



Assessing climate-induced agricultural vulnerable coastal communities of Bangladesh using machine learning techniques

Md. Jakariya^{a,*}, Md. Sajadul Alam^a, Md. Abir Rahman^a, Silvia Ahmed^b, M.M. Lutfel Elahi^b, Abu Mohammad Shabbir Khan^b, Saman Saad^a, H.M. Tamim^b, Tauseef Ishtiaq^b, Sheikh Mohammad Sayem^c, Mirza Shawkat Ali^d, Dilruba Akter^d

^a Department of Environmental Science and Management, North South University, Bangladesh

^b Department of Electrical and Computer Engineering, North South University, Bangladesh

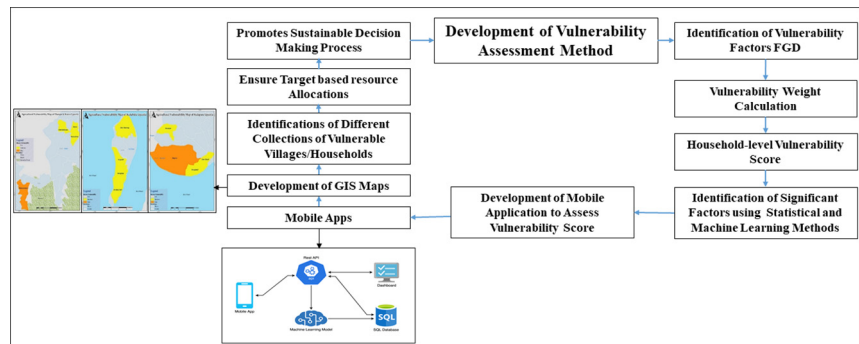
^c Department of Agricultural Statistics, Bangladesh Agricultural University, Bangladesh

^d Department of Environment, Government of the People's Republic of Bangladesh

HIGHLIGHTS

- The coastal areas are dynamic in nature and vulnerable to the changing climate.
- Agricultural practice is one of the main livelihood sources for coastal communities, which is highly dependent on nature.
- Use of machine learning techniques and mobile application can play an important role in assessing livelihood vulnerability.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 10 April 2020

Received in revised form 6 June 2020

Accepted 14 June 2020

Available online 16 June 2020

Editor: Fernando A.L. Pacheco

Keywords:

Coastal livelihood

Livelihood vulnerability index

Geographic information system

Regression analysis

Mobile application

ABSTRACT

The agricultural arena in the coastal regions of South-East Asian countries is experiencing the mounting pressures of the adverse effects of climate change. Controlling and predicting climatic factors are difficult and require expensive solutions. The study focuses on identifying issues other than climatic factors using the Livelihood Vulnerability Index (LVI) to measure agricultural vulnerability. Factors such as monthly savings of the farmers, income opportunities, damage to cultivable lands, and water availability had significant impacts on increasing community vulnerability with regards to agricultural practice. The study also identified the need for assessing vulnerability after certain intervals, specifically owing to the dynamic nature of the coastal region where the factors were found to vary among the different study areas. The development of a climate-resilient livelihood vulnerability assessment tool to detect the most significant factors to assess agricultural vulnerability was done using machine learning (ML) techniques. The ML techniques identified nine significant factors out of 21 based on the minimum level of standard deviation (0.03). A practical application of the outcome of the study was the development of a mobile application. Custom REST APIs (application programming interface) were developed on the backend to seamlessly sync the app to a server, thus ensuring the acquisition of future data without much effort and resources. The paper provides a methodology for a unique vulnerability assessment technique using a mobile application, which can be used for the planning and management of resources by different stakeholders in a sustainable way.

© 2020 Published by Elsevier B.V.

* Corresponding author.

E-mail address: md.jakariya@northsouth.edu (M. Jakariya).

1. Introduction

The Intergovernmental Panel on Climate Change (IPCC) highlights that the vast low-lying coastal region of Bangladesh is particularly vulnerable to risks from climate change (IPCC, 2014). Increase in temperatures, changing rainfall patterns, sea-level rise, and increase in frequency and intensity of extreme climatic events are negatively impacting agriculture, water resources, human health, and the ecosystem (Wheeler and Braun, 2013). Climate change has already had major impacts on the lives and livelihoods of the people in the coastal areas of Bangladesh (GoB, 2011, 2018).

The greater reliance of the people on the agricultural sector has made it more challenging today as a sustainable livelihood. Agriculture will become even more susceptible in the future due to changes in climate variables (IPCC, 2007; Islam et al., 2011). One predominant impact of climate change will be fluctuations in crop yield due to frequent changes in climatic variables (Mendelsohn and Dinar, 2003). Moreover, extreme climatic events, soil salinity in coastal areas, and incidence of pests and diseases due to the increased temperature and humidity may result in additional adverse effects on the agriculture sector (Rosenzweig et al., 2001). Despite technological development, climatic

factors are still fundamental dominators of agricultural productivity. Addressing the issues to find solutions to reduce agricultural vulnerability would require an integrated and comprehensive management plan with particular consideration for hazard vulnerability and the resilience of the coastal population to climate change (Sajjad and Chan, 2019). The authors recognize that long-term measures will be required to address the climatic factors of vulnerability and, therefore, the paper suggests methods to find the second most significant factors that can be addressed more easily and can reduce agricultural vulnerability in the short run. An effort was made to find the crop yield vulnerability of the farmers of the three coastal districts of Bangladesh by identifying the significant factors that have increased effects on the vulnerability score by Machine Learning models. This was done by first calculating the Vulnerability Livelihood Index of agriculture for each of the study areas.

Vulnerability Indicators help monitor and keep track of the changing vulnerability over time and space (Shah et al., 2013). The three components that characterize vulnerability include exposure, sensitivity, and adaptive capacity (IPCC, 2014). Local economic, social, and political factors can also significantly affect vulnerability components at the household level (Eriksen et al., 2005). Climate change is a widely-discussed

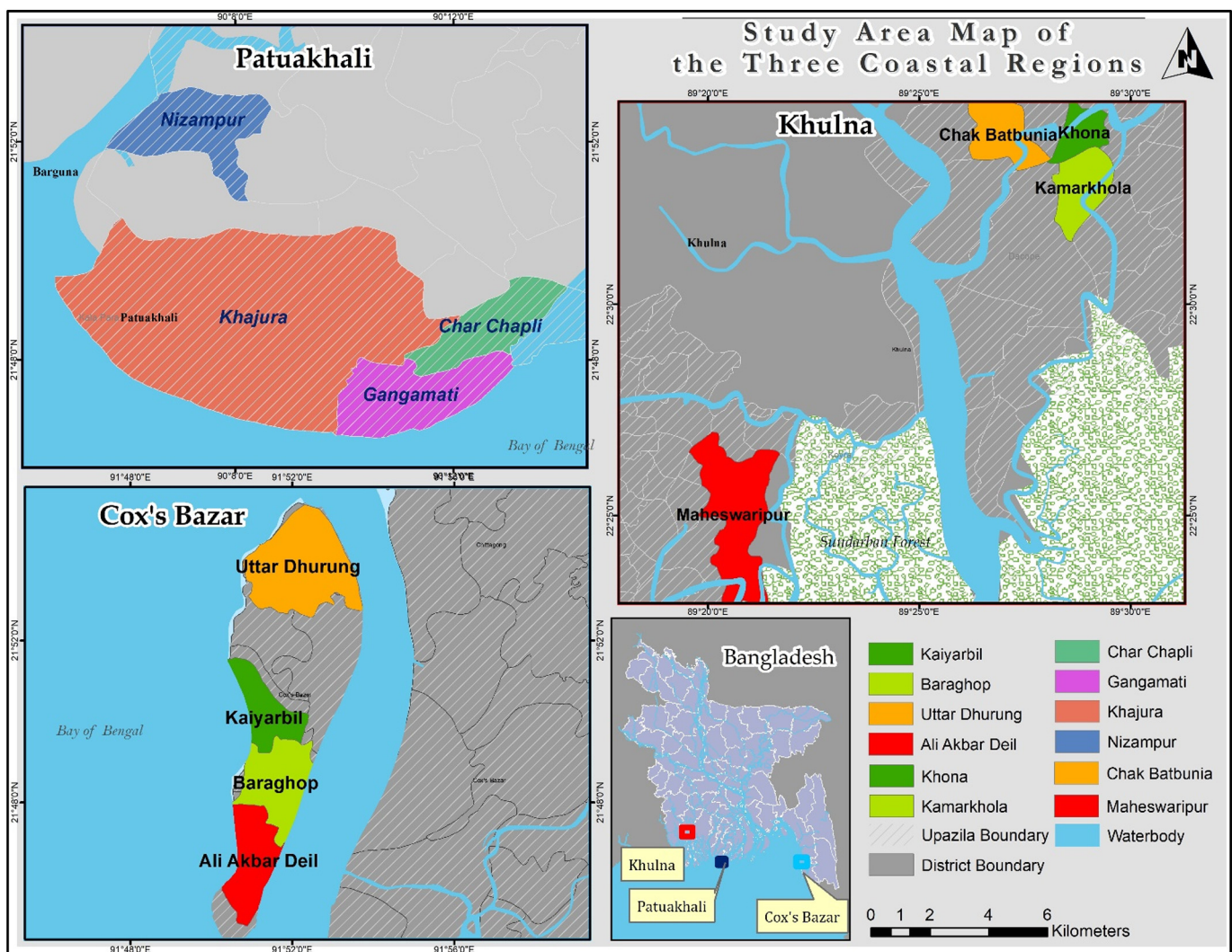


Fig. 1. Study area map.

topic and requires resources at a global scale but, for this paper, the authors focused on trying to find out adaptation strategies to cope with vulnerabilities related to rice production that would be specific for each region and could be managed locally with the help of mobile applications. Developing such technology-based solutions has become extremely important, especially considering the scarcity of global resources and, more importantly, the recent COVID-19 pandemic, which has emphasized the need to address the issue further. Bangladesh, as a country, with its unique geographical location, is prone to natural disasters and climatic effects. Therefore, at a time of such disasters, be it geographical or in the health-sector when mobility needs to be limited, it becomes next to impossible to maintain proper resource management. The work in this paper thus presents a complete system for the coastal areas of Bangladesh that work with data acquisition using mobile applications, data processing using machine learning techniques, and interpretation through web-based interfaces. In addition, the architecture that was designed for this system is quite generalized and can be adopted for work in other domains as well, with minimal modifications.

2. Methodology

2.1. Study area

Three coastal districts of Bangladesh, namely, Patuakhali, Kutubdia, and Khulna were selected for this research study. The maps of the study area are shown in Fig. 1.

2.2. Vulnerability assessment technique

The vulnerability assessment method used in the study was based on the vulnerability assessment method of GIZ Vulnerability Sourcebook (Fritzsche et al., 2014), which is built upon the IPCC Framework. It should be noted that the GIZ study only addressed specific methods of the vulnerability score calculation, but this study goes a step further and puts forth a practical application and usability of the score, giving policymakers a chance to apply the vulnerability information for functional purposes. Moreover, while GIZ only identified the main components of vulnerability (e.g., exposure, sensitivity, and potential

impacts), this research project identified the significant factors for all components of vulnerability, as shown in Fig. 2 through engaging in extensive discussions with local communities.

A weighted average method was used to calculate the Livelihood Vulnerability Index (LVI) (Eakin and Bojórquez-Tapia, 2008; Mathew et al., 2020). Each of the sub-components of vulnerability was measured using either the Likert scale or using primary or secondary data to facilitate broad comparability with past literature of the region. The variables were normalized on a numeric scale from 1 to 5. The calculation used for the component indices for exposure (E_N), sensitivity (S_N), and adaptive capacity (AC) are shown in Eqs. (1)–(4) below:

$$LVI = E_N + S_N - AC \quad (1)$$

$$E_N = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (2)$$

$$S_N = W \frac{\sum_{i=1}^N S_i}{N} \quad (3)$$

$$AC = W \frac{\sum_{i=1}^N AC_i}{N} \quad (4)$$

where, W = Weight and N = total number of factors.

The categories assigned for the vulnerability score is shown in Table 1. The regions were assigned into a category of low, medium, and high vulnerability based on the vulnerability score along with consultations with experts and local villagers (Schiffman and Kanuk, 2004).

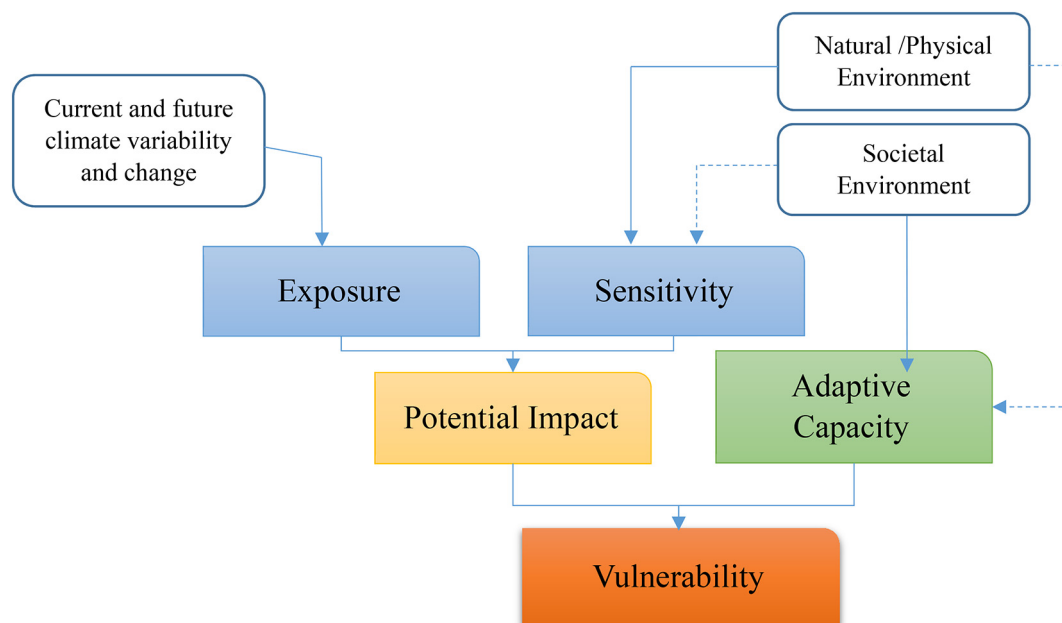


Fig. 2. Components of vulnerability.

2.3. Questionnaire and sampling design

Random sampling methods were used to select the study areas and study population (Bernard, 2002). A total of 930 households were selected to collect preliminary data in order to get a general overview of the study area. Broadly, socio-economic, climatic, water and sanitation, and disaster management related issues were covered in the questionnaire. A separate set of 297 samples were chosen from the preliminary selection of households to assess agricultural production and related vulnerability issues. The average vulnerability index and the spatial distribution of the houses according to vulnerability were prepared using the Geographic Information System (GIS) and Machine Learning techniques.

2.4. Development of a mobile application for vulnerability assessment

An online system comprised of a web and an Android application was developed for quick assessment of vulnerability and data-driven decision-making. The architecture of this system is depicted in Fig. 3. The mobile application contains a questionnaire to which coastal people can respond and the answers they provide are sent to a central web server through a REST API. The server then uses a machine learning model to get the predicted vulnerability of each respondent, stores the result in a database and sends the data to a web dashboard. The web dashboard summarizes results for all respondents and also provides individual level vulnerability scores. Administrators can use information from this dashboard to determine how resources can be optimally allocated to provide personalized help to each vulnerable individual.

3. Results & discussion

3.1. Identifying vulnerability factors

The factors related to three different variables of vulnerability, e.g., exposure, sensitivity, and adaptive capacity, were identified through focus group discussions (FGD) with the local farmers in each village. Climate change-related exposures are global issues, whereas sensitivity and adaptive capacity are location-specific and can be addressed with local interventions. The results of the FGDs are presented in Fig. 4.

3.2. Vulnerability weight calculation

The vulnerability weights reflect the perception of farmers regarding the factors related to vulnerability in the study areas. Weights measured from the ranking exercise conducted with farmers are displayed in Table 2. Across the coastal region of Bangladesh, the climatic conditions were among the factors with the highest weights, which illustrate their importance to assess vulnerability levels. The climatic conditions include average rainfall, average humidity, and average temperature (weight = 0.3), which holds a weight of particular importance that could disrupt farming activities. During FGD sessions, farmers often mentioned price and market conditions as vital factors for sustaining livelihoods. The factors shown in Table 2 are location-specific and were collected through focus group discussions with the local farmers.

The parameters with the highest vulnerability weights belong to exposure: average rainfall (weight = 0.3), average humidity (weight = 0.3), and average temperature (weight = 0.3). The parameters of sensitivity with the highest vulnerability weights calculated were the loss of crops (weight = 0.2648) and lack of water availability (weight =

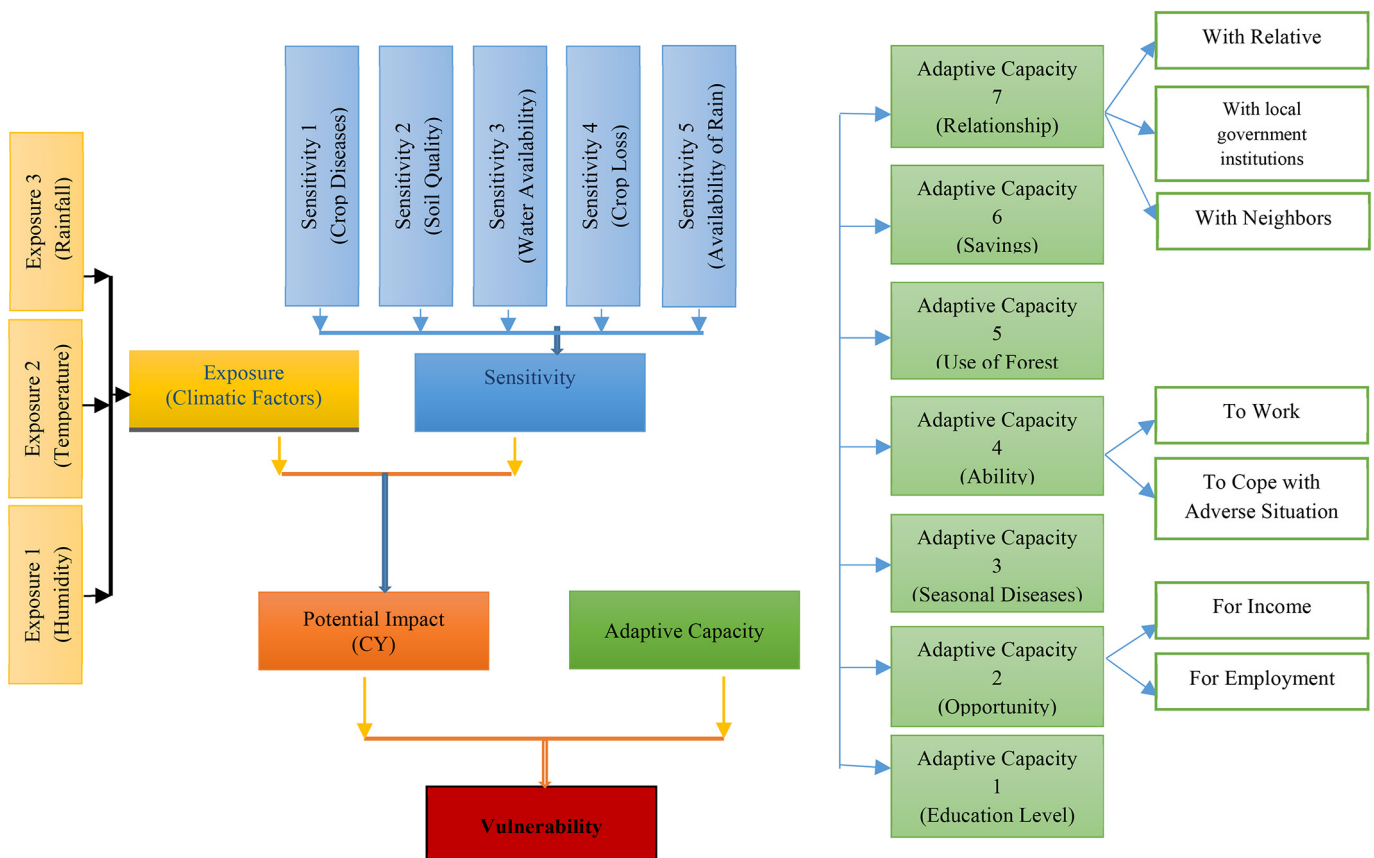


Fig. 3. Impact chain of the crop yield vulnerability livelihood index.

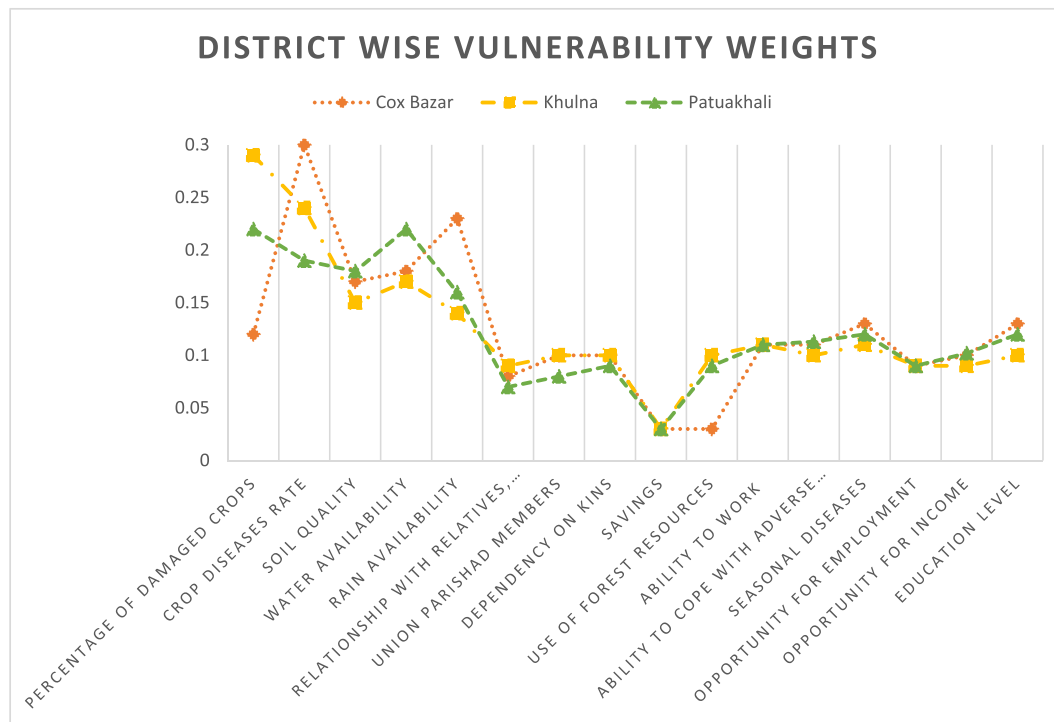


Fig. 4. Difference in district wise vulnerability weights of the factors of sensitivity and adaptive capacity.

Table 2
Variable weights of exposure, sensitivity and adaptive capacity.

Section	Required data	Data type	Remarks	Weight	Cox bazar	Khulna	Patuakhali	Standard deviation
Exposure	Average rainfall	Secondary	Number (mm)	0.30				
	Average humidity	Secondary	Humidity	0.30				
	Average temperature	Secondary	Number (OC)	0.30				
Sensitivity	Percentage of damaged crops	Secondary	Wind speed	0.26	0.12	0.29	0.22	0.07
	Crop diseases rate	Primary	Likert scale	0.17	0.30	0.25	0.19	0.04
	Soil quality	Primary	Likert scale	0.18	0.17	0.15	0.18	0.01
	Water availability	Primary	Likert scale	0.21	0.18	0.17	0.22	0.02
Adaptive Capacity	Rain availability	Primary	Likert scale	0.18	0.23	0.14	0.16	0.04
	Relationship with relatives, neighbors	Primary	Likert scale	0.09	0.08	0.09	0.07	0.01
	Union Parishad members	Primary	Likert scale	0.10	0.10	0.10	0.08	0.01
	Dependency on kins	Primary	Likert scale	0.08	0.10	0.10	0.09	0.00
	Savings	Primary	Likert scale	0.03	0.03	0.03	0.03	0.00
	Use of forest resources	Primary	Likert scale	0.08	0.03	0.10	0.09	0.03
	Ability to work	Primary	Likert Scale	0.11	0.11	0.11	0.11	0.00
	Ability to cope with adverse situation	Primary	Likert scale	0.11	0.11	0.10	0.11	0.01
	Seasonal diseases	Primary	Likert scale	0.12	0.13	0.11	0.12	0.01
	Opportunity for employment	Primary	Likert scale	0.10	0.09	0.09	0.09	0.00
	Opportunity for income	Primary	Likert scale	0.09	0.09	0.09	0.10	0.01
	Education level	Primary	Likert scale	0.12	0.13	0.10	0.12	0.01

0.2099). Moreover, the degradation of soil quality or higher crop disease rates can aggravate the situation. The parameters of adaptive capacity with the strongest influence on the vulnerability score calculation included seasonal diseases (weight = 0.1196) and the education level of farmers (weight = 0.1161).

Community response variations about different vulnerability factors in the three districts are shown in Fig. 5. Among the parameters for sensitivity, the parameter with the highest weight in Cox's Bazar and Patuakhali that affects agricultural productivity is rain availability (weight = 0.2287, weight = 0.2181 respectively), whereas, in Khulna and Patuakhali, the parameter that holds the highest weight in the vulnerability calculation is the percentage of damaged crops (weight = 0.2181).

The factors of adaptive capacity help to overcome the exposure and sensitivity factors when measuring the vulnerability of a community or household. The factors that hold the highest weight in the vulnerability calculation for Cox's Bazar were education level and seasonal crop diseases (weight = 0.134). The adaptive capacity factor with the highest weight in Khulna was also observed to be seasonal crop diseases (weight = 0.1081). In Patuakhali, the factor with the highest weight for adaptive capacity was the education level (weight = 0.1171).

3.3. Vulnerability score

Table 3 shows the state of crop yield vulnerability of the three coastal regions of Bangladesh, which is reflected in the vulnerability scores of different villages in the study area. Each village's vulnerability score

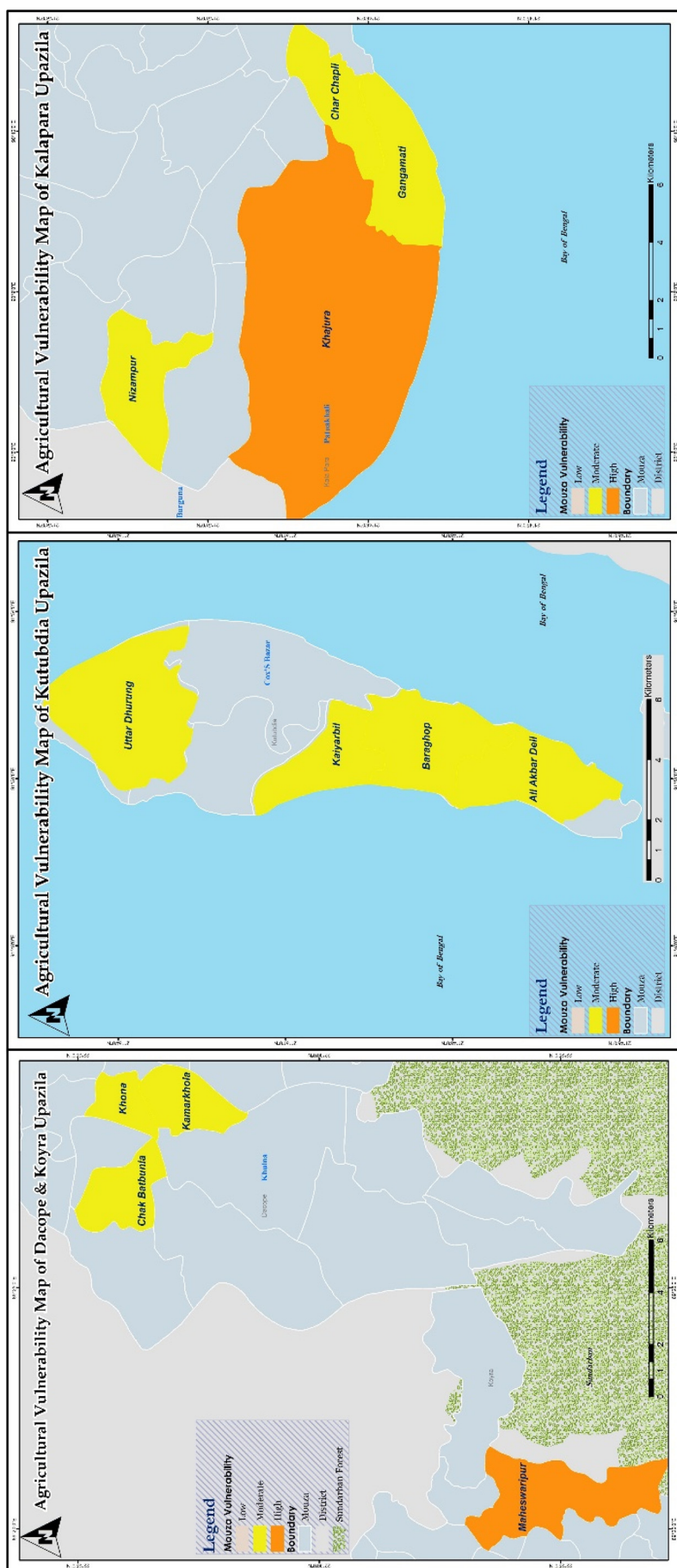


Fig. 5. Vulnerability maps of three coastal districts of Bangladesh.

Table 3
Crop yield vulnerability index.

District	Upazila	Village	E_N	S_N	AC_N	VI	Remarks
Khulna	Koyra Dacope	Maheswaripur	0.51	0.66	0.63	0.54	High
		Botbunia	0.54	0.54	0.60	0.48	Moderate
		Jaliakhali	0.54	0.54	0.58	0.49	Moderate
		Khona	0.54	0.59	0.63	0.50	Moderate
Cox's Bazar	Kutubdia	Ali Akbor Deil	0.49	0.53	0.61	0.42	Moderate
		Koiyarbil	0.49	0.58	0.64	0.44	Moderate
		Uttor Dhurong	0.49	0.54	0.64	0.40	Moderate
		BoroGhop	0.49	0.53	0.62	0.40	Moderate
Patuakhali	Kalapara	Gangamati	0.53	0.67	0.73	0.47	Moderate
		Nijampur	0.53	0.68	0.73	0.48	Moderate
		Khajura	0.53	0.65	0.67	0.51	High
		Char Chapli	0.52	0.63	0.67	0.48	Moderate

was derived from analyzing the individual score of each household. Every household's individual score was then examined and the cumulative score was achieved to obtain the vulnerability score. The maps in Fig. 6 of three coastal regions show geographical areas of vulnerability, which is the subject matter of the study. The spatial map shows the vulnerability level of the villagers according to the household survey. Similar to hazard maps, the vulnerability maps in Fig. 6 highlight the zones where farmers and farming land are most vulnerable to a variety of factors, which include social, physical, and economic aspects of rice production, as discussed above. (Bathrellos et al., 2011, 2013, 2017).

The overall average vulnerability level was found to be relatively moderate in all three study areas. This shows an overall similar vulnerability situation in the coastal region of the country. However, slight variations in terms of vulnerability were seen in Maheswaripur village in Khulna and Khajura village in Patuakhali, where relatively increased levels of vulnerability were observed.

It is apparent from the study that the vulnerability in the crop yield sector varies according to regional and temporal variations of natural disasters in the coastal areas of Bangladesh. It was observed that among the 297 farmers, about 46%, 34%, and 55% were vulnerable to the risk of humidity, temperature, and precipitation, respectively.

The study results were consistent with the national survey in showing that regions with lower income are more vulnerable to the impacts of climate change. Vulnerability to climate change varies between places, communities, and social classes (Adger, 2003; Smit and Wandel, 2006; CEGIS-FAO, 2006). There are important details in the findings which show the differential levels of vulnerability between

communities and the vulnerability of households within each community. The regions that show higher vulnerability have more households belonging in the lower-income group. In Maheswaripur village of Khulna district, about 60% of the surveyed population had a monthly income of less than USD 60. Similarly, in Khajura of Patuakhali district, 51.2% of the surveyed population was found to have an income level below USD 118.

3.4. Identification of significant factors

All the significant factors for vulnerability assessment were filtered using two different methods of statistical analysis and machine learning methods. Later, a comparative analysis of both the methods identified the best method to use for developing the mobile application. It was done with an understanding to develop a mobile application that was simple and convenient for the users in terms of handling fewer vulnerability factors for input. The following sections discuss both methods.

3.4.1. Statistical analysis

The multivariate logistic regression model was performed to screen out the non-significant factors of sensitivity and adaptive capacity (Tolles and Meurer, 2016; Brunner and Giannini, 2011). The goodness of fit of the model was high because the value of R^2 is 0.70 (Draper and Smith, 1998). On the basis of the Wald test, five variables for sensitivity and five variables for adaptive capacity showed significance (p -value < 0.01) out of a total of 19 variables, which were considered initially for the vulnerability score calculation (Fahrmeir et al., 2013; Ward and Ahlquist, 2018). The vulnerability score was measured by integrating the significant factors related to sensitivity and adaptive capacity without influencing the original score (Table 4).

3.4.2. Machine learning (ML) method

Though the statistical analysis gave a primary list of important variables, the correlation of the vulnerability scores generated only using these variables with those generated using the full list of variables under the GIZ framework was not high ($R^2 = 0.7$). So, using machine learning, we opted to find a better approach that would generate vulnerability scores closer to the original ones with fewer variables (Table 5).

The distribution of individual vulnerabilities, calculated as per the GIZ method, is plotted in Fig. 7. It can be seen that vulnerability follows a normal distribution and that there are no extreme vulnerability scores.

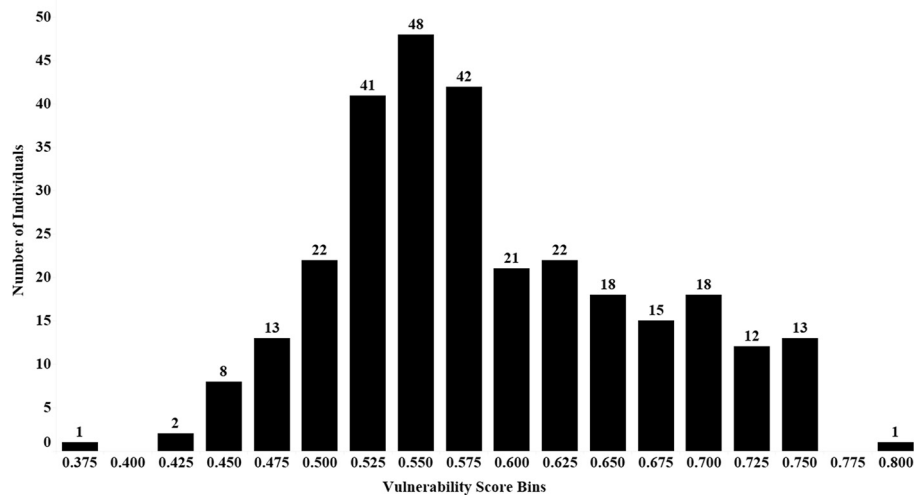


Figure 6: Vulnerability Distribution (2-column fitting image)

Fig. 6. Vulnerability distribution.

Table 4

Contribution of each independent variable to the model and its statistical significance.

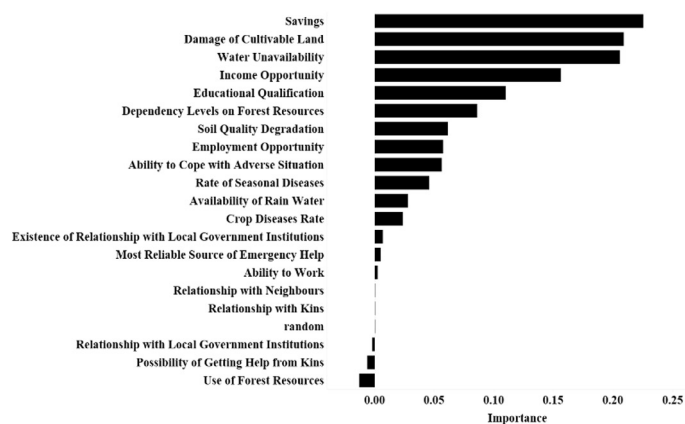
Factor	Variable	β	Wald	Sig.	Exp.(β)	R ²
Sensitivity	Crop diseases rate	10.564	15.022	0.000***	0.680	0.70
	Soil quality degradation	8.731	14.597	0.000***	0.056	
	Water unavailability	8.267	13.049	0.000***	0.087	
	Damage of cultivable land	9.116	16.182	0.000***	3.412	
	Availability of rainwater	6.429	14.386	0.000***	9.628	
Adaptive Capacity	Educational qualification	4.890	11.461	0.001***	0.008	
	Employment opportunity	3.303	9.757	0.002***	0.037	
	Income opportunity	3.481	9.160	0.002***	0.031	
	Rate of seasonal diseases	-5.877	12.850	0.000***	0.003	
	Ability to work	4.134	10.195	0.001***	0.016	

*** $P < .05$ at 95% Confidence interval.**Table 5**

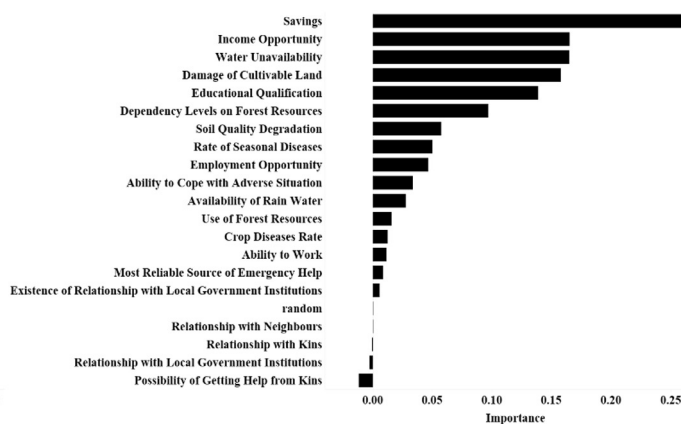
Performance of different machine learning models.

Machine learning model	R ² on training set	R ² on test set
Linear regression	0.93	0.90
Bayesian ridge regression	0.93	0.91
Random forest regression	0.97	0.83
XGB regression	1.00	0.83
Extremely randomized trees regression	1.00	0.81

3.4.2.1. Development of machine learning models. Before training the models, to ensure that there were no factors with high correlation, the Spearman's rank correlation coefficient between each pair of factors in the training set was calculated and no two factors with high correlation were found. Later, the vulnerability scores obtained by using the GIZ formula were taken as ground truths and five different regressors were tested to generate vulnerability scores as close as possible to the ones attained using the GIZ method. The models and their respective perfor-



(a): Feature Importance Obtained from Bayesian Ridge Regression



(b): Feature Importance Obtained from Linear Regression

Fig. 7. Permutation importance ranking of features.

It was assumed that the distribution is such because the agriculture-dependent coastal people are generally more or less vulnerable.

Of the 297 data points that were collected, three had null values for different factors, which might have been caused by erroneous data entries. Since the number of erroneous data points was very small, they were simply dropped and the remaining 294 data points were considered for the ML models. Moreover, in the dataset, there were only three distinct values for temperature, humidity, and precipitation. This occurred because each district was given a single value for each of these factors. As a result, there was minimal variance in the data for these factors and thus, was excluded from the ML models. Finally, to check whether any factor had little influence in predicting vulnerability, a column with random floating-point values taken from the half-open interval (0,1) was added, entitled "Random". The intention was to make an importance ranking of the factors where any factor ranked below "Random" could easily be disregarded. Thus, in the end, 294 data points having 20 actual factors and one random factor were considered. Randomly chosen, 80% of these data points were kept for training the models and the remaining 20% for testing the performance of the models.

mances are shown in Table 6.

Table 6

Combined ranking of factors.

Factor	Rank	Factor	Rank
Savings	1	Use of forest resources	12
Income opportunity	2	Availability of rain water	13
Damage of cultivable land	3	Most reliable source of emergency help	14
Water unavailability	4	Ability to work	15
Educational qualification	5	Existence of relationship with local government institutions	16
Dependency levels on forest resources	6	Random	17
Soil quality degradation	7	Relationship with neighbors	18
Employment opportunity	8	Relationship with kins	19
Rate of seasonal diseases	9	Relationship with local government institutions	20
Ability to cope with adverse situation	10	Possibility of getting help from kins	21
Crop diseases rate	11		

It can be seen that linear regression and Bayesian Ridge Regression performed well in predicting vulnerability scores while random forest regression, XGB regression, and extremely randomized trees regression overfitted the training data. The hyperparameters for the random forest, XGB, and extremely randomized trees regression (Breiman, 2001; Chen and Guestrin, 2016; Geurts et al., 2006) through Bayesian optimization were attempted to be tuned but were not successful in reducing the variance of these models without reducing their predictive capacity on the test set. This might be attributed to the fact that, in this case, these models are too complex for the small dataset being used. Linear regression and Bayesian Ridge Regression, on the other hand, did not require any hyperparameter tuning. As these models were functioning well, they were finally selected to generate the importance ranking of different factors.

3.4.2.2. Selection of the significant factors. In order to obtain the importance scores of different vulnerability factors, permutation importance was used, which works by measuring the R^2 score on the original set of factors for a model and then calculating the decrease in R^2 by randomly permutating the values of each of the factors one at a time (Altmann et al., 2010). In this way, the factors with a larger decrease in R^2 value are considered to be more important. Fig. 8(a) and (b) shows the factor importance rankings obtained from linear regression and Bayesian Ridge Regression, respectively.

Although linear regression and Bayesian Ridge Regression did not produce the same rankings, it was noticed that the two ranking schemes were similar in putting the same factors in higher or lower positions. To get a unified ranking, the ranks produced by the two regression models were summed up and sorted, with the factors in ascending order according to their sum of ranks. Later, the factors with smaller sums of ranks were considered more important than those with bigger sums of ranks. This unified ranking is shown in Table 6. The vulnerability factors with the lowest rank to the second-highest rank and so on were dropped one by one and trained new linear and Bayesian ridge regression models with continuously reducing sets of factors. Table 7 lists how the new models performed with the reduced sets of factors. It was noted that the factor “Random” was already dropped from our dataset before training the new models because it was no longer necessary.

It can be seen from Table 7 that up to 11 factors can be dropped and the vulnerability scores that have under 0.03 standard deviation from the original vulnerability scores while retaining a Pearson correlation coefficient of 0.93 can still be predicted. Since temperature, humidity, and precipitation were not included in the ML model like the original

vulnerability calculation, it can be stated that, in actuality, up to 14 factors can be reduced and reasonable predictive capacity of vulnerability scores can still be maintained by asking only 9 questions. The ML method demonstrated successfully in identifying significant factors for vulnerability score calculation than that of the statistical approach. The ML method also demonstrated that it could strategically identify the significant vulnerability factors with the highest rank for designing program intervention without considering all significant factors to reduce a specific community vulnerability in a resource constraint situation.

3.4.2.3. Use of a mobile application. A mobile application was developed after a successful reduction of non-significant factors, which later was used to assess the vulnerability scores. The design of the user interface (UI) and user experience (UX) were heavily considered while developing the mobile application so that people of any age with little educational background can use it. Farmers can log into this mobile application and answer the questions corresponding to the top 9 vulnerability factors, which were discussed in the previous section. To avoid any false input data, unrealistically large integer numbers of any input field can be filtered. Later, the responses of each individual will be sent to the central virtual server and vulnerability scores will be calculated for each individual household using the Bayesian Ridge Regression model.

To reduce agricultural vulnerability, it is important to consider the factors identified as being significant, such as soil and existence of groundwater, crop diseases, etc. along with the physical process of the area which provided issues related to agricultural vulnerability and this ultimately would help planners and policymakers to develop sustainable agricultural planning for the coastal communities of Bangladesh (Bathrellos et al., 2013).

Vulnerability assessment and planning are highly inter-dependent. In order to become more accurate in assessing farmers' vulnerability, it is important to consider environmental, social, economic, and other relevant factors such as culture, ethical issues, the proper understanding of the static relationships between man and nature, etc. while designing such an intervention (Ford et al., 2018; Bathrellos et al., 2011, 2013, 2017). A proper method of data collection, such as, if possible, an anthropogenic approach, could be applied for collecting such important information from the villagers. If the significant factors for vulnerability are identified properly, it would be easier for policymakers and planners to allocate scarce resources in a sustainable way. The mobile application, for example, provided a dynamic vulnerability score. The score, in reality, is not static and changes continuously. The application also allows

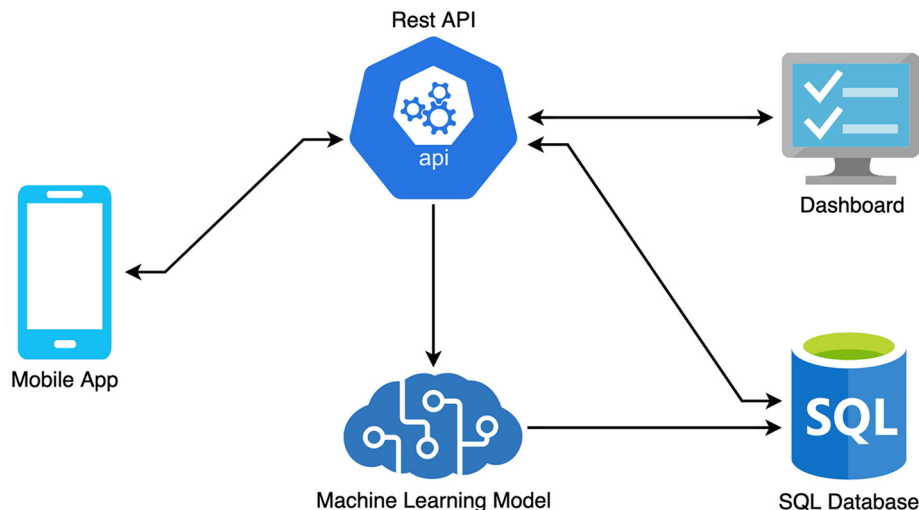


Fig. 8. System architecture of web-based and mobile application.

Table 7
Model performance after feature reduction.

Factors dropped	Actual factors present	Linear regression				Bayesian ridge regression			
		R ² train	R ² test	Std dev	Pearson corr coef	R ² train	R ² test	Std dev	Pearson corr coef
1	19	0.93	0.909	0.024	0.954	0.929	0.91	0.023	0.954
2	18	0.926	0.908	0.024	0.953	0.926	0.909	0.024	0.954
3	17	0.926	0.91	0.023	0.954	0.926	0.911	0.023	0.955
4	16	0.926	0.911	0.023	0.954	0.926	0.912	0.023	0.955
5	15	0.925	0.906	0.024	0.952	0.925	0.908	0.024	0.953
6	14	0.92	0.906	0.024	0.953	0.92	0.908	0.024	0.954
7	13	0.916	0.9	0.025	0.95	0.916	0.902	0.024	0.95
8	12	0.908	0.891	0.026	0.945	0.908	0.892	0.026	0.946
9	11	0.897	0.898	0.025	0.949	0.897	0.899	0.025	0.949
10	10	0.89	0.89	0.026	0.945	0.89	0.891	0.026	0.945
11	9	0.868	0.863	0.029	0.93	0.868	0.865	0.029	0.93
12	8	0.838	0.832	0.032	0.914	0.838	0.834	0.032	0.914
13	7	0.804	0.822	0.033	0.907	0.804	0.823	0.033	0.908
14	6	0.765	0.793	0.036	0.891	0.765	0.794	0.036	0.891
15	5	0.704	0.681	0.044	0.827	0.703	0.682	0.044	0.827
16	4	0.662	0.582	0.05	0.783	0.662	0.584	0.05	0.783
17	3	0.539	0.384	0.061	0.664	0.539	0.388	0.061	0.663
18	2	0.427	0.337	0.063	0.608	0.427	0.338	0.063	0.607
19	1	0.148	0.141	0.073	0.38	0.148	0.141	0.073	0.38

scores to be upgraded as frequently as required by relevant stakeholders to address a particular situation. As the application will be available to farmers for data input, the collected information will be more accurate while keeping expenses low as there will be no physical involvement for data collection and storage activities. The method of a mobile application-based quick vulnerability assessment technique can also be applied to assess other livelihood aspects simply by identifying significant factors responsible for vulnerability as the factors responsible for vulnerability are highly location and subject-specific.

4. Conclusion

Agricultural production in the coastal regions of Bangladesh is highly vulnerable to changes in weather conditions. The prevailing situation demands the development of a dynamic agricultural plan that considers the future consequences of climatic change and vulnerability to natural hazards. The study identified the most vulnerable agriculture-dependent households using a rapid and cost-effective method. The study focused on the development of a practical and community-friendly application to assess the vulnerability scores to aid local government institutions and other similar organizations with planning and management. Spatial maps were also prepared to show the locations of vulnerable households along with the extent to which they were vulnerable.

The study also provided a description for the application of the vulnerability scores, which can later be used to understand the vulnerability issues of certain livelihood options, such as agriculture, fishermen,

health, etc. Climate change-related exposures are global issues, whereas sensitivity and adaptive capacity are local issues and can be addressed with local interventions. Community response to the identified vulnerability factors of sensitivity and adaptive capacity varied significantly due to the diverse developmental profile and geographical characteristics of the study areas.

Statistical and machine learning methods were initially used to filter the most significant factors. The ML method was more successful in this aspect and was used to develop a mobile application that helped in vulnerability assessment by figuring out the factors which require immediate government intervention. The Climate Resilient Vulnerability Assessment Tool, which provides an authentic and faster process of identifying vulnerability, would be able to bring about a revolutionary change in resource distribution and, more importantly, the allocation of scarce resources in a sustainable way. The Vulnerability assessment process is usually highly technical, whereas the mobile application, which is based on a built-in system, provides a more user-friendly approach and can be used independently. The research findings provided an important starting point for directing future research into crop yield vulnerability to climate variability and change. It is expected that the output of the study can be used by policymakers and other stakeholders for better designing and targeting climate change adaptation policies and programs to ensure sustainability.

CRediT authorship contribution statement

Md. Jakariya: Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing, Supervision. **Md. Sajadul Alam:** Formal analysis, Writing - original draft, Investigation. **Md. Abir Rahman:** Formal analysis, Writing - original draft. **Silvia Ahmed:** Software, Investigation, Writing - review & editing. **M.M. Lutfel Elahi:** Software, Investigation, Writing - review & editing. **Abu Mohammad Shabbir Khan:** Writing - original draft, Software. **Saman Saad:** Writing - original draft, Writing - review & editing. **H.M. Tamim:** Software. **Taoseef Ishtiaq:** Software. **Sheikh Mohammad Sayem:** Formal analysis, Writing - review & editing. **Mirza Shawkat Ali:** Conceptualization. **Dilruba Akter:** Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research was supported and funded by the Climate Change Trust Fund (CTTF), the Government of the People's Republic of Bangladesh and the Department of Environment (DoE).

References

- Adger, W.N., 2003. Social capital, collective action, and adaptation to climate change. *Ecol. Econ.* 79, 384–404. https://doi.org/10.1007/978-3-531-92258-4_19.
- Altmann, A., Toloşi, L., Sander, O., Lengauer, T., 2010. Permutation importance: a corrected feature importance measure. *Bioinformatics* 26 (10), 1340–1347. <https://doi.org/10.1093/bioinformatics/btq134>.
- Bathrellos, G.D., Gaki-Papanastassiou, K., Skilodimou, H.D., Papanastassiou, D., Chousianitis, K.G., 2011. Potential suitability for urban planning and industry development by using natural hazard maps and geological - geomorphological parameters. *Environ. Earth Sci.* 66 (2), 537–548. <https://doi.org/10.1007/s12665-011-1263-x>.
- Bathrellos, G.D., Gaki-Papanastassiou, K., Skilodimou, H.D., Skianis, G.A., Chousianitis, K.G., 2013. Assessment of rural community and agricultural development using geomorphological - geological factors and GIS in the Trikala prefecture (Central Greece). *Stoch. Env. Res. Risk A.* 27 (2), 573–588. <https://doi.org/10.1007/s00477-012-0602-0>.
- Bathrellos, G.D., Skilodimou, H.D., Chousianitis, K., Youssef, A.M., Pradhan, B., 2017. Suitability estimation for urban development using multi-hazard assessment map. *Sci. Total Environ.* 575, 119–134. <https://doi.org/10.1016/j.scitotenv.2016.10.025>.
- Bernard, H.R., 2002. *Research Methods in Anthropology: Qualitative and Quantitative Approaches*. third ed. Altamira Press, Walnut Creek, CA.
- Breiman, L., 2001. Random forest. *Mach. Learn.* 45 (1), 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Brunner, H.J., Giannini, E.H., 2011. Trial design, measurement, and analysis of clinical investigations. In: Cassidy, J.T., Laxer, R.M., Petty, R.E., Lindsley, C.B. (Eds.), *Textbook of Pediatric Rheumatology*, sixth ed. Elsevier, pp. 127–156 <https://doi.org/10.1016/B978-1-4160-6581-4.10007-X>.
- CEGIS-FAO, 2006. *Study on Livelihood Systems Assessment, Vulnerable Groups Profiling, and Livelihood Adaptation to Climate Hazard and Long-term Climate Change in Drought Prone Areas of North-West Bangladesh*. Centre for Environmental Geographic Information Services (CEGIS) and Food and Agricultural Organization of the United Nations (FAO), Dhaka, Bangladesh.
- Chen, T., Guestrin, C., 2016. *XGBoost*. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining-KDD '16.
- Draper, N.R., Smith, H., 1998. *Applied Regression Analysis*. Wiley-Interscience <https://doi.org/10.1002/9781118625590>.
- Eakin, H., Bojórquez-Tapia, L.A., 2008. Insights into the composition of household vulnerability from multicriteria decision analysis. *Global Environ. Chang.* 18 (1), 112–127. <https://doi.org/10.1016/j.gloenvcha.2007.09.001>.
- Eriksen, S.H., Brown, K., Kelly, P.M., 2005. The dynamics of vulnerability: locating coping strategies in Kenya and Tanzania. *Geogr. J.* 171 (4), 287–305. <https://doi.org/10.1111/j.1475-4959.2005.00174.x>.
- Fahrmeir, L., Kneib, T., Lang, S., Marx, B., 2013. *Regression: Models, Methods, and Applications*. Springer, Berlin <https://doi.org/10.1007/978-3-642-34333-9>.
- Ford, J.D., Pearce, T., McDowell, G., et al., 2018. Vulnerability and its discontents: the past, present, and future of climate change vulnerability research. *Clim. Chang.* 151, 189–203. <https://doi.org/10.1007/s10584-018-2304-1>.
- Fritzsche, K., Schneiderbauer, S., Bubeck, P., Kienberger, S., Buth, M., Zebisch, M., Kahlenborn, W., 2014. The Vulnerability Sourcebook: Concept and Guidelines for Standardised Vulnerability Assessments. Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, Bonn and Eschborn Retrieved from. https://giz.de/ibt/var/app/wp342deP/1443/wp-content/uploads/filebase/va/vulnerability-guides-manuals-reports/Vulnerability_Sourcebook_-_Guidelines_for_Assessments_-_GIZ_2014.pdf.
- Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. *Mach. Learn.* 63 (1), 3–42.
- GoB, 2011. *Sixth Five Year Plan of Bangladesh*. General Economics Division, Planning Commission, Government of the People's Republic of Bangladesh, Dhaka.
- GoB, 2018. *Climate Change and Agriculture in Bangladesh: Information Brief*. Ministry of Environment and Forest, Government of the People's Republic of Bangladesh Retrieved from. <https://cmsdata.iucn.org/downloads/agriculture.pdf>. (Accessed 20 January 2019).
- IPCC, 2007. *Climate Change 2007: The Physical Science Basis*. Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK.
- IPCC, 2014. *Climate Change 2014: Fifth Assessment Synthesis Report*. Intergovernmental Panel on Climate Change, Geneva, Switzerland.
- Islam, M.B., Ali, M.Y., Amin, M., Zaman, S.M., 2011. Climate variations: farming systems and livelihoods in the high barind tract and coastal areas of Bangladesh. In: Lal, R., Sivakumar, M.V.K., Rahman, A.H.M.M., Islam, K.R. (Eds.), *Climate Change and Food Security in South Asia*. Springer Science+Business Media B.V.
- Mathew, M.J., Sautter, B., Ariffin, E.H., Menier, D., Ramkumar, M., Siddiqui, N.A., Delanoe, H., Del Estal, N., Traoré, K., Gensac, E., 2020. Total vulnerability of the littoral zone to climate change-driven natural hazards in north Brittany, France. *Sci. Total Environ.* 706, 135963. <https://doi.org/10.1016/j.scitotenv.2019.135963>.
- Mendelsohn, R., Dinar, A., 2003. Climate, water, and agriculture. *Land Econ.* 79 (3), 328–341. <https://doi.org/10.2307/3147020>.
- Rosenzweig, C., Iglesias, A., Yang, X.B., Epstein, P.R., Chivian, E., 2001. Climate change and extreme weather events: implications for food production, plant diseases, and pests. *Global Change and Human Health* 2 (2), 90–104. <https://doi.org/10.1023/A:1015086831467>.
- Sajjad, M., Chan, J.C.L., 2019. Risk assessment for the sustainability of coastal communities: a preliminary study. *Sci. Total Environ.* 671, 339–350. <https://doi.org/10.1016/j.scitotenv.2019.03.326>.
- Schiffman, L.G., Kanuk, L.L., 2004. *Consumer Behavior*. eighth ed. Prentice Hall, Englewood Cliffs, NJ.
- Shah, K.U., Dulal, H.B., Johnson, C.A., Baptiste, A.K., 2013. Understanding livelihood vulnerability to climate change: applying the livelihood vulnerability index in Trinidad and Tobago. *Geoforum* 47, 125–137. <https://doi.org/10.1016/j.geoforum.2013.04.004>.
- Smit, B., Wandel, J., 2006. Adaptation, adaptive capacity, and vulnerability. *Glob. Environ. Chang.* 16, 282–292. <https://doi.org/10.1016/j.gloenvcha.2006.03.008>.
- Tolles, J., Meurer, W.J., 2016. Logistic regression: relating patient characteristics to outcomes. *J. Am. Med. Assoc.* 316 (5), 533. <https://doi.org/10.1001/jama.2016.7653>.
- Ward, M.D., Ahlquist, J.S., 2018. *Maximum Likelihood for Social Science: Strategies for Analysis*. Cambridge University Press <https://doi.org/10.1017/9781316888544>.
- Wheeler, T., Braun, J.V., 2013. Climate change impacts on global food security. *Science* 341 (6145), 508–513. <https://doi.org/10.1126/science.1239402>.