Hybrid Tree-Based Wetland Vulnerability Modelling

Swades Pal and Satyajit Paul

Abstract Wetlands of the moribund region of the Ganga–Brahmaputra deltaic part experience extreme loss and degradation, which is the leading cause for our present study. In this study, the vulnerable situation, as a part of degradation, is explored using tree-based ML algorithms in python environment using eight conditioning parameters, namely: water presence frequency (WPF), change in WPF, hydro duration, water depth, agriculture presence frequency, proximity to the river, distance from the road network, and built-up proximity. Four tree-based machine learning algorithms, namely, bagging classification model, reduced error pruning tree (REP Tree), gradient boosting classification model (GBM), and AdaBoosting classification model (ADB), has been used to evaluate the vulnerability of wetlands for both phase II (1998–2007) and phase III (2008–2017). It is found that 23.92–25.01% and 44.67–46.99% area to total wetland area emerged as high to very high vulnerable zone in phase II, whereas 24.08–26.16% and 45.41–49.13% of wetland area identified as high to very high vulnerable zone in phase III. More than 45% of the total wetland area disappeared during phase II to phase III. The models have been validated using the following matrices like sensitivity, Precision F1-score, and MCC for justifying the best-suited model. With an average score of more than 91 for all the matrices, the gradient boosting classification model (GBM), and AdaBoosting classification model (ADB) exhibit more prediction capability and model accuracy than the bagging classification model, and Reduced Error Pruning (REP) Tree model. With the successful prediction, the study recommends tree-based ML algorithms for such or similar works. The study also warns about growing wetland habitat vulnerability and its negative consequences on socio-ecological benefits.

Keywords Wetland vulnerability · Tree-based algorithms · Moribund delta · Machine learning (ML)

S. Pal
Department of Geography, University of Gour Banga, Malda, India

S. Paul (✉)
Department of Geography, Gour Mahavidyalaya, Malda, India
e-mail: spaulofficial.geo@gmail.com

H. Sajjad et al. (eds.), *Challenges of Disasters in Asia*, Springer Natural Hazards,
https://doi.org/10.1007/978-981-19-3567-1_11

1 Introduction

The wetland contains a distinctive ecosystem system with significant hydro-ecological functions that have been altered significantly faster than any other known ecosystem (Dong et al. 2020; MEA 2005). Wetland provides 60% of global ecosystem services with only 6% global spatial extension (Finlayson and Davidson 2018). The wetland ecosystem also provides shelter for 40% of global species including some of the most endangered ones (Meng and Dong 2019). Despite immense ecosystem contribution, the wetland is one of the most threatened ecosystems due to rapid change in habitat ecology triggered by agricultural extension, infrastructural developments, population growth in the wetland area (Akpabio and Umoh 2021; Saha and Pal 2019a), hydrological modification (Pal and Debanshi 2021a, b; Talukdar and Pal 2019). Agricultural extension towards wetland areas, directly and indirectly, affects its habitat, and rapid urban growth and infrastructural development deteriorate its habitat completely (Xia et al. 2021). Davidson (2014) reported that 87% of the global wetland has been lost since 1700, with an increasing rate of up to 71% in late 1900. Space Application Centre (2018) reported that 32.5% of Indian wetlands shrink seasonally due to rainfall anomaly and water table fluctuation, which is responsible for ~3% annual wetland loss (Prasher 2018). In the floodplain region, population pressure towards the wetland area with vast agricultural practices leads to accelerate the rate of shrinking (Saha and Pal 2019b; Bassi et al. 2014). The study area moribund deltaic region of the Ganga–Brahmaputra delta is also facing similar situations (Paul and Pal 2020a). This region is enriched with many back swamps, sloughs, oxbow lakes, residual channels, and marshy lands of various sizes. Those wetlands are one of the major sources of various hydrological resources and also act as a corridor for many ecologically sensitive species (Bala and Mukherjee 2010). Studies like Paul and Pal (2020a) reported that 47.31% wetland of this region has been transformed seasonally due to extensive agricultural practices. This seasonal drying out process accelerates the rate of wetland conversion into agricultural lands and built-up land permanently (Paul and Pal 2020b). These factors are very crucial for determining the fate of the wetland, and therefore, these are considered conditioning parameters for wetland habitat modelling.

From the environmental perspective, wetland vulnerability is an ensemble of various natural and artificial factors like rainfall anomaly, lowering down of groundwater table, extensive land use/land cover (LU/LC) change, loss of connectivity with active recharge points, and climatic change (Pal and Talukdar 2019; Finlayson 2006). Current remote sensing (RS) and GIS have such capability to explore the nature of such change using various spatial models like the wetland vulnerability index (WVI) model (Defne et al. 2020), Pressure, State, Impact, Response (PSIR) framework (Mosaffaie et al. 2021), multivariate adaptive regression spline (MARS) (Adnan et al. 2021). Recently, data-driven and knowledge-driven models such as statistical index (SI) (Li et al. 2020), linear discriminant analysis (LDA) (Nie et al. 2020), artificial neural network (ANN) (Paul and Pal 2020b), support vector machines (SVM) (James et al. 2013), Bagging (Chhabra et al. 2021), Boosting (Pal and Paul 2021b), decision tree (DT) (Luo et al. 2021a, b), random forest (RF) (Granger et al. 2021), Random subspace (Talukdar et al. 2021), Reduced Error Pruning (REP) Tree (Pal and Paul 2020), boosted regression trees (BRT) (Shaziayani et al. 2021), Evidential belief function (EBF) (Ghosh 2021), deep belief network (DBN) (Scarpiniti et al. 2021), and naive Bayes (NB) (Costache et al. 2021) are also used to measure different spatial phenomenon like flood and landslide susceptibility mapping (Jacinth Jennifer and Saravanan 2021), groundwater potentiality mapping, climate forecasting (El-Magd and Eldosouky 2021; Lin et al. 2021). Studies show that the machine learning (ML) models have the capability to predict spatial phenomena on large datasets and give more accurate results than traditional statistical techniques (Pal and Paul 2020). Individual machine learning techniques also has some of their strengths and weaknesses, therefore, different ensemble techniques have been introduced to reduce such weaknesses (Rabbani et al. 2021; Pal and Paul 2021a). In a multi-model approach, validation of the model is necessary to check the accuracy level and also ensure the performance of the employed models (Ling et al. 2021). Studies like Mohana et al. (2021), and Qolipour et al. (2021), reported that tree-based ensemble ML algorithms have such capability to perform better than generic algorithms. Recently, various advanced machine learning (ML) algorithms have been incorporated for modelling various environmental phenomena, but the tree-based multi-model approach for wetland vulnerability mapping is rare, especially in this region. Therefore, in this present study, we have attempted to employ multiple tree-based machine learning techniques for modelling wetland vulnerability in the moribund deltaic part of India. Different matrices and an extensive field investigation have been done to validate the performance of the employed models.

Previously, it is mentioned that the moribund of the Ganga–Brahmaputra deltaic region is prone to rapid wetland loss and hydrological transformation due to rapid anthropogenic pressure and infrastructural developments. Previous studies like Mandal and Pal (2017), Paul and Pal (2020a, b), and Everard et al. (2019) focused on only the transformation and dynamic nature of wetlands but no study focused on the degree of risk faced by the habitat or habitat risk areas which have a long-term effect on habitat ecology and ecosystem services of the wetland. Therefore, this study attempts to identify vulnerable areas of wetland with the help of multiple tree-based machine learning algorithms.

2 Study Area

The present study area, moribund deltaic region, is a part of the great Ganges–Brahmaputra delta of Indo-Bangladesh. It extends from 24°30′ N/88° E to 23° N/89°45′ E with a total area of 7685.99 km^2. The extension of our present study is 23°24′35″ N/88°15′50″ E to 23°42′30″ N/88°33′20″ E with an area of 3927 km^2 (Fig. 1). The Ganges–Brahmaputra delta is divided into three geomorphological units and spread across an administrative unit of India and Bangladesh (Bagchi and Mukherjee 1983). The Wetland of this region is enriched by fertile alluvial soil

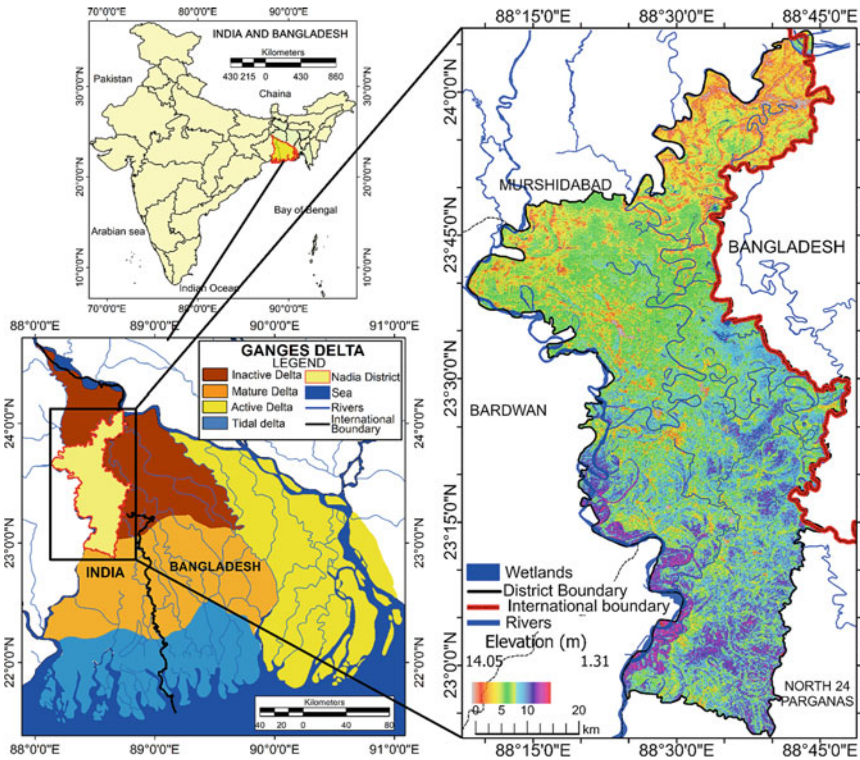

Fig. 1 Geolocation of the study area, moribund delta (Indian part)

and inundated water cane from rivers and depends on the seasonal rainfall regime, and therefore, most of the wetlands are seasonal (Bala and Mukherjee 2010). Since the region receives 80% of total annual rainfall (1450 mm) during the monsoon season, the wetland area gets its maximum areal extent during this season. With an interconnected network of streams like Bhagirathi–Hooghly, Jalangi, Ichamati, and Churni, numerous active riverine morphometric structures create a thick layer of fertile alluvium, silt, and clay layer in this region (Majumdar 1978). Agrarian economy determines the nature of wetland transformation to a large extent.

3 Materials and Methods

3.1 Materials

For this present study, Landsat TM 4–5 imageries ETM and OLI imageries from the United States Geological Survey (USGS) from 1988 to 2017 (Path/row: 138/43,44;

spatial resolution: 30 m) have been used to prepare wetland map, water depth map, seasonal dynamics of wetland, vegetation and agriculture cover map, and built-up area map. Open Street Map (OSM) has been used to prepare a road map. Survey of India (SOI) toposheets have been used to prepare river map. Whereas, the administrative map of Nadia district was used to demarcate the study area because all parts of the Nadia district come under the moribund part of the Ganga–Brahmaputra deltaic region. Extensive field surveys and high-resolution *Google Earth* imageries have been used to validate wetland map, built-up map, agricultural map, and road map. A total of 540 sites has been selected for validating the models.

3.2 Methods

3.2.1 Data Layers Preparation for Wetland Vulnerability Assessment (WVA)

Eight spatial data layers have been considered for this present study, among which five parameters are related to the hydrodynamics of the wetland, namely, water presence frequency (WPF), water depth, change in WPF, hydro duration, and proximity from the river. The remaining three parameters such as distance from the road network, built-up proximity, and agricultural presence frequency (APF) are related to LU/LC dynamics. For this present study, the range of available data (1988–2017) has been divided into three phases in a decadal manner such as phase I (1988–1997), phase II (1998–2007), and phase III (2008–2017). Due to the lack of availability of the change in the WPF layer for phase I, this phase has been excluded.

The normalized differences water index (NDWI) (McFeeter 1996) has been calculated for each Landsat image (1998–2017) to identify the wetlands. According to the studies conducted by Mandal and Pal (2017) and Das and Pal (2016), the NDWI technique of surface water detection gives better sensitivity for the Indo-Gangetic region. The NDWI value is higher in greater water depth areas. The NDWI map is used to prepare water presence frequency (WPF) and also wetland depth mapping for this study. Recent NDWI layers have been validated using 987 reference sites selected from Google earth images and field sites. The computed Kappa coefficient (K) value ranges from 0.86 to 0.95, which indicates an excellent match between image-based wetland map and ground reality. The equation to calculate NDWI is as follows:

$$\mathrm{NDWI} = \frac{b_{\mathrm{green}} - b_{\mathrm{NIR}}}{b_{\mathrm{green}} + b_{\mathrm{NIR}}} \tag{1}$$

where NDWI = Normalized Differences Water Index; the green band is indicated by b_{green}; and the near infra-red band is indicated by b_{NIR}. The NDWI value ranges from −1 to 1, where pixel value towards positive 1 indicates maximum availability of water.

Water presence frequency (WPF) indicates the frequency of appearance of water pixels within a selected temporal frame (Borro et al. 2014). Consistent appearance of water pixels considered high WPF and inconsistent appearance of water pixel considered as low WPF (Paul and Pal 2020a). Therefore, WPF can be an important indicator to determine the habitat health status of the wetland. To prepare the WPF layer, each NDWI layer has been converted to the binary image where the presence of water pixel is considered as 1 and non-water pixel is considered as 0. Thereafter, images of each decade have been summed up to prepare a decadal WPF map (Figs. 2, 3). The total frequency of water presence within a temporal span is considered as 100%. The WPF has been divided into three categories such as low WPF (<33%), moderate WPF (33–67%), and high WPF (>67%).

$$\mathrm{WPF} = \frac{\sum_{i=1}^{n} \mathrm{NDWI}}{N_I} \times 100 \tag{2}$$

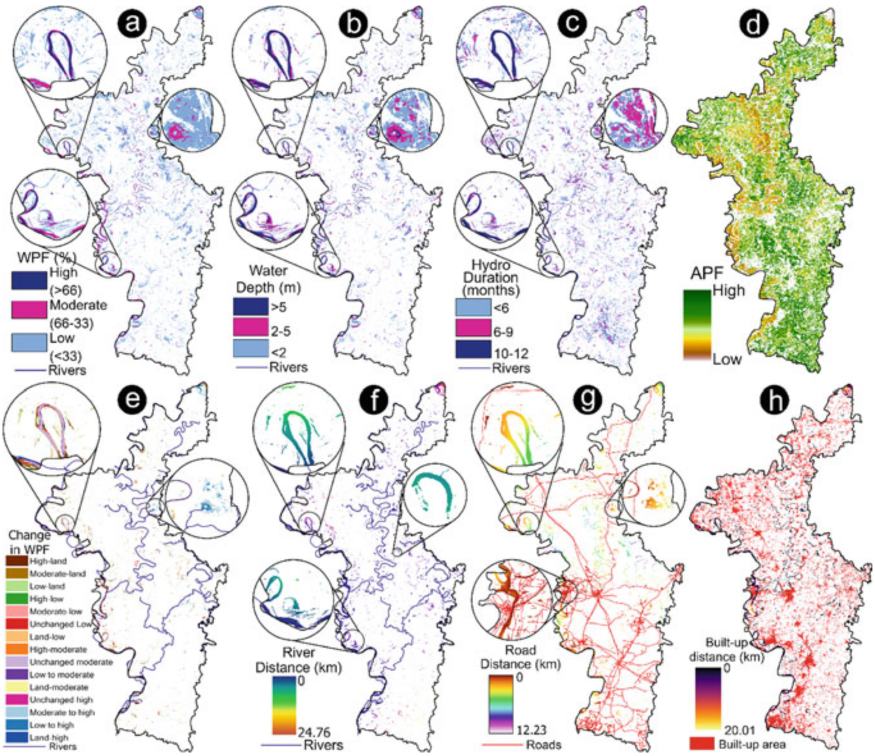

Fig. 2 Incorporated data layers for wetland vulnerability assessment of phase II **a** WPF, **b** water depth, **c** hydro duration, **d** APF, **e** change in WPF, **f** distance from river, **g** distance from road, and **h** distance from the built-up area

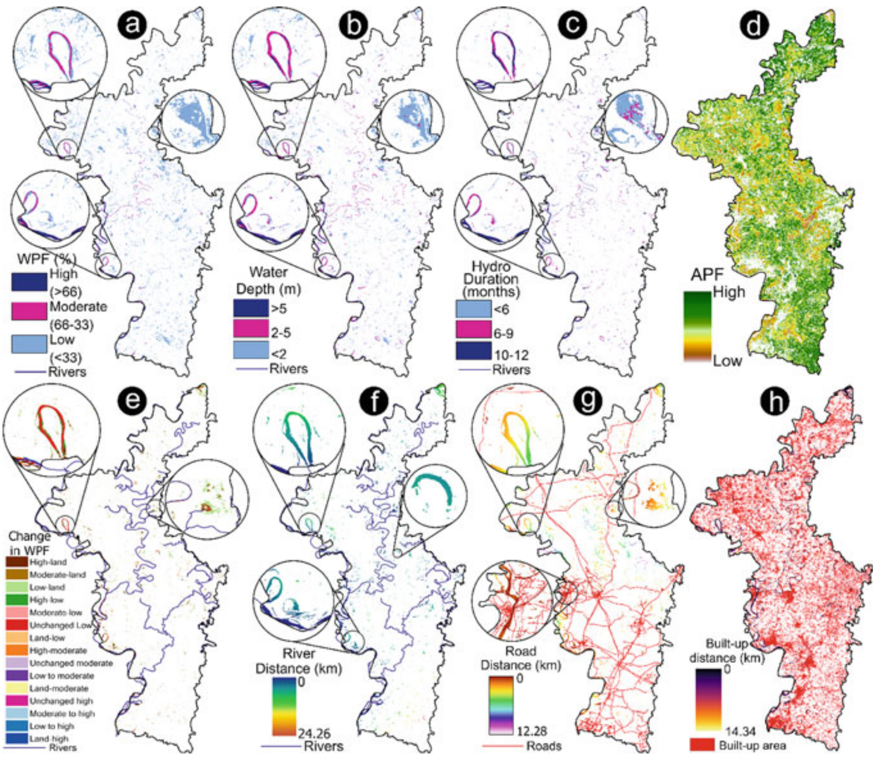

Fig. 3 Incorporated data layers for wetland vulnerability assessment of phase III **a** WPF, **b** water depth, **c** hydro duration, **d** APF, **e** change in WPF, **f** distance from river, **g** distance from road, and **h** distance from the built-up area

In this equation, I_{NDWI} is the frequency of water presence at the Ith pixel, and N is the total number of years.

Imaged-based change detection analysis has been done to detect the decadal change in WPF in the ArcGIS environment. Water depth indicates the potentiality of hydrological richness (Pal and Paul 2021b). NDWI maps have been used to prepare the depth map for this study as the pixel value varies with water depth (Paul and Pal 2021). 30 field-based wetland depth data have been used to calibrate the depth maps. Hydro-duration is another important parameter that indicates the presence and availability of saturated soil periodically (Pal and Sarda 2021). Hydro-duration map phase II and III prepared using 1999 and 2007 for phase and 2010 and 2017 averaging annual hydro-duration maps. This technique has been adopted due to the lack of monthly data for different phases.

In this floodplain deltaic region, wetlands in the riparian region often lost their link with feeding channels in the pre-monsoon season (Pal and Saha 2018). These small tie channels are very important for maintaining water availability and the hydro-ecological health of the wetland (Kundu et al. 2021). Studies made by Pal and Paul

(2021a) and Debanshi and Pal (2020) reported that the wetlands located far away from the perennial channels are generally unable to maintain their stable hydrological status throughout the year. Therefore, distance from rivers can be considered as a potential parameter to check the vulnerability status of the wetland. Phase-wise distance map has been prepared using the Euclidian tool in the ArcGIS environment. The rapid extension of infrastructure causes fragmentation of wetland scape and it is caused for disconnection among the fragmented wetland units and increasing anthropogenic pressure (Grzybowski and Glińska-Lewczuk 2019).

The current study region comprises some populated urban areas like Kalyani, Krishnanagar, Ranaghat, Santipur, and Nabadwip. These cities are well-connected with very dense road networks. Therefore, distance from the road and the built-up area is considered as a potential parameter for detecting the status of vulnerability due to anthropogenic pressure towards the wetland area (Islam et al. 2021). Phase-wise NDBI has been calculated to prepare the built-up distance map and Open Street Map (OSM) road layer is used to prepare the Euclidian road distance map (Figs. 2, 3). In the floodplain region, wetland capture through agricultural encroachment is very much visible (Pal and Saha 2018).

$$\mathrm{NDVI} = \frac{(b_{\mathrm{NIR}} - b_{\mathrm{red}})}{(b_{\mathrm{NIR}} + b_{\mathrm{red}})} \tag{3}$$

$$\mathrm{NDBI} = \frac{(b_{\mathrm{MIR}} - b_{\mathrm{NIR}})}{(b_{\mathrm{MIR}} + b_{\mathrm{NIR}})} \tag{4}$$

Normalized Difference Vegetation Index (NDVI) from 1998 to 2017 (Eq. 3) has been taken to prepare yearly vegetation and cropland maps. Thereafter, each NDVI map is converted into a binary map by providing1 for vegetation and cropland and 0 for the non-vegetated area. These maps are summed up phase-wise to prepare an agricultural presence frequency (APF) map. APF varies from 0 to 100%, where a value near 100% often indicates a consistent cropping area (Figs. 2, 3).

3.2.2 Modelling of Wetland Vulnerability

For this study, four tree-based ML classifiers have been used, namely: Reduced Error Pruning (REP) Tree, Gradient boosting classification model (GBM), AdaBoosting classification model (ADB), and Bagging classification model. A methodological overview of the employed models is discussed below.

Reduced Error Pruning Tree (REP Tree)

The reduced error pruning Tree or REP tree is considered to be a relatively fast decision-making algorithm that uses the pruning method to reduce the complexity of a data model and minimizes the model error (Pham et al. 2019; Sattari et al. 2021).

This pruning process simplifies the model tunning and structurization process and also saves more time during the training (Pham et al. 2019). The pruning process also reduces the overfitting problem and provides better accuracy to the model (Zharmagambetov et al. 2021). There are two types of pruning processes: pre-pruning and post-pruning. The pre-pruning is a relatively faster process with lesser accuracy than the post-pruning (Shahabi et al. 2021). For this study, we have applied the post-pruning technique to assess the wetland vulnerability.

Bagging Classification Model

Bagging or bootstrap algorithm uses bootstrapping technique to reduce the noise in a dataset and improve the performance of the model (Luo et al. 2021a, b; Song et al. 2021). It is one of the primitive ensemble models which generates multiple randomly to form a training set (Wen and Hughes 2020; Jain and Xu 2021; Ankar and Yadav 2021). The bagging model decreases the variance of classification error to improve classification accuracy. For this study, we have used the *SK-learn* package in python to run this model.

Gradient Boosting Classification Model (GBM)

The gradient boosting classification model (GBM) is an ensemble model which uses a decision tree algorithm under the hood (Zhang et al. 2021; Abdi 2020). The GBM algorithm uses boosting tree technique over the generic tree-based algorithm for optimizing the model accuracy and performance (Yang et al. 2021). The GBM algorithm replaces “best-fit” optimization with a “weak” learner model for staking the model and applying aggregation of the existing dataset (Jun 2021). GBM has high predictive power over RF, but for noisy data, sometimes it leads to overfitting (Yang et al. 2021). For this study, *GradientBoostingClassifie*r from the Scikit-learn ensemble package has been used for wetland vulnerability mapping.

AdaBoosting Classification Model (ADB)

Adaptive boosting or AdaBoosting is also a decision tree-based ensemble model designed to improve the performance and efficiency of binary classifiers (Zharmagambetov et al. 2021). Like other ensemble models, AdaBoosting also uses an iterative process to learn the mistakes of multiple weak classifiers and improve the model’s performance (Zharmagambetov et al. 2021; Walker 2021). AdaBoost classifier tools from the *SK-learn* ensemble library have been used in this study.

3.2.3 Data Preparation and Training of the Models

To construct a wetland vulnerability model, vulnerability conditioning factor layers water presence frequency (WPF), change in WPF, water depth, hydro duration and proximity from the river, wetland road distance, built-up proximity, and Agricultural presence frequency) have been converted to grid cell format with a spatial resolution of 30 m. Subsequently, the frequency ratio for each lower-class area of WPF (<33%) and depth map has been taken to identify poor wetland habitat areas for both phases. These maps have been used to extract training and validation datasets for the ML algorithms.

3.2.4 Parameter Optimization of the Models

K-fold cross-validation technique along with hyperparameter optimization technique like *GridSearch CV* method has been applied to optimize the model. All ML algorithms are optimized to a certain number of iterations using the grid search technique to generate hyper-parameters (Daviran et al. 2021). The training sets have been split into some equal random k-sets for training and validation of the model as a standard procedure (Wen and Hughes 2020). For better performance and accuracy, 5- and tenfold *K* iterative processes for 240 candidates have been run to generate 1200 and 2400 fits for each model.

3.2.5 Evaluation and Comparison Methods

In this study, the models have been evaluated using six matrices, namely: sensitivity, precision, FI-score, and MCC. The confusion matrix for the training and validation dataset consisted 2×2 contingency table from which four types of evaluation results have been categorized as, true positive or TP, false positive or FP, true negative or TN, and false-negative or FN. The TP and TN are correctly classified data, whereas the FP and FN part of datasets are incorrectly classified results. Based on these four classification results, sensitivity, precision, FI-score, and MCC are calculated using the following equations:

$$\text{Sensitivity} = \frac{Tp}{Tp + Fn} \tag{5}$$

$$\text{Precision} = \frac{Tp}{(Tp + Fn)} \tag{6}$$

$$F1 - \text{score} = \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{7}$$

$$\text{MCC} = \frac{Tp \times Tn - Fp \times Fn}{\sqrt{(Tp + Fp)(Tp + Fn)(Tn + Fp)(Tn + Fn)}} \tag{8}$$

3.2.6 Field-Based Validation Method to Determine Wetland Vulnerability

An extensive field investigation has been conducted to validate and assess the performance of the models and also for measuring the physical vulnerability status of wetlands. For this study, 30 wetlands have been studied from different parts of this region. The selection criteria of those wetlands include site, situation, wetland type, and distance from the feeding channel. A total of 12 vulnerability controlling factors like the connection with a nearby stream, wetland area change, quantity of natural and artificial inflows, quantity of surface outflows, hydrological period of wetland, depth (average), water level fluctuation (monthly), and wetland eutrophicated area are the physical factors, whereas cultivation of fish, presence or absence of agriculture practice, area encroached for agriculture, are considered as anthropogenic factors those are considered to evaluate the vulnerability status of the wetland. A composite rank score has been calculated to derive a factor-wise score and then the final score has been generated using the averaging technique in SPSS software to generate the final wetland vulnerability index (WVI).

4 Results

4.1 Characteristics of the Parameter Layers

Before assimilating all the eight layers, spatial variation of individual parameters can be quantified to comprehend the nature of each factor used for measuring wetland vulnerability. The overall wetland area has decreased from 150.38 to 80.63 km^2, which means more than 45% of the wetland area was lost between phase II to phase III. In case of water presence frequency (WPF), the area under moderate WPF has lowered from 60.57 to 24.59 km^2 between phase II to phase III during the post-monsoon season. The area under high (>5 m) wetland depth also decreases from 41.81 to 23.70 km^2, which indicates that many wetlands have been dried out during phase II to phase III. In the fragmentation dataset, the large core area decreases from 20.60 to 13.63 km^2 from phase II to phase III. The edge and patch areas also decrease from 55.59 to 5.94 km^2 and 24.22 to 13.63 km^2, which indicates growing wetland fragmentation and increasing pressure on the human landscape. The core area of wetlands is less affected as compared to edge and patch areas. In agricultural presence frequency, the area under the high APF zone is increased from phase II to phase III indicates that the low WPF areas area converted into permanent or semipermanent agricultural land by extending agricultural areas. Similarly, the built-up area towards the wetland is rapidly increased from phase II to phase III, which is also a cause for rapid wetland conversion. It is the fact that all the parameters are not concentrated in the same spatial unit or same spatial variability, therefore, to produce the final vulnerability model, it is necessary to integrate all the spatial layers.

4.2 Wetland Vulnerability Assessment (WVA) Modelling

Based on the tree-based machine learning technique, four vulnerability models for phase II and phase III have been developed. Each output model is classified into five subtypes (starting from very low to very high vulnerability) based on varying vulnerability intensities (Figs. 4, 5). In phase II, 102.88 km^2 (2.64%), 102.55 km^2 (2.63%), 104.55 km^2 (2.68%), and 106.96 km^2 (2.74%) areas are predicted as very high vulnerable category by the Bagging, REP Tree, ADB and GBM models, respectively (Table 1). These four vulnerability models indicate that more than 2.5% of wetland area belongs to a very high vulnerable zone in phase II. In phase III, the area under the very high vulnerable zone has declined to 53,67 km^2 (1.37%), 52.91 km^2 (1.36%), 55.37 km^2 (1.42%), and 56.12 km^2 (1.44%) for all four models in the same order. The area under high vulnerable area is almost twice than the very high vulnerable area. Wetland proximity to a perineal channel(s) tends to be hydrologically more secure than the wetland located away fromthe river. The overall wetland area under different vulnerable zones reduces from 426.28 km^2 (10.92%) to 215.14 km^2 (5.51%). This indicates that almost 50% of the wetland area lost since phase II among which most of the wetlands belong to the high to very highly vulnerable wetland category (Figs. 4, 5).

4.3 Assessing the Accuracy of the WVA Models

Table 2 depicts the model validation result for WVA using sensitivity, precision, F1-score, and MCC. The value for the matrices ranges from 0 to 100, where a value of 100 indicates good accuracy >88% accuracy level found in case of all the applied models. The model's sensitivity score is more than 89 in case of bagging and REP tree classifiers. Whereas, the sensitivity score increases to more than 90 for ADB and BGM models. Precision, recall, and MCC score has similarity to sensitivity scores for all the models (Table 2). The accuracy level for *k10* is lower than the *k5* value. The model's accuracy level tends to be higher in phase III as compared to phase II (Table 2). Among the four ML models, it is observed that the GBM and ADB models perform better in comparison to bagging and REP Tree models for both the phases II and III. Apart from this, the overall performance of all ML models is good for wetland vulnerability assessment and mapping.

4.4 Factor-Based Wetland Vulnerability Index (WVI) Analysis Using Filed Data

Based on the average rank score, the wetland vulnerability index (WVI) has been calculated on 30 selected wetlands in this region. The average score value has

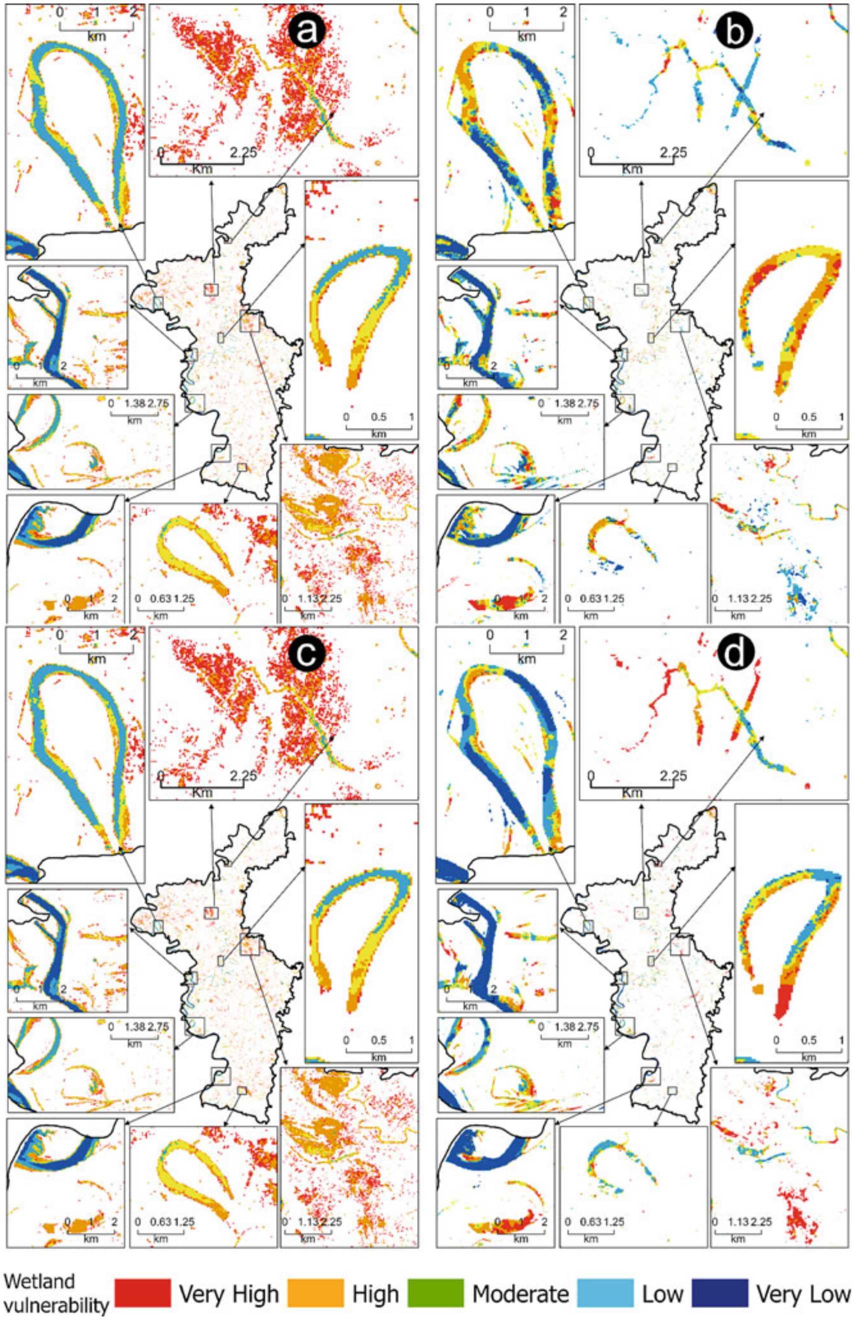

Fig. 4 Wetland vulnerability zones derived from bagging classifier **a** phase II, **b** phase III and REP Tree classifier, **c** phase II, and **d** phase III

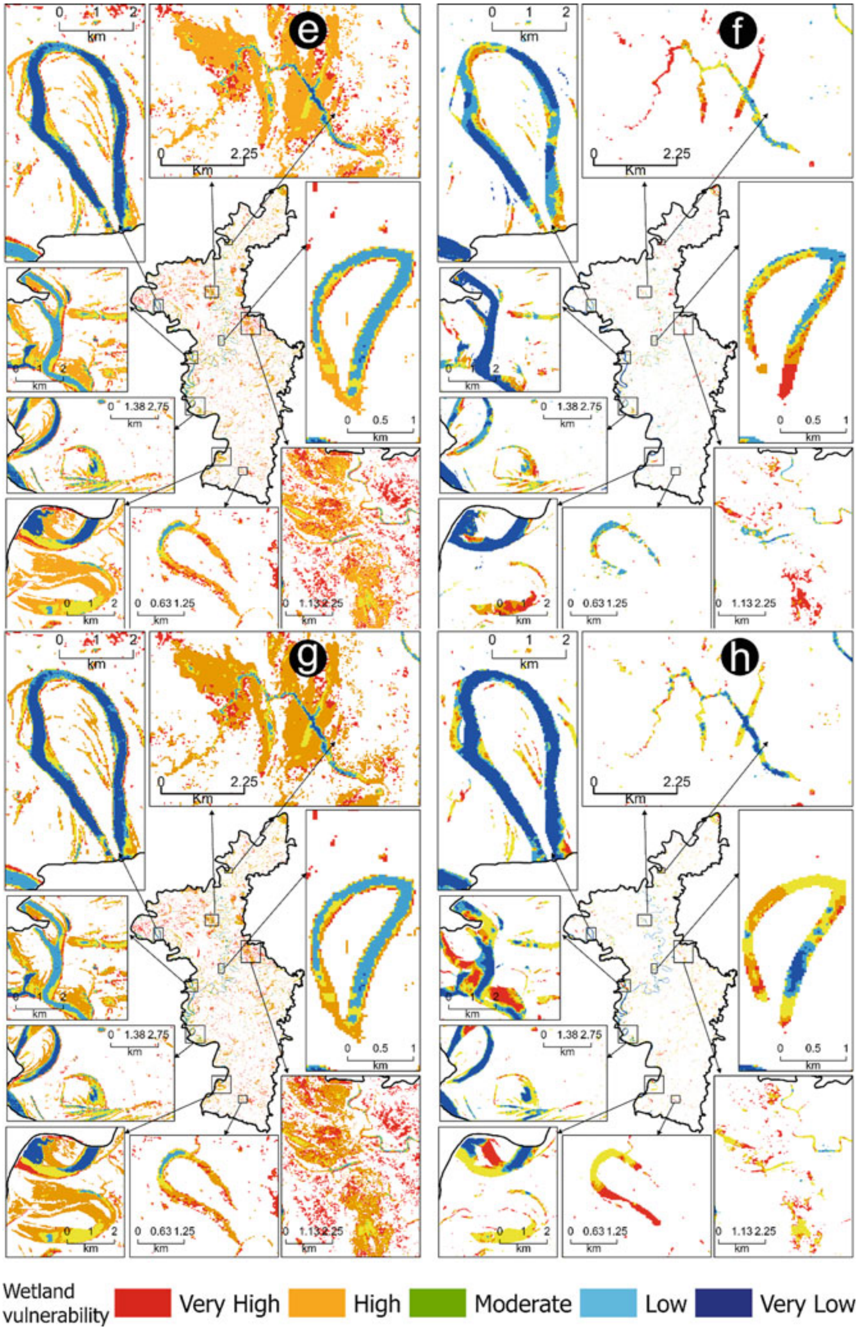

Fig. 5 Wetland vulnerability zones derived from ADB classifier **e** phase II, **f** phase III and GBM classifier, **g** phase II, and **h** phase III

Table 1 Phase-wise wetland vulnerability status based on different ML models

Phases	Vulnerability status	Very high		High		Moderate		Low		Very low	
	ML models	Area (km^2)	Percentage to the total area	Area (km^2)	Percentage to the total area	Area (km^2)	Percentage to the total area	Area (km^2)	Percentage to the total area	Area (km^2)	Percentage to the total area
Phase II	Bagging	102.88	2.64	197.65	5.06	59.51	1.52	41.43	1.06	19.75	0.51
	REP tree	102.55	2.63	183.16	4.69	62.88	1.61	42.98	1.10	18.43	0.47
	ADB	104.55	2.68	198.45	5.08	60.17	1.54	44.04	1.13	19.43	0.50
	GBM	106.96	2.74	210.16	5.38	63.78	1.63	45.68	1.17	20.67	0.53
Phase III	Bagging	53.67	1.37	94.19	2.41	26.88	0.69	20.81	0.53	11.64	0.30
	REP tree	52.91	1.36	91.83	2.35	25.40	0.65	20.33	0.52	11.74	0.30
	ADB	55.37	1.42	102.58	2.63	27.49	0.70	20.87	0.53	11.78	0.30
	GBM	56.12	1.44	114.49	2.93	29.67	0.76	21.03	0.54	11.74	0.30

Table 2 Ground truth accuracies of the models

Phase	Classifiers	K-fold	Sensitivity	Precision	F1-score	MCC	Support
Phase II	Bagging	5	89.33	0.89	0.87	0.87	40,000
		10	89.76	0.86	0.89	0.88	40,000
	REP tree	5	89.23	0.87	0.89	0.88	40,000
		10	90.11	0.84	0.90	0.89	40,000
	ADB	5	91.50	0.89	0.91	0.90	40,000
		10	91.45	0.88	0.88	0.87	40,000
	GBM	5	95.56	0.89	0.87	0.86	40,000
		10	93.21	0.87	0.88	0.89	40,000
Phase III	Bagging	5	89.52	0.89	0.88	0.89	40,000
		10	90.73	0.88	0.89	0.88	40,000
	REP tree	5	90.21	0.87	0.88	0.89	40,000
		10	89.71	0.86	0.89	0.90	40,000
	ADB	5	91.29	0.89	0.87	0.88	40,000
		10	92.46	0.88	0.87	0.87	40,000
	GBM	5	93.65	0.91	0.92	0.90	40,000
		10	92.86	0.89	0.92	0.91	40,000

been reclassified into five sub-categories similar to wetland vulnerability mapping (Table 3). Wetlands like Chaldoba Beel (wetland), Gupiyar Beel, Chuchokhola Beel, Sukna Beel, Mora Ganga, and Charganga 2 are identified as very high vulnerable wetlands with WVI ranges from 11.75 to 10.33. Whereas, wetlands like Bachamari Beel, Padmamala Beel, Chamta Beel, Boro Beel, Gorgore Beel, Nilkuri Beel, and Nabadwip Municipality Lake belongs to high vulnerable wetland category with a WVI score ranging from 9.67 to 9.00. The wetlands like Digri Beel, Chand Beel, Anjana, Chakla Beel, Bhomra Beel, Khalsi Beel, Charganga 1, Tungi Beel, Bahluka Beel, Majhdia Doapara Beel, and Hasnadanga Beel with well-connected recharge points and stable hydro-ecological characteristics belong to low to very low WVI category (Table 2). The correlation coefficient between WVI and WVA ranges from 0.88 to 0.93 which is significant at a 0.01 level of significance. The GBM model with a correlation value of 0.93 came out to be the most significant.

5 Discussion

Wetland risk or vulnerability assessment is a fundamental step towards wetland management and planning. In this present study of the Indian moribund deltaic floodplain region, seasonal hydrological alteration and human intervention towards the wetland area are found as the major triggering factors for exposing wetland habitat towards vulnerability. Intensive agricultural practices during the pre-monsoon season

Table 3 Calculated wetland vulnerability score of some selected wetlands

Wetlands	WVI	State of vulnerability	Wetlands	WVI	State of vulnerability
Chaldoba Beel	11.75	Very high	Arangsartsa Beel	8.42	Moderate
Gupiyar Beel	11.67		Chokar Beel	8.25	
Chuchokhola Beel	11.33		Arpara Beel	7.83	
Sukna Beel	10.83		Mathura Jhil	7.00	
Mora Ganga	10.33		Digri Beel	6.92	Low
Charganga 2	10.33		Chand Beel	6.58	
Bachamari Beel	9.67	High	Anjana	6.42	
Padmamala Beel	9.58		Chakla Beel	6.42	
Chamta Beel	9.42		Bhomra Beel	6.33	
Boro Beel	9.42		Khalsi Beel	6.25	
Gorgore Beel	9.17		Charganga 1	5.50	Very Low
Nilkuri Beel	9.08		Tungi Beel	5.08	
Nabadwip Municipality Lake	9.00		Bahluka Beel	4.42	
Gayshpurkhulia Jhil	8.83	Moderate	Majhdia Doapara Beel	4.17	
Muktaduar Beel	8.50		Hasnadanga Beel	3.92	

intensify the magnitude of wetland loss. The WPF change detection statistics indicate that there is a huge reclamation of agricultural land for crop cultivation which leads to extensive wetland loss. This is also supported by the statistics where more than 50% of the wetland area has been lost during phases II to III. Whereas, the agricultural and vegetation cover increased from 2497.85 to 2654.05 km^2 during phase II to phase III. Scholars like Sampson (2021), Fickas et al. (2016), and Saha and Pal (2019a) reported similar conversions of wetland areas due to agricultural extension in their studies across flood plain wetland. The present study also shows that the process of wetland conversion is deeply related to the rate of wetland fragmentation by which moderate to moderately large wetlands are divided into many small numbers of patches with small core areas. Small wetlands are the most vulnerable to such conversion and loss, whereas the large wetlands with relatively stable core areas somehow maintain their integrity. But the edge area of such large wetlands is significantly shrunken from phase II to phase III (Figs. 4, 5). The large core area decreases from 20.60 to 13.63 km^{2}during phase II to phase III. Extension of the built-up area and connected transportation networks increased from 913.58 to 1094.27 km^2 from phase II to phase III which is a reason behind wetland fragmentation in this region. Change in WPF also indicates that a large number of wetlands converted into ortho fluvial wetland from para fluvial wetland from phase I to II and also phase II to III (Figs. 2, 3). Rainfall is the only source of water for these orthofluvial wetlands, which rarely receive water from the river. Also, connectivity loss from the main feeder channels and lowering of groundwater table due to anthropogenic interferences negatively affects the habitat condition of these wetlands (Pal et al. 2020; Gómez-Baggethun et al. 2019).

Apart from the contemporary conditioning factor, historical hydrological evolution is also a reason behind the present geomorphic setting of the wetlands. Historically (around the sixteenth century), this part of the deltaic region had gone through massive hydrological alteration. In 1975, after the construction of the Farakka barrage, a massive wetland conversion in this region has been occurred (Hirst 1916; Pal 2011). Studies by Paul and Pal (2020a) reported a loss of 63.34% of wetland area from 1987 to 2017.

This present study successfully explored the predictability of tree-based WVA using four tree-based machine learning approaches. The result of four ML models is compared with field-based data to check the applicability of the models. Among the four models, ADB and GBM model performs better and gives better accuracy in comparison to the Bagging and REP Tree model (Table 2). The overall performance of the GBM model is better than the other three models. But overall, all four models perform more than satisfactory as found from the accuracy assessment matrix results (Table 2). The bagging and REP Tree model was reportedly given better results in the studies like Pal and Debanshi (2021a, b), Talukdar et al. (2021), Pal and Paul (2020), and Khatun et al. (2021). The accuracy level of all four models has been increased in phase III and also for fivefold *K* classification. The complex ensemble models have proved their superiority over the comparatively simple bootstrap algorithms. Studies made by Chen et al. (2018a, b, c) and Han et al. (2019) also reported better performance of ensemble models over the generic ML models. It should be mentioned that the spatial extension of WVA zones varies through their geospatial distribution which is similar to each other. Since the wetland habitat is a complex interactive system, and it is controlled by different controlling factors, all the factors do not equally impact control on the spatial extension of vulnerability. Hydrological factors are found more dominant than LU/LC factors. This finding has further clarified that hydrological modification is a dominant reason for wetland conversion and promoting land use transformation causing wetland loss. Pal and Paul (2021b), and Debanshi and Pal (2020), also reported that hydrological insecurity enhances landscape insecurity.

Wetland vulnerability assessment using such advanced tree-based ML algorithms is often rare but there is further scope for research in the future. The present study has focused on only mapping the state of vulnerability and changing the nature of such vulnerability without considering the parameters like river discharge, quality of water, and ecological productivity issues. The inclusion of those factors may uplift the quality of the result. The lack of such extensive spatial data availability restricted us to incorporate those datasets. In future studies, the incorporation of such extensive data regarding complex ecological phenomena can improve the acceptability of such studies a bit more.

6 Conclusion

This present work has assessed wetland vulnerability based on eight decisive parameters using four tree-based ML models. All the models have been validated using five matrices and found to be sensitive. Among all four models, AdaBoosting (ADB) and gradient boosting (GBM) are found as the most accurate to predict vulnerable wetland areas. Very high vulnerable areas have increased over the phases as per all the models. Wetland under greater exposure to the human landscape is more vulnerable to transformation. Hydrological parameters are found to be more important for explaining the vulnerability of wetlands. Hydrological transformation is found as a promoting factor behind land use transformation. From the management perspective, wetland vulnerable models are very important since it provides a database for the future wetland restoration plan. Moreover, since the study has identified hydrological factors are playing a decisive role in wetland habitat transformation, it will be very good information regarding the way of wetland conservation and restoration. In addition to this, the present study tries to build a methodological knowledge addition for wetland vulnerability study which will be helpful for other types of environmental risk assessment studies. In this consonance, the study recommends the use of hybrid tree-based ensemble ML models instead of simple ML models for similar works.

References

Abdi AM (2020) Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. Gisci Remote Sens 57(1):1–20. https://doi.org/10.1080/15481603.2019.1650447

Adnan RM, Khosravinia P, Karimi B, Kisi O (2021) Prediction of hydraulics performance in drain envelopes using Kmeans based multivariate adaptive regression spline. Appl Soft Comput 100:107008

Akpabio EM, Umoh GS (2021) The practical challenges of achieving sustainable wetland agriculture in Nigeria's Cross River basin. Water Int 46(1):83–97

Ankar SJ, Yadav A (2021) A high-speed protection strategy for bipolar CSC-based HVDC transmission system. Electric Power Comp Syst 49(1–2):48–66

Assessment ME (2005) Ecosystems and human well-being: wetlands and water

Bagchi K, Mukerjee KN (1983) Diagnostic survey of West Bengal(s). Dept. of Geography

Bala G, Mukherjee A (2010) Inventory of wetlands of Nadia district, West Bengal, India and their characterization AS. J Environ Sociobiol 7(2):93–106

Borro M, Morandeira N, Salvia M, Minotti P, Perna P, Kandus P (2014) Mapping shallow lakes in a large South American floodplain: a frequency approach on multitemporal Landsat TM/ETM data. J Hydrol 512:39–52

Bui DT, Pradhan B, Revhaug I, Tran CT (2014) A comparative assessment between the application of fuzzy unordered rules induction algorithm and J48 decision tree models in spatial prediction of shallow landslides at Lang Son City, Vietnam. In: Remote sensing applications in environmental research. Springer, Cham, pp 87–111

Chaffron S, Delage E, Budinich M, Vintache D, Henry N, Nef C, Eveillard D (2021) Environmental vulnerability of the global ocean epipelagic plankton community interactome. Sci Adv 7(35):eabg1921

Chen F, Yu B, Li B (2018a) A practical trial of landslide detection from single-temporal Landsat8 images using contour-based proposals and random forest: a case study of national Nepal. Landslides 15(3):453–464

Chen W, Li H, Hou E, Wang S, Wang G, Panahi M, Ahmad BB (2018b) GIS-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models. Sci Total Environ 634:853–867

Chen W, Shahabi H, Zhang S, Khosravi K, Shirzadi A, Chapi K, Ahmad BB (2018c) Landslide susceptibility modeling based on gis and novel bagging-based kernel logistic regression. Appl Sci 8(12):2540

Chhabra M, Shukla MK, Ravulakollu KK (2021) Bagging-and boosting-based latent fingerprint image classification and segmentation. In: International conference on innovative computing and communications. Springer, Singapore, pp 189–201

Codagnone C, Bogliacino F, Gómez C, Charris R, Montealegre F, Liva G, Veltri GA (2020) Assessing concerns for the economic consequence of the COVID-19 response and mental health problems associated with economic vulnerability and negative economic shock in Italy, Spain, and the United Kingdom. PLoS ONE 15(10):e0240876

Costache R, Arabameri A, Moayedi H, Pham QB, Santosh M, Nguyen H, Pham BT (2021) Flash-flood potential index estimation using fuzzy logic combined with deep learning neural network, naïve Bayes, XGBoost and classification and regression tree. Geocarto Int 1–28

Das RT, Pal S (2016) Identification of water bodies from multispectral landsat imageries of Barind Tract of West Bengal. Int J Innov Res Rev 4(1):26–37

Daviran M, Maghsoudi A, Ghezelbash R, Pradhan B (2021) A new strategy for spatial predictive mapping of mineral prospectivity: automated hyperparameter tuning of random forest approach. Comput Geosci 148:104688

Debanshi S, Pal S (2020) Wetland delineation simulation and prediction in deltaic landscape. Ecol Ind 108:105757

Defne Z, Aretxabaleta AL, Ganju NK, Kalra TS, Jones DK, Smith KE (2020) A geospatially resolved wetland vulnerability index: synthesis of physical drivers. PLoS ONE 15(1):e0228504

Dong X, Kattel G, Jeppesen E (2020) Subfossil cladocerans as quantitative indicators of past ecological conditions in Yangtze River Basin lakes, China. Sci Total Environ 728:138794

El-Magd SAA, Eldosouky AM (2021) An improved approach for predicting the groundwater potentiality in the low desert lands; El-Marashda area, Northwest Qena City, Egypt. J Afr Earth Sci 179:104200

Everard M, Kangabam R, Tiwari MK, McInnes R, Kumar R, Talukdar GH, Das L (2019) Ecosystem service assessment of selected wetlands of Kolkata and the Indian Gangetic Delta: multi-beneficial systems under differentiated management stress. Wetlands Ecol Manage 27(2):405–426

Fickas KC, Cohen WB, Yang Z (2016) Landsat-based monitoring of annual wetland change in the Willamette Valley of Oregon, USA from 1972 to 2012. Wetlands Ecol Manage 24(1):73–92

Finlayson M, Davidson N (2018) Global wetland outlook: Technical note on status and trends. Secretariat of the Ramsar Convention

Finlayson C (2006) Vulnerability assessment of important habitats for migratory species: examples from eastern Asia and northern Australia. In: Migratory species and climate change: impacts of a changing environment on Wild animals. UNEP/Earthprint, pp 18–25

Ghosh B (2021) Spatial mapping of groundwater potential using data-driven evidential belief function, knowledge-based analytic hierarchy process and an ensemble approach. Environ Earth Sci 80(18):1–19

Gómez-Baggethun E, Tudor M, Doroftei M, Covaliov S, Năstase A, Onără DF, Cioacă E (2019) Changes in ecosystem services from wetland loss and restoration: an ecosystem assessment of the Danube Delta (1960–2010). Ecosyst Serv 39:100965

Granger JE, Mahdianpari M, Puestow T, Warren S, Mohammadimanesh F, Salehi B, Brisco B (2021) Object-based random forest wetland mapping in Conne River, Newfoundland, Canada. J Appl Remote Sens 15(3):038506

Griffis-Kyle KL, Mougey K, Vanlandeghem M, Swain S, Drake JC (2018) Comparison of climate vulnerability among desert herpetofauna. Biol Cons 225:164–175

Grzybowski M, Glińska-Lewczuk K (2019) Principal threats to the conservation of freshwater habitats in the continental biogeographical region of Central Europe. Biodivers Conserv 28(14):4065–4097

Guo B, Cheng Z, Feng T (2020) Research on the influence of dual governance on the vulnerability of technology innovation network in major engineering projects. Int J Electr Eng Educ 0020720920940606

Han J, Park S, Kim S, Son S, Lee S, Kim J (2019) Performance of logistic regression and support vector machines for seismic vulnerability assessment and mapping: a case study of the 12 September 2016 ML5. 8 Gyeongju Earthquake, South Korea. Sustainability 11(24):7038

Henke J (2020) Regressing background characteristics on the self-assessed and the objective measure of economic vulnerability. In: Revisiting economic vulnerability in old age. Springer, Cham, pp 217–220

Hirst FC (1916) Report on the Nadia rivers, Calcutta, pp 1–29

Islam ARMT, Talukdar S, Mahato S, Ziaul S, Eibek KU, Akhter S, Linh NTT (2021) Machine learning algorithm-based risk assessment of riparian wetlands in Padma River Basin of Northwest Bangladesh. Environ Sci Pollut Res 1–22

Jacinth Jennifer J, Saravanan S (2021) Artificial neural network and sensitivity analysis in the landslide susceptibility mapping of Idukki district, India. Geocarto Int 1–23

Jain R, Xu W (2021) HDSI: High dimensional selection with interactions algorithm on feature selection and testing. PLoS ONE 16(2):e0246159

James G, Witten D, Hastie T, Tibshirani R (2013) An introduction to statistical learning, vol 112. Springer, New York, p. 18

Jun MJ (2021) A comparison of a gradient boosting decision tree, random forests, and artificial neural networks to model urban land use changes: the case of the Seoul metropolitan area. Int J Geograph Inform Sci 1–19

Khatun R, Talukdar S, Pal S, Saha TK, Mahato S, Debanshi S, Mandal I (2021) Integrating remote sensing with swarm intelligence and artificial intelligence for modelling wetland habitat vulnerability in pursuance of damming. Ecol Inform 101349

Kundu S, Pal S, Talukdar S, Mandal I (2021) Impact of wetland fragmentation due to damming on the linkages between water richness and ecosystem services. Environ Sci Pollut Res 1–20

Li Y, Liu B, Yu Y, Li H, Sun J, Cui J (2021) 3E-LDA: three enhancements to linear discriminant analysis. ACM Trans Knowl Discov Data (TKDD) 15(4):1–20

Li L, Nahayo L, Habiyaremye G, Christophe M (2020) Applicability and performance of statistical index, certain factor and frequency ratio models in mapping landslides susceptibility in Rwanda. Geocarto Int 1–19

Lin ML, Tsai CW, Chen CK (2021) Daily maximum temperature forecasting in changing climate using a hybrid of multi-dimensional complementary ensemble empirical mode decomposition and radial basis function neural network. J Hydrol Reg Stud 38:100923

Ling C, Wei X, Shen Y, Zhang H (2021) Development and validation of multiple machine learning algorithms for the classification of G-protein-coupled receptors using molecular evolution model-based feature extraction strategy. Amino Acids 53(11):1705–1714

Luo X, Wang F, Bhandari S, Wang N, Qiu X (2021a) Effectiveness evaluation and influencing factor analysis of pavement seal coat treatments using random forests. Constr Build Mater 282:122688

Luo X, Wen X, Zhou M, Abusorrah A, Huang L (2021b) Decision-tree-initialized dendritic neuron model for fast and accurate data classification. IEEE Trans Neural Netw Learn Syst

Majumdar D (1978) District Gazetteer, Nadia, Govt, of West Bengal, p. 7

McFeeters SK (1996) The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int J Remote Sens 17(7):1425–1432

Meng L, Dong J (2019) LUCC and ecosystem service value assessment for wetlands: a case study in Nansi Lake, China. Water 11(8):1597

Mohana RM, Reddy CKK, Anisha PR, Murthy BR (2021) Random forest algorithms for the classification of tree-based ensemble. Mater Today Proc

Mosaffaie J, Jam AS, Tabatabaei MR, Kousari MR (2021) Trend assessment of the watershed health based on DPSIR framework. Land Use Policy 100:104911

Myers MR, Cayan DR, Iacobellis SF, Melack JM, Beighley RE, Barnard PL, Page HM (2019) Santa Barbara area coastal ecosystem vulnerability assessment. California Sea Grant

Neto JG, Ozorio LV, de Abreu TCC, dos Santos BF, Pradelle F (2021) Modeling of biogas production from food, fruits and vegetables wastes using artificial neural network (ANN). Fuel 285:119081

Nie F, Wang Z, Wang R, Wang Z, Li X (2020) Adaptive local linear discriminant analysis. ACM Trans Knowl Discov Data (TKDD) 14(1):1–19

Pal S, Debanshi S (2021a) Developing wetland landscape insecurity and hydrological security models and measuring their spatial linkages. Eco Inform 66:101461

Pal S, Debanshi S (2021b) Machine learning models for wetland habitat vulnerability in mature Ganges delta. Environ Sci Pollut Res 28(15):19121–19146

Pal S, Paul S (2020) Assessing wetland habitat vulnerability in moribund Ganges delta using bivariate models and machine learning algorithms. Ecol Indicators 119:106866. https://doi.org/10.1016/j.ecolind.2020.106866

Pal S, Paul S (2021a) Stability consistency and trend mapping of seasonally inundated wetlands in Moribund deltaic part of India. Environ Dev Sustain. https://doi.org/10.1007/s10668-020-01193-z

Pal S, Paul S (2021b) Linking hydrological security and landscape insecurity in the moribund deltaic wetland of India using tree-based hybrid ensemble method in python. Ecol Inform 65:101422. https://doi.org/10.1016/j.ecoinf.2021.101422

Pal S, Saha TK (2018) Identifying dam-induced wetland changes using an inundation frequency approach: the case of the Atreyee River basin of Indo-Bangladesh. Ecohydrol Hydrobiol 18(1):66–81

Pal S, Talukdar S (2018) Application of frequency ratio and logistic regression models for assessing physical wetland vulnerability in Punarbhaba river basin of Indo-Bangladesh. Hum Ecol Risk Assess Int J 24(5):1291–1311

Pal S, Talukdar S (2019) Impact of missing flow on active inundation areas and transformation of parafluvial wetlands in Punarbhaba-Tangon river basin of Indo-Bangladesh. Geocarto Int 34(10):1055–1074

Pal S, Talukdar S, Ghosh R (2020) Damming effect on habitat quality of riparian corridor. Ecol Ind 114:106300

Pal S, Sarda R (2021) Modeling riparian flood plain wetland water richness in pursuance of damming and linking it with a methane emission rate. Geocarto Int 1–29

Pal S (2011) Wetland of Bengal basin: virtue and vulnerability, lower gangetic plain of India. Lap Lambert Academic Publishing, Saarbrücken, pp 63–87. ISBN 978-3-8473-2636-6

Paul S, Pal S (2020a) Exploring wetland transformations in moribund deltaic parts of India. Geocarto Int 35(16):1873–1894. https://doi.org/10.1080/10106049.2019.1581270

Paul S, Pal S (2020b) Predicting wetland area and water depth of Ganges moribund deltaic parts of India. Remote Sens Appl Soc Environ 19:100338. https://doi.org/10.1016/j.rsase.2020.100338

Pham BT, Prakash I, Singh SK, Shirzadi A, Shahabi H, Bui DT (2019) Landslide susceptibility modeling using reduced error pruning trees and different ensemble techniques: hybrid machine learning approaches. CATENA 175:203–218

Prasher K (2018) The state of India's disappearing 919 wetlands. The Weather Channel India. https://weather.com/en-IN/india/news/news/2018-11-08-the-case-of-indias-disappearingwetlands

Qolipour F, Ghasemzadeh M, Mohammad-Karimi N (2021) The predictability of tree-based machine learning algorithms in the big data context. Int J Eng 34(1):82–89

Rabbani M, Wang Y, Khoshkangini R, Jelodar H, Zhao R, Ahmadi SBB, Ayobi S (2021) A review on machine learning approaches for network malicious behavior detection in emerging technologies. Entropy 23(5):529

Saha TK, Pal S (2019a) Emerging conflict between agriculture extension and physical existence of wetland in post-dam period in Atreyee River basin of Indo-Bangladesh. Environ Dev Sustain 21(3):1485–1505

Saha TK, Pal S (2019b) Exploring physical wetland vulnerability of Atreyee river basin in India and Bangladesh using logistic regression and fuzzy logic approaches. Ecol Ind 98:251–265

Sampson SL (2021) Response of wetlands to impacts from agricultural land-use practices: implications for conservation, management, and rehabilitation in the Nuwejaars Catchment, Western Cape

Sattari MT, Feizi H, Colak MS, Ozturk A, Ozturk F, Apaydin H (2021) Surface water quality classification using data mining approaches: irrigation along the Aladag River. Irrigation and Drainage

Scarpiniti M, Colasante F, Di Tanna S, Ciancia M, Lee YC, Uncini A (2021) Deep belief network based audio classification for construction sites monitoring. Expert Syst Appl 177:114839

Shahabi H, Shirzadi A, Ronoud S, Asadi S, Pham BT, Mansouripour F, Bui DT (2021) Flash flood susceptibility mapping using a novel deep learning model based on deep belief network, back propagation and genetic algorithm. Geosci Front 12(3):101100

Shaziayani WN, Ul-Saufie AZ, Ahmat H, Al-Jumeily D (2021) Coupling of quantile regression into boosted regression trees (BRT) technique in forecasting emission model of PM10 concentration. Air Qual Atmosp Health 1–17

Song H, Liu A, Li G, Liu X (2021) Bayesian bootstrap aggregation for tourism demand forecasting. Int J Tourism Res

Talukdar S, Pal S (2019) Effects of damming on the hydrological regime of Punarbhaba river basin wetlands. Ecol Eng 135:61–74

Talukdar S, Eibek KU, Akhter S, Ziaul S, Islam ARMT, 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 Indicators 126:107612

Taser PY (2021) Application of bagging and boosting approaches using decision tree-based algorithms in diabetes risk prediction. In: Multidisciplinary digital publishing institute proceedings, vol 74, No. 1, p. 6

Trevisan DP, da Conceição Bispo P, Almeida D, Imani M, Balzter H, Moschini LE (2020) Environmental vulnerability index: an evaluation of the water and the vegetation quality in a Brazilian Savanna and Seasonal Forest biome. Ecol Ind 112:106163

Walker KW (2021) Exploring adaptive boosting (AdaBoost) as a platform for the predictive modeling of tangible collection usage. J Acad Librariansh 47(6):102450

Wen L, Hughes M (2020) Coastal wetland mapping using ensemble learning algorithms: a comparative study of bagging, boosting and stacking techniques. Remote Sensing 12(10):1683

Xia H, Ge S, Zhang X, Kim G, Lei Y, Liu Y (2021) Spatiotemporal dynamics of green infrastructure in an agricultural peri-urban area: a case study of Baisha District in Zhengzhou. China Land 10(8):801

Xiao H, Shahab A, Li J, Xi B, Sun X, He H, Yu G (2019) Distribution, ecological risk assessment and source identification of heavy metals in surface sediments of Huixian karst wetland, China. Ecotoxicol Environ Saf 185:109700

Yang Y, Chung H, Kim JS (2021) Local or neighborhood? Examining the relationship between traffic accidents and land use using a gradient boosting machine learning method: the case of suzhou industrial park, china. J Adv Transp

Zhang T, He W, Zheng H, Cui Y, Song H, Fu S (2021) Satellite-based ground PM2. 5 estimation using a gradient boosting decision tree. Chemosphere 268:128801

Zharmagambetov A, Carreira-Perpinán MA (2021) A simple, effective way to improve neural net classification: ensembling unit activations with a sparse oblique decision tree. In: 2021 IEEE international conference on image processing (ICIP). IEEE, pp 369–373

Zharmagambetov A, Hada SS, Gabidolla M, Carreira-Perpinán MA (2021) Non-greedy algorithms for decision tree optimization: an experimental comparison. In: 2021 international joint conference on neural networks (IJCNN). IEEE, pp 1–8