (mg/L)					1.000	0.127	-0.352	-0.397	0.651	0.339	-0.457	-0.529	-0.447	0.206	-0.010	-0.598	0.693
EC (µS/cm)						1.000	-0.639	-0.530	0.470	-0.021	0.233	0.421	0.328	0.797	0.382	0.309	0.303
FC (MPN/100ml)							1.000	0.906	-0.124	-0.460	-0.030	-0.176	0.062	-0.805	-0.605	-0.100	-0.764
F (mg/L)								1.000	-0.019	-0.440	-0.184	-0.157	-0.141	-0.615	-0.792	-0.088	-0.765
T (NTU)									1.000	0.155	-0.568	-0.476	-0.369	0.320	-0.378	-0.571	0.399
TC (MPN/100ml)										1.000	-0.499	-0.528	-0.403	0.000	0.238	-0.555	0.647
S (Total)											1.000	0.911	0.925	-0.004	0.664	0.916	-0.482
SAR												1.000	0.789	0.267	0.538	0.992	-0.462
pH													1.000	-0.090	0.637	0.775	-0.485
NO_3 (mg/L)														1.000	0.230	0.186	0.559
Hg (mg/L)															1.000	0.501	0.248
Pb (mg/L)																1.000	-0.525
Fe (mg/L)																	1.000

et al., 2015; Morrice et al., 2008; Wang et al., 2007) established that how anthropogenic stress makes spatiotemporal changes in surface water quality and also enhances the understanding of the positive relationship between both of them. In our present study, we have employed an environmental system research institute (ESRI) LULC map (2005 and 2022) of the study area to quantify vegetation, agricultural land, build-up area, and water body in the different periods; the spatiotemporal changes of these all parameters help to identify the anthropogenic stress on Chilka lake Ramsar region. According to Anderson (1976), LULC changes in a geographical region can both reduce and increase the environmental crisis. The saltwater ecosystem of the Chilka lake Ramsar site, including surface water, experienced a significant land al-

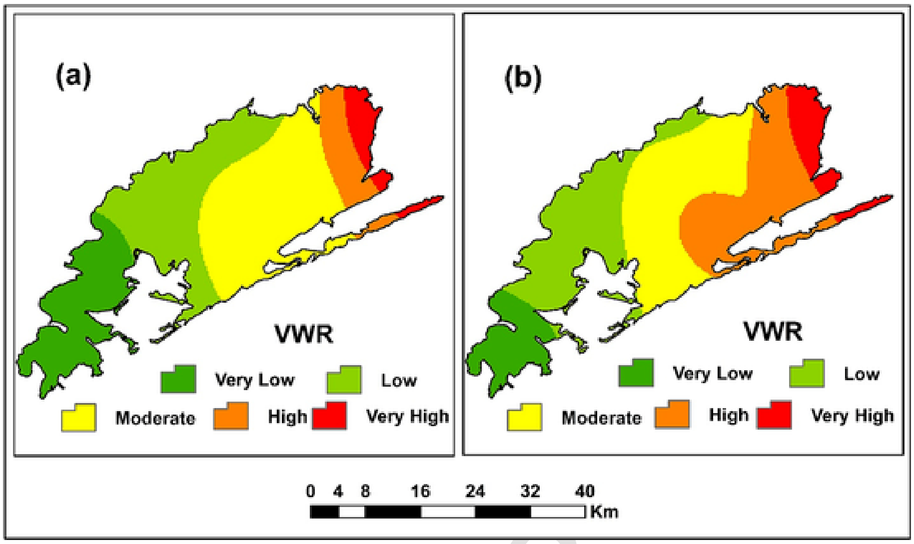


Fig. 5. VWR map of Chilka lake Ramsar site (a) RF method and (b) ANN method.

teration in the surrounding catchment area at an alarming rate which pushed this region into a threatening situation (Mishra, 2017; Panigrahi et al., 2007) and made the study area more vulnerable to water resource and ecological imbalance. The 2005 (Fig. 5a) and 2022 (Fig. 5b) ESRI LULC map show that the amount of agricultural land, settlement, vegetation cover, and build-up area significantly increased in the two time period that implying these changes brings remarkable consequences to the water quality of Chilka lake and make more vulnerable to water resource; which is shown in Fig. 6a and b.

3.4. Model performance evaluation

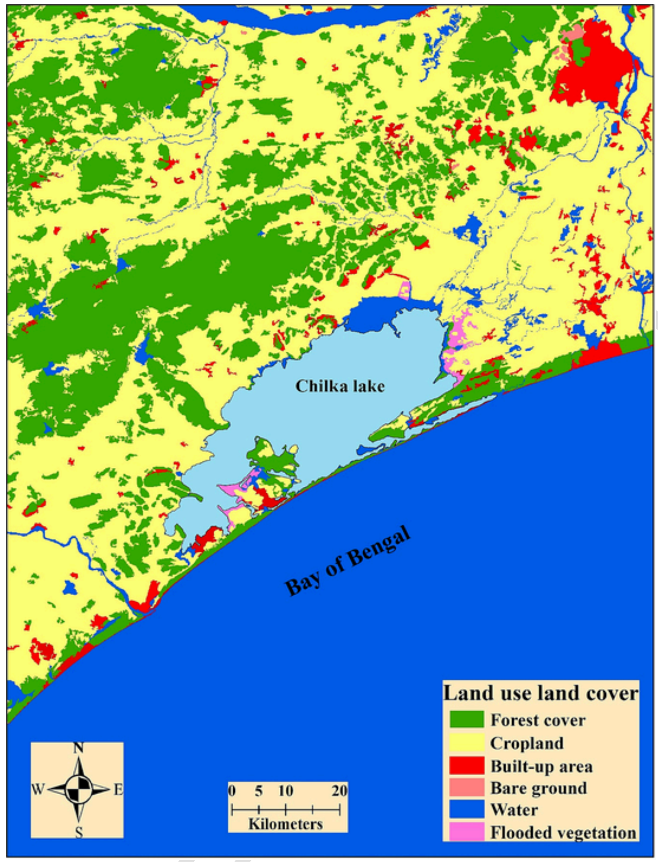
After applying all adopted modelling approaches by the establishment of training and testing dataset, the important procedure is to evaluate the model's performance in training and validating the dataset. This is necessary to ensure the closeness of predicted results and real-world situations. In our present study, several statistical techniques were used to validate the derived result, including specificity, sensitivity, PPV, NPV, Taylor diagram, and AUC-ROC. All the validation techniques show that both models play a significant role in analyzing, quantifying, and determining the VWR zone with an adequate accuracy level in this study area. Despite that, the model evaluation is mandatory due to its different data sources, which were remotely sensed and had some possibility of error. Fig. 7 represent the AUC-ROC analysis of the aforementioned two ML models, and it is clearly showing that both models have significant accuracy level in training (Fig. 6a) and validating stage (Fig. 6b); but based on the results, the ANN model has a high AUC value (training = 0.93, validating = 0.89) compare to RF modelling approach (training = 0.9, validating = 0.87) that implies ANN is the most suitable modelling approach than RF. Besides, the result of all other statistical techniques also shown in Table 4 which have remarkable performance values regarding both two adopted models; in RF, sensitivity (training = 0.86, validating = 0.85), specificity (training = 0.8, validating = 0.74), PPV (training = 0.82, validating = 0.78), NPV (training = 0.81, validating = 0.84) and in ANN, sensitivity (training = 0.92, validating = 0.90), specificity (training = 0.81, validating = 0.78), PPV (training = 0.85, validating = 0.78), NPV (training = 0.91, validating = 0.88). Apart from these, one important graphical validation technique, Taylor diagram (Fig. 8) shows RF (r = 0.92) and ANN (r = 0.94) gave significant performance. This clearly shows that ANN is the more reliable and preferable ML algorithm to develop the VWRM in this Chilka lake Ramsar site.

3.5. Variable importance of adopted factors

Two important ML algorithms (RF and ANN) were employed in this research work to assess and quantify the contribution in VWR of the aforesaid seventeen conditioning factors. Based on the results, these show the hierarchical position of each factor which has a significant role in making of VWRM of this region as is shown in Table 5 and Fig. 9. According to both results, the most (NO_3, FC, Hg, F^-, Fe) and least (NH_3, S, B, Pb, T) contributing factors among the adopted 17 factors in water resource vulnerability are the same; only have slight variability in their position. In RF method, the hierarchical order of adopted factor is
NO_3 > FC > Hg > F^- > Fe > BOD > Cd > pH > TC > DO > SAR > EC > NH_3 > S > B
whereas the highest influence of NO_3 (12.69), FC (11.08), Hg (10.47) and least influence of B (4.32), Pb (4.23), T (4.19). In ANN method the order is Hg > FC > NO3 > Fe > F- > BOD > pH > DO > Cd > EC > S > SAR > NH_3 > TC > T > B > Pb; whereas the most Hg (12.34), FC (12.32), NO3 (11.54) and less T (5.47), B (5.214), Pb (5.213) respectively. Thus, it is established that aforementioned all the causative factors have a higher to lower influence on VWR; therefore, all 17 factors were adopted to predict the VWRM of the Chilka lake Ramsar site.

4. Discussion

The essence of pollution-free lake water is the prominent requirement at the Chilka lake Ramsar site; therefore, our present work disclosed how the entire lake area was facing significant water resource deterioration due to inauspicious hydro-chemical properties of surface water as well as immense anthropogenic pressure on the lake by notable LULC change. However, it was crucial to develop an appropriate tool to assess the VWR and VWRM to identify the potentially vulnerable water resource zone. Forecasting VWRM is an essential and challenging task in water resource management studies. Globally, several researchers have done their work on the deteriorating nature of surface water and its impact on the entire aquatic ecosystem and also established how the VWR can be a threat to existing biodiversity (Bhateria and Jain, 2016b; De'ath and Fabricius, 2010; Njue et al., 2016) due to existence of Pb, Hg, Fe, F^-, Fe, etc. type of chronic poisoning heavy metals that brings several harmful effects on aquatic lives. Noori et al. (2021a, 2021b) also asses the complex dynamicity of water quality in the Sabalan reservoir by considering several hydro-chemical properties; it shows how nutrient enrichment helps in deteriorating the water qual-



(*caption on next page*)

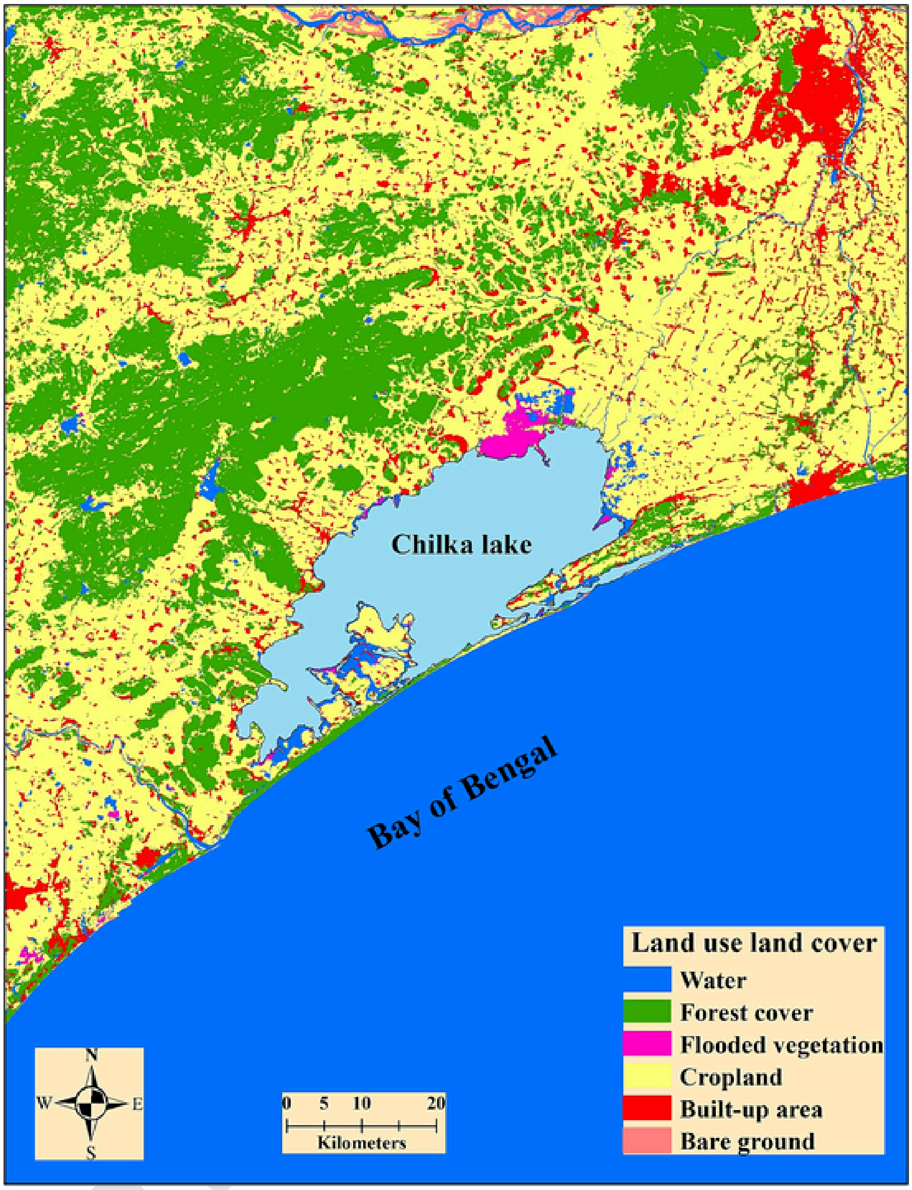


Fig. 6. (a) LULC map of study area (2005)
(b) LULC map of study area (2022).

ity significantly. Thus, the prediction of VWRM of the Chilka lake Ramsar site is a prime concern for proper water resource management with different hydro-chemical properties. At the same time, the adopted methodological approach substantially influences the reliability, accuracy, and precision level of derived results.

According to Ruidas et al. (2022b), in the present decade, some newly emerged ML algorithms such as bagging, ANN, RF, SVM, and logistic regression have derived significant results in place of the conventional methods. By taking into consideration the necessity of VWRM, we have employed two unique ML algorithms, namely ANN and RF, to assess the VWR. Earlier researchers, these two prominent ML algorithms have been used in their research work, and they have got more consistent results which meaningfully connected to reality (Ali et al., 2020; Campolo et al., 2003; Feng and Lu, 2010; Golkarian et al., 2018; Noori et al., 2011; Paschalidou et al., 2011); but in assessing of VWR, the application of these models are significantly less. These two adopted modelling approaches have the measurable capability and more consistency in prediction studies. In our research, the AUC-ROC technique has been used to evaluate the predicted result, where the ROC values of ANN are 0.93 (training) and 0.89 (validating), followed by RF. But in the case of large areal prediction, only one validating technique AUC-ROC is insufficient to ensure the derived result's reliability (Aguirre-Gutiérrez et al., 2013). Therefore, we have applied five other quantitative evaluating techniques, which also gave nearly the same result as AUC-ROC. Besides, the selection of VWR conditioning factors was the prime attention which significantly influenced water resource vulnerability. Hence, different researchers have used several methods in the identification of conditioning factors, namely VIF-TOL, regression technique, RF, and Pearson's correlation in their research (Ghosh and Maiti, 2021; Tehrany et al., 2019); but there is numerous conflict regarding factors identification procedure, although some techniques got universal acceptance in factor selection. This research employed VIF-TOL and Pearson's correlation techniques to identify the collinearity issues. In the same condition, RF and ANN, as used as variable importance techniques that showed NO_3, FC, and Hg are the most and B, Pb, and T are the least influential factors in VWRM. Specifically, our present studies show that the eastern part of this Ramsar site especially surrounded by Gagadala, Dharanikudi, Badabenakudi, Jaganathpur, and Subhadrapur is significantly characterized by very high-water resource vulnerability, whereas the western part has low VWR.

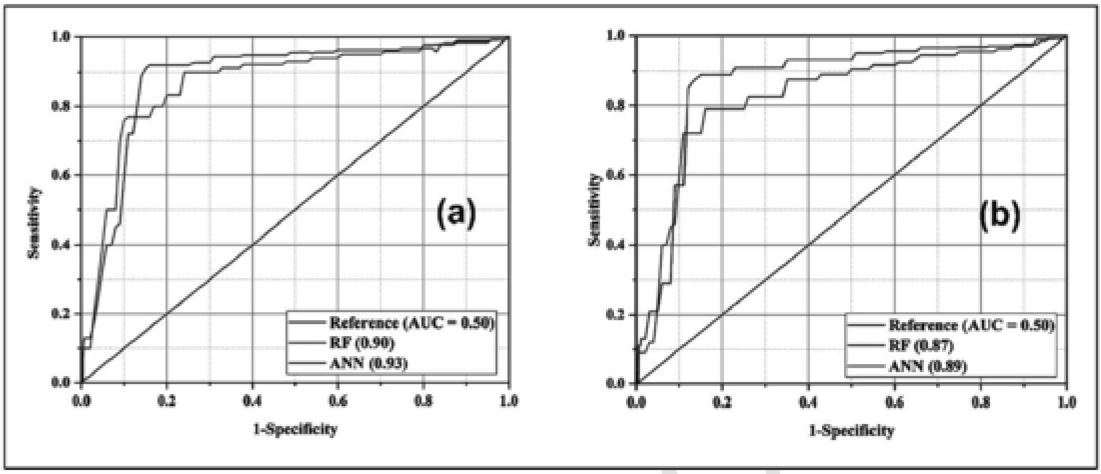


Fig. 7. AUC-ROC of (a) training and (b) validation of adopted models.

Table 4
Result of evaluation indices.

Models	Stage	Parameters				
		Sensitivity	Specificity	PPV	NPV	AUC
RF	Train	0.86	0.8	0.82	0.81	0.9
	Validation	0.85	0.74	0.78	0.84	0.87
ANN	Train	0.92	0.81	0.85	0.91	0.93
	Validation	0.9	0.78	0.78	0.88	0.89

Although similar to all other studies, our study also has some limitations which can't be avoided for future studies, i.e. firstly, we have ignored the climatic condition, soil properties, and underline rock structure of this lake region in making VWRM; secondly, surface water supply to the lake can be a significant perspective for water resource vulnerability which is overlooked here; thirdly, the water samples are collected only from the surrounding areas of this lake and not able to take the water samples from the central part due to vastness of this lake. In the future, several significant modelling approaches will emerge for hydrological research in this region. Furthermore, the accuracy and precession level of the employed modelling approach has unique imprints; thus, this may be suggested in delineating the VWRM of another similar lake region for water resource studies.

The quantification, as well as the demarcation of water resource vulnerable regions, is quite important economically in the present context to protect the environment; therefore, Whittington et al. (1996) state the significant relationship between the improvement of surface water quality and the economy of a country. In earlier times, the message of

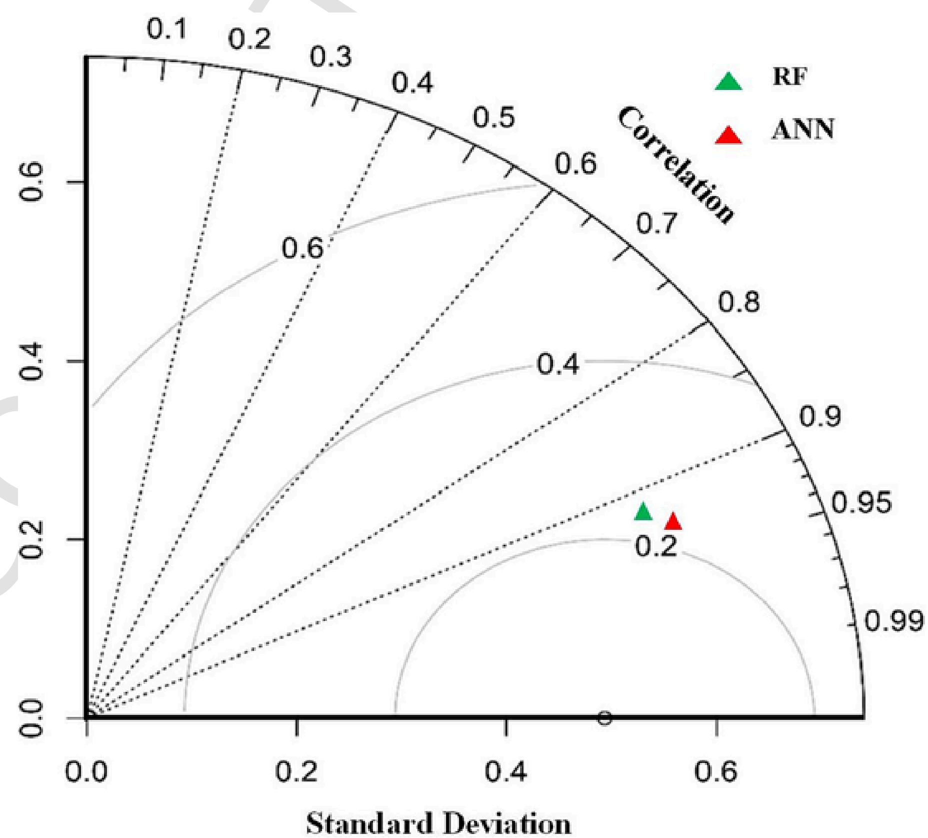


Fig. 8. Taylor diagram for adopted model validation.

Table 5
Relative importance adopted factors.

SL no.	Factor	Relative importance	
		RF	ANN
1	BOD (mg/L)	8.210	9.010
2	Cd (mg/L)	8.110	7.890
3	DO (mg/L)	7.020	8.360
4	EC (μS/cm)	6.320	7.120
5	TC (MPN/100 ml)	7.110	5.630
6	S (total)	5.020	7.120
7	Fe (mg/L)	10.210	9.610
8	NH_3 (mg/L)	5.360	6.320
9	B (mg/L)	4.320	5.214
10	FC (MPN/100 ml)	11.080	12.320
11	F (mg/L)	10.300	9.580
12	T (NTU)	4.190	5.470
13	SAR	6.340	7.120
14	pH	7.990	8.970
15	NO_3 (mg/L)	12.690	11.542
16	Hg (mg/L)	10.470	12.340
17	Pb (mg/L)	4.230	5.213

any developing country was not to take any significant effort in improving the environmental condition until those countries become economically sovereign, but the World Bank (1992) report said: "environmental quality should not be sacrificed for the economic growth". Thus, only sustainable environmental conditions can give us economically stable future well-being (Solow, 1993). So, the early prediction of our research work can reduce future disastrous impacts on society by eliminating probable causes. Henceforth, our study will encourage others to make VWRM by employing ANN and RF ML algorithms for more reliable results. All the validating techniques ensure that our predicted result is accurate and meaningful to reality which will be a very helpful tool to the policymakers and strategist to conserve the water resource of Chilka lake Ramsar site properly and helps to escape from vulnerability situation.

5. Conclusion

Lake water's water resource vulnerability can occur due to several reasons, including anthropogenic pressure, polluted water supply, and significant hydro-chemical properties. Therefore, VWRM is a beneficial tool in taking proper management strategies for controlling the VWR condition in a region. In our studies, we employed the most popular and reliable ML algorithms, such as ANN and RF, to build up the VWRM of the Chilka lake Ramsar site in Odisha. Here, 48 water samples were collected for spatial analysis of water resource vulnerability classified into five categories to specify the intensity of vulnerability; in addition, the LULC map of 2005 and 2022 was also prepared to show the anthropogenic pressure on this lake. Thus, the following outcomes are derived from our study:

- According to the all employed five validating techniques in training (ROC-AUC - 0.93, Sensitivity - 0.92, Specificity - 0.81, PPV - 0.85, NPV - 0.91) and validating phase (ROC-AUC - 0.89, Sensitivity - 0.90, Specificity - 0.78, PPV - 0.78, NPV - 0.88) with the Taylor diagram (r = 0.94), ANN emerge as a prominent and noteworthy learning algorithm for VWRM in this region and is also applicable to other regions in similar geo-hydrological conditions worldwide.
- NO_3, FC, and Hg are the most dominating factor in this region for water resource vulnerability.
- The VWRM of this lake region reveals that the eastern part is a remarkably high to very high vulnerable zone compared to the western part and covered an area of 261.07 km^2 (22.41 %) and 83.76 km^2 (7.19 %), respectively.
- In the last few years, this region experienced several disastrous effects on the aquatic ecosystem and biodiversity due to the

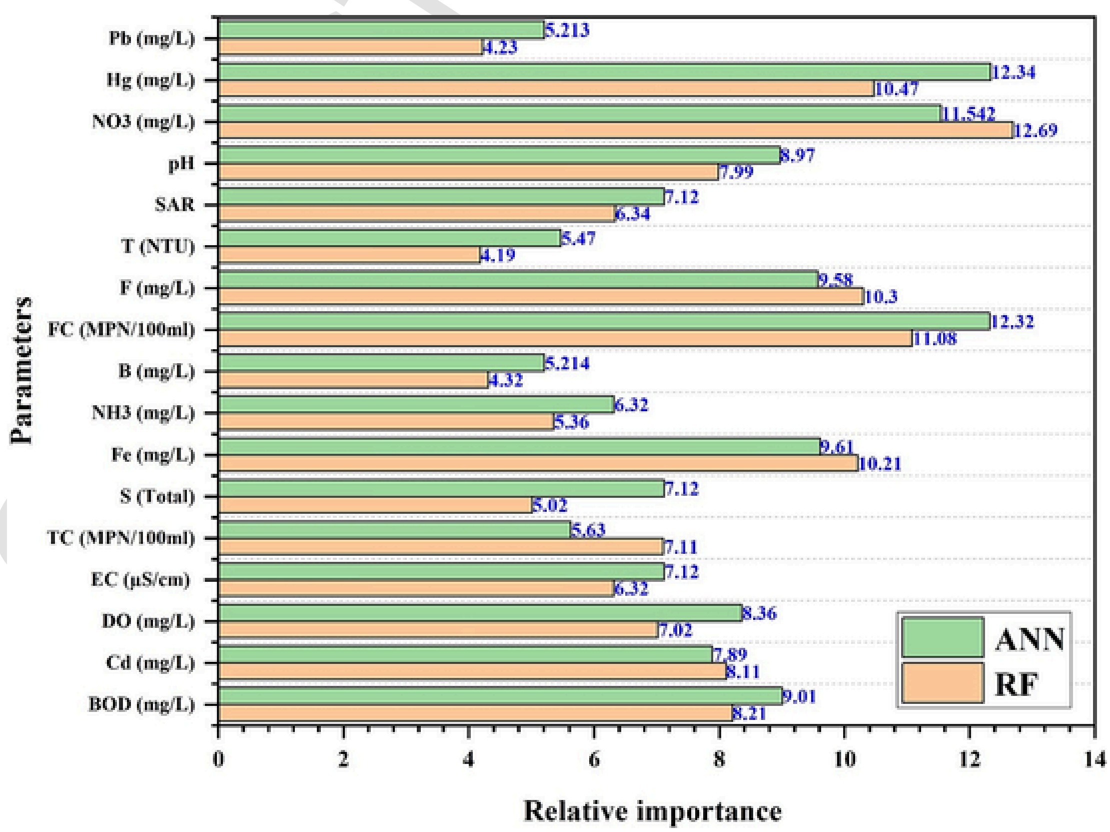


Fig. 9. quantification of Relative importance of conditioning factors.

deteriorating nature of surface water and increasing anthropogenic stress.

Therefore, the use of proper VWRM is an important tool to mitigate the degradation problem by identifying vulnerable regions. Our produced result can be used as a critical tool by the decision-makers and policymakers of government in reducing the deteriorating nature of water quality and also helps in the restoration of aquatic ecosystems and biodiversity.

CRediT authorship contribution statement

Dipankar Ruidas : Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Subodh Chandra Pal :** Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Asish Saha :** Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Indrajit Chowdhuri :** Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Manisa Shit :** Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing.

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.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.marpolbul.2022.114107.

References

Aguirre-Gutiérrez, J., Carvalheiro, L.G., Polce, C., van Loon, E.E., Raes, N., Reemer, M., Biesmeijer, J.C., 2013. Fit-for-purpose: species distribution model performance depends on evaluation criteria – Dutch hoverflies as a case study. PLOS ONE 8, e63708. https://doi.org/10.1371/journal.pone.0063708.

Ali, M., Prasad, R., Xiang, Y., Yaseen, Z.M., 2020. Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts. J. Hydrol. 584, 124647. https://doi.org/10.1016/j.jhydrol.2020.124647.

Alizadeh, M.J., Kavianpour, M.R., 2015. Development of wavelet-ANN models to predict water quality parameters in Hilo Bay, Pacific Ocean. Mar. Pollut. Bull. 98, 171–178. https://doi.org/10.1016/j.marpolbul.2015.06.052.

Al-Kalbani, M.S., Price, M.F., Abahussain, A., Ahmed, M., O'Higgins, T., 2014. Vulnerability assessment of environmental and climate change impacts on water resources in Al Jabal Al Akhdar, Sultanate of Oman. Water 6, 3118–3135. https://doi.org/10.3390/w6103118.

Anandhi, A., Kannan, N., 2018. Vulnerability assessment of water resources – translating a theoretical concept to an operational framework using systems thinking approach in a changing climate: case study in Ogallala Aquifer. J. Hydrol. 557, 460–474. https://doi.org/10.1016/j.jhydrol.2017.11.032.

Anderson, J.R., 1976. A Land Use and Land Cover Classification System for Use with Remote Sensor Data. U.S. Government Printing Office.

ASCE, A., 2000. Artificial neural networks in hydrology. I: preliminary concepts. J. Hydrol. Eng. 5, 115–123. https://doi.org/10.1061/(ASCE)1084-0699(2000)5:2(115).

Barica, J., 1978. Collapses of Aphanizomenon flos-aquae blooms resulting in massive fish kills in eutrophic lakes: effect of weather. SIL Proc. 1922–2010 (20), 208–213. https://doi.org/10.1080/03680770.1977.11896507.

Bartier, P.M., Keller, C.P., 1996. Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). Comput. Geosci. 22, 795–799. https://doi.org/10.1016/0098-3004(96)00021-0.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Bera, B., Bhattacharjee, S., Shit, P.K., Sengupta, N., Saha, S., 2021. Anthropogenic stress on a Ramsar site, India: study towards rapid transformation of the health of aquatic environment. Environ.Chall. 4, 100158. https://doi.org/10.1016/j.envc.2021.100158.

Bhat, S.A., Pandit, A.K., 2014. Surface water quality assessment of Wular Lake, a Ramsar site in Kashmir Himalaya, using discriminant analysis and WQI. J. Ecosyst. 2014.

Bhateria, R., Jain, D., 2016a. Water quality assessment of lake water: a review. Sustain.Water Resour.Manag. 2, 161–173.

Bhateria, R., Jain, D., 2016b. Water quality assessment of lake water: a review. Sustain. Water Resour. Manag. 2, 161–173. https://doi.org/10.1007/s40899-015-0014-7.

BIS, 2012. DRINKING WATER — SPECIFICATION (Second Revision). 16.

Bowen, J.L., Valiela, I., 2001. The ecological effects of urbanization of coastal watersheds: historical increases in nitrogen loads and eutrophication of Waquoit Bay estuaries. Can. J. Fish. Aquat. Sci. 58, 1489–1500. https://doi.org/10.1139/f01-094.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Bui, D.T., Khosravi, K., Karimi, M., Busico, G., Khozani, Z.S., Nguyen, H., Mastrocicco, M., Tedesco, D., Cuoco, E., Kazakis, N., 2020. Enhancing nitrate and strontium concentration prediction in groundwater by using new data mining algorithm. Sci. Total Environ. 715, 136836. https://doi.org/10.1016/j.scitotenv.2020.136836.

Campolo, M., Soldati, A., Andreussi, P., 2003. Artificial neural network approach to flood forecasting in the River Arno. Hydrol. Sci. J. 48, 381–398. https://doi.org/10.1623/hysj.48.3.381.45286.

Carbajal-Hernández, J.J., Sánchez-Fernández, L.P., Villa-Vargas, L.A., Carrasco-Ochoa, J.A., Martínez-Trinidad, J.Fco, 2013. Water quality assessment in shrimp culture using an analytical hierarchical process. Ecol. Indic. 29, 148–158. https://doi.org/10.1016/j.ecolind.2012.12.017.

Chakrabortty, R., Roy, P., Chowdhuri, I., Pal, S.C., 2021. Groundwater vulnerability assessment using random forest approach in a water-stressed paddy cultivated region of West Bengal, India. In: Groundwater Geochemistry. John Wiley & Sons, Ltd, pp. 392–410. https://doi.org/10.1002/9781119709732.ch20.

Chen, W., Tsangaratos, P., Ilia, I., Duan, Z., Chen, X., 2019. Groundwater spring potential mapping using population-based evolutionary algorithms and data mining methods. Sci. Total Environ. 684, 31–49. https://doi.org/10.1016/j.scitotenv.2019.05.312.

Chen, Y., Feng, Y., Zhang, F., Yang, F., Wang, L., 2020. Assessing and predicting the water resources vulnerability under various climate-change scenarios: a case study of Huang-Huai-Hai River Basin,China. Entropy 22, 333. https://doi.org/10.3390/e22030333.

Chowdhuri, I., Pal, S.C., Arabameri, A., Ngo, P.T.T., Chakrabortty, R., Malik, S., Das, B., Roy, P., 2020. Ensemble approach to develop landslide susceptibility map in landslide dominated Sikkim Himalayan region,India. Environ. Earth Sci. 79, 476. https://doi.org/10.1007/s12665-020-09227-5.

Chowdhuri, I., Pal, S.C., Saha, A., Chakrabortty, R., Roy, P., 2021. Evaluation of different DEMs for gully erosion susceptibility mapping using in-situ field measurement and validation. Ecol.Inform. 65, 101425. https://doi.org/10.1016/j.ecoinf.2021.101425.

Chung, C.-J., Fabbri, A.G., 2008. Predicting landslides for risk analysis — spatial models tested by a cross-validation technique. In: Geomorphology, GIS Technology and Models for Assessing Landslide Hazard and Risk, 94, pp. 438–452. https://doi.org/10.1016/j.geomorph.2006.12.036.

World Bank, W.collab, 1992. World Development Report 1992: Development and the Environment. The World Bank.

Criminisi, A., Shotton, J., 2013. Decision Forests for Computer Vision and Medical Image Analysis. Springer Science & Business Media.

Davis, R., Hirji, R., 2003. Water Quality: Assessment and Protection.

De`ath, G., Fabricius, K., 2010. Water quality as a regional driver of coral biodiversity and macroalgae on the Great Barrier Reef. Ecol. Appl. 20, 840–850. https://doi.org/10.1890/08-2023.1.

Dogan, E., Sengorur, B., Koklu, R., 2009. Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique. J. Environ. Manag. 90, 1229–1235. https://doi.org/10.1016/j.jenvman.2008.06.004.

Ducey, M.J., Johnson, K.M., Belair, E.P., Cook, B.D., 2018. The influence of human demography on land cover change in the Great Lakes States, USA. Environ. Manag. 62, 1089–1107. https://doi.org/10.1007/s00267-018-1102-x.

Elsayed, S., Gad, M., Farouk, M., Saleh, A.H., Hussein, H., Elmetwalli, A.H., Elsherbiny, O., Moghanm, F.S., Moustapha, M.E., Taher, M.A., Eid, E.M., Abou El-Safa, M.M., 2021a. Using optimized two and three-band spectral indices and multivariate models to assess some water quality indicators of Qaroun Lake in Egypt. Sustainability 13, 10408. https://doi.org/10.3390/su131810408.

Elsayed, S., Ibrahim, H., Hussein, H., Elsherbiny, O., Elmetwalli, A.H., Moghanm, F.S., Ghoneim, A.M., Danish, S., Datta, R., Gad, M., 2021b. Assessment of water quality in Lake Qaroun using ground-based remote sensing data and artificial neural networks. Water 13, 3094. https://doi.org/10.3390/w13213094.

Feng, L.-H., Lu, J., 2010. The practical research on flood forecasting based on artificial neural networks. Expert Syst. Appl. 37, 2974–2977. https://doi.org/10.1016/j.eswa.2009.09.037.

Feng, Y., Zheng, B.-H., Jia, H.-F., Peng, J.-Y., Zhou, X.-Y., 2021. Influence of social and economic development on water quality in Dongting Lake. Ecol. Indic. 131, 108220. https://doi.org/10.1016/j.ecolind.2021.108220.

Filik Iscen, C., Emiroglu, Ö., Ilhan, S., Arslan, N., Yilmaz, V., Ahiska, S., 2008. Application of multivariate statistical techniques in the assessment of surface water quality in Uluabat Lake, Turkey. Environ. Monit. Assess. 144, 269–276. https://doi.org/10.1007/s10661-007-9989-3.

Füssel, H.-M., Klein, R.J.T., 2006. Climate change vulnerability assessments: an evolution of conceptual thinking. Clim. Chang. 75, 301–329. https://doi.org/10.1007/s10584-006-0329-3.

Gad, M., Abou El-Safa, M.M., Farouk, M., Hussein, H., Alnemari, A.M., Elsayed, S., Khalifa, M.M., Moghanm, F.S., Eid, E.M., Saleh, A.H., 2021. Integration of water

quality indices and multivariate modeling for assessing surface water quality in Qaroun Lake,Egypt. Water 13, 2258. https://doi.org/10.3390/w13162258.
Gain, A.K., Rouillard, J.J., Benson, D., 2013. Can Integrated Water Resources Management Increase Adaptive Capacity to Climate Change Adaptation? A Critical Review. https://doi.org/10.4236/jwarp.2013.54A003.
Geng, Y., Yi, J., 2006. Integrated water resource management at the industrial park level: a case of the Tianjin Economic Development Area. Int. J. Sust. Dev. World Ecol. 13, 37–50. https://doi.org/10.1080/13504500609469660.
Ghosh, A., Maiti, R., 2021. Soil erosion susceptibility assessment using logistic regression, decision tree and random forest: study on the Mayurakshi river basin of Eastern India. Environ. Earth Sci. 80, 328. https://doi.org/10.1007/s12665-021-09631-5.
Ghosh, A.K., Pattnaik, A.K., 2005. Chilika Lagoon: experience and lessons learned brief. In: ILEC (Ed.), Managing Lakes and Their Basins for Sustainable Use: A Report for Lake Basin Managers and Stakeholders. International Lake Environment Committee Foundation, Kusatsu, pp. 115–132.
Golkarian, A., Naghibi, S.A., Kalantar, B., Pradhan, B., 2018. Groundwater potential mapping using C5.0, random forest, and multivariate adaptive regression spline models in GIS. Environ. Monit. Assess. 190, 149. https://doi.org/10.1007/s10661-018-6507-8.
Gu, Q., Li, J., Deng, J., Lin, Y., Ma, L., Wu, C., Wang, K., Hong, Y., 2015. Eco-environmental vulnerability assessment for large drinking water resource: a case study of Qiandao Lake Area,China. Front. Earth Sci. 9, 578–589. https://doi.org/10.1007/s11707-014-0472-5.
Hagan, M.T., Demuth, H.B., Beale, M., 1997. Neural Network Design. PWS Publishing Co.
Hastie, T., Friedman, J., Tibshirani, R., 2001. The elements of statistical learning. In: Springer Series in Statistics. Springer, New York, NY. https://doi.org/10.1007/978-0-387-21606-5.
Hernández-Terrones, L.M., Null, K.A., Ortega-Camacho, D., Paytan, A., 2015. Water quality assessment in the Mexican Caribbean: impacts on the coastal ecosystem. Cont. Shelf Res. 102, 62–72. https://doi.org/10.1016/j.csr.2015.04.015.
Hong, H., Liu, J., Bui, D.T., Pradhan, B., Acharya, T.D., Pham, B.T., Zhu, A.-X., Chen, W., Ahmad, B.B., 2018. Landslide susceptibility mapping using J48 Decision Tree with AdaBoost, Bagging and Rotation Forest ensembles in the Guangchang area (China). Catena 163, 399–413. https://doi.org/10.1016/j.catena.2018.01.005.
Hossain, A.K.M.A., Jia, Y., Chao, X., Altinakar, M., 2014. Application of advanced remote sensing techniques to improve modeling estuary water quality. In: Finkl, C.W., Makowski, C. (Eds.), Remote Sensing and Modeling: Advances in Coastal and Marine Resources, Coastal Research Library. Springer International Publishing, Cham, pp. 295–313. https://doi.org/10.1007/978-3-319-06326-3_12.
Hutton, C.W., Kienberger, S., Amoako Johnson, F., Allan, A., Giannini, V., Allen, R., 2011. Vulnerability to climate change: people, place and exposure to hazard. In: Advances in Science and Research. Presented at the BRAHMATWINN Results -, Copernicus GmbH, pp. 37–45. https://doi.org/10.5194/asr-7-37-2011.
Irandoost, F., Agah, H., Eslami, Z., Rossi, L., Colloca, F., Khalili, A., Costantini, M.L., 2021. Evaluation of nitrogen and heavy metal pollution in southern Caspian Sea: risk assessment and modeling approach. Mar. Pollut. Bull. 173, 113041. https://doi.org/10.1016/j.marpolbul.2021.113041.
Islam, A.R.Md.T., Pal, S.C., Chowdhuri, I., Salam, R., Islam, Md.S., Rahman, Md.M., Zahid, A., Idris, A.M., 2021a. Application of novel framework approach for prediction of nitrate concentration susceptibility in coastal multi-aquifers,Bangladesh. Sci. Total Environ. 801, 149811. https://doi.org/10.1016/j.scitotenv.2021.149811.
Islam, A.R.Md.T., Pal, S.C., Chowdhuri, I., Salam, R., Islam, Md.S., Rahman, Md.M., Zahid, A., Idris, A.M., 2021b. Application of novel framework approach for prediction of nitrate concentration susceptibility in coastal multi-aquifers,Bangladesh. Sci. Total Environ. 801, 149811. https://doi.org/10.1016/j.scitotenv.2021.149811.
Jamal, S., Ahmad, W.S., Ajmal, U., Aaquib, Md., Ashif Ali, Md., Babor Ali, Md., Ahmed, S., 2022. An integrated approach for determining the anthropogenic stress responsible for degradation of a Ramsar Site – Wular Lake in Kashmir, India. Mar. Geod. 1–28. https://doi.org/10.1080/01490419.2022.2034686.
Jaydhar, A.K., Chandra Pal, S., Saha, A., Islam, A.R.Md.T., Ruidas, D., 2022. Hydrogeochemical evaluation and corresponding health risk from elevated arsenic and fluoride contamination in recurrent coastal multi-aquifers of eastern India. J. Clean. Prod. 369, 133150. https://doi.org/10.1016/j.jclepro.2022.133150.
Juahir, H., Zain, S.M., Toriman, M.E., Mokhtar, M., Man, H.C., 2004. Application of artificial neural network models for predicting water quality index. Malays. J. Civ. Eng. 16.
Kazi, T.G., Arain, M.B., Jamali, M.K., Jalbani, N., Afridi, H.I., Sarfraz, R.A., Baig, J.A., Shah, A.Q., 2009. Assessment of water quality of polluted lake using multivariate statistical techniques: a case study. Ecotoxicol. Environ. Saf. 72, 301–309. https://doi.org/10.1016/j.ecoenv.2008.02.024.
Khan, S.S., Jehan, M., Rehman, N., Shah, S., 2013. MT and Din, I. (2013) . Drinking water quality and human health risk in Charsadda district,Pakistan. J. Clean. Prod. 60, 93–101.
Khosravi, K., Shahabi, H., Pham, B.T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H.-B., Gróf, G., Ho, H.L., Hong, H., Chapi, K., Prakash, I., 2019. A comparative assessment of flood susceptibility modeling using multi-criteria decision-making analysis and machine learning methods. J. Hydrol. 573, 311–323.
Kouadri, S., Elbeltagi, A., Islam, A.R.M., Kateb, S., 2021. Performance of machine learning methods in predicting water quality index based on irregular data set: application on Illizi region (Algerian southeast). Appl. Water Sci. 11, 1–20.
Kumar, R., Singh, R.D., Sharma, K.D., 2005. Water resources of India. Curr. Sci. 89, 794–811.
Li, X., Zhang, Y., Ma, N., Li, C., Luan, J., 2021. Contrasting effects of climate and LULC change on blue water resources at varying temporal and spatial scales. Sci. Total Environ. 786, 147488. https://doi.org/10.1016/j.scitotenv.2021.147488.
Liaw, A., Wiener, M., 2002. In: Classification and Regression by randomForest, 2, p. 5.
Luo, Z., Shao, Q., Zuo, Q., Cui, Y., 2020. Impact of land use and urbanization on river water quality and ecology in a dam dominated basin. J. Hydrol. 584, 124655.
Mabwoga, S.O., Chawla, A., Thukral, A.K., 2010. Assessment of water quality parameters of the Harike wetland in India, a Ramsar site, using IRS LISS IV satellite data. Environ. Monit. Assess. 170, 117–128. https://doi.org/10.1007/s10661-009-1220-2.
Mack, L., Andersen, H.E., Beklioğlu, M., Bucak, T., Couture, R.-M., Cremona, F., Ferreira, M.T., Hutchins, M.G., Mischke, U., Molina-Navarro, E., 2019. The future depends on what we do today–Projecting Europe's surface water quality into three different future scenarios. Sci. Total Environ. 668, 470–484.
Mansour, S., Kouz, T., Thaiki, M., Ouhadi, A., Mesmoudi, H., Hassani Zerrouk, M., Mourabit, T., Dakak, H., Cherkaoui Dekkaki, H., 2021. Spatial assessment of the vulnerability of water resources against anthropogenic pollution using the DKPR model: a case of Ghiss-Nekkour basin,Morocco. Arab. J. Geosci. 14, 699. https://doi.org/10.1007/s12517-021-06973-z.
May, R.J., Dandy, G.C., Maier, H.R., Nixon, J.B., 2008. Application of partial mutual information variable selection to ANN forecasting of water quality in water distribution systems. Environ. Model Softw. 23, 1289–1299. https://doi.org/10.1016/j.envsoft.2008.03.008.
McBride, G.B., 2005. Using Statistical Methods for Water Quality Management: Issues, Problems and Solutions. John Wiley & Sons.
McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. 5, 115–133. https://doi.org/10.1007/BF02478259.
McGrane, S.J., 2016. Impacts of urbanisation on hydrological and water quality dynamics, and urban water management: a review. Hydrol. Sci. J. 61, 2295–2311. https://doi.org/10.1080/02626667.2015.1128084.
Milano, M., Reynard, E., Köplin, N., Weingartner, R., 2015. Climatic and anthropogenic changes in Western Switzerland: impacts on water stress. Sci. Total Environ. 536, 12–24. https://doi.org/10.1016/j.scitotenv.2015.07.049.
Mishra, S., 2017. GEO-Climatic Abstractions of South Mahanadi Delta and Chilika Lagoon, India: Post Anthropogenic Interventions. https://doi.org/10.5829/idosi.wasj.2015.33.02.22174.
Mishra, S., Dwibedy, S., 2015. Geo-hydrology of South Mahanadi Delta and Chilika Lake, Odisha. Int. J. Adv. Res. 3.
Mohanty, P.K., Dash, S.K., Mishra, P.K., Murty, A.S.N., 1997. Heat and momentum fluxes over Chilka: a tropical lagoon. Oceanogr. Lit. Rev. 3, 187.
Morrice, J.A., Danz, N.P., Regal, R.R., Kelly, J.R., Niemi, G.J., Reavie, E.D., Hollenhorst, T., Axler, R.P., Trebitz, A.S., Cotter, A.M., Peterson, G.S., 2008. Human influences on water quality in great lakes coastal wetlands. Environ. Manag. 41, 347–357. https://doi.org/10.1007/s00267-007-9055-5.
Murray, K.E., Thomas, S.M., Bodour, A.A., 2010. Prioritizing research for trace pollutants and emerging contaminants in the freshwater environment. Environ. Pollut. 158, 3462–3471. https://doi.org/10.1016/j.envpol.2010.08.009.
Najafzadeh, M., Noori, R., Afroozi, D., Ghiasi, B., Hosseini-Moghari, S.-M., Mirchi, A., Torabi Haghighi, A., Kløve, B., 2021. A comprehensive uncertainty analysis of model-estimated longitudinal and lateral dispersion coefficients in open channels. J. Hydrol. 603, 126850. https://doi.org/10.1016/j.jhydrol.2021.126850.
Nisari, A.R., Sujatha, C.H., 2021. Assessment of trace metal contamination in the Kol wetland, a Ramsar site, Southwest coast of India. Reg. Stud. Mar. Sci. 47, 101953. https://doi.org/10.1016/j.rsma.2021.101953.
Njue, N., Sirmah, P., Koech, E., Hitimana, J., 2016. Influence of Land Use Activities on Riparian Vegetation, Soil and Water Quality: An Indicator of Biodiversity Loss, South West Mau Forest, Kenya.
Noori, R., Karbassi, A.R., Mehdizadeh, H., Vesali-Naseh, M., Sabahi, M.S., 2011. A framework development for predicting the longitudinal dispersion coefficient in natural streams using an artificial neural network. Environ. Prog. Sustain. Energy 30, 439–449. https://doi.org/10.1002/ep.10478.
Noori, R., Ansari, E., Bhattarai, R., Tang, Q., Aradpour, S., Maghrebi, M., Torabi Haghighi, A., Bengtsson, L., Kløve, B., 2021a. Complex dynamics of water quality mixing in a warm mono-mictic reservoir. Sci. Total Environ. 777, 146097. https://doi.org/10.1016/j.scitotenv.2021.146097.
Noori, R., Ansari, E., Jeong, Y.-W., Aradpour, S., Maghrebi, M., Hosseinzadeh, M., Bateni, S.M., 2021b. Hyper-nutrient enrichment status in the Sabalan Lake,Iran. Water 13, 2874. https://doi.org/10.3390/w13202874.
Ok, A.O., Akar, O., Gungor, O., 2012. Evaluation of random forest method for agricultural crop classification. Eur.J.Remote Sens. 45, 421–432. https://doi.org/10.5721/EuJRS20124535.
Olasoji, S.O., Oyewole, N.O., Abiola, B., Edokpayi, J.N., 2019. Water quality assessment of surface and groundwater sources using a water quality index method: a case study of a peri-urban town in Southwest,Nigeria. Environments 6, 23. https://doi.org/10.3390/environments6020023.
Olden, J.D., Kennard, M.J., Pusey, B.J., 2008. Species invasions and the changing biogeography of Australian freshwater fishes. Glob. Ecol. Biogeogr. 17, 25–37. https://doi.org/10.1111/j.1466-8238.2007.00340.x.
Pal, S.C., Chakrabortty, R., Arabameri, A., Santosh, M., Saha, A., Chowdhuri, I., Roy, P., Shit, M., 2022a. Chemical weathering and gully erosion causing land degradation in a complex river basin of Eastern India: an integrated field, analytical and artificial intelligence approach. Nat. Hazards 110, 847–879. https://doi.org/10.1007/s11069-021-04971-8.
Pal, S.C., Islam, A.R.M.T., Chakrabortty, R., Islam, M.S., Saha, A., Shit, M., 2022b. Application of data-mining technique and hydro-chemical data for evaluating vulnerability of groundwater in Indo-Gangetic Plain. J. Environ. Manag. 318, 115582.
Pal, S.C., Ruidas, D., Saha, A., Islam, A.R.Md.T., Chowdhuri, I., 2022c. Application of novel data-mining technique based nitrate concentration susceptibility prediction approach for coastal aquifers in India. J. Clean. Prod. 346, 131205. https://doi.org/10.1016/j.jclepro.2022.131205.
Palani, S., Liong, S.-Y., Tkalich, P., 2008a. An ANN application for water quality

forecasting. Mar. Pollut. Bull. 56, 1586–1597.
Palani, S., Liong, S.-Y., Tkalich, P., 2008b. An ANN application for water quality forecasting. Mar. Pollut. Bull. 56, 1586–1597. https://doi.org/10.1016/j.marpolbul.2008.05.021.
Panigrahi, S., Acharya, B.C., Panigrahy, R.C., Nayak, B.K., Banarjee, K., Sarkar, S.K., 2007. Anthropogenic impact on water quality of Chilika lagoon RAMSAR site: a statistical approach. Wetl. Ecol. Manag. 15, 113–126. https://doi.org/10.1007/s11273-006-9017-3.
Park, Y.S., Engel, B.A., 2016. Identifying the correlation between water quality data and LOADEST model behavior in annual sediment load estimations. Water 8, 368. https://doi.org/10.3390/w8090368.
Parry, M.L., Canziani, O.F., Palutikof, J.P., Linden, P.J.van der, Hanson, C.E., 2007. Climate change 2007: impacts, adaptation and vulnerability. In: Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Summary for Policymakers.
Paschalidou, A.K., Karakitsios, S., Kleanthous, S., Kassomenos, P.A., 2011. Forecasting hourly PM10 concentration in Cyprus through artificial neural networks and multiple regression models: implications to local environmental management. Environ. Sci. Pollut. Res. 18, 316–327. https://doi.org/10.1007/s11356-010-0375-2.
Patnaik, S., 1998. Estimation of sediment flow into the Chilika lake. In: Abstracts of the Workshop on Sustainable Development of Chilika Lake, Orissa (Bhubaneswar, India), 1998, 32.
Pavlidou, A., Kontoyiannis, H., Zarokanelos, N., Hatzianestis, I., Assimakopoulou, G., Psyllidou-Giouranovits, R., 2014. Seasonal and spatial nutrient dynamics in Saronikos Gulf: the impact of sewage effluents from athens sewage treatment plant. In: Ansari, A.A., Gill, S.S. (Eds.), Eutrophication: Causes, Consequences and Control, volume 2. Springer Netherlands, Dordrecht, pp. 111–130.
Pellerin, B.A., Stauffer, B.A., Young, D.A., Sullivan, D.J., Bricker, S.B., Walbridge, M.R., Clyde, Jr, G.A., Shaw, D.M., 2016. Emerging tools for continuous nutrient monitoring networks: sensors advancing science and water resources protection. JAWRA J.Am.Water Resour.Assoc. 52, 993–1008. https://doi.org/10.1111/1752-1688.12386.
Qin, J., Ding, Y.-J., Zhao, Q.-D., Wang, S.-P., Chang, Y.-P., 2020. Assessments on surface water resources and their vulnerability and adaptability in China. Adv. Clim. Chang. Res. 11, 381–391. https://doi.org/10.1016/j.accre.2020.11.002. . Special Issue on Climate change for coastal areas: risks, adaptation and acceptability.
Rahman, R., Dhruba, S.R., Ghosh, S., Pal, R., 2019. Functional random forest with applications in dose-response predictions. Sci. Rep. 9, 1628. https://doi.org/10.1038/s41598-018-38231-w.
Reid, K., 2018. Global Water Crisis: Facts, FAQs, and How to Help.
Rodriguez-Galiano, V., Mendes, M.P., Garcia-Soldado, M.J., Chica-Olmo, M., Ribeiro, L., 2014. Predictive modeling of groundwater nitrate pollution using random forest and multisource variables related to intrinsic and specific vulnerability: a case study in an agricultural setting (Southern Spain). Sci. Total Environ. 476–477, 189–206. https://doi.org/10.1016/j.scitotenv.2014.01.001.
Romshoo, S.A., Ali, N., Rashid, I., 2011. Geoinformatics for characterizing and understanding the spatio-temporal dynamics (1969 to 2008) of Hokersar wetland in Kashmir Himalayas. IJPS 6, 1026–1038. https://doi.org/10.5897/IJPS10.378.
Ruidas, D., Pal, S.C., 2022. Potential hotspot modeling and monitoring of PM2.5 concentration for sustainable environmental health in Maharashtra, India. Sustain. Water Resour. Manag. 8, 98. https://doi.org/10.1007/s40899-022-00682-5.
Ruidas, D., Pal, S.C., Islam, A.R.Md.T., Saha, A., 2021. Characterization of groundwater potential zones in water-scarce hardrock regions using data driven model. Environ. Earth Sci. 80, 809. https://doi.org/10.1007/s12665-021-10116-8.
Ruidas, D., Chakrabortty, R., Islam, A.R.Md.T., Saha, A., Pal, S.C., 2022a. A novel hybrid of meta-optimization approach for flash flood-susceptibility assessment in a monsoon-dominated watershed,Eastern India. Environ. Earth Sci. 81, 145. https://doi.org/10.1007/s12665-022-10269-0.
Ruidas, D., Pal, S.C., Towfiqul Islam, A.R.Md., Saha, A., 2022b. Hydrogeochemical evaluation of groundwater aquifers and associated health hazard risk mapping using ensemble data driven model in a water scares plateau region of eastern India. Expo Health. https://doi.org/10.1007/s12403-022-00480-6.
Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature 323, 533–536. https://doi.org/10.1038/323533a0.
Saha, A., Pal, S.C., Arabameri, A., Chowdhuri, I., Rezaie, F., Chakrabortty, R., Roy, P., Shit, M., 2021a. Optimization modelling to establish false measures implemented with ex-situ plant species to control gully erosion in a monsoon-dominated region with novel in-situ measurements. J. Environ. Manag. 287, 112284. https://doi.org/10.1016/j.jenvman.2021.112284.
Saha, A., Pal, S.C., Chowdhuri, I., Ruidas, D., Chakrabortty, R., Roy, P., Shit, M., 2021b. Impact of firecrackers burning and policy-practice gap on air quality in Delhi during Indian's great mythological event of Diwali festival. Cities 119, 103384. https://doi.org/10.1016/j.cities.2021.103384.
Saha, A., Pal, S.C., Santosh, M., Janizadeh, S., Chowdhuri, I., Norouzi, A., Roy, P., Chakrabortty, R., 2021c. Modelling multi-hazard threats to cultural heritage sites and environmental sustainability: the present and future scenarios. J. Clean. Prod. 320, 128713. https://doi.org/10.1016/j.jclepro.2021.128713.
Saha, A., Pal, S.C., Chowdhuri, I., Islam, A.R.Md.T., Chakrabortty, R., Roy, P., 2022a. Application of neural network model-based framework approach to identify gully erosion potential hotspot zones in sub-tropical environment. Geocarto Int. 1–27. https://doi.org/10.1080/10106049.2022.2091042.
Saha, A., Pal, S.C., Chowdhuri, I., Islam, A.R.M.T., Roy, P., Chakrabortty, R., 2022b. Land degradation risk dynamics assessment in red and lateritic zones of eastern plateau, India: a combine approach of K-fold CV, data mining and field validation. Ecol.Inf. 101653.
Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., Leemans, R., Lodge, D.M., Mooney, H.A., Oesterheld, M., Poff, N.L., Sykes, M.T., Walker, B.H., Walker, M., Wall, D.H., 2000. Global biodiversity scenarios for the year 2100. Science 287, 1770–1774. https://doi.org/10.1126/science.287.5459.1770.
Sarkar, S., Bhattacharya, A., Satpathy, K.K., Mohanty, A., Panigrahi, S., 2012. In: Chilika Lake – A Ramsar Site. pp. 148–156.
Sengorur, B., Koklu, R., Ates, A., 2015. Water quality assessment using artificial intelligence techniques: SOM and ANN—a case study of Melen River Turkey. Water Qual. Expo. Health 7, 469–490. https://doi.org/10.1007/s12403-015-0163-9.
Shabbir, R., Ahmad, S.S., 2016. Water resource vulnerability assessment in Rawalpindi and Islamabad, Pakistan using analytic hierarchy process (AHP). J.King Saud Univ.Sci. 28, 293–299. https://doi.org/10.1016/j.jksus.2015.09.007.
Shrestha, S., Kazama, F., 2007. Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan. Environ. Model. Softw. 22, 464–475. https://doi.org/10.1016/j.envsoft.2006.02.001. . Special section: Environmental Risk and Emergency Management.
Siddiqui, S., 1995. Limnology of Chilika Lake. Fauna of Chilika Lake.
Singh, K.P., Malik, A., Mohan, D., Sinha, S., 2004. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. Water Res. 38, 3980–3992. https://doi.org/10.1016/j.watres.2004.06.011.
Singh, N., Kaur, M., Kaur Katnoria, J., 2017. Spatial and temporal heavy metal distribution and surface water characterization of Kanjli wetland (a Ramsar site), India using different indices. Bull. Environ. Contam. Toxicol. 99, 735–742. https://doi.org/10.1007/s00128-017-2194-3.
Singh, A., Gogoi, A., Saikia, P., Karunanidhi, D., Kumar, M., 2021. Integrated use of inverse and biotic ligand modelling for lake water quality resilience estimation: a case of Ramsar wetland, (Deepor Beel), Assam,India. Environ. Res. 200, 111397. https://doi.org/10.1016/j.envres.2021.111397.
Smith, V.H., 2003. Eutrophication of freshwater and coastal marine ecosystems a global problem. Environ. Sci. Pollut. Res. 10, 126–139. https://doi.org/10.1065/espr2002.12.142.
Solow, R., 1993. An Almost Practical Step Toward Sustainability. 11.
Soonthornrangsan, J.T., Lowry, C.S., 2021. Vulnerability of water resources under a changing climate and human activity in the lower Great Lakes region. Hydrol. Process. 35, e14440. https://doi.org/10.1002/hyp.14440.
Sruthy, S., Ramasamy, E.V., 2017. Microplastic pollution in Vembanad Lake, Kerala, India: the first report of microplastics in lake and estuarine sediments in India. Environ. Pollut. 222, 315–322. https://doi.org/10.1016/j.envpol.2016.12.038.
Talukdar, S., Ghose, B., Shahfahad, Salam, R., Mahato, S., Pham, Q.B., Linh, N.T.T., Costache, R., Avand, M., 2020. Flood susceptibility modeling in Teesta River basin, Bangladesh using novel ensembles of bagging algorithms. Stoch. Environ. Res. Risk Assess. 34, 2277–2300. https://doi.org/10.1007/s00477-020-01862-5.
Talukdar, S., Eibek, K.U., Akhter, S., Ziaul, S.K., Islam, A.R.M.T., Mallick, J., 2021. Modeling fragmentation probability of land-use and land-cover using the bagging, random forest and random subspace in the Teesta River Basin, Bangladesh. Ecol. Indic. 126, 107612.
Tanaka, T., Sato, T., Watanabe, K., Wang, Y., Yang, D., Inoue, H., Li, K., Inamura, T., 2013. Irrigation system and land use effect on surface water quality in river, at lake Dianchi, Yunnan, China. J. Environ. Sci. 25, 1107–1116.
Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmos. 106, 7183–7192. https://doi.org/10.1029/2000JD900719.
Tehrany, M.S., Jones, S., Shabani, F., 2019. Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. Catena 175, 174–192. https://doi.org/10.1016/j.catena.2018.12.011.
Tian, Y., Jiang, Y., Liu, Q., Dong, M., Xu, D., Liu, Y., Xu, X., 2019. Using a water quality index to assess the water quality of the upper and middle streams of the Luanhe River, northern China. Sci. Total Environ. 667, 142–151.
Tiyasha, T., Tung, T.M., Bhagat, S.K., Tan, M.L., Jawad, A.H., Mohtar, W.H.M.W., Yaseen, Z.M., 2021. Functionalization of remote sensing and on-site data for simulating surface water dissolved oxygen: development of hybrid tree-based artificial intelligence models. Mar. Pollut. Bull. 170, 112639. https://doi.org/10.1016/j.marpolbul.2021.112639.
Tonmoy, F.N., El-Zein, A., Hinkel, J., 2014. Assessment of vulnerability to climate change using indicators: a meta-analysis of the literature. WIREs Clim.Chang. 5, 775–792. https://doi.org/10.1002/wcc.314.
Wang, X., Lu, Y., Han, J., He, G., Wang, T., 2007. Identification of anthropogenic influences on water quality of rivers in Taihu watershed. J. Environ. Sci. 19, 475–481. https://doi.org/10.1016/S1001-0742(07)60080-1.
Wang, W., Xu, D., Chau, K., Lei, G., 2014. Assessment of river water quality based on theory of variable fuzzy sets and fuzzy binary comparison method. Water Resour. Manag. 28, 4183–4200. https://doi.org/10.1007/s11269-014-0738-4.
Weaver, C.P., Lempert, R.J., Brown, C., Hall, J.A., Revell, D., Sarewitz, D., 2013. Improving the contribution of climate model information to decision making: the value and demands of robust decision frameworks. WIREs Climat. Chang. 4, 39–60. https://doi.org/10.1002/wcc.202.
Westerhuis, J.A., Hoefsloot, H.C.J., Smit, S., Vis, D.J., Smilde, A.K., van Velzen, E.J.J., van Duijnhoven, J.P.M., van Dorsten, F.A., 2008. Assessment of PLSDA cross validation. Metabolomics 4, 81–89. https://doi.org/10.1007/s11306-007-0099-6.
Whittington, D., Choe, K., Lauria, D., 1996. The economic benefits of surface water quality improvements in developing countries: a case study of Davao,Philippines. Land Econ. 72. https://doi.org/10.2307/3146913.
Wu, W., Dandy, G.C., Maier, H.R., 2014. Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling. Environ. Model Softw. 54, 108–127. https://

doi.org/10.1016/j.envsoft.2013.12.016.

Xu, X., Yang, G., Tan, Y., Tang, X., Jiang, H., Sun, X., Zhuang, Q., Li, H., 2017. Impacts of land use changes on net ecosystem production in the Taihu Lake Basin of China from 1985 to 2010. J. Geophys. Res. Biogeosci. 122, 690–707.

Yilmaz, N., Ozyigit, I.I., Demir, H.H., Yalcin, I.E., 2021. Assessment on phytoplankton composition and heavy metal pollution in a drinking water resource: Lake Terkos (Istanbul, Turkey). DWT 225, 265–274. https://doi.org/10.5004/dwt.2021.27221.

Zhongming, Z., Linong, L., Xiaona, Y., Wangqiang, Z., Wei, L., 2021. Globally, 3 Billion People at Health Risk Due to Scarce Data on Water Quality.