

# FLOOD MAPPING USING RANDOM FOREST AND IDENTIFYING THE ESSENTIAL CONDITIONING FACTORS; A CASE STUDY IN FREDERICTON, NEW BRUNSWICK, CANADA

M. Esfandiari<sup>1\*</sup>, S. Jabari<sup>1</sup>, H. McGrath<sup>2</sup>, D. Coleman<sup>1</sup>

<sup>1</sup> Department of Geodesy & Geomatics Engineering, University of New Brunswick, Fredericton, NB, Canada (phone: +1-506-260-7279; m.esfandiari@unb.ca).

<sup>1</sup> Department of Geodesy & Geomatics Engineering, University of New Brunswick, Fredericton, NB, Canada (phone: +1-506-447-3124; sh.jabari@unb.ca).

<sup>2</sup> Natural Resource Canada, Ottawa, ON Canada (heather.mcgrath@canada.ca).

<sup>1</sup>Department of Geodesy & Geomatics Engineering, University of New Brunswick, Fredericton, NB, Canada (email: dcoleman@unb.ca).

## Commission III, ICWG III/IVa

**KEY WORDS:** Machine Learning, Random Forest, Flood mapping, Conditioning factors, GIS, Remote Sensing

### ABSTRACT:

Flood is one of the most damaging natural hazards in urban areas in many places around the world as well as the city of Fredericton, New Brunswick, Canada. Recently, Fredericton has been flooded in two consecutive years in 2018 and 2019. Due to the complicated behaviour of water when a river overflows its bank, estimating the flood extent is challenging. The issue gets even more challenging when several different factors are affecting the water flow, like the land texture or the surface flatness, with varying degrees of intensity. Recently, machine learning algorithms and statistical methods are being used in many research studies for generating flood susceptibility maps using topographical, hydrological, and geological conditioning factors. One of the major issues that researchers have been facing is the complexity and the number of features required to input in a machine-learning algorithm to produce acceptable results. In this research, we used Random Forest to model the 2018 flood in Fredericton and analyzed the effect of several combinations of 12 different flood conditioning factors. The factors were tested against a Sentinel-2 optical satellite image available around the flood peak day. The highest accuracy was obtained using only 5 factors namely, altitude, slope, aspect, distance from the river, and land-use/cover with 97.57% overall accuracy and 95.14% kappa coefficient.

### 1. INTRODUCTION

Flood is one of the most destructive natural hazards that is rapidly growing as a result of global warming and climate change (Schiermeier 2011; Gaur, Gaur, and Simonovic 2018). There are different types of floods including coastal flood, flash flood, and river flood. One of the significant challenges in flood mapping is to provide an accurate estimation of flood extent and damage amount in affected areas. There are various techniques for estimating flood behaviour including *hydrodynamic models*, *simplified conceptual models*, and *empirical methods* (Teng et al. 2017). (1) The *hydrodynamic models* are mainly divided into 1D, 2D, and 3D and use complex mathematical equations to simulate fluid motion. Depending on the topography of a region and floodplain, and the required level of accuracy, different models can be selected for demonstrating flood damage and extent in affected areas (Teng et al. 2017). (2) *Simplified conceptual models* are not as detailed as hydrodynamic models and require less amount of data but were able to acquire highly acceptable results in many case studies (Teng et al. 2017; Momo 2014; Liu et al. 2016; Speckhann et al. 2018). (3) *Empirical methods* include all the flood maps which are generated using *observations*. These *observations* can be satellite images, aerial photographs, surveying, etc. The accuracy of the flood maps generated by the *observations* is totally dependent on the accuracy of the observations, which represents the limitations of empirical methods. Nevertheless, the output of empirical methods is used in a variety of ways for the validation of different models. On the other hand, machine learning algorithms and

statistical methods are being increasingly used for generating flood susceptibility maps in many research studies. Numerous researchers have implemented extensive investigations and applied different algorithms to various datasets (Tehrany, Jones, and Shabani 2019; Rahmati, Pourghasemi, and Melesse 2016; Youssef et al. 2016; Kia et al. 2012). Using machine learning, different factors, referred to as conditioning factors, are used to generate flood susceptibility maps (Tehrany, Pradhan, and Jebur 2014), or estimate the amount of damage (Tehrany, Pradhan, and Jebur 2014).

One of the major issues that researchers have been facing is the complexity and the number of conditioning factors that refer to either hydrological, topographical or geological layers. Also, it is possible to provide more conditioning factors to machine learning algorithms and expect to achieve better results. In this research, we have examined several different scenarios with different combinations of 12 conditioning factors including altitude, slope, aspect, distance from river, land-use/cover, terrain wetness index (TWI), terrain roughness index (TRI), stream power index (SPI), curvature, plan curvature, profile curvature, and height above the nearest drainage (HAND). The Random Forest algorithm (Ho 1995; Breiman 2001), which creates a multitude of decision trees and provides an estimation of the importance of the parameters in decision making, is used for this analysis. The algorithm's robustness, low bias, the capability of handling unbalanced data, high dimensional data, and its quick prediction makes it a useful tool for this research among other machine learning methods. In this paper, first the study area and the dataset are introduced in section 2, then the methodology,

\* Corresponding author

results and discussion, and the conclusion are presented in sections 3, 4, and 5, respectively.

## 2. STUDY AREA AND DATASET

This study focuses on the downtown and surrounding areas of the city of Fredericton, the capital of the province of New Brunswick, Canada. The city is located in the west-central part of the province along the Saint John River. Due to the geographical location of the city, every year the Saint John River gets frozen because of the cold winters. Usually, at the end of each April, the frozen river starts to meltdown, and it leads to a significant rise in the water level. In late April 2018, the water level at the Saint John River raised to the historic elevation of 8.13 meters. This event was recorded as one of the most damaging flooding events in the history of Fredericton, which affected a total of 12000 properties around the province. The general topography of Fredericton is flat in the areas close to the river and connected streams, but the elevation rises as we get farther from the river. The elevation ranges from around 190 m, west of Fredericton to just above sea level in the area of study. A dam, named Mactaquac dam, is located around 19 kilometres upstream from the city which contains a small pond and is not able to hold the melted ice for long so the water must be released to the river.

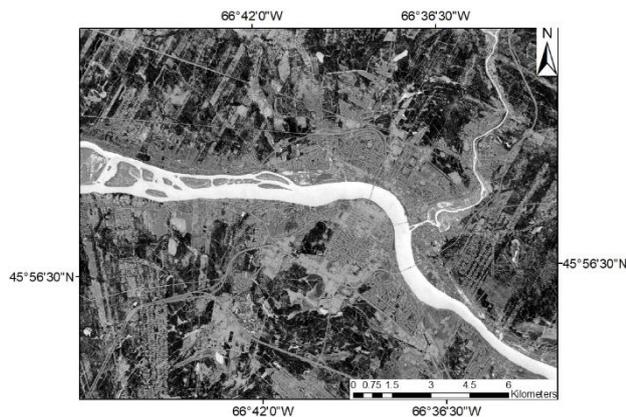


Figure 1. NDWI showing the water extent before the flood

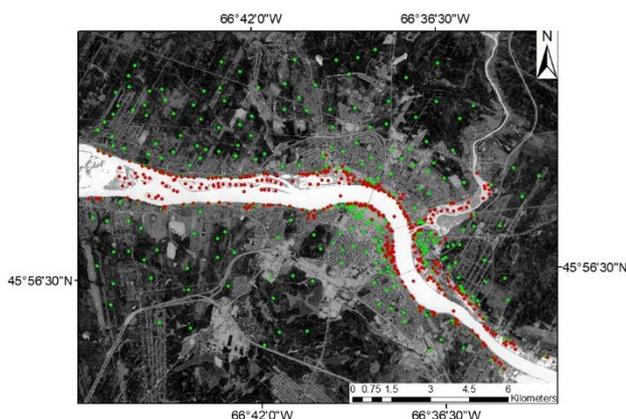


Figure 2. NDWI and sample points generated from Sentinel-2 satellite image at the flood peak

## 3. METHODOLOGY

Random Forest is one of the most robust, efficient, and highly flexible ensemble classifiers that creates a multitude of decision trees (Breiman 2001). The algorithm uses random bootstrapped

samples from training data to predict the probability of a pixel being flooded or not. The algorithm runs arbitrary binary trees that perform a subset of observations over the bootstrapping approach. From the original dataset, a random selection of the training data is considered for creating the model, and the disregarded data is described as out of bag (OOB) (Catani et al. 2013). Random Forest also predicts the importance of each variable as well. In the first step, the prediction error of the OOB part of the data, and then for permuting each variable is recorded. The difference between the two is averaged over all trees and normalized by the standard deviation of the differences. The second measure is the cumulative decrease in impurities of nodes from splitting onto the conditioning factor, averaged over all trees (Liaw and Wiener 2002).

### 3.1 Flood Conditioning Factors

In order to obtain the flood model using Random Forest, several conditioning factors that contribute to flooding were selected. For the analysis, various combinations of datasets were constructed. The selection of conditioning factors depends on the area of the study and its characteristics (Kia et al. 2012). For this research, the conditioning factors were selected based on expert's analysis and the information from the literature (Kia et al. 2012).

A total of 12 conditioning factors, all shown in Figure 3 and 4, were selected for flood mapping using Random Forest, namely: altitude, slope, aspect, distance from the river, land-use/cover, TWI, TRI, SPI, curvature, plan curvature, profile curvature, and HAND. High accuracy topographic data is one of the most important parameters required for precisely modelling the flood extent (Bates, Marks, and Horritt 2003). In this research we used the altitude layer (Digital Terrain Model, DTM), which was obtained from light detection and ranging (LiDAR) with 1m spatial resolution and slope, aspect, TWI, TRI, SPI, curvature, plan curvature, and profile curvature were derived from the altitude layer in ArcGIS 10.6.1 software. The distance from the river layer was generated using the Euclidean distance tool within the ArcGIS software. The distance was calculated from the river boundary polygon shapefile provided with GeoNB, the geographic data catalogue website of the province of New Brunswick. The Land-use/cover layer was made by overlaying available polygons within the catalogue GeoNB website containing seven classes of Urban, Forest, Gras Land, Bare Land, Roads, Water, and Wetlands. However, after overlaying all the polygons, there were still existing unclassified areas that were filled by classifying a Sentinel-2 satellite image. TWI is the cumulative upslope and it represents the potential of water accumulation in certain areas based on the tendency of gravitational forces; for the formula please refer to (BEVEN and Kirkby 1979). To express the amount of elevation difference specifically between the adjacent cells of a DTM, we have used TRI which is given in (Riley, DeGloria, and Elliot 1999). Stream's erosion is measured by SPI which is also seen as a conditioning factor reflecting the stability of an area; for the formula, please see (Moore, Grayson, and Ladson 1991). Curvature, plan curvature, and profile curvatures layers were considered as conditioning factors as parameters that show the level of flatness in the area. The formula for these parameters can be found in (Heerdegen and Beran 1982). HAND model is an adjusted elevation layer that is normalized toward the nearest stream (Rennó et al. 2008). The elevation of each pixel in the HAND layer is calculated based on the D-infinity flow direction (Tarboton 1997) and the elevation difference of each pixel (Nobre et al. 2016). The other conditioning factors are the derivatives of altitude, i.e., slope and aspect which play an important role in recognizing the vulnerable areas to flood.

The conditioning factors mentioned above are ordinal and nominal ones. Thus, for a better implication of Random Forest, all ordinal factors were normalized from 0 to 1 (Ihsan, Idris, and Abdullah 2013). All the conditioning factors were arranged to have the same extent containing the whole city of Fredericton and surrounding areas. The created databases of conditioning factors constructed grids of 22448 columns and 11533 rows (~ 258 km<sup>2</sup>). Generally, areas with lower elevation, flat surface, and rough surface with low potential for absorption are more prone to flood (Tehrany, Jones, and Shabani 2019).

### 3.2 Algorithm Training

The precision of the data used for generating a flood model has a very high impact on the accuracy of the flood model itself (Merz, Thielen, and Gocht 2007). Several sample points were collected through site visits at the time of flood events around the city. Also, Sentinel-2 satellite images were used for generating sample points that were taken at the time of the flood. For generating sample points from the satellite image, pre-flood (Figure 1.) and flood-peak (Figure 2.) images were used, which were taken on April 22<sup>nd</sup>, 2018 and May 02<sup>nd</sup>, 2018, respectively. To identify water pixels, a normalized difference water index

(NDWI) indicator was derived from the images; the formula for which can be found in (Gao 1996). Using the ground truth data and by visually inspecting the NDWI layer, a total of 740 flooded and non-flooded samples were generated. To prevent the class imbalance issue, an equal number of flooded and not-flooded points were generated, which were distributed evenly in the area close to the river boundary. Then, the sample points were randomly divided into two groups of training (70%) with 259 flooded and 259 not-flooded points, and testing (30%), with 111 flooded and 111 not-flooded points. The random selection of the points helps to avoid auto correlation.

### 3.3 Algorithm Implementation

The Random Forest algorithm was implemented in RStudio 1.2.1335. The hyperparameters and the implementation criteria were selected based on the literature (Rahmati, Pourghasemi, and Melesse 2016). To run the algorithm, it is necessary to define the number of parameters and trees (Youssef et al. 2016). In this research, each implemented scenario used a different number of conditioning factors, as parameters, and the number of trees was set to 1000 for all the different test scenarios.

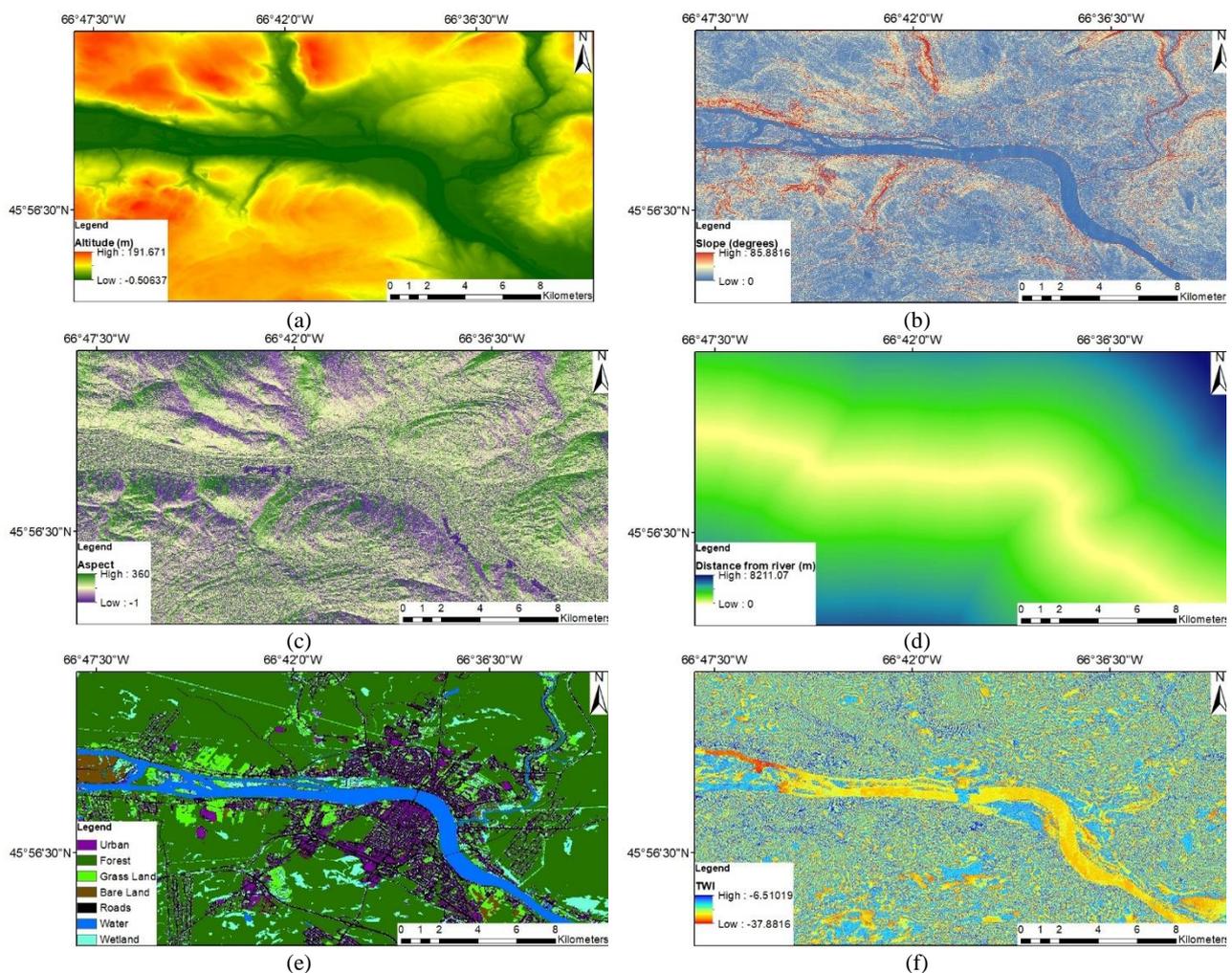


Figure 3. Conditioning factors used in the Random Forest classifier (part 1): (a) Altitude (b) Slope (c) Aspect (d) Distance from river (e) Land-use/cover (f) TWI 12

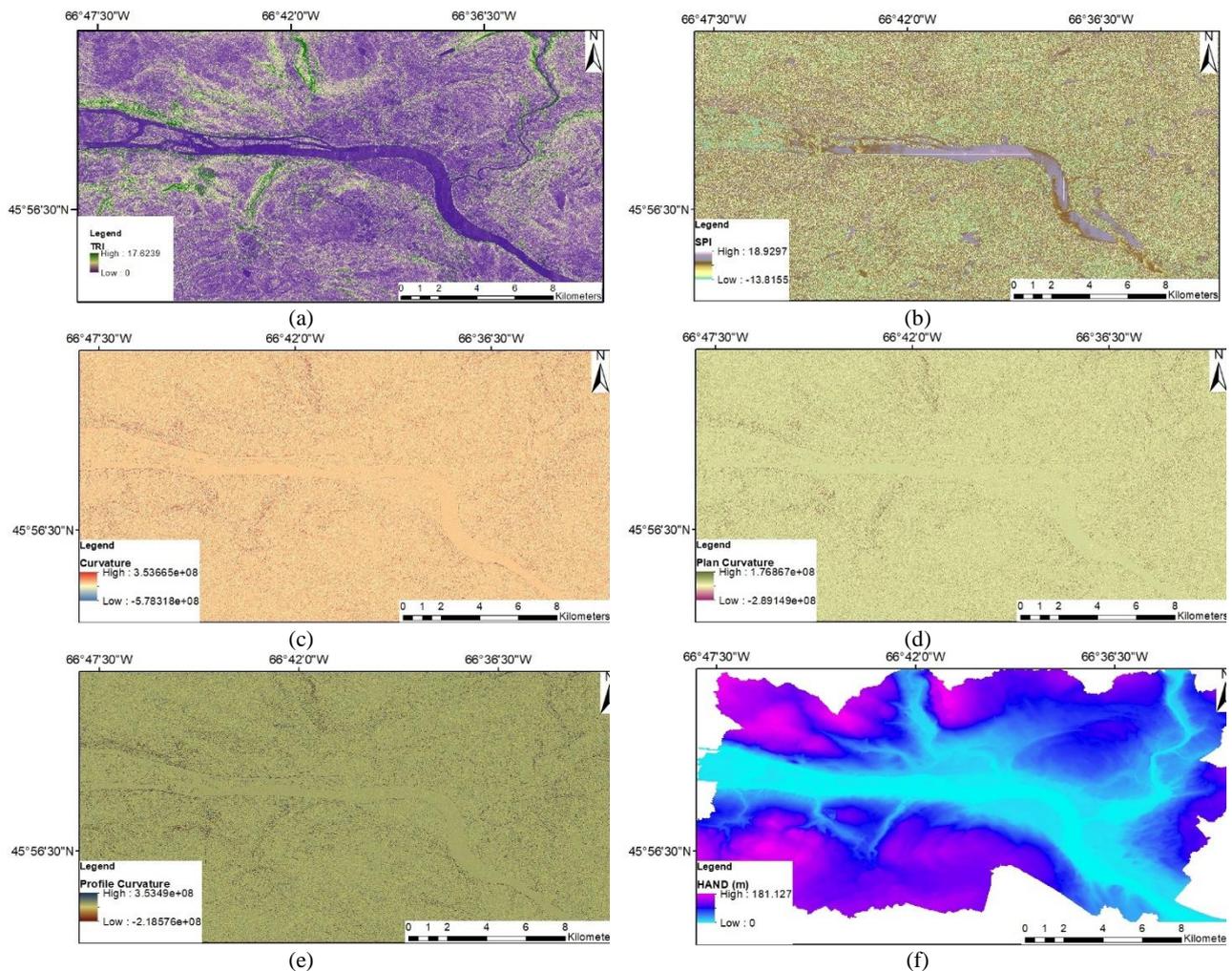


Figure 4. Conditioning factors used in the Random Forest classifier (part 2): (a) TRI (b) SPI (c) Curvature (d) Plan curvature (e) Profile curvature (f) HAND

Scenario 1	Conditioning Factors
1-a	Altitude-Slope-Aspect- Distance- Land-use/cover
1-b	Altitude-Slope-Aspect- Distance- Land-use/cover -TWI
1-c	Altitude-Slope-Aspect- Distance- Land-use/cover -TWI-TRI
1-d	Altitude-Slope-Aspect- Distance- Land-use/cover -TWI-TRI-SPI
1-e	Altitude-Slope-Aspect- Distance- Land-use/cover -TWI-TRI-SPI-Curvature
1-f	Altitude-Slope-Aspect- Distance- Land-use/cover -TWI-TRI-SPI-Curvature-Plan Curvature
1-g	Altitude-Slope-Aspect-Distance-Land-use/cover-TWI-TRI-SPI-Curvature-Plan Curvature-Profile Curvature
1-h	Altitude-Slope-Aspect-Distance-Land-use/cover-TWI-TRI-SPI-Curvature-Plan Curvature-Profile Curvature-HAND
1-i	Altitude-Slope-Distance- Land-use/cover -TWI-TRI-SPI-Curvature-Plan Curvature-Profile Curvature-HAND
1-j	Altitude-Slope-Distance- Land-use/cover -TWI-SPI-Curvature-Plan Curvature-Profile Curvature-HAND
1-k	Altitude-Slope-Distance- Land-use/cover -TWI-Curvature-Plan Curvature-Profile Curvature-HAND
1-l	Altitude-Slope- Distance- Land-use/cover -Curvature-Plan Curvature-Profile Curvature-HAND
1-m	Altitude-Slope-Distance- Land-use/cover -Curvature-Plan Curvature-HAND
1-n	Altitude-Slope- Distance- Land-use/cover -Plan Curvature-HAND
1-o	Altitude-Slope- Distance- Land-use/cover -HAND
1-p	Altitude-Distance- Land-use/cover -HAND

Table 1. Combinations of different conditioning factors implemented in Scenario 1

In order to identify the most important conditioning factors for flood mapping, different conditioning factors were tested in two separate scenarios. In Scenario 1, the Random Forest algorithm was trained using five conditioning factors which are most frequently used in the literature namely: altitude, slope, aspect, distance from river, and land-use/cover (Tehrany, Jones, and Shabani 2019; Tehrany, Pradhan, and Jebur 2014; Tehrany et al. 2014; Rahmati, Pourghasemi, and Melesse 2016). These five condition factors were used to predict the flooded pixels using random forest (Figure 5). Then, the remaining conditioning factors were added to the five condition factors one by one until we used all the 12 conditioning factors for training and prediction (Table 1 Scenarios 1-a to 1-h). The Random Forest algorithm prioritizes the conditioning factors based on their degree of importance. Therefore, for the next step, conditioning factors with the least degree of importance were removed from the combinations, and we continued removing the least important conditioning factors until there were four conditioning factors left only (Table 1 Scenarios 1-i to 1-p). In Scenario 2 (Table 2.), we grouped the correlated conditioning factors together and made sure that only one conditioning factor from each group is used in each combination for training and prediction. Altitude and HAND conditioning factors were grouped together as they both are elevation based. Slope, TWI, TRI, SPI, curvature, plan curvature, and profile curvature grouped together as well as they all are Slope-based conditioning factors.

Scenario	Conditioning factors
2	
2-a	Slope-Aspect-Distance- Land-use/cover -HAND
2-b	Aspect-Distance- Land-use/cover -TWI-HAND
2-c	Aspect-Distance- Land-use/cover -TRI-HAND
2-d	Aspect-Distance- Land-use/cover -SPI-HAND
2-e	Aspect-Distance- Land-use/cover -Curvature- HAND
2-f	Aspect-Distance- Land-use/cover -Plan Curvature- HAND
2-g	Aspect-Distance-Land-use/cover-Profile Curvature-HAND
2-h	Altitude-Aspect-Distance- Land-use/cover -TWI
2-i	Altitude-Aspect-Distance- Land-use/cover -TRI
2-j	Altitude-Aspect-Distance- Land-use/cover -SPI
2-k	Altitude-Aspect-Distance- Land-use/cover - Curvature
2-l	Altitude-Aspect-Distance- Land-use/cover -Plan Curvature
2-m	Altitude-Aspect-Distance- Land-use/cover -Profile Curvature

Table 2. Combinations of different conditioning factors for implementations in Scenario 2

After implementing the Random Forest algorithm to different combinations of conditioning factors, a probability map with values from 0 to 1 was generated from each implementation. The value of each pixel represents the probability of that pixel being flooded or not. The probability map of each scenario then was classified into 5 classes of very low, low, moderate, high, and very high using Jenks natural breaks classification method (North 2009). High and very high classes of the probability maps were considered to be flooded areas in this research.

#### 4. RESULTS & DISCUSSION

Through Random Forest analysis the flooded and not flooded areas were distinguished using several different combinations of

conditioning factors, through two different test scenarios, as shown in Tables 1 and 2.

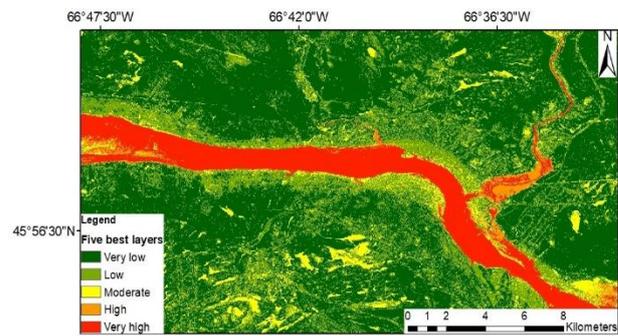


Figure 5. Random Forest output using altitude, slope, aspect, distance from the river, and land-use/cover

In order to model the 2018 flood map in Fredericton, we used the Random Forest algorithm by considering various combinations of 12 different conditioning factors contributing to flooding with different degrees of impact. As can be seen from Figure 6, Scenario 1-a distinguished flooded and not-flooded pixels accurately. However, as we kept adding the conditioning factors (Scenario 1-a to 1-h), the accuracy didn't increase. This shows that adding extra conditioning factors does not guarantee producing higher accuracies. This could be due to the negative importance of certain conditioning factors. Thus, in the next step, we kept removing the least important conditioning factors (Scenario 1-i to 1-p), while the accuracy didn't increase either. An explanation for that could be maybe since some conditioning factors were correlated, they could collectively degrade the accuracy. In Figure 7., the highest overall accuracy and kappa coefficient (Campbell and Wynne 2011) of 97.57% and 95.13% respectively, belong to Scenario 1-a where only five conditioning factors of altitude, slope, aspect, distance from river, and land-use/cover were used.

Since correlated conditioning factors could negatively affect the accuracy, In Scenario 2, we insured only uncorrelated conditioning factors are embedded in each combination. The best-acquired accuracy in Scenario 2 is achieved using slope, aspect, distance, land-use/cover, and HAND conditioning factors (Figure 8) in which flooded, and not-flooded pixels were distinguished with an overall accuracy and kappa coefficient of 97.57% and 95.14% respectively. This confirms that using altitude or the HAND does not change the final flood prediction accuracy.

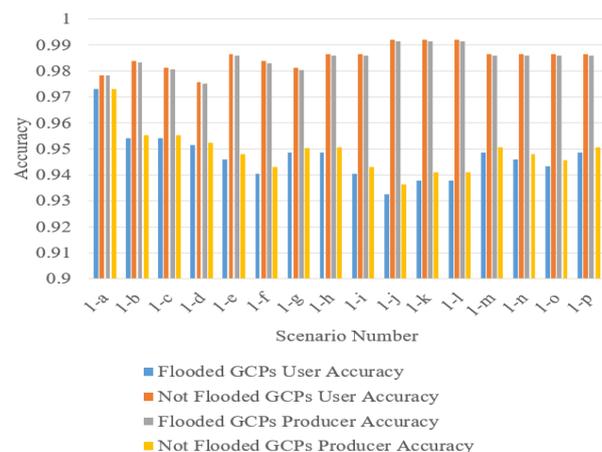


Figure 6. User and producer accuracy of flooded and not-flooded sample points in the first series of implementations

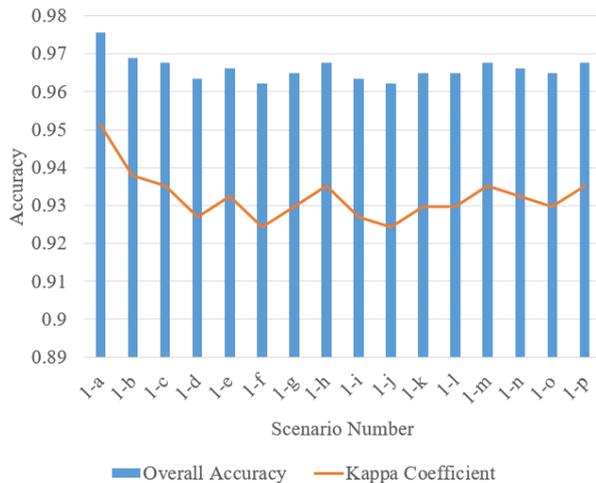


Figure 7. Overall Accuracy and Kappa Coefficient of the first series of implementations

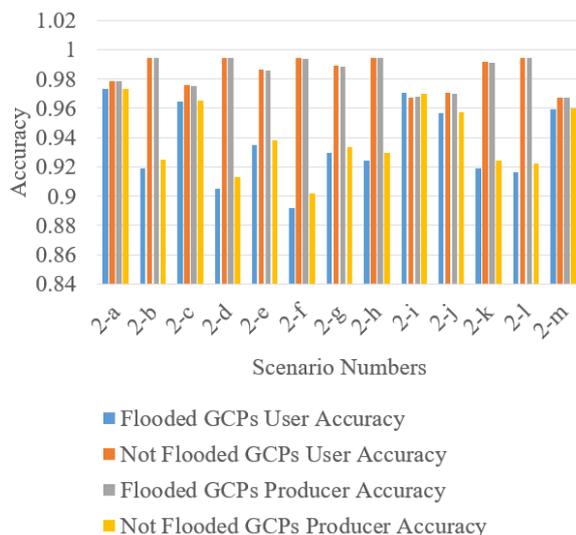


Figure 8. User and producer accuracy of flooded and not-flooded sample points

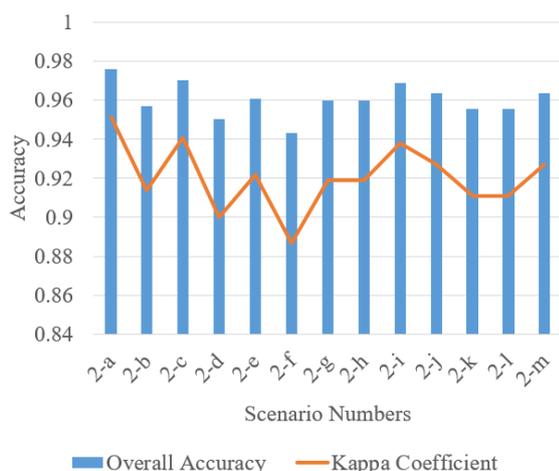


Figure 9. Overall Accuracy and Kappa Coefficient of the second series of implementations

Overall, among various conditioning factors, the most important conditioning factors for flood mapping, which produced the highest accuracies in both scenarios, were either altitude or HAND model, slope, aspect, distance from river, and land-use/cover.

## 5. CONCLUSION

Flood is one of the most catastrophic events that many countries around the world are experiencing. The city of Fredericton experienced a severe flood in 2018 and 2019 which caused considerable damage to urban infrastructures and residential buildings. There are several geological conditioning factors that contribute to flood mapping, but it is essential to identify the most effective ones for the analysis that could provide the best results. Applying similar research to another region will provide useful for impact assessment, prediction of vulnerable areas, and rescue assessment.

In this research, several combinations of conditioning factors were analyzed to find the combination that provides the most accurate flood model using the Random Forest algorithm. Results revealed that having correlated conditioning factors can degrade the prediction accuracy. Five conditioning factors of altitude or HAND model, slope, aspect, distance from the river, and land-use/cover provided the most accurate results. Furthermore, the following conclusions were achieved:

- Adding extra conditioning factors does not increase the accuracy of predictions.
- Including correlated layers decreases the accuracy of predictions.
- HAND and altitude layers are both major factors in flooding having similar effects on the final accuracy.

## REFERENCES

- Bates, P. D., K. J. Marks, and M. S. Horritt. 2003. "Optimal Use of High-Resolution Topographic Data in Flood Inundation Models." *Hydrological Processes* 17 (3): 537–557.
- BEVEN, Keith J., and Michael J. Kirkby. 1979. "A Physically Based, Variable Contributing Area Model of Basin Hydrology/Un Modèle à Base Physique de Zone d'appel Variable de l'hydrologie Du Bassin Versant." *Hydrological Sciences Journal* 24 (1): 43–69.
- Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32.
- Campbell, James B., and Randolph H. Wynne. 2011. *Introduction to Remote Sensing*. Guilford Press.
- Catani, F., D. Lagomarsino, S. Segoni, and V. Tofani. 2013. "Landslide Susceptibility Estimation by Random Forests Technique: Sensitivity and Scaling Issues." *Natural Hazards and Earth System Sciences* 13 (11): 2815.
- Gao, Bo-Cai. 1996. "NDWI—A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space." *Remote Sensing of Environment* 58 (3): 257–266.
- Gaur, Ayushi, Abhishek Gaur, and Slobodan P. Simonovic. 2018. "Future Changes in Flood Hazards across Canada under a Changing Climate." *Water* 10 (10): 1441.

- Heerdegen, Richard G., and Max A. Beran. 1982. "Quantifying Source Areas through Land Surface Curvature and Shape." *Journal of Hydrology* 57 (3–4): 359–373.
- Ho, Tin Kam. 1995. "Random Decision Forests." In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, 1:278–282. IEEE.
- Ihsan, Zohair, Mohd Yazid Idris, and Abdul Hanan Abdullah. 2013. "Attribute Normalization Techniques and Performance of Intrusion Classifiers: A Comparative Analysis." *Life Science Journal* 10 (4): 2568–2576.
- Kia, Masoud Bakhtyari, Saied Pirasteh, Biswajeet Pradhan, Ahmad Rodzi Mahmud, Wan Nor Azmin Sulaiman, and Abbas Moradi. 2012. "An Artificial Neural Network Model for Flood Simulation Using GIS: Johor River Basin, Malaysia." *Environmental Earth Sciences* 67 (1): 251–264.
- Liaw, Andy, and Matthew Wiener. 2002. "Classification and Regression by RandomForest." *R News* 2 (3): 18–22.
- Liu, Yan Y., David R. Maidment, David G. Tarboton, Xing Zheng, Ahmet Yildirim, Nazmus S. Sazib, and Shaowen Wang. 2016. "A CyberGIS Approach to Generating High-Resolution Height above Nearest Drainage (HAND) Raster for National Flood Mapping."
- Merz, Bruno, A. H. Thielen, and Martin Gocht. 2007. "Flood Risk Mapping at the Local Scale: Concepts and Challenges." In *Flood Risk Management in Europe*, 231–251. Springer.
- Momo, M. R. 2014. "Avaliação Da Aplicação Do Modelo HAND No Mapeamento de Áreas Suscetíveis a Inundação No Município de Blumenau." PhD Thesis, Tese de Mestrado em Engenharia Ambiental. Fundação Universidade Regional de ...
- Moore, Ian Donald, R. B. Grayson, and A. R. Ladson. 1991. "Digital Terrain Modelling: A Review of Hydrological, Geomorphological, and Biological Applications." *Hydrological Processes* 5 (1): 3–30.
- Naghibi, Seyed Amir, Kouros Ahmadi, and Alireza Daneshi. 2017. "Application of Support Vector Machine, Random Forest, and Genetic Algorithm Optimized Random Forest Models in Groundwater Potential Mapping." *Water Resources Management* 31 (9): 2761–2775.
- Nobre, Antonio Donato, Luz Adriana Cuartas, Marcos Rodrigo Momo, Dirceu Luís Severo, Adilson Pinheiro, and Carlos Afonso Nobre. 2016. "HAND Contour: A New Proxy Predictor of Inundation Extent." *Hydrological Processes* 30 (2): 320–333.
- North, Matthew A. 2009. "A Method for Implementing a Statistically Significant Number of Data Classes in the Jenks Algorithm." In *2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery*, 1:35–38. IEEE.
- Rahmati, Omid, Hamid Reza Pourghasemi, and Assefa M. Melesse. 2016. "Application of GIS-Based Data Driven Random Forest and Maximum Entropy Models for Groundwater Potential Mapping: A Case Study at Mehran Region, Iran." *Catena* 137: 360–372.
- Rennó, Camilo Daleles, Antonio Donato Nobre, Luz Adriana Cuartas, João Viane Soares, Martin G. Hodnett, and Javier Tomasella. 2008. "HAND, a New Terrain Descriptor Using SRTM-DEM: Mapping Terra-Firme Rainforest Environments in Amazonia." *Remote Sensing of Environment* 112 (9): 3469–3481.
- Riley, Shawn J., Stephen D. DeGloria, and Robert Elliot. 1999. "Index That Quantifies Topographic Heterogeneity." *Intermountain Journal of Sciences* 5 (1–4): 23–27.
- Schiermeier, Quirin. 2011. "Increased Flood Risk Linked to Global Warming: Likelihood of Extreme Rainfall May Have Been Doubled by Rising Greenhouse-Gas Levels." *Nature* 470 (7334): 316–317.
- Smith, Laurence C. 1997. "Satellite Remote Sensing of River Inundation Area, Stage, and Discharge: A Review." *Hydrological Processes* 11 (10): 1427–1439.
- Speckhann, Gustavo Andrei, Pedro Luiz Borges Chaffe, Roberto Fabris Goerl, Janete Josina de Abreu, and Juan Antonio Altamirano Flores. 2018. "Flood Hazard Mapping in Southern Brazil: A Combination of Flow Frequency Analysis and the HAND Model." *Hydrological Sciences Journal* 63 (1): 87–100.
- Strobl, Carolin, Anne-Laure Boulesteix, Achim Zeileis, and Torsten Hothorn. 2007. "Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution." *BMC Bioinformatics* 8 (1): 25.
- Tarboton, David G. 1997. "A New Method for the Determination of Flow Directions and Upslope Areas in Grid Digital Elevation Models." *Water Resources Research* 33 (2): 309–319.
- Tehrany, Mahyat Shafapour, Simon Jones, and Farzin Shabani. 2019. "Identifying the Essential Flood Conditioning Factors for Flood Prone Area Mapping Using Machine Learning Techniques." *Catena* 175: 174–192.
- Tehrany, Mahyat Shafapour, Mounq-Jin Lee, Biswajeet Pradhan, Mustafa Neamah Jebur, and Saro Lee. 2014. "Flood Susceptibility Mapping Using Integrated Bivariate and Multivariate Statistical Models." *Environmental Earth Sciences* 72 (10): 4001–4015.
- Tehrany, Mahyat Shafapour, Biswajeet Pradhan, and Mustafa Neamah Jebur. 2014. "Flood Susceptibility Mapping Using a Novel Ensemble Weights-of-Evidence and Support Vector Machine Models in GIS." *Journal of Hydrology* 512: 332–343.
- Teng, Jin, Anthony J. Jakeman, Jai Vaze, Barry FW Croke, Dushmanta Dutta, and S. Kim. 2017. "Flood Inundation Modelling: A Review of Methods, Recent Advances and Uncertainty Analysis." *Environmental Modelling & Software* 90: 201–216.
- Wang, Zhaoli, Chengguang Lai, Xiaohong Chen, Bing Yang, Shiwei Zhao, and Xiaoyan Bai. 2015. "Flood Hazard Risk Assessment Model Based on Random Forest." *Journal of Hydrology* 527: 1130–1141.
- Youssef, Ahmed Mohamed, Hamid Reza Pourghasemi, Zohre Sadat Pourtaghi, and Mohamed M. Al-Katheeri. 2016. "Landslide Susceptibility Mapping Using Random Forest, Boosted Regression Tree, Classification and Regression Tree, and General Linear Models and Comparison of Their Performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia." *Landslides* 13 (5): 839–856.