Artificial Intelligence in Geosciences 3 (2022) 179–191

Contents lists available at ScienceDirect

Artificial Intelligence in Geosciences

journal homepage: www.keaipublishing.com/en/journals/artificial-intelligence-in-geosciences

Estimation of the effectiveness of multi-criteria decision analysis and machine learning approaches for agricultural land capability in Gangarampur Subdivision, Eastern India

Sunil Saha, Prolay Mondal *

Department of Geography, Raiganj University, Raiganj, West Bengal, 733134, India

ARTICLE INFO

Keywords:
Analytic Hierarchy Process
Fuzzy Complex Proportional Assessment
Receiver Operating Characteristic
Gangarampur subdivision
Soil Capability

ABSTRACT

Land suitability analysis (LSA) is an evaluation method that measures the degree to which land is suitable for certain land use. The primary aims of this study are to identify potentially viable agricultural land in the Gangarampur subdivision (West Bengal) using Multiple Criteria Decision Making (MCDM) and machine learning procedures and to evaluate the efficacy of the employed methodologies. The Analytic Hierarchy Process (AHP) model was used to assign relative weights to the fifteen various criteria in this suitability analysis, and then the Fuzzy Complex Proportional Assessment (FCOPRAS) model was applied using the AHP's normalised pairwise comparison matrix, whereas the Waikato Environment for Knowledge Analysis (Weka) Software was used to apply machine learning algorithms to the field data. The Random Forest (RF) model, on the other hand, is a better fit for the locational study of soil potential. According to the RF findings, areas of 14.67 per cent (15368.46 ha) are excellent (ZONE V) for growing crops, approximately 22.30 per cent (23367.9 ha) are highly suitable (ZONE IV), and 23.63 per cent (24762.12 ha) are moderately suitable (ZONE III) for cultivation, respectively. The numbers for FCOPRAS are roughly 15.39% (16130.52 ha), 22.54% (23620.65 ha), and 19.79% (20733.26 ha). The Receiver Operating Characteristic (ROC) curve and accuracy measurements of the results indicate the high accuracy of the applied models, with Random Forest and FCOPRAS being the most popular and effective techniques. This study will make an important contribution to evaluations of soil fertility and site suitability. This will help local government officials, academics, and farmers scientifically use the land.

1. Introduction

Agriculture is the world's most primitive economic activity (Prakash, 2003), occupying one-third of the world's entire land surface (Grigg, 1995). Following industrialization (Davis and Langham, 1995), the world's population increased positively and food demand increased rapidly (Lambin and Meyfroidt, 2011; Scherer et al., 2018; Alexandratos and Bruinsma, 2012), resulting in the intensification and extensification (Yalew et al., 2016; Tscharntke et al., 2012) of the primary sector to sustain the supply and demand chain (Naik and Suresh, 2018). To sustain an increasing human population, the only approach is to boost agricultural output (Ibrahim et al., 2016) without endangering future generations' requirements (Feizizadeh et al., 2014; Turgut et al., 2013). The Dakshin Dinajpur district's economy is primarily based on the primary economic sector, more precisely on agriculture. According to census (2011) data, a total of 170,682 cultivators engage in agriculture cultivation, with 156,797 males and 13,885 women; and 174,690 individuals working as labourers on agricultural land, with 129,402 men and 45,288 women. Paddy, wheat, barley, and maize are the most widely farmed food grains. Agriculture's development is critical to the socio-economic development of the study area. Agriculture development has a huge impact on people's socioeconomic conditions, and it is critical to building optimal agriculture practices to enable people to self-sustain (Pramanik, 2016). Population pressure on agricultural land and strong demand (Cowie et al., 2018) for foods necessitate speedy production or high-yield agriculture. Due to the excessive use of chemical fertilizers and pesticides, insufficient irrigation infrastructure, insufficient transportation systems, depleted soil fertility, soil erosion, and other factors, agricultural production has fallen quickly (De la Rosa and Sobral, 2008).

Land suitability analysis (LSA) is a method of figuring out how well land can be used for a certain purpose (Mistri and Sengupta., 2020). The

* Corresponding author.
E-mail addresses: sahasunil2507@gmail.com (S. Saha), mon.prolay@gmail.com (P. Mondal).

https://doi.org/10.1016/j.aiig.2022.12.003
Received 12 September 2022; Received in revised form 21 December 2022; Accepted 22 December 2022
Available online 29 December 2022

land/soil suitability evaluation is required to decide which crop is most suited to a place with limited resources (Jamil et al., 2018). Land is a unique and significant element on the earth's surface (Rossiter, 1996), and every living entity interacts with it directly or indirectly. Potential Suitability Assessment of Land (PSAL) determines the land's suitability for a specified use (Bandyopadhyay et al., 2009; Hayashi, 2000). The basic purpose of "PSAL" is to ascertain people's satisfaction with a particular land use form (Hopkins, 1977), and it can assist policymakers in developing strategies to increase agricultural productivity by identifying potentially suitable land for cultivation (FAO, 1976; Malczewski, 2004). MCDM approaches are commonly used to estimate the potential and inherent properties of soil (Mendas and Delali, 2012) and to determine the optimal Zones for cultivation (Steiner et al., 2000). Mosadeghi et al. (2015) say that MCDM uses a lot of variables to predict how land will be used in the future.

The integration of remote sensing, GIS, and MCDM methodologies will always be promising for identifying new suitable places (Halder et al., 2020). Mokarram and Aminzadeh (2010) suggest that Geographic Information System (GIS) is a more adaptable system for acquiring, investigating, and evaluating data and outcomes. Researchers can fix small mistakes more quickly and get more accurate results by working together (Saaty, 1980; Duc, 2006).

There has been a significant amount of research on site suitability studies utilising MCDM and ML techniques, including Site Suitability Assessment for Rice Cultivation (Nyeko, 2012; Kihoro et al., 2013), Site suitability for agriculture (Zaredar and Jafari, 2010), for land management purposes (Rojas and Loubier, 2017) for land-use suitability (Joerin et al., 2001; Romeijn et al., 2016) for Land suitability analysis for various crops (Jamil et al., 2018), for avulsion potential Zone construction (Pal and Sarkar, 2018), for the recognition of touristic potential Zones (Ebrahimi et al., 2019), site suitability for conservation of water (Srdjevic and Medeiros, 2007) and water management (Badhe et al., 2019), Predicting the Soil Suitability (Jayaraman et al., 2021); Prediction of Cardiac Disease (Princy et al., 2020); Mapping the risk terrain for crime (Wheeler and Steenbeek, 2021) and executing surface and pressurized irrigation system (Azad et al., 2019), etc.

In this study, the FCOPRAS (Fuzzy Complex Proportional Assessment), AHP, and machine learning approaches, namely Random Forest and Multilayer Perceptron have been used to discover the potentiality of the soil for agriculture, and ROC and Accuracy techniques have been also applied to estimate the proficiency of the applied models and algorithms. Before arriving at the outcome, the weight of the alternatives for each variable was calculated (Malczewski and Boroushaki, 2008; Murayam and Bunruamkaew, 2011) using AHP. The opinions of three experts were considered in the whole process of the FCOPRAS. To aggregate the performing variables, the WLCM (Weighted Linear Combination Method) was used.

The land suitability assessment approach is a method for analysing land's potential to determine the best location for growing various crops. In planning and managing land resources, one of the most advantageous uses of the Geographic Information System (GIS) is the study of site suitability and preparation of land use maps. GIS and Multi-Criteria Decision Analysis (MCDA) can be used well enough for planning and managing how agricultural land is used (Adeniyi, 1993; Shearer and Xiang, 2009). GIS makes it easy to store, process, and analyse complex raw data at plenty of different levels. Combining MCDM and Machine Learning (ML) with Geographic Information Systems (GIS) is an excellent technique (Sánchez-Lozano et al., 2013) to make good judgments in complex and dynamic agricultural systems.

Agriculture is the economic backbone of West Bengal, employing about 60–70 per cent of the workforce and generating nearly 19.9 per cent of the state's GDP in 2020–21. In terms of food grain production, the state has achieved self-sufficiency. Only 48% of the Dakshin Dinajpur district's (West Bengal, India) geographic area is inside the research area. Because this district lacks an industrial sector, its economy is primarily agrarian. The overall quantity of land in the Gangrampur subdivision is not completely fertile; just a small portion is naturally acidic. Because of this, farming is much more expensive there. Another hand, this area is vulnerable to flooding or inundation in considerable portions. The fact that 80 per cent of the land is fragmented and belongs to marginal landholding categories is a major impediment to the implementation of an enhanced and mechanised farming system (Sarkar et al., 2021). To handle the difficulties of maintaining an expanding population, sustaining agriculture, and preserving the environment, it is vital to estimate the agricultural Land Capability to make long-term improvements as the Gangarampur Subdivision of Dakshin Dinajpur is significantly harming agricultural efficiencies due to insufficient soil nutrients (Halder et al., 2020). The cognitive goal of the present study is to estimate the efficiency of applied models and the utilitarian goal is to uncover the potentially suitable sites for agriculture in the Gangarampur subdivision.

2. Description of the study area

The Gangarampur Subdivision (Fig. 1) is located in the mid-eastern portion of West Bengal. Agricultural significance is enormous in the study area. Gangarampur Subdivision has historically been a major producer of paddy, wheat, pulses, and vegetables such as potatoes, onions, and green vegetables. Farmers were encouraged to produce cash crops or crops that allowed them to make more money. The importance of market-oriented crops has had a significant impact on grain production. Massive urban demand has resulted in significant growth in the production of green vegetables and seasonal fruits. In the last few decades, there have been more reports of changes in agricultural output in areas near rail lines or the National Highway. However, agricultural crop production is still important in the interior.

The geographical extension of the study area is a range between 25°17′04″ North to 25°36′12″ North and 88°10′05″ East to 88°42′01″ East with an elevation of 28–32 m above Mean Sea Level (Fig. 1). The district is bounded by Bangladesh on the north, the district of Malda on the south, the Balurghat Subdivision on the east, and the Raiganj Subdivision on the west. The study area spans a total of 104,790 ha of geographical area.

3. Database and adopted methods

3.1. Data sources

About 95 articles, 2 theses, and several government reports are reviewed for accurate research. The articles are reviewed based on their originality in methods. From the literature review, it is found that there is no such type of intensive research work in the same study area though some of the researchers are tried to evaluate the agricultural suitable zone at the district level. As the majority of the population of the study area is dependent on agriculture and related economic activities, it is essential to evaluate the proper agricultural area for the overall development of the study area.

Several sixteen variables, namely nitrogen (N), copper (Cu), manganese (Mn), zinc (Zn), sulphur (S), boron (B), phosphate, potassium (K), (Sarkar et al., 2021), electrical conductivity (EC) (Richards, 1954), Cation Exchange Capacity (CEC), soil PH, bulk density (Saha et al., 2021), soil moisture (Pramanik, 2016), soil organic carbon, soil depth, and soil texture are selected for estimation of potential site suitability analysis. After the laboratory testing of the soil samples (40 Spots) which have been collected in 2021, the Inverse Distance Weighted (IDW) interpolation technique has been used to generate the spatial distribution maps of selected variables in ArcMap 10.5. A brief account of the data source is provided in Table 1.

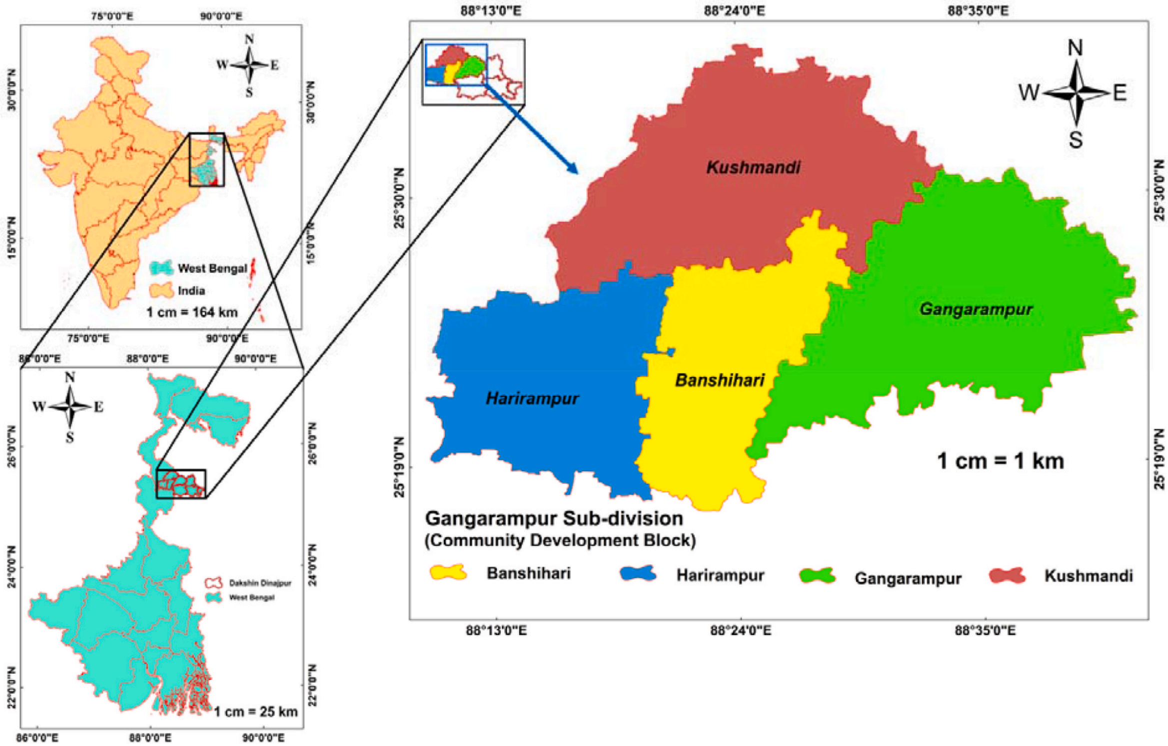

Fig. 1. Geographical location of the study area.

Table 1
Performing Variables, Sources, and their resolution.

Name of the Variable	Source	Website	Resolution (m)
Nitrogen (*N*)	Primary Data were collected from 40 spots from the Study Area (Laboratory Test)	Field Survey	30 × 30
Phosphate (PO_4^{3-})			
Potassium (*K*)			30 × 30
Zinc (*Zn*)			30 × 30
Sulphur (*S*)			30 × 30
Copper (*Cu*)			30 × 30
Boron (*B*)			30 × 30
Manganese (*Mn*)			30 × 30
Soil Depth (*Sd*)			30 × 30
Electrical Conductivity (*EC*)			30 × 30 (resample
Cation Exchange Capacity (*CEC*)	Secondary Data For MI, the NDWI was calculated	https://soilgrids.org/	
Potential of Hydrogen (pH)			
Bulk Density (*BD*)			
Moisture Index (*MI*)		Landsat 8 OLI $NDWI = \frac{(Band3 - Band5)}{(Band3 + Band5)}$	30 × 30
Soil Texture (*STEX*)			30 × 30

3.1.1. The rationale for selecting the variables

3.1.1.1. Nitrogen (N) and copper (Cu). Nitrogen is crucial for plant growth (structure), plant food processing (metabolism), and chlorophyll formation (Rawat and SanwalSaxena, 2016). Copper is essential for several enzymatic processes in plants, as well as for the formation of chlorophyll and seeds. Copper deficiency can increase susceptibility to diseases such as ergot, which can cause significant yield loss in tiny grains. The majority of Minnesota soils contain acceptable amounts of copper for crop cultivation. On the other hand, copper deficiency can arise in soils with high organic matter and sandy soils.

3.1.1.2. Manganese (Mn) and sulphur (S). Manganese is required by plants to flourish and mature appropriately. Sulphur is one of the most important nutrients for crop productivity, and it is required for plant protein synthesis.

3.1.1.3. Zinc (Zn) and boron (Br). Zinc is required for the formation of chlorophyll and carbohydrate metabolism, as well as for moving calcium through the plant. Boron is vital for controlling hormone levels in plants and encouraging appropriate growth. Boron promotes flower production and retention; pollen tube elongation and germination; seed and fruit development; and seed and fruit germination. Phosphorus concentrations in soils that are adequate stimulate speedy plant growth and development, increase fruiting or maturity, and often improve vegetation quality.

3.1.1.4. Potassium (K) and phosphate. Potassium (K) is an important nutrient for food crops since it is a vital ingredient for plant growth. Potassium, often known as potash, improves plant water use and drought resistance, as well as enhances fruits and vegetables. Potassium and phosphorus are two of the most important macronutrients for plant growth and metabolism. When plants don't get enough of them, their roots don't grow as well, they grow slowly, they aren't as resistant to disease, they mature later, they make fewer seeds, and their yields are lower (Rawat and SanwalSaxena, 2016).

3.1.1.5. Electric Conductivity (EC) and Cation Exchange Capacity (CEC). The ion concentration of the soil solution is assumed to be measured by the EC (Dedeoğlu and Dengiz, 2019), and crop yields, crop suitability, plant nutrient availability, and the activity of soil microbes, etc are all controlled by EC (Do Carmo et al., 2016). CEC is the total amount of exchangeable cations that a soil, soil component, or other material can hold at a certain pH (van Erp et al., 2001). Cation exchange capacity (CEC) is a helpful indication of soil fertility because it indicates the capacity of the soil to deliver three essential plant nutrients: calcium, magnesium, and potassium (Ross and Ketterings, 1995). Similar to the study area, a high concentration of salts (Higher value of EC) in the soil impacts the Kharif crop (Sown in Rainy Season) production negatively and essential variable for Rabi crops (Sown in Winter) cultivation (Francois et al., 1986).

3.1.1.6. Soil organic carbon (SOC) and soil pH. Soil organic carbon (SOC) accumulation is influenced in agricultural areas where the bare soil occurs frequently for several months, especially during hot summers and inconsistent rainfall seasons (Schillaci et al., 2019) and this variation in SOC controls the Copper content in the soil. Otherwise, they'll fail just as badly as if they were missing key components. Soil pH is thereby referred to as the "master soil variable" since it controls a myriad of soil biological, chemical, and physical qualities as well as processes that affect plant growth and biomass yield (Neina, 2019). The pH of the soil is a proxy for the activity or concentration of hydrogen ions ([H+]) in the soil solution. As H+ activity increases, the pH of the soil lowers. As soil pH drops, the most desirable crop nutrients become less available, while others, which are frequently undesirable, become more available and can reach toxic levels.

3.1.1.7. Bulk density and soil moisture. The availability of water, root growth, and the flow of air and water through the soil are all impacted by high bulk density. Soil bulk density is a basic though essential physical soil parameter related to soil porosity, soil moisture, and hydraulic conductivity, which is essential for soil quality evaluation and land use management (Li et al., 2019). Crop health is dependent on a sufficient supply of rainfall and soil nutrients, among other things. Plants' natural functions and growth are interrupted as moisture availability decreases, resulting in lower crop yields. The NDWI was calculated to determine the soil moisture of the study area (Pramanik, 2016) as it is a good indicator of soil moisture. To calculate NDWI the Landsat-8 imageries were collected from the USGS data portal with 30 m resolution (https://earthexplorer.usgs.gov/).

3.1.1.8. Soil depth and soil texture. The depth of the soil influences the types of plants that can survive there. Deeper soils often supply more water and nutrients to plants than shallower soils (Abd-Elmabod et al., 2017). The texture of the soil is the size distribution of grains composed of minerals. The soil texture is determined by three types of grains e.g. sand (2–0.05 mm), silt (0.05–0.002 mm), and clay (below 0.002 mm). Soil texture is an important factor in choosing which crops to cultivate because it affects how porous the soil is, how well it holds water, and how dense it is (Chakraborty and Mistri, 2015). In this research, the soil texture data has been collected from the 'Soil series of West Bengal (Vol. 89), National Bureau of Soil Survey and Land Use Planning, Indian Council of Agricultural Research which is the work of Nayak, D.C. (2001).

3.2. Methods

In the current study, MCDM methodologies such as AHP (Analytical Hierarchy Process), Fuzzy Complex Proportional Assessment (FCOPRAS) methods, as well as machine learning approaches such as Random Forest (RF) and Multilayer Perceptron (MLP) in the present study. When Weighted Linear Combination (WLCM) and MCDM (AHP, FCOPRAS) are used together, the best result is always reached.

3.2.1. Multi-criteria decision-making models

3.2.1.1. AHP (analytic hierarchy process). The primary pair-wise (Eq. (1)) matrix classification in AHP (Saaty, 1980; Malczewski, 2006) is based on a 1–9 scale of relative importance (Table 2), with 1 indicating "equally important" and 9 indicating "extreme importance" (Saaty, 1977; Ghosh and Maiti, 2021).

$$Matrix_{primary}\begin{pmatrix} a_1 & a_2 & a_3 \\ a_{1/2} & a_1 & a_2 \\ a_{1/3} & a_{1/2} & a_1 \end{pmatrix} \tag{1}$$

To correct irregularities, the pair-wise comparison matrix was normalised (Eq. (2)).

$$Matrix_{normalized} = \frac{Xij}{\sum_{j}^{i} Xij} \tag{2}$$

Further, the 'Weighted Sum Vector' was calculated to determine any inconsistencies among the parameters (Saaty, 2001).

The value of CR (Eq. (4)) should be below or equal to 10 per cent or 0.1.

$$CR = \frac{CI}{RI} \tag{3}$$

The Consistency Index and Random Index (Table 2) define the CR value, which is the ratio of CI and RI. The average score of the Consistency Vector (CV) (Eq. (4)) was used to measure the Consistency Index (CI) (Eq. (6)).

$$cv = \left[\frac{wsv_1}{vw_1}\right] \tag{4}$$

The consistency index effectiveness is based on the lambda (γ_{max}) value (Eq. (5)). The higher lambda value shows high inconsistency (Saaty, 1977).

$$\gamma_{max} = \frac{\sum_{i=1}^{n} CV_i}{n} \tag{5}$$

$$CI = \frac{\gamma_{max} - n}{n - 1} \tag{6}$$

where, **WSV:** Weighted Sum Vector; **VW:** Variable Weight; **CV:** Consistency Vector; **λ $_{max}$:** Highest Eigen Value; **n:** No. Of Criteria; **CI:** Consistency Index; **CR:** Consistency Ratio, and **RI:** Random Index.

3.2.1.2. FUZZY COPRAS method. The COPRAS (Complex Proportional Assessment) method is an MCDM method developed by Zavadskas and Kaklauskas in 1996. The matrix and variable weights are computed using Eqs. (7) and (8).

Table 2
The AHP Scale and Random Index for paired comparison.

Index (AHP)	Definition	Random Consistency Index	
		Size of matrix (n)	Random Consistency Index (RI)
1	Equally Important	4	0.90
3	Weak Importance	5	1.12
5	Strong Importance	6	1.24
7	Demonstrated Importance	7	1.32
9	Absolute Importance	8	1.41
2,4,6,8	Intermediate Values	9	1.45

$$V_{rating} = \begin{bmatrix} D_m^1 = l & m & u \\ D_m^2 = l & m & u \\ D_m^3 = l & m & u \end{bmatrix} \tag{7}$$

Where, V_{rating} is the rating assigned by several decision-makers of a particular variable; '$l''m''u$' is the lower, middle, and upper weight of the variable; 'D_m' is the decision-maker (Table 3; Table 4).

After formulation of the tabulation of the weights of the variables assigned by the decision-makers, the Fuzzy-aggregation technique has been employed (Turanoglu Bekar et al., 2016) (Eq. (8)).

$$F_A = \left(\frac{l+l+l}{n}\right); \left(\frac{m+m+m}{n}\right); \left(\frac{u+u+u}{n}\right) \tag{8}$$

Where, F_A is the Fuzzy Aggregation of the fuzzy weights; n is the number of decision-makers.

After the fuzzy aggregation, the AHP normalised matrix values have been used to form a fuzzy weighted normalised decision matrix (Eq. (9)).

$$x_m^w = \left[AHP^{n-matrix} \otimes F_A{}^{l,m,u}\right] \tag{9}$$

To normalise the primary pair-wise matrix Eq. (4) was applied to it. Using Eqs. (10) and (11), the values of the benefit and cost variables were estimated.

$$FB_i = \sum_{j=1}^{k} N_{ij} \tag{10}$$

$$FC_i = \sum_{j=k+1}^{m} N_{ij} \tag{11}$$

Where 'FB_i' is the Fuzzy benefit variable and 'FC_i' is the Fuzzy cost variable.

After the classification of variables into benefit and cost variables, the relative significance (R_i) of each alternative has been estimated using Eq. (12).

$$R_i = FB_i + \frac{\min(FC_i) \times \sum_{i=1}^{n} FC_i}{FC_i \times \sum_{i=1}^{n} \left(\frac{\min(FC_i)}{FC_i}\right)} \tag{12}$$

Where, 'R_i' is the relative significance.

Furthermore, the utility degree of each alternative (Table 5) has been calculated using Eq. (13).

$$U_i^d = \frac{R_i}{\max(R_i)} \times 100 \tag{13}$$

Where, 'U_i^d' is utility degree.

In this study, the normalised matrix from AHP (Table 5) was used in the FCOPRAS model to estimate the weight of the variables. The COPRAS in conjunction with the fuzzy set produces more accurate and refined outputs because fuzzy sets utilize good algorithms to direct incompetence, uncertainty, and dubiety (Zhang and Achari, 2010; Elaalem, 2012), and it improves output acceptance.

Table 3
Relative importance scale for FCOPRAS.

Importance	Linguistic form	Rating		
Extremely low	EL	0	0	0.1
Very low	VL	0	0.1	0.3
Low	L	0.1	0.3	0.5
Medium	M	0.3	0.5	0.7
High	H	0.5	0.7	0.9
Very High	VH	0.7	0.9	1
Extremely high	EH	0.9	1	1

Table 4
Assignment of rating to the Variable by Decision Makers.

Variable	Decision Maker I			Decision Maker II			Decision Maker III		
N	0.9	1.0	1.0	0.7	0.9	1.0	0.5	0.7	0.9
PO_4^{3-}	0.9	1.0	1.0	0.9	1.0	1.0	0.5	0.7	0.9
K	0.9	1.0	1.0	0.7	0.9	1.0	0.5	0.7	0.9
Zn	0.5	0.7	0.9	0.3	0.5	0.7	0.7	0.9	1.0
S	0.7	0.9	1.0	0.7	0.9	1.0	0.9	1.0	1.0
Sd	0.5	0.7	0.9	0.3	0.5	0.7	0.7	0.9	1.0
Cu	0.3	0.5	0.7	0.5	0.7	0.9	0.9	1.0	1.0
B	0.5	0.7	0.9	0.3	0.5	0.7	0.5	0.7	0.9
Mn	0.3	0.5	0.7	0.5	0.7	0.9	0.7	0.9	1.0
EC	0.3	0.5	0.7	0.0	1.0	0.3	0.5	0.7	0.9
CEC	0.5	0.7	0.9	0.1	0.3	0.5	0.5	0.7	0.9
p^H	0.9	1.0	1.0	0.7	0.9	1.0	0.9	1.0	1.0
BD	0.3	0.5	0.7	0.1	0.3	0.5	0.1	0.3	0.5
MI	0.1	0.3	0.5	0.3	0.5	0.7	0.5	0.7	0.9
STEX	0.5	0.7	0.9	0.3	0.5	0.7	0.9	1.0	1.0

3.2.1.3. Weighted Linear Combination Method. The weighted linear combination (WLC) technique (Eq. (14)) is a decision process used in GIS to generate composite maps. The scores for all of the alternatives are added up, and the classifier score is accepted.

$$SP_{WLCM} = \left(grid_{cr}^1 \otimes rating_{cr}^1\right) + ...\left(grid_{cr}^n \otimes rating_{cr}^n\right) \tag{14}$$

where, 'SP' is Soil Potentiality; 'WLCM' is Weighted Linear Combination Method.

3.2.2. Machine learning techniques

To implement the machine learning algorithms, agricultural related data were collected from the field survey. In this regard, the currently cultivated area assigns as a value of 1 and the non-cultivated area assigns as a value of 0. Thereafter, the potential points against the selected variables were collected in the GIS environment using the 'multi values to point' algorithm. Then the generated data were evaluated using the WEKA platform to estimate predicted values for potentiality analysis. After all, the Inverse Distance Weighted (IDW) interpolation technique has been used to generate the spatial distribution maps in ArcMap 10.5.

3.2.2.1. Random forest algorithms. Random forest is a supervised machine learning algorithm that is commonly used to address classification and regression problems (Park et al., 2013). It creates decision trees from various samples, using the supermajority for classification and the average for regression. One of the most essential characteristics of the Random Forest Algorithm is that it can handle data sets with both continuous and categorical variables, as in regression and classification. It produces superior results for classification difficulties. Leo Breiman's Random Forest (Breiman, 2001) is a collection of unpruned classification or regression trees derived from a random selection of training data samples. In the induction process, random features are selected. The ensemble's predictions are summed up (majority vote for classification, or average for regression) to make a prediction. In this study, the actual cropping status of the study area was collected from the agricultural and non-agricultural field visits. The agricultural land was coded as 1 and the non-agricultural land as 0. Thereafter, the suitability values against 1 and 0 for all selected variables were collected in ArcMap 10.5. The '*multi values to point*' algorithm was used the collect the values. There are few tuning options in random forests as the ***'ntree'*** '***mytr***', ***'sampsize',*** ***'nodsize',*** and ***'maxnodes'.*** The ***'ntree'*** is defining the number of trees. Adequate trees to normalise the error, but using too many trees is wasteful, especially when working with large data sets. The '***mytr***' is the number of variables to sample at random as candidates in each split. ***'sampsize',*** is the set of samples used to train. ***'nodsize'*** is the terminal nodes must have a minimum number of samples available. ***'maxnodes'*** is the terminal nodes must have a maximum number of samples available. In this study, '*ntree*' *i.e* the number of trees is 200, '*mytr*' i.e the

Table 5
Normalised pair-wise comparison matrix with criteria rating, CI and CR, AHP and FCOPRAS Weights.

Variable	*N*	PO_4^{3-}	*K*	*Zn*	*S*	*Sd*	*Cu*	*B*	*Mn*	*EC*	*CEC*	P^H	*BD*	*MI*	*STEX*	AHP Weight (*vw*)	FCOPRAS Weight (U_i^d)	Consistency Checking
N	0.09	0.08	0.08	0.13	0.08	0.08	0.14	0.08	0.14	0.08	0.07	0.09	0.08	0.07	0.07	0.09	1.00	**CI:**0.03
PO_4^{3-}	0.09	0.08	0.08	0.07	0.08	0.08	0.07	0.08	0.07	0.08	0.07	0.09	0.08	0.07	0.07	0.08	0.88	**CR:**0.02
K	0.09	0.08	0.08	0.07	0.08	0.08	0.07	0.08	0.07	0.08	0.07	0.09	0.08	0.07	0.07	0.08	0.88	Model's result Acceptable
Zn	0.04	0.08	0.08	0.07	0.04	0.08	0.07	0.08	0.07	0.08	0.07	0.09	0.08	0.07	0.07	0.07	0.80	
S	0.09	0.08	0.08	0.13	0.08	0.08	0.07	0.08	0.07	0.08	0.07	0.09	0.08	0.07	0.07	0.08	0.94	
Sd	0.09	0.08	0.08	0.07	0.08	0.08	0.07	0.08	0.07	0.08	0.07	0.09	0.08	0.07	0.07	0.08	0.88	
Cu	0.04	0.08	0.08	0.07	0.08	0.08	0.07	0.08	0.07	0.08	0.07	0.09	0.08	0.07	0.07	0.07	0.84	
B	0.09	0.08	0.08	0.07	0.08	0.08	0.07	0.08	0.07	0.08	0.07	0.09	0.08	0.07	0.07	0.08	0.88	
Mn	0.04	0.08	0.08	0.07	0.08	0.08	0.07	0.08	0.07	0.08	0.07	0.09	0.08	0.07	0.07	0.07	0.84	
EC	0.04	0.04	0.04	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.07	0.02	0.05	0.04	0.07	0.04	0.68	
CEC	0.09	0.08	0.08	0.07	0.08	0.08	0.07	0.08	0.07	0.04	0.07	0.04	0.05	0.07	0.07	0.07	0.96	
P^H	0.04	0.04	0.04	0.03	0.04	0.04	0.04	0.04	0.04	0.08	0.07	0.04	0.08	0.07	0.07	0.05	0.59	
BD	0.03	0.03	0.03	0.02	0.03	0.03	0.02	0.03	0.02	0.02	0.04	0.01	0.03	0.07	0.07	0.03	0.74	
MI	0.04	0.04	0.04	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.02	0.01	0.04	0.03	0.04	0.69	
STEX	0.09	0.08	0.08	0.07	0.08	0.08	0.07	0.08	0.07	0.04	0.07	0.04	0.03	0.07	0.07	0.07	0.99	

Where, *N*: Nitrogen; PO_4^{3-}: Phosphate; *K*: Potassium; *Zn*: Zinc; *S*: Sulphur; *Sd*: Soil Depth; *Cu*: Copper; *B*: Boron; *Mn*: Manganese; *EC*: Electrical Conductivity; *CEC*: Cation Exchange Capacity; P^H: Potential Hydrogen; *BD*: Bulk Density; *MI*: Moisture Index; *STEX*: Soil Texture.

number of variables to sample at random as candidates in each split is 1 (Liaw and Wiener, 2002), min*sampsize* i.e the set of samples used to train (min) is 3, max number of levels in each decision tree 80, 'maxnodes' i.e the terminal nodes (max) and 'nodsize' i.e the terminal nodes (min) is remain default.

3.2.2.2. Multilayer perceptron algorithms. One of the most basic feed-forward neural networks is the Multilayer Perceptron. Multilayer perceptrons are bidirectional neural networks in which the inputs are propagated forward and the weights are propagated backward (Breiman, 2001). A multilayer perceptron has an input layer and an output layer with one or more hidden layers. All MLPs connect all neurons in one layer to all neurons in the next layer. The input layer receives the input signals, and the output layer performs the intended task. All of the calculations are done by the hidden layers (Wang et al., 2021).

3.2.3. Effectiveness estimation of applied models/algorithms

3.2.3.1. 3.2.3.1. ROC (receiver operating characteristic curve). The Receiver Operating Characteristic Curve (ROC curve) is a graph that shows the output of a classification model at all classification levels. If the area under the ROC curve (AUC) values between 0.9 and 1 were considered excellent, the values between 0.8 and 0.9 were considered good, the values between 0.7 and 0.8 were considered fair, the values between 0.6 and 0.7 were considered poor, and the values between 0.5 and 0.6 were considered failed (Bradley, 1997; Sarkar et al., 2022).

3.2.3.2. Accuracy measurement. Accuracy is defined as the degree to which the outcome of measurement corresponds to the correct value or a standard and refers to the degree to which a measurement is near to its agreed-upon value, i.e., it is the degree of consistency between the outcome of measurement and the true value of the thing being measured. Equation 16 was used to measure the accuracy of the final outputs. The flow diagram and soil testing processes of the whole study area are in Fig. 2.

$$Accuracy^{overall} = \left(\frac{(\sum true\_values)}{(\sum predicted\_values)}\right) \times 100 \tag{15}$$

4. Result analysis

4.1. Execution of the AHP method for estimating relative weights of the variables

All the selected sixteen different variables including nitrogen (N), copper (Cu), manganese (Mn), zinc (Zn), sulphur (S), boron (B), phosphate, and potassium (K), electrical conductivity (EC), Cation Exchange Capacity (CEC), soil pH, bulk density, soil moisture, soil organic carbon, soil depth, and soil texture have been classified into suitable classes based on natural break classification in GIS environment. The Inverse Distance Weighted (IDW) interpolation (Chen and Liu, 2012) technique has been used to generate spatial distribution maps.

The spatial distribution of nitrogen (N) has been categorised (Fig. 3a) into five potential classes based on natural break classification (Febrianto et al., 2016) in the GIS environment, where the >70 Zone has the most potential, with 20.83 per cent (21832.44 ha) area. The 1.64–1.98 copper (Cu) Zone (Fig. 3b) is more suitable, with an 11.79 per cent (12354.32 ha) area. The thematic distribution of manganese (Mn) and zinc (Zn) is (Fig. 3c; Fig. 3d) also classified into five classes, where the 16.37–19.98 Mn Zone and the 0.87–0.99 Zn Zone have more potential, with 15.23 per cent (15955.27 ha) and 5.95 per cent (6234.64 ha) of the total area, respectively. The spatial database of the Sulphur (Fig. 3e) has been classified as 67.36–71.85, 71.85–79.76, 79.76–87.00, 87.00–96.01, and 96.01–113.36, respectively. Boron Zone 0.32–0.36 (Fig. 3f) has the most potential, with a 13.31 per cent (13951.9 ha) area. The Phosphate (Fig. 3g) Zone 255.74–310.30 has the most potential,

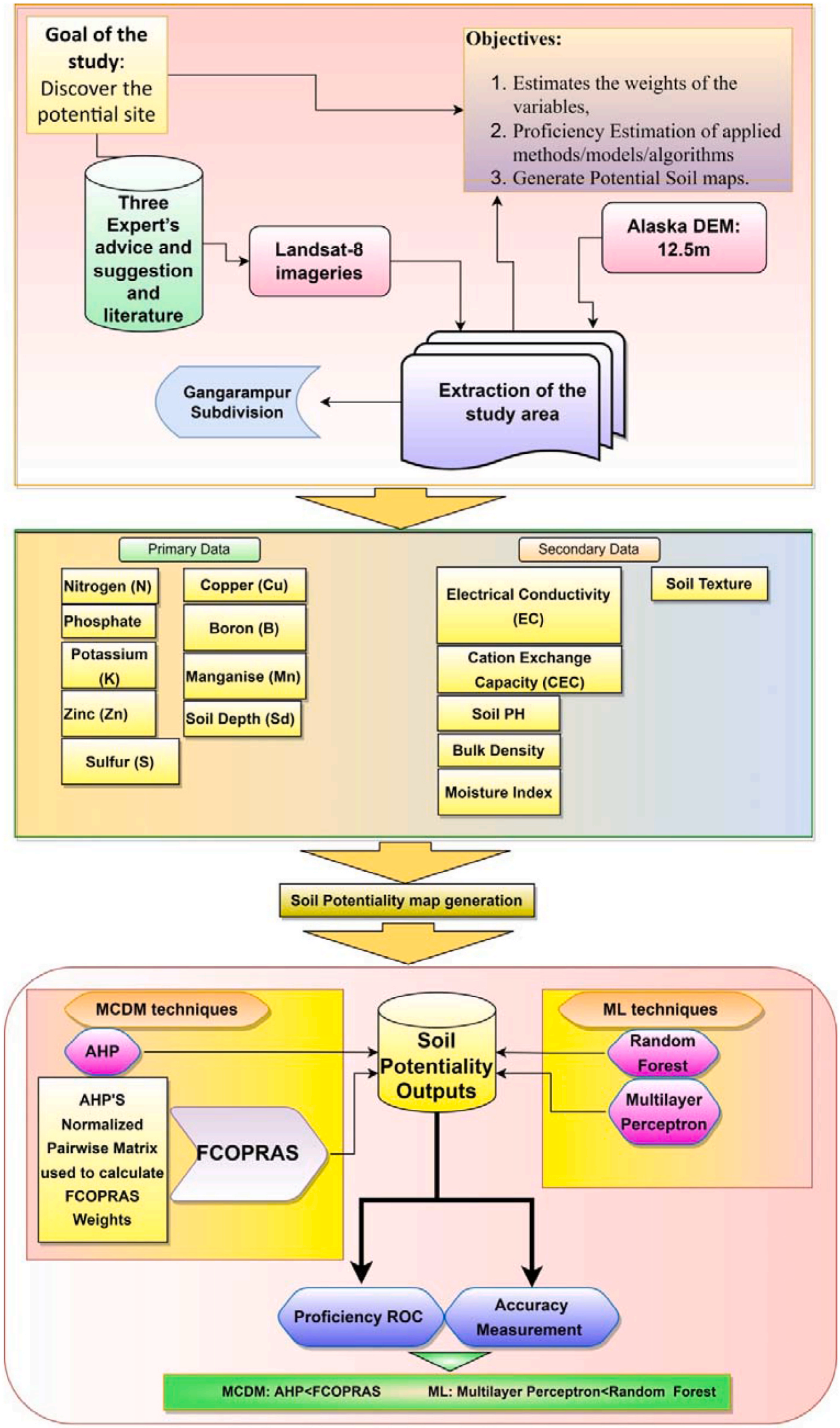

Fig. 2. Flow diagram of the Whole Study.

accounting for 7.01 per cent (7346.602 ha) of the total area. The spatial layer of the potassium (K) has (Fig. 3h) been classified into five potential classes, where the >70 Zone is the most suitable, with a 20.83 per cent (21832.44 ha) area. The Electrical Conductivity (Fig. 3i) Zone 1.41–2.29 and the Cation Exchange Capacity (CEC) Zone (Fig. 3g) 21.74–26.96 have a potential of 20.37 per cent (21350.92 ha) and 9.45 per cent (9901.33 ha) respectively. The pH Zone (Fig. 3k) of 6.2–6.3 and the bulk density Zone of 1.52–1.54 have the most potential for agriculture with 35.16 per cent (36848.48 ha) and 4.83 per cent (5061.38 ha) of the total area, respectively. The MI (Fig. 3m) Zone 0.50–0.78 has the most potential, with 20.22 per cent (23265.22 ha). The soil depth Zone 6652–8118 has the most potential with 1.23 per cent (1290.28 ha). The

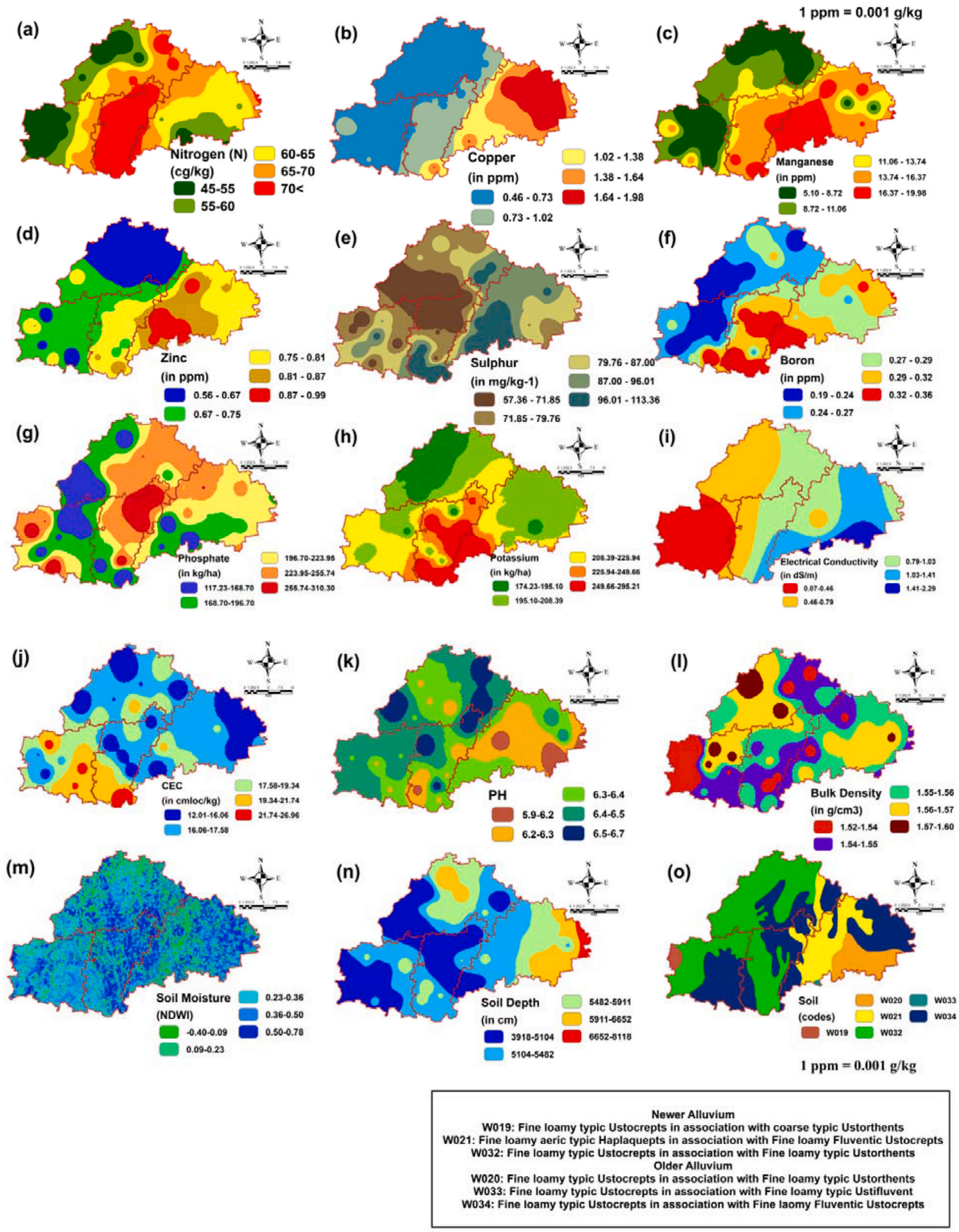

Fig. 3. Influecing variable of the Soil Fertility (a–o); (a) Nitrogen (N), (b) Copper, (c) Manganese, (d) Zinc, (e) Sulphur, (f) Boron, (g) Phosphate, (h) Potassium, (i) Electric Conductivity(EC), (j) CEC, (k) PH, (l) Bulk Density, (m) Soil Moisture, (n) Soil Depth, (o) Soil texture.

thematic layer of soil (Fig. 3o) has been classified into six classes: W019, W020, W021, W032, W033, and W034 respectively. W019, W021, and W032 belong to the new alluvial Zone, which has a high potential for cultivation with 38.18 per cent (40009.26 ha), 39.28 per cent (41156.81 ha), and 13.83 per cent (14490.93 ha) respectively.

4.2. Final soil potential map generation

4.2.1. MCDM techniques

The final potential suitability map using AHP weight has been generated in the GIS environment and classified (Fig. 4a) into five suitable classes as Zone I (Very Low) accounts for 22.50 per cent (23581.74 ha), Zone II (Low) accounts for 24.87 per cent (26062.27 ha), Zone III (Moderate) accounts for 19.57 per cent (20508.32 ha), Zone IV

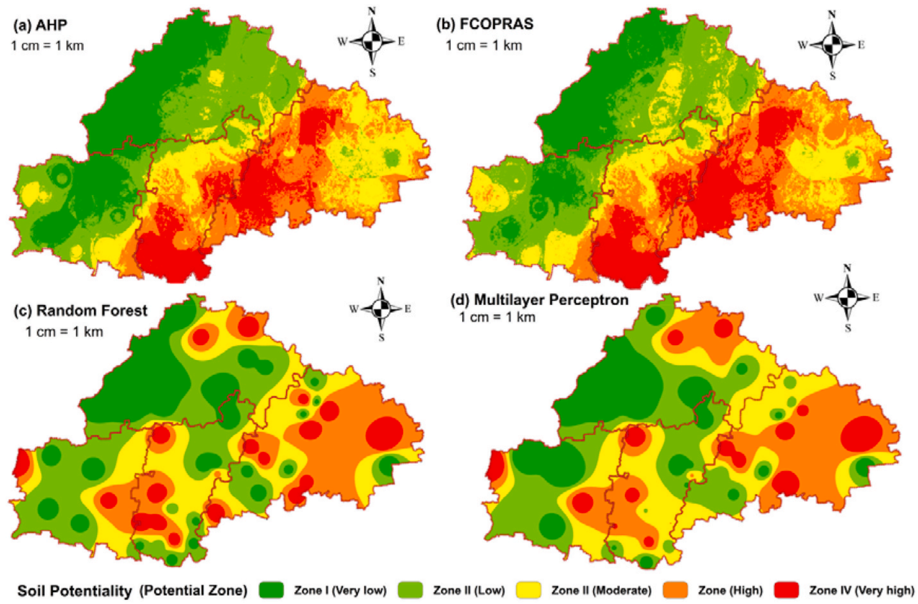

Fig. 4. Potential soil mapping (a–d); (a) AHP, (b) FCOPRAS, (c)Random forest, (d) multilayer perceptron.

(High) accounts for 20.52 per cent (21503.99 ha), and Zone V (Very High) accounts for 12.53 per cent (13133.67 ha).

The final potential suitability map using FCOPRAS weight has been generated in the GIS environment and classified (Fig. 4b) into five suitable Zones: Zone I (Very Low) with 17.51 per cent (18344.56 ha), Zone II (Low) with 24.77 per cent (25961.01 ha), Zone III (moderate) with 19.79 per cent (20733.26 ha), Zone IV (High) with 22.54 per cent (23620.65 ha), and Zone V (Very high) with 15.39 per cent (16130.52 ha) of the total area, respectively (Table 6).

Table 6
Area distribution of potential classes using MCDM and ML.

AHP			
Zone	Potentiality	Area (%)	Area (ha)
ZONE I	Very Low	22.50	23581.74
ZONE II	Low	24.87	26062.27
ZONE III	Moderate	19.57	20508.32
ZONE IV	High	20.52	21503.99
ZONE V	Very High	12.53	13133.67
FCOPRAS			
Zone	Potentiality	Area (%)	Area (ha)
ZONE I	Very Low	17.51	18344.56
ZONE II	Low	24.77	25961.01
ZONE III	Moderate	19.79	20733.26
ZONE IV	High	22.54	23620.65
ZONE V	Very High	15.39	16130.52
Random Forest			
Zone	Potentiality	Area (%)	Area (ha)
ZONE I	Very Low	18.46	19343.35
ZONE II	Low	20.94	21948.15
ZONE III	Moderate	23.63	24762.12
ZONE IV	High	22.30	23367.9
ZONE V	Very High	14.67	15368.46
Multilayer Perceptron			
Zone	Potentiality	Area (%)	Area (ha)
ZONE I	Very Low	10.53	11031.33
ZONE II	Low	26.02	27262.28
ZONE III	Moderate	35.99	37718.49
ZONE IV	High	22.01	23059.92
ZONE V	Very High	5.46	5717.977

4.2.2. *ML algorithms/techniques*

The calibration of current RF model is involved hyperparameter adjustment to find the best balance of the factors of n estimators, number of decision tree levels, number of features that can be used to split a node, and sampling method (with or without replacement. Bagging with 100 iterations of ten-fold cross validation was performed to facilitate model optimization and generalizability while avoiding over fitting on the test dataset.

The final potential suitability map using the Random Forest algorithm has been generated in Weka 3.8.5 and the GIS environment and classified (Fig. 4c) into five suitable classes as Zone I (Very Low) with 18.46 per cent (19343.35 ha), Zone II (Low) with 20.94 per cent (21948.15 ha), Zone III (moderate) with 23.63 per cent (24762.12 ha), Zone IV (High) with 22.30 per cent (23367.9 ha), and Zone V (Very high) with 14.67 per cent (15368.46 ha) of the total area, respectively.

The final potential suitability map using the Multilayer Perceptron algorithm has been generated in Weka 3.8.5 (WEKA platform was developed by the University of Waikato, New Zealand) and the GIS environment and classified (Fig. 4d) into five suitable classes as Zone I (Very Low) with 10.53 per cent (11031.33 ha), Zone II (Low) with 26.02 per cent (27262.28 ha), Zone III (moderate) with 35.99percent (37718.49 ha), Zone IV (High) with 22.01 per cent (23059.92 ha), and Zone V (Very high) with 5.46 per cent (5717.977 ha) of the total area, respectively (Table 6).

5. Effectiveness estimation of the applied models/methods/ algorithms

5.1. *Area changing matrix and proficiency estimation using ROC*

In the current research, 70 (70%) locations are employed for soil potentiality mapping, and 30 (30%) of the entire data set is used to justify the potentiality analysis. The AUC for the AHP model is 0.923, whereas the FCOPRAS model has an AUC of 0.938. On the other hand, the AUC for the Random Forest is 0.947, and for the Multilayer Perceptron, it is 0.932. The models used for the predicted maps are very accurate and sufficient. So, it's safe to say that Random Forest and FCOPRAS are better for analysing soil potentiality (Fig. 5).

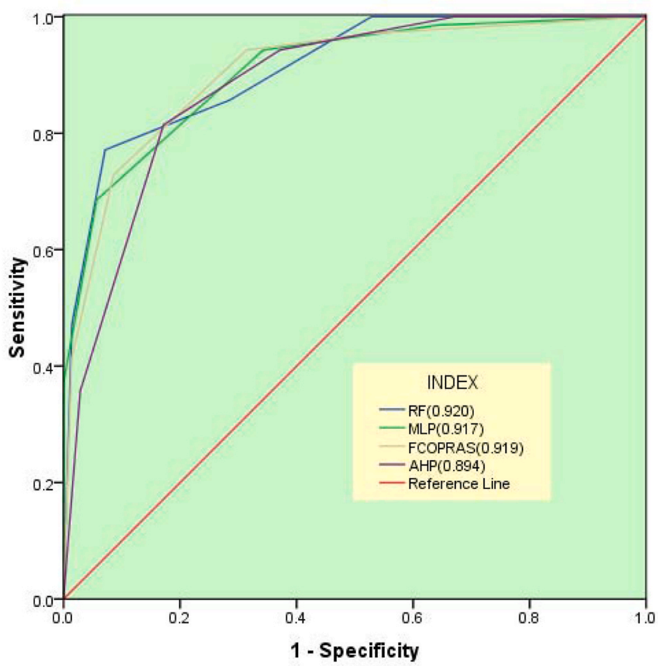

Fig. 5. Validation of the results applying ROC.

6. Discussion

6.1. Soil potentiality

Based on the results of the MCDM, Berail, Deul, Akcha G.P. (Khusmandi C.D. Block), Gokamo, and Bagichapur G.P. (Harirampur C.D. Block) have very low soil potential, while Udaypur, Kalikamera, and Maligaon G.P. have higher soil potential. Places with very high soil potential for agriculture include Ellahbad (Banshihari C.D. Block), Belbari, Gangarampur G.P. (Gangarampur C.D. Block), etc. From the ML outputs, it is found that Sukhdebpur G.P., Maligaon G.P., Basuria G.P. Nandanpur G.P., NW Saiyadpur G.P., Southern Sirshi G.P., Ganguria G. P., NW Karanji G.P. have the least potential, whereas Berail G.P., Deul G. P., Akcha G.P., Asokegram G.P., Kushmandi G.P., Northern Karanji G.P., etc have the highest potential for agriculture.

Deficiency in Nitrogen, Copper (Chaudhry and Loneragan., 1970), low Soil Moisture (Nagy et al., 2007), Deficiency in Manganese (Schmidt et al., 2016), Zinc, Sulphur, Boron, Phosphate, Potassium (Seth et al., 2018), High Bulk Density (Blake, 1965), Older Alluvial Soil, High pH, and least soil depth, etc are the main reason behind low to very low potentiality of Soil for agriculture. Whereas Optimum concentration of Nitrogen (>65 cg/kg^{-1}), Copper (1.02–1.98 ppm), Soil Moisture (0.36–0.78), Manganese (11.0–19.98 ppm), Zinc (0.75–0.99 ppm), Boron (0.25–0.38 ppm), Sulphur (75.0–113.36 mg/kg^{-1}), Phosphate (>195.0 kg/ha^{-1}), Potassium (>200 kg/ha^{-1}), low Bulk Density (<1.55 g/cm3), Older Alluvial Soil, ideal PH (6.5–7.5), low EC (<0.46 dS/m^{-1}), higher CEC activity (>20.0 cmloc/kg^{-1}) and ideal soil depth (<5000 cm) etc are the main reason behind low to very low soil positive potentiality for agriculture. **(1 ppm = 0.001 g/kg)**.

6.2. Model calibration, validation, and effectiveness

From the areal changing matrix, it is found that there is about +4.998 per cent variation for Zone I from the FCOPRAS to AHP, whereas it is −0.215 per cent for Zone III and −2.860 per cent for Zone V. The ROC and accuracy estimation indicate that the FCOPRAS is more valid and accurate than the AHP, with 91.9 percent and 84.31 percent, respectively. Another hand, it is found that there is about +1.021 per cent areal variation for Zone I from the FCOPRAS to Random Forest, whereas it is +3.781 per cent for Zone III and-0.742 per cent for Zone V. The ROC estimation indicates that Random Forest has higher validity and accuracy, with 92.0 percent and 85.88 percent. On the other hand, about −3.68 per cent areal variation of Zone I from FCOPRAS to Multilayer Perceptron, whereas it is-2.11 per cent for Zone III and +5.85 per cent for Zone V. The ROC estimation indicates FCOPRAS has 0.20 per cent higher validity than Multilayer Perceptron. The areal changing matrix indicates that there is about −3.98 per cent areal variation in Zone I from the AHP to Random Forest technique, whereas it is +3.10 per cent for Zone III and +2.12 per cent for Zone V. The ROC estimation indicates that the validity of the Random Forest is 92.0 per cent which is very much satisfying condition of the result (Table 7).

From the areal changing matrix, it is found that there is about +11.98 per cent areal variation for Zone I from the AHP to Multilayer Perceptron where it is −1.15 per cent for Zone II, 16.42 per cent for Zone III, and +7.08 per cent for Zone V. The ROC and accuracy estimation indicate that the MLP has higher validity and accuracy, with 91.7 per cent and 78.72 per cent, respectively. From the comparison between RF and MLP, it is found that about +7.93 per cent areal variation in Zone I, whereas it is −5.07 per cent for Zone II and +9.21 per cent for Zone V. The ROC estimation indicates that Random Forest (Shi et al., 2021) has higher acceptance than MLP (Thach et al., 2018) with 92.0 per cent and 91.7 per cent respectively.

The ROC for AHP is 0.894 (89.4%), for FCOPRAS is 0.919 (91.9%), for Random Forest is 0.920 (92.0%), for Multi-layer Perceptron is 0.917 (91.7%) whereas the Accuracy values for above mentioned techniques are 78.72%, 84.31%, 85.88%, and 83.33% respectively. In both ways, the accuracies of Random forest, and FCOPRAS has higher acceptance than other methods (Fig. 5).

From the overall ROC estimation, it is found that the Random Forest method is more proficient and effective to measure site suitability, followed by FCOPRAS though the ROC of applied techniques is above 90%. In the Second technique, the validity and accuracy of the Random Forest are higher than in other techniques.

7. Novelty and major findings of the study

- In this study, the AHP pair-wise comparison matrix is used in FCOPRAS for the estimation of the variable's weight.
- The effectiveness of the models has been checked using ROC and Accuracy estimation. It is revealed that the applied models are more or less strongly effective in potentiality analysis as the FCOPRAS and Random Forest approaches are more effective than other applied techniques.
- From the study, it is found that about 10–19 per cent of the total area is the least potential for agriculture and about 14–23 per cent of the area is the most potential for agriculture.
- This research will also help governments and non-governmental organisations in the region implement development plans and manage agricultural land.

8. Conclusion

Several distinct variables were chosen to assess land potential assessments, with scales ranging from nominal to ratio. The soil appropriateness of the research region was evaluated using a GIS-based multi-criteria decision-making technique and machine learning algorithms (Random Forest, Multilayer Perceptron) in this study. In this acceptability analysis, the AHP approach was utilised to assign relative weight to the fifteen different criteria, and thereafter the FCOPRAS model was applied using a normalised pairwise comparison matrix from AHP. The machine learning algorithms were applied using field data in the Weka software. Because it is a time-saving strategy, the "GIS" technique has been used to "Obtain the result," "Investigate the result," and "Analyse the data." The continuous, complicated, and uncertain information is

Table 7
Areal Changing Matrix among the applied techniques for Proficiency.

Changing Matrix	Potential Zone	ZONE I	ZONE II	ZONE III	ZONE IV	ZONE V	Area Changes (%)
FCOPRAS TO AHP	ZONE I	1.000000	0.000000	0.000000	0.000000	0.000000	4.10
	ZONE II	0.201700	0.798300	0.000000	0.000000	0.000000	0.10
	ZONE III	0.000000	0.257500	0.742500	0.000000	0.000000	-0.22
	ZONE IV	0.000000	0.000000	0.216500	0.783500	0.000000	-2.02
	ZONE V	0.000000	0.000000	0.000000	0.185800	0.814200	-2.86
RF TO MLP	ZONE I	0.157486	0.302111	0.376136	0.09056	0.073707	7.93
	ZONE II	0.117587	0.238813	0.323425	0.249753	0.070423	-5.07
	ZONE III	0.033267	0.164749	0.46348	0.322826	0.015678	-12.36
	ZONE IV	0.105843	0.327747	0.361057	0.157288	0.048063	0.29
	ZONE V	0.137107	0.288813	0.223202	0.270503	0.080375	9.21
AHP TO RF	ZONE I	0.089297	0.136498	0.128482	0.237723	0.408001	-3.98
	ZONE II	0.319210	0.258886	0.188157	0.200992	0.032754	-3.90
	ZONE III	0.117373	0.179420	0.425643	0.234080	0.043483	3.10
	ZONE IV	0.187911	0.230085	0.251662	0.209324	0.121018	1.76
	ZONE V	0.193541	0.257795	0.199592	0.243720	0.105352	2.12
FCOPRAS TO RF	ZONE I	0.068443	0.088551	0.116568	0.247058	0.479379	1.02
	ZONE II	0.307301	0.252792	0.165329	0.211720	0.062858	-3.80
	ZONE III	0.151701	0.226916	0.385488	0.208471	0.027425	3.78
	ZONE IV	0.166700	0.209058	0.299185	0.215987	0.109070	-0.26
	ZONE V	0.192102	0.257224	0.198763	0.241431	0.110480	-0.74
FCOPRAS TO MLP	ZONE I	0.117543	0.263939	0.263867	0.246819	0.107832	6.98
	ZONE II	0.072708	0.34743	0.313515	0.235952	0.030395	-1.24
	ZONE III	0.008263	0.210192	0.534505	0.236024	0.011016	-16.21
	ZONE IV	0.087597	0.211079	0.428551	0.212648	0.060125	0.54
	ZONE V	0.287889	0.244678	0.225435	0.15783	0.084167	9.94
AHP TO MLP	ZONE I	0.120461	0.263102	0.273037	0.241889	0.101511	11.98
	ZONE II	0.049384	0.347397	0.349756	0.236305	0.017158	-1.15
	ZONE III	0.009116	0.201458	0.54481	0.228693	0.015923	-16.42
	ZONE IV	0.108357	0.214086	0.391964	0.215493	0.070099	-1.48
	ZONE V	0.326129	0.240484	0.202888	0.146859	0.08364	7.08

represented in a simple, categorised map manner via GIS-based land potential analysis. According to the findings, the area is an appropriate agricultural region. The main factors are optimum organic carbon content, low bulk density, optimum soil pH, and cation exchange capability, which are behind the high potential. The primary causes of less productive areas are sluggish soil quality, lack of soil moisture, and so on. Although this location is naturally rich, there is a threat of crop devastation because of flooding during the rainy season. From the validity estimation, it is found that FCOPRAS and RF techniques are more proficient with 91.9 per cent and 92.0 per cent respectively. Farming can have a big impact on the environment. Even though the negative effects are serious and can include pollution and the degradation of soil, water, and air, agriculture can also have positive effects on the environment. For example, crops and soils can trap greenhouse gases. This suitability analysis helps the government instruct farmers on how to choose the best place for farming to preserve the environment. Government authorities should take steps to strengthen flood forecasting systems, flood management measures, marketing facilities, transportation, and soil management to minimise damage and promote growth in this area. This study will also assist governmental and non-governmental organisations in this region in executing development strategies and managing agricultural land.

Declaration of competing interest

The authors hereby declare that there is no conflict of interest. During the research work, no humans or animals are wounded or harmed in any way.

Acknowledgements

The authors are very much thankful to Dr. Gopal Chandra Debnath (Retired Professor, Visva Bharati, West Bengal, India) and Former Professor Dr. Narayan Chandra Ghosh (Rabindra Bharati University, West Bengal, India) for their valuable suggestions throughout the study.

References

Abd-Elmabod, S.K., Jordán, A., Fleskens, L., Phillips, J.D., Muñoz-Rojas, M., van der Ploeg, M., Anaya-Romero, M., El-Ashry, S., de la Rosa, D., 2017. Modeling agricultural suitability along soil transects under current conditions and improved scenario of soil factors. Soil Map. Pro. Mod. Sust. Land Use Manag. 193–219. https://doi.org/10.1016/B978-0-12-805200-6.00007-4.

Adeniyi, P.O., 1993. Integration of remote sensing and GIS for agricultural resource management in Nigeria. EARSeL Adv. Remote Sens. 2 (3), 6–21.

Alexandratos, N., Bruinsma, J., 2012. World agriculture towards 2030/2050: the 2012 revision (ESA working papers). ESA working paper No. 12-03. Agri. Dev. Econ. Div., Food and Agri Org United Nat (Rome).

Azad, N., Abajelu, R.E., Behmanesh, J., 2019. Locating the potential areas for executing surface and pressuerized irrigation systems using fuzzy analytical hierarchy process method in miandoab plain. JISE 41 (4), 119–132.

Badhe, Y., Medhe, R., Shelar, T., 2019. Site suitability analysis for water conservation using AHP and GIS techniques: a case study of upper sina river catchment. Ahmednagar (India) 3, 49–59. https://doi.org/10.21523/gcj3.19030201.

Bandyopadhyay, S., Jaiswal, R.K., Hegde, V.S., Jayaraman, V., 2009. Assessment of land suitability potentials for agriculture using a remote sensing and GIS based approach. Int. J. Rem. Sens. 30 (4), 879–895.

Blake, G.R., 1965. Bulk density. Methods of soil analysis: Part 1 physical and mineralogical properties. Including Statistics of Measurement and Sampling 9, 374–390.

Bradley, A.P., 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recogn. 30 (7), 1145–1159. https://doi.org/10.1016/S0031-3203(96)00142-2.

Breiman, L., 2001. Random forests. Mach learning 45 (1), 5–32.

Chakraborty, K., Mistri, B., 2015. Importance of Soil Texture in Sustenance of Agriculture: a Study in Burdwan-I CD Block, Burdwan, West Bengal, vol. 21. Eastern Geogra, pp. 475–482.

Chaudhry, F.M., Loneragan, J.F., 1970. Effects of nitrogen, copper, and zinc fertilizers on the copper and zinc nutrition of wheat plants. Aust. J. Agri. Research. 21 (6), 865. https://doi.org/10.1071/ar9700865.

Chen, F.-W., Liu, C.-W., 2012. Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. Paddy Water Environ. 10 (3), 209–222. https://doi.org/10.1007/s10333-012-0319-1.

Cowie, A.L., Orr, B.J., Castillo Sanchez, V.M., Chasek, P., Crossman, N.D., Erlewein, A., Welton, S., 2018. Land in balance: the scientific conceptual framework for Land Degradation Neutrality. Environ. Sci. Pol. 79, 25–35.

Davis, G.C., Langham, R.M., 1995. Agricultural industrialization and sustainable development: a global perspective. J. Agric. Appl. Econ. 21–34.

De la Rosa, D., Sobral, R., 2008. Soil Quality and Methods for its Assessment. Land Use Soil Resour, pp. 167–200.

Dedeoğlu, M., Dengiz, O., 2019. Generating of land suitability index for wheat with hybrid system aproach using AHP and GIS. Comput. Electron. Agric. 167, 105062 https://doi.org/10.1016/j.compag.2019.105062.
Do Carmo, D.L., De Lima, L.B., Silva, C.A., 2016. Soil fertility and electrical conductivity affected by organic waste rates and nutrient inputs. Rev. Bras. Cienc. Solo. 40.
Duc, T.T., 2006. Using GIS and AHP technique for land-use suitability analysis. Int. Symp. Geo. Informa. spatial inf. Dev. earth allied. sci. 1–6.
Ebrahimi, M., Nejadsoleymani, H., Mansouri Daneshvar, M.R., 2019. Land suitability map and ecological carrying capacity for the recognition of touristic Zones in the Kalat region, Iran: a multi-criteria analysis based on AHP and GIS. Asia-Pacific J. Regional Sci. 3 (3), 697–718.
Elaalem, M., 2012. Land suitability evaluation for sorghum based on boolean and fuzzy-multi-criteria decision analysis methods. Int. Environ. Sci. Dev. 3 (4), 357–361.
FAO, 1976. A Framework for Land Evaluation, vol. 32. FAO Soil. Bull.
Febrianto, H., Fariza, A., Nur Hasim, J.A., 2016. Urban flood risk mapping using analytic hierarchy process and natural break classification (Case study: surabaya, East Java, Indonesia). Int. Conf. Know. Creat. Int. Comp. https://doi.org/10.1109/kcic.2016.7883639.
Feizizadeh, B., Jankowski, P., Blaschke, T., 2014. A GIS based spatially-explicit sensitivity and uncertainty analysis approach for multi-criteria decision analysis. Comp. Geo sci. 64, 81–95.
Francois, L.E., Maas, E.V., Donovan, T.J., Youngs, V.L., 1986. Effect of salinity on grain yield and quality, vegetative growth, and germination of semi-dwarf and durum wheat. Agron. J. 78 (6), 1053–1058. https://doi.org/10.2134/agronj1986.00021962007800060023x.
Ghosh, A., Maiti, R., 2021. Development of new Ecological Susceptibility Index (ESI) for monitoring ecological risk of river corridor using F-AHP and AHP and its application on the Mayurakshi river of Eastern India. Ecol. Inf. 63, 101318.
Grigg, D., 1995. An Introduction Agricultural Geography, vol. 11. Rout. Pub.. New Fetter Lane, London EC4P 4EE.
Halder, B., Bandyopadhyay, J., Banik, P., 2020. Assessment of hospital sites' suitability by spatial information technologies using AHP and GIS-based multi-criteria approach of Rajpur-Sonarpur Municipality. Mod. Earth Syst. Environ. 6 (4), 2581–2596.
Hayashi, K., 2000. Multi criteria analysis for agricultural resource management: a critical survey and future perspectives. Eur. J. Oper. Res. 122, 486–500.
Hopkins, L., 1977. Methods for generating land suitability maps: a comparative evaluation. J. American Inst. Plann. 34 (1), 19–29.
Ibrahim, Y.H., Adeola, S.S., Ibrahim, L.H., 2016. Determinants of food insecurity among farming households in Katsina state, North Western Nigeria: an ordinal logit regression approach. J. Agric. Sci. 61 (3), 291–301.
Jamil, M., Sahana, M., Sajjad, H., 2018. Crop suitability analysis in the bijnor district, UP, using geospatial tools and fuzzy analytical hierarchy process. Agribiol. Res. 7 (8), 506–522.
Jayaraman, V., Sridevi, S., Monica, K.M., Lakshminarayanan, A.R., 2021. Predicting the soil suitability using machine learning techniques. Int. Conf. Dis. Tech. Multi-Dis. Res. Appl. 1, 200–202.
Joerin, F., Theriault, M., Musy, A., 2001. Using GIS and outranking multicriteria analysis for land-use suitability assessment. Int. J. Geogr. Inf. Sci. 15 (2), 153–174.
Kihoro, J., Bosco, N., Murage, H., 2013. Suitability analysis for rice growing sites using a multi criteria evaluation and GIS approach in great Mwea region, Kenya. Sprin. P. 2 (1), 1–9.
Lambin, E.F., Meyfroidt, P., 2011. Global land use change, economic globalization, and the looming land scarcity. Proc. Natl. Acad. Sci. USA 108 (9), 3465–3472.
Li, S., Li, Q., Wang, C., Li, B., Gao, X., Li, Y., Wu, D., 2019. Spatial variability of soil bulk density and its controlling factors in an agricultural intensive area of Chengdu Plain, Southwest China. J. Integr. Agric. 18 (2), 290–300. https://doi.org/10.1016/S2095-3119(18)61930-6.
Liaw, A., Wiener, M., 2002. Classification and regression by random Forest. R. News 2 (3), 18–22.
Malczewski, J., 2004. GIS-based land suitability: a critical overview. Prog. Plann. 62 (1), 3–65.
Malczewski, J., 2006. GIS-based multi-criteria decision analysis: a survey of the literature. Int. J. Geogr. Inf. Sci. 20 (7), 703–726.
Malczewski, J., Boroushaki, S., 2008. Implementing an extension of the analytical hierarchy process using ordered weighted averaging operators with fuzzy quantifiers in ArcGIS. Comput. Geosci. 34 (4), 399–410.
Mendas, A., Delali, A., 2012. Integration of multi-criteria decision analysis in GIS to develop land suitability for agriculture: application to durum wheat cultivation in the region of Mleta in Algeria. Comput. Electron. Agric. 83, 117–126.
Mistri, P., Sengupta, S., 2020. Multi-criteria decision-making approaches to agricultural land suitability classification of Malda district, eastern India. Nat. Resour. Res. 29, 2237–2256. https://doi.org/10.1007/s11053-019-09556-8.
Mokarram, M., Aminzadeh, F., 2010. GIS-based multi-criteria land suitability evaluation using ordered weight averaging with fuzzy quantifier: a case study in Shavur Plain, Iran. Int. Arch. Photo. Remote Sens. Spatial Infor. Sci. 38 (2), 508–512.
Mosadeghi, R., Warnken, J., Tomlinson, R., Mirfenderesk, H., 2015. Comparison of Fuzzy-AHP and AHP in a spatial multi-criteria decision making model for urban land-use planning. Comput. Environ. Urban Syst. 49, 54–65.
Murayam, Y., Bunruamkaew, K., 2011. Site suitability evaluation for ecotourism using GIS & AHP: a case study of surat thani province, Thailand. Pro. Soci. Beha. Sci. 21, 269–278.
Nagy, V., Stekauerová, V., Neményi, M., Milics, G., Koltai, G., 2007. The role of soil moisture regime in sustainable agriculture in both side of river Danube in 2002 and 2003. Cereal Res. Commun. 35 (2), 821–824.
Naik, G., Suresh, D.N., 2018. Challenges of Creating Sustainable Agri-Retail Chains. IIMB Management rev., pp. 1–28
Nayak, D.C., 2001. Soil Series of West Bengal, vol. 89. National Bureau of Soil Survey and Land Use Planning, Indian Council of Agricultural Research.
Neina, D., 2019. The role of soil pH in plant nutrition and soil remediation. App. Environ. Soil Sci 1–9. Article ID 5794869.
Nyeko, M., 2012. GIS and multi-criteria decision analysis for land use resource planning. J. Geogr. Inf. Syst. 4, 341–348.
Pal, S., Sarkar, D., 2018. Construction of avulsion potential zone model for kulik river of barind tract, India and Bangladesh. Environ. Monit. Assess. 190 (5).
Park, Y.S., Chung, Y.J., Moon, Y.S., 2013. Hazard ratings of pine forests to a pine wilt disease at two spatial scales (individual trees and stands) using self-organizing map and random forest. Ecol. Inf. 13, 40–46.
Prakash, T.N., 2003. Land suitability analysis for agricultural crops: a fuzzy multicriteria decision making, approach. Thesis (MSc). (ITC). In: Int. Inst. Geoinfor. Sci. Earth Observation. Enschede, the Netherlands.
Pramanik, M.K., 2016. Site suitability analysis for agricultural land use of Darjeeling district using AHP and GIS techniques. Mod. Earth Sys. Environ. 2, 56.
Princy, R.J.P., Parthasarathy, S., Jose, P.S.H., Lakshminarayanan, A.R., andJeganathan, S., 2020. Prediction of cardiac disease using supervised machine learning algorithms. Int. Conf. Int. Comp. Control Syst. 570–575.
Rawat, J., Sanwal, P., Saxena, J., 2016. Potassium and its role in sustainable agriculture. Potassium sol. micr. sust. agri. 235–253. https://doi.org/10.1007/978-81-322-2776-2_17.
Richards, L.A., 1954. Diagnosis and Improvement of Saline and Alkali Soils. United Stat. Depart. Agri., Washington., 1969.
Rojas, D., Loubier, J.C., 2017. Analytical Hierarchy Process coupled with GIS for land management purposes: a decision-making application. Int. Cong. Mod. Simula. 1482–1488.
Romeijn, H., Faggian, R., Diogo, V., Sposito, V., 2016. Evaluation of deterministic and complex analytical hierarchy process methods for agricultural land suitability analysis in a changing climate. Int. J. Geo-Inform. 5 (6), 99.
Ross, D.S., Ketterings, Q., 1995. Recommended methods for determining soil cation exchange capacity. Recom. soil test proced. northeastern United States. 493 (101), 62.
Rossiter, D.G., 1996. A theoretical framework for land evaluation. Geoderma 72 (3–4), 165–190.
Saaty, T.L., 1977. A scaling method for priorities in hierarchical structures. J. Math. psych. 15 (3), 234–281.
Saaty, T., 1980. In: Kobe, Japan (Ed.), The Analytic Hierarchy Process (AHP) for Decision Making, pp. 1–69.
Saaty, T.L., 2001. Fundamentals of the analytic hierarchy process. Nat. resource environ. Dec. making. 15–35.
Saha, S., Sarkar, D., Mondal, P., Goswami, S., 2021. GIS and multi-criteria decision-making assessment of sites suitability for agriculture in an anabranching site of sooin river, India. Mod. Earth Syst. Environ. 7 (1), 571–588. https://doi.org/10.1007/s40808-020-00936-1.
Sánchez-Lozano, J.M., Teruel-Solano, J., Soto-Elvira, P.L., García-Cascales, M.S., 2013. Geographical Information Systems (GIS) and Multi-Criteria Decision Making (MCDM) methods for the evaluation of solar farms locations: case study in south-eastern Spain. Ren. Sust. rev. 24, 544–556.
Sarkar, D., Saha, S., Maitra, M., Mondal, P., 2021. Site suitability for Aromatic Rice cultivation by integrating Geo-spatial and Machine learning algorithms in Kaliyaganj C.D. block, India. Art. Int. Geosci. 2, 179–191. https://doi.org/10.1016/j.aiig.2022.03.001.
Sarkar, D., Saha, S., Mondal, P., 2022. GIS-based frequency ratio and Shannon's entropy techniques for flood vulnerability assessment in Patna district, Central Bihar, India. Int. J. Environ. Sci. Tech. 19 (9), 8911–8932. https://doi.org/10.1007/s13762-021-03627-1.
Scherer, L.A., Verburg, P.H., Schulp, C.J.E., 2018. Opportunities for sustainable intensification in European agriculture. Global Environ. Change 48, 43–55.
Schillaci, C., Saia, S., Acutis, M., 2018. Modelling of soil organic carbon in the mediterranean area: a systematic map. Rend. Online Soc. Geol. Ital. 46, 161–166.
Schmidt, S.B., Jensen, P.E., Husted, S., 2016. Manganese deficiency in plants: the impact on photosystem II. Trends Plant Sci. 21 (7), 622–632.
Seth, A., Sarkar, D., Masto, R.E., Batabyal, K., Saha, S., Murmu, S., Das, R., Padhan, D., Mandal, B., 2018. Critical limits of Mehlich 3 extractable phosphorous, potassium, sulfur, boron and zinc in soils for nutrition of rice (Oryza sativa L.). J. Soil Sci. Plant Nutr. 18 (2), 512–523.
Shearer, K.S., Xiang, W.N., 2009. Representing multiple voices in landscape planning: a land suitability assessment study for a park land-banking program in Concord, North Carolina, USA. Landsc. Urban Plann. 93 (2), 111–122.
Shi, G.Y., Zhou, Y., Sang, Y.Q., Huang, H., Zhang, J.S., Meng, P., Cai, L.L., 2021. Modeling the response of negative air ions to environmental factors using multiple linear regression and random forest. Ecol. Inf. 66, 101464.
Srdjevic, B., Medeiros, Y.D.P., 2007. Fuzzy AHP assessment of water management plans. Water Resour. Manag. 22 (7), 877–894.
Steiner, F., McSherry, L., Cohen, J., 2000. Land suitability analysis for the upper Gila River watershed. Landsc. Urban Plann. 50 (4), 199–214.
Thach, N.N., Ngo, D.B.T., Xuan-Canh, P., Hong-Thi, N., Thi, B.H., Nhat-Duc, H., Dieu, T. B., 2018. Spatial pattern assessment of tropical forest fire danger at Thuan Chau area (Vietnam) using GIS-based advanced machine learning algorithms: a comparative study. Ecol. Inf. 46, 74–85.
Tscharntke, T., Clough, Y., Wanger, T.C., Jackson, L., Motzke, I., Perfecto, I., Whitbread, A., 2012. Global food security, biodiversity conservation and the future of agricultural intensification. Biol. Conserv. 151 (1), 53–59.

Turanoglu Bekar, E., Cakmakci, M., Kahraman, C., 2016. Fuzzy COPRAS method for performance measurement in total productive maintenance: a comparative analysis. J. Bus. Econ. Manag. 17 (5), 663–684.
Turgut, B., Özalp, A.Y., Akıncı, H., 2013. Agricultural land use suitability analysis using GIS and AHP technique. Comput. Electron. Agric. 97, 71–82.
van Erp, P.J., Houba, V.J.G., van Beusichem, M.L., 2001. Actual cation exchange capacity of agricultural soils and its relationship with pH and content of organic carbon and clay. Commun. Soil Sci. Plant Anal. 32 (1–2), 19–31. https://doi.org/10.1081/CSS-100102990.
Wang, P., Hafshejani, B.A., Wang, D., 2021. An improved multilayer perceptron approach for detecting sugarcane yield production in IoT based smart agriculture. Microprocess. Microsyst. 82, 103822 https://doi.org/10.1016/j.micpro.2021.103822.
Wheeler, A.P., Steenbeek, W., 2021. Mapping the risk terrain for crime using machine learning. J. Quant. Criminol. 37 (2), 445–480.
Yalew, S.G., Griensven, A.V., Mul, M.L., Zaag, P.V., 2016. Land suitability analysis for agriculture in the Abbay basin using remote sensing, GIS AHP Techniques. Mod. Earth Syst. Environ. 2, 1–14.
Zaredar, N., Jafari, S., 2010. Land Suityability analysis using multi attribute decision making approach. Int. J. Environ. Sustain Dev. 1 (5), 2010–2064.
Zavadskas, E.K., Kaklauskas, A., 1996. Pastatųsistemotechninisįvertinimas (Multiple Criteria Evaluation of Buildings). Technika, Vilnius.
Zhang, K., Achari, G., 2010. Uncertainty propagation in environmental decision making using random sets. Procedia. Environ. Sci. 2, 576–584.