

Article

Spatial Prediction of Current and Future Flood Susceptibility: Examining the Implications of Changing Climates on Flood Susceptibility Using Machine Learning Models

Navid Mahdizadeh Gharakhanlou  and Liliana Perez * 

Laboratory of Environmental Geosimulation (LEDGE), Department of Geography, University of Montreal, 1375 Avenue Thérèse-Lavoie-Roux, Montréal, QC H2V 0B3, Canada

* Correspondence: l.perez@umontreal.ca; Tel.: +1-514-343-8003

Abstract: The main aim of this study was to predict current and future flood susceptibility under three climate change scenarios of RCP2.6 (i.e., optimistic), RCP4.5 (i.e., business as usual), and RCP8.5 (i.e., pessimistic) employing four machine learning models, including Gradient Boosting Machine (GBM), Random Forest (RF), Multilayer Perceptron Neural Network (MLP-NN), and Naïve Bayes (NB). The study was conducted for two watersheds in Canada, namely Lower Nicola River, BC and Loup, QC. Three statistical metrics were used to validate the models: Receiver Operating Characteristic Curve, Figure of Merit, and F1-score. Findings indicated that the RF model had the highest accuracy in providing the flood susceptibility maps (FSMs). Moreover, the provided FSMs indicated that flooding is more likely to occur in the Lower Nicola River watershed than the Loup watershed. Following the RCP4.5 scenario, the area percentages of the flood susceptibility classes in the Loup watershed in 2050 and 2080 have changed by the following percentages from the year 2020 and 2050, respectively: Very Low = -1.68% , Low = -5.82% , Moderate = $+6.19\%$, High = $+0.71\%$, and Very High = $+0.6\%$ and Very Low = -1.61% , Low = $+2.98\%$, Moderate = -3.49% , High = $+1.29\%$, and Very High = $+0.83\%$. Likewise, in the Lower Nicola River watershed, the changes between the years 2020 and 2050 and between the years 2050 and 2080 were: Very Low = -0.38% , Low = -0.81% , Moderate = -0.95% , High = $+1.72\%$, and Very High = $+0.42\%$ and Very Low = -1.31% , Low = -1.35% , Moderate = -1.81% , High = $+2.37\%$, and Very High = $+2.1\%$, respectively. The impact of climate changes on future flood-prone places revealed that the regions designated as highly and very highly susceptible to flooding, grow in the forecasts for both watersheds. The main contribution of this study lies in the novel insights it provides concerning the flood susceptibility of watersheds in British Columbia and Quebec over time and under various climate change scenarios.

Keywords: climate change; machine learning (ML); geographical information systems (GIS); flood susceptibility mapping; natural hazards



Citation: Mahdizadeh Gharakhanlou, N.; Perez, L. Spatial Prediction of Current and Future Flood Susceptibility: Examining the Implications of Changing Climates on Flood Susceptibility Using Machine Learning Models. *Entropy* **2022**, *24*, 1630. <https://doi.org/10.3390/e24111630>

Academic Editors: Dionissios T. Hristopoulos and Emmanouil Varouchakis

Received: 30 September 2022

Accepted: 9 November 2022

Published: 10 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Floods have become the most prevalent natural catastrophe, accounting for 44% of all natural disasters and harming 1.6 billion people globally between 2000 and 2019 [1]. Floods are the most common natural disaster in Canada [2]. According to the Canadian Disaster Database [3], there were 241 flood disasters in Canada between 1900 and 2005, about five times as frequently as wildfires, the second most common natural hazard in Canada.

Climate change poses a significant peril to present and future generations. Climate change has made natural disasters more unpredictable, causing them to occur more frequently and with more significant impact [4]. Climate changes inducing hydrological changes and precipitation amounts affect the likelihood of flood occurrences. Accordingly, floods are more likely to occur in areas where the climate shifts toward more intense and

frequent precipitation [5]. Although flood avoidance is inescapable, accurate flood forecasting, which considers the impacts of climate changes through proper models, might aid in future damage reduction.

Due to the intrinsic complexity of the flood phenomenon and the influence of various variables on floods, simple models are insufficient for accurate flood prediction [6]. In general, flood susceptibility modeling and mapping methodologies have been developed using two main types of models: physically-based and data-driven; albeit certain studies to assess flood susceptibility employed Multi-Criteria Decision Analysis (MCDA), such as the Analytical Hierarchy Process (AHP) and Analytical Network Process (ANP) [7–9]. The main drawback of the MCDA-based flood models is that they are prone to be distorted due to their dependence on expert knowledge [10].

Although physical models have shown to be capable of investigating a wide range of phenomena (e.g., rainfall-runoff [11], hydraulic models of flow [12], and flood [13]), developing physical flood-prediction models requires using fundamentally complex equations and in-depth knowledge and expertise of the flood phenomenon [14,15]. Owing to the drawbacks of the physical models, the usage of advanced data-driven models has been increasingly popular in recent decades [16–18]. When compared to physical models, data-driven models have three key advantages: (1) nonlinearity and numerical formulating of the flood based on historical data without requiring knowledge about the underlying physical processes, (2) providing more straightforward implementation with low computation cost, high performance, and high accuracy [19], and (3) relatively minor complexity [20].

Among the data-driven models, various ML models have been suggested and implemented to assess flood susceptibility [21,22]. The most frequently used ML models are Decision Tree (DT) [23], Random Forest (RF) [24,25], Naïve Bayes (NB) [26–28], Multilayer Perceptron Neural Network (MLP-NN) [29,30], Adaptive Neuro-Fuzzy Inference System (ANFIS) [31,32], Support Vector Machine (SVM) [33], Gradient Boosting Machine (GBM) [34], Fuzzy Logic [35], etc. Although there is no consensus on which method or group of methodologies may produce the most accurate predictions [36], ML models have recently successfully assessed flood susceptibility with greater accuracy [37,38].

Flood susceptibility is described as a quantitative or qualitative assessment of a place with geographical distribution of floods and a high likelihood of flooding [39]. Flood susceptibility maps illustrate the susceptibility of places to flooding and highlight locations that are prone to flooding. Enhancing the accuracy of flood susceptibility maps is a concern for flood disaster management researchers and decision-makers. Flood susceptibility maps become more practical to local governments and policy-makers as flood estimations get more precise. In recent years, advances in data collection and preparation methods using RS and GIS have led to an increase in the reliability and accuracy of flood prediction models, and consequently flood susceptibility mapping [40–42]. RS provides a variety of data sources, data with excellent quality, day and night data gathering capabilities, and rapid analysis [43,44], and GIS is designed for the storage, retrieval, and analysis of geographically referenced data [45].

Numerous studies have been conducted using ML models to evaluate the impacts of climate change on the flood susceptibility [46–50]. However, ML models of MLP-NN, RF, NB, and GBM have not yet been used to investigate the impacts of climate change on the flood susceptibility for two watersheds in Canada, namely, Lower Nicola River, BC and Loup, QC. In doing so, we investigated the impacts of climate changes for the years of 2020, 2050, and 2080 under three different climate change scenarios of RCP2.6 (i.e., optimistic scenario), RCP4.5 (i.e., business as usual scenario), and RCP8.6 (i.e., pessimistic scenario) to contribute towards a dynamic estimation of flood susceptibility in these areas. Moreover, another highlight of the present study was to consider various topographic, hydrologic, environmental, and geologic flood conditioning factors yield by means of RS and GIS techniques.

2. Materials and Methods

2.1. Description of the Study Areas

This study focuses on two watersheds, one in Quebec (QC) and the other in British Columbia (BC) provinces. On 3 May 2017, Eastern Canada was flooded due to heavy rain, with QC being the worst hit. The Loup watershed located in QC was one of the watersheds affected by the May 2017 flood. Likewise, the watershed of the Lower Nicola River, located in south-central BC province, was struck by a flood triggered by heavy rain on 14 November 2021. The location of both watersheds was depicted in Figure 1.

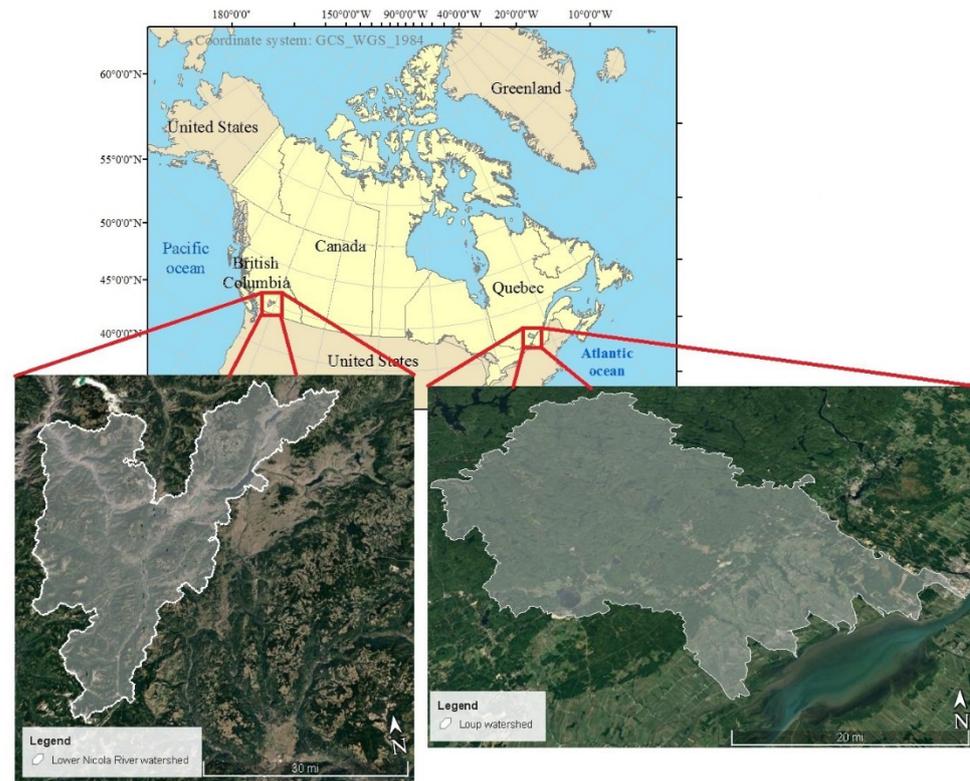


Figure 1. The geographical location of study watersheds, namely Loup, QC and Lower Nicola River, BC.

This study used four flood susceptibility assessment models to predict flood susceptibility in two different study regions correctly. The primary modeling technique for this investigation was broken down into six significant steps: (i) gathering and preparing the factors influencing flooding, (ii) iteratively picking flood and non-flood points in study areas and calculating Moran's I spatial autocorrelation for all factors; the set of points which the p -value for all factors was obtained extremely close to zero and less than the threshold (i.e., 0.05) was chosen to create the flood inventory map, (iii) assessing the correlation between flood occurrence and flood influencing factors using multicollinearity analysis and either including or excluding them in the following processes, (iv) training the ML models, evaluating and comparing their performance using three statistical metrics, and choosing the model with the highest accuracy, (v) gathering and preparing the annual precipitation data in three years of 2020, 2050, and 2080 under three climate change scenarios (i.e., RCP2.6, RCP4.5, and RCP8.5), and (vi) providing associated flood susceptibility maps with the years and scenarios. The methodology flowchart of the research was shown in Figure 2.

2.2. Flood Inventory Maps

The flood inventory map depicts the flooded and non-flooded locations that are used to train ML models. The basic premise behind the flood inventory map is that future floods will follow the same pattern as previous floods. The flood inventory map can

be produced by field survey, satellite images before and after the flooding, topographic maps, and Google Earth software. Accordingly, having performed the Moran's I spatial autocorrelation analysis for all flood influencing factors in the points picked, the flood inventory map was provided for each study watershed by ascertaining 120 flood sites using data from previous floods collated from satellite images, topographic maps, and Google Earth software in both watersheds (Figure 3). It is worth mentioning that the flood inventory maps created for the Loup and Lower Nicola River watersheds were a compilation of a single flood event that occurred on 3 May 2017, and 14 November 2021 respectively. Although, no specific literature on the number of flood-present and flood-absent places has been discovered, an approximately equal number of them is preferred for flood susceptibility mapping [27,51]. Accordingly, 120 non-flood sites were picked randomly in both watersheds, to establish a dichotomous dependent variable for modeling. Following the earlier studies [5,52], random selection was used to split two datasets into training (70%) and testing (30%) sets.

2.3. Flood Explanatory Factors and Their Preparation Processes

The intensity and severity of floods majorly rely on topographic, hydrologic, environmental, and geologic factors [53]. Accordingly, the sixteen factors reflecting topographic, hydrologic, environmental, and geologic attributes were recognized, and the data that describe each flood explanatory factor was compiled for both watersheds. Concerning a thorough review of the related literature [5,54–56], the following sets of factors were chosen: elevation, slope, aspect, plan curvature, profile curvature, roughness, Topographic Wetness Index (TWI), land cover, precipitation, distance from rivers, drainage density, lithology, soil, Stream Power Index (SPI), Normalized Difference Vegetation Index (NDVI), and Normalized Difference Moisture Index (NDMI).

The influences of topographic-related factors (i.e., elevation, slope, aspect, curvature, SPI, TWI, and roughness) derived from the digital elevation model (DEM) on flood occurrences have been well recognized in the literature [5,57,58]. The effects of elevation on hydrology and floods are significant [59]. Floods are uncommon in high-elevation places, while runoff gathers from above at lower altitudes, making floods more prevalent [60]. The DEM with a spatial resolution of 30 m was produced through Shuttle Radar Topography Mission (SRTM) and clipped to the study areas' border. Slope determines the rates of surface runoff [61]. Plan and profile curvatures reveal the concavity and convexity of slopes influence on flow velocity [5]. The aspect factor is connected with water flow convergence and directions [56]. The slope, aspect, plan and profile curvature layers were provided by applying the spatial analyst tools of Slope, Aspect, and Curvature on the DEM, respectively. SPI and TWI are also two common topographical factors affecting flow intensity and water accumulation [62]. The SPI and TWI were also obtained from the DEM layer using the Raster Calculator tool according to Equations (1) and (2), respectively.

$$SPI = \alpha * \tan \beta \quad (1)$$

$$TWI = \ln \left(\frac{\alpha}{\tan \beta} \right) \quad (2)$$

Here, α denotes the cumulative upstream discharge at one point, or flow accumulation (m^2m^{-1}), and β is the slope (in radian).

Roughness which indicates the elevation differences between neighboring pixels, is another factor affecting the surface runoff [63]. To generate the roughness layer from the DEM layer, the Focal Statistics tool was used three times to acquire the mean, minimum, and maximum focal statistics layers. Then, the Raster Calculator tool was applied to them using Equation (3).

$$Roughness = (FS_{mean} - FS_{min}) / (FS_{max} - FS_{min}) \quad (3)$$

Here, FS_{mean} , FS_{min} , and FS_{max} represent the mean, minimum, and maximum focal statistical layer, respectively.

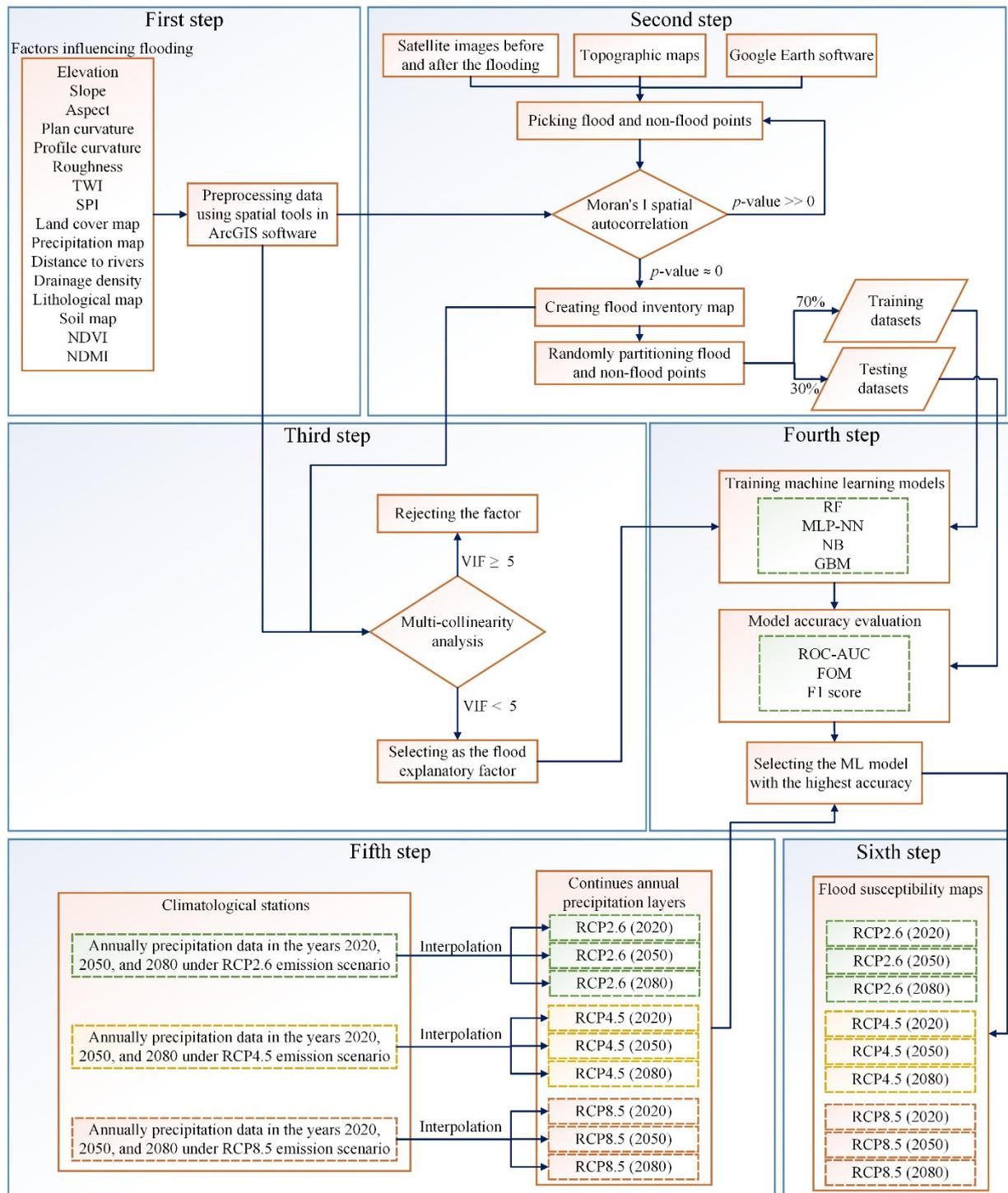


Figure 2. The methodology flowchart of research.

Another factor influencing the likelihood of flooding is stream density, which measures how much of a watershed is drained by stream channels [64]. Having prepared the stream layers from DEM using the Hydrology tools, the Line Density tool was used to obtain the stream density layer. Concerning the earlier studies [65,66], the likelihood of flooding is also influenced by proximity to a river. As the distance from rivers decreases, the chance of flooding increases. To acquire the layer of distance from rivers, the Euclidean Distance tool

was applied to the river polyline shapefile. Soil and land cover are also influencing factors due to their influence on the infiltration and runoff speed [67,68]. Likewise, geology which indicates underlying rock types impacts infiltration and runoff in watersheds [38]. To obtain the Soil, Land cover, and geology maps for the study areas, having downloaded the layers, they were clipped to the study areas' border. Precipitation is also a hydrologic factor that significantly influences the incidence of floods [69]. To provide continuous layers of average annual precipitation, first, annual precipitation was gathered at 10 climatological stations (both inside and outside the watersheds) for the period 2000–2020. After calculating the average annual precipitation at climatological stations, the Ordinary Kriging interpolation method was used. Vegetation, which on one hand impacts the evaporation process and hydrological cycle while also acting as a barrier to the flow of water over the ground, has a substantial impact on run-offs and floods. Accordingly, NDMI, a metric indicating the moisture content of vegetation, was employed in the modeling process [70]. There is also an inverse relationship between vegetation density and floods [34]. NDVI, an essential metric representing vegetation coverage, was also considered in the modeling process [71]. The NDMI and NDVI layers in the study area were obtained from the Landsat 8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) satellite images using the Raster Calculator according to Equations (4) and (5), respectively.

$$NDMI = \frac{NIR - SWIR}{NIR + SWIR} \quad (4)$$

$$NDVI = \frac{NIR - R}{NIR + R} \quad (5)$$

Here, *NIR* is the Near Infrared band (band 5 of Landsat 8), *SWIR* is the Short-Wave Infrared band (band 6 of Landsat 8), and *R* is the Red band (band 4 of Landsat 8).

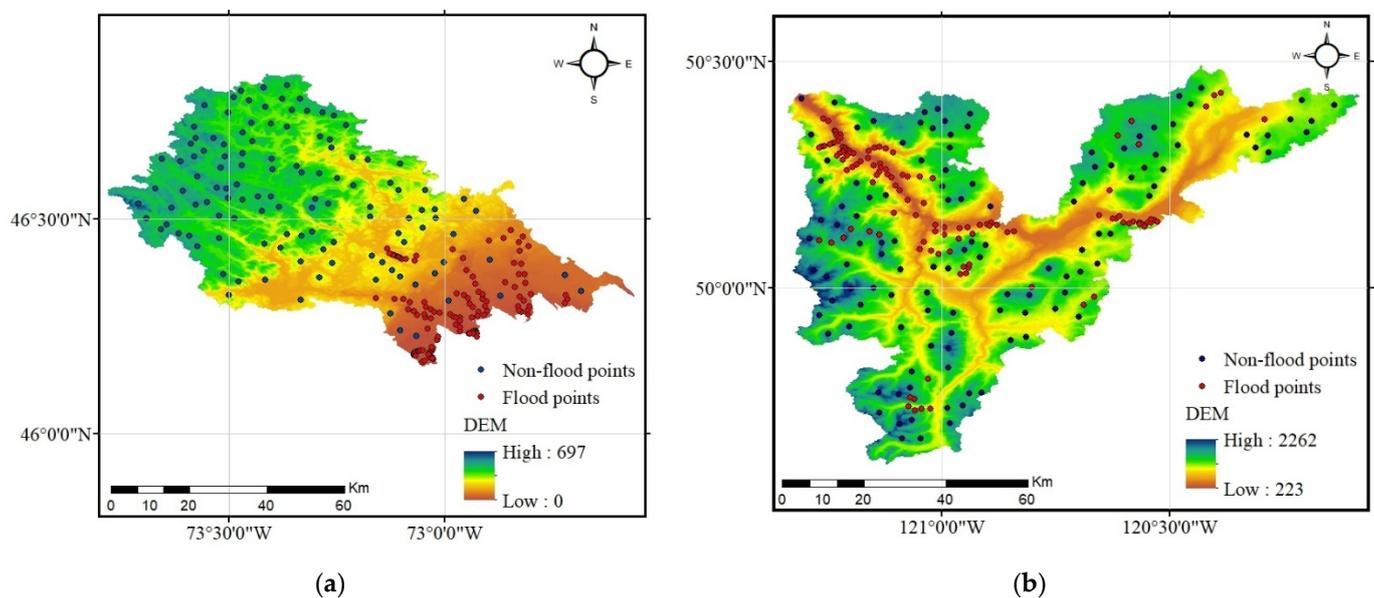


Figure 3. Flood inventory maps: (a) Loup watershed, and (b) Lower Nicola River watershed.

All flood explanatory data acquired for this investigation, along with their sources, were summarized in Table 1. The overall data preparation flowchart was given in Figure A1 (Appendix A). Moreover, each flood explanatory factor was plotted on a map after the preparation processes (Figures A2 and A3, Appendix B). All the factors were designed to have a comparable spatial scope of 30 m pixel size due to the spatial resolution of the land cover data (i.e., 30 m).

Table 1. Flood explanatory data along with their sources.

Primary Input Data	Original Format Sources	Source	Derived Map
Shuttle Radar Topography Mission (SRTM); DEM	Raster	United States Geological Survey (USGS); https://earthexplorer.usgs.gov/ (accessed on 1 March 2022)	Elevation
			Slope
			Aspect
			Plan curvature
			Profile curvature
			SPI
			TWI
			Roughness
Land cover map	Vector (i.e., polygons)	Government of Canada, Natural resources Canada; https://open.canada.ca/data/en/dataset/4e615eae-b90c-420b-adee-2ca35896ca66 (accessed on 1 March 2022)	Land cover map
Climatological stations	Vector (i.e., points)	Government of Canada, Environment and natural resources; https://climate.weather.gc.ca/historical_data/search_historic_data_e.html (accessed on 1 March 2022) and Climate Data Canada;	Precipitation map
Meteorological data	Numerical data	https://climatedata.ca/ (accessed on 1 March 2022)	
Streams, Rivers, and water bodies	Vector (i.e., polylines)	Government of Canada, Statistics Canada; https://open.canada.ca/data/en/dataset/448ec403-6635-456b-8ced-d3ac24143add (accessed on 1 March 2022)	Distance from rivers
			Drainage density
Geological map	Vector (i.e., polygons)	Government of Canada, Natural resources Canada, Geological Survey of Canada; https://geoscan.nrcan.gc.ca/starweb/geoscan/servlet.starweb?path=geoscan/download_e.web&search1=R=295462 (accessed on 1 March 2022)	Lithological map
Soil map	Vector (i.e., polygons)	Government of Canada, Agriculture and Agri-Food Canada; https://open.canada.ca/data/en/dataset/5ad5e20c-f2bb-497d-a2a2-440eec6e10cd (accessed on 1 March 2022)	Soil map
Landsat 8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS)	Raster	United States Geological Survey (USGS); https://earthexplorer.usgs.gov/ (accessed on 1 March 2022)	NDVI
			NDMI

2.4. Multicollinearity of Flood Explanatory Factors

Before implementing the models, a multicollinearity investigation of the independent variables is indispensable to reduce the risk of inaccuracy in flood susceptibility models. The multicollinearity analysis investigates whether the variables are affected by multicollinearity. In doing so, multicollinearity involves tightly coupling many independent variables in a multiple regression model and removing variables with significant collinearity. Variance inflation factors (VIF) and tolerance (TOL) are two exponents frequently used to analyze the multicollinearity of variables.

Table 2 shows the VIF and TOL calculated values for the proposed flood influencing factors in multicollinearity analysis in both watersheds. The TOL values less than 0.1 or VIF values greater than 10 indicate the multi-collinearity issue [72]. However, the threshold of 5 for VIF was taken into consideration in this study to choose significant independent predictors with a high degree of certainty. Accordingly, except for DEM factor

in the Loup watershed, the rest of the 15 explanatory variables were allowed for usage throughout the modeling process. In the Lower Nicola River watershed, on the other hand, the multicollinearity statistics indicated that all 16 explanatory variables could be included in the modeling process.

Table 2. VIF and Tolerance values in multi-collinearity analysis for all flood explanatory factors in both watersheds.

Predictors/Factors	Collinearity Statistics in the Loup Watershed		Collinearity Statistics in the Lower Nicola River Watershed	
	Tolerance	VIF	Tolerance	VIF
SPI	0.735	1.360	0.632	1.581
TWI	0.344	2.906	0.383	2.612
Precipitation	0.228	4.388	0.680	1.470
Drainage density	0.388	2.575	0.368	2.718
Distance from rivers	0.379	2.641	0.413	2.419
Lithology	0.356	2.811	0.805	1.242
Soil	0.511	1.955	0.578	1.731
Land cover	0.358	2.790	0.824	1.213
NDVI	0.231	4.338	0.458	2.182
NDMI	0.350	2.858	0.615	1.626
Roughness	0.444	2.254	0.856	1.168
Plan curvature	0.573	1.746	0.622	1.607
Profile curvature	0.572	1.747	0.663	1.509
Aspect	0.760	1.316	0.928	1.077
Slope	0.571	1.752	0.472	2.119
DEM	0.164	6.107	0.338	2.956

2.5. Predicting Future Precipitation Data

Many variables impact climate, most notably human activity, and greenhouse gas emissions. Although there are considerable uncertainties in climate forecasts due to the complex nature of the climate system, greenhouse gas emissions, and human activities, some aspects of this variability are thought to be predictable for a decade or more in advance. Emissions scenarios are one method of presenting a variety of possible futures depending on various future emissions. Accordingly, a collection of scenarios known as Representative Concentration Pathways (RCPs) is frequently used to investigate future climate change. RCPs are intended to offer probable future human emission trends. These include considering future greenhouse gas emissions, deforestation, population growth, and a variety of other factors. Based on best practices in the global science community, the Government of Canada typically offers three RCPs: RCP8.5 (high global emission scenario), RCP4.5 (medium global emission scenario), and RCP2.6 (low emission global scenario) [73].

In this study, Coupled Model Intercomparison Project Phase 5 (CMIP5) climate model datasets which were downscaled and bias-adjusted using the BCCAQv2 method were utilized. The preparation of future precipitation data was carried out in two steps: first, the annual precipitation data under three emission scenarios of RCP2.6 (optimistic scenario), RCP4.5 (business as usual scenario), and RCP8.5 (pessimistic scenario) were collected (from <https://climatedata.ca/> (accessed on 1 March 2022)) at climatological stations inside and outside of the study areas in the years 2020, 2050, and 2080; then, the annual precipitation amounts at stations were interpolated in ArcGIS using the Ordinary Kriging interpolation method to provide continuous annual precipitation layers for all three scenarios. Having prepared the precipitation layers that corresponded with each scenario for the years 2020, 2050, and 2080, they were used in the ML models to provide the associated flood susceptibility maps.

2.6. Methods for Flood Susceptibility Modeling

2.6.1. Multilayer Perceptron Neural Network (MLP-NN)

MLP-NN is classified as a feed-forward neural network trained using supervised and back-propagation learning methods. MLP-NN has been widely employed as a benchmark model in a variety of studies owing to its capabilities in the prediction and modeling of nonlinear and complicated phenomena [74,75]. Basically, the MLP-NN model comprises a system of simply interconnected neurons that are organized into three layers: an input layer, one or more hidden layers, and finally, an output layer. Neurons in each layer receive values, which are multiplied by corresponding weights, then summed up and passed through a nonlinear function (i.e., activation function) [76]. Using the activation function on the weighted sum enables the MLP-NN to account for the nonlinear relationship between the independent and dependent variables [77]. Accordingly, the MLP-NN model estimates the nonlinear connections between the independent variables (i.e., flood explanatory factors) and the dependent variable (i.e., flood occurrences).

The neurons in two sequential layers are linked by the unknown weights whose values are estimated through the iterative back-propagation learning technique. The back-propagation approach is generally an iterative gradient-based learning technique (e.g., Stochastic Gradient Descent (SGD)) that aims to reduce the discrepancy between the outputs of the network and actual target values (i.e., reduce the value of the cost function) by estimating the weights in each iteration [76]. The architecture of an MLP-NN was shown in Figure 4. The connections of the nodes from different layers are made using Equations (6)–(8).

$$t_j = \varphi \left(\sum_{i=1}^d w_{ij}x_i + w_{0j} \right) \tag{6}$$

$$y_k = \varphi \left(\sum_{j=1}^{n_{H_1}} w_{jk}t_j + w_{0k} \right) \tag{7}$$

$$z_v = f \left(\sum_{k=1}^{n_{H_2}} w_{kv}y_k + w_{0v} \right) \tag{8}$$

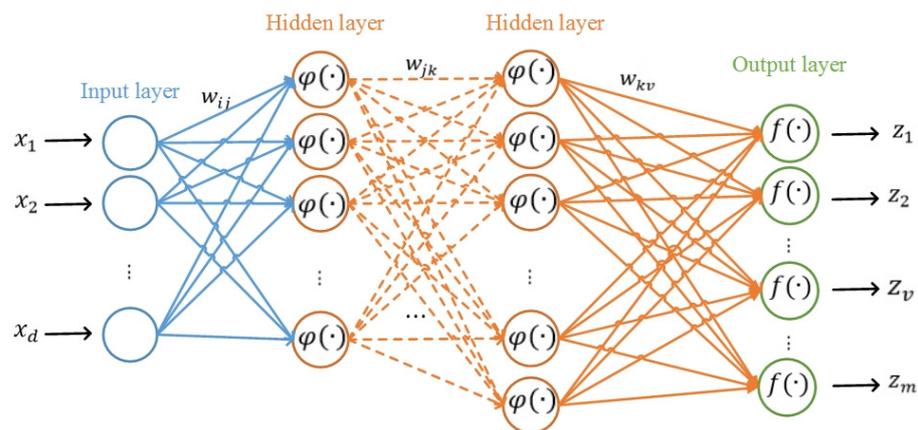


Figure 4. The structure of the MLP-NN model.

Here, presuming the MLP-NN made up of two hidden layers, d , n_{H_1} , and n_{H_2} are the number of nodes in the input, first hidden, and second hidden layer, respectively. w_{ij} , w_{jk} , and w_{kv} are the connection weights between two nodes from two consecutive layers. Moreover, w_{0j} , w_{0k} , and w_{0v} are the intercepts of the input, first hidden and second hidden layer, respectively. x_i , t_j , y_k , and z_v are the nodes in each input layer, first hidden, second hidden, and output, respectively. φ is the activation function of all layers except the output layer, and f is the activation function of the output layer.

2.6.2. Naïve Bayes (NB) Model

NB methods are a set of supervised learning algorithms based on Bayes' Theorem and the assumption of conditional independence between every pair of features. In other words, an NB classifier posits that the existence of one feature in a class is independent of the presence of any other feature. The Bayes theorem proposes a method for computing posterior probability $P(c|x_1, x_2, \dots, x_d)$ from $P(c)$, $P(x_1, x_2, \dots, x_d)$, and $P(x_1, x_2, \dots, x_d|c)$ (Equation (9)) given the naive conditional independence assumption (Equation (10)) [78].

$$P(c|x_1, x_2, \dots, x_d) = \frac{P(x_1, x_2, \dots, x_d|c) * P(c)}{P(x_1, x_2, \dots, x_d)} \quad (9)$$

$$P(x_1, x_2, \dots, x_d|c) = \prod_{i=1}^d P(x_i|c) = P(x_1|c) * P(x_2|c) * \dots * P(x_d|c) \quad (10)$$

Here, $P(c|x_1, x_2, \dots, x_d)$ is the posterior probability of class (c , target) given features, (x_1, x_2, \dots, x_d) , $P(c)$ is the prior probability of class, $P(x_1, x_2, \dots, x_d|c)$ is the likelihood which is the probability of predictor given class, and $P(x_1, x_2, \dots, x_d)$ is the prior probability of predictor.

The Gaussian NB method was chosen from among the various types of NB methods (e.g., Gaussian, Multinomial, Complement, Bernoulli) owing to its common use in classification. The Gaussian NB model assumes that features follow a normal distribution, and the likelihood of the features is calculated according to Equation (11). The parameters σ_c and μ_c in Equation (11) are estimated using maximum likelihood.

$$P(x_i|c) = \frac{1}{\sqrt{2\pi\sigma_c^2}} \exp\left(-\frac{(x_i - \mu_c)^2}{2\sigma_c^2}\right) \quad (11)$$

2.6.3. Random Forest (RF)

RF is one of the popular ML algorithms for addressing multi-classification and prediction issues [79]. The RF technique is a collection of DTs used to predict categorization or regression. The main procedure of the RF algorithm is to (1) resample the original data set using bootstrap (i.e., sampling with replacement) to generate various subsets with sizes equal to the original set, (2) use the subsets to construct DTs, and (3) combine the prediction or classification results of all the decision trees to obtain the final results [80]. One of the significant issues with DTs is that they are highly sensitive to training data and tend to over-fit the training datasets, consequently, perform poorly when an unknown dataset is given. Using the RF method to address this flaw is a viable option. Accordingly, a portion of the input records, as well as features, were picked at random, and DTs were created according to each set of inputs and features chosen.

2.6.4. Gradient Boosting Machine (GBM)

GBM is a supervised machine learning approach for classification and regression problems that use a prediction model in the form of an ensemble of weak prediction models. The central notion underlying it is a model built from a set of weak learners, commonly decision trees (DTs). The GBM is similar to functional gradient descent in that it applies a new learner to residual errors created by the prior learner to minimize a loss at each gradient descent step [81]. As with other boosting methods, various loss functions might be considered. The constructed decision tree is optimized with the gradient boosting approach in this model. Gradient boosting approaches create the solution and address the over-fitting problem by maximizing the loss functions in a stage-wise structure [82]. Presuming a custom base-learner $h(x, \theta)$ (such as a decision tree) and a loss function $\Psi(y, f(x))$; directly estimating the parameters is challenging; hence, an iterative model is recommended. The

model will be updated, and $h(x, \theta)$ will be chosen as the new base-learner feature, with the t increment driven by Equation (12) [81].

$$g_t(x) = E_y \left[\frac{\partial \Psi(y, f(x))}{\partial f(x)} \mid x \right]_{f(x)=\hat{f}^{t-1}(x)} \quad (12)$$

Instead of searching the function space for a general solution for the boost increment, one may just select the new function increment that is the most correlated with $g_t(x)$. This replaces the hard optimization problem with the standard least-squares optimization problem according to Equation (13) [81].

$$(p_t, \theta_t) = \arg \min_{p, \theta} \sum_{i=1}^N [-g_t(x_i) + \rho h(x_i, \theta)]^2 \quad (13)$$

The procedure of GBM algorithms include: (i) presuming that \hat{f}_0 is constant, (ii) calculating $g_t(x)$ and training $h(x_i, \theta)$ function, and (iii) finding element ρ_i and updating the function $\hat{f}_i = \hat{f}_{i-1} + \rho_i h(x_i, \theta)$.

2.7. Model Evaluation Metrics

2.7.1. Receiver Operating Characteristic (ROC) Curve

A ROC curve is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied. The ROC curve is a popular method for assessing the performance of predictive models [83]. This accuracy criterion has been increasingly utilized in a variety of susceptibility mapping applications using ML models, including landslide [84,85], earthquake [86,87], and flood [88–90]. For quantitative evaluation, this method employs the area under the ROC curve (AUC), which plots false positive rate (FPR) (Equation (14)) on the x-axis against true positive rate (TPR) (Equation (15)) on the y-axis [91]. A higher AUC value depicts a better goodness-of-fit of the model. Generally, AUC values ranging from 0.8 to 0.9 indicate extremely strong performance for the prediction model [88].

$$FPR = 1 - Specificity = \frac{FP}{FP + TN} \quad (14)$$

$$TPR = Sensitivity = \frac{TP}{TP + FN} \quad (15)$$

Here, TP (true positive) and TN (true negative) are test results that correctly indicate the presence and absence of a condition or characteristics, respectively. FP (false positive) and FN (false negative) are test results that incorrectly indicate the presence and absence of a condition or characteristics, respectively.

2.7.2. Figure of Merit (FOM)

The FOM is a statistical measure of sample set similarity and diversity. Equation (16) is used to calculate the FOM , which is the equivalent of the Jaccard index. The FOM has a value range of 0 to 1, with 1 being the ideal match [92].

$$FOM = \frac{TP}{TP + FP + FN} \quad (16)$$

2.7.3. F1 Score

In binary classification statistical analysis, the F-score measures a test's accuracy. The accuracy and recall (or sensitivity) of the test are used to calculate it, with precision equaling the number of true positive results divided by the total number of positive results, including those that were incorrectly identified. The harmonic mean of accuracy and recall is the F1 score (also known as the Dice similarity coefficient). The maximum possible F-score is 1,

which indicates flawless accuracy and recall, while the lowest possible F-score is 0 if neither precision nor recall is zero. F1 score is calculated according to Equation (17) and auxiliary Equations (18) and (19) [93].

$$F_1 = \frac{2 * precision * recall}{precision + recall} \quad (17)$$

$$precision = \frac{TP}{TP + FP} \quad (18)$$

$$recall = \frac{TP}{TP + FN} \quad (19)$$

3. Results

3.1. Model Validation and Performance Assessment

Before evaluating the findings, it is indispensable to determine the optimal values of various hyper-parameters in each ML model. Accordingly, to designate the ideal values for hyper-parameters and reduce the over-fitting issues, the 5-fold cross-validation method was used. Having employed the 5-fold cross-validation method, the ROC-AUC values for the training and validation dataset (Table 3) were calculated, considering the ideal values of corresponding hyper-parameters in each ML model.

Table 3. The ROC-AUC value achieved for the training and validation dataset using various ML models.

ML Models	The Loup Watershed		The Lower Nicola River Watershed	
	Training	Validation	Training	Validation
RF	1.0	1.0	1.0	0.9968
NB	0.9991	0.9805	0.9776	0.9571
MLP-NN	0.9644	0.9614	0.8817	0.8503
GBM	1.0	0.9978	1.0	0.9967

Three accuracy metrics of ROC-AUC, FOM, and F1 score were utilized to quantitatively measure and compare the efficiency of ML models in assessing flood susceptibility and validate the models (Table 4). It is worth emphasizing that these performance indicators were used to evaluate the models' spatial distribution performance since they reflect the degree to which the observed flood points overlap the flood susceptibility. Even though the differences in performance between the RF and GBM models were insignificant, the RF model outperformed the GBM model with a very trifling difference given three accuracy metrics. Moreover, the models were ranked in order of performance from the best to the worst: RF, GBM, NB, and MLP-NN. The ROC curves were also plotted in Figure 5 to evaluate and compare the classifiers' quality independent of the threshold.

Table 4. The calculated values of three accuracy metrics for each ML model.

ML Models	The Loup Watershed			The Lower Nicola River Watershed		
	ROC-AUC	FOM	F1 Score	ROC-AUC	FOM	F1 Score
RF	0.9992	0.9767	0.9882	0.9996	0.9787	0.9892
NB	0.9548	0.8864	0.9398	0.9884	0.8036	0.8911
MLP-NN	0.9643	0.7857	0.88	0.9411	0.6458	0.7848
GBM	0.9901	0.9333	0.9655	0.9979	0.9787	0.9892

3.2. Flood Susceptibility Map

Providing the flood susceptibility maps and investigating the impacts of climate change on flood susceptibility of study areas throughout time were the primary objectives of this study. Accordingly, the flood susceptibility maps were mapped for both Loup (Figure 6) and Lower Nicola River (Figure 7) watersheds in the years 2020, 2050, and 2080 under

three climate change scenarios of RCP2.6, RCP4.5, and RCP8.5. As briefly explained in Section 2.5, to provide flood susceptibility maps under climate change scenarios, the former precipitation layer was replaced with the intended climate change scenario precipitation layer; then, the intended precipitation layer was fed the ML model along with the rest flood explanatory factors. Finally, the resulted flood susceptibility values were stratified into five classes from very low to very high susceptibility using the natural break classification technique in ArcGIS 10.8 software. The area percentages of each flood susceptibility class given the year and climate change scenario were calculated for both the Loup watershed (Figure 8) and the Lower Nicola River watershed (Figure 9).

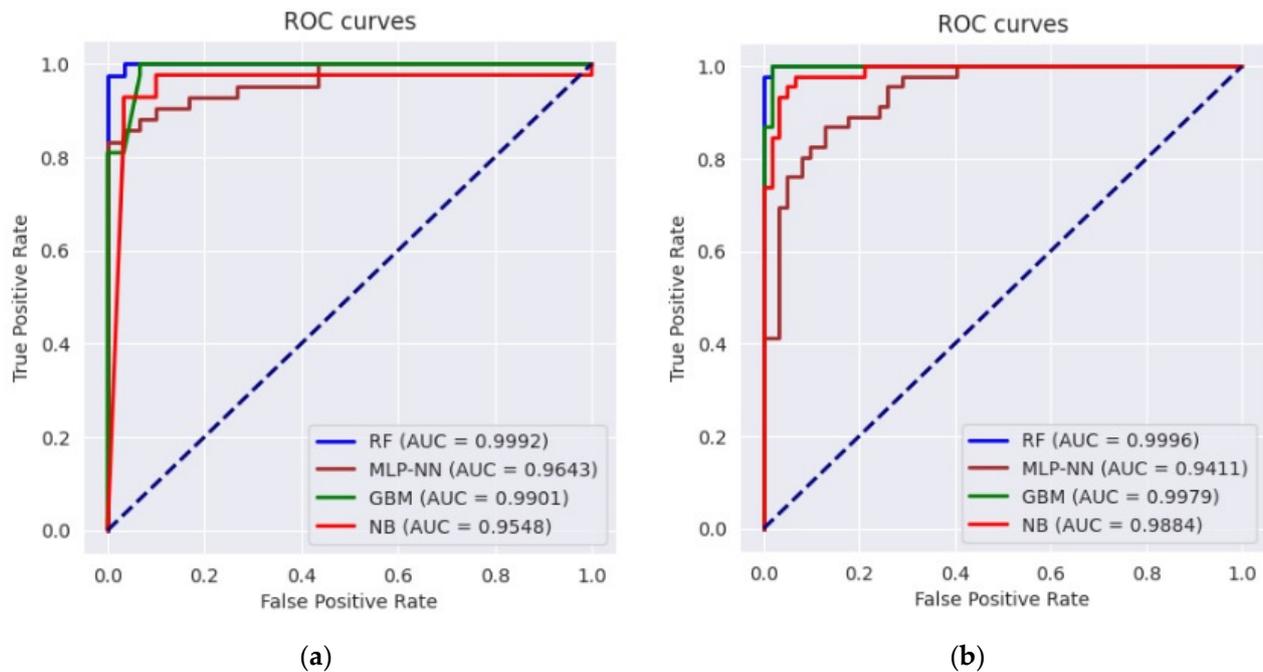


Figure 5. ROC curves along with the AUC values for each ML model: (a) the Loup watershed and (b) the Lower Nicola River watershed.

Considering the scenario RCP4.5 as the baseline scenario, the derived flood susceptibility map of the Loup watershed in the year 2020 revealed that the flood susceptibility was very low in 54.25% of the Loup watershed, low in 18.74%, moderate in 8.48%, high in 8.03%, and very high in 10.49%. Likewise, the resulted flood susceptibility map for the Lower Nicola River watershed in the year 2020 under the scenario RCP4.5 indicated that approximately 25.5%, 22.44%, 19.33%, 17.41%, and 15.32% of the watershed were in very low, low, moderate, high, and very high susceptibility classes, respectively. Assuming no changes in the emission scenario in the following years, in the year 2050, the flood susceptibility will be very low, low, moderate, high, and very high in 52.57%, 12.92%, 14.67%, 8.75%, 11.09% of the Loup watershed, respectively, and in the year 2080, in 50.96%, 15.9%, 11.18%, 10.03%, 11.92% of the watershed, respectively. Likewise, the area percentages of flood susceptibility classes for the Lower Nicola River in the year 2050 will be 25.12%, 21.63%, 18.38%, 19.13%, 15.74%, respectively, and in the year 2080 will be 23.81%, 20.28%, 16.57%, 21.5%, and 17.84%, respectively. Following the changes in the flood susceptibility classes in the Loup watershed, it can be concluded that irrespective of some fluctuations, the overall trend of changes is decreasing for the area percentages of very low and low flood susceptibility classes and increasing for the area percentages of the moderate, high, and very high flood susceptibility classes. As a result, the area percentages of the very low and low flood susceptibility classes were lowered and added to the area percentages of the moderate, high, and very high flood susceptibility classes, indicating that the flood susceptibility of the Loup watershed worsens over time. Similarly, the trend of changes

in the flood susceptibility classes in the Lower Nicola River watershed was decreasing in the very low, low, and moderate classes and increasing in the high and very high classes, indicating that the flood susceptibility of the Lower Nicola River watershed worsens over time, as well.

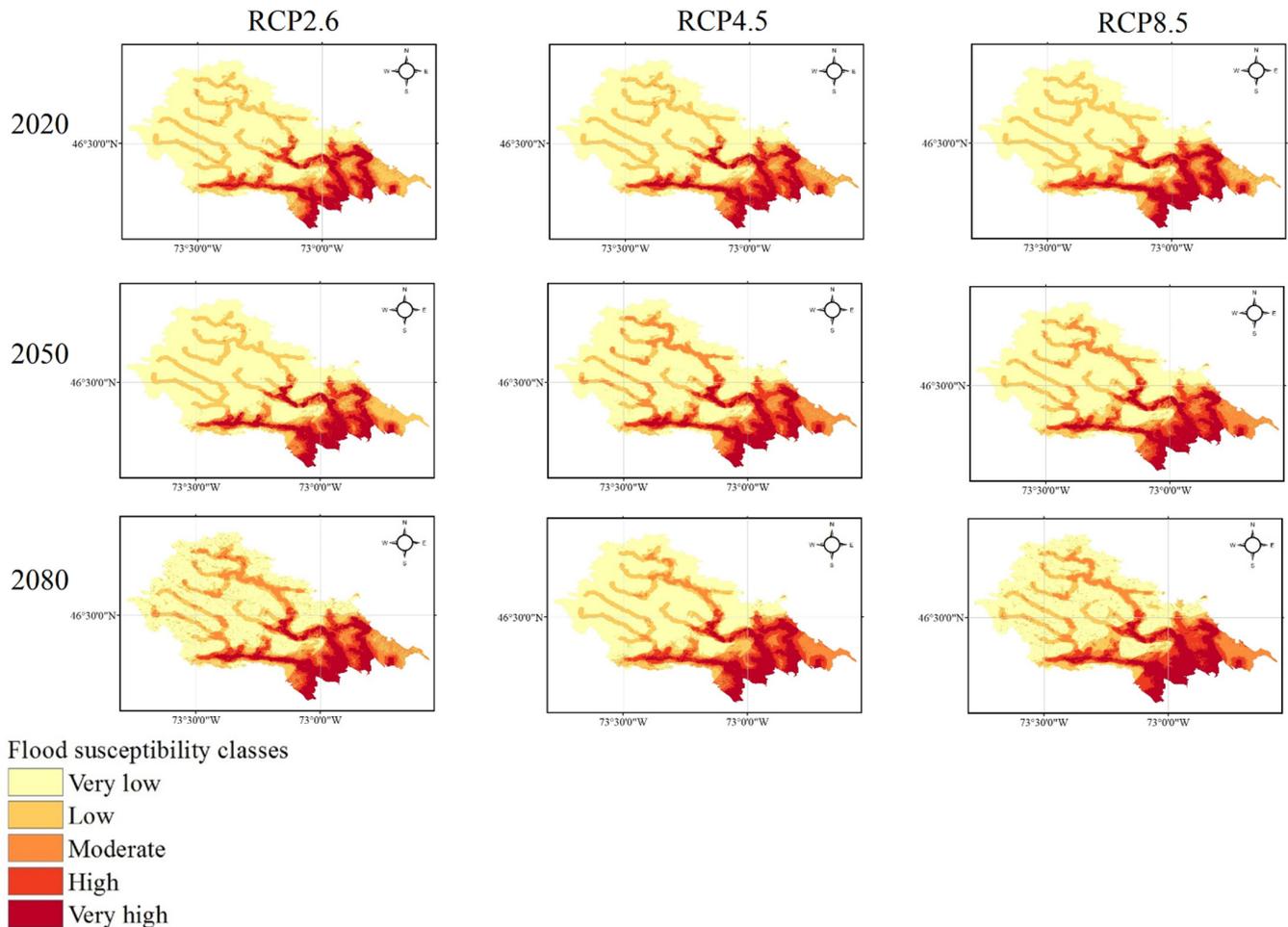


Figure 6. The flood susceptibility maps of the Loup, QC watershed in the years 2020, 2050, and 2080 under three climate change scenarios.

Comparing the area percentages of flood susceptibility classes in both watersheds revealed that the area percentages of the Loup watershed in the moderate, high, and very high flood susceptibility classes were relatively small compared to the area percentages of the same flood susceptibility classes in the Lower Nicola River watershed. This indicates flooding is more likely in the Lower Nicola River watershed than in the Loup watershed. Moreover, it can be concluded that the most flood-prone areas in the Loup watershed were in the southern and southeast, whereas in the Lower Nicola River watershed, the most flood-prone regions were in the center, northeast, and northwest. Furthermore, concerning the area percentages of flood susceptibility classes in both watersheds and their corresponding precipitation amounts, it can be concluded that despite the relatively high precipitation amounts in the Loup watershed compared to the Lower Nicola River watershed, significantly larger area of the Lower Nicola River watershed was susceptible to flooding. As a result, our findings indicated that climatological flood explanatory factors single-handedly are inadequate in identifying flood-prone regions and that topographic, hydrologic, environmental, and geologic factors must be considered and investigated in addition to them.

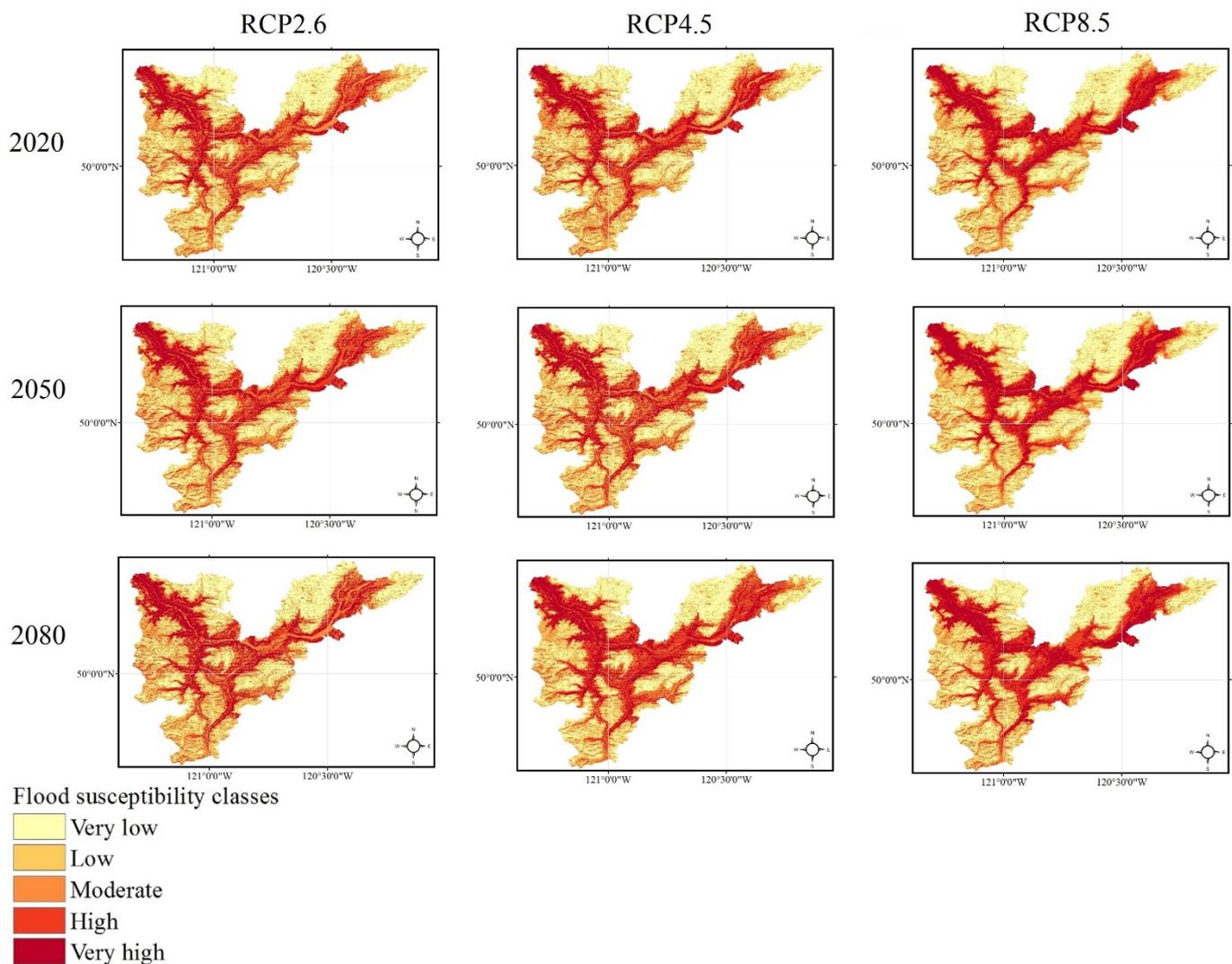


Figure 7. The flood susceptibility maps of the Lower Nicola River, BC watershed in the years 2020, 2050, and 2080 under three climate change scenarios.

The magnitude and direction of changes in the area percentages of flood susceptibility classes over three years (i.e., 2020, 2050, and 2080) were calculated and presented in Figure 10 (the Loup watershed) and Figure 11 (the Lower Nicola River). Following Figures 10 and 11, it can be concluded that despite some fluctuations in the area percentages of flood susceptibility classes, the most changes in the very low and low flood susceptibility classes were in the direction of decreasing and the majority of changes in the high and very high flood susceptibility classes were in the direction of increasing flood susceptibility in both watersheds. It is worth mentioning that the red color denoted the changes toward increasing the susceptibility (i.e., positive changes), and the blue color indicated the changes toward decreasing susceptibility (i.e., negative changes). Our findings indicated that climate change affects the flood susceptibility of watersheds, even though the changes in the flood susceptibility of watersheds are scant.

The variations in area percentages of flood susceptibility classes were plotted in Figure 12 to provide a better depiction of changes throughout time. The results in Figure 12 indicated that the changes in the Loup watershed mostly happened between the years 2050 and 2080 under the climate change scenario RCP2.6 and between the years 2020 and 2050 under the climate change scenarios RCP4.5 and RCP8.5. In the Lower Nicola River watershed, on the other hand, the most changes occurred between 2050 and 2080 in all climate change scenarios.

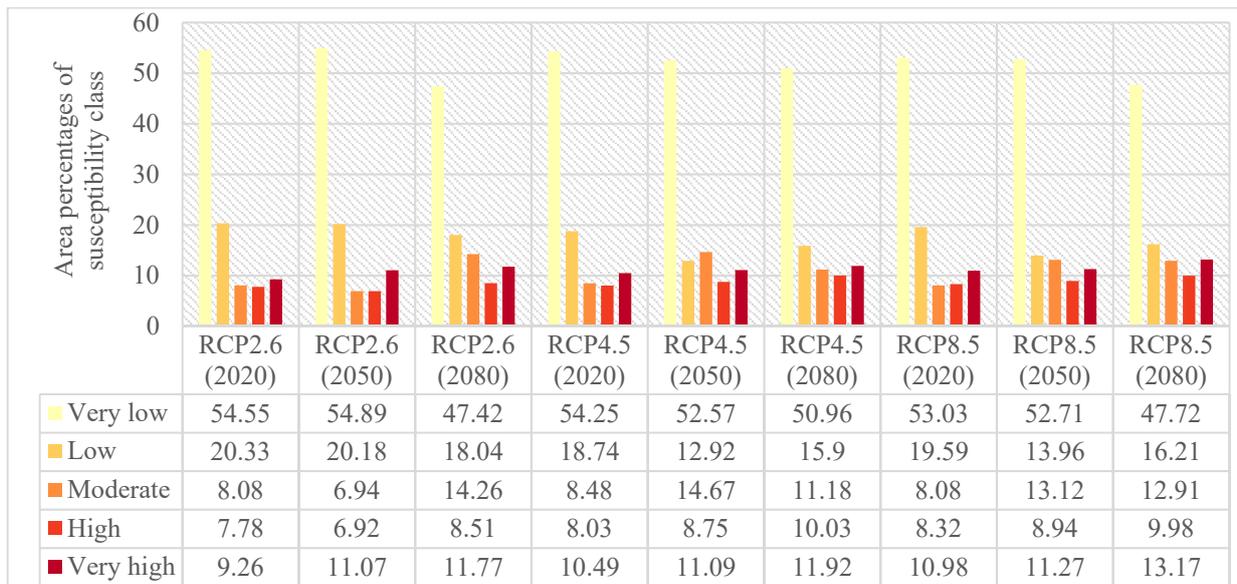


Figure 8. Area percentages of flood susceptibility classes under three climate change scenarios of RCP2.6, RCP4.5, and RCP8.5 in the Loup watershed.

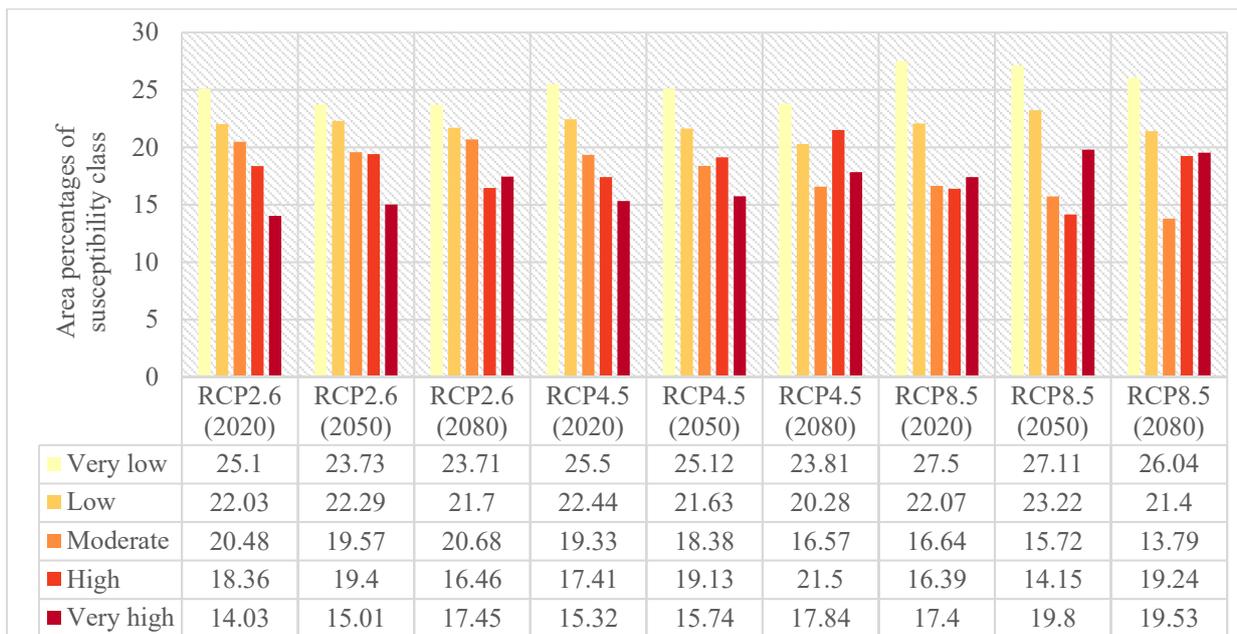


Figure 9. Area percentages of flood susceptibility classes under three climate change scenarios of RCP2.6, RCP4.5, and RCP8.5 in the Lower Nicola River watershed.

In addition to investigating changes over time, the changes in the area percentages of flood susceptibility classes were assessed depending on three climate change scenarios: optimistic (RCP2.6), business as usual (RCP4.5), and pessimistic (RCP8.5). The changes in the area percentages of flood susceptibility classes concerning the changes in the climate change scenarios in the Loup watershed and the Lower Nicola River watershed were calculated and illustrated in Figures 13 and 14, respectively. As with the changes over time, despite some fluctuations in area percentages of flood susceptibility classes, most of the changes in the high and very high flood susceptibility classes in both watersheds were toward rising as the scenarios changed from RCP2.6 to RCP4.5 and from RCP4.5 to RCP8.5. Following the findings in Figures 13 and 14, even though the changes in the area percentages of each flood susceptibility class seem trivial, the area percentages

of flood susceptibility classes in both watersheds were affected under various climate change scenarios.

Flood susceptibility classes	RCP2.6		RCP4.5		RCP8.5	
	2050 – 2020	2080 – 2050	2050 – 2020	2080 – 2050	2050 – 2020	2080 – 2050
Very low	0.34	-7.47	-1.68	-1.61	-0.32	-4.98
Low	-0.15	-2.14	-5.82	2.98	-5.63	2.25
Moderate	-1.14	7.32	6.19	-3.49	5.04	-0.21
High	-0.86	1.59	0.71	1.29	0.62	1.04
Very high	1.81	0.7	0.6	0.83	0.29	1.9

Figure 10. Changes (in %) in the area percentages of flood susceptibility classes based on the years in the Loup, QC watershed.

Flood susceptibility classes	RCP2.6		RCP4.5		RCP8.5	
	2050 – 2020	2080 – 2050	2050 – 2020	2080 – 2050	2050 – 2020	2080 – 2050
Very low	-1.37	-0.02	-0.38	-1.31	-0.39	-1.07
Low	0.26	-0.59	-0.81	-1.35	1.15	-1.82
Moderate	-0.91	1.11	-0.95	-1.81	-0.92	-1.93
High	1.04	-2.94	1.72	2.37	-2.24	5.09
Very high	0.98	2.44	0.42	2.1	2.4	-0.27

Figure 11. Changes (in %) in the area percentages of flood susceptibility classes based on the years in the Lower Nicola River, BC watershed.

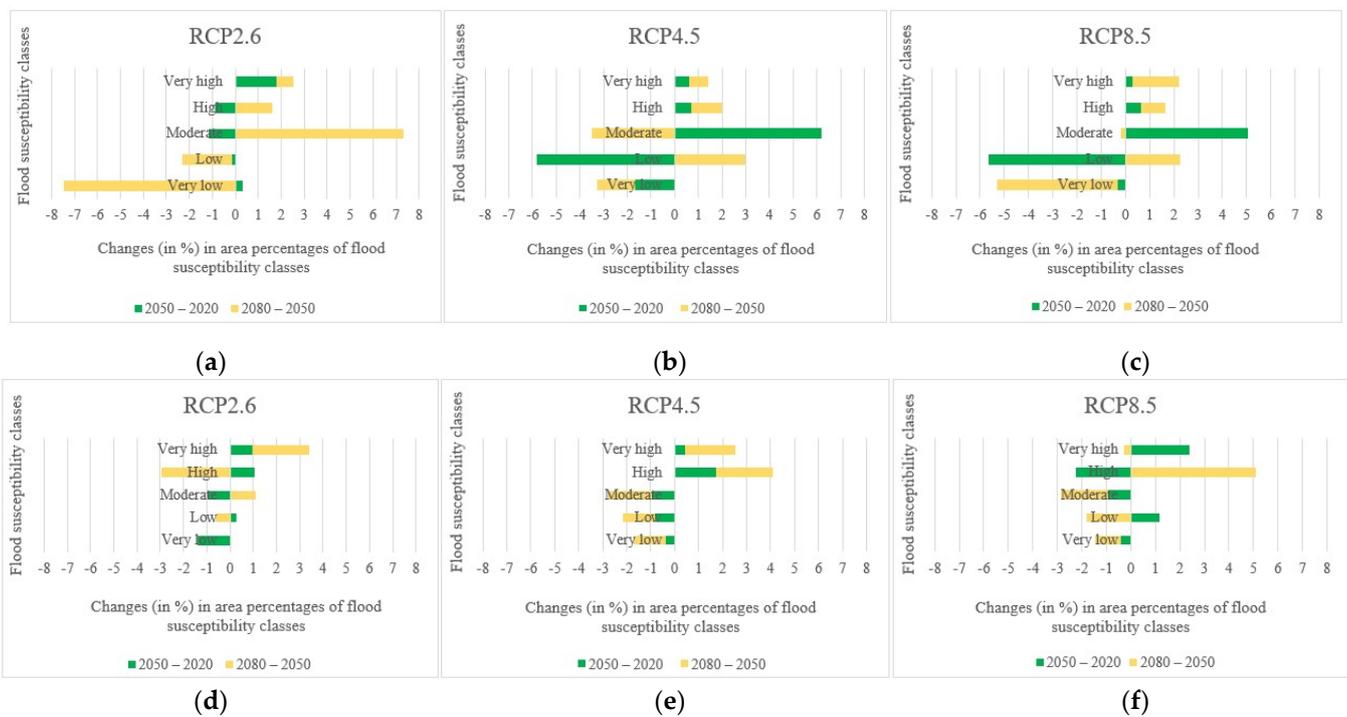


Figure 12. Changes (in %) in the area percentages of flood susceptibility classes over the years: (a) in the Loup watershed under scenario RCP2.6, (b) in the Loup watershed under scenario RCP4.5, (c) in the Loup watershed under scenario RCP8.5, (d) in the Lower Nicola River watershed under scenario RCP2.6, (e) in the Lower Nicola River watershed under scenario RCP4.5, (f) in the Lower Nicola River watershed under scenario RCP8.5.

Flood susceptibility classes	2020		2050		2080	
	RCP4.5 – RCP2.6	RCP8.5 – RCP4.5	RCP4.5 – RCP2.6	RCP8.5 – RCP4.5	RCP4.5 – RCP2.6	RCP8.5 – RCP4.5
Very low	-0.3	-1.22	-2.32	0.14	3.54	-3.23
Low	-1.59	0.85	-7.29	1.04	-2.14	0.31
Moderate	0.4	-0.4	7.73	-1.55	-3.08	1.73
High	0.26	0.28	1.83	0.19	1.53	-0.06
Very high	1.23	0.49	0.02	0.18	0.15	1.25

Figure 13. Changes (in %) in the area percentages of flood susceptibility classes based on the changes in the climate change scenarios in the Loup watershed.

Flood susceptibility classes	2020		2050		2080	
	RCP4.5 – RCP2.6	RCP8.5 – RCP4.5	RCP4.5 – RCP2.6	RCP8.5 – RCP4.5	RCP4.5 – RCP2.6	RCP8.5 – RCP4.5
Very low	0.4	2	1.39	1.99	0.1	2.23
Low	0.41	-0.37	-0.66	1.59	-1.42	1.12
Moderate	-1.15	-2.69	-1.19	-2.66	-4.11	-2.78
High	-0.95	-1.02	-0.27	-4.98	5.04	-2.26
Very high	1.29	2.08	0.73	4.06	0.39	1.69

Figure 14. Changes (in %) in the area percentages of flood susceptibility classes based on the changes in the climate change scenarios in the Lower Nicola River watershed.

To have a better representation of variations regarding the changes in the climate change scenarios, the changes in area percentages of flood susceptibility classes were plotted in Figure 15. The changes in the Loup watershed occurred mainly between the scenarios RCP2.6 and RCP4.5 in the years 2050 and 2080 and almost the same in the year 2020. The most changes in the Lower Nicola River watershed, on the other hand, occurred between RCP2.6 and RCP4.5 in 2080 and between RCP4.5 and RCP8.5 in the years 2020 and 2050.

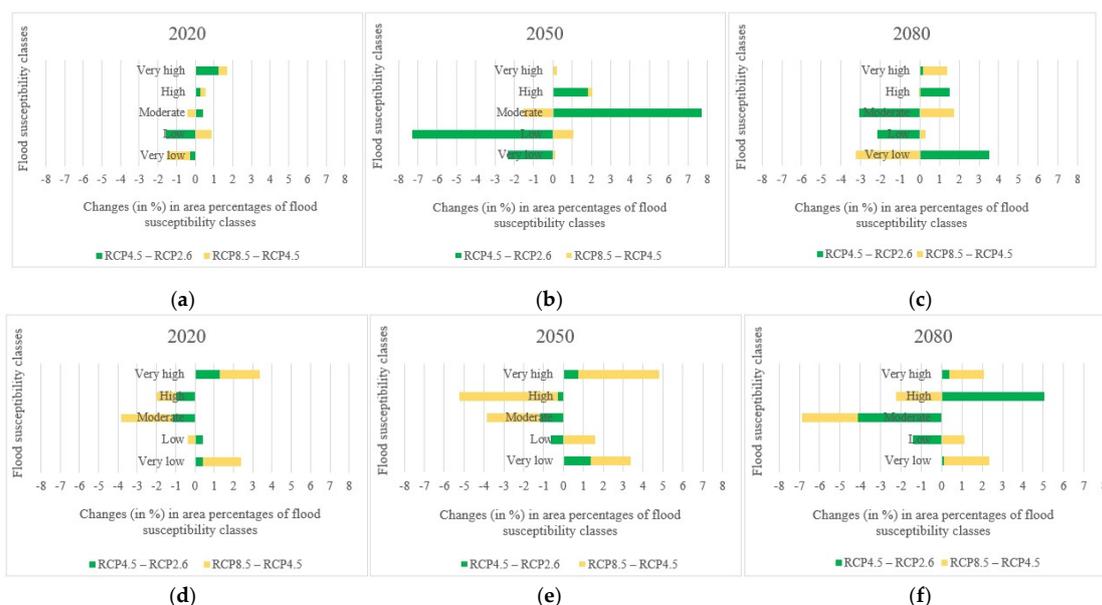


Figure 15. Changes (in %) in the area percentages of flood susceptibility classes regarding the changes in the climate change scenarios: (a) in the Loup watershed in the year 2020, (b) in the Loup watershed in the year 2050, (c) in the Loup watershed in the year 2080, (d) in the Lower Nicola River watershed in the year 2020, (e) in the Lower Nicola River watershed in the year 2050, (f) in the Lower Nicola River watershed in the year 2080.

4. Discussion

Floods are one of the most hazardous natural disasters that usually result in considerable loss of life and significant property damage [94]. Changing climates induced by global warming have affected the circulation patterns of the atmospheric and ocean currents and, consequently the spatial and temporal patterns of precipitation. Accordingly, changes in flood susceptibility are linked to climate changes. As the ability of the atmosphere to hold moisture increases due to global warming, more frequent and heavier precipitation events may occur, raising the peril of floods [95]. To the study by Houghton et al., (2001) [96], precipitation has increased by 0.5 to 1% every decade in much of the Northern Hemisphere's mid-to high latitudes over the last 100 years. As a result, evaluating flood-prone zones under future precipitation circumstances is crucial for gaining a thorough knowledge of future flood susceptibility patterns.

Having the capacity to predict the spatial patterns of flooding in watersheds and assess their flood susceptibility could improve the managers' abilities to reduce flood losses. Accordingly, the primary objective of this study was to develop various ML models to identify and predict current and future flood susceptible areas while considering the spatial and temporal impacts of climate change on floods. From a spatial perspective, there are three crucial elements in efficiently mapping flood susceptibility: (1) selection of appropriate flood explanatory factors, (2) spatial resolution of the flood explanatory factors, and (3) the accuracy and efficiency of data layer integration models [97]. Even though there is no conventional technique for selecting the factors that would best predict future floods, we chose various sets of factors regarding the literature review [5,54–56]. A variety of meteorological, hydrological, and geospatial flood explanatory factors were collected and prepared using RS and GIS techniques. Given the availability of data resources, various flood explanatory factors with a spatial resolution of 30 m were collected and prepared for both watersheds. Moreover, as the third substantial element in efficiently mapping flood susceptibility, four various models of MLP-NN, RF, NB, and GBM, as promising ML models, were employed to provide the flood susceptibility map.

Following the literature review, numerous ML models were formulated and developed to map flood susceptibility [27,57,60,98–103]. In this study, four potential ML models of MLP-NN, NB, RF, and GBM, were used to assess the flood susceptibility in two different watersheds in Canada. Moreover, three accuracy criteria were used to evaluate and compare the accuracy of the models. Regarding the results of three accuracy metrics, the RF model outperformed the rest of the employed models, which was consistent with the findings of other researchers who have described the RF model as a more accurate model [5,28].

After evaluating the efficiency and accuracy of the employed ML models, the model with the best accuracy was chosen and run using the precipitation data in the years 2020, 2050, and 2080 under three climate change scenarios: optimistic (RCP2.6), business as usual (RCP4.5), and pessimistic (RCP8.5). Accordingly, for each year as well as under each scenario, a flood susceptibility map was provided for both watersheds. The results of this study indicated that the flood susceptibility of both the Loup and Lower Nicola River watersheds worsens over time. Our findings were consistent with the study by Janizadeh et al., (2021) [34], demonstrating that flood susceptibility worsens over time.

Although the effects of climate change are still debatable, the impacts of climatic variability require more investigation. While precipitation is recognized as the most significant climatic factor for flooding in some places [104] and the runoff factor in flood events [105], regarding many earlier studies, the most influential factors for flood events include elevation [26,52], slope [52], distance from rivers [26,52,106,107], drainage density, and land cover/land use [52,97]. Comparing the area percentages of flood susceptibility classes in both watersheds given their corresponding precipitation maps demonstrated that even though the Loup watershed receives significantly more precipitation than the Lower Nicola River watershed, the area percentages of moderate, high, and very high flood susceptibility classes in the Loup watershed were much trivial compared to the area percentages of the same classes in the Lower Nicola River watershed. Consequently, our

findings indicated that the precipitation factor single-handedly is inadequate in identifying flood-prone regions and that topographic, hydrologic, environmental, and geologic factors must be considered and investigated in addition.

Datasets in ML models are divided into training and test datasets, where the training datasets are used for training and the performance of the models are evaluated based on the test dataset. The trained model may have poor generalization in space and time if the distributions between the training and test sets change (i.e., distribution shifts) or if there are inherent sample dependencies. The in-built assumption of independent and identical distribution (I.I.D) is automatically made when data is divided into training-test sets. Accordingly, the assumption of I.I.D is central to almost all ML algorithms. However, spatial autocorrelation and spatial heterogeneity (i.e., two intrinsic properties of spatial data) violate I.I.D assumption [108]. Due to the increased resemblance between neighboring data samples, spatial autocorrelation violates the independence principle. On the other hand, because the data generating processes frequently change with respect to space, spatial heterogeneity violates the identical distribution assumption [109]. To overcome spatial autocorrelation and spatial heterogeneity issues, the process of choosing flood and non-flood points was done iteratively and the set of points which the p -value of Moran's I for all factors was obtained extremely close to zero and less than the threshold (i.e., 0.05) was chosen to create the flood inventory map.

Although ML models have produced encouraging results in flood forecasting, essential uncertainties exist in their prediction results. Regardless of the ML model employed, numerous error sources influence the prediction results. The errors come from at least three significant sources. First, not only ML models but also all computational models are amplifications and approximations of a complex physical system due to mathematical and modeling restrictions. Second, ML models learn to extract patterns from the input data; therefore, training models with insufficient quality or scarce data will result in uncertainty in model predictions. Third, there will be more uncertainty since we cannot predict how the properties of real system could alter in the future. When extrapolating from the past to the future as is customary, uncertainty arises from both the imprecise depiction of the past and the degree to which the future will resemble the past [110].

Strengths and Limitations

The strengths of this study include: (i) considering 16 various meteorological, topographic, hydrologic, environmental, and geologic flood conditioning factors in the flood susceptibility assessment, (ii) using four various ML models, (iii) considering two alternative watersheds with diverse meteorological, topographic, hydrologic, environmental, and geologic conditions and comparing the flood susceptibility of them with each other, (iv) evaluating and comparing the performance of ML methods using three various accuracy criteria, and (v) evaluating the flood susceptibility in three years and under three climate change scenarios: optimistic (RCP2.6), business as usual (RCP4.5), and pessimistic (RCP8.5).

Notwithstanding these strengths, this study had few limitations in terms of future precipitation data and the preparation process of them. Future precipitation data have uncertainty and include some level of errors. On the other hand, the interpolation process also imports some level of errors in the data. Another limitation of this study was considering only precipitation as the dynamic flood explanatory factor following the primary goal of the study, while other factors such as land cover also vary over time. The creation of the flood inventory map for a single flood occurrence was another research constraint. A single flood event was considered in both study areas due to a paucity of data on previous flood occurrences in the study areas.

5. Conclusions

Flooding is one of the most devastating natural disasters that can result in death, injury, property destruction, loss of livelihoods and services, social and economic upheaval, and

environmental havoc. Climate change has the potential to exacerbate the runoff rates and patterns and the hydrological cycle, resulting in more intense precipitation and increases in flood intensity, frequency, and severity.

In this study, four various ML models, including MLP-NN, NB, RF, and GBM, were used to provide the current and future flood susceptibility maps in two different watersheds in Canada, one in Quebec province and the other in British Columbia province. Moreover, three RCP2.6, RCP4.5, and RCP8.5 climate change scenarios were examined to address the implications of climate change.

Regarding the accuracy metrics, the RF model had the highest accuracy and was chosen as the best ML model to provide the flood susceptibility maps. Regarding the provided flood susceptibility maps, flooding is more likely in the Lower Nicola River watershed than in the Loup watershed. The most flood-prone locations in the Loup watershed were in the southern and southeast, while the most flood-prone areas in the Lower Nicola River watershed were in the center, northeast, and northwest. The results of this study indicated that the flood susceptibility of both the Loup and Lower Nicola River watersheds worsens over time.

The contribution of this study lies in the identification of flood-prone areas over time and under various climate change scenarios in two different watersheds. The spatial forecasts provided by this research aim to assist disaster management agencies in making critical decisions and contributions to mitigate the damages caused by floods, and to better inform local communities and researchers on the important role and influence of climate change on flood susceptibility.

Author Contributions: Conceptualization, N.M.G. and L.P.; methodology, N.M.G. and L.P.; software, N.M.G.; validation, N.M.G. and L.P.; formal analysis, N.M.G. and L.P.; investigation, N.M.G. and L.P.; resources, N.M.G. and L.P.; data curation, N.M.G. and L.P.; writing—original draft preparation, N.M.G.; writing—review and editing, N.M.G. and L.P.; visualization, N.M.G.; supervision, L.P.; project administration, L.P.; funding acquisition, L.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada through the Discovery Grant number RGPIN/05396-2016 awarded to L.P. The funders had no role in the study design, data collection, and analysis, decision to publish, or preparation of the manuscript. There was no additional external funding received for this study.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data utilized in this research are openly accessible (at the time of writing this manuscript).

Acknowledgments: We appreciate the constructive comments received from the four reviewers on an earlier version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Data Preparation Flowchart

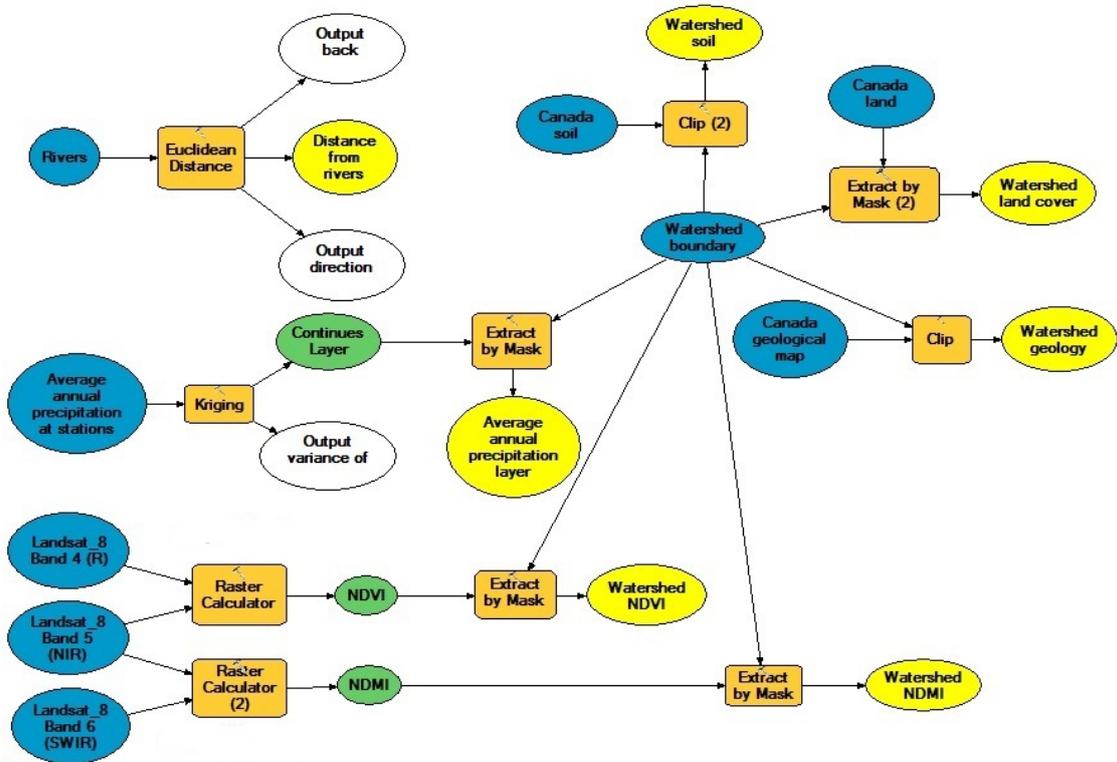
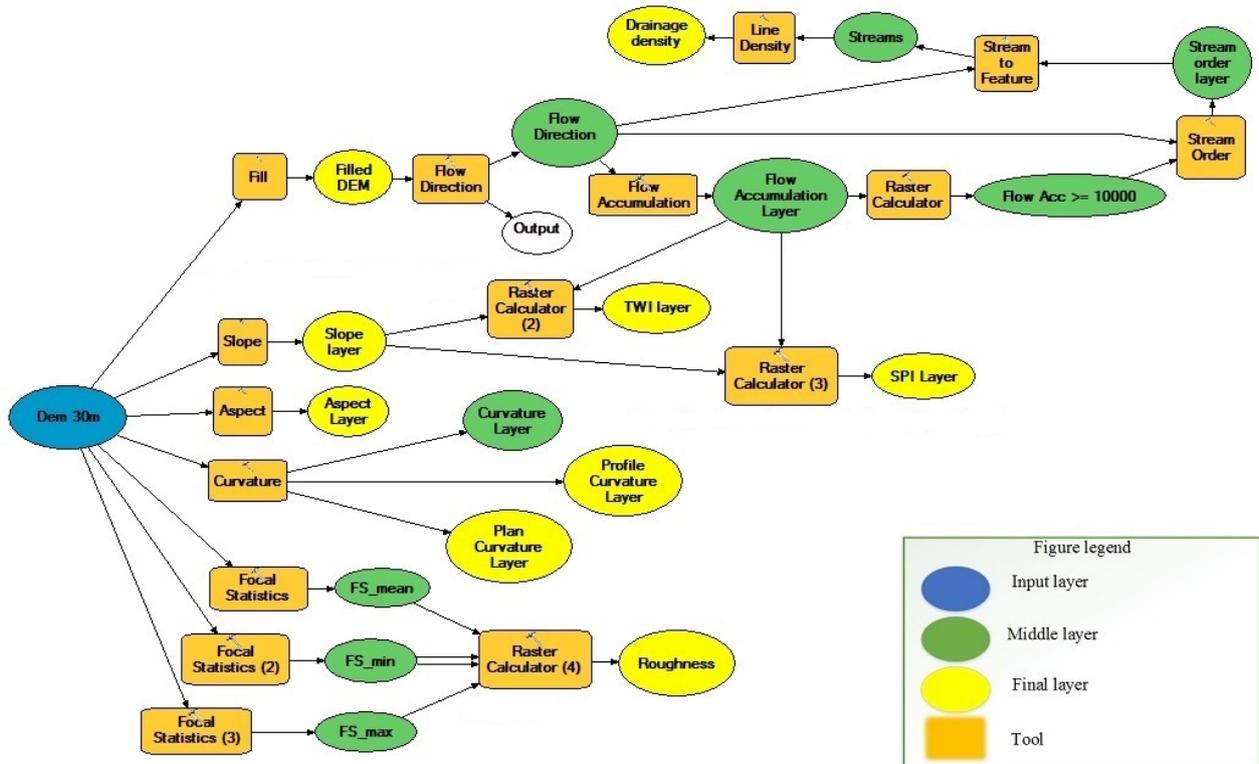


Figure A1. The overall data preparation flowchart using the Model Builder extension of ArcGIS.

Appendix B. Flood Explanatory Factors

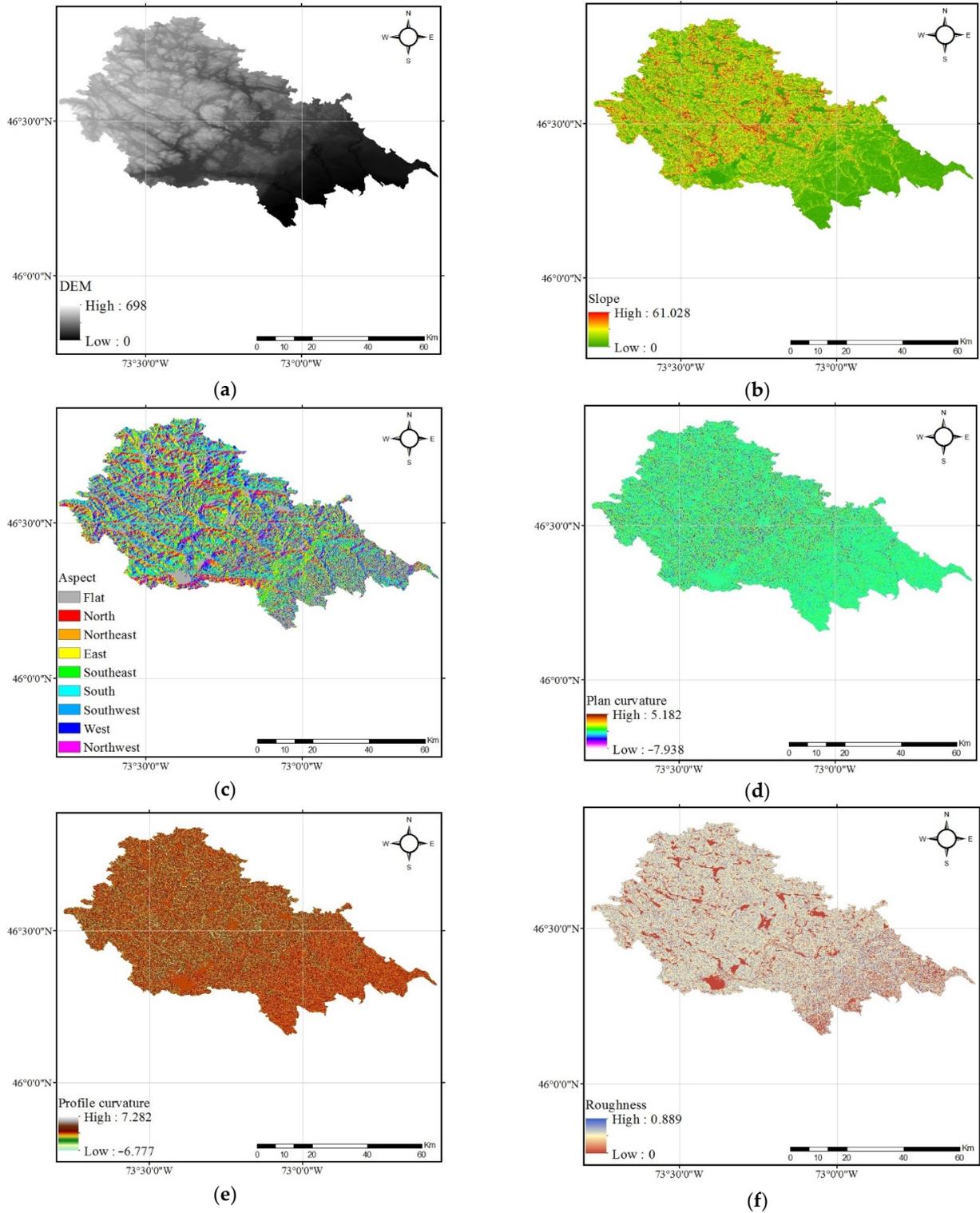


Figure A2. Cont.

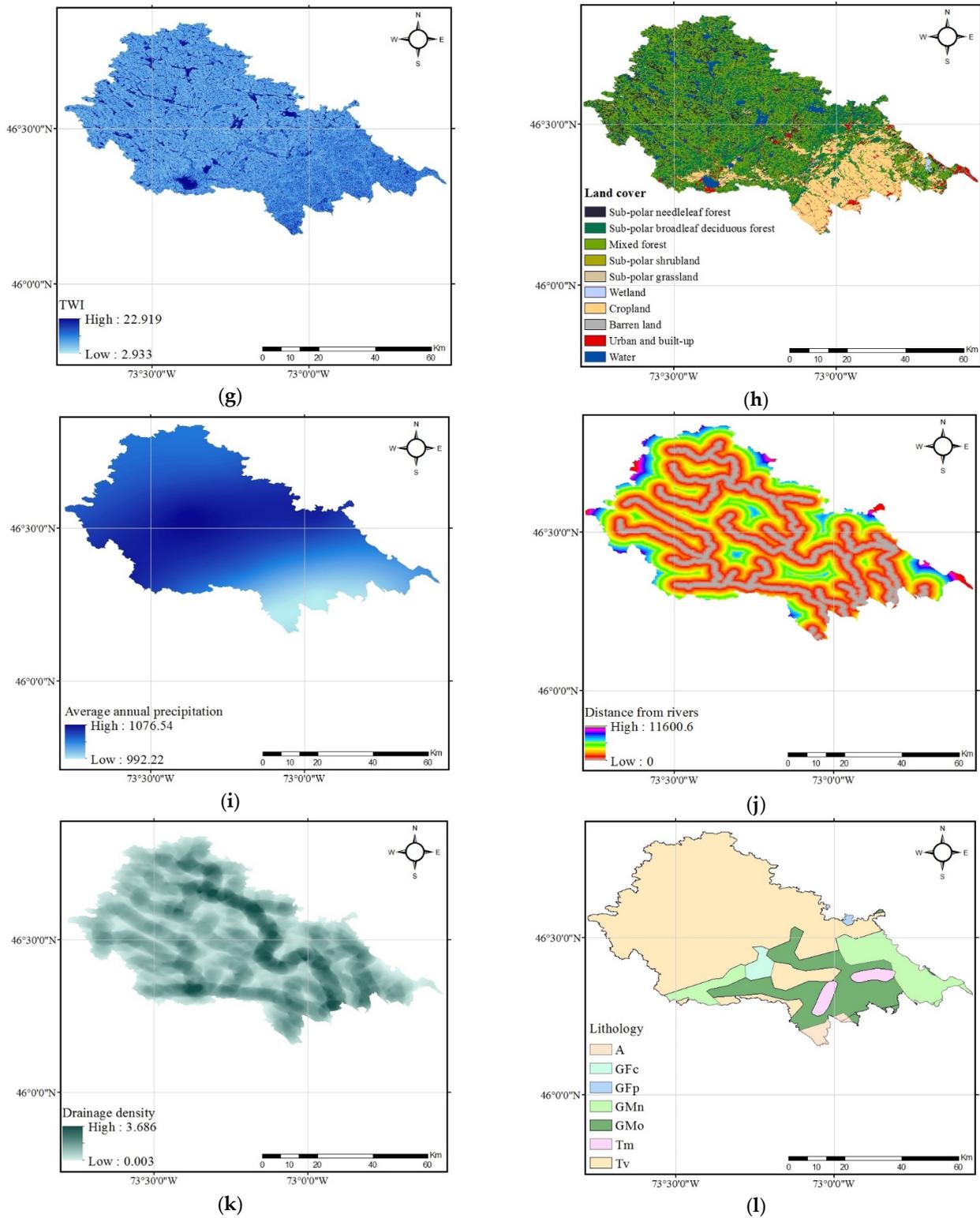


Figure A2. Cont.

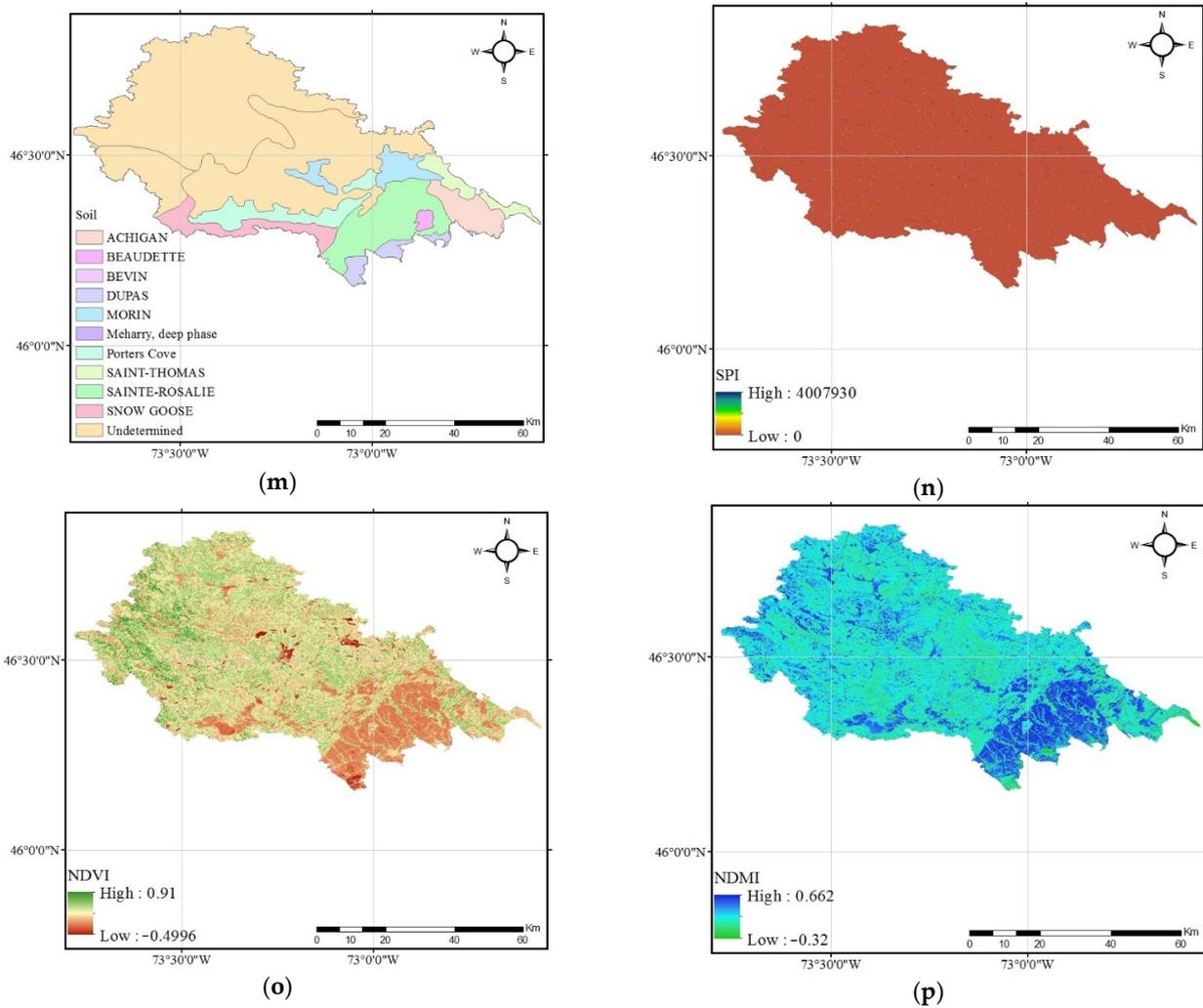


Figure A2. Thematic maps of flood explanatory factors in the Loup watershed: (a) DEM, (b) slope, (c) aspect, (d) plan curvature, (e) profile curvature, (f) roughness, (g) TWI, (h) land cover, (i) precipitation, (j) distance from rivers, (k) drainage density, (l) lithology, (m) soil, (n) SPI, (o) NDVI, (p) NDMI.

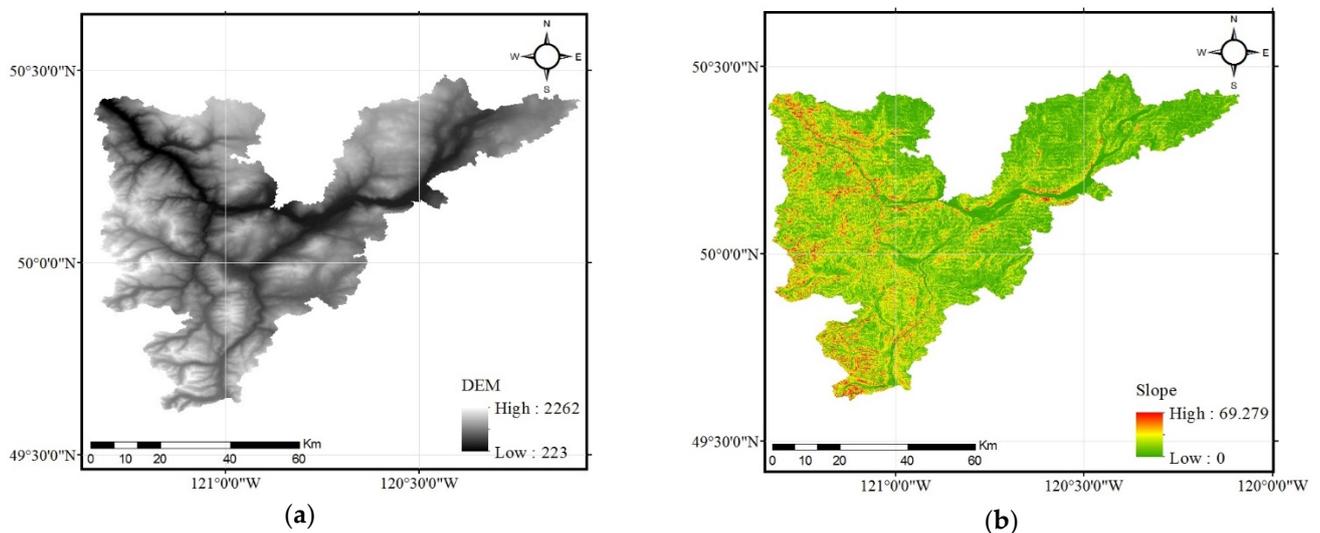


Figure A3. Cont.

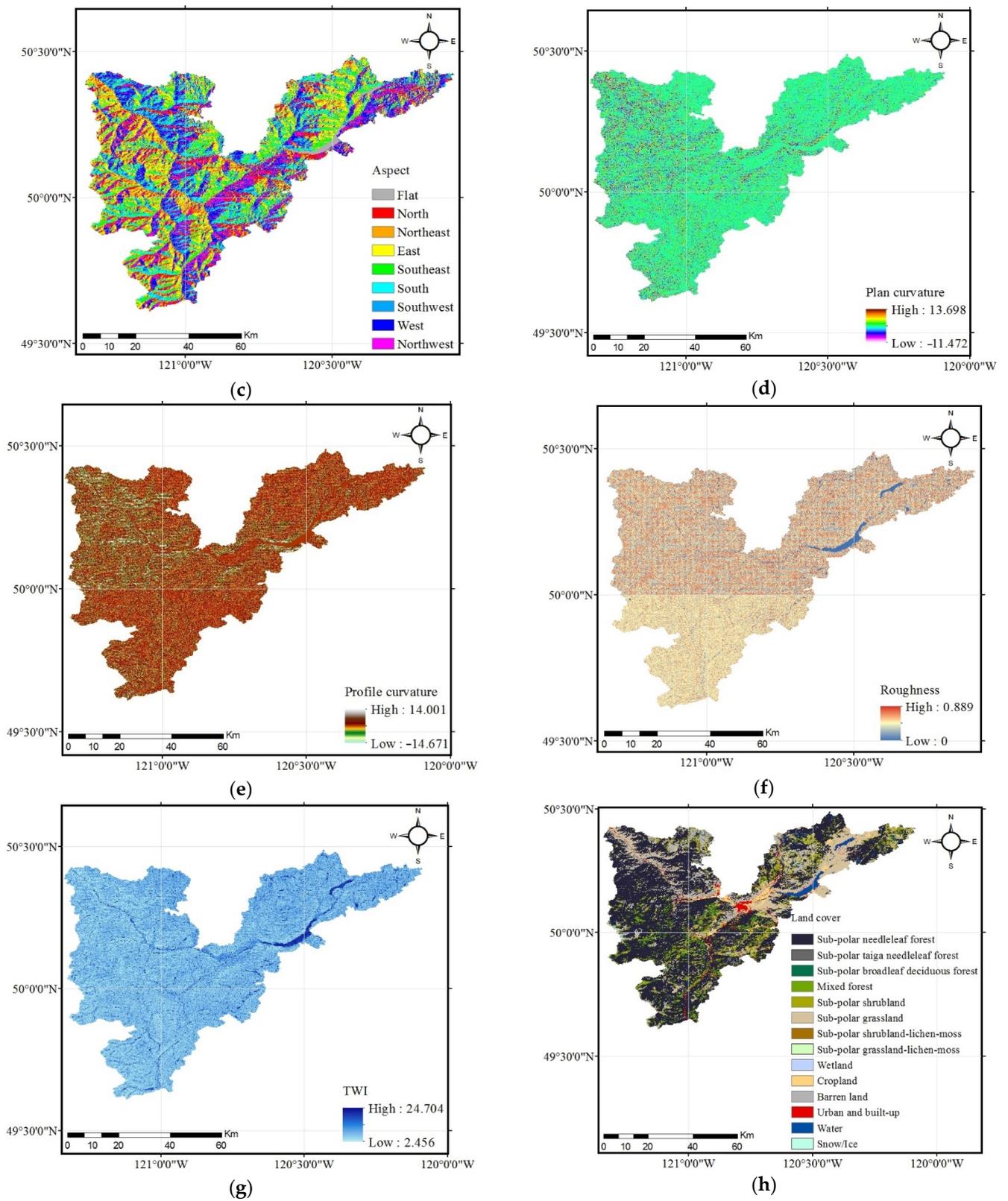
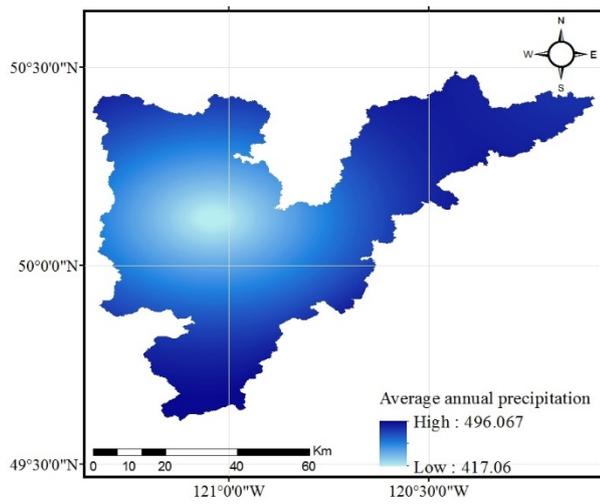
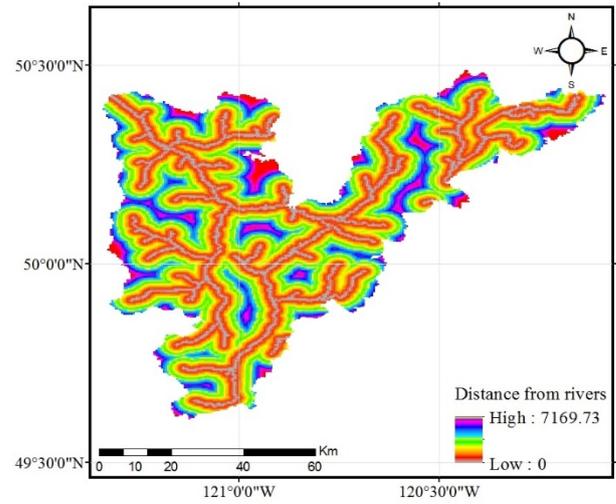


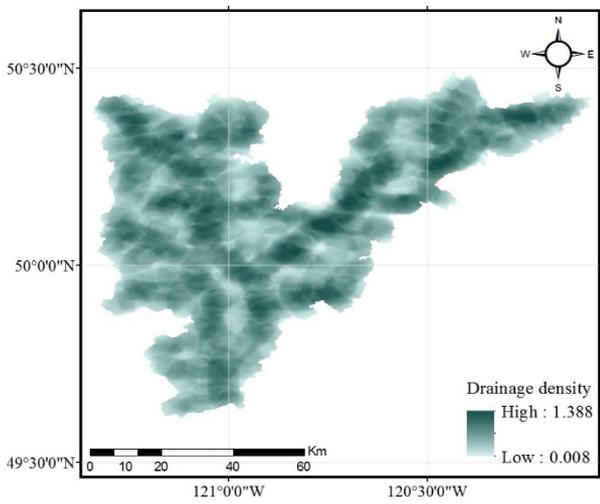
Figure A3. Cont.



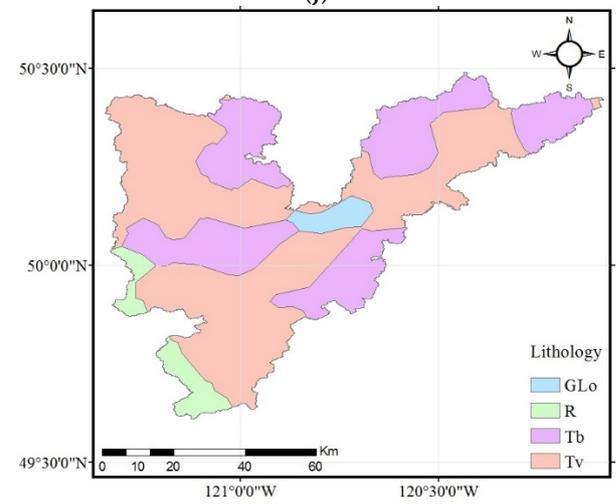
(i)



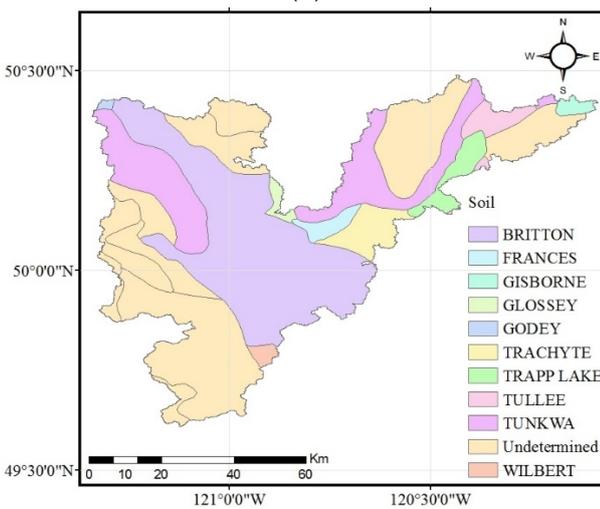
(j)



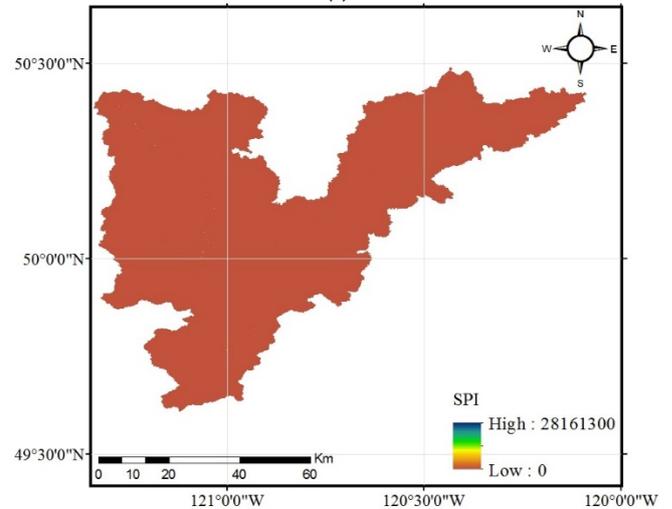
(k)



(l)



(m)



(n)

Figure A3. Cont.

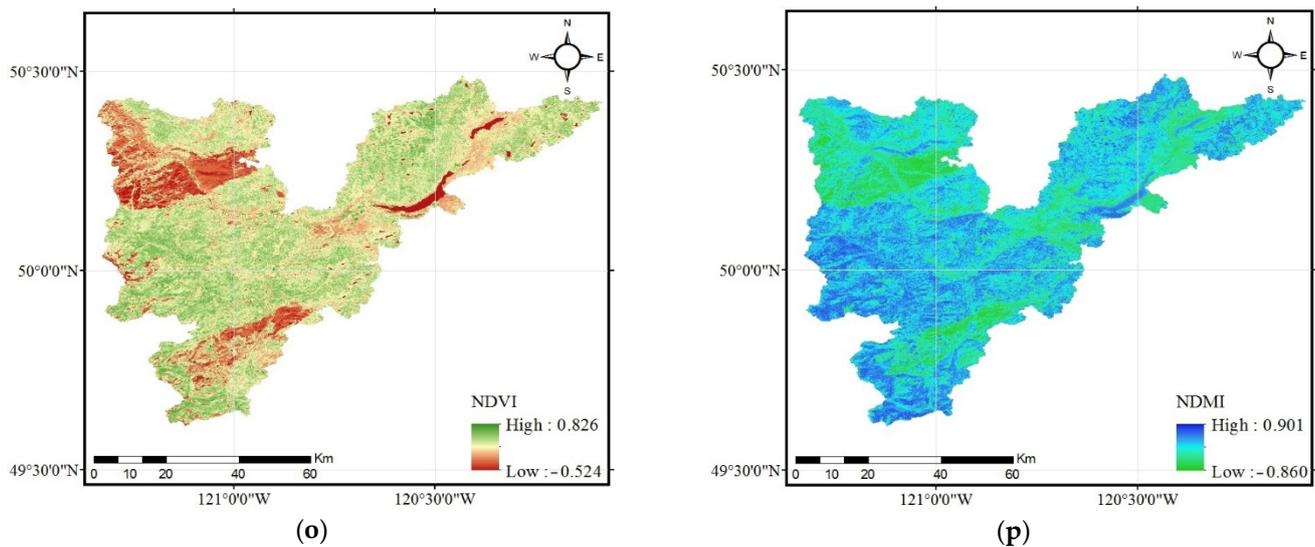


Figure A3. Thematic maps of flood explanatory factors in the Lower Nicola River watershed: (a) DEM, (b) slope, (c) aspect, (d) plan curvature, (e) profile curvature, (f) roughness, (g) TWI, (h) land cover, (i) precipitation, (j) distance from rivers, (k) drainage density, (l) lithology, (m) soil, (n) SPI, (o) NDVI, (p) NDMI.

References

- Centre for Research on the Epidemiology of Disasters CRED; UN Office for Disaster Risk Reduction. The Human Cost of Disasters: An Overview of the Last 20 Years (2000–2019). 2021. Available online: <https://reliefweb.int/sites/reliefweb.int/files/resources/Human%20Cost%20of%20Disasters%202000-2019%20Report%20-%20UN%20Office%20for%20Disaster%20Risk%20Reduction.pdf> (accessed on 15 January 2022).
- Sandink, D.; Kovacs, P.; Oulahan, G.; McGillivray, G. *Making Flood Insurable for Canadian Homeowners: A Discussion Paper*; Institute for Catastrophic Loss Reduction & Swiss Reinsurance Company Ltd.: Toronto, ON, Canada, 2010.
- Canada, P.S. Canadian Disaster Database. Public Safety Canada, Ottawa. Available online: <https://cdd.publicsafety.gc.ca/srchpg-eng.aspx?cultureCode=en-Ca&provinces=1&eventTypes=%27FL%27&eventStartDate=%2720000101%27%2c%2720201231%27&normalizedCostYear=1> (accessed on 15 January 2022).
- Kundzewicz, Z.W.; Kanae, S.; Seneviratne, S.I.; Handmer, J.; Nicholls, N.; Peduzzi, P.; Mechler, R.; Bouwer, L.M.; Arnell, N.; Mach, K. Flood risk and climate change: Global and regional perspectives. *Hydrol. Sci. J.* **2014**, *59*, 1–28. [[CrossRef](#)]
- Avand, M.; Moradi, H. Using machine learning models, remote sensing, and GIS to investigate the effects of changing climates and land uses on flood probability. *J. Hydrol.* **2021**, *595*, 125663. [[CrossRef](#)]
- Janizadeh, S.; Avand, M.; Jaafari, A.; Phong, T.V.; Bayat, M.; Ahmadisharaf, E.; Prakash, I.; Pham, B.T.; Lee, S. Prediction success of machine learning methods for flash flood susceptibility mapping in the Tafresh watershed, Iran. *Sustainability* **2019**, *11*, 5426. [[CrossRef](#)]
- Das, S. Flood susceptibility mapping of the Western Ghat coastal belt using multi-source geospatial data and analytical hierarchy process (AHP). *Remote Sens. Appl. Soc. Environ.* **2020**, *20*, 100379. [[CrossRef](#)]
- Yariyan, P.; Avand, M.; Abbaspour, R.A.; Torabi Haghighi, A.; Costache, R.; Ghorbanzadeh, O.; Janizadeh, S.; Blaschke, T. Flood susceptibility mapping using an improved analytic network process with statistical models. *Geomat. Nat. Hazards Risk* **2020**, *11*, 2282–2314. [[CrossRef](#)]
- Nachappa, T.G.; Piralilou, S.T.; Gholamnia, K.; Ghorbanzadeh, O.; Rahmati, O.; Blaschke, T. Flood susceptibility mapping with machine learning, multi-criteria decision analysis and ensemble using Dempster Shafer Theory. *J. Hydrol.* **2020**, *590*, 125275. [[CrossRef](#)]
- Miles, R.E.; Snow, C.C.; Fit, F. *The Hall of Fame; How Companies Succeed or Fail*; The Free Press: New York, NY, USA, 1994.
- Cea, L.; Garrido, M.; Puertas, J. Experimental validation of two-dimensional depth-averaged models for forecasting rainfall–runoff from precipitation data in urban areas. *J. Hydrol.* **2010**, *382*, 88–102. [[CrossRef](#)]
- Xia, X.; Liang, Q.; Ming, X.; Hou, J. An efficient and stable hydrodynamic model with novel source term discretization schemes for overland flow and flood simulations. *Water Resour. Res.* **2017**, *53*, 3730–3759. [[CrossRef](#)]
- Nayak, P.; Sudheer, K.; Rangan, D.; Ramasastri, K. Short-term flood forecasting with a neurofuzzy model. *Water Resour. Res.* **2005**, *41*. [[CrossRef](#)]
- Kim, B.; Sanders, B.F.; Famiglietti, J.S.; Guinot, V. Urban flood modeling with porous shallow-water equations: A case study of model errors in the presence of anisotropic porosity. *J. Hydrol.* **2015**, *523*, 680–692. [[CrossRef](#)]

15. Van den Honert, R.C.; McAneney, J. The 2011 Brisbane floods: Causes, impacts and implications. *Water* **2011**, *3*, 1149–1173. [[CrossRef](#)]
16. Mosavi, A.; Ozturk, P.; Chau, K.-w. Flood prediction using machine learning models: Literature review. *Water* **2018**, *10*, 1536. [[CrossRef](#)]
17. Costache, R.; Pham, Q.B.; Arabameri, A.; Diaconu, D.C.; Costache, I.; Crăciun, A.; Ciobotaru, N.; Pandey, M.; Arora, A.; Ali, S.A. Flash-flood propagation susceptibility estimation using weights of evidence and their novel ensembles with multicriteria decision making and machine learning. *Geocarto Int.* **2021**, 1–33. [[CrossRef](#)]
18. Costache, R.; Tin, T.T.; Arabameri, A.; Crăciun, A.; Ajin, R.; Costache, I.; Islam, A.R.M.T.; Abba, S.; Sahana, M.; Avand, M. Flash-flood hazard using deep learning based on H2O R package and fuzzy-multicriteria decision-making analysis. *J. Hydrol.* **2022**, *609*, 127747. [[CrossRef](#)]
19. Abbot, J.; Marohasy, J. Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks. *Atmos. Res.* **2014**, *138*, 166–178. [[CrossRef](#)]
20. Mekanik, F.; Imteaz, M.; Gato-Trinidad, S.; Elmahdi, A. Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. *J. Hydrol.* **2013**, *503*, 11–21. [[CrossRef](#)]
21. Sayers, W.; Savić, D.; Kapelan, Z.; Kellagher, R. Artificial intelligence techniques for flood risk management in urban environments. *Procedia Eng.* **2014**, *70*, 1505–1512. [[CrossRef](#)]
22. Solomatine, D.; See, L.M.; Abrahart, R. Data-driven modelling: Concepts, approaches and experiences. *Pract. Hydroinform.* **2009**, 17–30. [[CrossRef](#)]
23. Khosravi, K.; Pham, B.T.; Chapi, K.; Shirzadi, A.; Shahabi, H.; Revhaug, I.; Prakash, I.; Bui, D.T. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. *Sci. Total Environ.* **2018**, *627*, 744–755. [[CrossRef](#)]
24. Yu, P.-S.; Yang, T.-C.; Chen, S.-Y.; Kuo, C.-M.; Tseng, H.-W. Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting. *J. Hydrol.* **2017**, *552*, 92–104. [[CrossRef](#)]
25. Chen, J.; Li, Q.; Wang, H.; Deng, M. A machine learning ensemble approach based on random forest and radial basis function neural network for risk evaluation of regional flood disaster: A case study of the Yangtze River Delta, China. *Int. J. Environ. Res. Public Health* **2020**, *17*, 49. [[CrossRef](#)] [[PubMed](#)]
26. Janizadeh, S.; Vafakhah, M.; Kapelan, Z.; Dinan, N.M. Novel Bayesian Additive Regression Tree Methodology for Flood Susceptibility Modeling. *Water Resour. Manag.* **2021**, *35*, 4621–4646. [[CrossRef](#)]
27. Tang, X.; Li, J.; Liu, M.; Liu, W.; Hong, H. Flood susceptibility assessment based on a novel random Naïve Bayes method: A comparison between different factor discretization methods. *Catena* **2020**, *190*, 104536. [[CrossRef](#)]
28. Chen, W.; Li, Y.; Xue, W.; Shahabi, H.; Li, S.; Hong, H.; Wang, X.; Bian, H.; Zhang, S.; Pradhan, B. Modeling flood susceptibility using data-driven approaches of naïve bayes tree, alternating decision tree, and random forest methods. *Sci. Total Environ.* **2020**, *701*, 134979. [[CrossRef](#)]
29. Popa, M.C.; Diaconu, D.C. Flood and Flash Flood hazard mapping using the Frequency Ratio, Multilayer Perceptron and their hybrid ensemble. *Proc. Multidiscip. Digit. Publ. Inst. Proc.* **2019**, *48*, 6.
30. Wang, Y.; Fang, Z.; Hong, H.; Costache, R.; Tang, X. Flood susceptibility mapping by integrating frequency ratio and index of entropy with multilayer perceptron and classification and regression tree. *J. Environ. Manag.* **2021**, *289*, 112449. [[CrossRef](#)]
31. Vafakhah, M.; Mohammad Hasani Looor, S.; Pourghasemi, H.; Katebikord, A. Comparing performance of random forest and adaptive neuro-fuzzy inference system data mining models for flood susceptibility mapping. *Arab. J. Geosci.* **2020**, *13*, 417. [[CrossRef](#)]
32. Sahoo, A.; Samantaray, S.; Bankuru, S.; Ghose, D.K. Prediction of flood using adaptive neuro-fuzzy inference systems: A case study. In *Smart Intelligent Computing and Applications*; Springer: Singapore, 2020; pp. 733–739.
33. Sahana, M.; Rehman, S.; Sajjad, H.; Hong, H. Exploring effectiveness of frequency ratio and support vector machine models in storm surge flood susceptibility assessment: A study of Sundarban Biosphere Reserve, India. *Catena* **2020**, *189*, 104450. [[CrossRef](#)]
34. Janizadeh, S.; Pal, S.C.; Saha, A.; Chowdhuri, I.; Ahmadi, K.; Mirzaei, S.; Mosavi, A.H.; Tiefenbacher, J.P. Mapping the spatial and temporal variability of flood hazard affected by climate and land-use changes in the future. *J. Environ. Manag.* **2021**, *298*, 113551. [[CrossRef](#)]
35. Lohani, A.K.; Goel, N.; Bhatia, K. Improving real time flood forecasting using fuzzy inference system. *J. Hydrol.* **2014**, *509*, 25–41. [[CrossRef](#)]
36. Costache, R.; Arabameri, A.; Blaschke, T.; Pham, Q.B.; Pham, B.T.; Pandey, M.; Arora, A.; Linh, N.T.T.; Costache, I. Flash-flood potential mapping using deep learning, alternating decision trees and data provided by remote sensing sensors. *Sensors* **2021**, *21*, 280. [[CrossRef](#)] [[PubMed](#)]
37. Ngo, P.-T.T.; Hoang, N.-D.; Pradhan, B.; Nguyen, Q.K.; Tran, X.T.; Nguyen, Q.M.; Nguyen, V.N.; Samui, P.; Tien Bui, D. A novel hybrid swarm optimized multilayer neural network for spatial prediction of flash floods in tropical areas using sentinel-1 SAR imagery and geospatial data. *Sensors* **2018**, *18*, 3704. [[CrossRef](#)] [[PubMed](#)]
38. Talukdar, S.; Ghose, B.; Salam, R.; Mahato, S.; Pham, Q.B.; Linh, N.T.T.; Costache, R.; Avand, M. Flood susceptibility modeling in Teesta River basin, Bangladesh using novel ensembles of bagging algorithms. *Stoch. Environ. Res. Risk Assess.* **2020**, *34*, 2277–2300. [[CrossRef](#)]

39. Rahman, M.; Ningsheng, C.; Islam, M.M.; Dewan, A.; Iqbal, J.; Washakh, R.M.A.; Shufeng, T. Flood susceptibility assessment in Bangladesh using machine learning and multi-criteria decision analysis. *Earth Syst. Environ.* **2019**, *3*, 585–601. [[CrossRef](#)]
40. Jain, S.K.; Mani, P.; Jain, S.K.; Prakash, P.; Singh, V.P.; Tullios, D.; Kumar, S.; Agarwal, S.; Dimri, A. A Brief review of flood forecasting techniques and their applications. *Int. J. River Basin Manag.* **2018**, *16*, 329–344. [[CrossRef](#)]
41. Zaharia, L.; Costache, R.; Prăvălie, R.; Minea, G. Assessment and mapping of flood potential in the Slănic catchment in Romania. *J. Earth Syst. Sci.* **2015**, *124*, 1311–1324. [[CrossRef](#)]
42. Zaharia, L.; Costache, R.; Prăvălie, R.; Ioana-Toroimac, G. Mapping flood and flooding potential indices: A methodological approach to identifying areas susceptible to flood and flooding risk. Case study: The Prahova catchment (Romania). *Front. Earth Sci.* **2017**, *11*, 229–247. [[CrossRef](#)]
43. Bates, P.D. Remote sensing and flood inundation modelling. *Hydrol. Process.* **2004**, *18*, 2593–2597. [[CrossRef](#)]
44. Bates, P.D. Integrating remote sensing data with flood inundation models: How far have we got? *Hydrol. Process.* **2012**, *26*, 2515–2521. [[CrossRef](#)]
45. Murayama, Y. *Progress in Geospatial Analysis*; Springer Science & Business Media: Tokyo, Japan, 2012.
46. Roy, P.; Pal, S.C.; Chakraborty, R.; Chowdhuri, I.; Malik, S.; Das, B. Threats of climate and land use change on future flood susceptibility. *J. Clean. Prod.* **2020**, *272*, 122757. [[CrossRef](#)]
47. Chakraborty, R.; Pal, S.C.; Janizadeh, S.; Santosh, M.; Roy, P.; Chowdhuri, I.; Saha, A. Impact of climate change on future flood susceptibility: An evaluation based on deep learning algorithms and GCM model. *Water Resour. Manag.* **2021**, *35*, 4251–4274. [[CrossRef](#)]
48. Arnell, N.W.; Gosling, S.N. The impacts of climate change on river flood risk at the global scale. *Clim. Chang.* **2016**, *134*, 387–401. [[CrossRef](#)]
49. Garner, A.J.; Mann, M.E.; Emanuel, K.A.; Kopp, R.E.; Lin, N.; Alley, R.B.; Horton, B.P.; DeConto, R.M.; Donnelly, J.P.; Pollard, D. Impact of climate change on New York City’s coastal flood hazard: Increasing flood heights from the preindustrial to 2300 CE. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 11861–11866. [[CrossRef](#)] [[PubMed](#)]
50. Apurv, T.; Mehrotra, R.; Sharma, A.; Goyal, M.K.; Dutta, S. Impact of climate change on floods in the Brahmaputra basin using CMIP5 decadal predictions. *J. Hydrol.* **2015**, *527*, 281–291. [[CrossRef](#)]
51. Zhao, G.; Xu, Z.; Pang, B.; Tu, T.; Xu, L.; Du, L. An enhanced inundation method for urban flood hazard mapping at the large catchment scale. *J. Hydrol.* **2019**, *571*, 873–882. [[CrossRef](#)]
52. Avand, M.; Moradi, H. Spatial modeling of flood probability using geo-environmental variables and machine learning models, case study: Tajan watershed, Iran. *Adv. Space Res.* **2021**, *67*, 3169–3186. [[CrossRef](#)]
53. Diakakis, M.; Deligiannakis, G.; Pallikarakis, A.; Skordoulis, M. Factors controlling the spatial distribution of flash flooding in the complex environment of a metropolitan urban area. The case of Athens 2013 flash flood event. *Int. J. Disaster Risk Reduct.* **2016**, *18*, 171–180. [[CrossRef](#)]
54. Tien Bui, D.; Hoang, N.-D. A Bayesian framework based on a Gaussian mixture model and radial-basis-function Fisher discriminant analysis (BayGmmKda V1. 1) for spatial prediction of floods. *Geosci. Model Dev.* **2017**, *10*, 3391–3409. [[CrossRef](#)]
55. Ahmed, N.; Hoque, M.A.-A.; Arabameri, A.; Pal, S.C.; Chakraborty, R.; Jui, J. Flood susceptibility mapping in Brahmaputra floodplain of Bangladesh using deep boost, deep learning neural network, and artificial neural network. *Geocarto Int.* **2021**, 1–22. [[CrossRef](#)]
56. Bui, D.T.; Ngo, P.-T.T.; Pham, T.D.; Jaafari, A.; Minh, N.Q.; Hoa, P.V.; Samui, P. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. *Catena* **2019**, *179*, 184–196. [[CrossRef](#)]
57. Lee, M.-J.; Kang, J.-e.; Jeon, S. Application of frequency ratio model and validation for predictive flooded area susceptibility mapping using GIS. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 895–898.
58. Zhao, G.; Pang, B.; Xu, Z.; Peng, D.; Xu, L. Assessment of urban flood susceptibility using semi-supervised machine learning model. *Sci. Total Environ.* **2019**, *659*, 940–949. [[CrossRef](#)] [[PubMed](#)]
59. Pradhan, B.; Youssef, A. A 100-year maximum flood susceptibility mapping using integrated hydrological and hydrodynamic models: Kelantan River Corridor, Malaysia. *J. Flood Risk Manag.* **2011**, *4*, 189–202. [[CrossRef](#)]
60. Tehrany, M.S.; Pradhan, B.; Jebur, M.N. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *J. Hydrol.* **2014**, *512*, 332–343. [[CrossRef](#)]
61. Samanta, R.K.; Bhunia, G.S.; Shit, P.K.; Pourghasemi, H.R. Flood susceptibility mapping using geospatial frequency ratio technique: A case study of Subarnarekha River Basin, India. *Model. Earth Syst. Environ.* **2018**, *4*, 395–408. [[CrossRef](#)]
62. Martinez-Casasnovas, J.; Ramos, M.; Poesen, J. Assessment of sidewall erosion in large gullies using multi-temporal DEMs and logistic regression analysis. *Geomorphology* **2004**, *58*, 305–321. [[CrossRef](#)]
63. Evans, I.S. General geomorphometry, derivatives of altitude, and descriptive statistics. In *Spatial Analysis in Geomorphology*; Routledge: London, England, 2019; pp. 17–90.
64. Chapi, K.; Singh, V.P.; Shirzadi, A.; Shahabi, H.; Bui, D.T.; Pham, B.T.; Khosravi, K. A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environ. Model. Softw.* **2017**, *95*, 229–245. [[CrossRef](#)]
65. Bui, D.T.; Tsangaratos, P.; Ngo, P.-T.T.; Pham, T.D.; Pham, B.T. Flash flood susceptibility modeling using an optimized fuzzy rule based feature selection technique and tree based ensemble methods. *Sci. Total Environ.* **2019**, *668*, 1038–1054. [[CrossRef](#)] [[PubMed](#)]

66. Vojtek, M.; Vojteková, J. Flood susceptibility mapping on a national scale in Slovakia using the analytical hierarchy process. *Water* **2019**, *11*, 364. [[CrossRef](#)]
67. Ahmadlou, M.; Karimi, M.; Alizadeh, S.; Shirzadi, A.; Parvinnejhad, D.; Shahabi, H.; Panahi, M. Flood susceptibility assessment using integration of adaptive network-based fuzzy inference system (ANFIS) and biogeography-based optimization (BBO) and BAT algorithms (BA). *Geocarto Int.* **2019**, *34*, 1252–1272. [[CrossRef](#)]
68. Dammalage, T.; Jayasinghe, N. Land-use change and its impact on urban flooding: A case study on Colombo district flood on May 2016. *Eng. Technol. Appl. Sci. Res.* **2019**, *9*, 3887–3891. [[CrossRef](#)]
69. Tabari, H. Climate change impact on flood and extreme precipitation increases with water availability. *Sci. Rep.* **2020**, *10*, 13768. [[CrossRef](#)] [[PubMed](#)]
70. Taloor, A.K.; Manhas, D.S.; Kothiyari, G.C. Retrieval of land surface temperature, normalized difference moisture index, normalized difference water index of the Ravi basin using Landsat data. *Appl. Comput. Geosci.* **2021**, *9*, 100051. [[CrossRef](#)]
71. Helbich, M. Spatiotemporal contextual uncertainties in green space exposure measures: Exploring a time series of the normalized difference vegetation indices. *Int. J. Environ. Res. Public Health* **2019**, *16*, 852. [[CrossRef](#)] [[PubMed](#)]
72. Dormann, C.F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Marquéz, J.R.G.; Gruber, B.; Lafourcade, B.; Leitão, P.J. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **2013**, *36*, 27–46. [[CrossRef](#)]
73. Wuebbles, D.J.; Fahey, D.W.; Hibbard, K.A.; Dokken, D.J.; Stewart, B.C.; Maycock, T.K. *Climate Science Special Report: Fourth National Climate Assessment*; U.S. Global Change Research Program (USGCRP): Washington, DC, USA, 2017; Volume I, p. 470.
74. Faizollahzadeh Ardabili, S.; Najafi, B.; Shamsirband, S.; Minaei Bidgoli, B.; Deo, R.C.; Chau, K.-w. Computational intelligence approach for modeling hydrogen production: A review. *Eng. Appl. Comput. Fluid Mech.* **2018**, *12*, 438–458. [[CrossRef](#)]
75. Ahmad, A.; Anderson, T.; Lie, T. Hourly global solar irradiation forecasting for New Zealand. *Sol. Energy* **2015**, *122*, 1398–1408. [[CrossRef](#)]
76. Bishop, C.M.; Nasrabadi, N.M. *Pattern Recognition and Machine Learning*; Springer: New York, USA, 2006; Volume 4.
77. Gardner, M.W.; Dorling, S. Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences. *Atmos. Environ.* **1998**, *32*, 2627–2636. [[CrossRef](#)]
78. Zhang, H. The optimality of naive Bayes. *Aa* **2004**, *1*, 3.
79. Quiroz, J.C.; Mariun, N.; Mehrjou, M.R.; Izadi, M.; Mison, N.; Radzi, M.A.M. Fault detection of broken rotor bar in LS-PMSM using random forests. *Measurement* **2018**, *116*, 273–280. [[CrossRef](#)]
80. Zhang, H.; Wang, M. Search for the smallest random forest. *Stat. Its Interface* **2009**, *2*, 381.
81. Natekin, A.; Knoll, A. Gradient boosting machines, a tutorial. *Front. Neurobot.* **2013**, *7*, 21. [[CrossRef](#)] [[PubMed](#)]
82. Yuan, X.; Abouelenien, M. A multi-class boosting method for learning from imbalanced data. *Int. J. Granul. Comput. Rough Sets Intell. Syst.* **2015**, *4*, 13–29. [[CrossRef](#)]
83. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [[CrossRef](#)]
84. Panahi, M.; Gayen, A.; Pourghasemi, H.R.; Rezaie, F.; Lee, S. Spatial prediction of landslide susceptibility using hybrid support vector regression (SVR) and the adaptive neuro-fuzzy inference system (ANFIS) with various metaheuristic algorithms. *Sci. Total Environ.* **2020**, *741*, 139937. [[CrossRef](#)] [[PubMed](#)]
85. Bui, D.T.; Pradhan, B.; Lofman, O.; Revhaug, I.; Dick, O.B. Landslide susceptibility mapping at Hoa Binh province (Vietnam) using an adaptive neuro-fuzzy inference system and GIS. *Comput. Geosci.* **2012**, *45*, 199–211.
86. Han, J.; Nur, A.S.; Syifa, M.; Ha, M.; Lee, C.-W.; Lee, K.-Y. Improvement of earthquake risk awareness and seismic literacy of Korean citizens through earthquake vulnerability map from the 2017 pohang earthquake, South Korea. *Remote Sens.* **2021**, *13*, 1365. [[CrossRef](#)]
87. Umar, Z.; Pradhan, B.; Ahmad, A.; Jebur, M.N.; Tehrany, M.S. Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in West Sumatera Province, Indonesia. *Catena* **2014**, *118*, 124–135. [[CrossRef](#)]
88. Bui, D.T.; Pradhan, B.; Nampak, H.; Bui, Q.-T.; Tran, Q.-A.; Nguyen, Q.-P. Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood susceptibility modeling in a high-frequency tropical cyclone area using GIS. *J. Hydrol.* **2016**, *540*, 317–330.
89. Bui, D.T.; Hoang, N.-D.; Martínez-Álvarez, F.; Ngo, P.-T.T.; Hoa, P.V.; Pham, T.D.; Samui, P.; Costache, R. A novel deep learning neural network approach for predicting flash flood susceptibility: A case study at a high frequency tropical storm area. *Sci. Total Environ.* **2020**, *701*, 134413.
90. Lei, X.; Chen, W.; Panahi, M.; Falah, F.; Rahmati, O.; Uuemaa, E.; Kalantari, Z.; Ferreira, C.S.S.; Rezaie, F.; Tiefenbacher, J.P. Urban flood modeling using deep-learning approaches in Seoul, South Korea. *J. Hydrol.* **2021**, *601*, 126684. [[CrossRef](#)]
91. Chen, W.; Hong, H.; Panahi, M.; Shahabi, H.; Wang, Y.; Shirzadi, A.; Pirasteh, S.; Alesheikh, A.A.; Khosravi, K.; Panahi, S. Spatial prediction of landslide susceptibility using gis-based data mining techniques of anfis with whale optimization algorithm (woa) and grey wolf optimizer (gwo). *Appl. Sci.* **2019**, *9*, 3755. [[CrossRef](#)]
92. Yin, Y.; Yasuda, K. Similarity coefficient methods applied to the cell formation problem: A comparative investigation. *Comput. Ind. Eng.* **2005**, *48*, 471–489. [[CrossRef](#)]
93. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **2020**, *21*, 6. [[CrossRef](#)] [[PubMed](#)]

94. Zhao, G.; Pang, B.; Xu, Z.; Peng, D.; Zuo, D. Urban flood susceptibility assessment based on convolutional neural networks. *J. Hydrol.* **2020**, *590*, 125235. [[CrossRef](#)]
95. Goswami, B.N.; Venugopal, V.; Sengupta, D.; Madhusoodanan, M.; Xavier, P.K. Increasing trend of extreme rain events over India in a warming environment. *Science* **2006**, *314*, 1442–1445. [[CrossRef](#)]
96. Houghton, J.T.; Ding, Y.; Griggs, D.; Noguer, M.; Van Der Linden, P.; Dai, X.; Maskell, K.; Johnson, C. (Eds.) *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2001; p. 881.
97. Saha, T.K.; Pal, S.; Talukdar, S.; Debanshi, S.; Khatun, R.; Singha, P.; Mandal, I. How far spatial resolution affects the ensemble machine learning based flood susceptibility prediction in data sparse region. *J. Environ. Manag.* **2021**, *297*, 113344. [[CrossRef](#)] [[PubMed](#)]
98. Kia, M.B.; Pirasteh, S.; Pradhan, B.; Mahmud, A.R.; Sulaiman, W.N.A.; Moradi, A. An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. *Environ. Earth Sci.* **2012**, *67*, 251–264. [[CrossRef](#)]
99. Avand, M.; Khiavi, A.N.; Khazaei, M.; Tiefenbacher, J.P. Determination of flood probability and prioritization of sub-watersheds: A comparison of game theory to machine learning. *J. Environ. Manag.* **2021**, *295*, 113040. [[CrossRef](#)]
100. Pham, B.T.; Phong, T.V.; Nguyen, H.D.; Qi, C.; Al-Ansari, N.; Amini, A.; Ho, L.S.; Tuyen, T.T.; Yen, H.P.H.; Ly, H.-B. A comparative study of kernel logistic regression, radial basis function classifier, multinomial naïve bayes, and logistic model tree for flash flood susceptibility mapping. *Water* **2020**, *12*, 239. [[CrossRef](#)]
101. Farhadi, H.; Najafzadeh, M. Flood Risk Mapping by Remote Sensing Data and Random Forest Technique. *Water* **2021**, *13*, 3115. [[CrossRef](#)]
102. Chau, K.W.; Wu, C.; Li, Y.S. Comparison of several flood forecasting models in Yangtze River. *J. Hydrol. Eng.* **2005**, *10*, 485–491. [[CrossRef](#)]
103. Pradhan, B. Flood susceptible mapping and risk area delineation using logistic regression, GIS and remote sensing. *J. Spat. Hydrol.* **2010**, *9*, 2.
104. Kourgialas, N.; Karatzas, G. Gestion des inondations et méthode de modélisation sous SIG pour évaluer les zones d’aléa inondation-une étude de cas. *Hydrol. Sci. J.* **2011**, *56*, 212–225. [[CrossRef](#)]
105. Segond, M.-L.; Wheeler, H.S.; Onof, C. The significance of spatial rainfall representation for flood runoff estimation: A numerical evaluation based on the Lee catchment, UK. *J. Hydrol.* **2007**, *347*, 116–131. [[CrossRef](#)]
106. Avand, M.; Kuriqi, A.; Khazaei, M.; Ghorbanzadeh, O. DEM resolution effects on machine learning performance for flood probability mapping. *J. Hydro-Environ. Res.* **2022**, *40*, 1–16. [[CrossRef](#)]
107. Collins, E.L.; Sanchez, G.M.; Terando, A.; Stillwell, C.C.; Mitasova, H.; Sebastian, A.; Meentemeyer, R.K. Predicting flood damage probability across the conterminous United States. *Environ. Res. Lett.* **2022**, *17*, 034006. [[CrossRef](#)]
108. Atluri, G.; Karpadne, A.; Kumar, V. Spatio-temporal data mining: A survey of problems and methods. *ACM Comput. Surv.* **2018**, *51*, 1–41. [[CrossRef](#)]
109. Xie, Y.; He, E.; Jia, X.; Bao, H.; Zhou, X.; Ghosh, R.; Ravirathinam, P. A statistically-guided deep network transformation and moderation framework for data with spatial heterogeneity. In Proceedings of the 2021 IEEE International Conference on Data Mining (ICDM), Auckland, New Zealand, 7–10 December 2021; pp. 767–776.
110. Gaganis, P. Model calibration/parameter estimation techniques and conceptual model error. In *Uncertainties in Environmental Modelling and Consequences for Policy Making*; Springer: Dordrecht, The Netherlands, 2009; pp. 129–154.