Finding Out Suitable Index for Wetland Mapping in Barind Plain of India and Predicting Dynamics of Its Area and Depth

Pankaj Singha
University of Gour Banga
Swades Pal (swadespal2017@gmail.com)
University of Gour Banga https://orcid.org/0000-0003-4561-2783

Research Article

Keywords: Satellite images, wetland mapping, future prediction, ANN-CA, Adaptive exponential smoothing, Machine learning, Tangon river basin

Posted Date: November 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1003288/v1

Finding out suitable index for wetland mapping in Barind plain of India and predicting dynamics of its area and depth

Pankaj Singha[1] Swades Pal[2*]

[1]Senior Research Fellow, Department of Geography, University of Gour Banga, Malda, India. Email- pankajsingha2014@gmail.com (P.S.)

[2*]Professor, Department of Geography, University of Gour Banga, Malda, India. Email- swadespal2017@gmail.com (S.P.)

*** Corresponding author**

Abstract

Remote Sensing and GIS play an important role in mapping and monitoring natural resources and their management. The present study attempts to delineate wetland in the lower Tangon river basin in the Barind flood plain region using suitable water body extraction indices. The main objectives of this present study are mapping and monitoring the flood plains wetlands along with the future status of wetland areas of 2028 and 2038 using the advanced Artificial Neural Network-based Cellular Automata (ANN-CA) model. Apart from wetland area prediction, wetland depth simulation and prediction are also carried out using statistical (Adaptive Exponential Smoothing) as well as advanced machine learning algorithms such as Bagging, Random subspace, Random forest, Support vector machine, etc. for the year 2028. The result shows a remarkable change in the overall wetland area in the upcoming two decades. The small wetland patches away from the master stream are expected to dry out during the forecast period, where the major wetland patches nearer to the master stream with greater depth are rather sustainable but their depth of water may be reduced in the next decades. All models show satisfactory performance for wetland depth mapping, but the Random subspace model was identified as the best-suited depth predicting method and machine learning models explored better results that adaptive exponential smoothing. This recent study will definitely be very helpful for the policymakers for managing wetland landscape as well as the natural environment.

Keywords: Satellite images, wetland mapping, future prediction, ANN-CA, Adaptive exponential smoothing, Machine learning, Tangon river basin

1. Introduction:

The term "wetland" generally describes an area saturated or inundated by water at the dried month of the year (Tiner, 2016; Kaplan and Avdan, 2018). Wetland, particularly, terrestrial wetlands, holds a unique natural ecosystem that provide immense ecosystem services (ES) in comparison with other ecosystem such as forest, grassland, lakes, rivers, and coastal wetlands (Costanza et al., 1997, 2014; Ramsar Convention on Wetlands, 2018; Zhou et al., 2020; Song et al., 2021). Wetlands cover only 6% of the global surface but provide 40% of the global ecosystem services (Khatun et al., 2021). Wetlands provide various provisioning ES namely wood, food, freshwater, fodder, genetic materials, medicinal materials; regulating services such as global and local climate regulation, air purification, water flow regulation, waste treatment, nutrient regulation, improving soil fertility, pollination, soil erosion prevention, flood control; and also provide cultural services like tourism or recreation, natural diversity, cultural diversity and so on for human well-being (Keddy et al., 2000; M.A, 2005; Costanza et al., 2014; Lin et al., 2019; Pal and Sarda, 2020). Wetlands offer favourable habitats for a wide variety of plants and animals, such as mangroves, crabs, fishes, migratory birds, mammals, reptiles, amphibians, and invertebrate species (Ramsar Convention Secretariat, 2016; Wang et al., 2020; Balwan and Kour, 2021). Wetlands are also important for storing carbon, which helps to mitigate the warming effect of anthropogenic greenhouse gases of the atmosphere (Tiner et al., 2015; Nag et al., 2017).

In recent times, this valuable ecosystem is under growing threat due to an increasing human intervention and growing water demand for agriculture, livestock, industrial purposes, etc. (IUCN, 1996; Acreman et al., 2007; Xu et al., 2020; Hu et al., 2020; Li et al., 2021). The International Union for Conservation of Nature (IUCN) (1996) reported that 50% of the global wetlands have already lose their existence due to human unscientific activities, where agricultural encroachment is the main reason behind this situation (Johnston and McIntyre, 2019; Xu et al., 2019; Li et al., 2021). The tropical and subtropical countries have faced this situation since the 1950s (Slagter et al., 2020; Ghosh, 2021). In Asia, approximately 5000 km^2 of wetland area is vanished in every year due to human activities such as agriculture, expansion of built-up area, damming etc. (Singh and Sinha, 2019; Chandra and Kumar, 2020) and remaining 50% of the wetlands are under threat and continuously degraded due to those factors (Acreman et al., 2007; Singh and Sinha, 2019). According to Ramsar Convention on wetlands (2018), 35% of the global wetlands lost their existence between 1970-2015. The report also figured out climate change, increase of population, rapid urbanization are the main reasons behind wetland loss globally (Ramsar, 2018). The Effects of climate change, such as

an increase in temperature, precipitation patterns, are expected to have a significant impact on the occurrence, structure, and function of wetlands (Sarkar et al., 2020; Lee et al., 2020). Wetland degradation is a very common phenomenon in developing countries like India, where the population pressure is very high (Das and Basu, 2020; Debanshi et al., 2020; Saha et al., 2021). The deltaic or floodplain wetlands of India have also experienced immense anthropogenic pressure, mainly due to rapid urbanization and agricultural expansion (Saha and Pal, 2019; Paul and Pal, 2020). In India, many wetlands have witnessed a rapid detraction, including deterioration in water quality, squeezing wetland area, declined wetland depth, and loss of habitat for many species (Talukdar et al., 2017; Pal and Sarda, 2020). The present study area is the lower Tangon river basin in the Barind tract of India, which consists of numerous ox-bow lakes and floodplain marshes (Das and Pal, 2017; Chakraborty et al., 2018). These wetlands provide valuable ES, and a large section of economically marginal people solely depends on wetland for their livelihoods. Squeezing and degradation of wetlands insist on narrowing wetland dependent livelihood opportunities (Talukdar et al., 2020). This region is characterized by huge population density and infrastructural development such as road networks, an extension of built-up area, etc. (Pal and Singha, 2021). Apart from this, the Boda dam was built in 1989 in Panchagarh district in Bangladesh for irrigation purposes over the upper part of the river, which significantly reduced the water discharge level by 63% downstream of the Tangon river (Pal et al., 2019; Pal and Singha, 2021). This reduction of discharge is caused for hydrological changes in riparian wetlands (Chakraborty et al., 2018). Due to such modification, wetlands have been experiencing problems, like reducing in flow, decreasing water depth, narrowing hydro-period, uncertain water level fluctuation, etc., however these are very crucial factors for sustaining the habitat status of the river as well as riparian wetlands and concerned ecosystem (Saha and Pal, 2019; Talukdar et al., 2020; Debanshi and Pal, 2020; Islam et al., 2021). But for addressing these issues properly, it is highly necessary to demarcate the wetland boundary of a region.

Very few global and national scale wetlands inventories are available with a coarse resolution. In a finer resolution, it is almost absent, especially in developing countries (Ramsar Convention Secretariat, 2016). Water body extraction indices using remote sensing data are very helpful for wetland mapping (Debanshi et al., 2020; Talukdar and Pal, 2020). Previous studies like Hird et al. (2017), Rezaee et al. (2018), McCarthy et al. (2018), Li et al. (2018) etc. successfully applied satellite imageries for water body extraction using different spectral indices for wetland mapping purposes (Debanshi et al., 2020). In recent times, a good

number of water body extraction indices like Normalized Difference Water Index (NDWI) (McFeeters, 1996), Modified Normalized Difference Water Index (MNDWI) (Xu et al., 2006), Remodified Normalized Difference Water Index (RmNDWI) (Debanshi et al., 2020), Water Ration Index (WRI) (Shen and Li, 2010), Automated Water Extraction Index (AWEI), (Feyisa et al., 2014), etc. successfully applied for wetland mapping and monitoring. The above mentioned water body extraction methods are not equally applicable across the globe, they vary from one region to another (Debanshi and Pal, 2020). It is mentioned that NDWI is more suited for open water body delineation. Apart from an open water body, MNDWI is often suitable for identifying wetlands having wet soil, as well as clearly demonstrating ox-bow lakes, river scours and left channels (Debanshi and Pal, 2020). Recently, RmNDWI indices were successfully applied for wetland mapping (Debanshi and Pal., 2020). Debanshi and Pal (2020) stated that RmNDWI water indices are more suitable for floodplain and deltaic landscapes. Considering this space's respective uncertainty of the applicability of the indices, the present study has tried to identify the best suitable indies for wetland mapping using a multi-indices approach.

Mapping and monitoring of historical wetlands trends and its dynamics especially in terms of its areal extent and depth of water, play a vital role for implementation of any wetland management plans (Hird et al., 2017; Kaplan and Avdan, 2018; Mleczko et al., 2021). Wetland depth is very essential for sustaining different species (Lantz et al., 2011; Hamza and Selmi, 2018). Depth is a hydro-ecological aspect whose variation is linked to the species richness, diversity, comfortability, particularly to the fish species (Hamza and Selmi, 2018; Ouma, 2020). Wetland trend analysis (depth and area) is very essential for sustainable management planning. Lack of such information to the administration is also one of the primary causes of rapid wetland conversion around the world, particularly in floodplain regions (Talukdar and Pal, 2020). Therefore, predicting wetland area and depth of water is another major focus of this work. Regarding the originality issue, it is to be mentioned that area prediction at pixel scale using some advance methods like deep Convolutional Neural Network (CNN), Google Earth Engine (GEE), machine learning algorithms such as RF, Bag, ANN, SVM, Dagging, Random subspace etc. was performed by various researchers in worldwide (Hird et al., 2017; Liu et al., 2018; Martins et al., 2020; Du et al., 2020; Mahdianpari et al., 2020) But pixel scale depth prediction using advance machine learning technique from image data was not done so far. Not only pixel scale, monitoring depth data of wetland at least one gauge station per wetland is absent. So, there is no option rather than to use image data for such work. The present work has tried to apply advanced ML models

for pixel scale water depth prediction from time series image data in order to obtain precise depth data for future wetland protection and restoration. Since these ML models could successfully predicted events in different disciplines (such as landslide, floods etc.), the present study has also attempted to apply these successful models to this particular perspective. Successful application of robust machine learning methods at pixel scale can produce a reliable water depth state of wetland and it is really helpful for long-term ecosystem sustainability as a whole and fish habitability in particular.

2. Study area

The Tangon river, a downstream tributary of the Mahananda river, flows through Bangladesh and India (125 km in Bangladesh and 142 km in India) and built up a 2388.88 km^2 trans-boundary basin over the Indo-Bangladesh Barind region (1234.68 km^2 in Bangladesh and 1154.20 km^2 in India) extending from 26º 19' 56" N/88º 14' 14" E to 24º 57' 22" N/ 88º 29' 29" E. The climate in this region is dominated by subtropical monsoon types, with a high degree of seasonal variation in temperature and rainfall. The average rainfall per year of this region occurs between 1257 mm to 1508 mm (Pal et al., 2019), about 80% of which takes place in the monsoon season covering the month of June to September. It often inundated the depressed land. The topography is flat with an elevation range between 8 to 83 meters. Geographically, this region belongs to the North-Western part of Bangladesh and the Eastern part of India. The population density of the Bangladesh part is 728 persons/km^2 and the Indian part is 665 persons/km^2 and about 70% of people are involved in primary activities like farming, hunting, fishing, etc. (Pal et al., 2020). The land use consists of a water body, built-up area, agricultural land, and vegetation, where more than 60% of the area is covered by agricultural land. The downstream segment of the river basin experienced a reduction of river flow due to the construction of the Boda dam (Pal et al., 2019; Pal and Singha, 2021). Some researchers have already identified the effect of damming on wetland hydrological characteristics in this river basin (Pal et al., 2019; Pal et al., 2020). The riparian villages like Aiho, Chhatiangachi, Laxmipur, Dubapara, Shivganj, Bidhanghar, Soladanga, Jatradanga, Oltara, Nakail, Madhaipur, Bulbulchandi, etc. are largely dependent on fishing activities directly and indirectly. Hydrological changes have also forced a large number of people to change their occupation. The detailed study location is mentioned in the figure. 1.

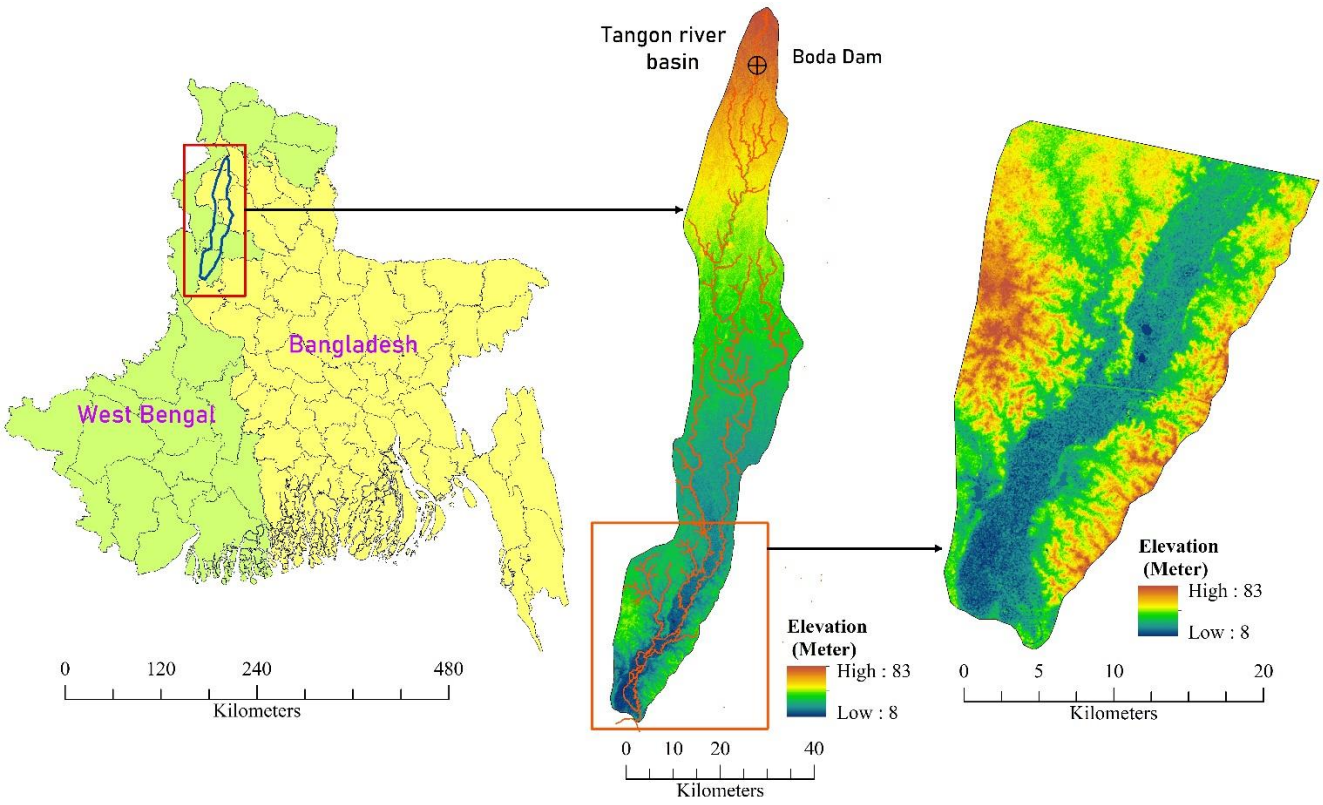

Figure. 1 Study location showing the entire river basin and lower part of the river basin with its elevation status.

3. Materials and methods

3.1 Materials used

Landsat 4-5 (TM) and Landsat (OLI) satellite data (cloud-free) with 30 meters spatial resolution (path/row: 139/43) have been taken from the USGS (United States Geological Survey) Earth Explorer website (https://earthexplorer.usgs.gov) for pre-monsoon seasons. Field-specific 33 sites of water depth data were used for wetland depth calibration. A digital elevation model (DEM) is used for base map preparation. In addition, Google Earth maps were used for validation purposes.

3.2 Methods

3.2.1 Wetland mapping using water indices and validation

Making a distinct difference between water and non-water pixels is very challenging using a common threshold due to spectral nearness of some land use land cover components (Ngoc et al., 2019). In recent times, numerous remote sensing spectral indices are available for delineation of water bodies as well as wetland mapping (Paul and Pal, 2020; Saha et al.,

2021). Spectral indices like NDWI (Mcfeeters, 1996), MNDWI (Xu, 2006), RmNDWI (Debanshi and Pal, 2020), AWEI (Feyisa et al., 2014), WRI (Shen and Li, 2010), and so on are frequently used for waterbody delineation. Hence, it will be justified to compare the results derived from the different indices (Das and Pal, 2017) and find out the most suitable one. There is some specification regarding the applicability of the water body extraction indices. Out of that three frequently used water body extraction NDWI (Eq. 1), MNDWI (Eq.2), and RmNDWI (Eq.3) techniques were applied for preparing wetlands maps for both pre- and post-monsoon periods.

$$NDWI = \frac{B_{Green} - B_{NIR}}{B_{Green} + B_{NIR}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

$$MNDWI = \frac{B_{Green} - B_{MIR}}{B_{Green} + B_{MIR}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$RmNDWI = \frac{B_{Red} - B_{MIR}}{B_{Ren} + B_{MIR}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

where, B_{Green}, B_{NIR}, B_{MIR}, B_{Red} refers to Green, Near-infrared, Middle-infrared, and Red bands of Landsat imagery. Hypothetically, the indices value ranges from 0 to 1, where closed to 0 indicates saturated water body having shallower depth, and closed to 1 indicates waterbody with high water availability. The indices values below 0 are not treated as water bodies (Das, 2017; Kaplan et al., 2019). Kappa statistics (K) and Area under the curve (AUC) of Receiver operating characteristics (ROC) and overall accuracy were successfully applied for validation of the wetlands map and selecting suitable technique.

3.2.2 Methods for simulating and predicting wetland area

Artificial neural network-based cellular automata (ANN-CA) methods were applied to simulate the and predict the wetland area. The ANN-CA algorithm explores the internal pattern of the input data set and provides output based on the intra pixel value (Debanshi and Pal, 2020; Saha et al., 2021). The flood plains wetlands are very dynamic in nature, where vegetation cover, agricultural land, and built-up area play a major role in determining the areal extent of wetlands in this region (Pal and Talukdar, 2018; Saha et al., 2021). In general, some other controlling factors, like rainfall occurrences, geology, types of soil, etc. may control wetland dynamics, but in this recent study unit, these were not found to be important controlling factors due their spatial similarity. Thus, conditioning parameters such as

agriculture status, vegetation status, and built-up status layer were used to predict the future wetlands area. Thus, presently, Normalized Difference Vegetation Index (NDVI) (Townshend and Justice, 1996) (Eq.4) for vegetation and agriculture land and Normalized difference built-up index (NDBI) (Zha et al., 2003) (Eq. 5) for built-up areas were successfully applied for the simulation and prediction of wetlands area. The details of conditioning parameters are given in the supplementary section (Figure. S2). For doing this, agriculture and built-up area maps were also simulated for the same years of wetland area prediction. While predicting wetland of a particular year, respective simulated agriculture and built-up maps were used. Controlling factors were predicted and the predicted factors maps were used for predicting wetland areas of 2028 and 2038. The training and testing datasets have been divided into 70 and 30 ratios.

$$NDVI = \frac{B_{NIR} - B_{\mathrm{Red}}}{B_{NIR} - B_{\mathrm{Red}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

$$NDBI = \frac{B_{MIR} - B_{NIR}}{B_{MIR} + B_{NIR}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (5)$$

where B_{NIR}= Near Infra-red band, B_{Red}= Red band, B_{MIR}=Middle Infra-Red band of Landsat imageries. The ANN simulation model was successfully developed in the QGIS software environment using the MOLUSCE plug-in tool. For validating the predicted wetland area, simulated wetland area maps of 1998, 2008 and 2018 were compared with existing wetland maps of the respective years. It was assumed that if the simulated wetland area maps were converged with actual wetland maps, predicted wetland maps of 2028 and 2038 could be treat as valid.

3.2.3 Methods for simulating and predicting wetland water depth

3.2.3.1 Adaptive Exponential Smoothing (AES)

AES techniques were used to simulate wetland depth for the years 1998, 2008, and 2018, and to predict the depth of the wetland for the year 2028. AES algorithms is one of the most commonly utilized forecasting approaches, which is widely used in different fields (de Oliveira et al., 2018; Yang et al., 2018; Smyl, 2020). The performance of the model was improved by applying this technique and forecasting reliability by smoothing the variables using simple computation (Helmstetter and Werner, 2014; Mi et al., 2018). Calibration was performed using NDWI, MNDWI and RmNDWI images and when calibrated images again

match with field based actual data, it was found NDWI is highly correlated with water depth. Therefore, NDWI images were taken for developing depth data and related simulation and prediction. A wetland depth map calibrating NDWI index of the decadal interval 1988 to 2018 has been developed using AES (Eq.6). techniques.

$$F_{t+1} = \acute{\alpha} X_t + (1 - \acute{\alpha}) Ft \quad \ldots\ldots\ldots\ldots (6)$$

where, F_t and F_{t+1} are treated as smoothing values of the water index score at t and $t+1$ time respectively, X_t *is* the actual water index score, and ά is the smoothing coefficient of X_t and Ft. The coefficient values range from 0 to 1.

The above mention approach is applied as a substitute to the usual AES, where the variable $\acute{\alpha}$ is updated along with the prediction. To increase the model performance, another variables β are applied in these models. The variables ά at each step is calculated using the equation 7.

$$\acute{\alpha}_{t+1} = \left| E_t / M \right| \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

where, E_t is represented as a smoothed error signal, which is calculated using the following equation 8.

$$E_t = \beta e_t + (1 - \beta) E_{(t-1)} \quad \ldots\ldots\ldots\ldots\ldots (8)$$

and M_t is represented as an absolute error signal, which is calculated using the following equation 9.

$$M_t = \beta \left| e \right| + (1 - \beta) M_{(t-1)} \quad \ldots\ldots\ldots\ldots\ldots (9)$$

Here, e_t is represented the deviation or error of the estimated value at point *t, which is* calculated by the following equation 10.

$$e_t = X_t - F_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (10)$$

3.2.3.2 Machine learning algorithms

3.2.3.2.1 Bagging (Bag)

The Bag algorithm is a very popular ML algorithm, which was widely used to develop ensemble ML models by coalescing with various algorithms (Prasad et al., 2006; Chapi et al., 2017). Breiman (1996) first introduced the bagging technique, which is a “bootstrap” (Efron

and Tibshirani, 1995) ensemble technique that generates multiple types of classifiers and also generated an aggregated classifier (Pham et al., 2018; Piao et al., 2015). The statistical bootstrapping techniques employ random sampling with replacement to obtain various samples from the training dataset. For the training datasets, a decision tree was developed based on each of the obtained subsets (Chen et al., 2018). The training dataset is continuously replaced by drawing random samples (Piao et al., 2015). The final model was developed by integrating all of the generated models (Chen et al., 2018). Numerous studies stated that bagging predicted environmental problems with very high accuracy (Chen et al., 2019; Wu et al., 2019; Shahabi et al., 2020). The bagging algorithm enhanced the model performance to predict any environmental issue by leasing the variance of the prediction errors (Prasad et al., 2006). Previous literature shows that Bag shows a higher accuracy for predicting flood (Chen et al., 2019; Arabameri et al., 2020), landslides (Truong et al., 2018), gully erosion (Dou et al., 2019) modelling. To estimate the out-of-bag (OBB) error, the training sample should be recorded for each base learner. For each base learner, the training sample should be recorded to estimate the out-of-bag (OBB) error. The OBB prediction denotes *H(X)* on the *X vector.* Only the learners not trained on X are involved in this stage, the OBB estimated using the following equation (Eq.11)

$$H(X)^{OOB} = \arg\max_{y \in Y} \sum_{t=1}^{T} \| (h_t(X) = y. \| (X \in N_t) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (11)$$

where, *X*= vector, *x*=variables, *y*= output spaces, *N*= data sample, *T*=number of base learners, *H*=learner and *II(.)* =indicator function (true=1, false=0).

The OOB error of bagging can be estimated using the following equation (Eq. 12).

$$Error^{OOB} \frac{1}{N} \sum_{ci(xj) \in N}^{T} (H(X)^{OOB} \neq y) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (12)$$

where, *x* is the variables

3.2.3.2.2 Random Forest (RF)

The RF classification algorithm is the modification of the CART (classification and regression of regression tree) model, which was proposed by Breiman (2001). The RF model is a very powerful ensemble learning method (Naghibi et al., 2017; Kim et al., 2018). RF is a novel non-parametric ML model which can handle a wide type of variables with wide applications (Sinha et al., 2019). For splitting each node, it uses a random selection method.

In RF models inbuilt cross-validation techniques are used to train a process which is called an 'out-of-bag (OBD) sample. During the bootstrapping process, the training data are replaced and some input data are omitted from the sample and it is continued with the OBD sample (Breiman, 2001). To successfully develop the RF model, n-tree (tress number) and m-tree (features in each split) are needed. In the RF model, classification trees are chosen by an individual choosing power or voting to the majority votes in the entire forest. The RF ML model is very powerful and accurate than the individual classifier due to its many advantages: (i) It can deal with a large number of datasets, (ii) it can handle thousands of input variables without eliminating a single variable, (iii) it creates unbiased estimation of the input variables, (iv) it can also determine the influence of each variable (Rodriguez-Galiano et al., 2012). In recent times, the RF ML algorithm was commonly applied in various fields of research and expressed excellent outcomes (Sinha et al., 2019; Sun et al., 2020; Ma et al., 2021). RF method is typically used for the purpose of prediction and interpretation of data and it is also applicable for modelling flood, landslide susceptibility etc. (Choubin et al., 2019; Tang et al., 2021; Sun et al., 2021). Worldwide various researchers have successfully applied RF techniques for flood susceptibility mapping, modelling human health vulnerability (Vafakhah et al., 2020; Mandal and Pal, 2020). The structure of RF generally depends on three steps and has been shown as follows (Eq.13):

$$h(x, i_k),\ k = 1, 2, ...n \quad \text{.....(13)}$$

where, i_k =conditioning variables, 1, 2,…n= input vectors x.

The error of RF algorithms can be estimated using the following equation (Eq. 14).

$$GE = P_{x,y}(mg(x, y) < 0) \quad \text{..... (14)}$$

where, x and y= conditioning variables, *mg*= margin function.

The margin function can be estimated as follows (Eq. 15).

$$mg(x, y) = a\upsilon_k I(h_k(x) = y) - \max_{j \neq i} a\upsilon_k I(h_k(x) = j) \quad \text{.....(15)}$$

3.2.3.2.3 Random subspace (RS)

The RS is a well-known random sampling-based ensemble machine learning algorithm developed by Ho (1998). The RS algorithms successfully enhance the performance of the weak classifier by merging them and are also used to help individual classifiers by improving

their classification accuracy (Pham et al., 2018). However, this model was widely used in different fields worldwide, such as environmental engineering, modelling natural hazards, and so on (Nhu et al., 2020; Mao et al., 2021). The RS algorithm is quite similar to bagging and both of them are influenced by bootstrapping and aggregation. The RS model bootstraps the feature space, while bagging generates outputs by bootstrapping the training samples (Chen et al., 2019). Four essential variables such as training set, base classifier, number of subspaces, and number of subsets were required to develop this model (Chen et al., 2021).

Giving the training sample set G of size n, set S= (S_1, S_2, …….S_n) with each training object Si (i=1,2……., n) to be a q-dimensional vector S_i= (S_{i1}, S_{i2}……S_{in}) described by q features. If selects $r<q$ features, then we have a r-dimensional random subspace of the original q-dimensional feature space. Therefore, each modified training object Si (s_{i1}, s_{i2}……s_{in}) ($i=1,.....n$) is a unit of training sample set S=(S_1,S_2……S_n). The algorithms of RS can be interpreted as follows (Eq. 16):

$$\gamma(s) = \arg\max \sum \delta_{sng}(C^d(s), y \text{ ……..(16)}$$

where $\delta_{i,j}$ is the Kronecker symbol, and y= (-1, 1) is a decision or class label of the classifier and $C^d(s)$ are the classifiers (d=1,2……. D).

3.2.3.2.4 Support vector machine (SVM)

SVM is well-known as a supervised non-parametric statistical ML algorithm. The concept of SVM is based on decision planes, which is defined as the plane of separation of different objectives or different class membership (Choubin et al., 2019). It can work with different types of variables like continuous, categorical and also linear and non-linear data sets in various class members. The main function of these techniques is hyper-plane separation and creating the training datasets. The mathematical function used to transform data is known as the kernel function (Tehrany et al., 2015). In the SVM model, to separate the original input space, an optimal hyper-plane is used for this purpose. The kernel function is used for data transformation, which is divided the entire dataset into flood and non-flood categories and it is determined by 1 and 0 respectively. Identification of proper kernel function shows the ability of the SVM model. Four kernel functions are available to develop the SVM model, such as polynomial kernel (PL), radial basis kernel (RBF), linear kernel (LN) and sigmoid (SM) kernel. Commonly in remote sensing environments, PL and RBF kernels have been

used (Sothe et al., 2020). RBF kernel is commonly used in SVM techniques for the higher performance and higher level of accuracy of the model than the other traditional methods (Zhang et al., 2019). The RBF kernel is frequently applied for solving different environmental problems such as flood, landslide, soil erosion etc. (Leong et al., 2019; Sahana et al., 2020). Therefore, the RBF kernel function was applied to simulate and predict the wetland water depth (Eq. 17, 18).

$$z = f(y) = \sum_{i=1}^{p} w_i \theta_i(y) = w\theta(y) \quad \text{.....(17)}$$

where, the model's output represents the linear P components and the non-linear model is given by $\varphi(y)$ to the converter.

$$z = f(y) = \left\{ \sum_{i=1}^{n} w_i k_i(y_i, y) - c \right\} \quad \text{.....(18)}$$

where, *K*=Kernel function, *wi* and *c* both are the parameters, *L*=Number of learning pattern, *yi*= data vector, *y*=independent vector.

3.2.4 Accuracy assessment

Accuracy assessment of the GIS modelling is the very essential part before applying it for management purposes (Rasyid et al., 2016). To validate the simulated and predicted the wetland depth, receiver operating characteristics, (ROC), kappa statistics, mean absolute deviation (MAD), mean square error (MSE), root mean square error (RMSE), mean absolute percentage error (MAPE) techniques were used. The ROC curve is a well-known validation method, offering the model performance accurately and very easily. This accuracy assessment technique has been widely applied by various researchers to solve different environmental problems (such as floods, landslides, etc.) (Singha et al., 2020; Pal and Singha, 2021). In recent times, the overall accuracy, kappa coefficient has also applied for validation of wetlands maps (Saha et al., 2021). Wetlands maps are validated using the area under the curve (AUC) of the ROC curve which is generated using the 300 randomly selected reference verification sites. The verification sites were collected from Google Earth imagery and ground truth collection from the field. Those points were further used to calculate kappa coefficient and overall accuracy of water indices as well as prediction wetlands maps, using equations 19 and 20.

$$K = \frac{N\sum_{i=1}^{r} xii - \sum_{i=1}^{r}(xi+ * xi+i)}{N^2 - \sum_{i=1}^{r}(xi+ * xi+i)} \quad \ldots\ldots\ldots\ldots\ldots\ldots (19)$$

where, N = total number of pixels; r = number of rows in the matrix; Xii = number of observations in row i and column i; $xi+$ and $x+i$ are the marginal totals for row i and column i, respectively.

$$O_{accuracy} = T_{sample} \times 100\% P_{sample} \ldots\ldots\ldots\ldots..(20)$$

where, $O_{accuracy}$ represents overall accuracy, T_{sample} is the total number of corrected samples and P_{sample} is the total number of samples.

Different error calculation methods were applied to identify the error or gap between the actual and predicted wetland depth (Equation 21-24).

$$MAD = \frac{\sum_{t=1}^{n} | R_t - P_t |}{n} \ldots\ldots\ldots\ldots\ldots\ldots..(21)$$

$$MSE = \frac{\sum_{t=1}^{n} (R_t - P_t)^2}{n} \ldots\ldots\ldots\ldots\ldots\ldots..(22)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (R_t - P_t)^2}{n}} \ldots\ldots\ldots\ldots\ldots..(23)$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \left| \frac{\sum_{t=1}^{n} R_t - P_t}{R_t} \right| \ldots\ldots\ldots\ldots..(24)$$

where, R_t and P_t is the actual and forecast value and n is the number of times the simulation iteration happens. The entire work was summarized in Figure 2.

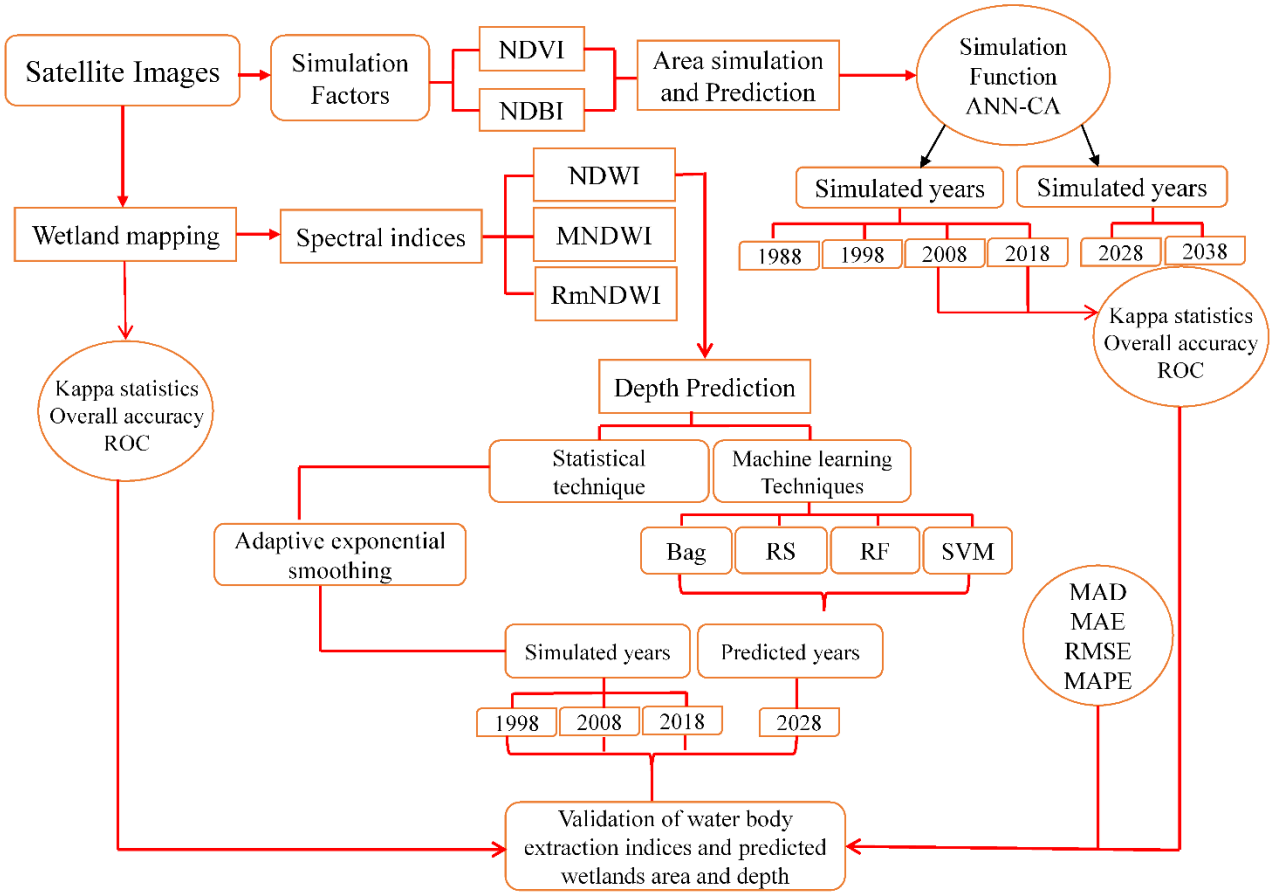

Figure 2: Flow chart of the entire work.

4. Result

4.1 Suitable water index for wetland mapping

Wetlands are primarily concentrated in the Tangon river basin along the lower portion of the stream, where the stream outflows during the monsoon season, which lasts from July to October. The time series wetland maps were developed using three water indices (NDWI, MNDWI, and RmNDWI) and such water indices provide quite different results (Figure 3 and Figure. S1 (Supplementary section)). However, when the time series trend was noticed, all the indices clarified the fact that there was a declining trend of wetland areas since 1988 to 2020 in both pre- and post-monsoon seasons. But it is highly necessary to know the most representative water indices showing wetland areas. Similar findings reported in previous flood plain or deltaic wetland environments inventoried by Saha and Pal, (2019), Pal and Talukdar (2018), and Saha et al. (2021).

Overall accuracy, kappa coefficient, and AUC of ROC statistical techniques were applied for the accuracy assessment of wetland area. The result of the accuracy assessment is represented

in Table 1. The overall accuracy, Kappa statistics, and AUC of ROC of NDWI spectral index were 92.90%, 0.89, and 0.81 for pre-monsoon and 96.06%, 0.91, and 0.89 for the pre- and post-monsoon seasons respectively. The details of accuracy assessment given in supplementary section (Table. S1). These values were higher than other applied indices like NDWI and MNDWI and therefore, RmNDWI could be treated as accepted. The RmNDWI maps were further used for wetlands area simulation and prediction.

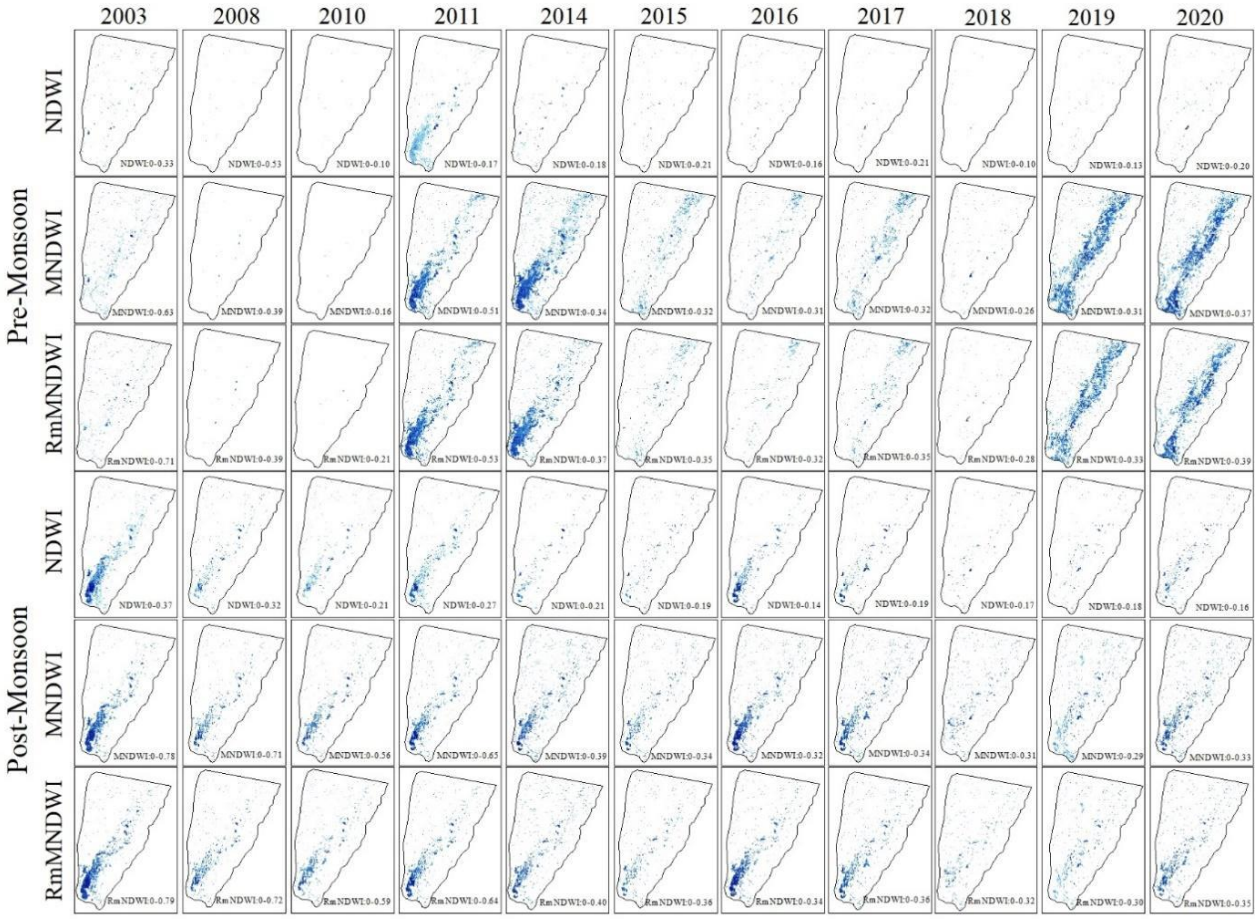

Figure. 3 Wetland mapping using NDWI, MNDWI, RmNDWI indices for both pre- and post-monsoon (1988-2020).

Table. 1 Average overall accuracy, Kappa coefficient, and AUC of ROC values of different water indices in pre-and post-monsoon seasons.

Pre-monsoon				Post-monsoon			
Indices	Overall accuracy	Kappa	ROC	Indices	Overall accuracy	Kappa	ROC
NDWI	92.09	0.84	0.73	NDWI	92.89	0.86	0.88

MNDWI	91.18	0.83	0.78	MNDWI	94.63	0.88	0.89
RmNDWI	92.90	0.89	0.81	RmNDWI	96.06	0.91	0.89

425

426 **4.2 Predicted wetland area and accuracy test**

427 The simulated wetland map derived from ANN-CA techniques, reported that wetland area
428 was rapidly declining in both in both pre and post monsoon season (Figure 4) and likely to be
429 declined in the next ten (2028) and twenty (2038) years. The wetland degradation rate is
430 higher in the pre-monsoon than the post-monsoon season. The predicted wetland area is
431 10.06 km^2 (1.97%) and 9.59 km^2 (1.87%) in 2028 and 2038 respectively, for the post-
432 monsoon season (Table. 2). The detailed areal extent of simulated and predicted wetlands is
433 mentioned in Table. 2.

434 The results of kappa statistics, overall accuracy, AUC of ROC were mentioned in Table 5.
435 The overall accuracy, kappa coefficient and AUC of ROC in 2008 were 85.3%, 0.93 and 0.91
436 respectively, while in 2018, these were 86.4%, 0.96 and 0.93 respectively. The details of
437 accuracy assessment given in supplementary section (Table. S2). These values state that the
438 simulated wetland map using ANN-CA technique reflects satisfactory performance. It could
439 be assumed that as the simulated wetland area of 2008 and 2018 was coincided with observed
440 wetland area, the predicted wetland area maps of 2028 and 2038 will also seem to be correct.

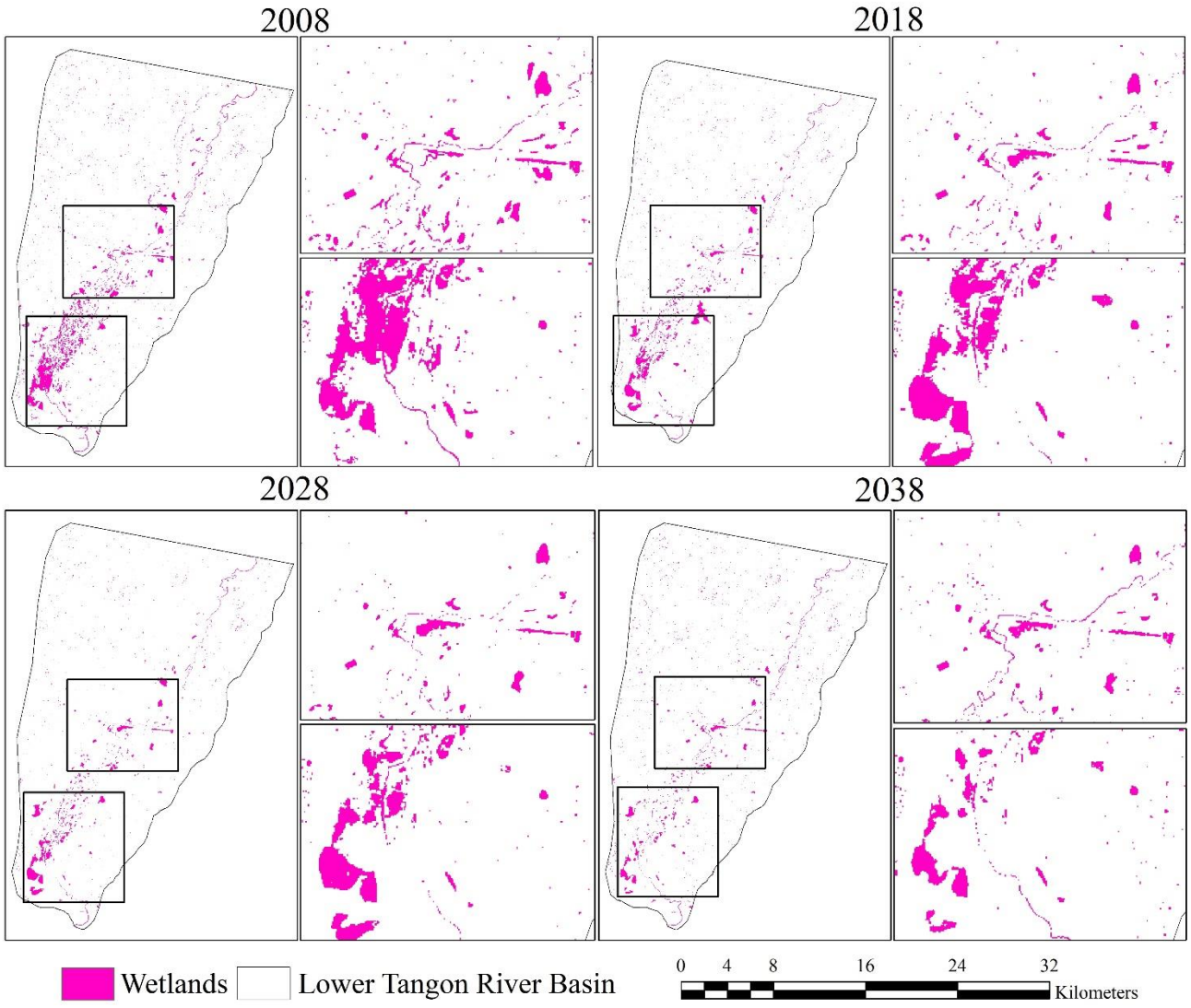

Figure. 4 Predicted wetlands areas of 2008, 2018, 2028 and 2038 using ANN-CA methods.

Table. 2 Actual and predicted areas of wetlands of post-monsoon season

Year	Water body			
	Actual area (km^2)	% of area	Predicted area ($km^{2)}$	Area (%)
2008	25.19	4.94	18.03	3.53
2018	16.21	3.18	14.44	2.83
2028	-	-	10.06	1.97
2038	-	-	9.59	1.88

4.3 Predicted wetland depth and accuracy assessment

Due to the absence of spatiotemporal datasets, time series wetland depth mapping is very difficult over a wide geographical area. The recent work calibrated NDWI map for wetland depth mapping. The calibration process was carried out using 33 field-specific wetlands depth data from different parts of the study unit. The linear regression method was used to estimate the wetland depth. The computed regression slope (β) used for depth calibration is 14.1051 meter from the NDWI image. Some researchers around the world used a similar approach for depth mapping (Donchyts et al., 2016; Khatun et al., 2021). The details of the year wise wetland depth map are given in the supplementary section (Figure. S3). The simulated and predicted wetland depth map was developed using one statistical and four machine learning algorithms. The simulated and predicted wetlands depth map based on AES and ML techniques indicate that the wetlands depth is likely to be decreased in the predicted periods. Figure 3, stated that the magnitude of depth was declined from 1988 to 2018 and likely to be declined further in 2028. Figure. 5 shows such a decreasing trend of wetland depth. The predicted wetland depth is likely to be declined from 2.49 to 1.69, 2.28 to 1.66, 2.71 to 2.18, 2.19 to 2.02, and 4.40 to 3.47 meters from 2018 to 2028 using Bag, RF, RS, SVM, and AES techniques, respectively. The simulated and predicted wetland depth maps are mentioned in Figure 5. The predicted maps showed that, many small wetland patches away from the master stream are expected to dry out during the forecast period, where the major wetland patches nearer to the master stream with greater depth are rather sustainable but their depth of water may be reduced.

To execute the best representative model from the applied five models, it is necessary to measure the accuracy of the simulated wetland water depth from observed depth. Table 3 represents the calculated error for five predictive models. Results show that the random subspace model is the best model for wetland depth prediction since the error values are minimum in this case in all the simulated years. Table 4 represents the actual and simulated depth statistics, their difference and standard deviation of five simulation techniques. From this also random subspace model can be treated as the best representative model for wetland depth prediction followed by random forest, bagging, support vector machine and adaptive exponential smoothing methods (Table. 3 and 4).

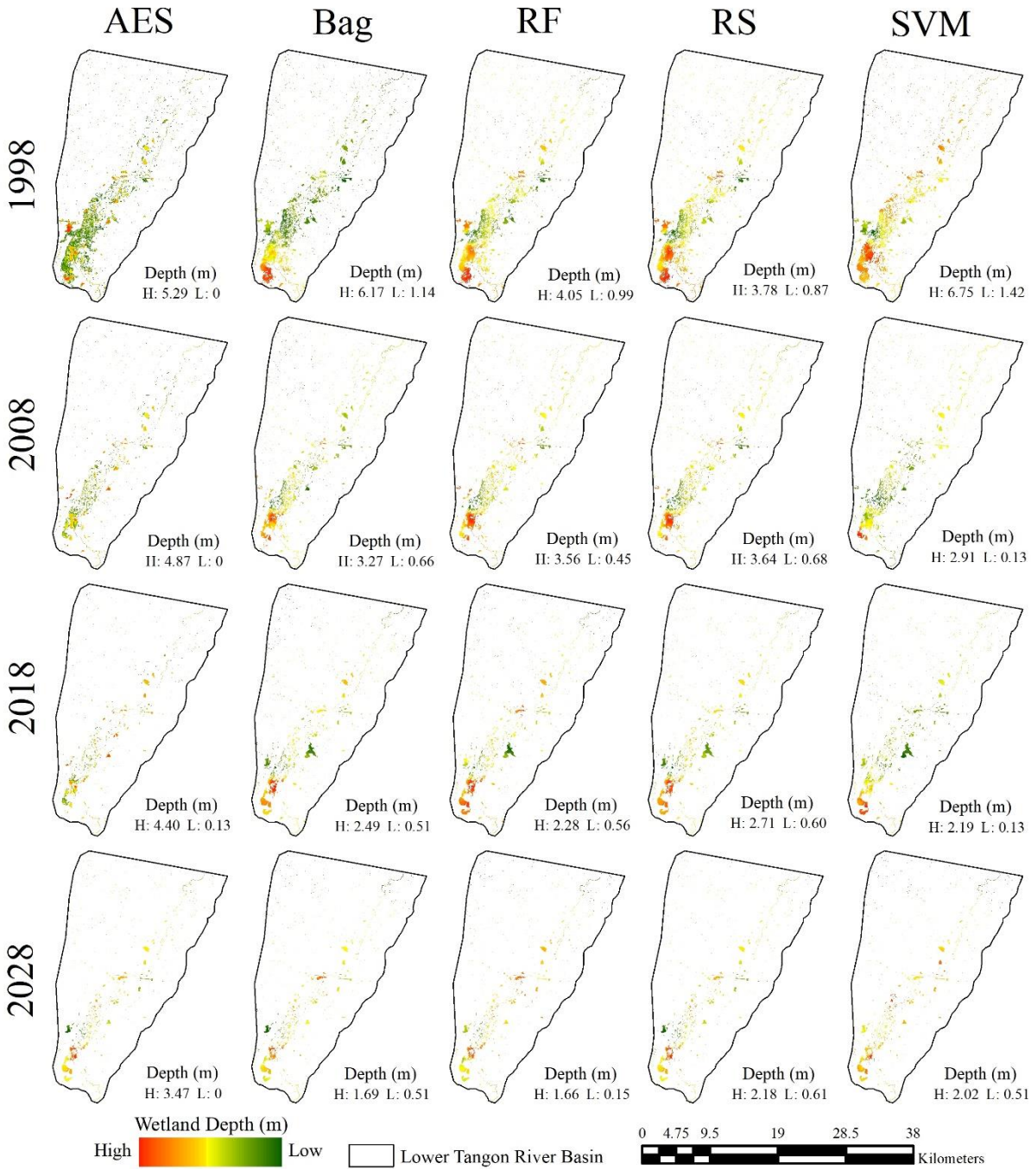

Figure. 5 Wetlands depth prediction using advanced machine learning algorithms and statistical techniques like AES, Bag, RF, RS, SVM methods of post-monsoon Season.

484 **Table. 3** Different error measurements between actual and predicted values for wetland depth simulation.

Algorithms	Predicted years											
	1998				2008				2018			
	MAD	MSE	RMSE	MAPE	MAD	MSE	RMSE	MAPE	MAD	MSE	RMSE	MAPE
AES	0.383	1.827	1.351	15.67%	0.362	1.609	1.268	12.40%	0.249	0.786	0.886	22.20%
Bag	0.095	0.144	0.380	3.01%	0.091	0.141	0.375	4.04%	0.068	0.067	0.26	4.40%
RF	0.087	0.118	0.344	2.87%	0.081	0.119	0.345	4.03%	0.060	0.055	0.235	4.53%
RS	0.086	0.115	0.339	2.89%	0.077	0.111	0.334	4.02%	0.056	0.047	0.217	5.51%
SVM	0.118	0.226	0.475	3.37%	0.134	0.305	0.552	4.63%	0.196	0.469	0.684	14.68%

485

486 **Table. 4** Actual and simulated depth statistics using different simulation techniques.

Years	Actual avg. depth	Standard deviation	Predicted avg. depth					Standard deviation					Differences				
			AES	Bag	RF	RS	SVM	AES	Bag	RF	RS	SVM	AES	Bag	RF	RS	SVM
1998	1.10	1.09	4.78	3.02	2.69	1.46	2.61	2.01	1.09	0.60	0.59	0.23	3.68	1.92	1.59	0.36	1.51
2008	1.30	1.08	4.93	1.96	2.19	0.98	2.30	1.91	0.47	0.56	0.56	0.33	3.63	0.66	0.89	0.32	1.00
2018	0.64	0.47	3.63	1.52	1.55	1.35	1.74	0.11	0.42	0.38	0.39	0.11	2.99	0.88	0.91	0.71	1.10

487

5. Discussion

Wetland demarcation is a very important task for the implementation of proper management strategies to reduce the fast rate of wetland conversion, especially in floodplain regions (Debanshi and Pal, 2020; Saha et al., 2021). The command area and depth of the floodplain or rain-faded wetlands fluctuate every year in both the pre-and post-monsoon seasons. So, monitoring wetlands area, depth and, its prediction is highly important for improving management strategies. The majority of previous studies (Tong et al., 2014; Zheng et al., 2019) reflect aerial fluctuations of wetlands from a numerical standpoint, which cannot accurately reflect the spatial situation of wetlands loss. This numerical analysis is not so important for the implementation of proper management strategies. So, wetlands area, depth, prediction at the pixel level can provide more reliable outcomes both in numerical and spatial levels and this could be very useful from management point of view.

Previous studies mostly explored the temporal changes of the wetland area but rarely addressed the issue of simulation and prediction of wetland area. From the ANN-CA declining trend of wetland area was predicted in the Barind flood plain of India and this finding is accordant with the findings of Punarbhaba river basin of Inadia and Bangladesh (Talukdar and Pal, 2020), Atryee river basin of India and Bangladesh (Saha et al., 2021), lower Ganges river basin (Sarkar et al., 2020), Danube delta (Gómez-Baggethun et al., 2019), Sanjiang plain (Yan and Zghang, 2019). Traditionally, CA was frequently used for area prediction but considering more successful prediction outcomes by ANN-CA (Debanshi et al., 2020; Saha et al., 2021), this method was adopted in this present work.

Predicted wetland area using ANN-CA methods shows significant decline in wetland area in the upcoming years of 2028 and 2038. Smaller wetland patches like the left channel, ox-bow lakes, etc. away from the mainstream are likely to be missing in predicted years (2028 and 2038). Trapping of such smaller units of wetland by anthropogenic landscape is the main reason behind death of the wetland (Kundu et al., 2021). Increasing edge area ratio, growing fragmentation of wetland area due to perforation of anthropogenic landscape insist such loss. It is also observed that the large core of wetlands is relatively safe, secure, and having a relatively higher depth compared to surrounding wetlands. The core wetlands sometimes are facing vulnerable conditions due to the drying out and connectivity loss of surrounding wetlands from the main river. The rate of wetland conversion is very high due to agriculture encroachment followed by expansion of built-up area (Pal et al., 2019). Previous studies like

Saha and Pal (2018), Talukdar and Pal (2020), Saha et al. (2021) reported the same cause behind wetland loss in the Punarbhaba, the Atreyee river basin of Indo-Bangladesh. In 1989, a dam was constructed over the upper reach of the Tangon river to divert water for irrigation purpose and it can be treated as a crucial reason for reducing the discharge in the river (Pal et al., 2019). It is also caused for the reduction of lateral inundation extent, flood frequency in this river basin (Pal and Singha, 2021). It is also treated as the main reason for reducing availability of water in the riparian wetland since flood water is also a major water source of wetlands in the flood plain regions (Talukdar and Pal, 2020; Saha et al., 2021). It may create a disturbance of the wetland ecological functions as well as wetlands ecosystem services (Khatun et al., 2021). Apart from this, lifting of water from rivers or wetlands for irrigation purposes may be responsible for the negative trend of water level. Moreover, harvesting ground water for supplying green water and associated lowering of groundwater level is caused for vertical disconnection between ground water and wetland (Chakraborty et al., 2018) In the study region, this is also a dominant fact enforces wetland transformation and loss (Das and Pal, 2017; Pal and Talukdar, 2018).

Water depth prediction is also a vital issue of long-term wetland management planning and restoration. But the major barrier behind such work is data lack. Since this issue, the recent work attempted to mine depth data from image-based water indices successfully calibrated with field data. Its a good attempt to fill the data gap. Besides, ground-based monitoring of such data at pixel level is almost impossible. Precise prediction of wetland water depth at spatial scale is also a very important but difficult task (Paul and Pal, 2020). Simple linear regression, adaptive exponential smoothing are not just sufficient for obtaining the precise result (Paul and Pal, 2020). No studies so far attempted to do this at pixel level. Therefore, the present work could be treated as a landmark in this approach since the study applied multiple ML models for predicting water depth precisely at pixel scale. Here lies the novelty of this work.

Whatever may the absolute difference in prediction result using different ML models, the overall trend shows declining water depth which is of immense ecological concern. It may cause loss of suitable ecological habitat for many species in general and livelihood centric fish in particular. So, declining depth of water is also linked with growing hardship to livelihood. Khatun and Pal (2021) worked on transformation of fish habitat state in post-dam condition in flood plain wetland and reported significant loss of ambient habitat for economically remunerative fish species.

However, along with the novelty and merits of the work, some limitations and future scope should be mentioned for further progress of research. First of all, for developing depth data from satellite images, calibration was done using s relatively smaller number of ground control points. This needs to be increased as much as possible for obtaining improved results. In this work, field data-based calibration was done for the recent four years and the computed regression slope was used for developing depth data of the previous years. Year specific field calibrated depth surface would be good. For calibration linear regression was applied here but new non-linear, region specific, hydrological status specific calibration may remove the generalization effect. Research is further required in this roadway.

6. Conclusion

This study delineated the water bodies using three water body extraction indices and identified RmNDWI as the suitable index for wetland mapping in the Barind flood plain. Wetland area and depth are likely to be declined. All the computed models are validated using well-known validating approaches like Kappa statistics, AUC of ROC and overall accuracy. All the models provide satisfactory results for wetland depth and area simulation and prediction with varying degrees. Random subspace was identified as the best suited depth predicting method and ML models explored better results that AES. Considering the success of ML models for predicting water depth, the study recommends application of ML models for water depth prediction from calibrated image data. Calibrated image data is a good way to fill in the pixel specific depth data scarcity. The study further recommends to work further for improving the depth calibration method. Moreover, application of high resolution image data could provide more improved results. Since the study clearly reported the squeezing of wetland water and shallowing water depth, it is clear that habitat shrinking and qualitative degradation of habitat triggered by depth reduction may create ecological hardship and concerned livelihood stress. Spatial maps on this could help policy makers to design prioritize planning of wetland conservation and restoration.

Declarations

Ethical approval and consent to participate: Not applicable.

Consent to publish: All the co-authors agreed to publish the manuscript.

Author's Contribution: Conceptualization, Swades Pal and Pankaj Singha; Formal analysis, Swades Pal and Pankaj Singha; Methodology, Swades Pal, and Pankaj Singha; Software, Pankaj Singha; Supervision, Swades Pal; Validation: Pankaj Singha; Writing – original draft, Swades Pal and Pankaj Singha; Writing – review & editing, Swades Pal, and Pankaj Singha.

Competing interests: The authors declare that they have no conflict of interests.

Availability of data and materials All the data and materials related to the manuscript Are published with the paper, and available from the corresponding author upon request (swadespal2017@gmail.com).

Funding There is no funding source to declare. Not applicable.

Acknowledgements: The first author of the article would like to thank the University Grants Commission (UGC Ref. No.3267/(SC)(NET-JAN-2017), New Delhi, India for providing financial support as a Junior Research Fellowship (JRF) to conduct the research work presented in this paper.

Conflict of Interest: "No potential conflict of interest was reported by the authors**".**

References

Acreman, M. C., Fisher, J., Stratford, C. J., Mould, D. J., & Mountford, J. O. (2007). Hydrological science and wetland restoration: some case studies from Europe. *Hydrology and Earth System Sciences*, *11*(1), 158-169.

Arabameri, A., Chen, W., Blaschke, T., Tiefenbacher, J. P., Pradhan, B., & Tien Bui, D. (2020). Gully head-cut distribution modeling using machine learning methods—A case study of nw iran. *Water*, *12*(1), 16.

Balwan, W. K., & Kour, S. (2021). Wetland-An Ecological Boon for the Environment.

Breiman, L. (1996). Bagging predictors. *Machine learning*, *24*(2), 123-140.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Chakraborty, R., Talukdar, S., Basu, T., & Pal, S. (2018). Habitat identity crisis caused by the riparian wetland squeeze in Tangon River Basin, Barind Region, India. *Spatial Information Research*, *26*(5), 507-516.

Chandra, P., & Kumar, M. (2020). Contribution of microbes in the renovation of wetlands. In *Restoration of Wetland Ecosystem: A Trajectory Towards a Sustainable Environment* (pp. 101-124). Springer, Singapore.

Chapi, K., Singh, V. P., Shirzadi, A., Shahabi, H., Bui, D. T., Pham, B. T., & Khosravi, K. (2017). A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environmental modelling & software*, *95*, 229-245.

Chen, W., Shahabi, H., Zhang, S., Khosravi, K., Shirzadi, A., Chapi, K., ... & Ahmad, B. B. (2018). Landslide susceptibility modeling based on gis and novel bagging-based kernel logistic regression. *Applied Sciences*, *8*(12), 2540.

Choubin, B., Moradi, E., Golshan, M., Adamowski, J., Sajedi-Hosseini, F., & Mosavi, A. (2019). An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *Science of the Total Environment*, *651*, 2087-2096.

Costanza, R., d'Arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., ... & Van Den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *nature*, *387*(6630), 253-260.

Costanza, R., De Groot, R., Sutton, P., Van der Ploeg, S., Anderson, S. J., Kubiszewski, I., ... & Turner, R. K. (2014). Changes in the global value of ecosystem services. *Global environmental change*, *26*, 152-158.

Das, A., & Basu, T. (2020). Assessment of peri-urban wetland ecological degradation through importance-performance analysis (IPA): A study on Chatra Wetland, India. *Ecological Indicators*, *114*, 106274.

Das, K. (2017). NDVI and NDWI based Change Detection Analysis of Bordoibam Beelmukh Wetlandscape, Assam using IRS LISS III data. *ADBU Journal of Engineering Technology*, *6*(2).

Das, R. T., & Pal, S. (2017). Exploring geospatial changes of wetland in different hydrological paradigms using water presence frequency approach in Barind Tract of West Bengal. *Spatial Information Research*, *25*(3), 467-479.

Davidson, N. C. (2018). Ramsar convention on wetlands: scope and implementation. In *The Wetland Book I:: Structure and function, management, and methods* (pp. 451-458). Springer.

de Oliveira, E. M., & Oliveira, F. L. C. (2018). Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. *Energy*, *144*, 776-788.

Debanshi, S., & Pal, S. (2020). Wetland delineation simulation and prediction in deltaic landscape. *Ecological Indicators*, *108*, 105757.

Donchyts, G., Baart, F., Winsemius, H., Gorelick, N., Kwadijk, J., & Van De Giesen, N. (2016). Earth's surface water change over the past 30 years. *Nature Climate Change*, *6*(9), 810-813.

Dou, J., Yunus, A. P., Bui, D. T., Merghadi, A., Sahana, M., Zhu, Z., ... & Pham, B. T. (2020). Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. *Landslides*, *17*(3), 641-658.

Du, L., McCarty, G. W., Zhang, X., Lang, M. W., Vanderhoof, M. K., Li, X., ... & Zou, Z. (2020). Mapping forested wetland inundation in the Delmarva Peninsula, USA using deep convolutional neural networks. *Remote Sensing*, *12*(4), 644.

Efron, B., & Tibshirani, R. J. (1995). *Cross-validation and the bootstrap: Estimating the error rate of a prediction rule*. Division of Biostatistics, Stanford University.

Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment*, *140*, 23-35.

Ghosh, K. (2021). A framework to assess the benefits and challenges of ecosystem services of a reservoir-based wetland in the himalayan foothills. *Environmental Development*, 100669.

Gómez-Baggethun, E., Tudor, M., Doroftei, M., Covaliov, S., Năstase, A., Onără, D. F., ... & Cioacă, E. (2019). Changes in ecosystem services from wetland loss and restoration: An ecosystem assessment of the Danube Delta (1960–2010). *Ecosystem services*, *39*, 100965.

Hamza, F., & Selmi, S. (2018). Diversity of waterbirds wintering in Douz wetlands (south Tunisia): factors affecting wetland occupancy and species richness. *Ecological research*, *33*(5), 917-925.

Handbook, R. (2016). An Introduction to the Ramsar Convention on Wetlands. Gland, Switzerland: Ramsar Convention Secretariat.

Helmstetter, A., & Werner, M. J. (2014). Adaptive smoothing of seismicity in time, space, and magnitude for time-dependent earthquake forecasts for California. *Bulletin of the Seismological Society of America*, *104*(2), 809-822.

Hird, J. N., DeLancey, E. R., McDermid, G. J., & Kariyeva, J. (2017). Google Earth Engine, open-access satellite data, and machine learning in support of large-area probabilistic wetland mapping. *Remote sensing*, *9*(12), 1315.

679 Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE*
680 *transactions on pattern analysis and machine intelligence*, *20*(8), 832-844.
681 Hu, T., Liu, J., Zheng, G., Zhang, D., & Huang, K. (2020). Evaluation of historical and future
682 wetland degradation using remote sensing imagery and land use modeling. *Land Degradation*
683 *& Development*, *31*(1), 65-80.
684 Islam, A. R. M. T., Talukdar, S., Mahato, S., Ziaul, S., Eibek, K. U., Akhter, S., ... & Linh, N.
685 T. T. (2021). Machine learning algorithm-based risk assessment of riparian wetlands in
686 Padma River Basin of Northwest Bangladesh. *Environmental Science and Pollution*
687 *Research*, 1-22.
688 IUCN, 1996. Guidelines for Aid Agencies for Improved Conservation and Sustainable Use of
689 Tropical and Sub-tropical Wetlands. OECD.
690 Johnston, C. A., & McIntyre, N. E. (2019). Effects of cropland encroachment on prairie
691 pothole wetlands: numbers, density, size, shape, and structural connectivity. *Landscape*
692 *Ecology*, *34*(4), 827-841.
693 Kaplan, G., & Avdan, U. (2017). MAPPING AND MONITORING WETLANDS USING
694 SENTINEL-2 SATELLITE IMAGERY. *ISPRS Annals of Photogrammetry, Remote Sensing*
695 *& Spatial Information Sciences*, *4*.
696 Kaplan, G., & Avdan, U. (2018). Monthly analysis of wetlands dynamics using remote
697 sensing data. *ISPRS International Journal of Geo-Information*, *7*(10), 411.
698 Keddy, P., Gaudet, C., & Fraser, L. H. (2000). Effects of low and high nutrients on the
699 competitive hierarchy of 26 shoreline plants. *Journal of Ecology*, 413-423.
700 Khatun, R., & Pal, S. (2021). Effects of hydrological modification on fish habitability in
701 riparian flood plain river basin. *Ecological Informatics*, *65*, 101398.
702 Khatun, R., Talukdar, S., Pal, S., & Kundu, S. (2021). Measuring dam induced alteration in
703 water richness and eco-hydrological deficit in flood plain wetland. *Journal of Environmental*
704 *Management*, *285*, 112157.
705 Kim, J. C., Lee, S., Jung, H. S., & Lee, S. (2018). Landslide susceptibility mapping using
706 random forest and boosted tree models in Pyeong-Chang, Korea. *Geocarto*
707 *international*, *33*(9), 1000-1015.
708 Kundu, S., Pal, S., Talukdar, S., & Mandal, I. (2021). Impact of wetland fragmentation due to
709 damming on the linkages between water richness and ecosystem services. *Environmental*
710 *Science and Pollution Research*, 1-20.

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. *Environmental Science and Pollution Research*, 1-20.

Lantz, S. M., Gawlik, D. E., & Cook, M. I. (2011). The effects of water depth and emergent vegetation on foraging success and habitat selection of wading birds in the Everglades. *Waterbirds: The International Journal of Waterbird Biology*, 439-447.

Lee, O., Kim, H. S., & Kim, S. (2020). Hydrological simple water balance modeling for increasing geographically isolated doline wetland functions and its application to climate change. *Ecological Engineering*, *149*, 105812.

Leong, W. C., Kelani, R. O., & Ahmad, Z. (2020). Prediction of air pollution index (API) using support vector machine (SVM). *Journal of Environmental Chemical Engineering*, *8*(3), 103208.

Li, H., Li, T., Sun, W., Zhang, W., Zhang, Q., Yu, L., ... & Zha, X. (2021). Degradation of wetlands on the Qinghai-Tibetan Plateau causing a loss in soil organic carbon in 1966–2016. *Plant and Soil*, 1-13.

Li, N., Lu, D., Wu, M., Zhang, Y., & Lu, L. (2018). Coastal wetland classification with multiseasonal high-spatial resolution satellite imagery. *International Journal of Remote Sensing*, *39*(23), 8963-8983.

Lin, Q., Zhang, K., Shen, J., & Liu, E. (2019). Integrating long-term dynamics of ecosystem services into restoration and management of large shallow lakes. *Science of the Total Environment*, *671*, 66-75.

Liu, T., Abd-Elrahman, A., Morton, J., & Wilhelm, V. L. (2018). Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. *GIScience & remote sensing*, *55*(2), 243-264.

MA (Millennium Ecosystem Assessment), 2005. Ecosystems and Human Well-being: Synthesis. Island Press/World Resources Institute, Washington, DC.

Ma, R., Ban, J., Wang, Q., Zhang, Y., Yang, Y., He, M. Z., ... & Li, T. (2021). Random forest model based fine scale spatiotemporal O3 trends in the Beijing-Tianjin-Hebei region in China, 2010 to 2017. *Environmental Pollution*, *276*, 116635.

Mahdianpari, M., Jafarzadeh, H., Granger, J. E., Mohammadimanesh, F., Brisco, B., Salehi, B., ... & Weng, Q. (2020). A large-scale change monitoring of wetlands using time series

743 Landsat imagery on Google Earth Engine: a case study in Newfoundland. *GIScience &*
744 *Remote Sensing*, *57*(8), 1102-1124.
745 Mandal, I., & Pal, S. (2020). Modelling human health vulnerability using different machine
746 learning algorithms in stone quarrying and crushing areas of Dwarka river Basin, Eastern
747 India. *Advances in Space Research*, *66*(6), 1351-1371.
748 Mao, X., Zhang, F., Wang, G., Chu, Y., & Yuan, K. (2021). Semi-random subspace with Bi-
749 GRU: Fusing statistical and deep representation features for bearing fault
750 diagnosis. *Measurement*, *173*, 108603.
751 Martins, V. S., Kaleita, A. L., Gelder, B. K., Nagel, G. W., & Maciel, D. A. (2020). Deep
752 neural network for complex open-water wetland mapping using high-resolution WorldView-
753 3 and airborne LiDAR data. *International Journal of Applied Earth Observation and*
754 *Geoinformation*, *93*, 102215.
755 McCarthy, M. J., Radabaugh, K. R., Moyer, R. P., & Muller-Karger, F. E. (2018). Enabling
756 efficient, large-scale high-spatial resolution wetland mapping using satellites. *Remote Sensing*
757 *of Environment*, *208*, 189-201.
758 McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the
759 delineation of open water features. *International journal of remote sensing*, *17*(7), 1425-
760 1432.
761 Mi, J., Fan, L., Duan, X., & Qiu, Y. (2018). Short-term power load forecasting method based
762 on improved exponential smoothing grey model. *Mathematical Problems in*
763 *Engineering*, *2018*.
764 Mleczko, M., Mróz, M., & Fitrzyk, M. (2021). Riparian Wetland Mapping and Inundation
765 Monitoring Using Amplitude and Bistatic Coherence Data From the TanDEM-X
766 Mission. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote*
767 *Sensing*, *14*, 2432-2444.
768 Murry, B., Bowden, J., Branoff, B., García-Bermúdez, M., Middleton, B. A., Ortiz-Zayas, J.
769 R., ... & Terando, A. (2019). Perspective: developing flow policies to balance the water needs
770 of humans and wetlands requires a landscape scale approach inclusive of future scenarios and
771 multiple timescales. *Wetlands*, *39*(6), 1329-1341.
772 Nag, S. K., Liu, R., & Lal, R. (2017). Emission of greenhouse gases and soil carbon
773 sequestration in a riparian marsh wetland in central Ohio. *Environmental monitoring and*
774 *assessment*, *189*(11), 1-12.

Naghibi, S. A., Ahmadi, K., & Daneshi, A. (2017). Application of support vector machine, random forest, and genetic algorithm optimized random forest models in groundwater potential mapping. *Water Resources Management*, *31*(9), 2761-2775.

Ngoc, D. D., Loisel, H., Jamet, C., Vantrepotte, V., Duforêt-Gaurier, L., Minh, C. D., & Mangin, A. (2019). Coastal and inland water pixels extraction algorithm (WiPE) from spectral shape analysis and HSV transformation applied to Landsat 8 OLI and Sentinel-2 MSI. *Remote Sensing of Environment*, *223*, 208-228.

Nhu, V. H., Khosravi, K., Cooper, J. R., Karimi, M., Kisi, O., Pham, B. T., & Lyu, Z. (2020). Monthly suspended sediment load prediction using artificial intelligence: testing of a new random subspace method. *Hydrological Sciences Journal*, *65*(12), 2116-2127.

Ouma, D. (2020). *The Influence of Flooding on Fish Species Diversity and Fisheries Patterns in the Floodplain Lakes of Lower Tana River, Kenya* (Doctoral dissertation, University of Nairobi).

Pal, S., & Sarda, R. (2020). Damming effects on the degree of hydrological alteration and stability of wetland in lower Atreyee River basin. *Ecological Indicators*, *116*, 106542.

Pal, S., & Singha, P. (2021). Analyzing sensitivity of flood susceptible model in a flood plain river basin. *Geocarto International*, 1-34.

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. *Human and Ecological Risk Assessment: An International Journal*, *24*(5), 1291-1311.

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. *Human and Ecological Risk Assessment: An International Journal*, *24*(5), 1291-1311.

Pal, S., Saha, A., & Das, T. (2019). Analysis of flow modifications and stress in the Tangon river basin of the Barind tract. *International Journal of River Basin Management*, *17*(3), 301-321.

Pal, S., Talukdar, S., & Ghosh, R. (2020). Damming effect on habitat quality of riparian corridor. *Ecological Indicators*, *114*, 106300.

Paul, S., & Pal, S. (2020). Predicting wetland area and water depth of Ganges moribund deltaic parts of India. *Remote Sensing Applications: Society and Environment*, *19*, 100338.

Pham, B. T., Bui, D. T., & Prakash, I. (2018). Bagging based support vector machines for spatial prediction of landslides. *Environmental earth sciences*, *77*(4), 1-17.

Piao, Y., Piao, M., Jin, C. H., Shon, H. S., Chung, J. M., Hwang, B., & Ryu, K. H. (2015). A new ensemble method with feature space partitioning for high-dimensional data classification. *Mathematical Problems in Engineering*, *2015*.

Prasad, R. S., Al-Ani, M. S., & Nejres, S. M. (2015). An efficient approach for human face recognition. *International Journal of Advanced Research in Computer Science and Software Engineering*, *5*(9), 133-136.

Ramsar Convention on Wetlands, 2018. Global Wetland Outlook: State of the World's Wetlands and Their Services to People. Ramsar Convention Secretariat,

Rasyid, A. R., Bhandary, N. P., & Yatabe, R. (2016). Performance of frequency ratio and logistic regression model in creating GIS based landslides susceptibility map at Lompobattang Mountain, Indonesia. *Geoenvironmental Disasters*, *3*(1), 1-16.

Rezaee, M., Mahdianpari, M., Zhang, Y., & Salehi, B. (2018). Deep convolutional neural network for complex wetland classification using optical remote sensing imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(9), 3030-3039.

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *67*, 93-104.

Saha, T. K., & Pal, S. (2019). Emerging conflict between agriculture extension and physical existence of wetland in post-dam period in Atreyee River basin of Indo-Bangladesh. *Environment, Development and Sustainability*, *21*(3), 1485-1505.

Saha, T. K., Pal, S., & Sarkar, R. (2021). Prediction of wetland area and depth using linear regression model and artificial neural network based cellular automata. *Ecological Informatics*, *62*, 101272.

Sahana, M., Rehman, S., Sajjad, H., & Hong, H. (2020). Exploring effectiveness of frequency ratio and support vector machine models in storm surge flood susceptibility assessment: A study of Sundarban Biosphere Reserve, India. *Catena*, *189*, 104450.

Sarkar, U. K., Bakshi, S., Lianthuamluaia, L., Mishal, P., Ghosh, B. D., Saha, S., & Karnatak, G. (2020). Understanding enviro-climatological impact on fish biodiversity of the tropical floodplain wetlands for their sustainable management. *Sustainable Water Resources Management*, *6*(5), 1-12.

Shahabi, H., Shirzadi, A., Ghaderi, K., Omidvar, E., Al-Ansari, N., Clague, J. J., ... & Ahmad, A. (2020). Flood detection and susceptibility mapping using sentinel-1 remote sensing data and a machine learning approach: Hybrid intelligence of bagging ensemble based on k-nearest neighbor classifier. *Remote Sensing*, *12*(2), 266.

Shen, L., & Li, C. (2010, June). Water body extraction from Landsat ETM+ imagery using adaboost algorithm. In *2010 18th International Conference on Geoinformatics* (pp. 1-4). IEEE.

Singh, M., & Sinha, R. (2019). Evaluating dynamic hydrological connectivity of a floodplain wetland in North Bihar, India using geostatistical methods. *Science of the Total Environment*, *651*, 2473-2488.

Singha, P., Das, P., Talukdar, S., & Pal, S. (2020). Modeling livelihood vulnerability in erosion and flooding induced river island in Ganges riparian corridor, India. *Ecological Indicators*, *119*, 106825.

Sinha, P., Gaughan, A. E., Stevens, F. R., Nieves, J. J., Sorichetta, A., & Tatem, A. J. (2019). Assessing the spatial sensitivity of a random forest model: Application in gridded population modeling. *Computers, Environment and Urban Systems*, *75*, 132-145.

Slagter, B., Tsendbazar, N. E., Vollrath, A., & Reiche, J. (2020). Mapping wetland characteristics using temporally dense Sentinel-1 and Sentinel-2 data: A case study in the St. Lucia wetlands, South Africa. *International Journal of Applied Earth Observation and Geoinformation*, *86*, 102009.

Smyl, S. (2020). A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. *International Journal of Forecasting*, *36*(1), 75-85.

Song, F., Su, F., Mi, C., & Sun, D. (2021). Analysis of driving forces on wetland ecosystem services value change: A case in Northeast China. *Science of The Total Environment*, *751*, 141778.

Sothe, C., De Almeida, C. M., Schimalski, M. B., La Rosa, L. E. C., Castro, J. D. B., Feitosa, R. Q., ... & Tommaselli, A. M. G. (2020). Comparative performance of convolutional neural network, weighted and conventional support vector machine and random forest for classifying tree species using hyperspectral and photogrammetric data. *GIScience & Remote Sensing*, *57*(3), 369-394.

Sun, D., Wen, H., Wang, D., & Xu, J. (2020). A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm. *Geomorphology*, *362*, 107201.

873 Sun, D., Xu, J., Wen, H., & Wang, D. (2021). Assessment of landslide susceptibility mapping
874 based on Bayesian hyperparameter optimization: A comparison between logistic regression
875 and random forest. *Engineering Geology*, *281*, 105972.
876 Talukdar, S., & Pal, S. (2017). Impact of dam on inundation regime of flood plain wetland of
877 punarbhaba river basin of barind tract of Indo-Bangladesh. *International Soil and Water*
878 *Conservation Research*, *5*(2), 109-121.

879 Talukdar, S., & Pal, S. (2020). Modeling flood plain wetland transformation in consequences
880 of flow alteration in Punarbhaba river in India and Bangladesh. *Journal of Cleaner*
881 *Production*, *261*, 120767.
882 Tang, X., Machimura, T., Liu, W., Li, J., & Hong, H. (2021). A novel index to evaluate
883 discretization methods: A case study of flood susceptibility assessment based on random
884 forest. *Geoscience Frontiers*, 101253.
885 Tehrany, M. S., Pradhan, B., Mansor, S., & Ahmad, N. (2015). Flood susceptibility
886 assessment using GIS-based support vector machine model with different kernel
887 types. *Catena*, *125*, 91-101.
888 Tiner, R. W. (2016). *Wetland indicators: A guide to wetland formation, identification,*
889 *delineation, classification, and mapping*. CRC press.

890 Tiner, R. W., Lang, M. W., & Klemas, V. V. (Eds.). (2015). *Remote sensing of wetlands:*
891 *applications and advances*. CRC press.

892 Tong, L., Xu, X., Fu, Y., & Li, S. (2014). Wetland changes and their responses to climate
893 change in the "three-river headwaters" region of China since the 1990s. *Energies*, *7*(4), 2515-
894 2534.

895 Townshend, J. R., & Justice, C. O. (1986). Analysis of the dynamics of African vegetation
896 using the normalized difference vegetation index. *International journal of remote*
897 *sensing*, *7*(11), 1435-1445.
898 Truong, X. L., Mitamura, M., Kono, Y., Raghavan, V., Yonezawa, G., Truong, X. Q., ... &
899 Lee, S. (2018). Enhancing prediction performance of landslide susceptibility model using
900 hybrid machine learning approach of bagging ensemble and logistic model tree. *Applied*
901 *Sciences*, *8*(7), 1046.
902 Vafakhah, M., Mohammad Hasani Loor, S., Pourghasemi, H., & Katebikord, A. (2020).
903 Comparing performance of random forest and adaptive neuro-fuzzy inference system data
904 mining models for flood susceptibility mapping. *Arabian Journal of Geosciences*, *13*, 1-16.

Wu, L., Zhou, H., Ma, X., Fan, J., & Zhang, F. (2019). Daily reference evapotranspiration prediction based on hybridized extreme learning machine model with bio-inspired optimization algorithms: Application in contrasting climates of China. *Journal of Hydrology*, *577*, 123960.

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International journal of remote sensing*, *27*(14), 3025-3033.

Xu, T., Weng, B., Yan, D., Wang, K., Li, X., Bi, W., ... & Liu, Y. (2019). Wetlands of international importance: Status, threats, and future protection. *International journal of environmental research and public health*, *16*(10), 1818.

Xu, X., Chen, M., Yang, G., Jiang, B., & Zhang, J. (2020). Wetland ecosystem services research: A critical review. *Global Ecology and Conservation*, *22*, e01027.

Yan, F., & Zhang, S. (2019). Ecosystem service decline in response to wetland loss in the Sanjiang Plain, Northeast China. *Ecological Engineering*, *130*, 117-121.

Yang, H. F., Dillon, T. S., Chang, E., & Chen, Y. P. P. (2018). Optimized configuration of exponential smoothing and extreme learning machine for traffic flow forecasting. *IEEE Transactions on Industrial informatics*, *15*(1), 23-34.

Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International journal of remote sensing*, *24*(3), 583-594.

Zhang, T. Y., Han, L., Zhang, H., Zhao, Y. H., Li, X. A., & Zhao, L. (2019). GIS-based landslide susceptibility mapping using hybrid integration approaches of fractal dimension with index of entropy and support vector machine. *Journal of Mountain Science*, *16*(6), 1275-1288.

Zheng, Y., Zhang, G., Wu, Y., Xu, Y. J., & Dai, C. (2019). Dam effects on downstream riparian wetlands: the Nenjiang River, Northeast China. *Water*, *11*(10), 2038.

Zhou, J., Wu, J., & Gong, Y. (2020). Valuing wetland ecosystem services based on benefit transfer: A meta-analysis of China wetland studies. *Journal of Cleaner Production*, *276*, 122988.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterials.docx