

Multi-resource potentiality and multi-hazard susceptibility assessments of the central west coast of India applying machine learning and geospatial techniques

Pankaj Prasad^{1*}, Victor Joseph Loveson¹, Sourav Mandal², Priyankar Chandra³, Luc Cimusa Kulimushi^{4,5}

¹Geological Oceanography Division, CSIR- National Institute of Oceanography, Dona Paula 403004, Goa, India

²Ocean Engineering, CSIR- National Institute of Oceanography, Dona Paula 403004, Goa, India

³Department of Geography, Institute of Science, Banaras Hindu University, Varanasi-221005, Uttar Pradesh, India

⁴Faculty of Agronomy, Université Catholique de Bukavu, Bukavu, Democratic Republic of Congo.

⁵Centre Régional d'Etudes Interdisciplinaires Appliquées au Développement Durable (CEREIAD), Université Catholique de Bukavu, Bukavu, Democratic Republic of Congo

*Corresponding Author:Pankaj Prasad

Email: ppankaj.earthscience@gmail.com Mobile No.: +919883329432

Orcid Id: 0000-0002-3118-2201

Abstract

The understating of natural resources and hazards is very fundamental from spatial perspective to reduce resource losses and human lives. For this, the main aim of the current research is to synthesize the three natural resources (groundwater, wetland, and forest) into one multi-resource (MR) potentiality map and three natural hazards (flood, landslide, and shoreline erosion) into one multi-hazard (MH) susceptibility map. In order to achieve this goal, several inventories, geo-environmental factors, satellite bands, and indices have been taken as input datasets. Random forest (RF) model has been considered for individual mapping (except shoreline erosion mapping) due to its high precision. For shoreline erosion mapping, digital shoreline analysis system (DSAS) model has been employed. The accuracies of the model have been determined from the area under receiver operating characteristic (AUROC) curve, producer accuracy (PA), and user accuracy (UA). RF model has more than 90% of prediction accuracy in different resource and hazard mapping whereas for the forest mapping PA and UA are 90% and 80%, respectively. The results of this study for multi-resource mapping show forest (24.05%) and groundwater (22.28%) as the major resources whereas flood (17.27%) is the most destructive in comparison to other hazards. The multi-resource and multi-hazard maps of the research area provide an important tool to land managers and policymakers for sustainable development and management.

Keywords :Multi-resource; Multi-hazard; GIS; Remote Sensing; Random Forest.

1. Introduction

The demands of natural resources and frequency of natural disasters have dramatically increased worldwide in the last few decades (UNISDR, 2015; Prasad et al., 2020; Rocchi et al., 2022). The development of any area depends on the availability and sustainable utilization of natural resources such as groundwater, mineral, wetland, soil, forest, etc. The growing population in the coastal areas of India has caused an increasing demand for resources. Therefore, it is required to prepare resource potentiality or probabilistic maps for better planning and management of these resources. Resource potentiality refers to utility, availability (reserve or occurrence), and accessibility of natural resources. On the other hand, probabilistic defines the chance of occurrence of an element (resource or hazard) (usually expressed as a value between 0 and 1). For example, it is used in probabilistic wetland mapping (Maxwell et al., 2016; Hird et al., 2017).

On the other side, natural hazards pose severe threats to civilization all over the world. Natural phenomena like floods, landslides, debris flows, rockfalls, soil, and shoreline erosion have been active in geological time but their intensity varies spatially and temporally (Castillo and Gómez, 2016; Javidan et al., 2021). Globally, at the present time, the number of natural hazards and resulting casualties have been increasing at an alarming rate particularly in developing countries (Rahmati et al., 2019; Angeli et al., 2022). In the context of hazard study, the terms hazard; vulnerability; and risk are closely related to each other. Hazard is defined as an event, phenomenon, process, situation, or activity that may potentially be harmful to the affected population and damaging to society and the environment (Marker, 2013). Hazards may be characterized by location, intensity, geometry, time, and frequency. For instance, a flood hazard includes spatial extension of flood, flood depth, rainfall, frequency, and geomorphic characteristics of the region. Vulnerability expresses the sensitivity of exposure of people and assets to a hazard of a given magnitude. Risk describes the impacts of hazardous events with regard to resource loss, economic disruption, injuries, and deaths. Therefore, risk only exists when both hazard and vulnerability are present (Dewan, 2013). In search of suitable inhabitable places sometimes people end up settling in inherently risky places. Human activity also makes such places more dangerous for themselves and eventually leads to mishaps of imprudence (Javidan et al., 2021). For this, hazard mapping is needed to demarcate the vulnerable and suitable areas. In hazard mapping like landslide and flood, the word susceptibility is extensively used to mean the spatial aspects of these hazards without taking into account of potential victims and economic losses (Cuesta, 2013). The susceptibility

map of any particular hazard can be prepared using different methods and classified based on the probability values. Often susceptibility is interchangeably used with vulnerability (Marker, 2013). In order to reduce the loss of human lives and resources caused by hazards, it is required to prepare multi-hazard map rather than individual hazard maps (Javidan et al., 2021; Rocchi et al., 2022). The term multi-hazard was introduced in Agenda 21 in connection with sustainable development (UNEP, 1992; Pourghasemi et al., 2019). The goal of this Agenda 21 is to research on multi-hazard in vulnerable regions for environmental risk mitigation. On the other side, multi-resource as conceptualized in this research refers to the selection of multiple significant natural resources in the study area and delineation of their zones of occurrence and concurrence. This proposed multi-resource approach offers a holistic and realistic appraisal of resource endowments in sub-regional level. The resource potentiality of a region can be redefined by considering the interaction or coincidence (common coverage) of different natural resources. In the present work, three hazards including flood, landslide, shoreline erosion and three resources namely groundwater, wetland, and forest have been selected based on the literature and field knowledge for MH and MR mapping, respectively. This study proceeds with two vital research questions: firstly, how multi-resource potentiality and multi-hazard susceptibility maps can be generated and secondly, how the multi-resource potential and multi-hazard susceptible zones spatially vary as compared to the separate resource potential and hazard prone zone.

As a result of remarkable technological advancement, the method of spatial data generation in the form of map has shifted from traditional topographical survey to satellite remote sensing. Geospatial technology has enabled terrestrial mapping to be easier, more accurate, cost-effective, time-saving, and updated. For handling high-dimensional datasets and preparing prediction mapping, machine learning is a robust technique at the present time (Cracknell and Reading, 2014; Arora et al., 2021). Among the machine learning methods, RF model has many advantages and is more accurate in different resource and hazard mapping (Breiman, 2001; Abdi, 2020; Denize et al., 2019). For this reason, the present work has applied the RF model for MR and MH mapping. In the case of shoreline erosion mapping, the popular DSAS model has been applied. In recent years, geospatial and machine learning techniques have been employed in groundwater (Naghbi et al., 2019), land use (Talukdar et al., 2020), wetland (Wen and Hughes, 2020), flood (Arora et al., 2021), landslide (Chen et al., 2020), land subsidence (Arabameri et al., 2020), gully erosion (Kulimushi et al., 2023), forest fire (Pham et al., 2020) mapping. The number of research on

particular resources and hazards has substantially increased in this decade, though very few studies are devoted to combine them together for multi-hazard mapping (Pourghasemi et al., 2020 a,b; Javidan et al., 2021; Pouyan et al., 2021; Saha et al., 2021; Angeli et al., 2022; Arvin et al., 2023). No research work on multiple natural resources has been carried out so far. This study makes itself distinct with this modest novel attempt of mapping multi-hazard and multi-resource in a single platform.

Considering the above-mentioned research gaps, the aim of the present study is to prepare the multi-resource (groundwater, wetland, and forest) and multi-hazard (flood, landslide, and shoreline erosion) maps of the central west coast of India. These maps have been prepared with the help of geospatial and machine learning techniques and subsequently synthesized. These maps as an important tool for decision support systems will be of immense help to policymakers for sustainable development, environmental planning and management, judicious resource utilization, and hazard mitigation in the area under study. This approach can be opted for other regions of the world suffering from rapid natural resource depletion, acute shortage of resources, poverty, and the occurrence of various hazards as well.

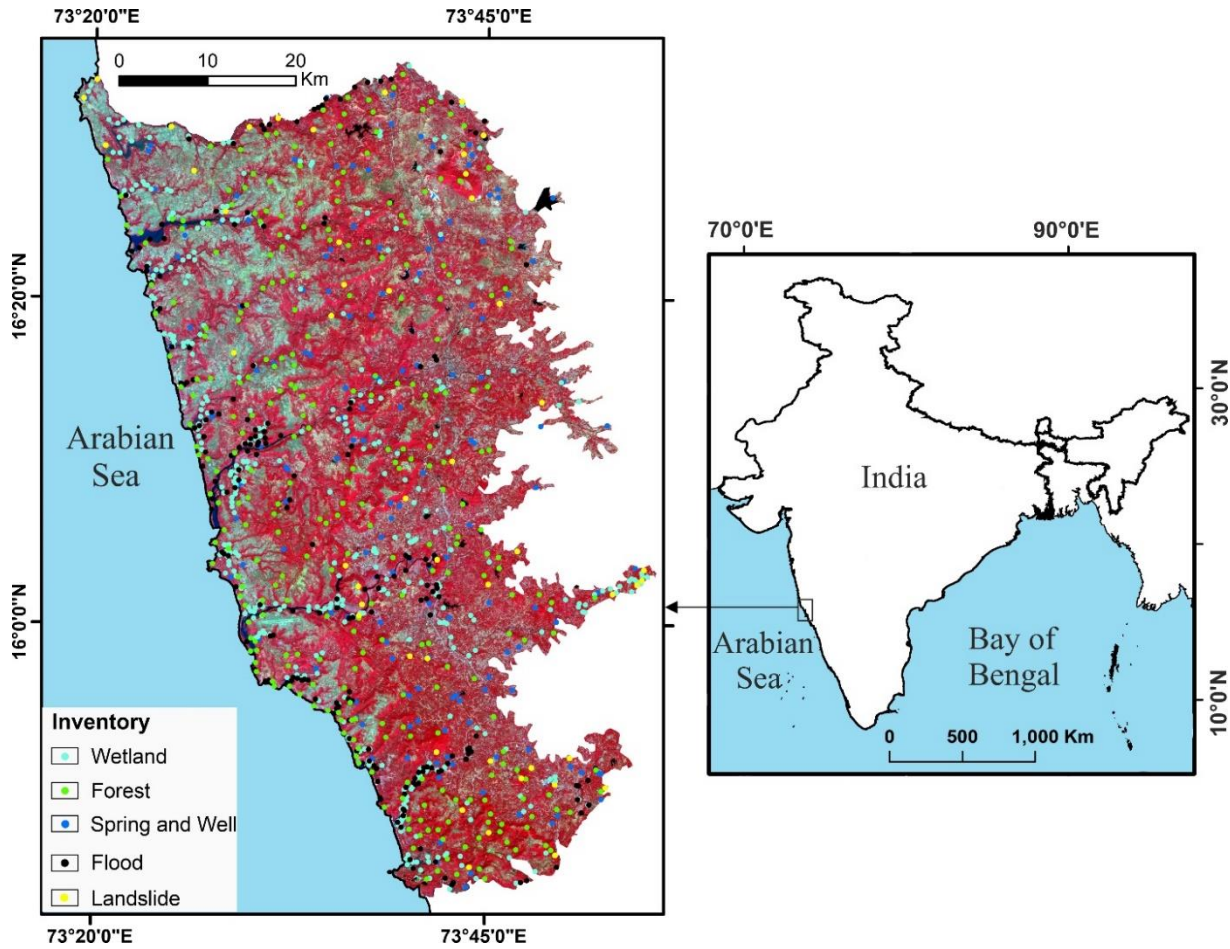


Fig. 1 Location map of the study area

2. Study area

The study area is confined within $15^{\circ}43'11.43''\text{N}$ to $16^{\circ}33'45.63''\text{N}$ and from $73^{\circ}18'36.53''\text{E}$ to $73^{\circ}55'50.07''\text{E}$ (Fig. 1). Flanked by the central west coast of India, this area is representative of a typical coastal environment. The study area rises from seashore (0 m) to landward dissected hill



Fig. 2 Field photos of (a) flood, (b) landslide, (c) shoreline erosion, (d) wetland, (e) forest, (f) spring and well.

ranges (450 m) with an average slope of about 15° . The geological formation of the study area spans from the Archean to the Recent age. Granite, basalt, lateritic, and alluvial deposits are the main constituents of lithology. More than 30% of the study area is comprised of hard rock. From the geomorphological aspect, the structural origin highly and moderately dissected plateau occupies more than 50% of the area. The principal rivers of the region namely Karli, Gad, Waghotan, Sukhandi, and Tillari originate from the Western Ghats of India and discharge into the Arabian Sea. These rivers and their tributaries have formed the trellis and parallel types of drainage patterns which are primarily controlled by lineaments. The study region belongs to the subtropical monsoon climate with a mean annual temperature of approximately 25°C and average annual rainfall of 3000 mm (Prasad et al., 2022). Maximum rainfall occurs from June to September owing to the arrival of the southwest monsoon which recharges the groundwater. Apart from the rainfall, several factors influence the groundwater potentiality in the region. The wetlands, one of the important attractions of the region, are characterized by profound diversity of flora and fauna. The thick evergreen forest is mainly found along the boundary of the Western Ghats. It has get scattered

in the coastal area. Being situated in the transition of the Arabian Sea and Western Ghats mountain range, the study area faces several hazards including flood, landslide, and shoreline erosion. In consideration of the availability of natural resources and occurrence of hazards, this study region has been selected for MR and MH mapping.

Table 1 Details of the database used in multi-resource and multi-hazard mapping

Datasets	Source of data	Acquisition dates	Derived variables	Scale or Resolution (m)
Optical	Sentinel-2 https://scihub.copernicus.eu/	11.11.2019, 21.11.2019, 26.11.2019, 16.12.2019, 31.12.2019, 30.01.2020.	Blue, Green, Red, Red edge (5) Red edge (6) Red edge (7) Red edge (8A) NIR, SWIR (11) SWIR (12), NDVI,	10 (Sentinel-2)
	Landsat-8 OLI https://earthexplorer.usgs.gov/	19.10.2016, 17.12.2016.	GNDVI, MNDWI, Land use	30 (Landsat-8)
Synthetic aperture radar (SAR)	Sentinel-1 https://scihub.copernicus.eu/	12.11.2019, 24.11.2019, 06.12.2019, 23.01.2020.	VV, VH, VV/VH, VV contrast, VV variance, VH variance, VV correlation, VH correlation, VV mean, VH mean, NPol	10
Inventory of natural resources and hazards	Prasad et al. 2020 Conservation and Management department, Maharashtra state https://sindhudurg.nic.in/en/ Prasad et al. 2021a Google earth image, Cartosat image, LISS III and LISS IV images, Field survey	2007-2016, 2018-2020, 2018-2020, 2005-2020, 2005-2020	Spring and well, Wetland, Forest, Landslide, Flood.	-
Topo-graphical	SRTM-Digital Elevation Model (DEM) https://earthexplorer.usgs.gov/	23.09.2014	Elevation, slope, aspect, plan and profile curvature, slope length, topographical TWI, TPI, TRI	30
Hydro-logical	SRTM-DEM	23.09.2014	SPI, distance from streams, drainage density,	30 -

	Department of agriculture Maharashtra state http://maharain.gov.in/	2013-2018	Rainfall	
Geological	Geological Survey of India https://www.gsi.gov.in SRTM-DEM	2001	Lithology, Distance from Lineaments, Lineament density	1:250,000 30
Geomorphology	National remote sensing centre https://bhuvan.nrsc.gov.in SRTM-DEM Landsat-8 OLI Field survey	2005-2006	Geomorphology	1:50000
Soil	National bureau of soil survey and land use planning https://www.nbsslup.in/	1996	Soil	1:500,000
Road	Topographical maps http://www.surveyofindia.gov.in/ / Google earth images	2005 2019-2022	Road density	1:50000 0.5-1

3. Data collection and processing

3.1 Inventory of resources and hazards

On the basis of available documents, relevant work and familiarity with the area of interest, the research has considered three potential natural resources namely groundwater, wetland, forest, and three common natural hazards such as flood, landslide, and shoreline erosion (Fig. 2). The major resources lying in the transitional zone of terrestrial and marine environments of the study region has direct or indirect impact on its sustainable development. There are many advantages of these resource conservation, for example, shoreline stabilization, climate regulation, air quality control, groundwater recharge, water cycling and purification, timber production, landscape aesthetics, flood control as well as cultural recreation, and spiritual enrichment (Wen. and Hughes, 2020). So, it is necessary to map such natural resources. On the other side, flood has been a grave problem in the study region. Mainly heavy monsoon rain in the Western Ghats and coastal plain causes the rivers to spill over their banks resulting in severe flash flood. The basin areas of the major rivers including Gad, Karli, Sukhnadi, Tillari, and Waghota witness flood at least once a year (ED and MPCB, 2020). As a terrible consequence, the region suffers crops damage, disruption of communication link, private and public properties loss etc. In 2005, flood inundated many parts of

research area and estimated damaged was as much as about nine crores Rs (DDMA, 2014). Besides, landslide, another crucial hazard in this region, mainly occurs during the monsoon rainfall due to slope failure. Here, rock fall and debris fall are the dominant types of landslides. Landslide events are not uniform across the area of investigation. It mostly takes place in the eastern part i.e., the break of slope of Western Ghats and steep slopes around along the riverbanks. Shoreline erosion is another growing problem in the study region. There are many anti sea erosion bunds constructed by the government in order to abate beach erosion. Some of the places which are exposed to higher erosion, have been identified in the present research and therefore it needs requisite erosion preventive action. The first and key step for the resource potentiality and hazard susceptibility mapping is to collect the inventory data. These datasets have been assembled from the existing works and field surveys (Table 1). The details of the inventory data have been given in table 2. For modelling, locations of spring and well (165), wetland (378), forest (300), landslide (90), flood (203), are plotted in point format (latitude and longitude) rather than polygon (Fig. 1). On the other side, the same number of absence points of the resources and hazards under consideration has been taken to train machine learning model. There are many classifications ratio of the datasets such as 70:30, 50:50, 60:40, 80:20 have been employed in various studies. Among these, most of the researchers have been used 70:30 as a standard partition of training and testing, respectively (Chapi et al., 2017; Naghibi et al., 2019; Arora et al., 2021). In the current research also, the total points of each inventory have been classified in the ratio of 70:30. The methodology of this study is depicted in figure 3.

Table 2 Details of inventory of natural resources and hazards

Inventory of presence	Numbers (Y)	Inventory of absence	Numbers (N)	Total (Y+N)	Training (70%)	Testing (30%)
Spring and well	165	Non-spring and well	165	330	231	99
Wetland	378	Non-wetland	378	756	529	227
Forest	300	Non-forest	300	600	420	180
Landslide	90	Non-Landslide	90	180	126	54
Flood	210	Flood	210	420	294	126

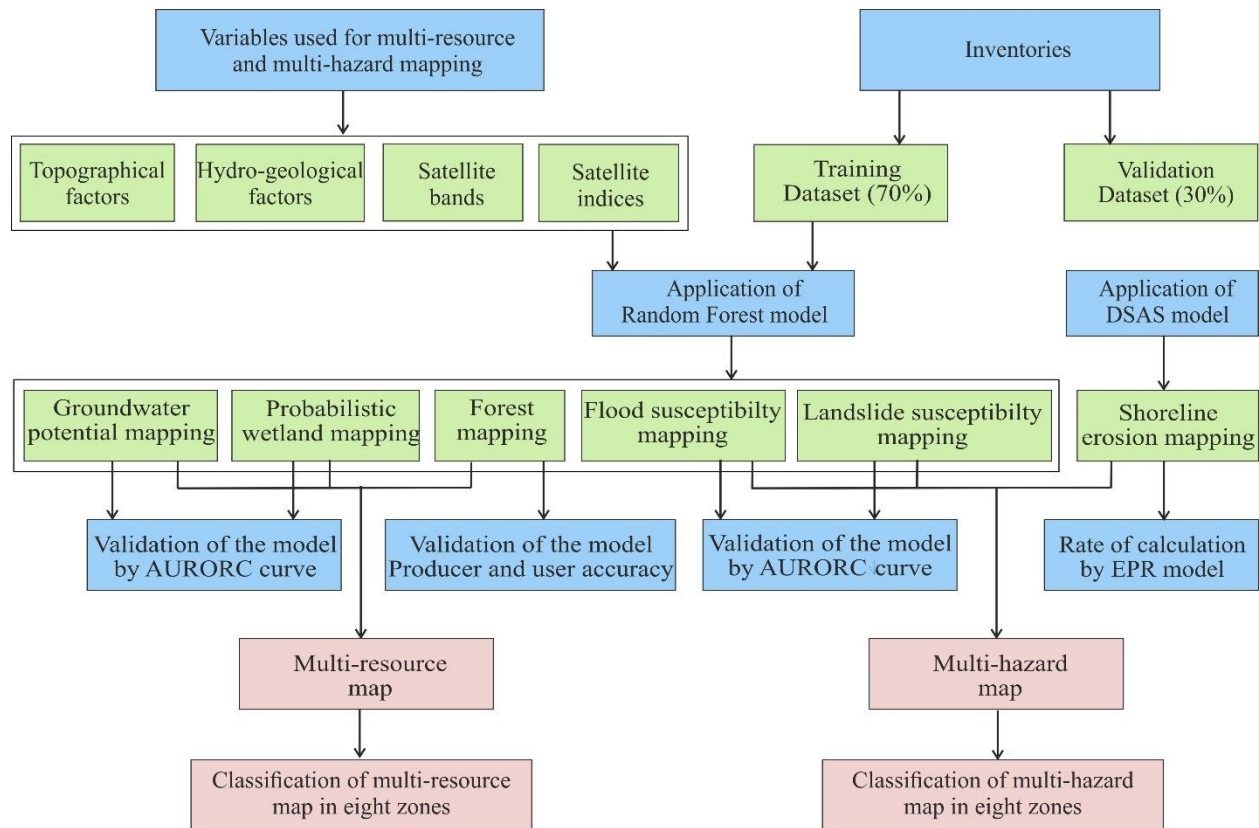


Fig. 3 Flowchart of the present study for multi-resource and multi-hazard mapping

3.2 Conditioning factors of multi-resource and multi-hazard

The appropriate selection of influencing factors is necessary for individual resource and hazard mapping, though, there are no specific guidelines for selecting conditioning variables of natural hazards. Therefore, based on the previous literature and field knowledge a number of factors including topographical, hydrological, geological, satellite bands, and indices have been selected for MH and MH mapping (Table 3).

The SRTM digital elevation model (DEM) data has been used to extract the data of topographic factors. Elevation is the primary conditioning factor in resource potential and hazard susceptibility mapping. It has been seen that the importance of elevation is highest in many studies as it determines temperature, moisture, and wind direction (Pourghasemi et al., 2020a). Slope is another key factor in groundwater potentiality and the occurrence of wetland, landslide, and flood. The range of elevation and slope of the study region is 0-450 m and 0-51°, respectively. The variable

Table 3 The variables used in multi-resource and multi-hazard mapping

Groundwater	Wetland	Forest	Flood	Landslide
Elevation	Elevation	VV	Elevation	Elevation
Slope	Slope	VH	Slope	Slope
Slope Length	Curvature	VV/VH	Aspect	Slope Length
Aspect	Drainage density	VV contrast	Topographical roughness index	Topographical roughness index
Distance from streams	Topographical position index	VV variance	Stream power index	Stream power index
Topographical wetness index	Topographical wetness index	VV correlation	Topographical wetness index	Topographical wetness index
Plan curvature	VV mean	VV mean	Distance from streams	Drainage density
NDVI	VH mean	VH mean	NDVI	NDVI
Rainfall	NPol	VH correlation	Rainfall	Rainfall
Lithology	BLUE (2)	BLUE band (2)	Lithology	Lithology
Soil	GREEN band (3)	GREEN band (3)	Soil	Soil
Geomorphology	RED band (4)	RED band (4)	Geomorphology	Geomorphology
Landuse	NIR band (8)	NIR band (8)		Road density
Profile curvature	MNDWI	MNDWI		Lineament density
Distance from lineaments	GNDVI	GNDVI		
		NDBI		
		RED edge (5)		
		RED edge (6)		
		RED edge (7)		
		RED edge (8A)		
		SWIR (11)		
		SWIR (12)		
		VH variance		
		VH contrast		

VV= Vertical Transmit-Horizontal Receive Polarisation
 VH= Vertical Transmit-Horizontal Receive Polarisation
 Npol=Normalized Polarization
 NDVI=Normalized Difference Vegetation Index
 NDBI=Normalized Difference Built-up Index
 GNDVI=Green Normalized Difference Vegetation Index
 MNDWI= Modified Normalized Difference Water Index

curvature shows the shape of the slope which influences the physical characteristics of the drainage basin (ESRI, 2016). Curvature can be expressed as plan curvature (perpendicular to the slope) and profile curvature (parallel to the slope). Aspect defines the downslope direction of the pixels and it determines the groundwater availability in different slope directions (0-360°). In this study, aspect is used only for groundwater potential mapping because it is exposed to sunlight, winds, and rainfall (Chen et al. 2018). Slope length (LS) is another topographical factor that demonstrates

the length (L) and steepness (S) of the slope which is employed in groundwater and landslide mapping. Applying the following formula, LS value is obtained which differs from 0 to 47.2.

$$LS = (fa \times cellsize/22.13)^{0.4} \times (\sin\theta/0.0896)^{1.3} \quad (1)$$

where *fa* refers to flow accumulation and θ represents the slope in degrees.

Topographic position index (TPI) defines the relative elevational position of an area. TPI value (-21.6-20) is calculated from the difference between particular pixel of the DEM (*Ep*) and mean value of neighboring cells(*Em*) (Reu et al., 2013).

$$TPI = Ep/Em \quad (2)$$

Topographic Ruggedness Index (TRI) illustrates the degree of changing topographic gradient between adjacent cells of a DEM. It is derived using the following formula.

$$TRI = Abs(maximum^2 - minimum^2) \quad (3)$$

here, maximum and minimum represent the absolute (Abs) elevation values of DEM. The lowest TRI value is 0 while the highest value is 22.5. Topographical wetness index (TWI) describes the topographic control on wetness conditions of an area. It has been adopted in groundwater, wetland, flood, and landslide mapping. Stream power index (SPI) is a vital factor in hazard studies that reveals the erosive power of the flowing water (De Rosa et al., 2019). For this, the present study has included SPI in flood and landslide susceptibility mapping. The variables TWI and SPI have been generated from the catchment area (A) and slope (β) using equations 4 and 5 (Moore et al. 1991):

$$TWI = \ln(A/\tan\beta) \quad (4)$$

$$SPI = A \times \tan\beta \quad (5)$$

The calculated TWI and SPI values range from 2.8-24.8 and 0-25.9, respectively. Areas with high values of these indexes indicate high probability of flood and landslide events.

In the monsoon affected region, rainfall is one of the essential conditioning factors because it readily reflects on groundwater condition, flood, and landslide occurrences. The six years (2013-2018) average annual rainfall (2100-4300 mm) data from twenty-eight weather stations have been collected. The inverse distance weighted (IDW) interpolation method has been applied to prepare the rainfall map of the research area using spatial analyst tool of ArcGIS. IDW interpolation

method calculates the unknown raster cell values using a linearly weighted combination of known points of rainfall stations. Lineament is an affecting factor in groundwater potentiality and landslide zonation mapping. Lineament can be quantified as a variable in terms of distance from lineaments (from the spring inventory) or lineament density (landslide mapping). Like lineament, distance from streams (from the spring and flood locations) and drainage density (landslide and wetland mapping) are used due to their sensitivity to moisture. The factor road density is very meaningful for landslide mapping as the high density of the roads in high altitude regions are prone to landslides (Huang et al., 2022).

The categorical variable namely geomorphology, lithology, and soil, have been used in the present study for groundwater, flood, and landslide mapping. The different lithological characteristics influence porosity, permeability, surface runoff, and rock stability which are directly related to groundwater potentiality, landslide, and flood events. With regard to geomorphology, it controls groundwater storage and movement, slope instability, and flow patterns (Saha et al., 2021). The property of soil affects infiltration, surface runoff, and erosional capacity that highly impact groundwater availability, spatial extension of flood and frequency of landslide. The nature and characteristics of the soil of the study are given in table 4.

Table 4 Nature and characteristics of soil in the study region (Source: NBSSLUP, 1996)

Soil types	Nature and characteristics
Ultic Typic Haplustalfs	Moderately deep, well drained loamy soils on gently sloping valley lands with moderate erosion; associated with deep, well drained, clayey soils with moderate erosion.
Typic Ustropepts	Very shallow, well drained loamy soils on moderately sloping undulating lands with mesas and narrow valleys with severe erosion and moderate stoniness; associated with shallow, well drained, loamy soils with severe erosion and strong stoniness.
Ultic Haplustalfs	Moderately deep, well drained, loamy soils on moderately sloping undulating lands with mesas and narrow valleys with moderate erosion and moderate stoniness; associated with shallow, well drained, loamy soils with moderate erosion and strong stoniness.
Lithic Ustorthents	Extremely shallow, somewhat excessively drained, loamy soils on moderately sloping lands with residual hills with severe erosion and strong stoniness; associated with extremely shallow, somewhat excessively drained, loamy soils on gently sloping lands with severe erosion and strong stoniness.

The land use and land use changes affect the groundwater potentiality (Chen et al., 2018). The land use map of the study region has been prepared from the Landsat-8 and used in groundwater potential mapping.

Forest cover map has been generated from the satellite data. For this, different optical bands (BLUE band-2, GREEN band-3, RED band-4, NIR band-8, RED edge-5, RED edge-6, RED edge-7, RED edge-8A, SWIR -11, SWIR-12), indices (GNDVI, NDBI, MNDWI), microwave variables (VV, VH, VV/VH, VV contrast, VH contrast, VV variance, VH variance, VV correlation, VH correlation, VV mean, VH mean, Npol) have been considered (Table 3). Some of the optical bands and indices have also been employed in probabilistic wetland mapping (Table 3). The indices have been measured based on the sensitivity of the bands. The values derived from the indices vary between -1 and +1. The following equations (6 to 10) are used to measure the above indices:

$$NDVI = (NIR - R)/(NIR + R) \quad (6)$$

$$GNDVI = (Green - Red)/(Green + Red) \quad (7)$$

$$MNDWI = (Green - SWIR1)/(Green + SWIR1) \quad (8)$$

$$NDBI = (SWIR - NIR)/(SWIR + NIR) \quad (9)$$

$$NPol = (VH - VV)/(VH + VV) \quad (10)$$

The input layers have various scales. For the modelling, all the input and output thematic maps of this study have been transformed into uniform pixel sizes (30 × 30m) using geospatial information system (GIS) software. The transformed thematic layers of the study region consist of 3,530,426 pixels with 2201 columns and 3159 rows. The maps of the conditioning variables (excluding satellite bands) have been shown in figure 4.

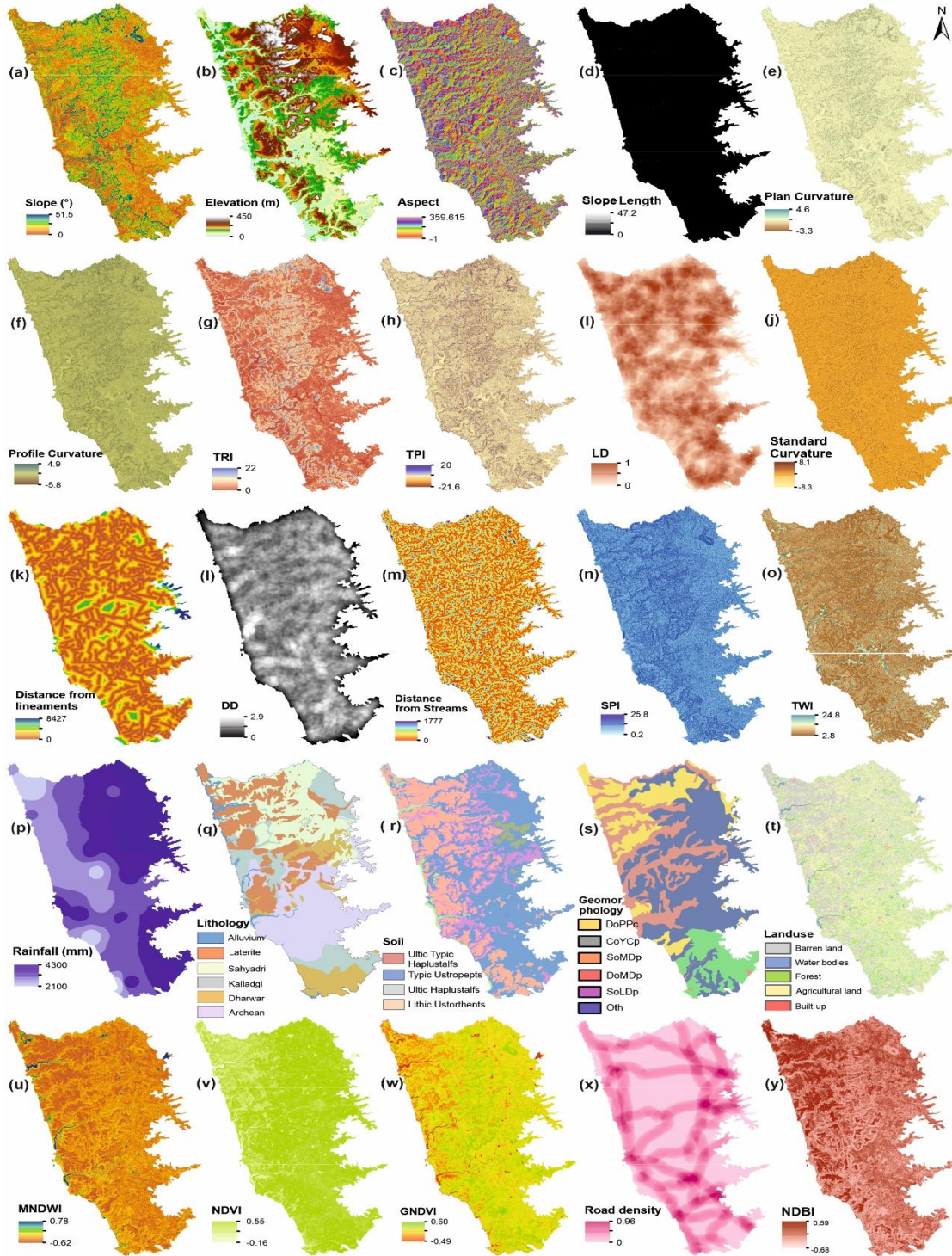


Fig. 4 Spatial distribution and classes of categorical variables (a) slope, (b) elevation, (c) aspect, (d) slope length, (e) plan curvature, (f) profile curvature, (g) TRI, (h) TPI, (i) LD, (j) standard curvature, (k) distance from lineaments, (l) DD, (m) distance from Streams, (n) SPI, (o) TWI, (p) rainfall (q) lithology, (r) soil, (s) geomorphology, denudational origin-pediment pediplain complex (DoPPc); coastal origin-younger coastal plain (CoYCp); structural origin moderately dissected plateau (SoMDp); denudational origin moderately dissected plateau (DoMDp); structural origin-low dissected plateau (SoLDp); others (Oth), (t) land use, (u) MNDWI, (v) NDVI, (w) GNDVI, (x) road density, (y) NDBI.

4. Model building strategies

4.1 Multicollinearity diagnosis

Multicollinearity is a statistical method that defines the linear relationship between two or variables (Alin, 2010). Many methods such as variance inflation factors (VIF), spearman coefficient, pearson's correlation coefficients are employed to check the multicollinearity. In the current research, variance inflation factor (VIF) method has been applied for multicollinearity test of groundwater, flood and landslide conditioning factors. For wetland and forest mapping, multicollinearity test has not been conducted because most of the variables are different satellite bands and their indices. The VIF values greater than 10 and tolerance values less than 0.1 have been considered for the rejection of the variable (Nohani et al., 2019; Wang et al., 2019). The VIF values have been measured using the following equations:

$$Tolerance = R_C^2 \quad (11)$$

Here, R_C^2 represents the coefficient of determination of regression of explanatory C on all the other explanatories.

$$VIF = \frac{1}{Tolerance} \quad (12)$$

4.2 Random Forest (RF)

Random forest is an ensemble based non-parametric model developed by Breiman (2001) for both classification and regression problems (Feizizadeh et al., 2021). RF model uses two robust methods, i.e. bagging and random subspace selection for improving the precision of the model (Adam et al., 2014; Talukdar et al., 2020). This algorithm generates many decision trees by bootstrap samples for classification. Other than bootstrap samples, the rest are recognized as out of bag (OOB) samples which are used to calculate misclassification errors. The OOB error (OOB_e) is measured as:

$$OOB_e = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (13)$$

where y_i and \hat{y}_i are the observed and model output, respectively; n represents the total number of out of bag samples. The main advantages of the RF model are i) it can enhance the precision of the classification; ii) it manages large input datasets with lesser processing time and also handles different types of datasets; iii) it escapes the errors of bias and overfitting of the input datasets; iv)

it is able to measure the importance of each selected factors in the classification; v) rescaling and transformation of the datasets are not required (Berhane et al., 2018; Prasad et al., 2020; Talukdar et al., 2020).

4.3 RF model construction

The higher variability of the values among the selected factors may affect the model result. To escape this problem, the pixel values of the variables are normalized between 0 and 1 using min-max normalization. For the preparation of RF model, the inventory data have been divided into Y and N, where Y represents the presence of spring, well, wetland, forest, landslide, flood, and N denotes the absence of these resources and hazards (Table 2). Afterwards, the input datasets have been grouped into two categories i.e., training (70% of the dataset) and testing (30% of the dataset). The training dataset has been used for model building with both inventory and controlling factor datasets. On the other hand, testing dataset consisting of variables data has been kept reserved for verification purposes.

It is vital to apply the cross-validation method in machine learning model to avoid bias and noise (Nematzadeh et al., 2015; Amiri et al., 2019). Several methods namely leave-p-out cross-validation, leave-one-out cross-validation, validation set approach, k-fold cross-validation, and stratified k-fold cross-validation are used for cross-validation. Among them, k-fold cross-validation is popular and widely applied method in machine learning (Chapi et al., 2017; Chen et al., 2020; Prasad et al., 2021b). For this, the current research has applied the 10 cross-validation method (k-fold cross-validation) to training datasets. To execute this method, at first, the training datasets have been distributed into ten equal portions. Then, each part of these data is served as a validation set and the final accuracy is derived from the mean results of the 10 iterations. At last, this training model has been used to predict the testing dataset (Fig. 5).

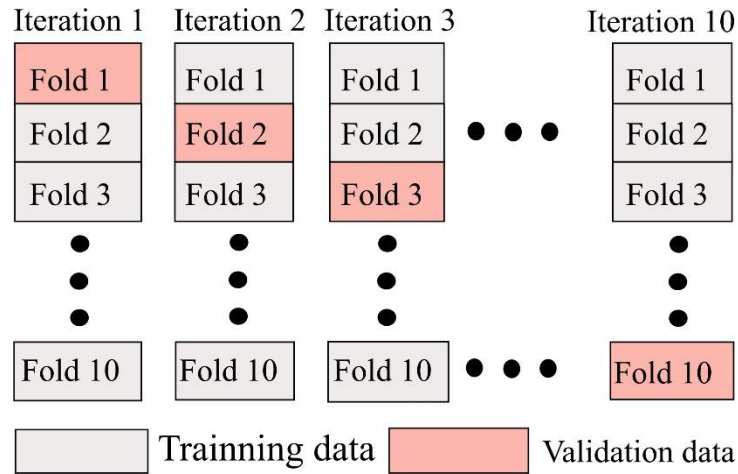


Fig. 5 10 Cross-validation of the tanning data

There are number of tuning parameters namely number of trees node size, mtry, sample size, splitting rule, and replacement involved in the RF model (Table 5). The default values of the hyperparameters in RF algorithm are well-known to provide good accuracy compared to the other algorithms in machine learning (Adam et al., 2014; Noi and Kappas, 2018; Probst et al., 2019). Moreover, according to Breiman (2001), more number trees do not harm the model but unnecessarily it can take more time for processing the input datasets. For this reason, the current research has employed the default parameters in each resource and hazard modelling (Table 5). The classification tasks of the model have used the random forest package in the R environment.

Table 5 Different hyperparameters of applied RF model

4.4 DSAS model for shoreline erosion mapping

There are several methods which were followed in previous studies including automatic, manual

Hyperparameter	Description	Values in resource and hazard mapping
number of trees	Number of trees in the forest	500
node size	Minimum number of observations in a terminal node	1
mtry	Number of drawn candidate variables in each split	\sqrt{p} here p is the number of conditioning factors.
sample size	Number of observations that are drawn for each tree	n
splitting rule	Splitting criteria in the nodes	Random (num.random.splits)
replacement	Draw observations with or without replacement	TRUE (with replacement)

digitization for shoreline extraction (Yamano et al., 2006; Mahapatra et al., 2014; Yasir et al., 2020). Automatic shoreline delineation is bewildering because of the water saturated zone lying in the transition of land and water (Maiti and Bhattacharya 2009; Nandi et al., 2016; Mondal et al., 2017). To avoid this problem, in this research, tasseled cap transformation (TCT) technique has been employed for shoreline delineation. To implement this procedure, satellite bands and various indices namely NDVI, brightness, greenness, and wetness for land and water boundary demarcation have been taken into account. After that, shoreline pixels have been transformed into a binary image. At last, this image has been converted into the polyline representing the shoreline. This procedure has been repeated for all the satellite images. The transects interval taken for this study area is 500 m.

End point ratio (EPR) is an extensively used method by the reasons of its robustness and simplicity for shoreline change and future prediction measurement (Mukhopadhyay et al., 2012; Mondal et al., 2017). This method measures the per year coastline shifting. The EPR method is estimated as follows:

$$EPR = \frac{D_1 - D_2}{T_1 - T_2} \quad (14)$$

D1 and D2 define the distance (m) whereas T1 and T2 represent the time between the oldest and most recent shorelines. Linear regression (LR) method is also used to measure the coastline change rate. It is determined by fitting a least-squares regression line to all shoreline points for a particular transect.

4.5 Model verification

The final verification step in machine learning is crucial to evaluate the model. The precision of the RF model has been done by area under the receiver operating characteristic curve. The detailed theoretical background of ROC and AUC methods is provided in Chapi et al. (2017). The AUC value differs from 0 to 1 where 0 indicates that model has failed to predict the output and 1 denotes the correct prediction of all the datasets. Hosmer (2000) has proposed three classes including 0.7-0.8 (acceptable), 0.8-0.9 (excellent), and 0.9-1 (outstanding) of AUC value based on the model performance. In land-cover studies, most of the existing research has used producer and user accuracy for land-cover class validation (Adam et al., 2014; Steinhausen et al., 2018; Abdi, 2019).

In line with the current practice, the study has also preferred producer and user accuracy for the verification of forest mapping.

4.6 Multi-resource and multi-hazard maps preparation

The outputs of three resource potentiality (groundwater, wetland, and forest) and three hazard susceptibility (flood, landslide, and shoreline erosion) maps have been used to prepare multi-resource and multi-hazard maps. The high potential and susceptible zones of the separate resource and hazard maps respectively, are considered to prepare the MH and MR maps. Firstly, these individual raster images have been transformed into vector maps with the same resolution. These vector images have been used as input layers for preparing MR and MH maps. Secondly, multi-resource and multi-hazard maps of the entire study region have been generated by synthesizing single resource potential and hazard susceptibility maps respectively, using union tool. At last, both maps of the entire study region have been classified into eight classes and the area under each class has been calculated. The same series of steps have been followed in other multi-hazard mapping (Pourghasemi et al., 2019; Rusk et al., 2022). The whole procedure of producing these maps has been carried out in the ArcGIS environment (Fig 6).

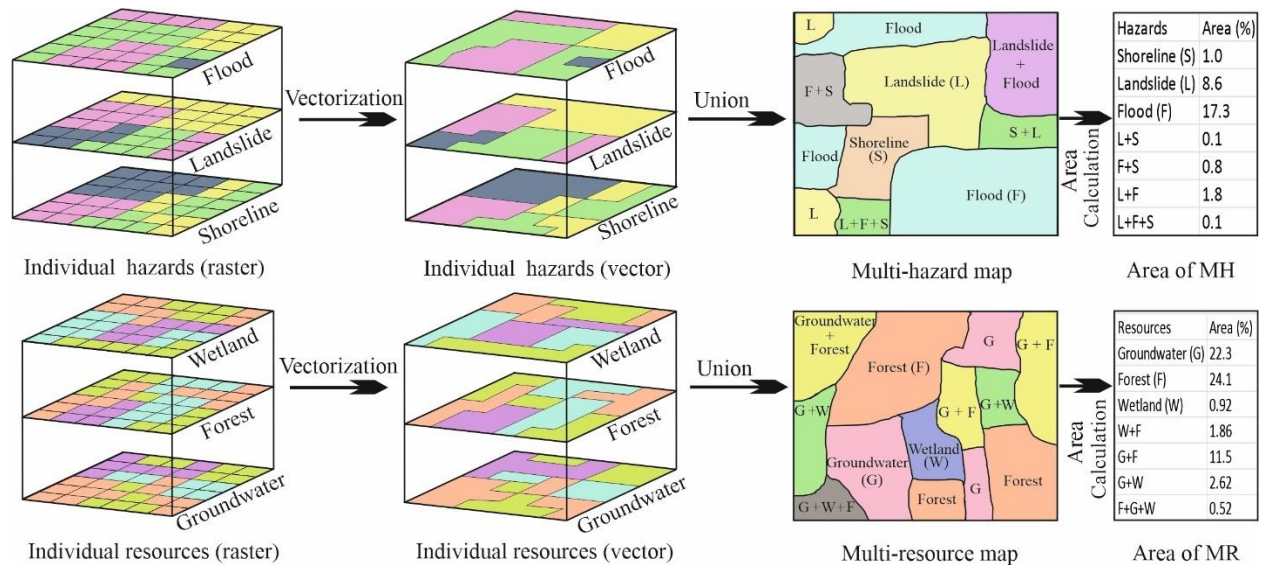


Fig. 6 Steps of preparation of multi-hazard and multi-resource maps

5. Results

In this study, VIF method has been used to check whether multicollinearity exists or not among the variables. The results of multicollinearity diagnosis of different variables in groundwater, flood and landslide have been provided in table 6. Since the VIF value is less than 10 for each variable, all the primarily chosen affecting factors have been retained.

Table 6 Multicollinearity analysis

Groundwater	VIF	Flood	VIF	Landslide	VIF
Elevation	1.94	Elevation	2.07	Elevation	3.09
Slope	2.17	Slope	4.89	Slope	4.51
Slope Length	1.50	Aspect	1.21	Slope Length	2.20
Aspect	1.08	TRI	4.60	TRI	3.79
Plan curvature	1.43	SPI	1.31	SPI	1.30
TWI	1.82	TWI	2.09	TWI	4.60
Distance from streams	1.40	Distance from streams	1.26	Drainage density	1.35
NDVI	1.75	NDVI	1.33	NDVI	1.98
Rainfall	1.39	Rainfall	1.39	Rainfall	1.28
Lithology	1.13	Lithology	1.12	Lithology	1.40
Soil	1.14	Soil	1.12	Soil	1.48
Geomorphology	1.98	Geomorphology	1.59	Geomorphology	1.73
Landuse	1.29			Road density	1.10
Profile curvature	1.40			Lineament density	1.35
Distance from lineaments	1.18				

The present study takes into account three important resources and three major hazards of the study region for MR and MH mapping (Fig. 7 and 8). The resources have been mapped by means of geospatial and machine learning techniques. In machine learning, RF method has been employed to predict the potentiality of single resources and susceptibility of hazards except for shoreline erosion. Among the many classification methods, in the current research, natural break method has been adopted for individual natural resource and hazard maps classification owing to its superiority over other methods (Chapi et al., 2017). The area of interest has been classified into five zones: very low, low, medium, high, and very high based on the outputs of all pixels. Among the five classes, only very high class has been considered for the preparation of MH and MR maps. The total areas of very high potential resources are shared by groundwater, wetland, and forest resources with 37.97%, 6.02%, 37.06% of the research region, respectively. On the other side, very high susceptible hazard zones of flood, landslide, and shoreline erosion occupy 16.25%, 10.73%, 1.85% areas of the study region, respectively.

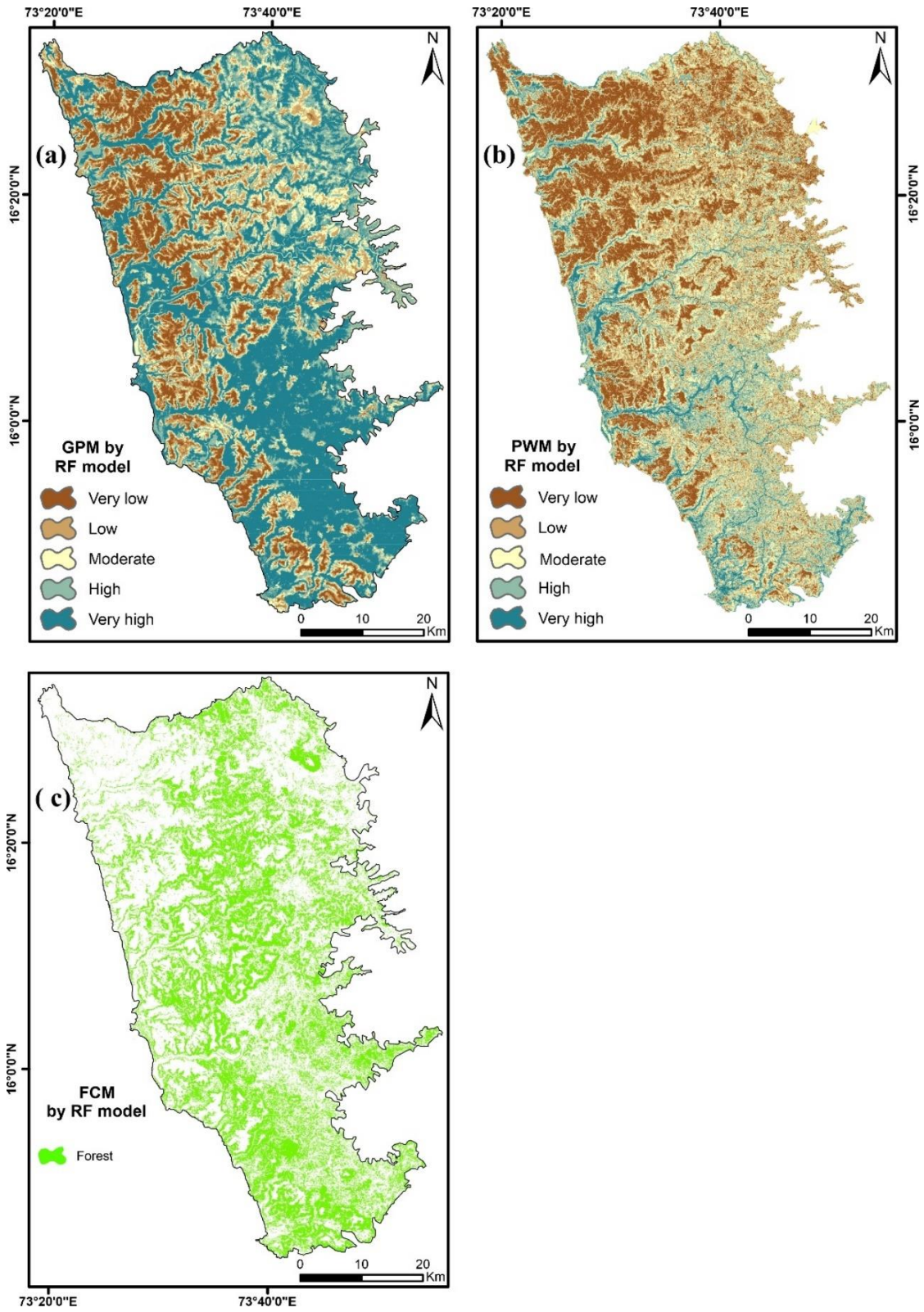


Fig. 7 (a) groundwater potential, (b) probabilistic wetland, and (c) forest cover maps derived from the RF model

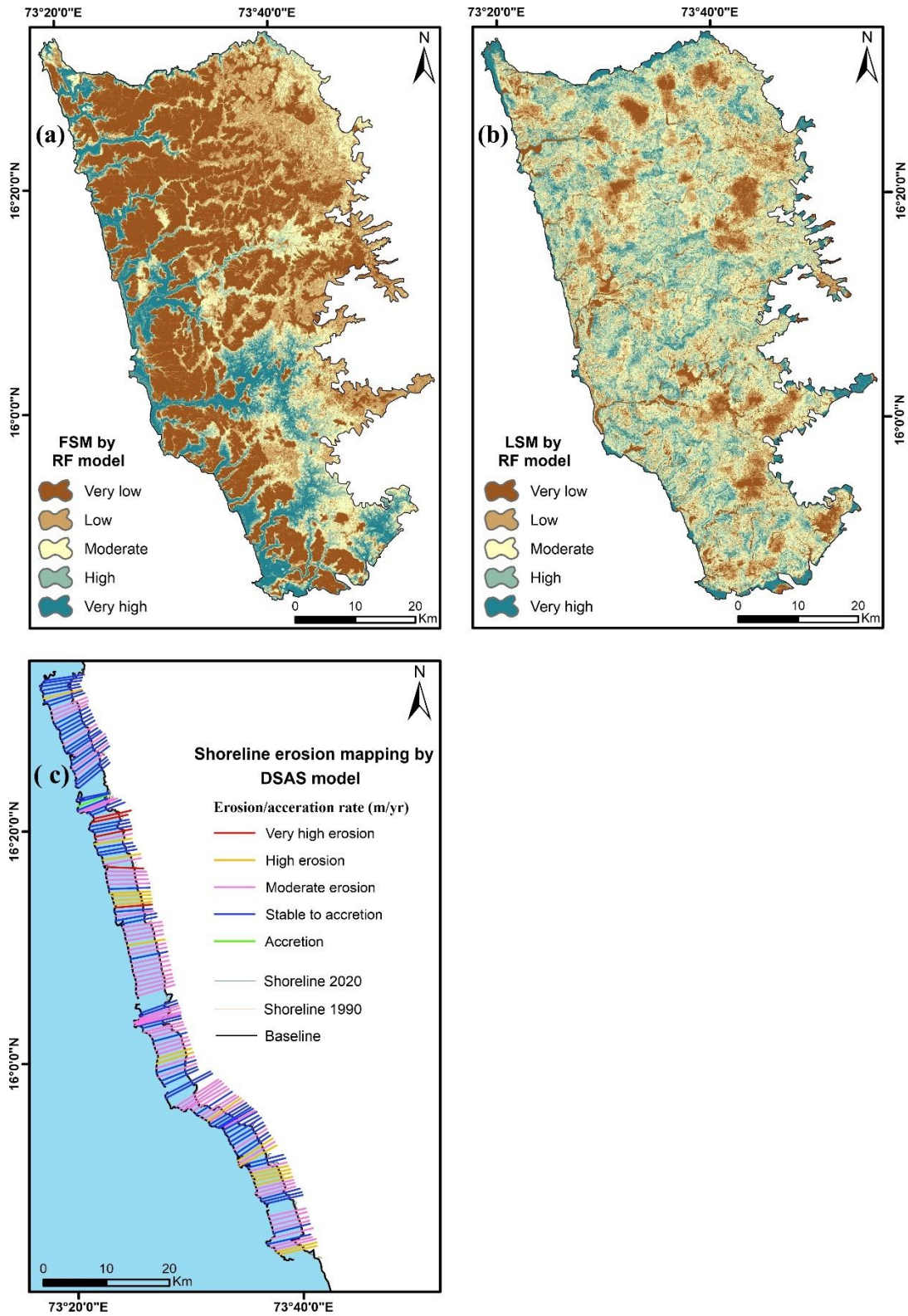


Fig. 8 (a) flood susceptibility, (b) landslide susceptibility maps derived from the RF model and (c) shoreline erosion map from the DSAS model

The rate of shoreline changes from 1990 to 2020 along the central west coast of India has been demonstrated in this study (Fig. 8c). The total length of the shoreline is about 130 km including both rocky and sandy beaches. The landward migration of shoreline with reference to the baseline is referred to as erosion (-) whereas shifting toward seaward is regarded as accretion (+) at each transect of the segments (Fig. 8c). Shoreline change rate between 1990 and 2020 using EPR and LR methods has been depicted in table 7. Based on EPR model, statistics such as mean erosion and accretion rate, accretion, erosion, standard deviation, maximum and minimum rate of both accretion and erosion have been quantified from 1990 to 2020 at 10 years intervals.

Table 7 Details of transects and shoreline rate (m/yr) for the period of 1990-2000, 2000-2010, 2010-2020, and 1990-2020 in study region.

Parameters /years	1990-2000	2000-2010	2010-2020	1990-2020
No of transects	156	155	157	157
Accretion transects	49	37	19	10
Erosion transects	107	128	138	147
Mean erosion and accretion change rate (m/yr)	-0.85	-1.24	-1.42	-1.85
Accretion (m/yr)	3.48	3.10	2.94	2.76
Erosion (m/yr)	-1.39	-1.82	-1.92	-2.16
Standard deviation	2.45	3.25	2.89	3.01
Maximum rate of accretion (m/yr)	23.24	17.42	18.47	17.88
Maximum rate of erosion (m/yr)	14.25	22.23	17.45	18.68
Minimum rate of accretion (m/yr)	0.07	0.03	0.1	0.08
Minimum rate of erosion (m/yr)	0.08	0.05	0.06	0.1

For calculating erosion and accretion rate, a maximum of 157 transects are drawn. It has been seen that the accretion transects continuously decrease whereas erosion transects show an increasing trend (Table 7). The mean change (erosion/accretion) of all transects are -0.85 m (1990-2000), -1.24 m (2000-2010), and -1.42 m (2010-2020). Comparatively, the accretion rate remains higher than the erosion rate throughout the study period. The maximum accretion and erosion stand at

17.88 m and 18.68 m, respectively while the minimum lie at 0.1 m and 0.08 m, respectively. According to the erosion and accretion rate, transects of the shoreline have been classified into very high erosion, high erosion, moderate erosion, stable to accretion and accretion (Fig. 8c).

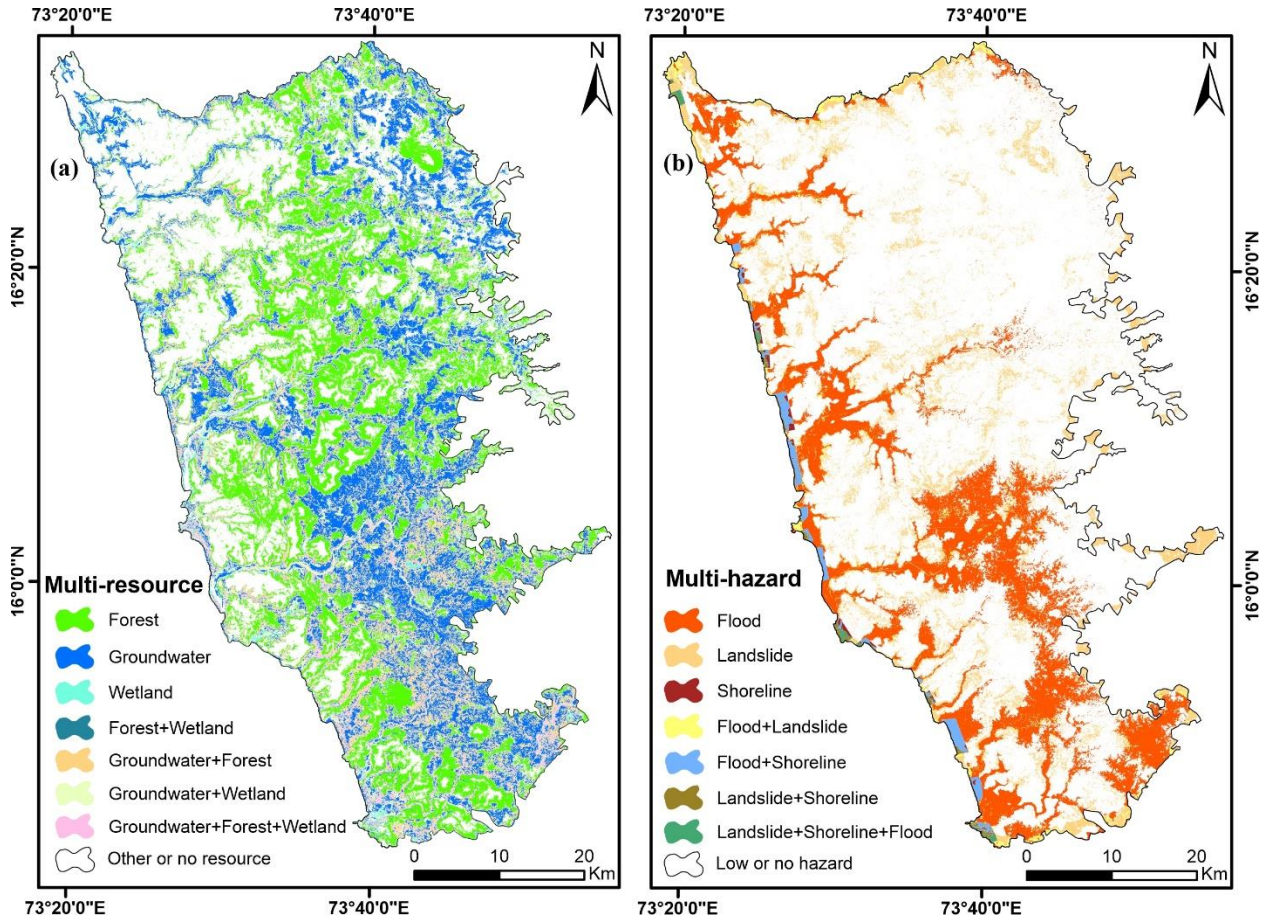


Fig. 9 (a) Multi-resource and (b) multi-hazard maps of the area of interest

5.1 Multi-resource potential map

In order to produce multi-resource map, the above three potential resources have been superimposed with each other (Fig. 9a). The areas under these resources have been calculated and classified into eight classes individually and collectively as well (Fig. 10a). As it can be seen in the figure, forest alone covers 24.05% of the total research area followed by high groundwater potential area (22.28%), groundwater with forest (11.5%), groundwater with wetland (2.62%), wetland with forest (1.86%), wetland (0.92%), and common coverage of the trio (0.52%). The result reveals that more than 50% of the area is potential for these natural resources. The outputs point out that the availability of the resources like forest, groundwater, and wetland have greatly

influenced the life of inhabitants of this region. Though, some parts of the regions are deprived of these benefits.

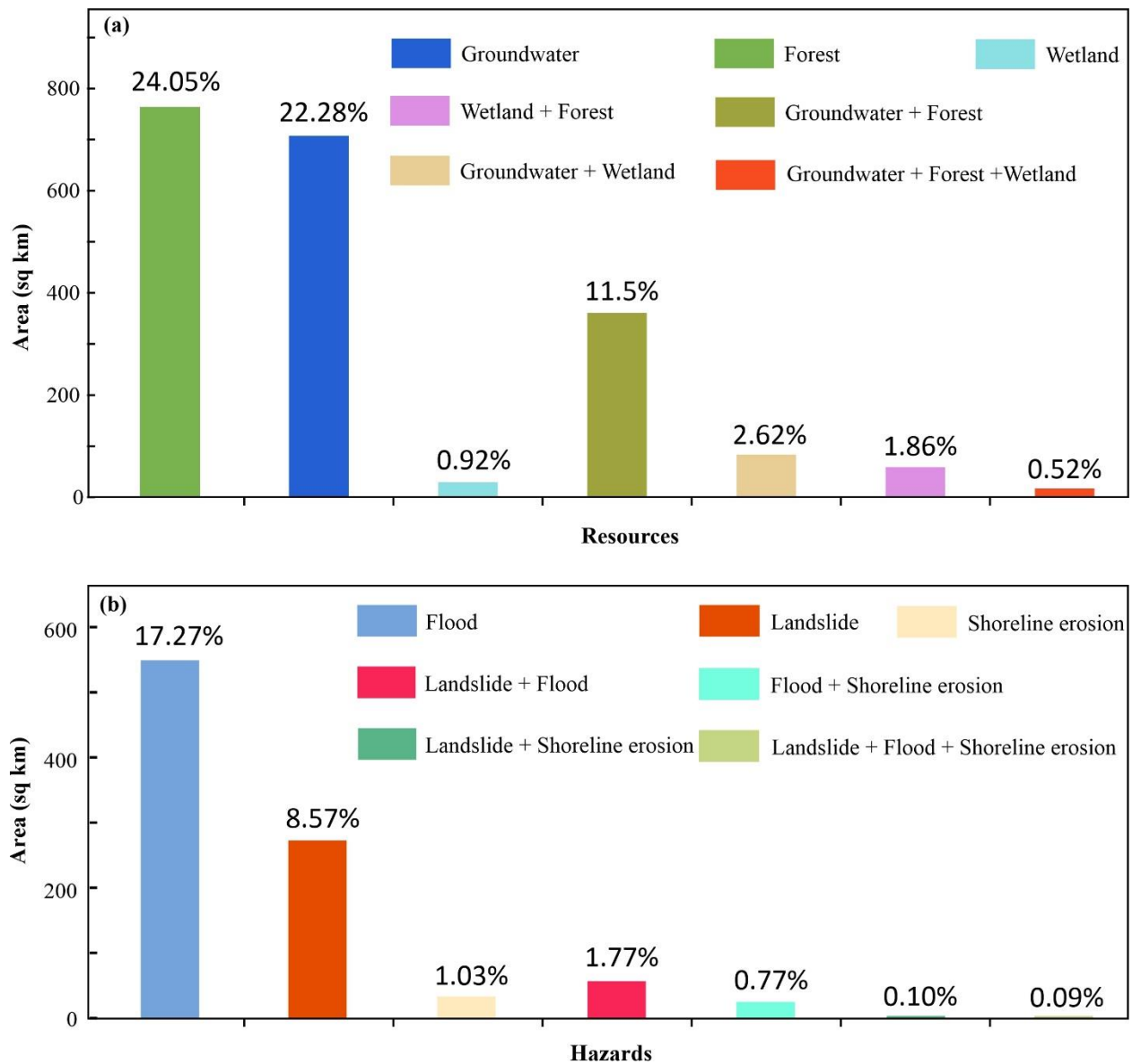


Fig. 10 Areal coverage of individual and multiple (a) resource and (b) hazard

5.2 Multi-hazard susceptibility map

The natural features of the earth undergo continuous change. Sometimes, the cumulative impacts of these changes are manifested both intensively and extensively. In view of this, the current work has made a modest attempt to map the major hazards prevailing in the study area including flood, shoreline erosion, and landslide. To understand the spatial pattern of the multi-hazards, the study

area has been categorized into eight different classes which include separate proportion of particular hazard and their combination (Fig. 9b). The major portion of the study area are susceptible to flood hazard (17.27%), followed by landslide (8.57%), flood with landslide (1.77%), shoreline erosion (1.03%), flood with shoreline erosion (0.77%), landslide with shoreline erosion (0.10%) overlap of threesome (0.09%) (Fig. 10b). This result has found that around 10% of the study region is vulnerable to the hazard. It is also observed that these hazards are mostly seen in coastal tract of the region of concern where human population is more compared to the interior region.

Table 8 Spatial distribution of multi-resource and multi-hazard with total area of each sub-class of categorical variables

Lithology	Area (%)			Geomorphology	Area (%)		
	Total	MR	MH		Total	MR	MH
Alluvium	3.4	2.4	7.3	DoPPc	49.7	64.0	65.8
Laterite	23.4	13.3	12.8	CoYCp	2.3	2.8	7.6
Sahyadri	20.1	17.9	9.8	SoMDp	18.7	21.0	14.0
Kalladgi	14.4	18.2	22.0	DoMDp	26	9.5	10.0
Dharwar	11.3	13.7	13.8	SoLDp	2.0	1.5	0.3
Archean	27.4	34.5	34.3	Others	1.3	1.2	2.3
Total	100	100	100		100	100	100

Soil	Area (%)			Land use	Area (%)		
	Total	MR	MH		Total	MR	MH
Ultic Typic Haplustalfs	25.3	27.2	34.0	Forest	15.1	21.8	21.5
Lithic Ustorthents	18.2	8.7	9.9	Agricultural land	26.1	22.4	23.7
Ultic Haplustalfs	10.9	15.1	17.3	Barren land	55	53.3	47.8
Typic Ustrophepts	45.6	49.0	38.8	Built-up	2.0	0.8	1.6
				Water-bodies	1.8	1.7	5.2
Total	100	100	100		100	100	100

The frequency, intensity of MH and spatial distribution of MR are highly influenced by sub-classes of each categorical variable such as lithology, geomorphology, soil, and land use. These sub-types of the categorical variables have been overlapped with MR and MH maps to compute the spatial distribution of each class (Table 8). It is better to analysis the area of each class with respect to the total area of that class. With regards to lithology, alluvium zone is more vulnerable whereas laterite and Sahyadri regions are less vulnerable as far as multi-hazard is concerned. In the case of geomorphology, it has contributed around 65% for both MR and MH. The area of coastal zone (CoYCp) is 3 times larger as compared to total area while DoMDp has 2.5 times less in relation to MH. In the soil types, Ultic Typic Haplustalfs and Ultic Haplustalfs are more prone to MH. On the other side, Lithic Ustorthents has covered half of its total area. Regarding land use classes, forest land with 15.1% of total area accounts for 21.8% and 21.5% for MR and MH, respectively while water bodies has contributed 5.2% in MH which is about 3 times than total areal coverage.

Table 9 Accuracies of natural resources and hazards mapping

Natural resources/hazards	Accuracy (%)
Groundwater potential mapping	94.00
Forest cover mapping	85.00
Probabilistic wetland mapping	95.30
Flood susceptibility mapping	92.69
Landslide susceptibility mapping	90.00

5.3 Accuracy assessment

In the present study, different statistical techniques have been employed to validate the model results. AUROC curve has been applied for the evaluation of RF model in groundwater potential mapping (GPM), probabilistic wetland mapping (PWM), landslide susceptibility mapping (LSM), and whereas for forest cover mapping (FSM) user and producer accuracies have been utilized. This research has only considered testing dataset for model evaluation as it defines the predictive capability of the model. The AUC values of the RF model are 94.0% in GWM, 95.30% in PWM, 85% in FCM, 92.69% in FSM, and 90% in LSM (Table 9). For forest mapping, producer and user accuracies are 90.0% and 80.0%, respectively. In the case of shoreline erosion mapping, accretion and erosion beaches were identified from the field and literature knowledge (SAC, 2014). The

AUC values of more than 90% in GPM, PWM, LSM, and FSM suggest the high precision mapping of these selected aspects in the region of interest

6. Discussion

Multi-resource and multi-hazard mapping in an area are the best practice for the sustainable development of natural resources and effective management of hazards. Researchers have prepared multi-hazard map on the basis of frequency and intensity of a particular hazard in a specific region (Kappes et al., 2012; Pourghasemi et al., 2020a; Pouyan et al., 2021; Rusk et al., 2022). It may not be possible to bring all the hazards and resources in MH and MR maps, respectively. For that, the present work has focused on the potential resources and prevailing hazards in the area of investigation. Many researchers have done mapping of particular resource (groundwater, wetland, mineral, forest) and hazard (floods, earthquake, gully erosion, shoreline erosion, landslide, and wildfire) in different parts of the world. It is quite difficult and time-consuming for policy-maker to analyze each resource and hazard map separately. On the contrary, MR and MH maps are easy to interpret for efficient management measures. Few researchers have attempted to prepare MH map (Pourghasemi et al., 2019; Rahmati et al., 2019; Nachappa et al., 2020; Pourghasemi et al., 2020b; Javidan et al., 2021; Pouyan et al., 2021; Rocchi et al., 2022). But, no study has been done so far combining MR and MH. For this, the present study has endeavoured to map both multi-resource and multi-hazard.

There are several machine learning models which have been used in numerous studies. Among the machine learning methods, RF model has many merits (multiple decision trees, not overfitted, no need of data scaling) and has proven to be more accurate in different resource and hazard mapping (Breiman, 2001; Abdi, 2020; Denize et al., 2019). For this reason, the present work has applied the RF model for MR and MH mapping. From the individual natural resource mapping, it has been observed that the high groundwater potential zone is concentrated in the south and southeastern parts of the study region covering around 37.9% of the total area. Prasad et al. (2020) found that 33.31% of areas with very high groundwater potential in the same region. In the case of forest mapping, though forest occupies about 37% of the total area, it mostly has grown in the break of slope of the Western Ghats and high elevation areas. Regarding wetland, it has been estimated that 6% of the entire study region is embodied with wetland primarily developed in the coastal sector. Increasing human encroachment on this coastal wetland is a pressing problem. On the other hand,

in regard to particular hazard mapping, flood has been a grave problem in comparison with landslide and shoreline erosion in the study region. It shows that 16.25% of the region is highly vulnerable to floods. From the study of Prasad et al. (2021a) on flood susceptibility mapping using ensemble of adabag and rotation forest models in similar study region, it has been found that 18.2% of the area is under highly vulnerable zone. Noticeably, the spatial pattern of the present research matches with the research of Prasad et al. (2021a). The heavy monsoon rain in the Western Ghats and coastal plain causes the rivers to spill over their banks resulting in severe flash floods (Prasad et al., 2021a). The basin areas of the major rivers including Gad, Karli, Sukhnadi, Tillari, and Waghota witness flood at least once a year. In 2005, flood inundated many parts of the research area and estimated damage was as much as about ninety million rupees (ED and MPCB, 2020). As a terrible consequence, the region suffers damage of crops, private and public properties loss, disruption of communication links, loss of lives, etc. Concerning landslide, 10.7% of the study area comes under the occurrence of landslides. It mainly occurs during monsoon rainfall due to slope failure. Landslide events are not spatially uniform across the area of investigation. It mostly takes place in the eastern part comprising the break of slope of the Western Ghats, steep slopes along the river banks and coastal tract. Regarding shoreline erosion, though it shares small portion (1.8%) in contrast to other two hazards, it is another growing problem in the study region. In comparison to the other studies on different beaches of India (Kuleli et al., 2011; Nandi et al., 2016; Mondal et al., 2017; Nassar et al., 2019), the erosion rate estimated in the present study is relatively less. However, some parts of the region have appeared as a matter of worry because of increasing rate of erosion. From the results, it can be envisioned that in future the coastal erosion may emerge as one of the major problems in this region. Pisolkar and Chaudhary, (2016) identified many research gaps and problems including non-availability of land cover/land use map, cadastral map in digital format, shoreline change map, etc. for sustainable management along this coast. Several parts of the region which are exposed to higher erosion, have been marked in the current research. Requisite actions should be taken for the restoration of these areas (Fig. 8c). The pace and patterns of erosion and accretion vary spatially as well as temporally. The northern portion of the region has withstood erosion and remained stable by virtue of rocky beaches. In contrast, the southern part near the Redi and Vengurla contains highly erodible beaches. It is also noted that the rate of erosion in some of the transects is visibly high but with the lapse of time it has decreased. It may be because of the changes in physical parameters. The used Landsat images may generate

some error due to medium spatial resolution (30 m). To escape this shortcoming, in future, this research can be continued with available high resolution optical and microwave datasets for better precision. Multi-hazard and multi-resource environments differ from individual resource and hazard environments by the mechanistic connections between the controlling factors (Rusk et al., 2022). The preparation of MH and MR maps has greater favorable significance in comparison to the separate resource and hazard mapping, for instance, i) it gives a comprehensive picture at a time as all the information is contained in a single map e.g., all the major hazards in multi-hazard map, ii) it demarcates the most vulnerable region where multiple hazards occurred, iii) it enables planner for priority based planning and management. On the other side, the main disadvantages of these mapping are i) it needs different types and huge amount of datasets, ii) it is complex to understand. In this research, after preparing the individual resource and hazard maps, multi-resource and multi-hazard maps have been generated by overlapping the selected resources and hazards. From the results of MR map, it has been seen that forest and high groundwater potential zone individually spread over most of the region whereas wetland region is significantly less. As per the combination of resources, forest and groundwater have jointly shared the highest area (11.5%) in compared to the other combinations. It is also observed that the combined area of wetland with forest and groundwater is not larger than wetland alone. It is because of the fact that the high groundwater and forest cover in low altitude regions are favourable for wetland development. On the other hand, the results of multi-hazard map have manifested that the area under flood is maximum than other selected hazards. These hazards rarely tend to overlap. This is mainly due to the different mechanism and lesser interaction of the respective responsible factors for instance; flood mainly happens in lower river basin and estuary regions whereas landslide mostly occurs at higher altitude. In the Hindu Kush Himalaya, Rusk et al. (2022) has prepared multi-hazard map (including landslide, wildfire, and flood) and pointed out that urbanization processes are likely to increase the exposure to multi-hazard risk. From the results of spatial distribution of the categorical variables related to MR and MH, it is revealed that the role of each sub-classes of lithology, geomorphology, soil and land use are crucial in occurrence of resources and hazards under study. But, it is needed to compare with high-resolution datasets for more information.

7. Conclusions and management recommendations

For the first time, this study has attempted to deal with both MR and MH in a single platform using geospatial and machine learning model. The applied RF model has performed outstanding with more than 90% accuracy in preparation of individual maps. With regard to the resources, very high groundwater potential regions and probabilistic wetland areas have mostly covered east, southeast, south, and coastal parts of the study region. Out of the three hazards, flood has caused the most concern since the final map readily exhibits that major parts of the study region are flood affected and it is ascertained that management deserves the first priority. The MR and MH maps of the study region may be utilized for conservation and sustainable development of these resources, and also for planning and management in hazards susceptible regions. Moreover, it is also suggested that similar type of works are required in hazard affected areas around the world to minimize loss and in naturally resource scarce region to increase resource potentiality. The present work offers a robust assessment of MR and MH, albeit there are some limitations in terms of method and datasets. The following limitations and future scopes of this research are i) it has relied on only random forest model but other machine learning or ensemble models can be implemented for comparative performance assessment and high precision, ii) quality of the groundwater resource can be assessed and monitored. iii) in addition, based on the MR and MH maps of the study area along with other collateral information, land suitability and land capability maps can be prepared for better policy making. Despite of these limitations, the outcome of this work will prove to be highly significant for planning and development of the area as well as the methodology of the research will be a roadmap for future work in other geographic regions.

The following measures are proposed for the multiple resource and hazard management in the study area, i) Preservation and creation of wetlands in the coastal sector can reduce sediment load deposited by both marine and fluvial processes in the monsoon season, ii) Construction of flood embankments, ring circuit bunds, raised platforms and river channel diversion around the flood affected towns and villages with proper planning can decrease the severity of floods, iii) The government ought to make policy and discourage new development in landslide, flood, and shoreline erosion prone areas, iv) Beach nourishment and building artificial sand dunes are methods to protect the shoreline movement and backshore property. Besides, it also possesses many advantages like recreation site (tourism) for its aesthetic appeal and lesser cost of management in comparison with the structural method.

Declaration of statement

All authors are declared no actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations.

Data availability statement

The dataset is included in the manuscript.

Acknowledgements

We acknowledge the financial support from the University Grant Commission (3160 NET-June 2015). We also acknowledge the NASA, ESA, and ISRO for providing the satellite data. The authors are thankful to the Director CSIR-NIO for time to support. The authors are also grateful to Editor and anonymous reviewers for their critical comments and constructive suggestions to improve the manuscript. Field assistance from the survey team members is thankfully acknowledged. The NIO contribution number is.....

References

- Abdi, A.M., 2020. Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience & Remote Sensing*, 57(1), 1-20. <https://doi.org/10.1080/15481603.2019.1650447>
- Adam, E., Mutanga, O., Odindi, J. and Abdel-Rahman, E.M., 2014. Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers. *International Journal of Remote Sensing*, 35(10), 3440-3458. <https://doi.org/10.1080/01431161.2014.903435>
- Alin, A., 2010. Multicollinearity. *Wiley interdisciplinary reviews: computational statistics*, 2(3), pp.370-374.
- Amiri, M., Pourghasemi, H.R., Ghanbarian, G.A. and Afzali, S.F., 2019. Assessment of the importance of gully erosion effective factors using Boruta algorithm and its spatial modeling and mapping using three machine learning algorithms. *Geoderma*, 340, 55-69. <https://doi.org/10.1016/j.geoderma.2018.12.042>
- Angeli, S. D., Malamud, B.D., Rossi, L., Taylor, F.E., Trasforini, E. and Rudari, R., 2022. A multi-hazard framework for spatial-temporal impact analysis. *International Journal of Disaster Risk Reduction*, 73, 102829. <https://doi.org/10.1016/j.ijdrr.2022.102829>
- Arabameri, A., Saha, S., Roy, J., Tiefenbacher, J.P., Cerda, A., Biggs, T., Pradhan, B., Ngo, P.T.T. and Collins, A.L., 2020. A novel ensemble computational intelligence approach for the spatial prediction of land subsidence susceptibility. *Science of The Total Environment*, p.138595. <https://doi.org/10.1016/j.scitotenv.2020.138595>

- Arora, A., Arabameri, A., Pandey, M., Siddiqui, M.A., Shukla, U.K., Bui, D.T., Mishra, V.N. and Bhardwaj, A., 2021. Optimization of state-of-the-art fuzzy-metaheuristic ANFIS-based machine learning models for flood susceptibility prediction mapping in the Middle Ganga Plain, India. *Science of the Total Environment*, 750, 141565. <https://doi.org/10.1016/j.scitotenv.2020.141565>
- Arvin, M., Beiki, P., Hejazi, S.J., Sharifi, A. and Atashafrooz, N., 2023. Assessment of infrastructure resilience in multi-hazard regions: A case study of Khuzestan Province. *International Journal of Disaster Risk Reduction*, 103601. <https://doi.org/10.1016/j.ijdr.2023.103601>
- Berhane, T.M., Lane, C.R., Wu, Q., Autrey, B.C., Anenkhonov, O.A., Chepinoga, V.V. and Liu, H., 2018. Decision-tree, rule-based, and random forest classification of high-resolution multispectral imagery for wetland mapping and inventory. *Remote sensing*, 10(4), 580. <https://doi.org/10.3390/rs10040580>
- Breiman, L., 2001. Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Castillo, C. and Gómez, J.A., 2016. A century of gully erosion research: Urgency, complexity and study approaches. *Earth-Science Reviews*, 160, 300-319. <https://doi.org/10.1016/j.earscirev.2016.07.009>
- Chapi, K., Singh, V.P., Shirzadi, A., Shahabi, H., Bui, D.T., Pham, B.T. and Khosravi, K., 2017. A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environmental modelling & software*, 95, 229-245. <https://doi.org/10.1016/j.envsoft.2017.06.012>
- Chen, W., Chen, Y., Tsangaratos, P., Ilia, I. and Wang, X., 2020. Combining evolutionary algorithms and machine learning models in landslide susceptibility assessments. *Remote Sensing*, 12(23), 3854. <https://doi.org/10.3390/rs12233854>
- Chen, W., Li, H., Hou, E., Wang, S., Wang, G., Panahi, M., Li, T., Peng, T., Guo, C., Niu, C. and Xiao, L., 2018. GIS-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models. *Science of the Total Environment*, 634, 853-867. <https://doi.org/10.1016/j.scitotenv.2018.04.055>
- Cracknell, M.J. and Reading, A.M., 2014. Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Computers & Geosciences*, 63, 22-33. <https://doi.org/10.1016/j.cageo.2013.10.008>
- Cuesta, M.J.D., (2013). Susceptibility, p. 988. In: Bobrowsky P.T. (eds) *Encyclopedia of Natural Hazards*. Encyclopedia of Earth Sciences Series. Springer, Dordrecht, 1135 pp.
- Denize, J., Hubert-Moy, L., Betbeder, J., Corgne, S., Baudry, J. and Pottier, E., 2019. Evaluation of using sentinel-1 and-2 time-series to identify winter land use in agricultural landscapes. *Remote Sensing*, 11(1), 37. <https://doi.org/10.3390/rs11010037>

- De Rosa, P., Fredduzzi, A. and Cencetti, C., 2019. Stream power determination in gis: an index to evaluate the most 'Sensitive' points of a river. *Water*, 11(6), 1145. <https://doi.org/10.3390/w11061145>
- Dewan, A., 2013. *Floods in a megacity: geospatial techniques in assessing hazards, risk and vulnerability*. Dordrecht: Springer. <https://doi.org/10.1007/978-94-007-5875-9>
- DDMA (District Disaster Management Authority, Sindhudurg) 2014, District Disaster Management Plan.
- ED and MPCB (Environment Department and Maharashtra Pollution Control Board, Maharashtra), 2020. District Environment Plan.
- ESRI. 2016. "ArcGIS for Desktop." <http://desktop.arcgis.com>
- Feizizadeh, B., Omarzadeh, D., Kazemi Garajeh, M., Lakes, T. and Blaschke, T., 2021. Machine learning data-driven approaches for land use/cover mapping and trend analysis using Google Earth Engine. *Journal of Environmental Planning and Management*, 1-33. <https://doi.org/10.1080/09640568.2021.2001317>
- Hird, J.N., DeLancey, E.R., McDermid, G.J. and 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. <https://doi.org/10.3390/rs9121315>
- Hosmer, D.W., 2000. *Wiley series in probability and statistics, Chap. 2. Multiple logistic regression. Applied logistic regression*, Wiley, pp.31-46.
- Huang, F., Pan, L., Fan, X., Jiang, S.H., Huang, J. and Zhou, C., 2022. The uncertainty of landslide susceptibility prediction modeling: suitability of linear conditioning factors. *Bulletin of Engineering Geology and the Environment*, 81(5), 1-19. <https://doi.org/10.1007/s10064-022-02672-5>
- Javidan, N., Kavian, A., Pourghasemi, H.R., Conoscenti, C., Jafarian, Z. and Rodrigo-Comino, J., 2021. Evaluation of multi-hazard map produced using MaxEnt machine learning technique. *Scientific reports*, 11(1), 1-20. <https://doi.org/10.1038/s41598-021-85862-7>
- Kappes, M.S., Keiler, M., von Elverfeldt, K. and Glade, T., 2012. Challenges of analyzing multi-hazard risk: a review. *Natural hazards*, 64(2), 1925-1958. <https://doi.org/10.1007/s11069-012-0294-2>
- Kuleli, T., Guneroglu, A., Karsli, F. and Dihkan, M., 2011. Automatic detection of shoreline change on coastal Ramsar wetlands of Turkey. *Ocean Engineering*, 38(10), 1141-1149. <https://doi.org/10.1016/j.oceaneng.2011.05.006>
- Kulimushi, L.C., Bashagaluke, J.B., Prasad, P., Heri-Kazi, A.B., Kushwaha, N.L., Masroor, M., Mohammed, S. 2023. Soil erosion susceptibility mapping using ensemble machine learning models: A case study of upper Congo river sub-basin. *Catena* 222, 106858 <https://doi.org/10.1016/j.catena.2022.106858>

- Mahapatra, M., Ratheesh, R. and Rajawat, A.S., 2014. Shoreline change analysis along the coast of South Gujarat, India, using digital shoreline analysis system. *Journal of the Indian Society of Remote Sensing*, 42(4), 869-876. <https://doi.org/10.1007/s12524-013-0334-8>
- Maiti, S. and Bhattacharya, A.K., 2009. Shoreline change analysis and its application to prediction: A remote sensing and statistics based approach. *Marine Geology*, 257(1-4), 11-23. <https://doi.org/10.1016/j.margeo.2008.10.006>
- Marker, B.R., (2013). Hazard and risk mapping, p. 426. In: Bobrowsky P.T. (eds) *Encyclopedia of Natural Hazards*. Encyclopedia of Earth Sciences Series. Springer, Dordrecht, 1135 pp.
- Maxwell, A.E., Warner, T.A. and Strager, M.P., 2016. Predicting palustrine wetland probability using random forest machine learning and digital elevation data-derived terrain variables. *Photogrammetric Engineering & Remote Sensing*, 82(6), 437-447. <https://doi.org/10.14358/PERS.82.6.437>
- Mondal, I., Bandyopadhyay, J. and Dhara, S., 2017. Detecting shoreline changing trends using principle component analysis in Sagar Island, West Bengal, India. *Spatial Information Research*, 25(1), 67-73. DOI 10.1007/s41324-016-0076-0
- Moore, I.D., Grayson, R.B., and Ladson, A.R., 1991. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, 5(1), 3-30. <https://doi.org/10.1002/hyp.3360050103>
- Mukhopadhyay, A., Mukherjee, S., Mukherjee, S., Ghosh, S., Hazra, S. and Mitra, D., 2012. Automatic shoreline detection and future prediction: A case study on Puri Coast, Bay of Bengal, India. *European Journal of Remote Sensing*, 45(1), 201-213. <https://doi.org/10.5721/EuJRS20124519>
- Nachappa, T.G., Ghorbanzadeh, O., Gholamnia, K. and Blaschke, T., 2020. Multi-hazard exposure mapping using machine learning for the State of Salzburg, Austria. *Remote Sensing*, 12(17), 2757. <https://doi.org/10.3390/rs12172757>
- Naghibi, S.A., Dolatkordestani, M., Rezaei, A., Amouzegari, P., Heravi, M.T., Kalantar, B. and Pradhan, B., 2019. Application of rotation forest with decision trees as base classifier and a novel ensemble model in spatial modeling of groundwater potential. *Environmental Monitoring and Assessment*, 191(4), 248. <https://doi.org/10.1007/s10661-019-7362-y>
- Nandi, S., Ghosh, M., Kundu, A., Dutta, D. and Baksi, M., 2016. Shoreline shifting and its prediction using remote sensing and GIS techniques: a case study of Sagar Island, West Bengal (India). *Journal of coastal conservation*, 20(1), 61-80. DOI 10.1007/s11852-015-0418-4
- Nassar, K., Mahmood, W.E., Fath, H., Masria, A., Nadaoka, K. and Negm, A., 2019. Shoreline change detection using DSAS technique: Case of North Sinai coast, Egypt. *Marine Georesources & Geotechnology*, 37(1), 81-95. <https://doi.org/10.1080/1064119X.2018.1448912>

- NBSSLUP (National Bureau of Soil Survey and Land use planning), 1996. Soil Map of Maharashtra.
- Nematzadeh, Z., Ibrahim, R. and Selamat, A., 2015, May. Comparative studies on breast cancer classifications with k-fold cross validations using machine learning techniques. In 2015 10th Asian control conference (ASCC) (pp. 1-6). IEEE. [10.1109/ASCC.2015.7244654](https://doi.org/10.1109/ASCC.2015.7244654)
- Nohani, E., Moharrami, M., Sharafi, S., Khosravi, K., Pradhan, B., Pham, B.T., Lee, S. and M. Melesse, A., 2019. Landslide susceptibility mapping using different GIS-based bivariate models. *Water*, 11(7), p.1402. <https://doi.org/10.3390/w11071402>
- Noi, P. T. and Kappas, M., 2018. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors*, 18(1), 18. <https://doi.org/10.3390/s18010018>
- Pham, B.T., Jaafari, A., Avand, M., Al-Ansari, N., Dinh Du, T., Yen, H.P.H., Phong, T.V., Nguyen, D.H., Le, H.V., Mafi-Gholami, D. and Prakash, I., 2020. Performance evaluation of machine learning methods for forest fire modeling and prediction. *Symmetry*, 12(6), 1022. <https://doi.org/10.3390/sym12061022>
- Pisolkar, Y. and Chaudhary, N., 2016. The Problems of Coastal Tourism, Environment and Local Sustainable Development along Sindhudurg District, Coastal Maharashtra India. *Annual Research Journal of SCMS, Pune*, 4.
- Pourghasemi, H.R., Gayen, A., Edalat, M., Zarafshar, M. and Tiefenbacher, J.P., 2020a. Is multi-hazard mapping effective in assessing natural hazards and integrated watershed management?. *Geoscience Frontiers*, 11(4), 1203-1217. <https://doi.org/10.1016/j.gsf.2019.10.008>
- Pourghasemi, H.R., Kariminejad, N., Amiri, M., Edalat, M., Zarafshar, M., Blaschke, T. and Cerda, A., 2020b. Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. *Scientific reports*, 10(1), 1-11. <https://doi.org/10.1038/s41598-020-60191-3>
- Pourghasemi, H.R., Sadhasivam, N., Kariminejad, N. and Collins, A.L., 2020c. Gully erosion spatial modelling: Role of machine learning algorithms in selection of the best controlling factors and modelling process. *Geoscience Frontiers*, <https://doi.org/10.1016/j.gsf.2020.03.005>
- Pourghasemi, H.R., Gayen, A., Panahi, M., Rezaie, F. and Blaschke, T., 2019. Multi-hazard probability assessment and mapping in Iran. *Science of the total environment*, 692, 556-571. <https://doi.org/10.1016/j.scitotenv.2019.07.203>
- Pouyan, S., Pourghasemi, H.R., Bordbar, M., Rahmanian, S. and Clague, J.J., 2021. A multi-hazard map-based flooding, gully erosion, forest fires, and earthquakes in Iran. *Scientific Reports*, 11(1), 1-19. <https://doi.org/10.1038/s41598-021-94266-6>

- Prasad, P., Loveson, V.J., Chandra, P. and Kotha, M., 2022. Evaluation and comparison of the earth observing sensors in land cover/land use studies using machine learning algorithms. *Ecological Informatics*, 68, 101522. <https://doi.org/10.1016/j.ecoinf.2021.101522>
- Prasad, P., Loveson, V.J., Das, B. and Kotha, M., 2021a. Novel ensemble machine learning models in flood susceptibility mapping. *Geocarto International*, 1-23. <https://doi.org/10.1080/10106049.2021.1892209>
- Prasad, P., Loveson, V.J., Das, S. and Chandra, P., 2021b. Artificial intelligence approaches for spatial prediction of landslides in mountainous regions of western India. *Environmental Earth Sciences*, 80(21), 1-20. <https://doi.org/10.1007/s12665-021-10033-w>
- Prasad, P., Loveson, V.J., Kotha, M., and Yadav, R., 2020. Application of machine learning techniques in groundwater potential mapping along the west coast of India. *GIScience & Remote Sensing*, 57(6), 735-752. <https://doi.org/10.1080/15481603.2020.1794104>
- Probst, P., Wright, M.N. and Boulesteix, A.L., 2019. Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: data mining and knowledge discovery*, 9(3), e1301. <https://doi.org/10.1002/widm.1301>
- Rahmati, O., Yousefi, S., Kalantari, Z., Uuemaa, E., Teimurian, T., Keesstra, S., Pham, T.D. and Tien Bui, D., 2019. Multi-hazard exposure mapping using machine learning techniques: A case study from Iran. *Remote Sensing*, 11(16), 1943. <https://doi.org/10.3390/rs11161943>
- Rocchi, A., Chiozzi, A., Nale, M., Nikolic, Z., Riguzzi, F., Mantovan, L., Gilli, A. and Benvenuti, E., 2022. A Machine Learning Framework for Multi-Hazard Risk Assessment at the Regional Scale in Earthquake and Flood-Prone Areas. *Applied Sciences*, 12(2), 583. <https://doi.org/10.3390/app12020583>
- Rusk, J., Maharjan, A., Tiwari, P., Chen, T.H.K., Shneiderman, S., Turin, M. and Seto, K.C., 2022. Multi-hazard susceptibility and exposure assessment of the Hindu Kush Himalaya. *Science of the total environment*, 804, 150039. <https://doi.org/10.1016/j.scitotenv.2021.150039>
- SAC (Space Application Centre), 2014 Shoreline change atlas of the Indian coast (Maharashtra and Goa) Volume-2. <https://vedas.sac.gov.in/vedas/downloads/atlas/Seashore/Vol-2-Full.pdf>
- Saha, A., Pal, S.C., Santosh, M., Janizadeh, S., Chowdhuri, I., Norouzi, A., Roy, P. and Chakraborty, R., 2021. Modelling multi-hazard threats to cultural heritage sites and environmental sustainability: The present and future scenarios. *Journal of Cleaner Production*, 320, 128713. <https://doi.org/10.1016/j.jclepro.2021.128713>
- Steinhausen, M.J., Wagner, P.D., Narasimhan, B. and Waske, B., 2018. Combining Sentinel-1 and Sentinel-2 data for improved land use and land cover mapping of monsoon regions. *International journal of applied earth observation and geoinformation*, 73, 595-604. <https://doi.org/10.1016/j.jag.2018.08.011>

- Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y.A. and Rahman, A., 2020. Land-use land-cover classification by machine learning classifiers for satellite observations—A review. *Remote Sensing*, 12(7), 1135. <https://doi.org/10.3390/rs12071135>
- UNEP, 1992. Agenda 21.Tech. rep., United Nations Environment Programme. http://www.un.org/esa/dsd/agenda21/res_agenda21_07.shtml,
- UNISDR, C., 2015. The human cost of natural disasters: A global perspective. <http://repo.floodalliance.net/jspui/handle/44111/1165>
- Wang, Y., Fang, Z. and Hong, H., 2019. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. *Science of the total environment*, 666, pp.975-993. <https://doi.org/10.1016/j.scitotenv.2019.02.263>
- Wen, L. and Hughes, M., 2020. Coastal wetland mapping using ensemble learning algorithms: A comparative study of bagging, boosting and stacking techniques. *Remote Sensing*, 12(10), 1683. <https://doi.org/10.3390/rs12101683>
- Yamano, H., Shimazaki, H., Matsunaga, T., Ishoda, A., McClennen, C., Yokoki, H., Fujita, K., Osawa, Y., Kayanne, H., 2006. Evaluation of various satellite sensors for waterline extraction in a coral reef environment: Majuro Atoll, Marshall Islands. *Geomorphology* 82, 398–411. <https://doi.org/10.1016/j.geomorph.2006.06.003>
- Yasir, M., Sheng, H., Fan, H., Nazir, S., Niang, A.J., Salauddin, M. and Khan, S., 2020. Automatic coastline extraction and changes analysis using remote sensing and GIS technology. *IEEE Access*, 8, 180156-180170. [10.1109/ACCESS.2020.3027881](https://doi.org/10.1109/ACCESS.2020.3027881)