International Journal of Multidisciplinary Research and Growth Evaluation.

River bank erosion induced social vulnerability index assessment using factor analysis and logistic regression model: A study on Manikchak block in West Bengal, India

Samrat Majumdar [1*], Sujit Mandal [2]

[1] Research Scholar, Department of Geography, University of Gour Banga, West Bengal, India

[2] Professor and HoD, Department of Geography, Diamond Harbour Women's University, West Bengal, India

* Corresponding Author: **Samrat Majumdar**

Article Info

ISSN (online): 2582-7138
Volume: 03
Issue: 04
July-August 2022
Received: 03-07-2022;
Accepted: 21-07-2022
Page No: 299-309

Abstract

Social vulnerability indices are a way of gathering information on people who may be impacted by disasters such as river bank erosion. The development and validation of a social vulnerability map of population characteristics towards bank erosion covering Manikchak blocks in Malda district in West Bengal, India, is the goal of this project. This map is based on a composite index of three key indices of social vulnerability in Diara blocks: fragility, socioeconomic conditions, and geographic location. A factor analysis of selected demographic variables derived from the District Statistical Handbook was used to identify these indicators. As a result, depending on a reputable data source, these indicators can be updated annually. An impartial second data set is used to confirm the susceptibility patterns discovered by factor analysis. The second data set's interpretation suggests that a real extreme river bank erosion event reveals vulnerability and that the patterns of the presumed vulnerability match the observations of a real occurrence. It includes a survey of erosion-affected Manikchak families. It is shown that the theoretically assumed signs of susceptibility are right and that the indicators are valid using logistic regression. It has been demonstrated that particular social groups, such as the elderly, the financially disadvantaged, and city dwellers, are at increased risk.

Keywords: iver bank erosion, Factor analysis, Logistic regression, Social Vulnerability Index (SVI), Rotated Component Matrix

1. Introduction

Natural catastrophes typically result in unequal economic losses and mortality across and within nations, regions, communities, and individual groups. Vulnerable groups are those who are more likely to be affected disproportionately by hazardous events. The vulnerability of populations and communities to natural catastrophes is defined not only by their proximity to the source of risk, but also by their social vulnerability status. Social vulnerability refers to population characteristics that influence a community's ability to plan for, respond to, and recover from disasters (Cannon 1994) [1]. During disasters, socially marginalized populations are less likely to have access to crucial supplies. As a result, knowing social vulnerability can help explain why various communities may react differently to the same hazard event (Morrow 2008). Understanding the diverse effects of hazard events is essential for reducing the harmful consequences of natural catastrophes. As a result, over the last two decades, there have been a number of attempts to quantify social vulnerability, including conceptual, methodological, and practical issues (McCarthy et al. 2003; Birkmann 2006) [2, 3]. However, there is currently no agreement on the optimum quantitative methodology for assessing social vulnerability. For example, poverty (Fothergill and Peek 2004; Long 2007) [4, 5], race and ethnicity (Fothergill et al. 1999; Peacock et al. 2000) [6, 7], gender (Enarson and Morrow 1998; Enarson et al. 2006) [8, 9], and age (Enarson and Morrow 1998; Enarson et al. 2006) [8, 9] all influence vulnerability (Bolin and Klenow 1983; Ngo 2001; Anderson 2005; Phillips and Hewett 2005; Kar 2009; Smith et al. 2009) [10, 11, 12, 13, 14, 15]. Without all or most of these characteristics, an assessment of vulnerability is likely to be insufficient; consequently, this measure must be a composite measure or index (Adger et al. 2004; Gall 2007; Barnett et al. 2008) [16, 32, 17].

Bank erosion is a critical component of fluvial dynamics, affecting a wide range of physical, ecological, and socioeconomic issues. These include the establishment and evolution of river and floodplain morphology and associated habitats (e.g., Hooke, 1980; Millar and Quick, 1993; Darby and Thorne, 1996a; Barker et al., 1997; Millar, 2000) [18-22], turbidity issues (e.g., Bull, 1997; Eaton et al., 2004) [26, 30], sediment, nutrient, and contaminant dynamics (e.g., Romanelli et al., 2004) [47], (e.g. Simon, 1995). On average, the Ganga carries 800 million tonnes of sediments upstream of the barrage each year, with an estimated 13 lakh ha-m of sediments already deposited upstream of the barrage, resulting in the formation of several shoals/bed bars, meandering, cross-slope, strong flow curvature, and lateral flow instability (Sanyal, 1980) [53]. The stability of minor rills, subchannels, and rivers is generally improved by the root system created in the banks by vegetation nurtured by human population. However, in big systems when river currents affect the banks below this root and the added subcharge weight of the vegetation contributes to bank failure, this may have an undesirable effect. When trees fall into rivers, the debris may cause currents to be deflected, producing significant erosion along the banks. In some parts of the study region, particularly along the banks where tree planting is common, this erosion acceleration mechanism is highly important (Gupta and Nandy, 1982) [33]. Agriculture is practised, and the land along the banks is extremely fertile, attracting large-scale farming operations. The surface of the soil or land becomes loose during the preparation of agriculture fields. As a result, during the rainy season, rain water easily penetrates the soil particle and acts as seepage pressures, causing the soil particles to unconsolidated and lose their cohesiveness. As a result, fast degradation could occur (Sesoren, 1984) [55]. Bank geological qualities are one of the most important determinants of bank stability (Rogers et al., 1989) [46]. The height of the water level is related to all other elements that contribute to bank erosion.

The existence of Calcutta Port is dependent on the Farakka Barrage. The construction of the Farakka barrage began in 1962 and ended in 1971. The feeder canal took four more years to complete, and the project was dedicated to the country on May 21, 1975. (Banerjee 1999) [22]. The 2.64-kilometer-long Farakka Barrage was built to channel 40,000 cusecs (or 1,133 cumecs) of Ganga water toward the Bhagirathi River in order to flush sediment into the estuary's deeper parts and restore the port's navigational status (Rudra 2004) [51]. It was constructed with the intention of causing water to flow into the Hugli River. The dam appears to be causing the river to go her own way, mostly upstream of the Farakka barrage. Malda and Murshidabad, two West Bengal districts, are the hardest hit (Banerjee and Chakroborty 1983; Banerjee 1999) [21, 22]. In West Bengal, the change in Ganga river course and the accompanying river bank failure is a long-term natural calamity. As evidenced by numerous studies, this has been a persistent problem since the early 1960s, and it has grown to enormous proportions in the subsequent four decades (Rudra 1996a, b, 2004; Banerjee 1999; Mukhopadhyay 2003) [49, 22, 41]. Showkat investigated the socio-economic effects of floods and accompanying land erosion in this area in depth (2010). The majority of the erosion-affected region in the upstream lies within the Malda district, namely four blocks: Manikchak, Kaliachak-II, Kaliachak-III, and Ratua-I. Since the 1960s, this river bank erosion hazard has eroded a large region.

The river's flow rate causes high and frequent bank erosion, resulting in a large quantity of river bank cutting and population relocation from the river's vicinity every year (Iqbal 2010) [36]. The shift in river course was followed by changes in river morphometry, such as overall width, sinuosity, braiding features, and so on (Parua 2002) [44]. It also changes the geometry of the boundary between Malda and Murshidabad districts, as well as between West Bengal and Jharkhand, and India and Bangladesh. It also causes social and political turmoil. 2009 (Parua). In Malda district, there is yet no comprehensive profile of socioeconomic vulnerability or sub-national index map for the entire river system. The Social Vulnerability Index (SVI) in relation to river-floods will fill this gap.

Despite the fact that social vulnerability indices are becoming more widely used, there have been little attempts to validate them. Validation can be done with a second, independent data set, preferably with a finer spatial resolution. Alternatively, statistical tests can be used to assess the model's internal reliability in terms of sensitivity and uncertainty. This demonstrates that an index's validity has two aspects: conceptual and methodological validity. The importance of conceptual validity in the substance of social vulnerability is growing, but it is becoming more difficult to achieve and assess. It necessitates terminological clarity, empirical proof, and a social vulnerability theoretical framework.

Attempting to validate a social vulnerability index is limited by a number of factors. To begin with, scientific evidence on social vulnerability is difficult to come by. Social vulnerability is frequently concealed, complicated, and nested in a variety of human features and variables tied to many levels of society. Second, the concept of vulnerability is conceived in at least two ways. On the one hand, it is seen as a comprehensive and generic notion that encompasses a wide range of intricate interrelationships. On the other hand, it is viewed as a more monolithic idea, focused on a particular item in relation to a certain hazard, such as river-flood evacuation needs. Vulnerability is better suited for experimental tests depending on the conceptualization. Third, estimating societal vulnerability is challenging due to methodological issues. Indicators and indices are numerical substitutes for real-world occurrences. In order to obtain computation and comparability, quantitative assessments of qualitative phenomena are subjected to generalizations.

Because of these difficulties, it's not unexpected that most social vulnerability index examinations focus primarily on index generation rather than extra validation. The availability of data limits the building of an index (King, 2001) [37], and independent second data sets are limited. Technical validation of such indices is becoming more common, for example, through sensitivity and uncertainty analysis (Wu et al., 2002) [58] or random simulation testing of index robustness using Monte Carlo simulations (Gall, 2007) [32]. Finding empirical evidence of social vulnerability is often the focus of validation at other spatial levels. When the danger settings, as well as the spatial and cultural environment, are the same, this strategy can be used (O'Brien et al., 2004) [43]. When the same theoretical framework is used, the feasibility and success of such a validation at several scales is even more feasible.

The objective behind estimation of SVI and its validation is to find evidence whether the construction of a Social Vulnerability Index without direct relation to disaster impact or hazard parameters is valid. That means that

- The identification and analysis of various factors of SVI (social vulnerability index) using factor analysis.
- Assessment of dependent and independent variables and their validation through various common indicators.
- The Correlation between two variables (SVI) using logistic regressions and assessment of accuracy result.
- The identification of various mouzas of Manikchak Diara of Malda district having high SVI, moderate SVI and low SVI.

2 The Study Area

The district is located between Latitudes 24°40'20''N and 25°32'8''N, and Longitudes 87°45'50''E to 88°28'10''E (Fig. 1), and is bordered on the south by Murshidabad district, on the north by Uttar Dinajpur district, on the east by Bangladesh, on the west by the state of Bihar, on the north east by Dakshin Dinajpur district, and on the southwest by Jhar The district covers around 3,733 square kilometres and is located 365 kilometres from Kolkata, the state capital. Malda is divided into 15 blocks and two sub-divisions, Sadar and Chanchal. Englishbazar is the district headquarters.

Within the district, the Diara includes the Manikchak, Kaliachak I, II, and III blocks, as well as the Englishbazar block. Among them, the Manikchak block has been considered. In the transitional zone between the upland and the swampy Tal track, the Diara is a reasonably well-drained flatland formed by alluvial deposition of younger alluvium. The Diara tract is an eight-mile-wide stretch that runs along the western and southern edges of the District. It was formed by millennia of fluvial action by the Ganges, whose old channels may still be traced, starting at the current course of the Bhagirathi River at Gour and extending westwards in stages. The soil is pale in colour and has a sandy texture. This tract covers 112188 hectares.

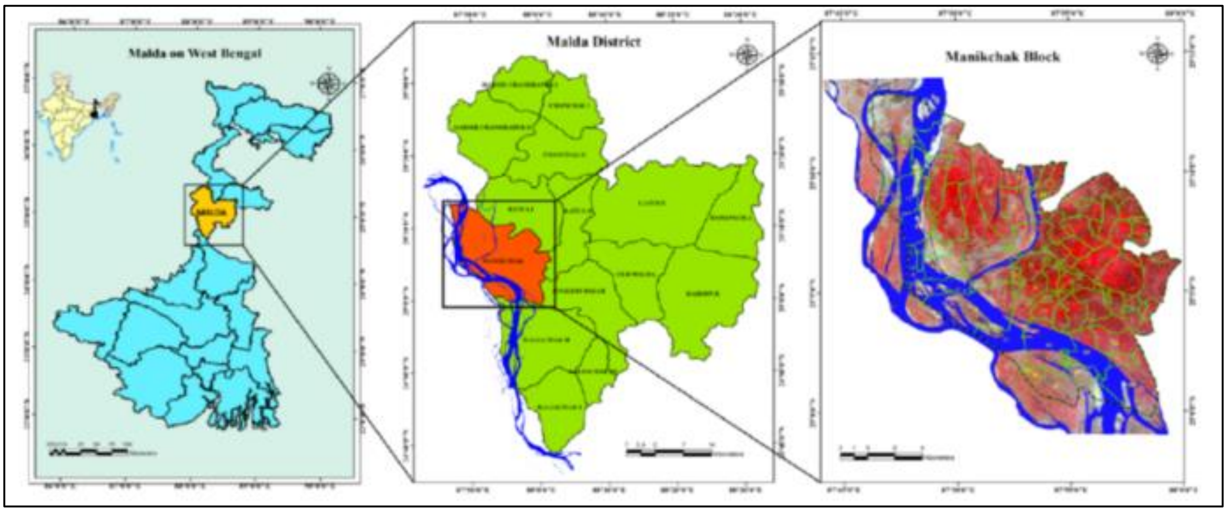

Fig 1: Location of Manikchak Block and river erosion affected areas

Thousands of individuals become landless and lose their livelihood every monsoon season. As a result, their demand for reallocation to the government far exceeds the government's available land resources. As a result, victims must look for property on their own, resulting in neo-refugees with several social issues. Non-affected residents are not eager to welcome the victims. Crime can sometimes rise as a result of the poverty that has been produced. The neo-refugees prefer to build their huts on public works department (PWD) roadside since breaking their huts requires a lengthy procedure, and these roadside huts also result in a high rate of vehicle accidents. Following the reallocation, social groupings as well as conflict between victims and non-affected people have emerged. Previously, there was often a variety of religions, but following reallocation, they all chose the same religion, and religious clustering begins (Thakur et al. 2011) [57].

3 Materials and Methods

3.1 Social Vulnerability Factors

District statistics Handbook and field survey data were utilized to estimate social vulnerability at the district level in West Bengal. The data was first evaluated using a factor analysis, and then the data was validated using a logistic regression model in the second stage. The goal of component analysis is to create profiles of social groupings based on particular traits such as income, gender, and age, which can be related to building type, rural location, and medical treatment. The fundamental goal of component analysis is to decrease variables so that a set of variables those summaries social vulnerability characteristics may be created. In addition, the structures of variable groups can be elicited in order to create a social vulnerability profile.

3.2 Factor analysis

Factor analysis is a multivariate analysis technique for identifying information packing that takes into account all variables' interdependencies (Bernard 2006: 495) [24]. For data reduction and to find variable groups, the factor analysis is performed in SPSS version 14.0 with a principal component analysis. The factor analysis methodology follows a consistent procedure (e.g. Nardo et al., 2005) [42]. The goal of principle component analysis is to discover a linear combination of variables that accounts for as much variance as feasible in the original variables. By rotating the axes of the components perpendicular to each other, a Varimax rotation with Kaiser Normalisation is applied to the component matrix to make interpretation easier (Schneiderbauer, 2007: 55) [54]. This stage separates the various components from one another as much as feasible. All of the extracted communalities are greater than 0.5, indicating that the extracted components accurately describe the variables. Only eigenvalues bigger than one are used in the interpretation, and absolute loading values less than 0.30

are suppressed (Nardo et al., 2005: 40, 43; B uhner, 2006: 200, 211; Bernard, 2006: 677) [42, 25, 24]. The standardized variance associated with a given factor is known as the Eigen value. Another criterion for limiting the number of elements is the scree plot. The components on the steep slope up to the curve's "scree elbow" are particularly capable of explaining the majority of the data.

The factor analysis follows the idea of variance maximization, in which the factors that explain the most variance of all items are sought after (Buhner, 2006: 182). The variable selection is adequate for factor analysis, according to the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) of 0.905. The KMO illustrates how much of the variance in the variables is explained by underlying factors. KMO-scores greater than 0.60 indicate an acceptable level of compatibility with the test, and values greater than 0.80 suggest an excellent level of compatibility (as cited in Buhner, 2006: 207) [24]. The Bartlett's test of sphericity with a result less than 0.05 rejects the hypothesis that the variables are unrelated and thus unsuitable for structure discovery.

The factor analysis uses 41 variables as input (Table 1) to reveal seven latent factors and characterize relationships between all variables, accounting for 76.6 percent of the total variance. Only the first three factors (or components) have more than two loading values, as indicated in the first three columns, namely region, socioeconomic state, and fragility (Table 1). The values of each variable that load the highest in the seven factors matrix are referred to as three factors values. Within the value loadings, factor analysis identifies major latent groupings. These variables can be used to calculate the composite index. The first component is region, which includes factors related to regional urban or rural aspects such as population density and dwelling type (Table 1). This component is made up of variables that include both negative and positive aspects of river bank erosion risk, such as medical care density and population density. The second component represents socioeconomic conditions that are linked to financial deficits or income resources, allowing for the interpretation of specific age and employment groups. In contrast to middle-aged persons, the third component highlights the elderly physical infirmity or need for assistance.

Table 1 Rotated component matrix of the factor analysis showing the computed value loadings

Input Value with presumed direction towards vulnerability: '-'more vulnerable, '+' more capacities	**Component**						
	1	2	3	4	5	6	7
-Residents below age 6		0.773	-0.423				
+Residents from age 30 to 50			-0.85				
-Residents age 65 and older		-0.318	0.882				
-Persons in need of care			0.586			0.377	
-Handicapped unemployed		0.629					
-Female gender	0.632		0.545				
+Income per household		0.767			-0.343		
-Unemployment		0.83	-0.33				
+Female employed	0.821						
+Foreign employed		0.705					
+High qualification employed	0.737	-0.307				-0.329	
-Foreign females		0.828					
-Social welfare recipients	0.433				0.655		
-Rent Subsidies		-0.811					
-Graduate without basic education		-0.415			0.38		0.54
+Graduate with high school graduation	0.74	-0.337					
+University students	0.719						-0.454
-Foreigners	0.597	0.618					
-Residents per doctor	-0.829						
+Hospital beds	0.707					0.348	
+Rural Population	-0.724			0.303			
-Population per settlement area	0.833					-0.358	
+Open space	-0.735				-0.383		
+Building land prices	0.634	0.484					
-Commuters in	0.734						
+New apartment		0.35	-0.681				
+One and two family homes	-0.819						
-Small apartment	0.824			0.378			
+Living space per person	-0.351	-0.483				0.444	
-Person per household	-0.756		-0.376				
-New residents	0.697	0.34	-0.369				
-Municipality debts per residents					0.567		
-Tourist overnight stays				0.904			
+GDP per labour force		0.637					0.396
-Key funds allocation		-0.8					
+Fixed investment	-0.375	-0.613			-0.359		
-Day care centre		-0.866					
-Rehabilitation centre per residents				0.84			
-Elementary school per residents	-0.649						
-Medical care centre			0.451			0.618	
-Population projection age 60+		-0.736	0.58				
-Interpretation							

-Positive value loading	Rural	Young Income	Old Fragile	Tourism	Welfare Debates	Care centres	Low Education
-Negative value loading	Rural	Financial Deficiencies	Middle age, Home worker				
-Percent variance explained	26.10%	22.10%	10.90%	5.40%	4.70%	4.60%	2.80%
Total=76.6%							
Factor name	Region	Socio Economic Condition	Fragility				

Exclusion or selection of variables for input can affect factor analysis. However, assessment of the factor analysis with stepwise exclusion of variables in the model reveals that the 41 set of variables (Table 1) is internally sound. In addition, the factor analysis was done step by step, starting with two variables and gradually adding one more. The patterns found have also been verified through stepwise testing. However, there is no clear correlation between this outcome and danger or disaster factors. An indicator developed by factor analysis that requires either internal statistical reliability analysis, such as Cronbach's alpha, or, better yet, validation by a separate data set.

3.3 Validation of social vulnerability by factor's impact analysis

The endpoint of data analysis for the production of a social vulnerability index is usually the outcome of factor analysis. The results of the factor analysis are tested in this chapter by testing the patterns of the factor analysis on a real-life occurrence. A second data set was chosen to evaluate the social vulnerability profiles on a genuine bank erosion incident. The research topic is whether an actual extreme bank erosion event discloses some of the potential social vulnerability predicted by Table 1's vulnerability criteria.

A testing category must be identified in order to uncover proof that the hypothesized social vulnerability concept and profiles play a role in catastrophe outcome. The question "Did you have to leave your home owing to the flood?" was recognized as a discriminator of those badly affected by the flood in terms of social vulnerability for the purposes of this study. This question encompasses a broader breadth of exposure, susceptibility, and capacities than just the economic standpoint. People who were forced to flee their homes (765 of 1697) were more vulnerable to floods, had to cope with temporary accommodation, and required a recovery period following the flood. They required not just financial resources, but also social networks such as relatives and friends.

A small percentage of those forced to flee their homes (N=765) needed to seek emergency refuge (N=70). This is a particularly intriguing subgroup because it is likely that these individuals lacked other social networks or financial means. These are merely assumptions, given the questionnaire contains no questions concerning the precise reasons behind each individual's decision in the poll. They can, however, be linked to data from evacuation groups on social vulnerability (Cutter et al., 2003; Chakraborty et al., 2005) [28, 27]. As a result, the test categories "people forced to leave their house" and "people compelled to seek emergency shelter" were chosen to elicit various social group profiles. It allows comparisons to be made between those who had to flee and those who were able to stay in their homes despite being flooded.

The final test category is based on the following question: "Are you pleased with the current state of damage regulation?" The responses were graded on a scale of one to six, with one being the most positive and six being the most negative. For bivariate comparison, this range is converted to binary coding. This dependent variable can identify indirect financial demands and satisfaction with administration to some extent. As a result, this sort of susceptibility measure is used in conjunction with the other two dependent variables that capture exposure and evacuation requirements.

3.4 Logistic Regression model

The logistic regression is calculated for each of the three binary dependent variables, i.e. leave home, emergency shelter, and regulation, using the second data set. Each dependent variable is examined using the same subset of independent variables that has been pre-selected (Table 2). The fundamental goal of logistic regression is to determine whether or not the independent variables differ significantly. The dependent variables contain binary yes/no situations, while the independent variables represent demographic vulnerability characteristics (e.g., age of people). For example, inside the full logistic regression model, independent variables like age are compared to the dependent variable leave home to see if age is a feature that distinguishes human groups as more vulnerable.

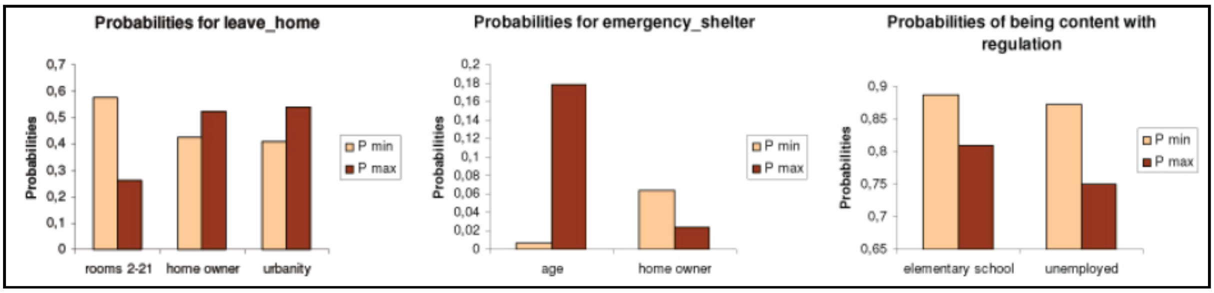

Fig 2: Minimum and maximum probabilities of the three logistic regression

The logistic regression provides two sorts of measurements, both of which are relevant in this context. Rooms, home ownership, urbanity, age, home for emergency refuge, primary school, and unemployment are all independent factors in this model. To begin, the regression model identifies which independent variables are significant across the entire model; only these are used to calculate probabilities (Table 3). Second, the probabilities estimated for each independent variable's minimum and maximum values indicate the dependent variable's influence direction (Table 3 and Fig. 2). This direction can be either positive or negative, implying that flood impact rises with increasing values, such as more income, or is inversely connected.

Table 2: Dependent and independent variables used for all three logistic regressions

Dependent Variables	Independent Variables	Sub Variables
leave home	age	unemployed
emergency shelter	gender	pensioner
regulation	high school degree	residents up to 14 years
	elementary school	persons per household
	income very high	rooms
	income low	home ownership
	high qual employed	urbanity

3.5 Process of Validating Social Vulnerability Index:

The validation technique consists of two steps: first, the independent variables of the factor analysis (Table 1) were tested for validity by running a logistic regression model with the independent variables of the independent second data set. Because the second data collection did not capture all of the same demographic factors as the first, only a few independent variables from the first data set are available from the second data set at the same time (Table 4). Rurality, house ownership, rooms, age, unemployment, and primary school are six independent factors in the second data set that can be used to measure discriminatory vulnerability, according to the logistic regression analysis. These six variables capture demographic and spatial features that are also captured by the first data set's nine independent variables. It signifies that specific variables were validated as having a significant impact on identifying susceptibility in the first step.

Table 3: Calculated significances, probabilities and confidence intervals for the three dependent variables

Dependent Variables	Independent Variables	Significance			Probabilities			
		Sig.	95.0% Lower CI	95.0% Upper CI	P min	P max	95.0% CI Change min	95.0% Change max
leave home	rooms	0.024	0.877	0.991	0.5755	0.2624	−0.5506	−0.0756
	home ownership	0.019	1.066	2.053	0.4272	0.5245	0.0167	0.1779
	urbanity	0.000	1.272	2.261	0.4091	0.5399	0.0608	0.201
emergency shelter	age	0.012	1.010	1.081	0.0067	0.1785	−0.0537	0.3974
	home ownership	0.003	0.175	0.707	0.0636	0.0233	−0.0752	−0.0052
regulation	elementary school	0.019	0.325	0.905	0.8873	0.8102	−0.1453	−0.0089
	unemployed	0.020	0.221	0.881	0.8719	0.7502	−0.2457	0.0024

This knowledge can be applied to a subset of nine variables from the first data set's 41 variables. It would be dangerous to assume that this approach validates the entire 41-variable model. At least nine variables from the first data set, such as population/settlement area, one and two family homes, rural population, small apartments, residents (age 30-50), residents (age more than 65), unemployment, living space, and graduates with an elementary education, can be assumed to describe vulnerability. The other 32 independent variables are either not significant in the regression model or are not testable because they are not included in the second data set. This isn't to say that they can't be relevant in another model or that they aren't meaningful. The factor analysis is conducted with a subset of nine independent variables from the Federal Statistics in the second step of the validation. The goal of the second step of validation is to see if the components (or social vulnerability indicators) that were obtained without any direct disaster-relation are disclosed in the same way by the limited set of nine validated variables.

The three factors produced from the nine variables as grouped in the rotated component matrix (Table 5) exhibit the identical factors that were discovered with the complete variable set of 41 variables with the entire federal statistics set as a result of the validation (Table 1). The fact that the factors are generally valid is revealed by this outcome.

Table 4: Comparison of the nine variables of the Federal Statistics with the according variables of the logistic regression

Variables of the logistic regression	Variables of the factor analysis from the first data set
Rurality (urban areas have more than 150 persons per km2 per municipality)	Population per settlement area
home ownership	One and two family homes
Rurality (rural areas have less than 150 persons per km2 per municipality)	Rural population
rooms [2; 21]	Small apartments
age	Residents from age 30 to 50
age	Residents age 65 and older
unemployed	Unemployment
rooms [2; 21]	Living space pp
elementary school	Graduates with only elementary education

Table 5: Rotated Component Matrix of the nine variables of the federal statistics that are validated by the logistic regression

Component	1	2	3
Population per settlement area	−.951		
One and two family homes	856		−.358
Rural population	.831		
Small apartments	−.788		
Residents from age 30 to 50		−.935	

Residents age 65 and older		.913	
Unemployment		.383	.853
Living space pp	.416		−.716
Graduates with only			.697
elementary education			
Factor name	Region	Fragility	Socioeconomic conditions

The selection and aggregation of the Social Vulnerability Index is based on the elements produced from factor analysis. Each element produces one indication, and the index is made up of the aggregated indicators. Each component must have the ability to indicate both positive and negative indications of vulnerability to the same degree in order to enable both positive and negative indicators of vulnerability to the same degree. First, the variables are normalized to equal intervals between zero and one. Missing values are substituted with the variable's average value, ensuring that there is no trend in the average of all negative or positive variables. Each county's susceptibility is varied as a result of this. The three factors are utilized as social vulnerability indicators.

4 Results

4.1 Analysis of Social Vulnerability Index (SVI)

The SVI (Fig 3) identifies demographic patterns of river-flood sensitivity, capacity, and prospective exposure. It's calculated by adding three indicators: fragility, socioeconomic conditions, and region. The indicator fragility is the ratio of old citizens (>64 years); the indicator socio-economic circumstances is the living space per person, the (un)employment ratio, and the kind of education; and the indicator region is the population density and dwelling type. The results are displayed in ratios per county as equal intervals from zero to one, using data from the Statistical Handbook, Malda.

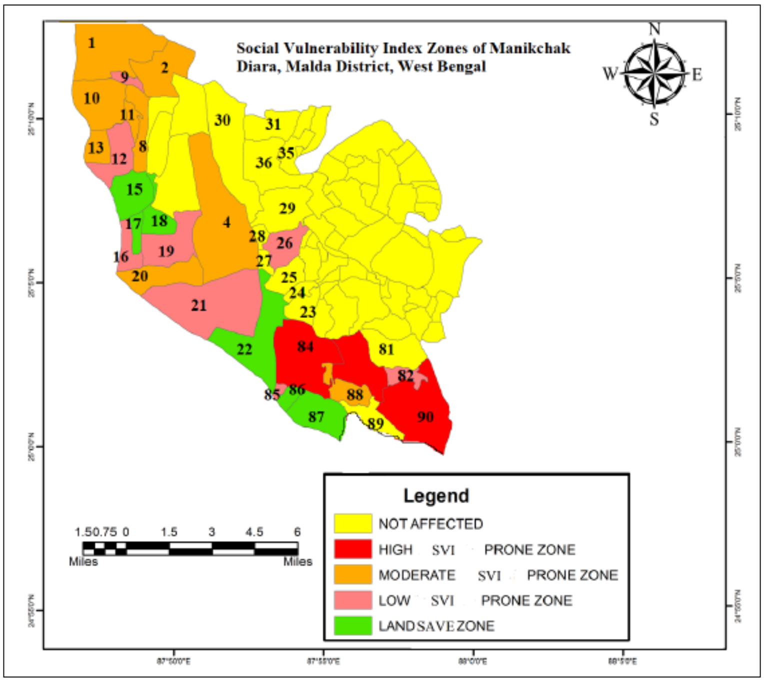

Fig 3: Social Vulnerability Index Zone Map of Manikchak Diara of Malda district, West Bengal

Low SVI areas (Fig. 3) are known for their proclivity for river bank erosion. Existing capacities for erosion mitigation, such as financial capacities for private preparatory measures and recovery from erosion by high-income sources, are examples of these strengths. These counties have no signs of possible erosion, such as high population density. Elderly people's susceptibility, as well as their physical fragility, is often low. Counties with a high SVI have a lot of flaws when it comes to bank erosion. Lack of capacities, high levels of vulnerability, and signs of exposure potential are among these flaws.

The SVI detects future county strengths and weaknesses, not actual river-flood danger or hazard. There is no danger information in the SVI; hence there is no actual exposure. The SVI, on the other hand, is not a natural hazard index because the variables are chosen and aggregated only after flood effect evidence. The indicator input variables are developed once the flood impact on various social groups and settlement types has been validated. Counties have different social vulnerability profiles based on demographics, population density, and settlement form. The SVI's strength is that it is not reliant on direct hazard information. It identifies important components of flood effect and danger that hazard evaluations miss.

For their ability to evaluate multivariate interdependencies between the selected variables, the conventional statistical methods of factor analysis and logistic regression were used. This appears to be an appropriate measure for capturing the multiple interdependencies in the case of social vulnerability, and it allows for the elicitation of previously unknown interrelationships between demographic and spatial variables on the mouza scale in different places of Manikchak Diara of Malda District. While the choice of input variables can cause changes in the component composition, the display of the matrix with the value loadings allows for a straightforward and comprehensible interpretation. Multiple interdependencies of the input factors can be analyzed using logistic regression against particular test categories of social vulnerability outcome. This technique, like factor analysis, is sensitive to the choice of input parameters. The stepwise inclusion and exclusion of input variables, as well as the sensitivity tests, revealed that the main pattern of interdependencies is robust for both methodologies. This combination of methodologies successfully indicates that latent interdependencies of social vulnerability may be identified, grouped into indicators, and a way for validating these indicators can be found. The validation demonstrates that the selected factors are valid in identifying societal vulnerability as a critical component in the context of river-floods; second, composite indicators can be constructed without having a direct relationship to hazard characteristics.

4.2 Social Vulnerability analysis of Manikchak Diara using Social Vulnerability Index

More than 15 mouzas in Manikchak diara have been impacted by channel shifting and river bank erosion. A big number of individuals from rural backward areas are facing socio-economic vulnerability as a result of severe bank erosion in the Diara region of Malda district. People's livelihoods and social security are eroding day by day. They are being pushed to migrate from their own countries to other countries in search of work and a way of life. Hjaubona, Rezakpur, Kamaluddinpur, Daridiar Jhaubona, Dharampur, Manikchak, and Gopalpur mouzas in Manikchak are characterised by high socioeconomic vulnerability, with frequent bank erosion and flooding (Table 6 and Fig 4). Shukurullapur, Darijayrampur, Samastipur, Dakshin Chandipur, Paschim Narayanpur, and Ranigani have low bank erosion susceptibility because the population have the ability to cope with the problem. Hamidpur, Shobhapur, Paranpur, Narayanpur, Gobindapur, and Rostampur, as well as a remote area of the Manikchack diara with JL. Nos. 23, 24, 25, 27, 28, 29, 30, 31, 35, 36, and 81, are land safe mouzas.

Table 6: Mouza Level status of SVI in Manikchak Diara

JL. No.	Name of the Mouzas	SVI
3, 4, 5, 6, 83, 84, 90	Jhaubona, Rezakpur, Kamaluddinpur, Daridiar Jhaubona, Dharampur, Manikchak, Gopalpur	High
1, 2, 8, 10, 11, 13, 20, 04, 88	Gadai, Kesharpur, Jot Kasturi, Birodhi, Shobhanathpur, Masaha, Duani Tafir, Chandipur, Mirpur	Moderate
9, 12, 16, 19, 21, 85	Shukurullapur, Darijayrampur, Samastipur, Dakshin Chandipur, Paschim Narayanpur, Ranigani	Low
15, 17, 18, 22, 86, 87	Hamidpur, Shobhapur, Paranpur, Narayanpur, Gobindapur, Rostampur	Land safe zone
23, 24, 25, 27, 28, 29, 30, 31, 35, 36, 81	All these mozas are located far distance from the river bank	Not affected by bank erosion

Certain validation conditions and restrictions must be emphasized. Despite the fact that the survey's research area is very extensive, covering the Manikchak Diara Block, generalizing the results to entire river channels is problematic. More case studies are needed to cover all of Malda's regions. Although the questionnaire covers important data categories, it was not created with the objective of validating a social vulnerability indicator or the findings of this study. As a result, not all variables can be tested for validity.

The dependent variables were chosen based on the idea that having to leave one's house or seek emergency refuge has a significant impact. Although this type of measure has been utilized in the literature to detect social vulnerability (Chakraborty et al., 2005) [27], the extent to which it exposes social vulnerability in Malda area has not been thoroughly investigated. It is, nevertheless, a good opportunity for the exploratory nature of this pilot project, given the lack of analogous initiatives. To extract coping problems of an indirect economic, administrative, and perception nature, regulation satisfaction was used. The author's assumptions and decisions influence the selection of variables, the exclusion of sub-variables, and the establishment of thresholds.

In compared to all interviewees affected by the flood scenario, the results of the analyses suggest that home owners were more likely to leave their home for a length of time but less likely to seek emergency shelter. This might be viewed as home owners being more exposed than the norm, but having more financial and social resources on hand in the event of an evacuation. It's possible that home owners have more financial resources to plan for hazards privately, but this relationship isn't evident from the data presented in this chapter, therefore such an assumption should be avoided. Because urban residents had to leave their homes more frequently than rural residents, it's possible that population density plays a factor. Many senior residents were forced to seek emergency refuge due to a lack of financial or social resources to find alternative shelter or lodging with relatives, friends, or motels. Damage regulation is more unsatisfied by persons with a low educational level and those who are unemployed. This may be consistent with findings from other social vulnerability studies indicating persons with low qualifications have limited access to damage compensation, which could be due to a variety of factors. As with any interpretations in this case, one must exercise extreme caution, because the underlying reasons for each individual answer to the interview are not documented. As a result, while this semi-quantitative technique allows for the identification of social vulnerability patterns, it does not seek causal explanations.

5 Discussion

The number of rooms, home ownership, and degree of rurality are capable of explaining the distribution of individuals who had to leave their home and those who did not in the regression with the dependent variable leave home. For each independent variable, the likelihood of the dependent variable (leave home, answer "yes") can grow or decrease.

The probability relationship between the variable rooms is inverse. The lower the proportion of those in the group who had to leave their home, the higher the number of rooms. In other words, persons who lived in smaller flats had to leave their homes more frequently. The lower the proportion of those in the group who had to leave their home, the higher the number of rooms.

The larger the proportion of homeowners to tenants, the more likely it was that these families were forced to flee their homes owing to the flood. People who live near rivers are less harmed than those who live far away.

Age (from 16–95) and home ownership are likely to explain the distribution of those who had to seek emergency shelter against those who did not in the regression with the dependent variable emergency shelter. It was common for people of a certain age to seek emergency shelter.

The larger the proportion of homeowners to tenants, the less likely these households was to be forced to seek emergency refuge as a result of the flood. This violates the dependent variable leave home's ownership prediction direction.

Elementary school and unemployment explain the distribution of satisfaction with damage regulation in the regression with the dependent variable regulation. Damage regulation dissatisfies people with a low educational background (elementary school). Unemployed folks are subjected to the same observation.

Low SVI mouzas are characterized by river-flood resistance. Existing capacities for river-flood mitigation, such as financial capacities for private preparedness measures and flood recovery by high-income sources, are examples of these strengths.

Counties with a high SVI have a lot of flaws when it comes to river flooding. Lack of capacities, high levels of vulnerability, and signs of exposure potential are among these flaws.

6 Conclusion

The social vulnerability index (SVI) (Fig. 3) is made up of three variables (fragility, socioeconomic conditions, and region) that were produced from census data factor analysis and validated by a second data set. While most sub national quantitative vulnerability assessments end with the generation of vulnerability indicators, this work goes on to show how to validate vulnerability indicators further. In the first step, impact analysis is used to assess the vulnerability factors discovered by factor analysis in a real-life flood event. The logistic regression demonstrates that some demographic vulnerability variables do distinguish between those who are afflicted and those who are not. The factor analysis is rerun with the subset of variables that could be validated. The approach's validity is further supported by the emergence of the same variables. Region, social conditions, and fragility are the three key criteria identified and used as markers of social vulnerability. The SVI identifies German counties with a potential for high or low social vulnerability to flooding. Because social vulnerability is considered independent of individual river flood hazard, this index does not include hazard data. This index provides the foundation for a comprehensive risk assessment that includes both hazard and vulnerability studies.

7. References

1. Blaikie P, T Cannon, I Davis, B Wisner. At Risk: Natural hazards, People's vulnerability, and disasters. London, Routledge, 1994.
2. Turner BL II, Roger E Kasperson, Pamela A Matson, James J McCarthy, Robert W Corell, Lindsey Christensen, *et al*. A framework for vulnerability analysis in sustainability science. PNAS (Proceedings of the National Academy of Sciences of the United States of America). 2003; 100(14):8074-8079.
3. Birkmann J. Measuring vulnerability to natural hazards: towards disaster resilient societies. United Nations Publications, New York, 2006.
4. Fothergill A, Peek LA. Poverty and disasters in the United States: a review of recent sociological findings. Nat Hazards. 2004; 32(1):89-110.
5. Long AP. Poverty is the new prostitution: race, poverty, and public housing in post-Katrina New Orleans. J Am Hist. 2007; 94:795-803.
6. Fothergill A, Maestras E, Darlington J. Race, ethnicity, and disasters in the United States: a review of the literature. Disasters. 1999; 23(2):156-173.
7. Peacock WG, Morrow BH, Gladwin H. Hurricane Andrew and the reshaping of Miami: ethnicity, gender, and the socio-political ecology of disasters. Florida International University, International Hurricane Center, Miami, FL, 2000.
8. Enarson E, Morrow BH. The gendered terrain of disaster: through women's eyes. Praeger, Westport, 1998.
9. Enarson E, Fothergill A, Peek L. Gender and disaster: Foundations and directions. In: Quarantelli EL, Dynes R, Rodriguez H (eds) Handbook of Disaster Research. Springer, New York, 2006, 130-146.
10. Bolin R, Klenow D. Response of the elderly to disaster: an age-stratified analysis. Int J Aging Hum Dev. 1983; 16:283-296
11. Ngo EB. When disasters and age collide: reviewing vulnerability of the elderly. Nat Hazards Rev. 2001; 2(2):80-89.
12. Anderson WA. Bringing children into focus on the social science disaster research agenda. Int J Mass Emerg Disasters. 2005; 23(3):159-175.
13. Phillips BD, Hewett PL. Home alone: disasters, mass emergencies and children in self-care. J Emerg Manag. 2005; 3(2):31-35.
14. Kar N. Psychological impact of disasters on children: review of assessment and interventions. World J Pediatr. 2009; 5(1):5-11.
15. Smith S, Tremethick MJ, Johnson P, Gorski J. Disaster planning and response: considering the needs of the frail elderly. Int J Emerg Manag. 2009; 6(1):1-13.
16. Adger WN, Brooks N, Bentham G, Agnew M, Eriksen S. New indicators of vulnerability and adaptive capacity. Tyndall Centre for Climate Change Research, Norwich,

2004, 7.
17. Barnett J, Lambert S, Fry I. The hazards of indicators: insights from the environmental vulnerability index. Ann As Am Geogr. 2008; 98(1):102-119.
18. Hooke JM. Magnitude and distribution of rates of river bank erosion. Earth Surface Processes and Landforms. 1980; 5:143-157. https://onlinelibrary.wiley.com/doi/abs/10.1002/esp.3760050205.
19. Amiri-Tokaldany E, Darby SE, Tsswell P. Bank stability analysis for predicting reach scale land loss and sediment yield. Journal of the American Water Resources Association. 2003; 39(4):897-909. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1752-1688.2003.tb04414.x
20. Backhaus K, Erichson B, Plinke W, Weiber R. Multivariate Analysemethoden, 11th edn., Springer, Berlin, 2006.
21. Banerjee SN, Chakroborty P. Some observations on recent trends of shifting of the Ganga between Rajmahal and Ahiron. J Geol Soc India. 1983; 24:318-321.
22. Banerjee M. A report on the impact of Farakka Barrage on the human fabric, On behalf of South Asian network on dams, rivers and people (SANDRP), 1999. http://www.sandrp.in/dams/impact_frka_wcd.pdf. Accessed on 12.2.2008.
23. Barker R, Dixon L, Hooke J. Use of terrestrial photogrammetry for monitoring and measuring bank erosion. Earth Surface Processes and Landforms. 1997; 22:1217-1227. https://doi.org/10.1002/(SICI)1096-9837(199724)22:13%3C1217::AID-ESP819%3E3.0.CO;2-U.
24. Bernard HR. Research Methods in Anthropology, Qualitative and Quantitative Approaches, 4th edn., Altamira Press, Oxford, 2006.
25. B¨uhner M. Einf¨uhrung in die Test- und Fragebogenkonstruktion, 2nd edn. Pearson Studium, M¨unchen, 2006.
26. Bull LJ. Magnitude and variation in the contribution of bank erosion to the suspended sediment load of the River Severn, UK. Earth Surface Processes and Landforms. 1997; 22:1109-1123. DOI: 10.1002/(SICI)1096-9837(199712)22:12<1109::AID-ESP810>3.0.CO;2-O.
27. Chakraborty J, Tobin GA, Montz BE. Population Evacuation: Assessing Spatial Variability in Geophysical Risk and Social Vulnerability to Natural Hazards, Natural Hazards Review. 2005; 23-33.
28. Cutter SL, Boruff BJ, Shirley WL. Social Vulnerability to Environmental Hazards, Soc. Sci. Quart. 2003; 84:242-261.
29. Darby SE, Thorne CR. Numerical simulation of widening and bed deformation of straight sand-bed rivers. II: Model evaluation. ASCE Journal of Hydraulic Engineering. 1996; 122(4):194-202. https://www.sciencebase.gov/catalog/item/50538fabe4b097cd4fce4052.
30. Eaton BC, Millar RG. Optimal alluvial channel width under a bank stability constraint. Geomorphology. 2004; 62:35-45. DOI: 10.1016/j.geomorph.2004.02.003.
31. Fromm S. Bin¨are logistische Regressionsanalyse, Eine Einf¨uhrung f¨ur Sozialwissenschaftler mit SPSS f¨ur Windows, Bamberg, 2005.
32. Gall M. Indices of social vulnerability to natural hazards: A comparative evaluation, PhD dissertation, Department of Geography, University of South Carolina, 2007.
33. Gupta SD, Nandi DR. Seismicity and Tectonics of Meghalaya Plateau, North India, vii symposium on Earthquake Engineering, University of Roorkee, 19-24.
34. Conference: VII Symposium on Earthquake Engineering, University of Roorkee, At Roorkee, India, Volume: I.
35. Hooke JM. Magnitude and distribution of rates of river bank erosion. Earth Surface Processes and Landforms. 1980; 5:143-157. https://onlinelibrary.wiley.com/doi/abs/10.1002/esp.3760050205.
36. Iqbal S. Flood and erosion induced population displacements: a socio-economic case study in the Gangetic riverine tract at Malda district, West Bengal, India. J Hum Ecol. 2010; 30(3):201-211.
37. King D. Uses and Limitations of Socioeconomic Indicators of Community Vulnerability to Natural Hazards: Data and Disasters in Northern Australia, Natural Hazards. 2001; 24:147-156.
38. Millar RG, Quick MC. Effect of bank stability on geometry of gravel rivers', Journal of Hydraulic Engineering. 1993; 119:1343-1363. https://scite.ai/reports/effect-of-bank-stability-on-8KypK9.
39. Millar RG. Influence of bank vegetation on alluvial channel patterns. Water Resources Research. 2000; 36(4):1109-1118. https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/1999WR900346.
40. Moore DS, McCabe GP. Introduction to the practice of statistics, Freeman, New York, 2006.
41. Mukhopadhyay S. River bank erosion and land degradation, a study of river Ganga, land degradation and desertification. Rawat Publication, New Delhi, 2003, 366-373.
42. Nardo M, Saisana M, Saltelli A, Tarantola S, Hoffman A, Giovannini E. Handbook on constructing composite indicators: methodology and user guide, 2005.
43. O'Brien K, Leichenko R, Kelkar U, Venema H, Aandahl G, Tompkins H, *et al*. Mapping vulnerability to multiple stressors: climate change and globalization in India, Global Environ. Chang. 2004; 14:303-313.
44. Parua PK. Fluvial geomorphology of the river Ganga around Farakka. J Inst Eng. 2002; 82:193-196.
45. Parua PK. The Ganga water use in the Indian subcontinent. Water Sci Technol Libr. 2009; 64:391. doi:10.1007/978-90-481-3103-7
46. Rogers P, Lydon P, Seekter D. Eastern Water Studies; Strategies to manage flood and drought in the Ganga-Brahmaputra Basin, 1989, 5-75. http://lib.icimod.org/record/3476.
47. Rossi Romanelli L, Rinaldi M, Darby SE, Luppi L, Nardi L. Monitoring and modeling river bank processes: a new methodological approach. In: Greco M., Carravetta A. & Della, 2004.
48. Morte R. (Eds), Proceedings River Flow, Taylor & Francis Group, London, 2004, 993-998.
49. Rudra K. Problems of river bank erosion along the Ganga in Murshidabad district of West Bengal. Indian J Geogr Environ. 1996a; 1:25-32.
50. Rudra K. The Farakka barrage- an interruption to Fluvial regime. Indian J Landsc Syst Ecol Stud. 1996b; 19(2):105-110.

51. Rudra K.The encroaching Ganga and social conflicts: the case of West Bengal. India. Independent Broadcasting Associates, Littleton, 2004, 40.
52. http://www.ibaradio.org/India/ganga/extra/resource/Rudra.pdf. Accessed on 11 Oct 2010.
53. Sanyal N. Effect of Embankment of River Kosi Proc. Int. Workshop on 'Alluvial Rivers' org. by Univ. of Roorkee, 1980, 18-20.
54. Schneiderbauer S. Risk and Vulnerability to Natural Disasters-from Broad View to Focused Perspective, Doctoral thesis, FU Berlin, 2007.
55. Sesoren A. Geological interpretation of landsat imagery of the Bangladesh Ganges Delta, ITC Journal, 1984 29-232.
56. Showkat I. Flood and erosion induced population displacements: A socio-economic case study in the Gangetic riverine tract at Malda District, West Bengal, India. J Hum Ecol. 2010; 30(3):201-211. Kamla-Raj
57. Thakur *et al*. River bank erosion hazard study of river Ganga, upstream of Farakka barrage using remote sensing and GIS, Nat Hazards, 2011. DOI 10.1007/s11069-011-9944-z.
58. Wu SY, Yarnal B, Fisher A. Vulnerability of coastal communities to sea-level rise: a case study of Cape May County, New Jersey, USA, Climate Res. 2002; 22:255-270.