Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Comparative Assessment of Padma River Bank Erosion Susceptibility Using Ahp, Fuzzy Ahp, Random Forest, Support Vector Machine, and XGBoost in a Gis-Based Framework

Harun Ar Rashid

24271005@student.duet.ac.bd

Dhaka University of Engineering & Technology

Md. Tofiquzzaman

River Research Institute, Faridpur, Bangladesh

Research Article

Keywords: Riverbank erosion susceptibility, Padma River, MCDM, Machine learning, GIS

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11793017/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Riverbank erosion in Bangladesh poses serious threat to thousands of people, loss of agricultural land, infrastructure sharing losses every year and one of the most destructive geomorphic processes of Bangladesh. This study fills this gap by designing and testing five susceptibility mapping methods fostering integration between Google Earth Engine (GEE) and ArcGIS namely Analytical Hierarchy Process (AHP), Fuzzy AHP, Selecting Random Forest (RF), Support Vector Machine (SVM) and eXtreme Gradient Boosting (XGBoost). Eleven geomorphic and hydroclimatic conditioning factors encompassed terrain, vegetation, hydraulic and climatic parameters were extracted from GEE. Four epochs of multi-temporal Landsat imagery were used to identify erosion and accretion polygons totaling 168.59 km^2 of erosion and 212.60 km^2 of accretion in the study reach. The 4807 sample points (with 3262 points from the erosion and 1545 points from the non-erosion samples) were split in a ratio of 70:30 to train and validate the machine learning (ML) models. XGBoost achieved superior performance with accuracy of 87.18%, F1-score of 0.909, ROC-AUC of 0.903, and Cohen's Kappa of 0.693, followed by RF (accuracy 85.86%, Kappa 0.659) and SVM (accuracy 80.67%, Kappa 0.561). The top three variables identified by RF were NDVI, flood frequency and water occurrence. The AHP consistency ratio of 0.027 ($0.10 < CR < 0.15$) indicated that the experts were reliable in assigning weights. Maps produced from each of the five methods showed that MCDM methods create wider distance buffer hazard maps, as does ML, suggesting complementary applications. The GEE–GIS–ML framework can be adopted elsewhere for other large alluvial. rivers in South and Southeast Asia.

1. INTRODUCTION

The Padma River, the major distributary of the Ganges River in Bangladesh, experiences some of the highest rates of lateral migration in the river system (500–1,000 m per year) of any in South Asia, and is one of the most destructive geomorphic processes in Bangladesh[1].

Susceptibility mapping methods fall into two broad categories: expert-driven MCDM approaches, such as the AHP [2] and Fuzzy AHP[3], and data-driven ML algorithms, including RF [4], SVM [5], and XGBoost [6]. MCDM methods provide theory-based frameworks and weights for the factors that are easily interpreted and transparent, and ML algorithms are able to capture the complex non-linear relationships between factors and quantify the accuracy of such factors. Although both paradigms are widely adopted, there is still limited comparative analysis of both approaches side by side for bank erosion susceptibility mapping of the Padma River. This study addresses that gap with four objectives: (i) quantifying the dynamics of erosion and accretion along the Padma River from 2010 to 2025; (ii) creating susceptibility maps with all five methods; (iii) validating and comparing the performance of the models; and (iv) identifying the most influential conditioning factors.

2. METHODOLOGY

2.1 Study Area

The study area includes the approximately 115 km reach of the Padma River in west-central Bangladesh (88.50°E–89.40°E, 24.30°N–23.45°N) in the districts of Rajbari, Faridpur, Kushtia and Pabna.

2.2 Erosion Mapping and Conditioning Factors

For this study, the multi-temporal Landsat imagery (Landsat 5/8/9, USGS Collection 2 Level-2) was processed through the GEE for four dry seasons (2010, 2015, 2020, 2025) with the condition of cloud cover < 10% and median compositing. Water body extraction (MNDWI = (Green − $SWIR_1$)/(Green + $SWIR_1$); threshold = 0.0) was performed and banklines were manually digitised for each epoch[7]. Consecutive banklines were intersected to create erosion and accretion polygons. Eleven conditioning factors were extracted from GEE: elevation and slope (SRTM 30 m) [8], curvature, Topographic Wetness Index {TWI = ln(A/tanβ)} [9], Stream Power Index (SPI = A × tanβ) [10], NDVI (Landsat 8/9), LULC (Dynamic World) [11], flood frequency (JRC Monthly Water History) [12], water occurrence (JRC Global Surface Water), distance from river, and rainfall (CHIRPS) [13].

Table 1
Summary of eleven erosion conditioning factors and data sources.

Factor	Source / GEE Dataset	Resolution	Category
Elevation, Slope, Curvature	USGS	30 m	Terrain morphology
TWI, SPI	Derived from SRTM in GEE	30 m	Terrain morphology
NDVI	LANDSAT/LC08/C02/T1_L2	30 m	Vegetation
LULC	GOOGLE/DYNAMICWORLD/V1	10 m	Land cover
Flood Frequency	JRC/GSW1_4/Monthly History	30 m	Hydraulic dynamics
Water Occurrence	JRC/GSW1_4/Global Surface Water	30 m	Hydraulic dynamics
Distance from River	Derived (fast Distance Transform)	30 m	Hydraulic dynamics
Rainfall	UCSB-CHG/CHIRPS/DAILY	~ 5.6 km	Climate

2.3 AHP and Fuzzy AHP

AHP was adopted based on the work by Saaty [2]. Using Saaty's 1–9 scale, a 9×9 pairwise comparison matrix was made with the 9 major conditioning factors. The factor weights (wi) were computed as average of each row in the normalized matrix. The Consistency Ratio (CR = CI/RI, where CI = (λmax − n)/(n − 1) and RI = 1.45 for n = 9) was calculated to ensure that the matrices are coherent (threshold CR < 0.10). Susceptibility index.

$$SI = \sum W_i \times R_i \ (i = 1, 2, \ldots \ldots \ldots ., 9) \ (1)$$

In the above formula, w_i represents the AHP weight, and R_i represents the reclassified raster value (scale 1–5) for factor i. In ArcGIS all nine factor rasters were reclassified to a 1–5 scale of erosion risk.

Chang's extent analysis method was adopted for implement fuzzy AHP [3]. A crisp Saaty value was converted into a triangular fuzzy number (TFN) l. All factor pairs were used to calculate the fuzzy synthetic extent S_i and the degree of possibility $V(S_i \geq S)$ of the factors, and then the minimum possibility values were normalized to obtain the final Fuzzy AHP weights.

2.4 Machine Learning Models

Sample points were stratified randomly within confirmed erosion and confirm stable bank polygons, to generate a total of 6,906 sample points (3,453 erosion, 3,453 non-erosion). The Extract Multi Values to Points tool in ArcGIS was used to extract the values of the conditioning factors at every point location. 4,807 complete points were left after the removal of points outside of the valid extent of the image, including 3,262 erosion and 1,545 non-erosion points. To compensate for this class ratio, the Approximate 2.1:1 ratio for the classes was corrected using balanced class weighting in the SVM Classifier. The dataset was the split into 70% training (n = 3,364) and 30% testing (n = 1,443) in Fig. 4 with stratified random sampling with fixed seed (42) in order to maintain the distribution of the classes in both sets.

3. RESULTS AND DISCUSSION

3.1 Erosion and Accretion

Table 2
Padma River erosion and accretion statistics (2010–2025)

Period	Erosion (km²)	Accretion (km²)	Net Change (km²)	Erosion Rate (km²/yr)
2010–2015	33.93	109.32	+ 75.39	6.79
2015–2020	64.64	50.74	−13.90	12.93
2020–2025	70.02	52.54	−17.48	14.00
Total 2010–2025	168.59	212.60	+ 44.01	11.24

3.2 AHP, Fuzzy AHP, SVM, XGBoost and RF Susceptibility Maps

The consistency (CR) was found to be less than 0.1 (very good) from the pairwise matrix. Distance from river received the highest AHP weight (0.2801; 28.01%), followed by flood frequency (0.1952; 19.52%), elevation and slope (0.1397 each; 13.97%), water occurrence (0.0880), NDVI (0.0558), rainfall (0.0384),

LULC (0.0326), and curvature (0.0205). The hydraulic factors had smaller curve values in LULC and curvature, whereas it had larger curve values in the most important hydraulic factors, which is normal

because the most important hydraulic factors have more expert judgments and thus become more influential in the fuzzy method. Both AHP and Fuzzy AHP maps have demonstrated a good agreement with the

observed erosion record in accordance with the zonal statistics calculation, in which 99.24% (188,354 pixels) of the erosion registered in Class 4 (High susceptibility class) corresponds to the values shown in

the maps. The three ML models were then used on the entire 11-factor raster stack to create the spatially continuous susceptibility surfaces.

The probabilities of erosion shown on the maps were adapted to the same five susceptibility levels of the MCDM maps, (Very Low to Very High), for ease of comparison with visual interpretation. The five susceptibility maps are shown in Fig. 3.

The spatial difference between the two paradigms can be easily derived from the dominant weighting mechanisms that inform them. Distance along the river is assigned weights of 28.01% and 29.00% respectively by AHP and Fuzzy AHP, and this produces very broad radial buffer patterns. The flood frequency, water occurrence, combined with NDVI (27.8% for RF), instead regulate both RF and XGBoost, which provide channel-confined susceptibility zones. They are both eleven-factor models and SVM does not include an intrinsic measure of feature importance, but again generates a spatial pattern much similar to that produced by RF and XGBoost.

3.3 Machine Learning Performance

The detailed performance of all three ML models is shown in Table 3.

Table 3
ML model performance comparison.

Model	Accuracy	Precision	Recall	F1-Score	Kappa
XGBoost	0.8718	0.8767	0.9438	0.9090	0.6931
Random Forest	0.8586	0.8638	0.9397	0.9002	0.6594
SVM	0.8067	0.8631	0.8498	0.8564	0.5606

3.4 Discussion

A comparison of the methods shows complementary strengths in both MCDM and the ML approaches. In terms of quantitative classification accuracy, XGBoost and RF performed better than MCDM methods, which are demonstrated in recent benchmark studies (15). AHP and Fuzzy AHP, on the other hand, demonstrate better spatial validation (99.2% zonal agreement) and offer susceptibility frameworks as

transparent and theory-grounded models, which can be used in the operational planning of water sectors. Compared to the sensitivity values (0.850 to 0.944), the always lower specificity values (0.688 to 0.720) are due to the fact that the cleaned dataset retains some imbalances in the residuals (2.1:1 erosion to non-erosion) thereby biasing classifiers toward the majority class. The top three factors for RF that have been found are NDVI (27.8%), flood frequency (16.3%) and the occurrence of water (13.3%) show in Fig. 6.

This is subsequently conflicting with the distance from river in AHP expert weights which ranked highest (28.01%). This difference is actually attributed to the fact that, as a consequence of multicollinearity, a flood frequency and a water occurrence provides information about the spatial level of their occurrence, with a higher degree of spatial specificity, than only the Euclidean distance is able to provide. RF synthesizes part of the variance contribution of the distance factor to that of a more detailed representation of the position. There was a direct observable difference in the weighting in the spatial encompassment of the bands of susceptibility incurred by the different models, in terms of the actual distribution of susceptibilities in the x-y plane, shown graphically in Fig. 3. AHP and Fuzzy AHP resulted in smooth radial bands of susceptibility extending several kilometres away from the channel; RF, SVM, and XGBoost concentrated a high susceptibility in the channel margin only.

An AUC going from 0.863 to 0.396 if the data was NOT scaled using StandardScaler makes this point seem very important when scaling is assumed in the RBF kernel $K(x_i, x') = \exp(-\gamma\|x_i - x'\|^2)$. Without normalization, the SPI factor ranged between 0.09 and 1.7×10^9, which totally eclipsed Euclidean distance calculations and rendered all other conditioning factors numerically irrelevant. They have operational implications: using MCDM maps for regional setback protection and uses maps in the construction of land use frameworks, and using ML maps for locations with a high spatial resolution in which structural interventions on a specific erosion hotspot can be considered.

4. Conclusion

The susceptibility from the Padma River banks was comprehensively concluded using AHP, Fuzzy AHP, RF, SVM and XGBoost. The findings are: (i) Erosion of the Padma River increased by 106%, from 33.93 km2 (2010–2015) to 70.02 km2 (2020–2025); (ii) The optimum classification model to achieve best ML performance in this study was XGBoost with an accuracy of 87.18% and Kappa of 0.693; (iii) The highest AUC was achieved by the RF model, which contains the highest percentage of spatial variance (70.0%); (iv) AHP has a good level of weight consistency (CR = 0.027) and the spatial validation was a success (99.2%); (v) The RF model has the highest level of spatial variance (70.0%) and (vi) The feature standardization is important for performing the SVM model. The reasons for this may derive from the different mechanisms of factor weighting that are used by MCDM and ML maps systematic versus channel-confined indicating that the maps are not mutually exclusive from an operation perspective. The framework of GEE–ArcGIS–ML integrated system developed here may be applied to other data-scarce braided meandering river systems found in South and Southeast Asia.

Declarations

ACKNOWLEDGEMENTS

The authors thank the Google Earth Engine platform and USGS for making their data freely available. The authors are thankful to the river research institute (RRI), Faridpur for institutional support.

References

1. Sarker MH, Huque I, Alam M, Koudstaal R (2003) Rivers, chars and char dwellers of Bangladesh. Int J River Basin Manag 1(1):61–80. 10.1080/15715124.2003.9635193
2. Saaty TL (1980) The Analytic Hierarchy Process. McGraw-Hill
3. Chang DY (1996) Applications of the extent analysis method on fuzzy AHP. Eur J Oper Res 95(3):649–655. 10.1016/0377-2217(95)00300-2
4. Breiman L (2001) Random Forests. Mach Learn 45(1):5–32. 10.1023/A:1010933404324
5. Vapnik V (1995) The Nature of Statistical Learning Theory. Springer
6. Chen T, Guestrin C, XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [Internet]. San Francisco California USA: ACM; 2016 [cited 2026 Aug 10]. pp. 785–94. Available from: https://dl.acm.org/doi/10.1145/2939672.2939785doi:10.1145/2939672.2939785
7. Xu H (2006) Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. Int J Remote Sens 27(14):3025–3033. 10.1080/01431160600589179
8. Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S et al (2007) The Shuttle Radar Topography Mission. Rev Geophys 45(2):2005RG000183. 10.1029/2005RG000183
9. Beven KJ, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrol Sci Bull 24(1):43–69. 10.1080/02626667909491834
10. Moore ID, Grayson RB, Ladson AR (1991) Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. Hydrol Process 5(1):3–30. 10.1002/hyp.3360050103
11. Brown CF, Brumby SP, Guzder-Williams B, Birch T, Hyde SB, Mazzariello J et al (2022) Dynamic World, Near real-time global 10 m land use land cover mapping. Sci Data 9(1):251. 10.1038/s41597-022-01307-4
12. Pekel JF, Cottam A, Gorelick N, Belward AS (2016) High-resolution mapping of global surface water and its long-term changes. Nature 540(7633):418–422. 10.1038/nature20584
13. Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S et al (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci Data 2(1):150066. 10.1038/sdata.2015.66

Figures

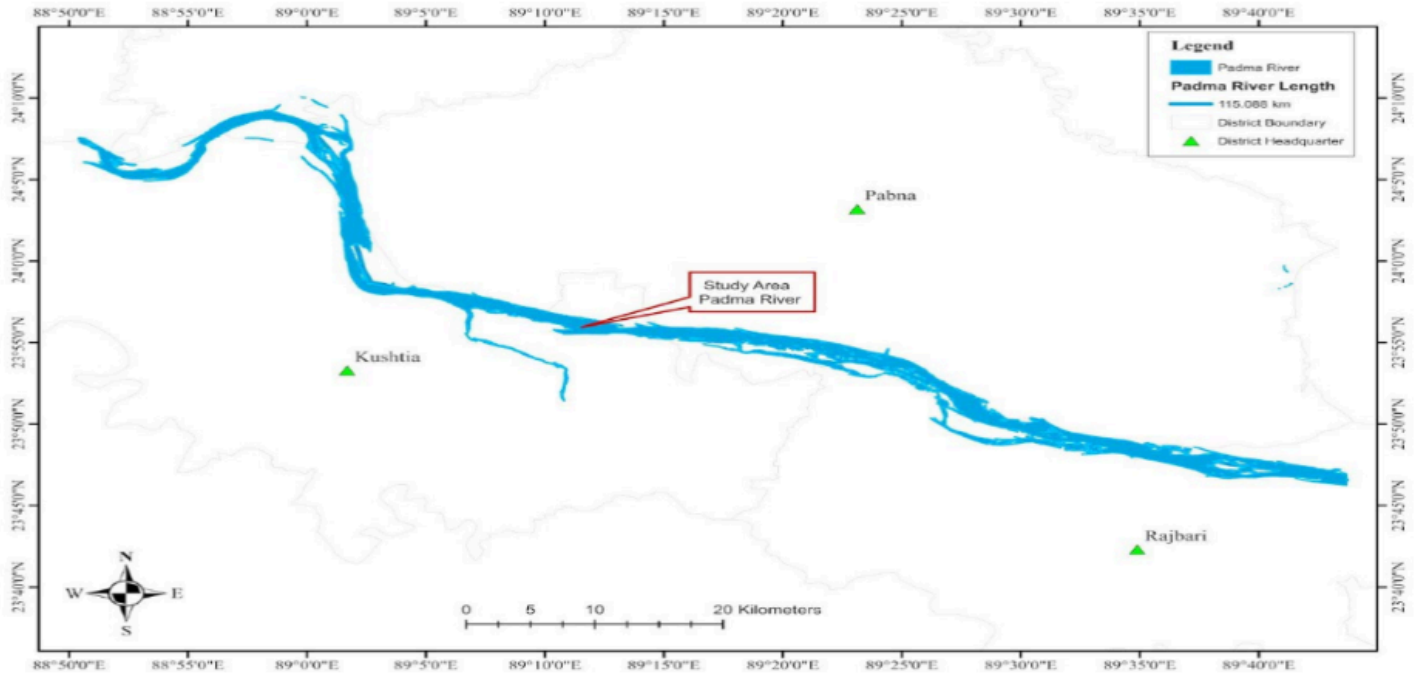


Figure 1

Study area

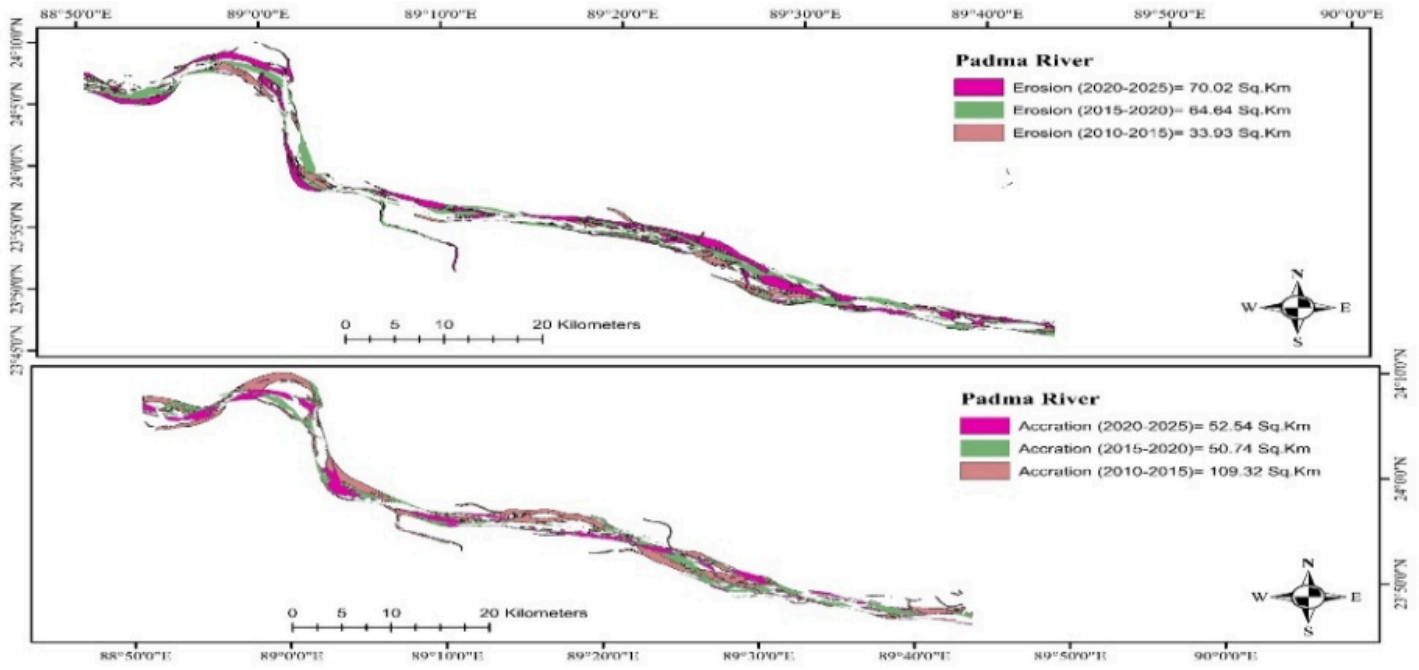


Figure 2

Erosion and Accretion ArcGIS view

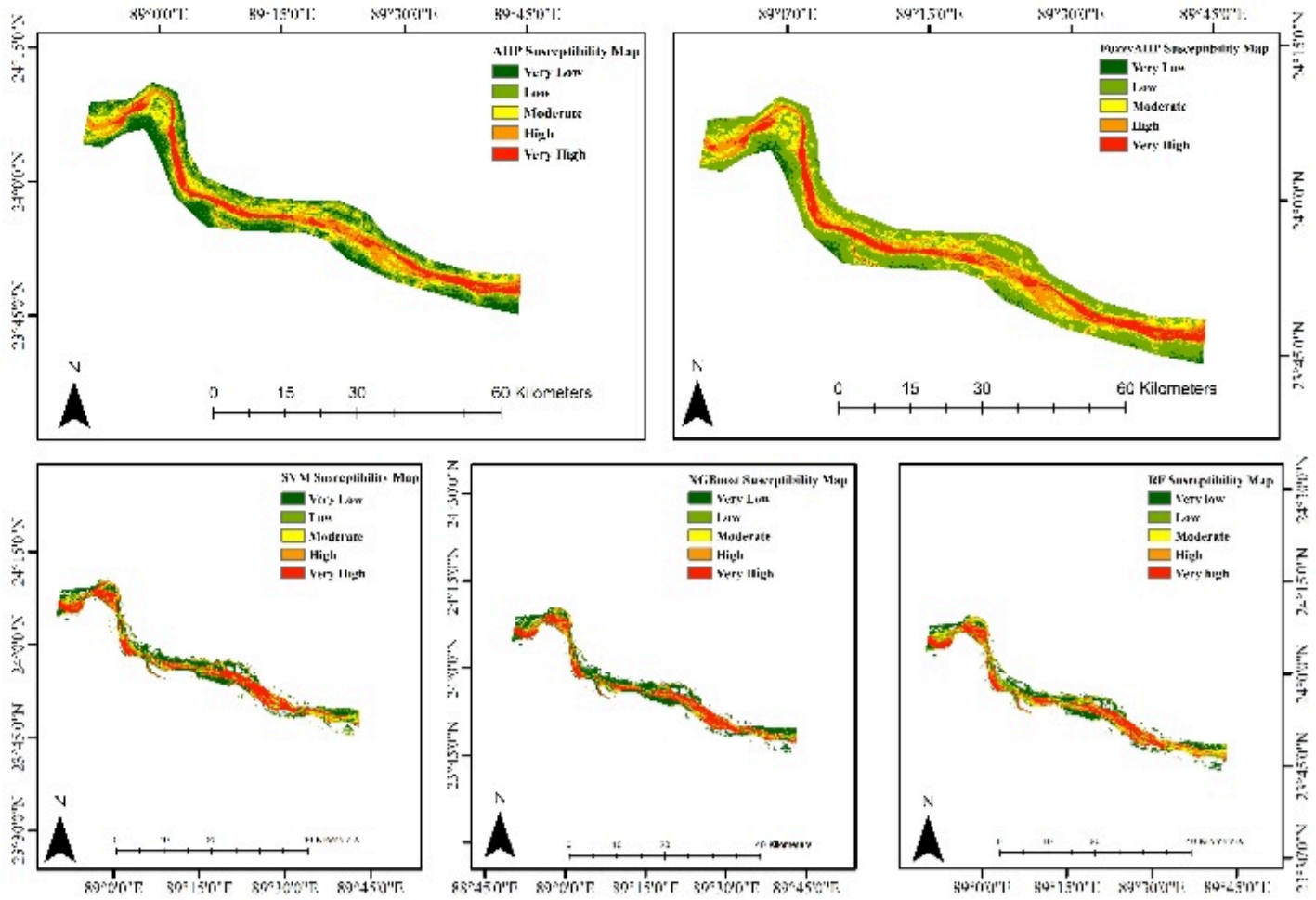


Figure 3

Bank erosion susceptibility maps

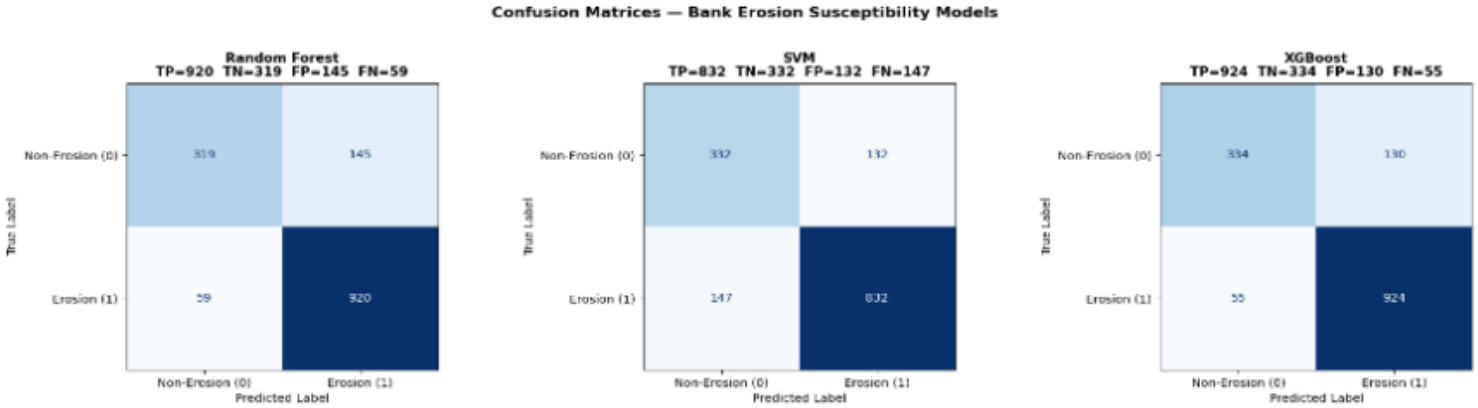


Figure 4

Confusion matrices

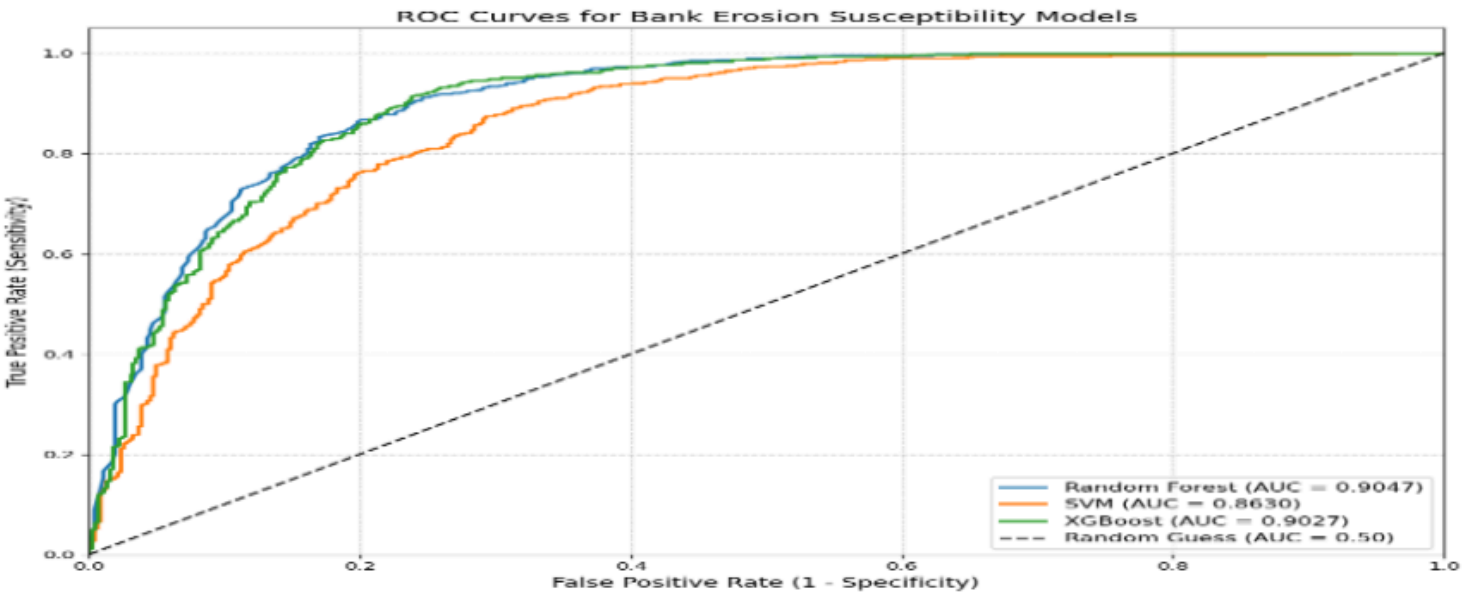


Figure 5

ROC curves for RF, SVM, and XGBoost

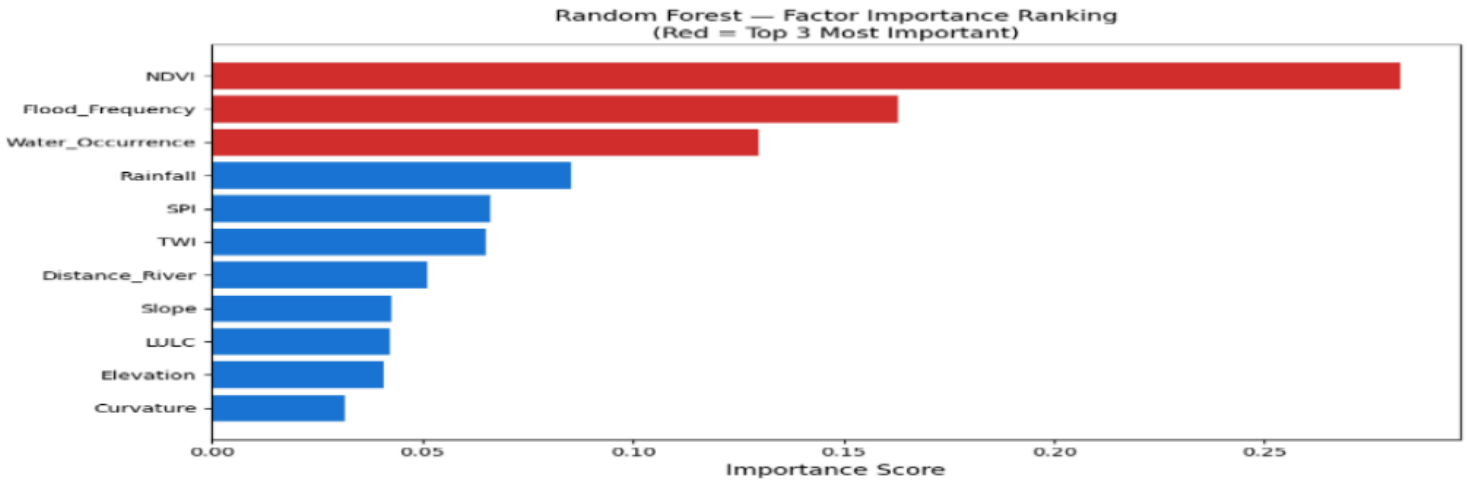


Figure 6

Random forest feature importance ranking