Centre for Research on Settlements and Urbanism

Journal of Settlements and Spatial Planning

Journal homepage: https://jssp.reviste.ubbcluj.ro

Unravelling the Role of Socio-Physical Drivers for Potential Built-up Site Selection in the Kumaun Himalayas Using GIS-Based Fuzzy-AHP and Machine Learning

Akash TIWARI[1], Manish KUMAR*[1], Syed Irtiza MAJID[1], Sourav BHADWAL[1], Naresh Kumar VERMA[2], Dinesh Kumar TRIPATHI[3], Subhash ANAND[4]

** Corresponding author*

[1] *Central University of Haryana, School of Basic Sciences, Department of Geography, Jant-Pali, Mahendragarh, Haryana, INDIA*

[2] *Jawaharlal Nehru University, Special Centre for National Security Studies, New Delhi, INDIA*

[3] *Rana Pratap Post Graduate College, Sultanpur, Uttar Pradesh, INDIA*

[4] *University of Delhi, Delhi School of Economics, Department of Geography, Delhi, INDIA*

researchakashgeo@gmail.com 0000-0003-0066-781X

manish.ks1@gmail.com 0000-0002-2588-6499

sirtiza34@gmail.com 0000-0002-7889-4281

bhadwalsourav@gmail.com 0000-0002-9555-7957

nareshrai.jnu@gmail.com 0009-0001-8621-2615

tripathidk.geoinformatics@gmail.com 0000-0001-8274-5118

sanandpvs@gmail.com 0000-0001-8274-5118

DOI: https://doi.org/10.24193/JSSP.2024.1.03

Received: 05 November 2023

Received in revised form: 21 April 2024

Accepted for publication: 11 June 2024

Available online: 18 June 2024

Keywords: site suitability analysis, Fuzzy-AHP, support vector machine, Bayesian Network, Kumaun Himalayas

ABSTRACT

Rapid and uncontrolled urban growth in the Kumaun Himalayas in absence of proper land use policy has pushed built-up areas towards the tectonically and ecologically sensitive regions, reducing the availability of suitable built-up land while simultaneously increasing the vulnerability of both communities and environment. The identification of areas for sustainable built-up growth is of paramount importance to address the challenges arising from unregulated urban expansion. In this study GIS-based Fuzzy-AHP technique and machine learning algorithms (SVM and BN) were employed to delineate the potential built-up sites selection in Hawalbagh Block, Uttarakhand (India) using nine socio-physical drivers, including slope, aspect, LU/LC, distance to road, distance to drainage, distance to lineament, distance to landslide, distance to settlement, and lithology. The suitability maps generated by the three methods were validated using AU-ROC analysis, which demonstrated that each approach produces outstanding results with AU-ROC values more than 0.90. The comparison of the approaches shows that SVM (AUROC=0.99) outperforms BN (0.95) and GIS-based Fuzzy-AHP (0.90). The suitability maps were classified into five suitability classes. Assuming that very high and high suitability classes are acceptable for built-up expansion, the study identified potential built-up locations in the study region covering an area of 148.86 km^2, 85.23 km^2, and 55.25 km^2 according to the Fuzzy-AHP technique, SVM model, and BN model, respectively. The suitability zonation in this study can serve as a foundation for the development of land-use policy or the formulation of master plans aimed at achieving a sustainable mountain ecology in the Kumaun Himalayas.

1. INTRODUCTION

The Himalaya is recognized as the most densely populated and rapidly urbanized mountain ecosystem in the world. Compared to other Himalayan states in India, Uttarakhand has the largest proportion of urban population (approximately 30%), which is higher than the average of 25% for all Himalayan states. During the last twenty years, the built-up areas of Uttarakhand towns have witnessed a 33 percent increase. However, due to the absence of an official land-use policy, more than 60 percent of this construction is concentrated in environmentally unsafe zones, specifically areas that are prone to landslides and floods (Tiwari et al., 2023).

This rapid urbanization or built-up growth is driven by various factors, including the increasing trend of rural outmigration, improved road connectivity, the presence of rural markets, remarkable tourism growth etc. However, the construction of buildings or infrastructures on slopes with inclinations ranging from 35 to 45 degrees often neglects essential safety considerations, leading to significant devastation during earthquakes and heavy rainfall (Xie et al., 2003). This devastation is further intensified due to the alignment of the constructed houses, causing a chain effect of damage.

The lack of proper planning in expanding settlements creates regional disparities and associated issues, emphasizing the importance of identifying suitable sites for comprehensive development and the sustainable use of available land resources (Chen and Wu, 2009; Kumar and Shaikh, 2013; Madu, 2007). In mountainous provinces, land suitability for built-up expansion is primarily determined by physical factors and is a decisive concern for land-use planning and management (Kumar and Biswas, 2013). The availability of land for construction in mountainous regions is already limited due to topographical and geographical constraints. Therefore, conducting built-up site suitability analysis plays a critical role in identifying the most appropriate sites while ignoring less suitable and unusable areas within a given region.

Over the past few years, data-driven models such as Machine Learning (ML), Multi-Criteria Decision-Making (MCDM), Soft Computing (SC), Computational Intelligence (CI), Data Mining (DM), Intelligent Data Analysis (IDA) and Knowledge Discovery in Databases (KDD) have demonstrated considerable potential in various fields of predictive sciences (Solomatine and Shrestha, 2009). In recent decades, the integration of various physical factors in analyses of built-up suitability has led to the increased utilization of quantitative methodologies, particularly MCDM techniques (Arciniegas et al., 2011; Chang et al., 2008; Chen et al., 2010; Greene et al., 2010; Kordi and Brandt, 2012). MCDM techniques offer the advantage of addressing the challenges involved in analysing multiple criteria or drivers. They enable the efficient processing of large volumes of data and complex information, making them valuable tools in the decision-making process (Zopounidis and Doumpos, 2002; Zopounidis and Pardalos, 2010).

The utilization of MCDM techniques in geographical contexts has the ability to enhance the efficiency and investigative robustness of land-use decision-making processes (Dunning et al., 2000; Hajkowicz and Collins, 2007). The Analytic Hierarchy Process (AHP), also known as pair-wise comparison method, designed by Saaty, 1984, is recognized as a highly effective and widely adopted approach that significantly reduces time and effort. AHP, as a widely employed method falls under the category of Multi-Criteria Decision Analysis (Thill, 2019). It is utilized for analysing intricate alternatives with multiple attributes and conflicting objectives involving multiple stakeholders. This approach may give rise to incongruities when making comparisons across pairs, hence posing difficulties in accurately articulating the specific preference inside the comparison matrix (Cheng, 1997; Das and Pal, 2019; Kahraman et al., 2003; Saaty and Tran, 2007). To address this limitation, the present research utilizes the Fuzzy Analytical Hierarchy Process (Fuzzy-AHP), which was introduced by Van Laarhoven and Pedrycz in 1983. To leverage the benefits of both fuzzy logic and AHP, the concept of fuzzy logic is incorporated into the AHP (Li et al., 2009). Fuzzy-AHP combines the AHP with fuzzy set theory to solve issues encountered in MCDM. The Fuzzy-AHP method adopts fuzzy numbers instead of precise numeric values to derive the results. Fuzzy set theory is used to allocate weights to each driver and its sub-classes (Das and Pal, 2019; Mohamed and Elmahdy, 2017; Şener et al., 2018; Tan et al., 2014). Many researchers have applied the Fuzzy-AHP method in their studies to address the MCDM issues (Das and Pal, 2019; Şener et al., 2018). The Fuzzy-AHP procedure proves highly effective in determining weight information by considering both fuzziness and uncertainty, which cannot be achieved solely by AHP (Chaudhry et al., 2021).

With the advances in computational programming, the application of AI in geostatistical analyses particularly in site suitability analysis has recently gained more attention than other methods (Almansi et al., 2021; Gharaibeh et al., 2023; Sarkar et al., 2021). This approach has garnered greater attention compared to other methods due to its ability to overcome the limitations of conventional statistical analysis. Additionally, AI enhances computational efficiency when dealing with extensive data sets involving multiple attributes in complex phenomena. ML algorithms such as Support Vector Machine (SVM), Logistic Regression (LR), Bayesian Network (BN), Artificial Neural Network (ANN), Boosted Regression Tree (BRT), and Multivariate Adaptive Regression Splines (MARS) have been widely used in flood susceptibility mapping (Mojaddadi Rizeei, 2018; Singh and Singh, 2017; Tehrany et al., 2015, 2019; Xiong et al., 2019), landslide susceptibility mapping (Pandey et al., 2020; Pham et al., 2017; Tien Bui et al., 2016) and land suitability analysis (Ornella and Tapia, 2010), wildlife habitat suitability (Hamilton et al., 2015; Johnson et al., 2010; Tantipisanuh et al., 2014). In this study, two ML algorithms, viz. SVM and Bayesian Network (BN) were selected by auto-classifier in Statistical Package for Social Sciences (SPSS) Modeler for delineating the built-up site suitability. The auto classifier comparatively estimates the models for nominal (set) or binary (yes/no) target data and provides the best three models (S. Singh et al., 2022; Verma et al., 2018, 2019). SVM is one of the well-known and powerful classifiers stated by Cortes and Vapnik (1995).

The advantages of SVM lie in its ability to classify input data with optimum output accuracy, and the flexibility of the process to optimize the classification through hyper parameter tuning and cross-validation. BN are Directed Acyclic Graphs (DAG) which connect variables using conditional probabilities based on Bayes' Theorem (Fenton and Neil, 2018; Koski and Noble, 2011). The first portrayal by BN effectively illustrates the challenges, thus facilitating a deeper comprehension and broader perspective on the ambiguities and intricacies involved (Marcot and Penman, 2019). This study intends to check the predictive precision of the MCDM technique and compare them with the ML algorithms, SVM and BN. The objective of this study was to conduct a built-up site suitability analysis in the relatively more inhabited Kumaun Himalayas with a case study of Hawalbagh Block in Uttarakhand, India, considering nine socio-physical drivers, namely slope, aspect, LU/LC, distance to road, distance to drainage, distance to lineament, distance to landslide, distance to settlement, and lithology. In this regard, MCDM and ML models were used. The validation and comparison of the used models were conducted using area under the receiver operating characteristic curve (AUROC) analysis method.

2. THEORY AND METHODOLOGY

This study aims to undertake a built-up site suitability analysis for the region of Hawalbagh Block, Uttarakhand, India. It also tries to appraise the predicted accuracy of the MCDM technique and compare it with two popular machine learning algorithms, SVM and BN, for the better estimation of site suitability analysis. The methodology implemented to fulfil the aims of the study is outlined in Figure 1.

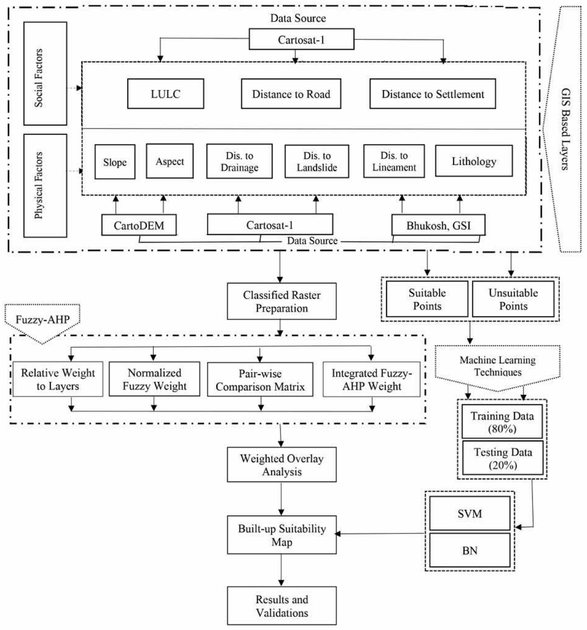

Fig. 1. Flowchart of the methodology.

2.1. Study area

The Hawalbagh Block, situated within the Almora district of Uttarakhand, India, is recognized as one of the development blocks in the state. The study area (Fig. 2) geographically spans from 29°32'30" N to 29°44'23" N latitude and 79°31'11" E to 79°43'50" E longitude.

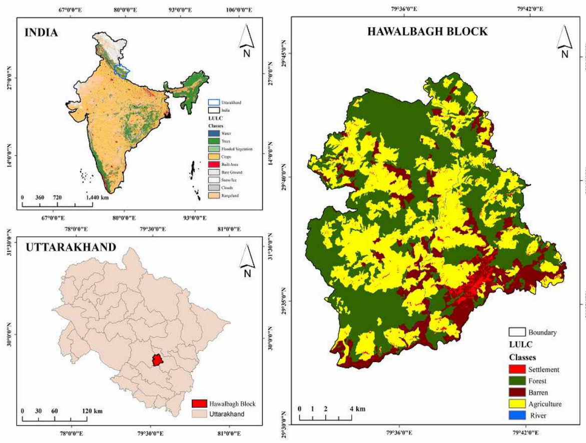

Fig. 2. Location map of the study area.

This block is located in the Lesser Himalayas at an elevation of 1405 m (above MSL), covering an area of 267.53 km². It is positioned within the Kumaun Himalayas region of Uttarakhand. In terms of lithology, the central and southern parts of the Hawalbagh Block contain rocks from the 'Almora Group', while the northern part consists of rocks from the 'Damtha Group' (Valdiya, 1980). The region is traversed by the main channel and tributaries of the Kosi River System. Hawalbagh Block experiences a cool temperate climate with an average annual temperature of 20°C and an average rainfall of 1095 mm per year.

The land use/land cover (LU/LC) of Hawalbagh Block has undergone modifications influenced by both natural and human factors across different temporal and spatial dimensions (Rawat and Kumar, 2015). The region exhibits a substantial extent of vegetation cover (146.49 km²), followed by agricultural land (84.73 km²), barren land (31.17 km²), built-up land (2.72 km²), and water bodies (2.42 km²). The relatively low coverage of built-up areas (1.01% of the total area) can be attributed to the physiographic constraints of the region, which limit the suitability of a large portion of the study area for built-up development. Additionally, the study area's location in the Kumaun Himalayas, characterized as a young and tectonically active fold mountain system, presents additional challenges for built-up growth.

2.2. Preparation of the spatial database

This study focuses on analysing the site suitability for built-up expansion in the mountainous landscape of the Kumaun Himalayas, specifically in the Hawalbagh Block of Uttarakhand, India. This analysis considered nine socio-physical drivers that have a substantial effect on determining the site suitability for built-up areas. The data layers, sources and specifications of these socio-physical drivers are provided in Table 1.

Table 1. Databases: data layer, sources and specifications.

Data layer		Data source	Source specification
Physical factors	Slope (2020)	CartoDEM	Simultaneous capture of two images, one taken by the forward-facing camera (PAN FORE) at +26° and the other by the aft-facing camera (PAN AFT) at -5°, to obtain near-instantaneous stereo data Spatial resolution: 2.5 m Stereo Strip Triangulation of 500×27 km Base-to-height Ratio: 0.63
	Aspect (2020)		
	Distance to Drainage (2020)	Bhukosh, GSI	Scale: 1:50,000
	Distance to Landslide (2020)		
	Distance to Lineament (2020)	Cartosat-1	Spatial Resolution: 2.5 m Spectral Range: 0.5 - 0.85 µ Radiometric resolution: 10 bit Temporal Resolution: 5 days
	Lithology (2020)		
Social factors	LU/LC (2020)		
	Distance to Road (2020)		
	Distance to Settlement (2020)		
Ancillary data	Base map, Block boundary (2020)	SOI Toposheet	Topographical sheet No. 53O/10 Scale: 1:50,000

2.2.1. Slope

Slope refers to the angular gradient of the terrain from the hilltop to the valley bottom (Kumar et al., 2023). The slope map of the study area (Fig. 3(a)) was generated using CartoDEM in ArcGIS software, ranging between 0˚ and 75.3˚. Typically, flat, and stable ground is considered more suitable for built-up development. To align with this approach, small fuzzy membership function has been employed to standardize slope map of the study area, assigning higher weightage to areas with lesser slope and vice versa (Fig. 4(a)).

2.2.2. Aspect

Aspect, or slope aspect denotes the direction in which a mountain flank is oriented. An aspect map of the study area was created using CartoDEM in ArcGIS software, as depicted in Figure 3(b). In the northern hemisphere, southward slopes receive greater solar radiation than northward slopes. Therefore, southern slopes are considered more favorable for settlement. Moreover, it is worth noting that slopes facing east are exposed to sunlight in the morning, a time when temperatures tend to be lower. Conversely, slopes facing west get sunshine in the second half of the day, a period when temperatures often rise. Consequently, the eastern slope tends to be cooler, while the western slope tends to be warmer, making the western slopes more suitable for human habitation. The aspect map of the Hawalbagh Block was standardized using Gaussian fuzzy membership function. This function assigns higher weightage to intermediate aspect values and relatively lower weightage to lower and higher aspect values (Fig. 4(b)).

2.2.3. Land Use/Land Cover map

Global environmental change is significantly influenced by anthropogenic activities, whether directly or indirectly. Understanding the present landscape and assessing the potential for development and regional planning requires a comprehensive LU/LC map. In this context, on-screen digitization of Cartosat-1 imagery was performed using ArcGIS software to create an LU/LC map of the Hawalbagh Block. Figure 3(c) illustrates the five LU/LC categories in the study area, namely settlement, forest, barren land, agriculture, and rivers. The forest and agricultural land make up the majority of the area, accounting for 46% of the total. To standardize the LU/LC map, linear membership function was employed. This function assigns higher weightage to higher values and vice versa, ensuring a consistent representation of the data (Fig. 4(c)).

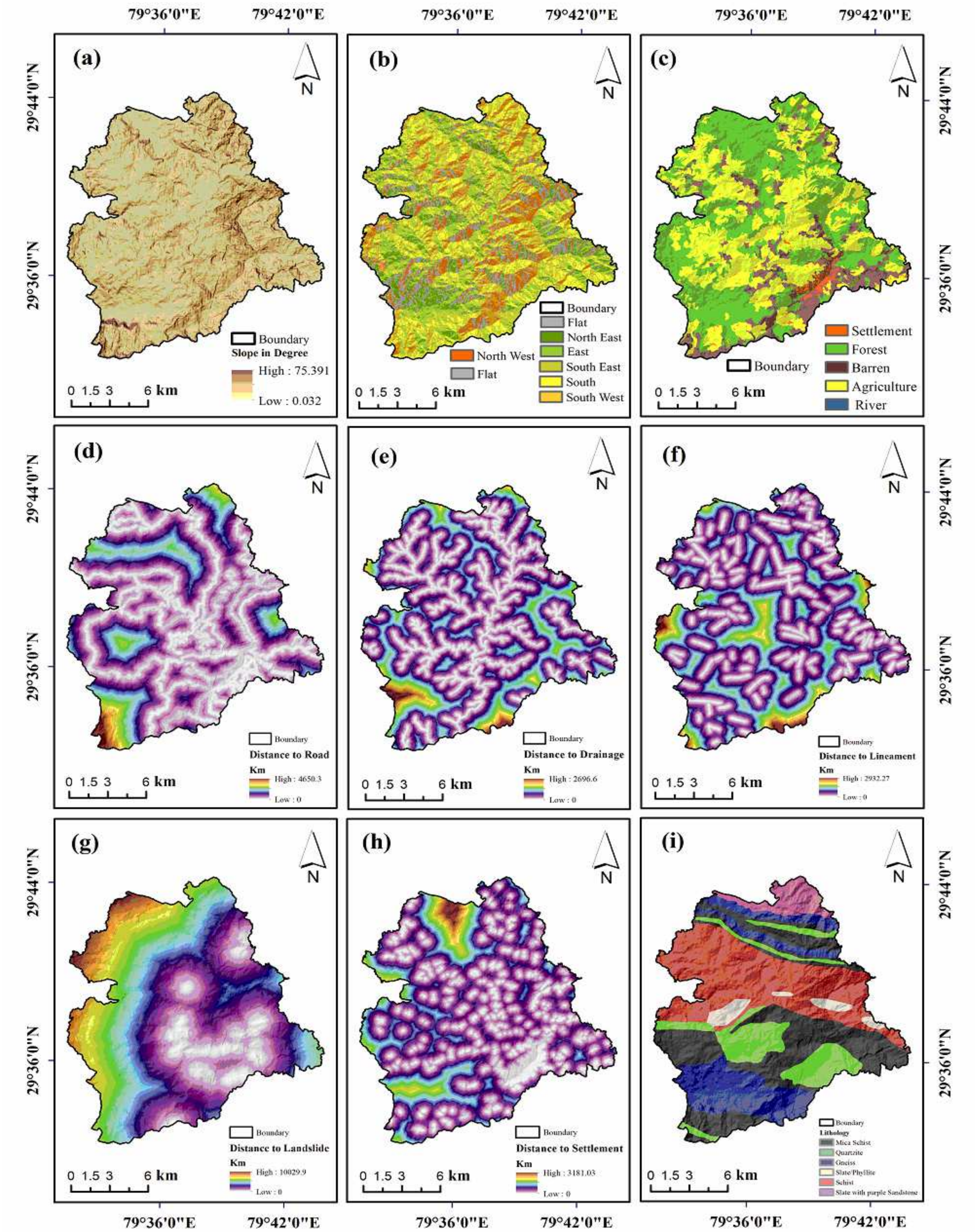

Fig. 3. Built-up suitability drivers: (a) slope; (b) aspect; (c) LU/LC; (d) distance to road; (e) distance to drainage; (f) distance to lineament; (g) distance to landslide; (h) distance to settlement; (i) lithology.

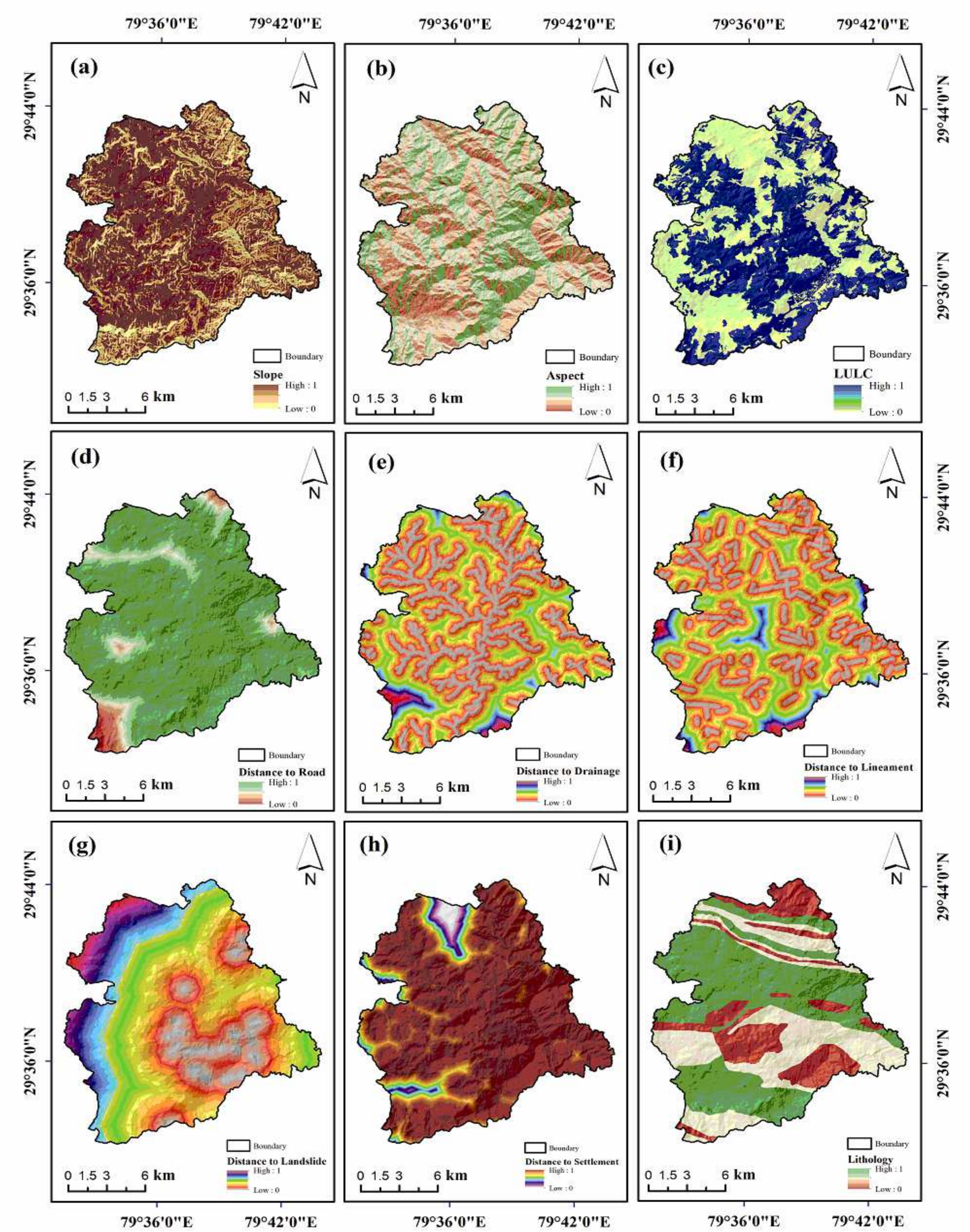

Fig. 4. Standardized built-up suitability drivers: (a) slope; (b) aspect; (c) LU/LC; (d) distance to road; (e) distance to drainage; (f) distance to lineament; (g) distance to landslide; (h) distance to settlement; (i) lithology.

2.2.4. Distance to road

Road infrastructure plays a crucial role in urban development as it not only connects major cities, towns, and villages but also enables access to essential goods and services such as education, healthcare, and employment opportunities. Consequently, the suitability of sites for built-up development is inversely related to the distance from transportation links, particularly roads. The distance to road map of Hawalbagh Block is shown in Figure 3(d). To establish the relationship between built-up site suitability and the distance to roads, small fuzzy membership function has been used to standardize distance to road map of Hawalbagh Block, Almora, Uttarakhand, India (Fig. 4(d)).

2.2.5. Distance to drainage

In hilly areas, the proximity to drainage plays a critical role due to its association with increased runoff, which can lead to landslides and flooding. Previous studies (Gökceoglu, 2001; Gökceoglu and Aksoy, 1996; Pachauri et al., 1998) have demonstrated the importance of proximity to drainage lines in controlling the occurrence of landslides, particularly in relation to intensive gully erosion. With these factors in mind, a drainage proximity map (Fig. 3(e)) was generated to assess the suitability. Areal extents near drainage are generally measured to be unsafe for built-up. Therefore, a large fuzzy membership function has been employed to standardize the distance to drainage map (Fig. 4(e)).

2.2.6. Distance to lineament

Lineament pertains to the linear faults of earth's crust, serving as indicators of regions where the bedrock is relatively less resistant (Manjare, 2014). The stability of the bedrock tends to increase as the distance from the lineament increases. Therefore, sites located away from lineaments are generally considered more suitable for built-up development, while those in closer proximity are less desirable. The lineament map (Fig. 3(f)) of Hawalbagh Block was taken from the Geological Survey of India and standardized using large fuzzy membership function (Fig. 4(f)).

2.2.7. Distance to landslide

Regions in proximity to areas prone to landslides are unsuitable for urban development, whereas locations situated farther from these susceptible zones are more suitable for construction. In the case of Hawalbagh Block, the distance to landslide-prone areas (Fig. 3(g)) was standardized using large fuzzy membership function (Fig. 4(g)).

2.2.8. Distance to settlement land or built-up land

Humans, as social beings, strive to live in communities that provide a sense of neighborhood, public spaces, civic amenities, basic infrastructure, and easy accessibility to a range of services. Consequently, developed areas tend to attract new settlements in their vicinity. However, in this study, built-up areas are deemed unsuitable for future development due to the long lifespan of constructed buildings, typically lasting for a minimum of 50-75 years (Kumar and Shaikh, 2013). To assess the distance to settlement land ((Fig. 3(h)), the LU/LC map of Hawalbagh Block was utilized. Subsequently, small fuzzy membership function has been applied to standardize distance to settlement land (Fig. 4(h)).

2.2.9. Lithology map

The study area encompasses six distinct lithological categories, namely mica schist, quartzite, gneiss, slate/phyllite, schist, and slate, with purple sandstone depicted in Figure 3(i). Hard rock lithology is generally regarded as more stable and therefore more suitable for built-up development compared to soft rock lithology. In line with this understanding, the lithological map was reclassified and standardized using linear fuzzy membership function. Figure 4(i) represents the standardized map of Hawalbagh Block's lithology.

2.3. Application of GIS-based Fuzzy-AHP Technique

In this study, the GIS-based Fuzzy-AHP MCDM technique was employed. The Analytic Hierarchy Process (AHP) is widely recognized as one of the most important techniques for analysing multiple criteria (Saaty, 1984). However, AHP involves uncertainty associated with expert judgments. To address this uncertainty, Fuzzy-AHP was introduced. Fuzzy-AHP aims to clarify uncertainty by generating a scale of values, allowing experts to select the value that corresponds to their level of confidence. A rigorous approach was adopted by incorporating experts' ideas from diversified fields (e.g. urban planning, environmental science, civil engineering, geology, data science and sociology etc.) to enhance the strength and effectiveness of the Fuzzy-AHP method.

To assess the fuzziness of the drivers associated with the built-up site selection, Triangular Fuzzy Numbers (TFN) were utilized in this study. TFN are the fuzzy numbers in which the membership values are determined by a trio of real values i.e., l, m, and u (Deng, 1999; Wu et al., 2004). Mathematically, TFN is described by Cox, 1995 in equation (1);

$$\mu_m(x) = \begin{cases} \dfrac{(x-l)}{(m-l)}, l \le x \le m \\ \dfrac{(u-x)}{(u-m)}, m \le x \le u \\ 0, \textit{otherwise} \end{cases} \tag{1}$$

The real values *l, m* and *u* correspond to the lower, middle, and upper limit of TFN respectively. In this study, Fuzzy Extent Analysis (Chang, 1996) was employed to calculate the weight for each criterion for it is a computationally simple procedure. The linguistic variables and their respective TFN (Torabi-Kaveh et al., 2016) are given in Table 2. Based on their knowledge, the experts assigned suitable rank to each criterion to develop the fuzzy suitability matrix.

Table 2. Triangular fuzzy number of linguistic variables *(source: Torabi-Kaveh et al., 2016)*.

Linguistic Variables	Triangular Fuzzy Numbers	Reciprocal TFN
Extremely Strong	(9, 9, 9)	(1/9, 1/9, 1/9)
Very Strong	(6, 7, 8)	(1/8, 1/7, 1/6)
Strong	(4, 5, 6)	(1/6, 1/5, 1/4)
Moderately Strong	(2, 3, 4)	(1/4, 1/3, 1/2)
Equally Strong	(7, 8, 9); (5, 6, 7); (3, 4, 5); (1, 2, 3)	(1/9, 1/8, 1/7); (1/7, 1/6, 1/5); (1/5, 1/4, 1/3); (1/3, 1/2, 1)

Subsequently, fuzzy synthetic extent for the i^th object was calculated based on fuzzy suitability matrix. Normalized fuzzy weights for each criterion were calculated with the help of equations (2) and (3).

$$S_i = \sum_{j=1}^{n} \overset{\cap}{a}_{ij} \otimes \left[\sum_{i=1}^{n} \sum_{j=1}^{n} \overset{\cap}{a}_{ij} \right]^{-1} \quad (2)$$

$$V(M_2 \geq M_1) = \sup_{y \geq x}\left[\min(\mu_{M_1}(x), \mu_{M_2}(y))\right] \quad (3)$$

where, M_1 (l_1, m_1, n_1) and M_2 (l_2, m_2, n_2) are triangular fuzzy numbers. The likelihood of a convex fuzzy number being larger than k other convex fuzzy numbers, M_i (i = 1, 2 .... k) can be assessed using equation (4):

$$V(M \geq M_1, M_2, \ldots M_k) = V[(M \geq M_1), (M \geq M_2), \ldots (M \geq M_k)] = \min V(M \geq M_i), i=1, 2, \ldots, k \quad (4)$$

Assuming, $d'(A_i) = \min V(S_i \geq S_k)$

For k = 1, 2,....., n; k≠i. The weight vector (W') and normalized weight vector (W) are specified by equation (5) and equation (6), respectively:

$$W' = (d'(A_1), d'(A_2), \ldots, d'(A_n))^T \quad (5)$$

$$W = \left(\frac{d'(A_1), d'(A_2), \ldots, d'(A_n)}{\sum_{i=1}^{n} d'(A_i)} \right) \quad (6)$$

where, appropriateness of the object is represented by A_i (i = 1, 2,........, n).

The final suitability map was prepared using the weighted sum of all the standardised conditioning socio-physical drivers according to equation (7):

$$\sum_{i=1}^{n} X_i W_i \quad (7)$$

where, X_i is the i^th standardised conditioning factor and W_i is its relative Fuzzy-AHP weight.

2.4. Application of Machine Learning Algorithms

In the present study, two ML algorithms, SVM and BN, have been utilized for the analysis of built-up site suitability. For training these ML algorithms, 2000 point locations suitable for built-up usage were identified based on the slope map, while 2000 point locations unsuitable for built-up use were identified using a distance to drainage map. These two datasets were merged into a single dataset, where the IDs 1 and 0 were assigned to represent suitable and unsuitable locations, respectively. The socio-physical driver attributes (independent variables) for each of the 4000-point locations (dependent variable) were extracted in ArcGIS software and compiled in Excel format. The dataset was fragmented into training and testing subsets in the ratio 80% (3200 Points) and 20% (800 Points), respectively. The application of the ML algorithms for built-up site suitability analysis in this study was conducted using the SPSS Modeler 18.0.

2.4.1. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised ML algorithm used for built-up site suitability analysis in this study. The selection of the kernel plays a crucial role in obtaining accurate results when using the SVM model. Kernel functions in SVM are utilized to transform the input data into a higher-dimensional feature space, enabling the handling of complex non-linear relationships (Naghibi et al., 2017). Four commonly used kernel functions in SVM are the Radial Basis Function (RBF), Linear Kernel, Polynomial Kernel, and Sigmoid Kernel. Among these, the RBF kernel function is generally favoured due to its high learning capacity (Abdollahi et al., 2019). Therefore, the RBF kernel was chosen for this study. Another significant step in applying the SVM model is optimizing the kernel parameters or hyper-parameter tuning to enhance the efficiency of the model. The most employed method for hyper-parameter tuning is k-fold cross-validation, which was also applied in this study.

2.4.2. Bayesian Networks (BN)

Bayesian Networks (BN), alternatively referred to as Bayesian Belief Networks (BBN), are a machine learning algorithm based on the graphical structure and Bayes' theorem. The BN framework is widely accepted for representing subjective or objective uncertain knowledge (Abebe et al., 2018; Guarnieri et al., 2015). It encompasses both qualitative and quantitative components (Chen et al., 2019). As mentioned earlier in the introduction, the Directed Acyclic Graph (DAG) represents the qualitative component, while sets of conditional probabilities constitute the quantitative component of the BN model. Equation (8) represents the complete BN model:

$$N = \langle G, P \rangle \quad (8)$$

where, G (G = <V, E>) represents the structure graph of the BN. Nodes which depict the criteria for the site-suitability analysis are represented by V, whereas casual dependencies between the parent nodes and child nodes is represented by E. Conditional Probability Distribution (CPD) of each parameter is represented by P.

2.5. Synthesis of Final Suitability Maps

Fuzzy-AHP, as discussed previously, is used to assign the weights to the socio-physical drivers and prepare their raster layers. The raster layers are integrated with the help of overlay tools to synthesize the final suitability maps. Several overlay methods are provided in GIS environment for synthesizing the final maps from the multiple raster layers of socio-physical drivers. In this study, weighted overlay was used to obtain the final built-up suitability map. ML algorithms are self-sufficient as they undergo training to understand the influence of each socio-physical driver on built-up site suitability and apply that information to integrate the data from multiple socio-physical drivers and produce the probability range of suitability of each unknown location. Resultant point data are converted into raster on ArcGIS software.

After synthesizing the built-up suitability maps from the GIS-based Fuzzy-AHP technique and ML algorithms, quantile method was used to classify the final output raster into five categories i.e., Very High, High, Medium, Low and Very Low. The quantile method is an effective classification technique that evaluates the pixels of raster values present in the input layer, to statistically determine the equivalent representation for each corresponding class (Mayfield, 2015).

2.6. Validation using the Area Under the Receiver Operating Characteristic Curve (AUROC) Analysis

Assessing the accuracy and precision of a model before using it for planning and management is essential to ensure ethical practices (Sarkar and Mondal, 2019). To evaluate accuracy, various researchers utilize Area under the Receiver Operating Characteristic Curve (AUROC), which considers several parameters including sensitivity, specificity, and area under the curve (Maity et al., 2022; Sarkar et al., 2021; Wu et al., 2019).

In this study, the GIS-based Fuzzy-AHP, SVM, and BN were validated by generating an AUROC curve using training and testing data.

Table 3. Fuzzy-AHP weight for socio-physical drivers.

Socio-Physical Drivers	Fuzzy-AHP Weight
Slope	0.24
Distance to Road	0.21
LU/LC	0.17
Lithology	0.18
Distance to Drainage	0.01
Aspect	0.00
Distance to Lineament	0.00
Distance to Landslide	0.13
Distance from Developed Land	0.17

3. RESULTS

The final built-up suitability map was created using the MCDM and ML approaches, considering the nine criteria outlined earlier. It is represented in Fig. 5. The visualization of the amount of constructions is depicted using Google Earth Pro imagery across three distinct zones within the study area, as illustrated in Figure 5.

3.1. Performance of Fuzzy-AHP as Multi-Criteria Decision-Making Model for Site Suitability Analysis

All nine socio-physical drivers were assigned normalized fuzzy weights based on their respective importance, as calculated using Equation (6). The Fuzzy-AHP weights for each driver are presented in Table 3. Among the drivers, slope received the highest weightage (0.24), indicating its significant influence. On the other hand, distance to drainage, aspect, and distance to lineament were assigned a normalized fuzzy weight of zero, suggesting their negligible impact in the analysis.

The final suitability map (Fig. 5(a)) was prepared by combining the weighted values of all socio-physical drivers. Based on this map, the Hawalbagh Block was divided into five categories representing different levels of suitability for built-up use. These categories include very high suitability, covering an area of 64 km^2 (24.15% of the total area), high suitability, covering 84.86 km^2 (31.72% of the total area), medium suitability, covering 62.80 km^2 (23.47% of the total area), less suitability, covering 42.87 km² (16.03% of the total area), and very low suitability, covering 12.39 km² (4.6% of the total area). The suitable areas determined using the Fuzzy-AHP approach can be found in Table 4.

Table 4. Model wise suitability area.

Class	BN		SVM		Fuzzy-AHP	
	Area (km^2)	Area (%)	Area (km^2)	Area (%)	Area (km^2)	Area (%)
Very Low	151.15	56.50	50.99	19.06	12.39	4.63
Low	28.84	10.78	67.30	25.16	42.87	16.03
Medium	32.31	12.08	64.01	23.93	62.80	23.47
High	27.58	10.31	52.17	19.50	84.86	31.72
Very High	27.67	10.34	33.06	12.36	64.60	24.15

3.2. Performance of Machine Learning Algorithms for Site Suitability Analysis

3.2.1. Support Vector Machine (SVM)

SVM, a powerful ML algorithm, has been extensively utilized in site suitability analysis. In this study, all nine socio-physical drivers were used to create the final built-up suitability map. By applying the SVM model with an RBF kernel, the relative importance and influence of these factors were determined. Among them, slope was found to have the highest predictor importance, accounting for 88% of the influence in determining built-up site suitability. The site suitability map generated from the SVM model is presented in Fig. 5(b), displaying five distinct built-up site suitability classes. These classes include areas with very high suitability, covering 33.06 km² (12.36% of the total area), high suitability, covering 52.17 km² (19.50% of the total area), medium suitability, covering 64.01 km² (23.93% of the total area), low suitability, covering 67.30 km² (25.16% of the total area), and very low suitability, encompassing approximately 51 km² (19.06% of the total area). The suitable areas determined through the SVM model can be found in Table 4.

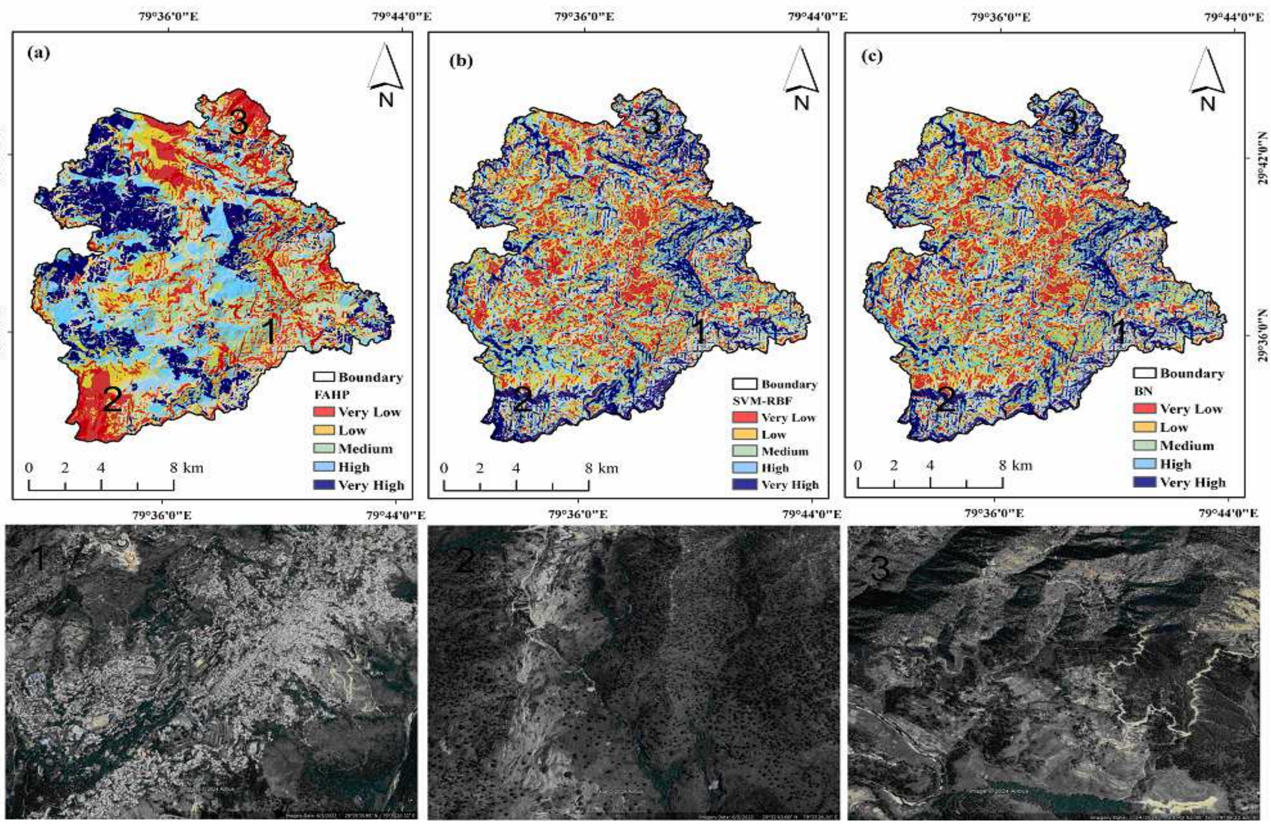

Fig. 5. Built-up suitability map: (a) GIS-based Fuzzy-AHP, (b) SVM, (c) BN.

3.2.2. Bayesian Network (BN)

The BN model serves as a graphical representation of probabilities, depicting a set of variables and their conditional interdependencies through a Directed Acyclic Graph (DAG). In the DAG, edges represent conditional dependencies, while nodes represent individual random variables. The DAG for the suitability analysis is illustrated in Figure 6. Among the nine conditioning factors, the highest percentage influence of 20% in determining built-up suitability was assigned to slope. The site suitability map resulting from the BN model is presented in Figure 5(c), which delineates five distinct built-up site suitability classes. These classes include areas with very high suitability, covering 27.67 km² (10.34% of the total area), high suitability, covering 27.58 km² (10.31% of the total area), medium suitability, covering 32.31 km² (12.08% of the total area), less suitability, covering 28.84 km² (10.78% of the total area), and very low suitability, encompassing 151.15 km² (56.50% of the total area). The suitable areas determined through the BN approach can be found in Table 4.

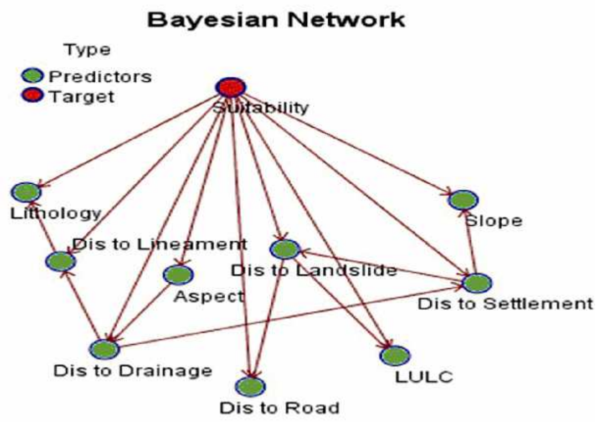

Fig. 6. DAG for the suitability analysis.

3.3. Model validation and accuracy assessment

The AUROC analysis is commonly employed to validate spatially modelled maps. This method utilizes the receiver operating characteristic curve, which represents sensitivity and specificity of the model. 1-specificity and sensitivity values are plotted on the x-axis and y-axis, respectively (Shahabi and Hashim, 2015). The area under the curve provides a quantitative measure for predicting the occurrence or non-occurrence of an event and ranges between 0.5 and 1. A value closer to 1 indicates high accuracy, while a value around 0.5 signifies lower accuracy (Shirzadi et al., 2017).

The AUROC curves generated in this study illustrate the validation of the employed methodologies. Figure 7(a) represents the success rate curve obtained from the training dataset, while Figure 7(b) displays the prediction rate curve obtained from the testing dataset. These curves demonstrate that all three models converge near the top left corner of the graph, indicating a high level of overall accuracy. Among the models, SVM exhibits the highest success rate with an AUROC value of 0.99, surpassing BN with an AUROC of 0.96 and Fuzzy-AHP with an AUROC of 0.91. Similarly, in terms of prediction rate, SVM performs the best with an AUROC of 0.99, followed by BN with an AUROC of 0.95 and Fuzzy-AHP with an AUROC of 0.90.

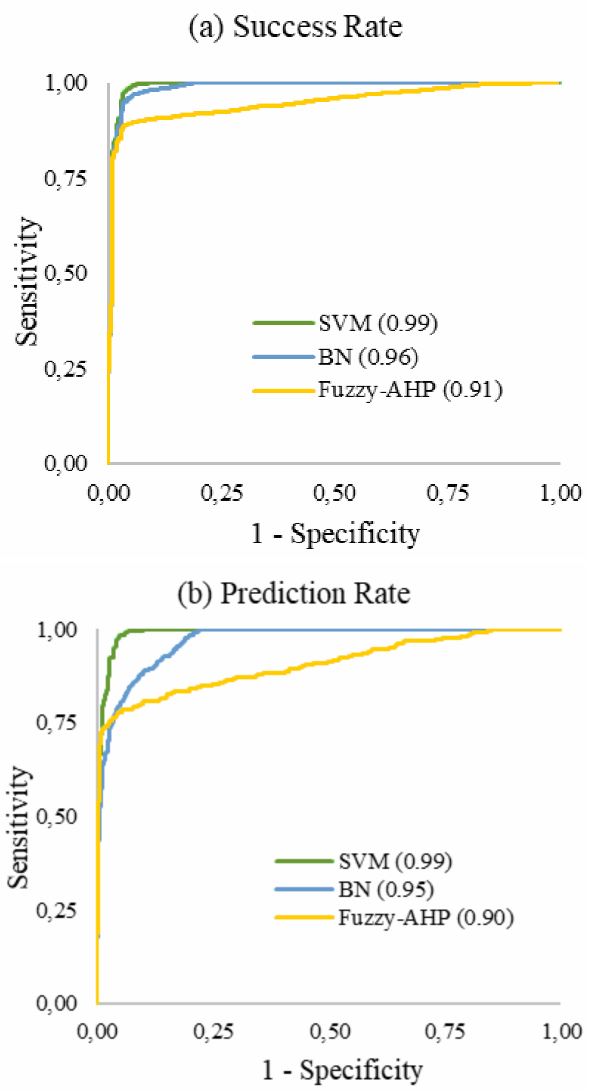

Fig. 7. AUROC curve: (a) success rate; (b) prediction rate.

While the AUROC analysis reveals the comparative superiority of the ML algorithms over the GIS-based Fuzzy-AHP, it is important to acknowledge that Fuzzy-AHP still emerges as an “excellent” predictive model with higher AUROC values. The Fuzzy-AHP model exhibits high efficiency. However, the utilization of ML algorithms in this study has overshadowed its performance standards.

4. DISCUSSION

The rapid population growth and the presence of diverse natural hazards have exerted significant pressure on the land resources in Hawalbagh Block. To alleviate this pressure, it is essential to implement systematic development strategies. One such strategy involves the identification of appropriate locations for built-up expansion. The primary objective of site suitability analysis is to determine sites that meet the requirements of various stakeholders and are conducive to the desired development objectives (Thill, 2019).

Physiography has always had a significant control on human settlements, especially in the hilly areas. Therefore, nine important socio-physical drivers were taken into consideration while determining a suitable built-up region based on prior research and pioneering work (Ananda and Herath, 2008; Chen et al., 2011; Mosadeghi et al., 2015).

In the process of built-up suitability mapping, slope emerged as a significant conditioning factor among the nine factors considered in all three methodologies. The findings indicate that the central and western sections of Hawalbagh Block are highly stable and hence fit for development owing to their gentle gradation, ground stability, road accessibility, and developed fields. However, all three models suggest that the northern and southern margins and east edges of the block are not fit for built-up area development in the near future. These findings may prove valuable to policymakers, researchers and other stakeholders when deciding on the potential sites for hillside development and planning. While GIS-based Fuzzy-AHP has several advantages, it also has significant limitations as variable weighting in Fuzzy-AHP requires strong assumptions from decision-makers, which can subjectively lead to lower accuracy in the results (Al Mamun et al., 2019; Almansi et al., 2021).

In this study, the suitability assessment and mapping of built-up areas in the study region were compared between the GIS-based Fuzzy-AHP technique and ML algorithms. The ML algorithms employed hyper-parameter tuning and 5-fold cross-validation to address over-fitting concerns and improve predictive accuracy (Maglogiannis, 2007). As discussed earlier, the accuracy of the models was evaluated with AUROC analysis. The results indicate that the ML algorithms (SVM & BN) outperformed the GIS-based Fuzzy-AHP technique in assessing the suitability of built-up land.

Data restrictions are a major limitation in these types of studies. High resolution satellite data are more relevant for suitability assessments. Data with low resolution may lead to more generalization. A significant constraint in MCDM-based technique is the identification of the optimal rank for the influencing criterion in order to minimize sensitivity of the weight to optimize the outcome. One major challenge in developing ML algorithms is the ample availability of inventory data for training and validating the model. The complex geographic structure of the study region poses a significant obstacle in collecting such data. Therefore, it is recommended that future studies explore the integration of ML algorithms with MCDM techniques such as Fuzzy-SVM, Fuzzy-RT, or other techniques to overcome this limitation.

5. CONCLUSIONS

The Himalayas range represents the world's most rapidly urbanized mountain system, characterized by fast but unplanned and unregulated urban growth. This accelerated urbanization amplifies the vulnerability of densely populated and extensively modified slopes to active mass movements and land sliding (Tiwari et al., 2023).

The proliferation of urban settlements in the Himalayas cannot be prevented since these are centres of social and economic development. However, the challenges of unsystematic expansion may be averted via feasible and comprehensive policies, particularly land-use policy. Consequently, it is imperative to optimize the utilization of available land, develop new land, and establish substantial land reserves to meet the growing demand for suitable land in society. The Hawalbagh Block in the Kumaun Himalayas has observed a rapid and unplanned urban expansion, accompanied by the mismanagement of natural resources, resulting in significant issues. Over a span of two decades, from 1990 to 2010, the built-up area of the Hawalbagh Block has undergone substantial growth, increasing from 2.72 km² to 12.20 km² (Rawat and Kumar, 2015). The findings of this research work deliver significant insights for the Uttarakhand state government and the municipality of Almora in promoting sustainable built-up area development, including proper land-use development, planning, and the systematic management and conservation of natural resources.

Considering the densely populated nature of the eastern part of Hawalbagh Block, the discoveries of this research work also open possibilities for the impending relocation of people in this region. The best suitable zonation requires high-quality data availability, the best subjective evaluation of MCDM criteria, and proper training of ML-based models. The suitability zonation may be used to establish a land-use policy in line with The Forest (Conservation) Act, 1980, The Environmental (Protection) Act, 1986 and National Action Plan on Climate Change (NAPCC) for a sustainable mountain ecology in the Kumaun Himalayas. Proper and systematic built-up area development is a precondition for sustainable development, and the findings of this study can contribute to achieving this goal by guiding the actions of the Uttarakhand state government and the municipality of Almora in their efforts towards sustainable built-up area development, including land-use planning and management, and the conservation of natural resources. Some examples of good practices for appropriate and sustainable development that can be applied to Almora could be mixed-use development, green building design, preservation of cultural heritage etc.

6. ACKNOWLEDGEMENTS

The authors express their gratitude to the Department of Geography at the Central University of Haryana, Mahendragarh, IN for providing a technically robust platform, supporting software services (ArcGIS and SPSS), and a supportive atmosphere for the successful execution of this research.

REFERENCES

Abdollahi S., Pourghasemi H. R., Ghanbarian G. A., Safaeian R. (2019), Prioritization of effective factors in the occurrence of land subsidence and its susceptibility mapping using an SVM model and their different kernel functions. Bulletin of Engineering Geology and the Environment, 78(6), 4017–4034. DOI: https://doi.org/10.1007/s10064-018-1403-6

Abebe Y., Kabir G., Tesfamariam S. (2018), Assessing urban areas vulnerability to pluvial flooding using GIS applications and Bayesian Belief Network model. Journal of Cleaner Production, 174, 1629–1641. DOI: https://doi.org/10.1016/j.jclepro.2017.11.066

Al Mamun M. A., Howladar M. F., Sohail M. A. (2019), Assessment of surface water quality using Fuzzy Analytic Hierarchy Process (FAHP): A case study of Piyain River's sand and gravel quarry mining area in Jaflong, Sylhet. Groundwater for Sustainable Development, 9, 100208. DOI: https://doi.org/10.1016/j.gsd.2019.03.002

Almansi K. Y., Shariff A. R. M., Abdullah A. F., Syed Ismail S. N. (2021), Hospital Site Suitability Assessment Using Three Machine Learning Approaches: Evidence from the Gaza Strip in Palestine. Applied Sciences, 11(22), 11054. DOI: https://doi.org/10.3390/app112211054

Ananda J., Herath G. (2008), Multi-attribute preference modelling and regional land-use planning. Ecological Economics, 65(2), 325–335. DOI: https://doi.org/10.1016/j.ecolecon.2007.06.024

Arciniegas G., Janssen R., Omtzigt N. (2011), Map-based multicriteria analysis to support interactive land use allocation. International Journal of Geographical Information Science, 25(12), 1931–1947. DOI: https://doi.org/10.1080/13658816.2011.556118

Chang D. Y. (1996), Applications of the extent analysis method on fuzzy AHP. European Journal of Operational Research, 95(3), 649–655. DOI: https://doi.org/10.1016/0377-2217(95)00300-2

Chang N.-B., Parvathinathan G., Breeden J. B. (2008), Combining GIS with fuzzy multicriteria decision-making for landfill siting in a fast-growing urban region. Journal of Environmental Management, 87(1), 139–153. DOI: https://doi.org/10.1016/j.jenvman.2007.01.011

Chaudhry A. K., Kumar K., Alam M. A. (2021), Mapping of groundwater potential zones using the fuzzy analytic hierarchy process and geospatial technique. Geocarto International, 36(20), 2323–2344. DOI: https://doi.org/10.1080/10106049.2019.1695959

Chen H., Wood M. D., Linstead C., Maltby E. (2011), Uncertainty analysis in a GIS-based multi-criteria analysis tool for river catchment management. Environmental Modelling & Software, 26(4), 395–405. DOI: https://doi.org/10.1016/j.envsoft.2010.09.005

Chen J., Zhong P.-A., An R., Zhu F., Xu B. (2019), Risk analysis for real-time flood control operation of a multi-reservoir system using a dynamic Bayesian network. Environmental Modelling & Software, 111, 409–420. DOI: https://doi.org/10.1016/j.envsoft.2018.10.007

Chen Y., Yu J., Khan S. (2010), Spatial sensitivity analysis of multi-criteria weights in GIS-based land suitability evaluation. Environmental Modelling & Software, 25(12), 1582–1591. DOI: https://doi.org/10.1016/j.envsoft.2010.06.001

Cheng C.-H. (1997), Evaluating naval tactical missile systems by fuzzy AHP based on the grade value of membership function. European Journal of Operational Research, 96(2), 343–350. DOI: https://doi.org/10.1016/S0377-2217(96)00026-4

Cortes C., Vapnik V. (1995), Support-vector networks. Machine Learning, 20(3), 273–297. DOI: https://doi.org/10.1007/BF00994018

Cox E. D. (1995), Fuzzy logic for business and industry. Charles River Media, Inc. ISBN: 1886801010

Das B., Pal S. C. (2019), Combination of GIS and fuzzy-AHP for delineating groundwater recharge potential zones in the critical Goghat-II block of West Bengal, India. HydroResearch, 2, 21–30. DOI: https://doi.org/10.1016/j.hydres.2019.10.001

Deng H. (1999), Multicriteria analysis with fuzzy pairwise comparison. International Journal of Approximate Reasoning, 21(3), 215–231. DOI: https://doi.org/10.1016/S0888-613X(99)00025-0

Dunning D. J., Ross Q. E., Merkhofer M. W. (2000), Multiattribute utility analysis for addressing Section 316(b) of the Clean Water Act. Environmental Science & Policy, 3(1), 7–14. DOI: https://doi.org/10.1016/S1462-9011(00)00022-8

Fenton N., Neil M. (2018), Risk Assessment and Decision Analysis with Bayesian Networks (2nd ed.). Chapman and Hall/CRC. New York. DOI: https://doi.org/10.1201/b21982

Gharaibeh A. A., Jaradat M. A., Kanaan L. M. (2023), A Machine Learning Framework for Assessing Urban Growth of Cities and Suitability Analysis. Land, 12(1), Article 1, 214. DOI: https://doi.org/10.3390/land12010214

Gokceoglu C. (2001), Discussion on "Landslide hazard zonation of the Khorshrostam area, Iran" by A. Uromeihy and M.R. Mahdavifar. Bulletin of Engineering Geology and the Environment, 60, 78-80. DOI: https://doi.org/10.1007/s100640000081

Gökceoglu C., Aksoy H. (1996), Landslide susceptibility mapping of the slopes in the residual soils of the Mengen region (Turkey) by deterministic stability analyses and image processing techniques. Engineering Geology, 44(1–4), 147–161. DOI: https://doi.org/10.1016/S0013-7952(97)81260-4

Greene R., Luther J. E., Devillers R., Eddy B. (2010), An approach to GIS-based multiple criteria decision analysis that integrates exploration and evaluation phases: Case study in a forest-dominated landscape. Forest Ecology and Management, 260(12), 2102–2114. DOI: https://doi.org/10.1016/j.foreco.2010.08.052

Guarnieri A., Masiero A., Vettore A., Pirotti F. (2015), Evaluation of the dynamic processes of a landslide with laser scanners and Bayesian methods. Geomatics, Natural Hazards, and Risk, 6(5–7), 614–634. DOI: https://doi.org/10.1080/19475705.2014.983553

Hajkowicz S., Collins K. (2007), A Review of Multiple Criteria Analysis for Water Resource Planning and Management. Water Resources Management, 21(9), 1553–1566. DOI: https://doi.org/10.1007/s11269-006-9112-5

Hamilton S. H., Pollino C. A., Jakeman A. J. (2015), Habitat suitability modelling of rare species using Bayesian networks: Model evaluation under limited data. Ecological Modelling, 299, 64–78. DOI: https://doi.org/10.1016/j.ecolmodel.2014.12.004

Johnson S., Mengersen K., de Waal A., Marnewick K., Cilliers D., Houser A. M., Boast L. (2010), Modelling cheetah relocation success in southern Africa using an Iterative Bayesian Network Development Cycle. Ecological Modelling, 221(4), 641–651. DOI: https://doi.org/10.1016/j.ecolmodel.2009.11.012

Kahraman C., Ruan D., Doğan I. (2003), Fuzzy group decision-making for facility location selection. Information Sciences, 157, 135–153. DOI: https://doi.org/10.1016/S0020-0255(03)00183-X

Kordi M., Brandt S. A. (2012), Effects of increasing fuzziness on analytic hierarchy process for spatial multicriteria decision analysis. Computers, Environment and Urban Systems, 36(1), 43–53. DOI: https://doi.org/10.1016/j.compenvurbsys.2011.07.004

Koski T., Noble J. (2011), Bayesian networks: an introduction. John Wiley & Sons. ISBN: 978-0-470-74304-1

Kumar M., Biswas V. (2013), Identification of Potential Sites for Urban Development Using GIS Based Multi Criteria Evaluation Technique. A Case Study of Shimla Municipal Area, Shimla District, Himachal Pradesh, India. Journal of Settlements and Spatial Planning, 4(1) 45-51. URL: https://geografie.ubbcluj.ro/ccau/jssp/arhiva_1_2013/05JSSP012013.pdf

Kumar M., Shaikh V. R. (2013), Site Suitability Analysis for Urban Development Using GIS Based Multicriteria Evaluation Technique. Journal of the Indian Society of Remote Sensing, 41(2), 417–424. DOI: https://doi.org/10.1007/s12524-012-0221-8

Kumar M., Singh R. B., Singh A., Pravesh R., Majid S. I., Tiwari A. (2023), Case Study 3: Identification of Potential Sites for Housing Development Using GIS-Based Multi-criteria Evaluation Technique. In Geographic Information Systems in Urban Planning and Management. Springer Nature Singapore, 171-189. DOI: https://doi.org/10.1007/978-981-19-7855-5_11

Li L., Shi Z.-H., Yin W., Zhu D., Ng S. L., Cai C.-F., Lei A.-L. (2009), A fuzzy analytic hierarchy process (FAHP) approach to eco-environmental vulnerability assessment for the danjiangkou reservoir area, China. Ecological Modelling, 220(23), 3439–3447. DOI: https://doi.org/10.1016/j.ecolmodel.2009.09.005

Madu I. A. (2007), The underlying factors of rural development patterns in the Nsukka region of southeastern Nigeria. Journal of Rural and Community Development, 2(1), 110-122. ISSN: 1712-8277

Maglogiannis I. G. (2007), Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in EHealth, HCI, Information Retrieval and Pervasive Technologies. IOS Press. ISBN: 978-1-58603-780-2

Maity B., Mallick S. K., Das P., Rudra S. (2022), Comparative analysis of groundwater potentiality zone using fuzzy AHP, frequency ratio and Bayesian weights of evidence methods. Applied Water Science, 12(4), 63. DOI: https://doi.org/10.1007/s13201-022-01591-w

Manjare B. S. (2014), Identification of groundwater prospecting zones using Remote Sensing and GIS techniques in upper Vena river watersheds Nagpur district, Maharashtra, India. In *15th ESRI India user conference* (pp. 1-14)

Marcot B. G., Penman T. D. (2019), Advances in Bayesian network modelling: Integration of modelling technologies. Environmental Modelling & Software, 111, 386-393. DOI: https://doi.org/10.1016/j.envsoft.2018.09.016

Mayfield C. (2015), Automating the Classification of Thematic Rasters for Weighted Overlay Analysis in GeoPlanner for ArcGIS [Master of Science, University of Redlands]. DOI: https://doi.org/10.26716/redlands/master/2015.6

Mohamed M. M., Elmahdy S. I. (2017), Fuzzy logic and multi-criteria methods for groundwater potentiality mapping at Al Fo'ah area, the United Arab Emirates (UAE): An integrated approach. Geocarto International, 32(10), 1120–1138. DOI: https://doi.org/10.1080/10106049.2016.1195884

Mojaddadi Rizeei H. (2018), Flood risk assessment using multi-sensor remote sensing, geographic information system, 2D hydraulic and machine learning based models [Thesis]. DOI: https://opus.lib.uts.edu.au/handle/10453/133315

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. Computers, Environment and Urban Systems, 49, 54–65. DOI: https://doi.org/10.1016/j.compenvurbsys.2014.10.001

Naghibi S. A., Ahmadi K., Daneshi A. (2017), Application of Support Vector Machine, Random Forest, and Genetic Algorithm Optimized Random Forest Models in Groundwater Potential Mapping. Water Resources Management, 31(9), 2761–2775. DOI: https://doi.org/10.1007/s11269-017-1660-3

Ornella L., Tapia E. (2010), Supervised machine learning and heterotic classification of maize (Zea mays L.) using molecular marker data. Computers and Electronics in Agriculture, 74(2), 250–257. DOI: https://doi.org/10.1016/j.compag.2010.08.013

Pachauri A. K., Gupta P. V., Chander R. (1998), Landslide zoning in a part of the Garhwal Himalayas. Environmental Geology, 36(3–4), 325–334. DOI: https://doi.org/10.1007/s002540050348

Pandey V. K., Pourghasemi H. R., Sharma M. C. (2020), Landslide susceptibility mapping using maximum entropy and support vector machine models along the highway corridor, Garhwal Himalaya. Geocarto International, 35(2), 168–187. DOI: https://doi.org/10.1080/10106049.2018.1510038

Pham B. T., Tien Bui D., Prakash I., Dholakia M. B. (2017), Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. CATENA, 149 (Part 1), 52–63. DOI: https://doi.org/10.1016/j.catena.2016.09.007

Rawat J. S., Kumar M. (2015), Monitoring land use/cover change using remote sensing and GIS techniques: A case study of Hawalbagh block, district Almora, Uttarakhand, India. The Egyptian Journal of Remote Sensing and Space Science, 18(1), 77–84. DOI: https://doi.org/10.1016/j.ejrs.2015.02.002

Saaty T. L. (1984), The Analytic Hierarchy Process: Decision Making in Complex Environments. In: Avenhaus, R., Huber, R.K. (eds) Quantitative Assessment in Arms Control. Springer, Boston, MA, 285-308. DOI: https://doi.org/10.1007/978-1-4613-2805-6_12

Saaty T. L., Tran L. T. (2007), On the invalidity of fuzzifying numerical judgments in the Analytic Hierarchy Process. Mathematical and Computer Modelling, 46(7-8), 962–975. DOI: https://doi.org/10.1016/j.mcm.2007.03.022

Sarkar D., Mondal P. (2019), Flood vulnerability mapping using frequency ratio (FR) model: A case study on Kulik river basin, Indo-Bangladesh Barind region. Applied Water Science, 10(1), 17. DOI: https://doi.org/10.1007/s13201-019-1102-x

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. Artificial Intelligence in Geosciences, 2, 179–191. DOI: https://doi.org/10.1016/j.aiig.2022.03.001

Şener E., Şener Ş., Davraz A. (2018), Groundwater potential mapping by combining fuzzy-analytic hierarchy process and GIS in Beyşehir Lake Basin, Turkey. Arabian Journal of Geosciences, 11(8), 187. DOI: https://doi.org/10.1007/s12517-018-3510-x

Shahabi H., Hashim M. (2015), Landslide susceptibility mapping using GIS-based statistical models and Remote sensing data in tropical environment. Scientific Reports, 5(1), Article 9899. DOI: https://doi.org/10.1038/srep09899

Shirzadi A., Bui D. T., Pham B. T., Solaimani K., Chapi K., Kavian A., Shahabi H., Revhaug I. (2017), Shallow landslide susceptibility assessment using a novel hybrid intelligence approach. Environmental Earth Sciences, 76(2), 60. DOI: https://doi.org/10.1007/s12665-016-6374-y

Singh A., Singh K. K. (2017), Satellite image classification using Genetic Algorithm trained radial basis function neural network, application to the detection of flooded areas. Journal of Visual Communication and Image Representation, 42, 173–182. DOI: https://doi.org/10.1016/j.jvcir.2016.11.017

Singh S., Kulkarni V., Briggs M., Mahajan D., Farchi E. (2022), Using Fuzzy Matching of Queries to optimize Database workloads. Cornell University. DOI: https://doi.org/10.48550/ARXIV.2207.06820

Solomatine D. P., Shrestha D. L. (2009), A novel method to estimate model uncertainty using machine learning techniques. Water Resources Research, 45(12). DOI: https://doi.org/10.1029/2008WR006839

Tan R. R., Aviso K. B., Huelgas A. P., Promentilla M. A. B. (2014), Fuzzy AHP approach to selection problems in process engineering involving quantitative and qualitative aspects. Process Safety and Environmental Protection, 92(5), 467–475. DOI: https://doi.org/10.1016/j.psep.2013.11.005

Tantipisanuh N., Gale G. A., Pollino C. (2014), Bayesian networks for habitat suitability modeling: A potential tool for conservation planning with scarce resources. Ecological Applications, 24(7), 1705–1718. DOI: https://doi.org/10.1890/13-1882.1

Tehrany M. S., Kumar L., Shabani F. (2019), A novel GIS-based ensemble technique for flood susceptibility mapping using evidential belief function and support vector machine: Brisbane, Australia. PeerJ, 7, e7653. DOI: https://doi.org/10.7717/peerj.7653

Tehrany M. S., Pradhan B., Jebur M. N. (2015), Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. Stochastic Environmental Research and Risk Assessment, 29(4), 1149–1165. DOI: https://doi.org/10.1007/s00477-015-1021-9

Thill J.-C. (ed.) (2019), Spatial Multicriteria Decision Making and Analysis: A geographic information sciences approach. First edition. Routledge. London. DOI: https://doi.org/10.4324/9780429436628

Tien Bui D., Tuan T. A., Klempe H., Pradhan B., Revhaug I. (2016), Spatial prediction models for shallow landslide hazards: A comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides, 13(2), 361–378. DOI: https://doi.org/10.1007/s10346-015-0557-6

Tiwari A., Kumar M., Thakur S., Majid S. I. (2023), Groundwater quality assessment for drinking purpose in Gurugram district of Haryana, India. Arabian Journal of Geosciences, 16(8), 469. DOI: https://doi.org/10.1007/s12517-023-11571-2

Torabi-Kaveh M., Babazadeh R., Mohammadi S., Zaresefat M. (2016), Landfill site selection using combination of GIS and fuzzy AHP, a case study: Iranshahr, Iran. Waste Management & Research, 34(5), 438-448. DOI: https://doi.org/10.1177/0734242X16633777

Valdiya K. (1980), Stratigraphic scheme of the sedimentary units of the kumaun lesser himalaya. URL: http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=PASCALGEODEBRGM8220008288

Van Laarhoven P. J. M., Pedrycz W. (1983), A fuzzy extension of Saaty's priority theory. Fuzzy Sets and Systems, 11(1-3), 229–241. DOI: https://doi.org/10.1016/S0165-0114(83)80082-7

Verma C., Illés Z., Stoffová V. (2019), Age Group Predictive Models for the Real Time Prediction of the University Students using Machine Learning: Preliminary Results. 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 1–7. DOI: https://doi.org/10.1109/ICECCT.2019.8869136

Verma C., Tarawneh A. S., Illes Z., Stoffova V., Dahiya S. (2018), Gender Prediction of the European School's Teachers Using Machine Learning: Preliminary Results. 2018 IEEE 8th International Advance Computing Conference (IACC), Greater Noida, India, 213–220. DOI: https://doi.org/10.1109/IADCC.2018.8692100

Wu F.-G., Lee Y.-J., Lin M.-C. (2004), Using the fuzzy analytic hierarchy process on optimum spatial allocation. International Journal of Industrial Ergonomics, 33(6), 553–569. DOI: https://doi.org/10.1016/j.ergon.2004.01.002

Wu Z., Shen Y., Wang H., Wu M. (2019), Assessing urban flood disaster risk using Bayesian network model and GIS applications. Geomatics, Natural Hazards, and Risk, 10(1), 2163–2184. DOI: https://doi.org/10.1080/19475705.2019.1685010

Xie M., Esaki T., Zhou G., Mitani Y. (2003), Geographic Information Systems-Based Three-Dimensional Critical Slope Stability Analysis and Landslide Hazard Assessment. Journal of Geotechnical and Geoenvironmental Engineering, 129(12), 1109–1118. DOI: https://doi.org/10.1061/(ASCE)1090-0241(2003)129:12(1109)

Xiong J., Li, J., Cheng W., Wang N., Guo L. (2019), A GIS-Based Support Vector Machine Model for Flash Flood Vulnerability Assessment and Mapping in China. ISPRS International Journal of Geo-Information, 8(7), Article 7, 297. DOI: https://doi.org/10.3390/ijgi8070297

Zopounidis C., Doumpos M. (2002). Multicriteria classification and sorting methods: A literature review. European Journal of Operational Research, 138(2), 229–246. DOI: https://doi.org/10.1016/S0377-2217(01)00243-0

Zopounidis C., Pardalos P. M. (eds.) (2010), Handbook of multicriteria analysis (Vol. 103). Springer Science & Business Media. DOI: https://doi.org/10.1007/978-3-540-92828-7