Identification and validation of potential flood hazard area using GIS-based multi-criteria analysis and satellite data-derived water index

Pratik Dash\(^1\) | Jishnu Sar\(^2\)

\(^1\)Department of Geography, School of Science, Adamas University, Kolkata, India
\(^2\)Department of Geography, Banaras Hindu University, Varanasi, India

**Correspondence**
Pratik Dash, Department of Geography, School of Science, Adamas University, Kolkata 700126, India.
Email: pratik.bidyut@gmail.com

**Abstract**
This article identifies potential flood hazard areas through multi-criteria analysis in Allahabad district, India. The study has incorporated eight criteria, namely, flow accumulation, draining capability, elevation, groundwater depth, land use, runoff coefficient, slope, and geology for preparing hazard index. The weights of the criteria were obtained through the analytical hierarchy process (AHP) method based on their relative importance for occurring floods. Finally, a flood hazard index (FHI) was prepared by combining the parameter ratings and corresponding weights. The credibility of the present methodology was tested through validation with the satellite-based inundation map of August 20, 2016. A normalized difference water index (NDWI) was prepared from Landsat-8 OLI data and the inundated area was delineated by a binary classification of NDWI based on a threshold calculated following Otsu’s method. The analysis found 81% of inundation is associated with high to very high flood hazard zones. Agricultural land is more prone to flood than other land use types. The results showed that the GIS-based multi-criteria analysis framework could be effectively applied for flood hazard analysis to support decision making in disaster management.

**KEYWORDS**
analytical hierarchy process, flood hazard index, GIS, Landsat-8 OLI, normalized difference water index

1 | INTRODUCTION

Flooding is one of the most recurring and devastating natural hazards that threaten society from its immense impacts on human lives and socio-economic conditions (Qi & Altinakar, 2011). The effects of floods are severely felt in developing countries (like India), in particular by the local rural community. Flooding problems, regardless of topographical and meteorological settings, are becoming acute due to environmental changes, like land use change (Du, Shi, Van Rompaey, & Wen, 2015), rapid urbanization (Suriya & Mudgal, 2012), and climate change (Detrembleur et al., 2015). Since they are not avoidable, flood risks could be minimized through adopting proper management and mitigation strategies. Therefore, the assessment of flood risks is a prerequisite for adopting mitigation strategies.
Flood risk is a measure of exposure to damage and loss from flooding, commonly estimated accounting physio-climatic, hydrodynamic, economic, social, and ecological aspects. Flood risk is calculated by combining hazard and vulnerability considering either their summation (Wang, Li, Tang, & Zeng, 2011) or multiplication (Li, Xiang, Tong, & Wang, 2013). The analytical calculation of flood risk is very complex and challenging due to data constraints (Kron, 2005). Instead, hazard assessment through numerical modeling (Dutta et al., 2007; Liu, Zhang, & Cui, 2012), and index-based analysis (Chen, Yeh, & Yu, 2011; Kazakis, Kougias, & Patsialis, 2015) are quite popular as a substitute for full flood risk analysis. Hence, it is very effective to calculate the hazard for developing and making a strategy in disaster mitigation.

Hydrological and hydrodynamic models are widely used for flood simulation in terms of magnitude, frequency, and extent of floods at basin scale (Kuldeep et al., 2016; Liu et al., 2012; Ullah, Farooq, Sarwar, Tareen, & Wahid, 2016). Nevertheless, the application of hydrodynamic models in limited-data conditions and at an administrative level is a difficult task. Thus, an index-based approach coupling critical parameters is a valid solution for assessing the magnitude of hazard.

Past studies on flood hazard assessment have applied index-based approach using various parameters based on digital elevation model (DEM)-derived geomorphological and hydrological characteristics (Noman, Nelson, & Zundel, 2001; Samela et al., 2016), land use and urbanization information (Dewan, Islam, Kumamoto, & Nishigaki, 2007; Xiao, Yi, & Tang, 2016), economic activity, infrastructure, and demographic aspects (Mojtahedi & Oo, 2016; van der Veen & Logtmeijer, 2005), and so on while some studies tried to focus on mitigation policy and rehabilitation issues (Burby, Deyle, Logtmeijer, & Olshansky, 2000). However, the assessment of spatial exposure to flood risk is essential for providing early warning and disaster risk reduction.

In the last two decades, a lot of studies have used multi-criteria decision analysis (MCDA) to estimate index-based flood hazard and risk through understanding the role of parameters controlling floods (Hazarika, Barman, Das, Sarma, & Borah, 2016; Kazakis et al., 2015; Papaioannou, Vasilides, & Loukas, 2015; Wu et al., 2015; Xiao et al., 2016). GIS-based MCDA examines complex decision problems by organizing the criteria in a hierarchical manner (Chen et al., 2011). A literature review conducted by de Brito and Evers (2016) found that around 82% of the study used MCDA, most of which studied in Asian and European countries (85%), and about 72% of the study used the analytical hierarchy process (AHP) in MCDA framework (de Brito & Evers, 2016; Madhuri, Aniruddha, & Rahul, 2013; Wang et al., 2011), though the perspective and number of parameters used in these studies were not fixed. For example, Chen, Ito, Sawamukai, and Tokunaga (2015) used six criteria namely, river system, elevation, depression area, ratio of impermeable area, detention ponds, and precipitation; while Kazakis et al. (2015) prepared a flood hazard index (FHI) at the administrative level by analyzing seven hydrogeological and geomorphological parameters. However, most studies have ignored validation which is essential as the AHP model is subjected to uncertainty from multiple sources (de Brito & Evers, 2016). Few studies attempted to validate results with historical data of flood events at some selected points (Kazakis et al., 2015; Wu et al., 2015). However, location-based event data is not sufficient to validate the MCDA-based hazard map, rather spatial maps of historical floods can be a better option for validation; in this context, validation of MCDA approach using spatial inundation maps extracted from satellite data is not well established.

Spatial mapping of historical floods using spaceborne data is becoming popular due to advancements in sensor capability and repetitive coverage. The synthetic aperture radar (SAR) data is quite useful in flood mapping due to its cloud penetration capability (Giustarini et al., 2016), though it is very costly. Therefore, freely available multi-spectral imageries like Landsat 8 OLI images are suitable for land surface water mapping (Du et al., 2014; Xu, 2006). Various techniques have been developed to classify water pixels by processing surface reflectance captured by different spectral bands of optical sensors (Feyisa, Meilby, Fensholt, & Proud, 2014). Many studies used normalized difference water index (NDWI) for extracting surface water bodies (Du et al., 2014; Li, Du et al., 2013) and flood inundation (Pandey, Singh, & Nathawat, 2010).

The Gangetic plain is one of the densely populated areas in the world that are repeatedly affected by floods. The recurring flood causes massive damage to infrastructure and agricultural production. The farming community, mostly smallholders and marginal farmers are most vulnerable to floods (Aggarwal, 2008). Mapping of potential flood hazard areas at the local level (block and district) can benefit more to vulnerable rural communities (Asare-Kyei, Forkuor, & Venus, 2015). This study tried to prepare an integrated flood hazard map for Allahabad district located in the middle Ganga plain, India. The objectives of this study are to (a) prepare a flood hazard map for identifying hazard zone where mitigation measure should be applied at a priority basis, and (b) derive flood inundation extent from multi-spectral satellite image for validating hazard map. Here, GIS-based spatial multi-criteria evaluation framework has been used to prepare an index-based flood hazard map by combining selected eight criteria. An inundation map of 2016 was
prepared from Landsat 8 OLI image through NDWI to validate the index-based flood hazard map. We then reconcile hazard mapping with land use mapping to find which types of land use are more vulnerable to floods. The proposed method can be used as a useful tool in flood hazard mapping for data-limited conditions. Besides, the overall methodology applied for satellite data-derived inundation mapping may contribute to flood management, especially in the agricultural regions.

2 STUDY AREA

Allahabad district is located in the eastern part of Uttar Pradesh, India extending between 24°48'36" N to 25°45'18" N latitudes 81°30'42" E to 82°20'41" E longitudes with an area of about 5,482 km² (Figure 1). Elevation in the district varies from 70 to 380 m, with an average height of 103 m. Quaternary alluvium is the primary lithological unit of the study area as most part lies in the Ganga floodplain. The study area is drained by the major rivers like Ganga, Yamuna, Tons, and Belon. The traces of paleochannels and river shifting still exist in this area in the form of oxbow lakes. Many low-lying areas and ponds are evident in the district. Meteorologically, the district lies in a humid subtropical climate. The monthly average temperature varies between 41.5°C (May) and 9.3°C (January). The area receives about 934 mm average annual rainfall, contributed mostly by south-west monsoon with a maximum in August (about 280 mm). Agricultural land is the dominant land use type. As per the 2011 census report, the total population of the district is 5,954,391, of which 75.3 and 24.7% are shared by rural and urban populations, respectively (Census of India, 2011). Between 2001 and 2011, the decadal growth of population is 20.6%.

By its location in the floodplain with several paleochannels, the district is more vulnerable to flooding. Nevertheless, the confluence of the Yamuna river with the Ganga river near Allahabad city makes flooding as a recurrent phenomenon during peak monsoon season. The district has experienced massive floods in 1948, 1956,
1967, 1978, 1983, and 2016 that caused immense losses of agricultural production, lives, and property (City Development Plan 2015). The previous study found about 64 villages of the district were affected by flooding during 2010, even though the magnitude of rainfall and flood in that year was low as compared to the years of an extreme flood (Bhatt & Rao, 2016). In the recent flood of 2016, the water level of Ganga river crossed the danger mark (84.73 m) and flowed above 86.12 m which flooded a vast area of Ashok Nagar, Baghada, Ganganagar, Mehndauri, Rajapur, Salori, Daraganj in and around Allahabad city (Mani, 2016). Besides the geographical location, the severity of floods has increased with solid waste disposal, encroachment, and mismanagement (Tiwari, 2013).

3 MATERIAlS AND METHODS

The present study used GIS-based spatial data for multicriteria analysis to identify flood hazard zones. The relative weights of the criteria were calculated through the AHP method (Saaty, 1980) by analyzing the relationships between selected criteria in a structured way considering their influence on the flood. The AHP method includes six necessary steps (Lee, Chen, & Chang, 2008; Papaioannou et al., 2015): Define the unstructured problem, establishment of the AHP hierarchy, formation of the pairwise comparison table, calculation of the relative weights, consistency check, and validation of the method.

The thematic layers of the selected criteria were prepared, and their value range was clustered for assigning ratings. Jenks’ natural breaks clustering technique (Jenks, 1967) was used for all criteria, except qualitative parameters like land use type, geology, and so on that have predefined discrete class values. The Natural breaks classification technique is popularly applied in MCDA-based flood studies because it divides the data range into clusters based on natural groups (Huan, Wang, & Teng, 2012; Papaioannou et al., 2015; Stefanidis & Stathis, 2013). This study used natural breaks technique because it reduces variance within the classes and maximizes interclass variability (Huan et al., 2012). Therefore, this method minimizes the mean deviation of each class from its mean concomitant with maximizing deviation from the means of other class groups (Stefanidis & Stathis, 2013). Combining the relative weights of the criteria and parameter ratings, the flood hazard map was prepared in GIS environment.

3.1 Criteria identification

The selection of appropriate criteria is most important in the MCDA-based semiquantitative method. Previous studies have used various criteria in correspondence to local geographical settings that influence the physical process of flood mechanism (Chen et al., 2015; Kazakis et al., 2015; Papaioannou et al., 2015; Wu et al., 2015). Since the present study area is characterized by low-lying flood plain dominated with fertile alluvial agricultural land, eight criteria related to physiographic, hydrologic, and surface characteristics were identified. These criteria are flow accumulation, draining capability, elevation, groundwater depth, land use, runoff coefficient, slope, and geology.

Flow accumulation defines the cumulative flow downslope; therefore, it is very beneficial for identifying the flood-prone area. In addition to flooding from a spill of bankfull discharge, a significant part of the low-lying area is flooded due to insufficiency in water outflow. Therefore, the draining capability was considered as relevant parameters as it reflects the ability to drain out excess rainfall. Likewise, elevation and slope were selected as they control water flow over the surface; while groundwater depth may contribute to control infiltration and runoff rates. In this study, the groundwater depth parameter used for the pre-monsoon period rather than the monsoon period, because the pre-monsoon conditions influence antecedent moisture of soil that, in turn, may control the rate of infiltration during peak rainfall. Land use is another critical parameter that controls the rate of runoff generation (Nie et al., 2011). Since the interaction of rainfall with the land surface is a very complex process, the land use parameter is not sufficient alone to emphasize runoff potential because soil types and slope also control flood susceptibility (Asare-Kyei et al., 2015). Therefore, a spatial variation in runoff coefficient was used to incorporate the complex interactions of land use, soil, and slope with rainfall by decomposing their individual influence.

The indices of selected criteria were prepared in GIS environment with the same spatial resolution of ~30 m. The thematic layers of flow accumulation, elevation, and slope were prepared from DEM (ASTER GDEM) of 1 arc-second (~30 m) spatial resolution. Flow accumulation was estimated from DEM using the Hydrology tool of ArcGIS, and slope (in degree) layer was derived using the “surface analysis” extension of Spatial Analyst tool in ArcGIS. The thematic layer of draining capability was prepared by overlaying vector layers of drainage density and elevation zone; where drainage density layer obtained through the interpolation of DEM-derived drainage density values calculated for 1 km square grid cells, and elevation zones prepared by reclassifying the DEM value range. Each combination was assigned to a rating and converted to a raster layer. The parameter describing the depth of groundwater below surface
(hereafter groundwater depth) in the pre-monsoon period was prepared through vectorization of the map published by Central Ground Water Board, India (Pandey, 2009). The land use land cover (LULC) map was prepared from Landsat 8 OLI image (date of acquisition February 11, 2016). Maximum-likelihood based supervised classification along with preprocessing and post-processing techniques were applied to prepare a reliable LULC map. Rural settlements that are more vulnerable to floods were digitized carefully from Google Earth Image and used in post-processing of the classified image. The runoff coefficient was prepared by overlay analysis of LULC, soil, and slope layers. A look-up table was generated for the overlay output and each combination was assigned a coefficient value based on literature (Mahmoud, Mohammad, & Alazba, 2014; Merz, Blöschl, & Parajka, 2006). The spatial layer of lithologic conditions was prepared by digitizing the geological map published by the Geological Survey of India.

3.2 | Experts’ scoring for criteria weights

The AHP method is a semiquantitative approach that allows criteria weight estimation relying on the experts’ view of the relative importance of criteria against another for flooding. The competence of the AHP method depends on the quality of the subjective judgment of the experts. This study has taken the judgment of five experts from the field of hydrology and disaster management, based on that pairwise comparison tables were prepared. For each comparison matrix, consistency ratio (CR) was calculated and allowed the experts to change their score until CR is below 10% as suggested in past studies (Papaioannou et al., 2015; Wu et al., 2015). The final pairwise matrix (Table 1) was prepared by taking a round-off of the average score given by the experts. The score matrix was then normalized to obtain the corresponding weight of each parameter. Finally, the weight of each parameter was calculated by taking the mean value of the corresponding row in the normalized matrix (Annexure 1).

3.3 | Consistency check

The AHP method requires a consistency check of the weight assigned to criteria (de Brito & Evers, 2016). The following Equation (1) has been used for the evaluation of consistency (Saaty, 1980).

\[
CR = \frac{CI}{RI}
\]

Where CR is the consistency ratio, CI is the consistency index, and RI is the random index that varies with the number of criteria considered. Saaty (1980) provided RI values for a given number of criteria. Accordingly, for this study, RI is 1.41 as eight criteria were chosen for comparison. CI is calculated from the following Equation (2), where \(\lambda_{\text{max}}\) the maximum eigenvalue of the comparison matrix, and \(n\) is the total number of criteria.

\[
CI = \frac{\lambda_{\text{max}}}{n - 1}
\]

The detail mathematical explanation for calculation of \(\lambda_{\text{max}}\) and consistency check is found in Saaty (1990). In

| Parameters                  | Flow accumulation | Draining capability | Elevation | Groundwater depth | Land use | Runoff coefficient | Slope | Geology |
|-----------------------------|-------------------|---------------------|-----------|-------------------|----------|-------------------|-------|---------|
| Flow accumulation           | 1                 | 2                   | 2         | 3                 | 5        | 7                 | 7     | 9       |
| Draining capability         | 1/2               | 1                   | 1         | 3                 | 4        | 5                 | 6     | 7       |
| Elevation                   | 1/2               | 1                   | 1         | 2                 | 3        | 5                 | 5     | 7       |
| Groundwater depth           | 1/3               | 1/3                 | 1/2       | 1                 | 3        | 4                 | 5     | 6       |
| Land use                    | 1/5               | 1/4                 | 1/3       | 1/3               | 1        | 2                 | 4     | 5       |
| Runoff coefficient          | 1/7               | 1/5                 | 1/5       | 1/4               | 1/2      | 1                 | 3     | 5       |
| Slope                       | 1/7               | 1/6                 | 1/5       | 1/5               | 1/4      | 1/3               | 1     | 3       |
| Geology                     | 1/9               | 1/7                 | 1/7       | 1/6               | 1/5      | 1/5               | 1/3   | 1       |
this study, we found $\lambda_{\text{max}} = 8.52$, $n = 8$, CI = 0.074, and CR = 0.05. The AHP theory suggested the weights of a pairwise matrix is consistent if CR value lies below 0.1. Therefore, the CR value affirms the consistency of priority values assigned in this study.

### 3.4 Categorization of parameters for rating

The value range of each parameter was categorized to increase objectivity. The quantitative parameters like...
flow accumulation, draining capability, elevation, and runoff coefficient were clustered using the grading method of natural breaks. On the other hand, qualitative parameters like groundwater depth, land use, and geological settings were classified according to local conditions. The slope map was grouped into classes following the classification of Demek (1972) as it was found very similar to the natural break method in this study. Each of the eight parameters was grouped into five classes, excluding geology that has only three lithological units. Considering the influence of each parameter in flooding, the pixels in each group for each parameter were assigned a value between 2 (for minimum influence) and 10 (for maximum influence; Table 2). For example, pixels with very high flow accumulation were assigned a value of 10.

### 3.5 Flood hazard index

After calculating the weights and ratings of each parameter, the following Equation (3) has been used to prepare flood hazard index (FHI) with the help of the Raster Calculator tool in ArcGIS.

\[
FHI = \sum_{i=1}^{j=1}^{5} r_{ij} \cdot w_{ij}
\]

Where \( r_{ij} \) is the rating of \( i \)th parameter for \( j \)th category, \( w_j \) is the weight of the \( i \)th parameter obtained through AHP, and \( n \) is the number of criteria (parameters). Finally, the values of the FHI raster layer were categorized into five classes (natural break method) to obtain hazard zones.

### 3.6 Inundation mapping

Landsat 8 OLI data (acquisition date August 20, 2016) that was captured during peak flood period has been used for mapping the flood extent. Atmospheric correction was not performed as more than 98% of the image covering the study domain was cloud-free. First, the digital number (DN) of the data was converted into surface reflectance. Followed by NDWI was calculated from Band 3 (Green) and Band 6 (SWIR1) as a ratio between two bands (Green-SWIR)/(Green+SWIR). Instead of Band 5 (NIR), the study used SWIR because it is capable to separate built-up from water (Du et al., 2014). NDWI, therefore, was classified into a binary raster layer of water and nonwater classes based on a threshold. The threshold was derived following the Otsu method (Otsu, 1979) as it was successfully used in many studies for surface water mapping (Du et al., 2014; Li, Du, et al., 2013). The optimal threshold (\( t^* \)) for NDWI was calculated with the help of python programming using the following Equation (4) (Otsu, 1979).

\[
\begin{align*}
\sigma^2 &= P_{nw} \cdot (M_{nw} - M)^2 + P_{w} \cdot (M_{w} - M)^2 \\
M &= P_{nw} \cdot M_{nw} + P_{w} \cdot M_{w} \\
P_{nw} + P_{w} &= 1 \\
t^* &= \text{Arg Max}_{0 \leq t \leq 1} \left\{ P_{nw} \cdot (M_{nw} - M)^2 + P_{w} \cdot (M_{w} - M)^2 \right\}
\end{align*}
\]

Where \( \sigma \) is the between-class variance of the water and non-water, \( M \) is mean value of NDWI, \( P_{nw} \) is the probability of a non-water class, \( P_{w} \) is the probability of water class, \( M_{nw} \) is mean value of the non-water class, and \( M_{w} \) is the mean value of water class.

The monsoon period (June–September) is a peak season for harvesting paddy crop in India. The presence of moisture in crop leaf, and water in the agricultural field during the monsoon season is quite normal. Therefore, to avoid ambiguity in separating cropland from inundated area, a mask of the non-vegetated area was prepared from a normalized difference vegetation index (NDVI). It is noted that the non-vegetated area in the NDVI may also include crop fields that are more or less flooded as the image was captured during peak flood. The NDVI was prepared from the same satellite image as a ratio of NIR and IR bands as \((\text{NIR-IR})/(\text{NIR} + \text{IR})\). Similar to NDWI, a threshold for NDVI was estimated using the Otsu method described in Equation (4) by replacing water and non-water with vegetation and non-vegetation. The threshold was used for the binary classification of NDVI layer into vegetation (above threshold) and non-vegetation (below threshold) classes. Finally, the inundated area was identified from NDVI layer using a query that finds a pixel as flooded if NDVI value is higher than the threshold but must have an NDVI value below the NDVI threshold (i.e., non-vegetate pixel). The delineated inundation map was used for validation of the MCDA approach.

### 4 RESULTS

#### 4.1 Parameters

##### 4.1.1 Flow accumulation

The flow accumulation was found as the most important parameter through the expert survey as it indicates the degree of surface flow concentration. The flow accumulation values increase in the downstream as concentration...
of flow increases. Therefore, the downstream is more prone to flood as their flow accumulation is very high. The calculation of flow accumulation at an administrative level is erroneous as it is estimated at basin-scale. The study did not estimate flow accumulation for the whole Ganga basin, rather for a subset of the study domain. Because the DEM processing for this large river basin is cumbersome and time-consuming, and the actual value of flow accumulation is not essentially required for this study. Result shows that the highest flow accumulation value concentrated at the outflow of the Ganga river as most of the tributaries drain into it (Figure 2). This result is literally true though the absolute value of accumulation was not obtained.

4.1.2 Draining capability

The concept of draining capability is introduced in this study as flooding in the Ganga floodplain is commonly caused by waterlogging due to lack in draining excess water. Thus, the use of drainage density alone, as many previous studies applied (Ouma & Tateishi, 2014; Wu et al., 2015), may delude the analysis. Here, areas with a combination of very low to low elevation and very low to moderate drainage density were considered a very high priority for flooding, while the combination of very high elevation with low to very high drainage density was assigned lowest ranking for flooding (Table 2). The spatial variation in draining capability (Figure 2) shows that a large part of the district has a low draining capability.

4.1.3 Elevation

The elevation is one of the most crucial parameters for flood susceptibility in floodplains (Ouma & Tateishi, 2014). Areas with low altitude are more vulnerable to flood due to logging of water drained from surrounding uplands. In this study, the elevation range was categorised into five groups and assigned a value between 2 and 10 at an equal interval (Figure 2 and Table 2). Areas along the valley of Ganga, Yamuna and Tons rivers have very low elevation (69–90 m), and a large portion of the central and northern part of the district has a low elevation (90–101 m). These parts are considered as high potential for flooding, therefore assigned by ratings of 8 and 10 for low and very low elevation zone, respectively.

**FIGURE 2** Thematic map of parameters used for multi-criteria analysis
4.1.4 | Groundwater depth

The depth of groundwater is intrinsically related to infiltration capacity (Moussa, Voltz, & Andrieux, 2002). Low depth of the water table from surface may allow less water to infiltrate down during peak rainfall period as compared with that of high depth. Therefore, for a surface with low groundwater depth may cause high overland after heavy rainfall that may increase the probability and duration of waterlogging. We assigned a very high priority (10) for a groundwater depth less than 2 m, and very low priority (2) for a groundwater depth greater than 8 m (Table 2 and Figure 2).

4.1.5 | Land use

Land use controls hydrological processes in many ways, like forest cover favors interception and infiltration, thus less runoff generation, while built-up area helps to generate surface runoff (Kazakis et al., 2015). We assigned water as a very high priority for flooding as the maximum flow is concentrated in the river, while built-up and sandbar assigned as a high priority for runoff generation, as well as for flooding (Table 2 and Figure 2). A significant part of the study area is covered by cropland that has been assigned with six considering moderate influence for the flood.

4.1.6 | Runoff coefficient

The runoff coefficient is a measure that expresses the ratio between direct runoff and rainfall during an event (Viglione, Merz, & Blöschl, 2009). Hence, it is considered as one of the crucial criteria for controlling flood mechanism. This parameter is mainly used to include the influence of soil and slope on runoff generation besides land use. We combined land use, slope, and soil for calculating runoff coefficient that varies from 0.08 to 1 (Table 2). The very high value of runoff coefficient was found for water body and built-up area, whereas a very low coefficient value was found for vegetation cover in low slope and high permeable soil (Figure 2). The Very high coefficient value indicates the maximum probability of flood, and vice versa.

4.1.7 | Slope

The surface slope controls the velocity of overland flow, as well as the concentration of flow. High slope helps to drain water quickly, while low slope may lead to the stagnation of water, and causing flooding. As the study area is located over the floodplain region, surface slope is very less. Therefore, most of the district has a low slope (<5%) that was assigned as high to very high parameter ratings (Figure 2).

4.1.8 | Geology

The capacity of infiltration and runoff of an area is also controlled by the underlying geological structure. As alluvium has a better permeability than sandstone, it was assigned a rating of 4 while Upper Vindhyan sandstone was rated as 8 (Figure 2). The study has found geology as least significant in flood generation as it may be incapable to exert strong control on spatial variability.

4.2 | Flood hazard index

The weights of criteria obtained through AHP for MCDA are presented in Table 2. Combining eight parameters (Figure 2) and their corresponding weights, FHI was created. The hazard index was classified into five classes (Figure 3a). By majority 37.64% of the study area lies in the high flood hazard zone, followed by very high-hazard and moderate-hazard zones that cover 27.17 and 21.56% area, respectively. Only 5.80 and 7.85% area lie in very low and low flood hazard zones, respectively.

The distribution of land use types on each flood hazard zone is presented in Figure 4. The analysis shows that 75% of the area in very high hazard zone is shared by agricultural land, while built-up and barren land share 6 and 7% area, respectively. Overall, agricultural land was found most vulnerable for floods in all classes of hazard zones. Following to agricultural land, forest cover was found more vulnerable to flood lying in very low to high hazard zones. Most of the built-up area lies between moderate and very high hazard zones.

4.3 | Inundation mapping

A ground scenario of inundation was identified from Landsat-8 OLI image for the peak flood period on August 20, 2016. After testing several band combinations for NDWI calculation, we found Green and SWIR1 bands very useful to extract water pixels for its capability to remove the noise like built-up, shallow cloud, and so on. The estimated thresholds obtained by the Otsu method are 0.014 for NDWI and 0.21 for NDVI. Finally, the inundated area was identified from the area where
NDWI value is greater than 0.014, and NDVI value is less than 0.21.

The inundation map shows that a large portion surrounding the Ganga valley and south-southwestern part of the district were affected by flooding (Figure 3b). The analysis of inundation signifies that majority of the flooded area was agricultural land (68%), followed by forest cover (12.5%) and sandbar (9.4%). The area of inundation covered by barren land and built-up area was each by 1.7%.

4.4 Validation

The flood hazard zone developed from the MCDA method was tested for validation with the satellite-based historical inundation map. The comparison shows that about 32 and 49% area of total inundation lie in high hazard and very high hazard zones, respectively; while, about 5 and 10% area of inundation comes under low and moderate hazard zones, respectively. Only 4% of the inundated area lies in the very-low hazard zone.

The comparisons in spatial distributions of land use types over index-based hazard zones and satellite data-derived inundation map indicate that agricultural area is most vulnerable to floods for both scenarios, followed by forest cover. The average share of each land use type (excluding water and sandbar) for moderate to very high hazard zone matches very closely ($R^2 = 0.99$) with the same for satellite data-derived inundation map. As sandbar and water mostly lie in a very high hazard zone, the comparisons are made for only this group. The share of sandbar and water in a very high hazard zone is 7 and 5%, respectively, while respective statistics for inundation are about 9 and 7%. Overall, the analysis showed that the estimated index-based flood hazard zones are quite realistic and closely match with the inundation event.

5 DISCUSSION

The present study used the MCDA approach for mapping flood hazard areas in an agricultural region located in the Middle Ganga plain. The methodology adopted for satellite data-derived flood mapping of 2016 has shown a good conformity with the index-based hazard map.

The most commonly applied approaches of flood hazard mapping are parametric (index-based) approach and
numerical methods. Both of these methods have advantages and disadvantages in terms of data quality and quantity, model structure, uncertainty and application. The index-based approach is more suitable to understand the flood susceptibility and vulnerability rather than the simulating spatial extent (Balica, Popescu, Beevers, & Wright, 2013). MCDA-based parametric approach is used in this study due to the simplicity of the methods and focus on subjectivity which is very essential in flood management.

The spatial distribution of flood hazard zones showed a very close resemblance with the criteria layers of draining capability and elevation (Figure 2 and 3). The study area is commonly affected by the floods as it located in floodplain characterized by small scale topographic differences. The study may recommend for considering the draining capability as an important criterion for hazard mapping in floodplain regions. It may be argued that the inclusion of a parameter for preparing different criteria may affect the final decision. For example, the study used land use and runoff coefficient as two different parameters, where land use is a common parameter. Land use is used as a single parameter to emphasize the role of land use as a single parameter, while the runoff coefficient is used for incorporation of the influence of soil types and slope to determine flood susceptibility. It is noteworthy that agricultural land contributes a significant amount of runoff during peak flood. But, as it lies in almost flat surface, its runoff coefficient is very less (Figure 2) that may affect the final hazard output. Hence, to make a more realistic hazard map the study used the runoff coefficient along with land use by decomposing its direct effect.

One of the major conflicts in the MCDA approach arises in data scaling, either to choose a clustering technique or to choose a normalization technique. Papaioannou et al. (2015) found linear normalization to be the best approach for data scaling rather than the
clustering technique. But, the difficulty arises when the study incorporates qualitative parameters that could not be normalized. However, the application of clustering techniques (like natural breaks) is also subject to limitations as it reduces data variability (de Brito & Evers, 2016). In the AHP-based MCDA approach criteria, weights are also dependent on the opinion of the experts. Therefore, the reliability of the estimated hazard index depends on the subjective knowledge of the expert. Another limitation of this study is not employing sensitivity analysis that was used in a few studies (Kazakakis et al., 2015). However, we found a very good accordance between the index-based hazard map and satellite-based inundation maps.

This study has validated the MCDA approach with satellite data-derived inundation map of a single event of extreme floods, rather than a composite layer of flood events (Skakun, Kussul, Shelestov, & Kussul, 2014). But, this approach is very limited for the present study area as the extreme events had occurred before 1985 for which satellite data is rarely available. The spatial extent of flooding for several return periods could be derived using numerical modeling that may improve the validation. But, hydrodynamic modeling in floodplains is very difficult due to the complexity in water flow simulation under minute elevation differences and local depressions (Yamazaki et al., 2012).

### 6 | CONCLUSION

The present paper has adopted a GIS-based multi-criteria evaluation framework to identify flood hazard zones at the administrative level for Allahabad district. The study has tried to incorporate AHP-based criteria weights and parameters ranking for decision-making in complex relations. Overlaying the thematic layers of eight criteria a hazard index was prepared and validated with historical inundation map prepared from remote sensing data.

Eight parameters related to hydro-geomorphological characteristics were prepared in GIS environment and clustered into five groups for assigning ratings based on their influence on floods. The criteria and their weights were combined linearly to obtain the final hazard map. The analysis showed that about 65% of the study area lies in high to a very high hazard zone of which agricultural land is most vulnerable to floods.

An inundation map was prepared from NDWI by processing a Landsat 8 OLI image that was captured during a peak flood in 2016. To minimize the ambiguity in wrongly identifying of non-flooded water in cropland as water, a raster query was applied using the thresholds of NDWI and NDVI. The inundation map shows that about 21% of the study area was flooded, of which 81% of the area belongs to high and very high hazard zone. Besides, the patterns of affected areas under each land use type in the inundation map and flood hazard zones are quite similar. Therefore, the analysis confirms the credibility of the present method, that is, MCDA-based hazard index to identify flood hazard areas.

Application of remote sensing data coupled with GIS tool is very efficient for preliminary analysis of flood hazard. But, the availability of suitable datasets and the data resolution plays vital role in the reliability and applicability of the present method. The study used ASTER DEM of ~30 m spatial resolution that may be considered as coarser for the floodplain region. The analysis could be fine-tuned if higher resolution data were incorporated. The accuracy of the present approach would be increased with incorporation of intensive field data like soil properties, groundwater depth, local drainage, and so on. Similarly, the derivation of the flooded area from satellite images also requires field comparisons to increase its reliability. Moreover, the unavailability of remote sensing data at high temporal scale is a major hurdle for flood mapping. The inundation mapping in terms of depth, magnitude, and extent of flooding for a substantial period could be incorporated in further studies for a comprehensive justification of the adopted methodology.

The present study suggests that the GIS-based MCDA method can be very effective for mapping flood hazards that may be beneficial for decision making in flood management. The methodology employed here can also be applied for data-limited areas anywhere in the world.

### ACKNOWLEDGEMENT

The authors are grateful to Chancellor, Adamas University, and Prof. Milap Punia (JNU, India) for providing support and valuable suggestions in this research work. NASA-METI and USGS are highly acknowledged for providing freely available data. We also acknowledge the Editor in Chief (Dr. Paul Samuels), Associate Editor and anonymous reviewers for their valuable comments and efforts to improve the quality of the manuscript.

### DATA AVAILABILITY STATEMENT

The processed data will be provided to the applicant considering the need and purpose of utilization.

### ORCID

Pratik Dash https://orcid.org/0000-0001-5290-2784

### REFERENCES

Aggarwal, P. K. (2008). Global climate change and Indian agriculture: Impacts, adaptation and mitigation. *Indian Journal of Agricultural Sciences*, 78(11), 911.
Asare-Kyei, D., Forkuor, G., & Venus, V. (2015). Modeling flood hazard zones at the Sub-District level with the rational model integrated with GIS and remote sensing approaches. *Water, 7*(7), 3531–3564. https://doi.org/10.3390/w7073531

Balica, S. F., Popescu, I., Beever, L., & Wright, N. G. (2013). Parametric and physically based modelling techniques for flood risk and vulnerability assessment: A comparison. *Environmental Modelling & Software, 41*, 84–92.

Bhatt, C. M., & Rao, G. S. (2016). Ganga floods of 2010 in Uttar Pradesh, North India: A perspective analysis using satellite remote sensing data geomatics. *Natural Hazards and Risk, 7*(2), 747–763.

Burby, R. J., Deyle, R. E., Godschalk, D. R., & Olshansky, R. B. (2000). Creating hazard resilient communities through land-use planning. *Natural Hazards Review, 1*, 99–106. https://doi.org/10.1016/(S1106-9488(00)10129-9)

Census of India (2011). *District census handbook—Allahabad*. Directorate of Census Operations, Ministry of Home Affairs, Government of India. Retrieved from http://censusindia.gov.in/2011-common/censusdataonline.html.

Chen, H., Ito, Y., Sawamukai, M., & Tokunaga, T. (2015). Flood hazard assessment in the Kujukuri plain of Chiba prefecture, Japan, based on GIS and multicriteria decision analysis. *Natural Hazards, 78*(1), 105–120.

Chen, Y. R., Yeh, C. H., & Yu, B. (2011). Integrated application of the analytic hierarchy process and the geographic information system for flood risk assessment and floodplain management in Taiwan. *Natural Hazards, 59*, 1261–1276. https://doi.org/10.1007/s11069-011-9831-7

de Brito, M. M., & Evers, M. (2016). Multi-criteria decision-making for flood risk management: A survey of the current state of the art. *Natural Hazards and Earth System Sciences, 16*(4), 1019–1033.

Demek, J. (1972). *Manual of detailed geomorphological mapping*. Prague: Academia.

Detrembleur, S., Stilmant, F., Dewals, B., Erpicum, S., Archambeau, P., & Pirotton, M. (2015). Impacts of climate change on future flood damage on the river Meuse, with a distributed uncertainty analysis. *Natural Hazards, 77*, 1533–1549. https://doi.org/10.1007/s11069-015-1661-6

Dewan, A. M., Islam, M. M., Kumamoto, T., & Nishigaki, M. (2007). Evaluating flood hazard for land-use planning in greater Dhaka of Bangladesh using remote sensing and GIS techniques. *Water Resources Management, 21*, 1601–1612. https://doi.org/10.1007/s11269-006-9116-1

Du, S., Shi, P., Van Rompaey, A., & Wen, J. (2015). Quantifying the impact of impervious surface location on flood peak discharge in urban areas. *Natural Hazards, 76*, 1457–1471. https://doi.org/10.1007/s11069-014-1463-2

Du, Z., Li, W., Zhou, D., Tian, L., Ling, F., Wang, H., ... Sun, B. (2014). Analysis of Landsat-8 OLI imagery for land surface water mapping. *Remote Sensing Letters, 5*, 672–681. https://doi.org/10.1080/2150704X.2014.960606

Dutta, D., Alam, J., Umeda, K., Hayashi, M., & Hironaka, S. (2007). A two-dimensional hydrodynamic model for flood inundation simulation: a case study in the lower Mekong river basin. *Hydrological Processes: An International Journal, 21*(9), 1223–1237.

Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated water extraction index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment, 140*, 23–35. https://doi.org/10.1016/j.rse.2013.08.029

Giustarini, L., Hostache, R., Kavetski, D., Chini, M., Corato, G., Schlaffer, S., & Matgen, P. (2016). Probabilistic flood mapping using synthetic aperture radar data. *IEEE Transactions on Geoscience and Remote Sensing, 54*, 6958–6969. https://doi.org/10.1109/TGRS.2016.2592951

Hazarika, N., Barman, D., Das, A. K.,arma, A. K., & Borah, S. B. (2016). Assessing and mapping flood hazard, vulnerability and risk in the upper Brahmaputra River valley using stakeholders’ knowledge and multicriteria evaluation (MCE). *Journal of Flood Risk Management, 11*, S700–S716. https://doi.org/10.1111/jfr3.12237

Huan, H., Wang, J., & Teng, Y. (2012). Assessment and validation of groundwater vulnerability to nitrate based on a modified DRASTIC model: A case study in Jilin City of Northeast China. *Science of the Total Environment, 440*, 14–23. https://doi.org/10.1016/j.scitotenv.2012.08.037

Jenks, G. F. (1967). The data model concept in statistical mapping. *International Year Book of Cartography, 7*, 186–190.

Kazakis, N., Kougias, I., & Patsialis, T. (2015). Assessment of flood hazard areas at a regional scale using an index-based approach and analytical hierarchy process: Application in Rhodope-Evros region, Greece. *Science of the Total Environment, 538*, 555–563. https://doi.org/10.1016/j.scitotenv.2015.08.055

Kron, W. (2005). Flood risk = hazard values vulnerability. *Water International, 30*(1), 58–68.

Kuldeep, Garg, P. K., & Garg, R. D. (2016) Geospatial techniques for flood inundation mapping. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). pp. 4387–4390.

Lee, A. H. I., Chen, W.-C., & Chang, C.-J. (2008). A fuzzy AHP and BSC approach for evaluating performance of IT department in the manufacturing industry in Taiwan. *Expert Systems with Applications, 34*, 96–107. https://doi.org/10.1016/j.eswa.2006.08.022

Li, G. F., Xiang, X. Y., Tong, Y. Y., & Wang, H. M. (2013). Impact assessment of urbanization on flood risk in the Yangtze River Delta. *Stochastic Environmental Research and Risk Assessment, 27*(7), 1683–1693.

Li, W., Du, Z., Ling, F., Zhou, D., Wang, H., Gui, Y., ... Zhang, X. (2013). A comparison of land surface water mapping using the normalized difference water index from TM, ETM+ and ALI. *Remote Sensing, 5*, 5530–5549. https://doi.org/10.3390/rs5115530

Liu, Y., Zhang, W., & Cui, X. (2012) Flood emergency management using hydrodynamic modelling. In: Procedia Engineering. pp 750–753.

Madhuri, B., Aniruddha, G., & Rahul, R. (2013). Identification and classification of flood prone areas using AHP, GIS and GPS. *Disaster Advances, 6*, 120–131.

Mahmoud, S. H., Mohammad, F. S., & Alazba, A. A. (2014). Determination of potential runoff coefficient for Al-Baha region, Saudi Arabia using GIS. *Arabian Journal of Geosciences, 7*, 2041–2057. https://doi.org/10.1007/s12517-014-1303-4

Mani, R. (2016). Flood situation worsens in Allahabad. *The Times of India*. Retrieved from: https://timesofindia.indiatimes.com/city/allahabad/Flood-situation-worsens-in-Allahabad/articleshow/53812688.cms.
