DYNAMIC FLOOD MAPPING USING HYDROLOGICAL MODELING AND MACHINE LEARNING

by

Morteza Esfandiari

B.Sc. Geomatics Engineering, Islamic Azad University of South Tehran Branch, 2017

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Science in Engineering

in the Graduate Academic Unit of Geodesy and Geomatics Engineering

Supervisors:	Shabnam Jabari, Ph.D., Geodesy and Geomatics Engineering
	David Coleman, Ph.D., Geodesy and Geomatics Engineering
Examining Board:	Robert Kingdon, Ph.D., Geodesy and Geomatics Engineering
	Heather McGrath, Ph.D., Geodesy and Geomatics Engineering
	Lloyd Waugh, Ph.D., Civil Engineering

This thesis is accepted by the
Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

July 2021

ABSTRACT

Flooding is one of the most devastating natural hazards around the globe. Having access to abundant sources of data such as Light Detection and Ranging (LiDAR) and satellite images in Geographic Information System (GIS), it is possible to estimate the geospatial extent of floods. Currently, Machine Learning plays an essential role in GIS applications and flood mapping. In this study, the aim was to provide a precise flood model by improving a hydrological model called Height Above Nearest Drainage (HAND) using one of the most robust machine learning algorithms, Random Forest (R.F.). In this study, first, the essential conditioning factors contributing to flooding were identified using optical satellite images as a reference. Then, using the most efficient conditioning factors, an R.F. classifier was trained to predict flooded areas with training data selected using the HAND model. However, since the HAND model has uncertainties in flood mapping, the Random Sample Consensus (RANSAC) paradigm is used along with the essential conditioning factors to remove outliers. Since the proposed method uses the HAND model predictions as pseudo training points, it is called flood mapping using Pseudo Supervised Random Forest (PS-RF). The accuracy of PS-RF for flood extent prediction was tested in 5 different flood events in Fredericton, NB, and one event in Ottawa, ON, confirming that PS-RF improves the flood mapping results of the HAND model without requiring any ground truth training data.

DEDICATION

To my family for their love and support.

ACKNOWLEDGEMENTS

I would like to extend sincere gratitude to all people who made this work achievable.

I am highly grateful to my supervisor Dr. Shabnam Jabari for her endless guidance and support throughout this work. Her understanding, encouragement, enthusiasm, and patience have provided a good basis for the present thesis. I would have been lost without her.

I am truly thankful to Dr. Heather McGrath for her time, guidance, and support during this research. She provided constructive lessons about hydrological models, encouragement, sound advice, great feedback, and lots of ideas. I also really appreciate Dr. David J. Coleman for his time and help in reviewing my papers and thesis.

I would also like to thank Dr. Ghasem Abdi who has always kindly offered his assistance and guidance.

I acknowledge the Natural Sciences and Engineering Research Council (NSERC) and the New Brunswick Innovation Foundation (NBIF) for financially supporting this research.

Many thanks to David Fraser, Greg Smith, Alicia Farnham, Lorry Hunt, and Michelle Ryan for their kind assistance all the time.

Table of Contents

List of Tables

List of Figures

List of Symbols, Nomenclature or Abbreviations

TWI – Terrain Wetness Index

TRI – Terrain Roughness Index

SPI – Stream Power Index

HAND – Height Above Nearest Drainage

NDWI – Normalized Difference Water Index

DTM – Digital Terrain Model

LiDAR – Light Detection and Ranging

RANSAC – Random Sample Consensus

TPR – True Positive Rate

TNR – True Negative Rate

FPR – False Positive Rate

FNR – False Negative Rate

KNN – K-Nearest Neighbors

ANN – Artificial Neural Networks

SVM – Support Vector Machines

R.F. – Random Forest

GP – Genetic Programming

FR – Frequency Ratio

LR – Logistic Regression

GIS – Geographic Information System

PS-RF – Pseudo Supervised Random Forest

TOA – Top of the Atmosphere Correction

HRDM – High Resolution Digital Elevation Model

TauDEM – Terrain Analysis Using Digital Elevation Models

MCC – Matthews Correlation Coefficient

T.P. – True Positive

T.N. – True Negative

F.P. – False Positive

F.N. – False Negative

Chapter 1: INTRODUCTION

This chapter provides background information on recent flood mapping techniques, challenges in flood mapping, objectives of this research, the proposed methodology for accurate, dynamic, and efficient flood mapping, and explains the structure of this thesis.

1.1 Research Background

Flooding is one of the most devastating phenomena and its occurrence is drastically increasing as a consequence of global warming and climate change (Schiermeier 2011; Gaur, Gaur, and Simonovic 2018). There are different types of flood, including coastal flood, flash flood, and river flood -- in which one of the major challenges is to provide an accurate estimation of flood extent and amount of damage to surrounding areas based on flood level. Early warning capabilities are necessary in order to prevent any sort of human casualties or intensive damage to sensitive urban infrastructure.

The complexity of water flow and the importance of flood maps has brought many researchers from different fields to employ various techniques for estimating flood behavior (Paprotny et al. 2020; Mosavi, Ozturk, and Chau 2018; Montero et al. 2013). In general, the approaches for flood mapping can be classified into 5 categories:

- Hydrodynamic models
- Empirical methods
- Simplified conceptual methods
- Machine learning approaches
- Flood conditioning factors

These approaches are discussed in the following sections.

1.1.1 Hydrodynamic models

Traditional hydrodynamic models are mainly divided into one-, two-, and three-dimensional categories. They use intricate mathematical equations to simulate fluid motion. Depending on the topography of a region and the required level of accuracy, different models could be selected for demonstrating flood extent and damage in affected areas.

One-dimensional (1D) models, which are the simplest hydrodynamic models, are generated by solving mass and momentum conservation equations between two cross-sections (Teng et al. 2017a). Multiple parameters are required to solve these equations, including flow discharge, average velocity, time, water depth, gravitational acceleration, friction slope, and channel bed slope. However, 1D models are unable to estimate flood distribution due to lateral flood waves, and they only consider the topography between two cross-sections instead of a continuous surface.

Two-dimensional (2D) models are generated by solving Shallow Water Equations. In 2D models, the floodplain is represented in a two-dimensional plane assuming that the third dimension, the water depth, is shallow in comparison with other dimensions. 2D models are extensively used in flood mapping and it is possible to estimate inundated extent with high accuracy, but the fact that they are computationally expensive cannot be neglected.

Three-dimensional (3D) models were developed to better address the effect of flood on vertical features (Ye and McCorquodale 1998a). These models are mainly derived from three-dimensional Navier-Stokes equations which describe the motion of fluid substances. 3D models are generally considered to be unnecessary since the required parameters are costly and hard to obtain, and 2D models can provide sufficient accuracy most of the time.

However, in many case studies, when the water motion is intensive and abnormal due to catastrophic floods such as tsunami, flash floods, and dam breaks, 3D models could be used to model vortices, vertical turbulence, and spiral flow (Teng et al. 2017a).

1.1.2 Empirical methods

Empirical methods are mainly based on "observations" in which the flood extents are derived from different types of data such as on-ground measurements, satellite imagery, aerial photography, and even public interviews. The quality of these observations in terms of accuracy and spatial resolution contribute significantly to the quality of final flood maps. However, they are reliant on the time of the data acquisition (often the date and time of a given satellite pass does not coincide with peak flood). One example of empirical methods could be applying remote sensing techniques (Smith 1997) to extract flood extent from satellite imagery. Empirical methods are also commonly used for accuracy assessment purposes and validation of other models.

1.1.3 Simplified conceptual methods

Another group of popular approaches for flood mapping is based on simplified conceptual models. These models are not as detailed as hydrodynamic models and require less data for flood simulation, but have nevertheless been able to obtain precise results in many case studies (Teng et al. 2017a). One notable example of these simplified models is Height Above Nearest Drainage (HAND). The HAND model uses a hydrologically conditioned digital terrain model and normalizes the elevation of each specific cell of a raster dataset in relation to the closest drainage feature (Nobre et al. 2011). Then, the extent of the flood is estimated based on the river water level. Therefore, cells with values equal to or lower than the river level will be considered as flooded, and an approximate depth at each cell will be acquired (Nobre et al. 2016a). Two main parameters that highly contribute to the model's correctness are the stream network, which is usually derived from the DTM, and gauge readings at that specific stream. Due to the simplified nature of the HAND model, external factors such as a hydrologically conditioned DTM, precise water level, and even natural features can significantly affect the outcomes. For example, in a study carried out by McGrath (McGrath et al. 2018a), the HAND model generated for rapid flood inundation mapping was able to accurately detect inundated areas compared to the actual flood extent. However, in (Nobre et al. 2016b), the generated HAND model has performed poorly due to the influence of an existing waterfall in the area which affects the flood pressure. Therefore, the generated model was not able to precisely map the inundated areas in the vicinity of the waterfall, whilst there was significant overestimation in flat areas.

1.1.4 Machine learning approaches

Different types of machine learning and statistical methods have recently been used in this field as well to generate flood susceptibility maps. Machine learning models consider the effects of hydrological, geological, and topographical factors that contribute to flooding and analyze the feature values of several ground truth points from previous flood events to identify flood-prone areas (Tehrany et al., 2019). These factors are referred to as flood conditioning factors, and different machine learning models such as Artificial Neural Networks (ANNs) (Falah et al. 2019a), Support Vector Machines (SVM) (Tehrany et al. 2015), and Random Forest (S. Lee et al. 2017) have been applied over different datasets for generating flood susceptibility maps.

The first efforts of this research focused on investigating the effectiveness of various conditioning factors used in machine learning models that contribute to flooding and identifying the most important ones that provide the most accurate estimations. Once multiple different scenarios were tested, and the most important conditioning factors were identified, a new method for flood mapping was generated by combining the HAND model and Random Forest (R.F.) called Pseudo Supervised Random Forest (PS-RF). PS-RF takes advantage of both approaches and generates more accurate flood maps than the HAND model without requiring ground truth samples for training the R.F. model.

1.1.5 Flood conditioning factors

Many spatial methods have also been proposed by scientists for flood hazard and risk mapping and identifying flood-prone areas. Flood hazard maps require several flood-related parameters (Tehrany et al. 2014). These flood-related parameters, more commonly referred to as *flood conditioning factors*, should be optimized and properly selected to achieve reliable results. For example, soil type affects the degree of water infiltration, impacts surface run-off generation, and consequently influences the flooding process (Rahmati et al. 2016). Different composition of the surface soils also considerably affects the infiltration of water which results in the level of water movement permitted by the soil type (Pla et al. 2017). As reported by (Huang, Wu, and Zhao 2013), the degree of infiltration has a direct impact on the levels of surface erosion and run-off, as well as the recharging of soil water content and groundwater levels.

1.2 Research Challenges

During this research, several challenges have been dealt with which can be summarized as:

- **Feature Selection**: Despite the capability of machine learning models to analyze extensive amounts of data, these models do not always result in better decision making. Correlated or irrelevant features, in our case flood conditioning factors, can lead to misclassification and accordingly degrade the performance of machine learning. Feature selection is also a function of the machine learning algorithm used.
- **Reference Ground Truth Data**: Training data is essential for machine learning methods. Using machine learning for flood mapping requires dynamic Flooded and Not Flooded ground truth samples for training the model so it can distinguish the extent of flood dynamically. However, in the majority of recent research studies, multiple past flood events were used as input data into a machine learning model and a general flood susceptibility map of a region was generated. Hence, machine learning cannot still be used for flood forecasting and prediction. Also, possible future flood events with unprecedented water levels cannot be estimated.
- **Integration of a hydrological and a machine learning model**: Despite the capability of hydrological models in producing dynamic flood maps, including the HAND model, they have limited accuracy.

1.3 Research Objectives

The main objective of this study is to provide a flood mapping system that can be used to generate flood maps with any given water level. It is expected to have a dynamic system which can be used for both generating flood events that occurred in the past, and possible future floods with unprecedented water levels. The efficiency of this system in terms of required data and processing time is significantly important, as well as its accuracy when compared to ground truth observations.

To achieve this goal, this study attempts to:

- Use machine learning to dynamically estimate the inundated areas based on the river water depth.
- Improve the accuracy of the HAND model in flood mapping using machine learning (R.F. model).
- Identify the essential conditioning factors (feature selection) that contribute to flood mapping using R.F. model as a classifier.
- Evaluate the performance of the dynamic flood mapping model in different study areas.

1.4 Overview of the Chapters

This is an article-based thesis comprising the following peer-reviewed articles:

- Article 1 (Published)

  Esfandiari, M., S. Jabari, H. McGrath, and D. Coleman. "FLOOD MAPPING USING RANDOM FOREST AND IDENTIFYING THE ESSENTIAL CONDITIONING FACTORS; A CASE STUDY IN FREDERICTON, NEW BRUNSWICK, CANADA." *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 3 (2020): 609-615.

- Article 2 (Published)

  Esfandiari, M., Abdi, G., Jabari, S., McGrath, H. and Coleman, D., 2020. Flood Hazard Risk Mapping Using a Pseudo Supervised Random Forest. *Remote Sensing*, 12(19), p.3206.

Chapter 2 gives an overview of the essential conditioning factors that contribute to flooding. The effectiveness of twelve conditioning factors of Altitude, Slope, Aspect, Distance from River, Land use/cover, TWI, TRI, SPI, Curvature, Plan Curvature, Profile Curvature, and HAND was investigated through implementing R.F. over different combinations of these conditioning factors. The scenarios were validated against the 2018 flood event that occurred in Fredericton, New Brunswick, Canada.

Chapter 3 presents the integration of the HAND model as a hydrological model and R.F. for dynamic flood mapping. The HAND model was generated to provide primary labels for R.F. Then, the RANSAC algorithm was used to identify outliers from the HAND model, and the best training subset of modeling R.F. was identified. After that, the hyperparameters of R.F. with the best training subset were then adjusted through cross-validation. Then, the model was used to predict inundated areas in five flood events in 2014, 2016, 2018, and 2019 in Fredericton, and in 2019 in Gatineau.

In chapter 4, the contribution of this study is summarized, a comparison between the flood maps released by GeoNB and the proposed method in this study is carried out, and some recommendations for future research are provided.

1.5 References

Falah, F., Rahmati, O., Rostami, M., Ahmadisharaf, E., Daliakopoulos, I. N., & Pourghasemi, H. R. (2019). Artificial neural networks for flood susceptibility mapping in data-scarce urban areas. In Spatial modeling in GIS and R for Earth and Environmental Sciences (pp. 323-336). Elsevier.

Gaur, A., Gaur, A., & Simonovic, S. P. (2018). Future changes in flood hazards across Canada under a changing climate. Water, 10(10), 1441.

Huang, J., Wu, P., & Zhao, X. (2013). Effects of rainfall intensity, underlying surface and slope gradient on soil infiltration under simulated rainfall experiments. Catena, 104, 93-102.

Lee, S., Kim, J. C., Jung, H. S., Lee, M. J., & Lee, S. (2017). Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. Geomatics, Natural Hazards and Risk, 8(2), 1185-1203.

McGrath, H., Bourgon, J. F., Proulx-Bourque, J. S., Nastev, M., & El Ezz, A. A. (2018). A comparison of simplified conceptual models for rapid web-based flood inundation mapping. Natural Hazards, 93(2), 905-920.

Montero, R. A., Schwanenberg, D., Hatz, M., & Brinkmann, M. (2013). Simplified hydraulic modelling in model predictive control of flood mitigation measures along rivers. Journal of Applied Water Engineering and Research, 1(1), 17-27.

Mosavi, A., Ozturk, P., & Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. Water, 10(11), 1536.

Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A., & Saleska, S. (2011). Height Above the Nearest Drainage–a hydrologically relevant new terrain model. Journal of Hydrology, 404(1-2), 13-29.

Nobre, A. D., Cuartas, L. A., Momo, M. R., Severo, D. L., Pinheiro, A., & Nobre, C. A. (2016). HAND contour: a new proxy predictor of inundation extent. Hydrological Processes, 30(2), 320-333.

Paprotny, D., Vousdoukas, M. I., Morales-Nápoles, O., Jonkman, S. N., & Feyen, L. (2020). Pan-European hydrodynamic models and their ability to identify compound floods. Natural Hazards, 1-25.

Pla, C., Cuezva, S., Martinez-Martinez, J., Fernandez-Cortes, A., Garcia-Anton, E., Fusi, N., ... & Benavente, D. (2017). Role of soil pore structure in water infiltration and CO2 exchange between the atmosphere and underground air in the vadose zone: A combined laboratory and field approach. Catena, 149, 402-416.

Rahmati, O., Pourghasemi, H. R., & Zeinivand, H. (2016). Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. Geocarto International, 31(1), 42-70.

Schiermeier, Q. (2011). Increased flood risk linked to global warming.

Smith, L. C. (1997). Satellite remote sensing of river inundation area, stage, and discharge: A review. Hydrological processes, 11(10), 1427-1439.

Tehrany, M. S., Jones, S., & Shabani, F. (2019). Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. Catena, 175, 174-192.

Tehrany, M. S., Lee, M. J., Pradhan, B., Jebur, M. N., & Lee, S. (2014). Flood susceptibility mapping using integrated bivariate and multivariate statistical models. Environmental earth sciences, 72(10), 4001-4015.

Tehrany, M. S., Pradhan, B., Mansor, S., & Ahmad, N. (2015). Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. Catena, 125, 91-101.

Teng, J., Jakeman, A. J., Vaze, J., Croke, B. F., Dutta, D., & Kim, S. (2017). Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. Environmental Modelling & Software, 90, 201-216.

Ye, J., & McCorquodale, J. A. (1998). Simulation of curved open channel flows by 3D hydrodynamic model. Journal of Hydraulic Engineering, 124(7), 687-698.

Chapter 2: Flood Mapping Using Random Forest and Identifying the Essential Conditioning Factors; A Case Study in Fredericton, NB, Canada

This chapter introduces the essential conditioning factors that contribute to flood mapping which were obtained through the ranking capability of the Random Forest algorithm. The main focus is on generating flood maps with different combinations of conditioning factors and selecting the ones with the highest degree of impact. For this research, the 2018 flood event in Fredericton, New Brunswick, Canada was considered as a case study, and twelve different conditioning factors were selected for analysis. It was discussed that a higher number of conditioning factors does not necessarily increase the accuracy of flood maps, and the highest accuracy is obtained by only five conditioning factors.

The chapter is published as a peer-reviewed conference proceedings paper:

Esfandiari, M., Jabari, S., McGrath, H. and Coleman, D., 2020. FLOOD MAPPING USING RANDOM FOREST AND IDENTIFYING THE ESSENTIAL CONDITIONING FACTORS; A CASE STUDY IN FREDERICTON, NEW BRUNSWICK, CANADA. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, *5*(3).

Minor modifications are applied to the original text for the notation, terminology, and formating to be consistent with the rest of the thesis. Footnotes are added to further explain some details in the text.

Abstract

Flood is one of the most damaging natural hazards in urban areas in many places around the world. Recently, the city of Fredericton, New Brunswick, Canada 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.

2.1 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. 2017a).

(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 flood plain, and the required level of accuracy, different models can be selected for demonstrating flood damage and extent in affected areas (Teng et al. 2017a).

(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. 2017a; Momo 2014; Liu et al. 2016; Speckhann et al. 2018).

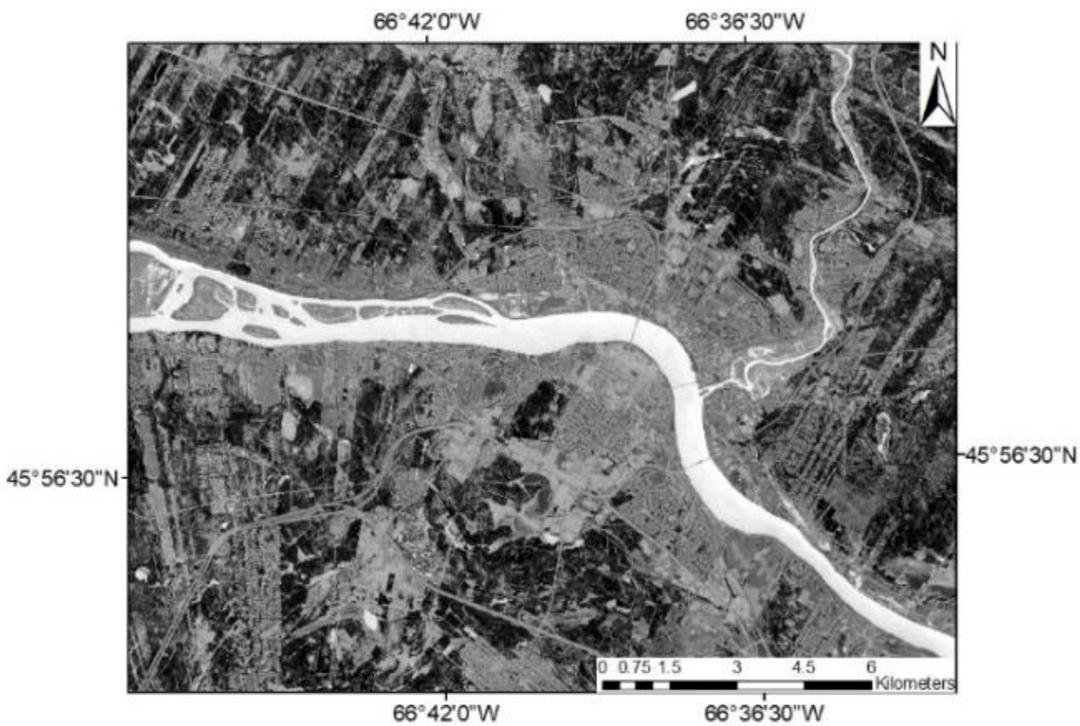
(3) *Empirical methods* include all the flood maps which are generated using *observations*. These *observations* can be made from satellite images, aerial photographs, surveying, etc. The accuracy of the flood maps generated by the *observations* is totally dependent on the accuracy of the observations, 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, 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, results and discussion, and the conclusion are presented in sections 3, 4, and 5, respectively.

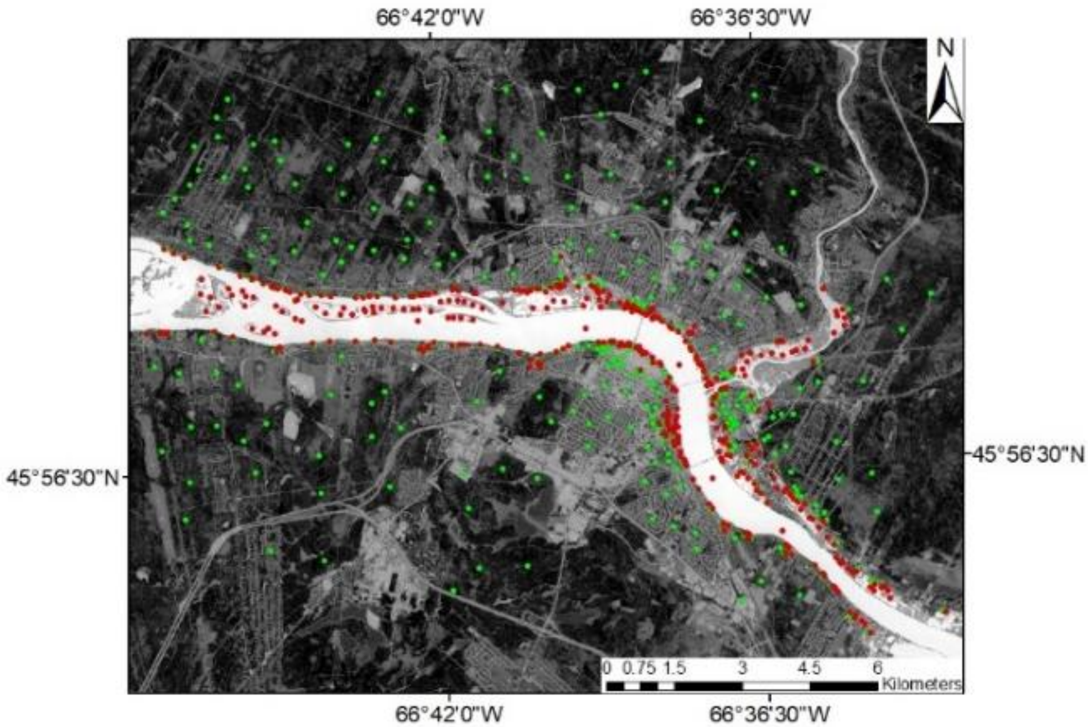
2.2 Study Area and Dataset

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 and surrounding snow cover 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 at Fredericton 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. Located approximately 19 kilometres upstream from the city, the Mactaquac hydroelectric dam controls a relatively small headpond and is not able to hold the melted ice for long so the water must be released to the river.



(a)

Figure 2-1 Cont



(b)

Figure 2-1 The generated NDWIs representing the Fredericton area before flood (a) and the collected sample points at the flood peak (b). Both images are from Sentinel-2 Satellite captured on May 02, 2018.

2.3 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 being "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).

2.3.1 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 experts' analysis and the information from the literature (Kia et al. 2012). A total of 12 conditioning factors, all shown in Figure 2-2 and 2-3, 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, Grass 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[1] 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[2] which is given in (Riley, DeGloria, and Elliot 1999). Stream's erosion is measured by SPI[3] 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 1997a) and the elevation difference of each pixel (Nobre et al. 2016a). The other conditioning factors are the

---

[1] TWI = ln $(A_s / \tan \beta)$ where A_s represents the area of catchment (m^2) and β (radians) the gradient of the slope

[2] TRI = Y $\left[\sum (x_{ij} - x_{00})^2\right]^{1/2}$ where x_{ij}= elevation of each neighbor cell to cell (0,0)

[3] SPI = $A_s \tan \beta$

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 2013a). 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^2).

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).

2.3.2 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, Thieken, 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 2-1) and flood-peak (Figure 2-2.) images were used, which were taken on April 22nd, 2018 and May 2nd, 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.

2.3.3 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.

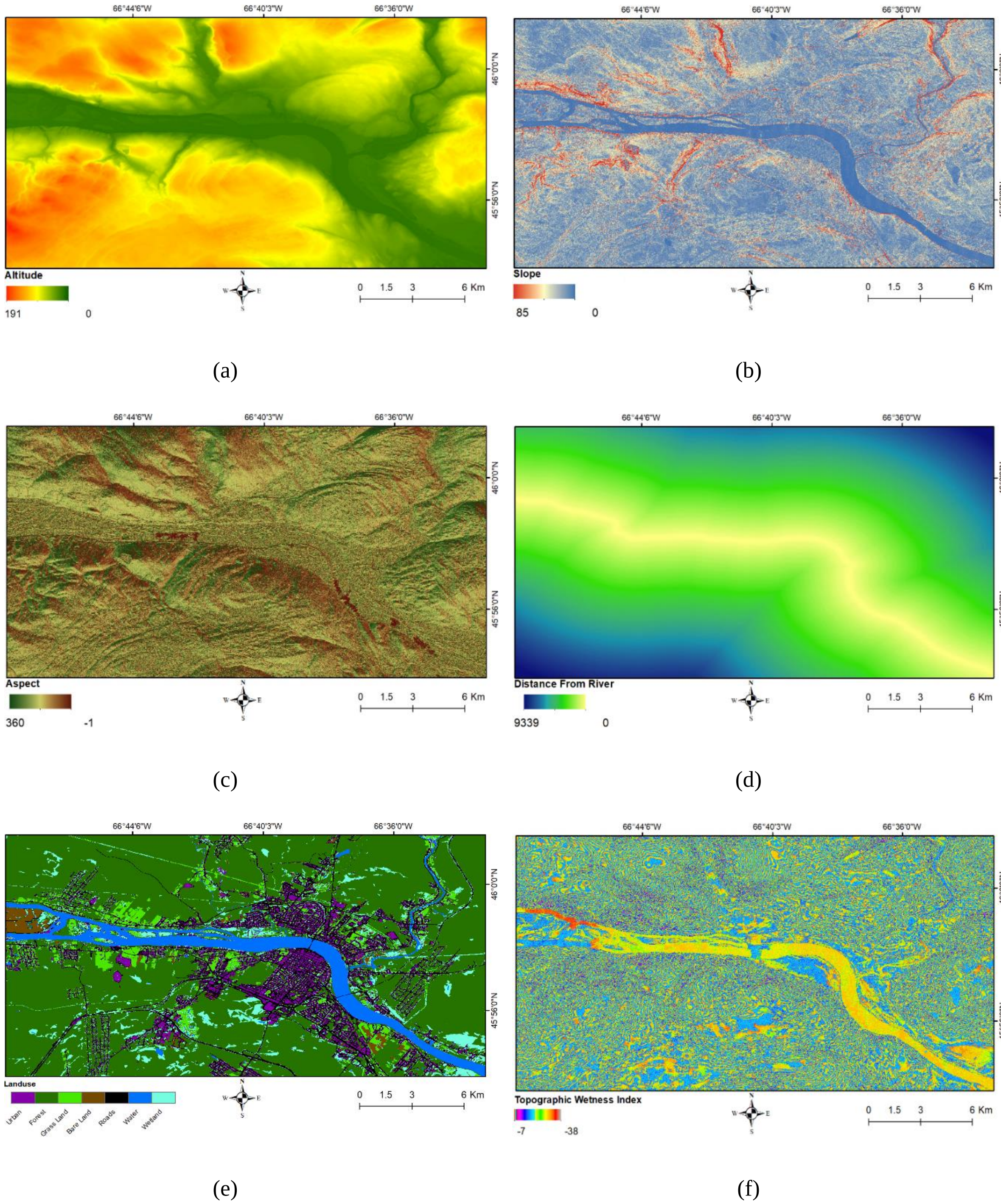


Figure 2-2 Conditioning factors used in the Random Forest classifier (part 1): (a) Altitude, (b) Slope, (c) Aspect, (d) Distance from River, (e) Land use/cover, (f) TWI

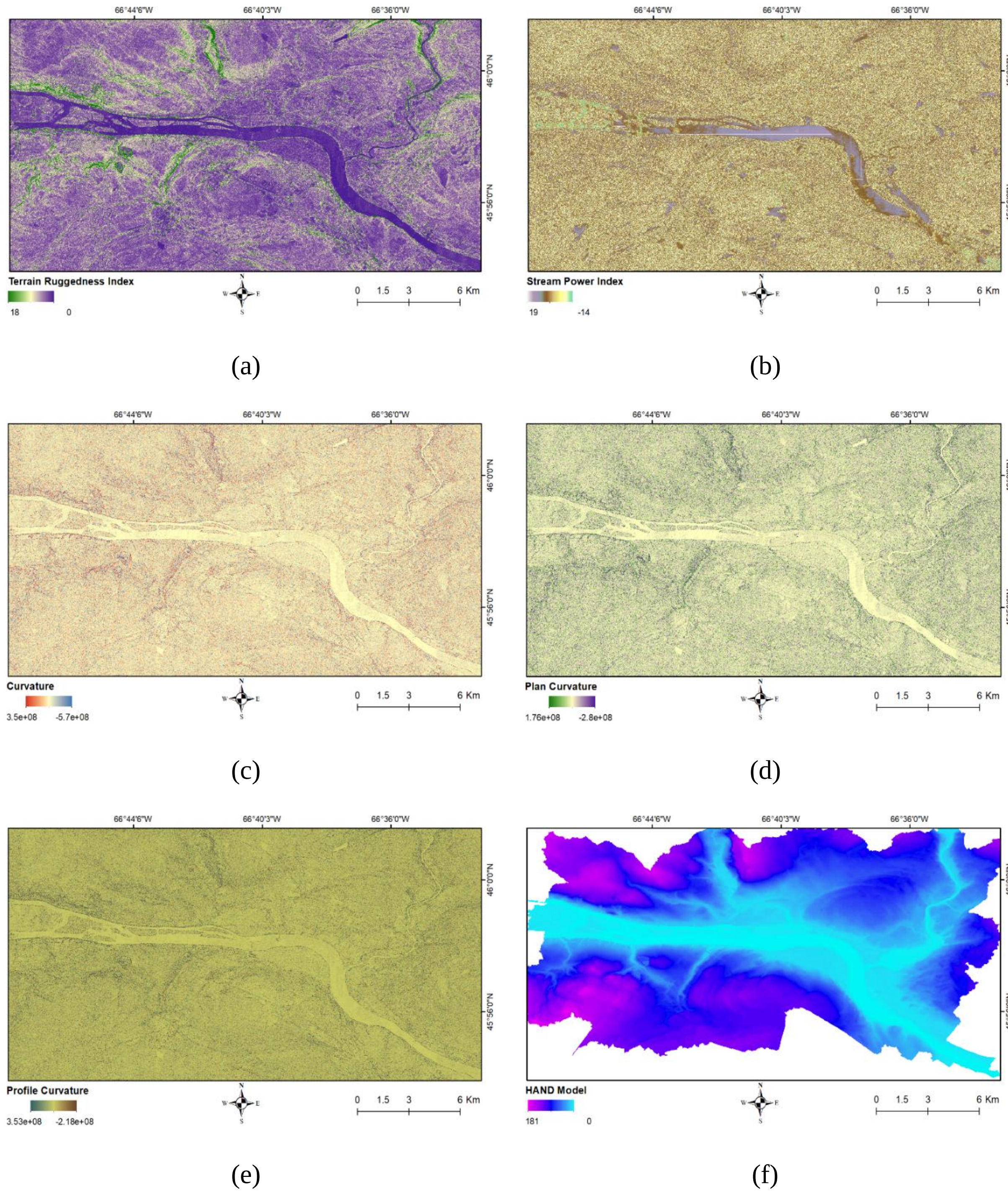


Figure 2-3 Conditioning factors used in the Random Forest classifier (part 2): (a) TRI, (b) SPI, (c) Curvature, (d) Plan Curvature, (e) Profile Curvature, (f) HAND model

Table 2-1 Combinations of different conditioning factors implemented in Scenario 1

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

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 2-4). 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 2-1 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 2-2), 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.

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

Scenario 2	Conditioning factors
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

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.

2.4 Results and 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 2-1 and 2-2.

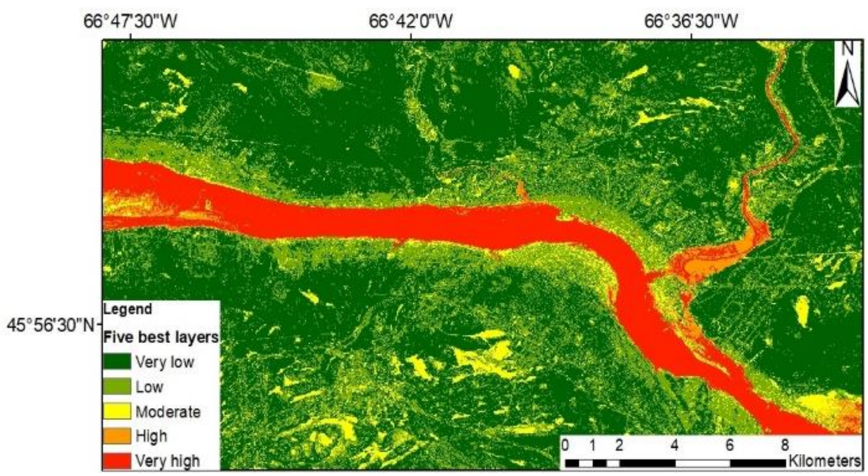


Figure 2-4 Random Forest output using altitude, slope, aspect, distance from river, and land use/cover. The very high class represents areas that are flooded with the highest probability, and very low class represent areas with the least probability of being flooded.

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 2-5, 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 *might* be that, since some conditioning factors were correlated, they could collectively degrade the accuracy. In Figure 2-6, 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 rive and land-use/cover were used.

Since correlated conditioning factors could negatively affect the accuracy, we ensured in Scenario 2 that only uncorrelated conditioning factors were 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 2-7) 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.

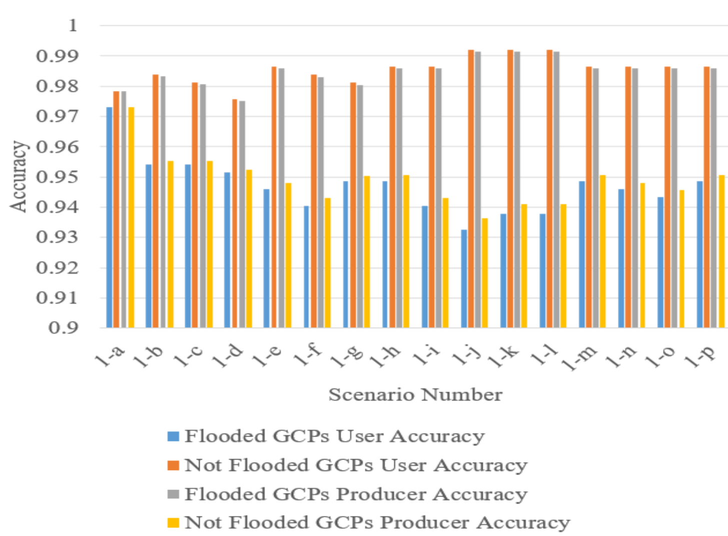


Figure 2-5 User and producer accuracy of flooded and not-flooded sample points in the first series of implementations

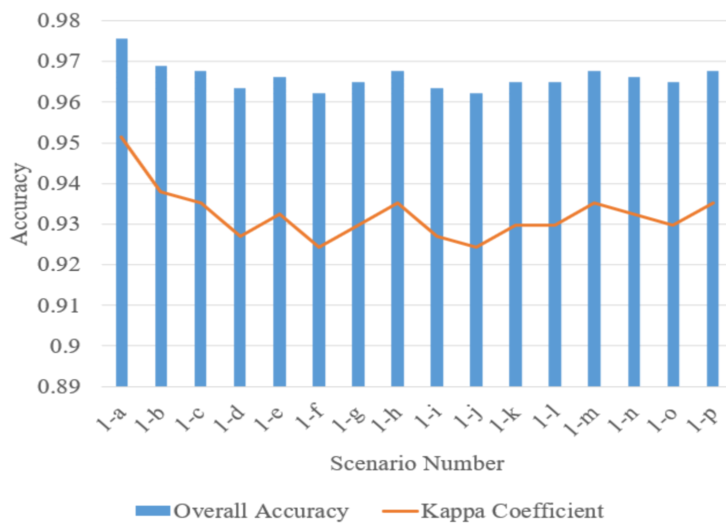


Figure 2-6 Overall accuracy and Kappa Coefficient of the first series of implementations

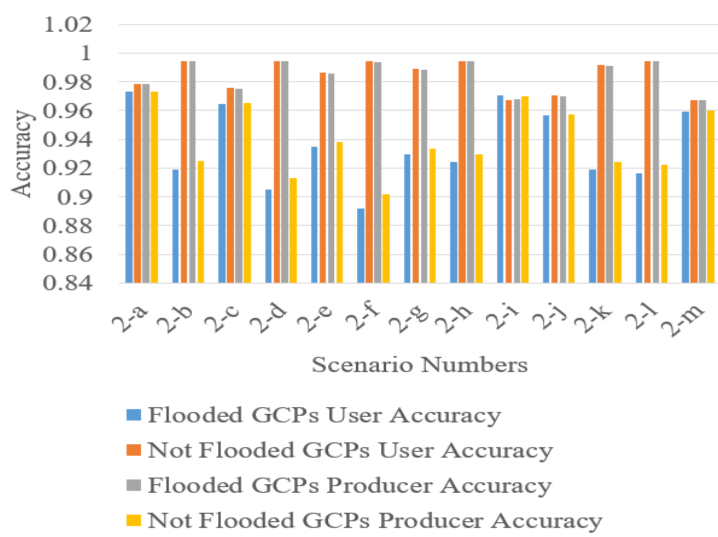


Figure 2-7 User and producer accuracy of flooded and not-flooded sample points of the second series of implementations

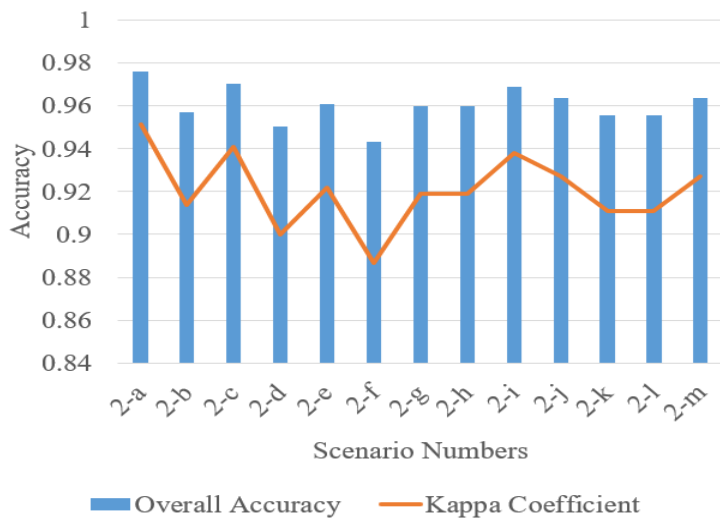


Figure 2-8 Overall accuracy and Kappa Coefficient of the second series of implementations

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.

2.5 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 conditioning factors decreases the accuracy of predictions.
- HAND and altitude layers are both major factors in flooding having similar effects on the final accuracy.

2.6 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, K. J., & Kirkby, M. J. (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, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Campbell, J. B., & Wynne, R. H. (2011). Introduction to remote sensing. Guilford Press.

Catani, F., Lagomarsino, D., Segoni, S., & Tofani, V. (2013). Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. Natural Hazards and Earth System Sciences, 13(11), 2815-2831.

Gao, B. C. (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, A., Gaur, A., & Simonovic, S. P. (2018). Future changes in flood hazards across Canada under a changing climate. Water, 10(10), 1441.

Heerdegen, R. G., & Beran, M. A. (1982). Quantifying source areas through land surface curvature and shape. Journal of Hydrology, 57(3-4), 359-373.

Ho, T. K. (1995, August). Random decision forests. In Proceedings of 3rd international conference on document analysis and recognition (Vol. 1, pp. 278-282). IEEE.

Ihsan, Z., Idris, M. Y., & Abdullah, A. H. (2013). Attribute normalization techniques and performance of intrusion classifiers: A comparative analysis. Life Science Journal, 10(4), 2568-2576.

Kia, M. B., Pirasteh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A., & Moradi, A. (2012). An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. Environmental Earth Sciences, 67(1), 251-264.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R news, 2(3), 18-22.

Liu, Y. Y., Maidment, D. R., Tarboton, D. G., Zheng, X., Yildirim, A., Sazib, N. S., & Wang, S. (2016). A CyberGIS approach to generating high-resolution height above nearest drainage (HAND) raster for national flood mapping.

Merz, B., Thieken, A. H., & Gocht, M. (2007). Flood risk mapping at the local scale: concepts and challenges. In Flood risk management in Europe (pp. 231-251). Springer, Dordrecht.

Momo, M. R., Pinheiro, A., Severo, D. L., Cuartas, L. A., & Nobre, A. D. (2016). Desempenho do modelo HAND no mapeamento de áreas suscetíveis à inundação usando dados de alta resolução espacial. RBRH, 21(1), 200-208.

Moore, I. D., Grayson, R. B., & Ladson, A. R. (1991). Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. Hydrological processes, 5(1), 3-30.

Nobre, A. D., Cuartas, L. A., Momo, M. R., Severo, D. L., Pinheiro, A., & Nobre, C. A. (2016). HAND contour: a new proxy predictor of inundation extent. Hydrological Processes, 30(2), 320-333.

North, M. A. (2009, August). 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 (Vol. 1, pp. 35-38). IEEE.

Rahmati, O., Pourghasemi, H. R., & Melesse, A. M. (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, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., & Tomasella, J. (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, S. J., DeGloria, S. D., & Elliot, R. (1999). Index that quantifies topographic heterogeneity. intermountain Journal of sciences, 5(1-4), 23-27.

Schiermeier, Q. (2011). Increased flood risk linked to global warming.

Speckhann, G. A., Borges Chaffe, P. L., Fabris Goerl, R., Abreu, J. J. D., & Altamirano Flores, J. A. (2018). Flood hazard mapping in Southern Brazil: a combination of flow frequency analysis and the HAND model. Hydrological Sciences Journal, 63(1), 87-100.

Tarboton, D. 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, M. S., Jones, S., & Shabani, F. (2019). Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. Catena, 175, 174-192.

Tehrany, M. S., Lee, M. J., Pradhan, B., Jebur, M. N., & Lee, S. (2014). Flood susceptibility mapping using integrated bivariate and multivariate statistical models. Environmental earth sciences, 72(10), 4001-4015.

Tehrany, M. S., Pradhan, B., & Jebur, M. N. (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, J., Jakeman, A. J., Vaze, J., Croke, B. F., Dutta, D., & Kim, S. (2017). Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. Environmental Modelling & Software, 90, 201-216.

Youssef, A. M., Pourghasemi, H. R., Pourtaghi, Z. S., & Al-Katheeri, M. M. (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.

Chapter 3: Flood Hazard Risk Mapping Using a Pseudo Supervised Random Forest

This chapter introduces a novel flood mapping procedure through the combination of hydrological and machine learning approaches. This study aims to develop a dynamic flood mapping system that can efficiently generate accurate flood maps. By combining the Height Above Nearest Drainage (HAND) hydrological model and Random Forest algorithm, the required training data for Random Forest is obtained and more accurate flood maps compared to the HAND model are generated. It is also discussed that Random Sample Consensus (RANSAC) was used as an outlier detection method to maintain the most reliable training data for Random Forest.

The chapter is published as a peer-review journal article:

Esfandiari, M., Abdi, G., Jabari, S., McGrath, H. and Coleman, D., 2020. Flood Hazard Risk Mapping Using a Pseudo Supervised Random Forest. *Remote Sensing*, *12*(19), p.3206. https://doi.org/10.3390/rs12193206

Minor modifications are applied to the original text for the notation, terminology, and formatting to be consistent with the rest of the thesis. Footnotes are added to further explain some details in the text.

Abstract

Devastating floods occur regularly around the world. Recently, machine learning models have been used for flood susceptibility mapping. However, even when these algorithms are provided with adequate ground truth training samples, they can fail to predict flood extends reliably. On the other hand, the height above nearest drainage (HAND) model can produce flood prediction maps with limited accuracy. The objective of this research is to produce an accurate and dynamic flood modeling technique to produce flood maps as a function of water level by combining the HAND model and machine learning. In this paper, the HAND model was utilized to generate a preliminary flood map; then, the predictions of the HAND model were used to produce pseudo training samples for a R.F. model. To improve the R.F. training stage, five of the most effective flood mapping conditioning factors are used, namely, Altitude, Slope, Aspect, Distance from River and Land use/cover map. In this approach, the R.F. model is trained to dynamically estimate the flood extent with the pseudo training points acquired from the HAND model. However, due to the limited accuracy of the HAND model, a random sample consensus (RANSAC) method was used to detect outliers. The accuracy of the proposed model for flood extent prediction, was tested on different flood events in the city of Fredericton, NB, Canada in 2014, 2016, 2018, 2019. Furthermore, to ensure that the proposed model can produce accurate flood maps in other areas as well, it was also tested on the 2019 flood in Gatineau, QC, Canada. Accuracy assessment metrics, such as overall accuracy, Cohen's kappa coefficient, Matthews correlation coefficient, true positive rate (TPR), true negative rate (TNR), false positive rate (FPR) and false negative rate (FNR), were used to compare the predicted flood extent of the study areas, to the extent estimated by the HAND model and the extent imaged by

Sentinel-2 and Landsat satellites. The results confirm that the proposed model can improve the flood extent prediction of the HAND model without using any ground truth training data.

3.1 Introduction

Floods are a devastating natural hazard but are inadequately understood (Tehrany et al. 2014) and controlled (Thistlethwaite et al. 2018). Floods cause severe damage to people, their health and properties, city infrastructure, ecological systems, agricultural lands and economic activities (Messner and Meyer 2006; Ghoneim and Foody 2013; Nicholls et al. 2015). Global warming, urbanization along rivers and coasts and climate change are causing rapid growth in flood events around the world (Gaur, Gaur, and Simonovic 2018; Schiermeier 2011) and it is crucial for the public, the emergency management community and decision-makers to have access to an accurate estimation of the flood extent. There are different approaches for flood prediction including (1) hydrodynamic models (Brunner 2016; DHI 2003; 2012; Moulinec et al. 2011; Prakash, Rothauge, and Cleary 2014; Vacondio et al. 2012); (2) empirical methods (Smith 1997; Schumann et al. 2009); (3) simplified conceptual models (Lhomme et al. 2008; Nobre et al. 2011); and more recently (4) machine learning and statistical models (Tehrany, Pradhan, and Jebur 2014; Rahmati, Pourghasemi, and Melesse 2016; Zhao et al. 2019; Pourghasemi et al. 2020).

(1) Hydrodynamic models use mathematical equations to simulate fluid motion and usually require high computational power (Teng et al. 2017b). These models require significant data inputs such as flow discharge, water depth, gravitational acceleration and channel bed slope, to simulate flow in their attempt to replicate flow patterns and estimate flow velocity and spatial extents of inundation. Hydrodynamic models are mainly divided into three

categories depending on their dimensionality, 1D, 2D and 3D. 1D models consider the floodplain flow as one-dimensional along the river central line (Brunner 2016; DHI 2003) and by solving mass and momentum conservation equations, model the flood over a floodplain on multiple cross-sections. In 2D models, an intensive collection of topographical data are used for flood simulation assuming that the third dimension, water depth, is shallow (Teng et al. 2017b). 2D models are probably the most common approach among hydrodynamic models for flood mapping (Teng et al. 2017b). Furthermore, 1D/2D solutions are becoming very popular, where the 1D component models flow in the river and 2D flow modeled on the floodplain (Brunner et al. 2015). 3D hydrodynamic models generate a complex three-dimensional representation of floodplain which is usually regarded unnecessary because of their extensive required parameters and maybe comparable accuracy with 2D models (Teng et al. 2017b). For modeling more complicated flood events, which occur due to tsunami and dam breaks and require modeling of vertical turbulence and spiral flow, 3D models provide better representation because of their capability in modeling vertical dimensions (Monaghan 1994; Ye and McCorquodale 1998b). The higher dimensionality of a model, the more parameters required for flood simulation and longer computation time. In general, hydrodynamic models are extremely complex, have sensitive outputs based on site-specific parameters, require a substantial amount of data and their calibration is costly (Sidrane et al. 2019).

(2) Empirical methods generate flood maps mainly using observations which could be obtained through satellite imagery, aerial photography, land survey and so forth. Depending on the credibility and resolution of the observations, empirical methods generate flood maps with varying accuracy and are mainly used as a benchmark for

modeling and assessment of other approaches (Smith 1997). For example, using remote sensing techniques it is possible to map areas covered by water and through a pre- and post-event comparison, affected areas will be identified. However, there are certain limitations involved with this approach such as the often coarse spatial resolution of freely available satellite images, availability of the data at the time of the flood, high cost of data collection through surveying and various types of observations which require knowledge for interpretation (Teng et al. 2017b).

(3) Simplified conceptual models generally involve less physical detail than hydrodynamic models and are based on simplified hydraulic concepts. One example of these types of models is Rapid Flood Spreading Method (RFSM) which separates the floodplain into distinct areas representing the depressions. Then, based on the flood volume and filling/spilling process, inundated areas are identified (Lhomme et al. 2008) . Teng in Reference (Teng et al. 2017b) categorized the height above nearest drainage (HAND) model (Nobre et al. 2011) in the simplified conceptual class, which is essentially an elevation model normalized toward the nearest stream (Rennó et al. 2008). In the HAND model, the elevation of each pixel is calculated from the nearest stream based on a series of steps including producing flow direction, accumulation area and so forth. Then, by considering the water level in the stream, the inundated areas can be recognized. This capability of the HAND model can be used to predict inundated areas as a function of water level in the nearby stream. Different hydrological models can estimate the amount of water in the streams each year and with the HAND model it is possible to estimate inundated areas and manage further damages. However, depending on the topography of the region of interest, existence of natural or artificial infrastructures that affect the water discharge

like waterfalls or dams, the HAND model can produce different accuracies. In the studies reported by Momo (Momo 2014) and McGrath et al. (Heather McGrath et al. 2018b), the HAND model was able to obtain flood extents that compared closely to actual flooded areas; however, in (Nobre et al. 2016a), the generated HAND model considerably overestimated the real flooded extent. Despite the capability of the HAND model to predict the flood extent and its straightforward approach, the model is highly susceptible to the accuracy of hydrologically conditioned Digital Terrain Models (DTM). Due to its dependence on the external factors, such as precise water level of the river and a hydrologically conditioned DTM, the HAND model can predict flood extents with limited accuracy.

(4) Machine learning and statistical models investigate the probability of an area being flooded or not by studying past flood events and generate flood susceptibility maps. The analysis is based on different topographical, hydrological and geological conditioning factors, which can vary depending on the area of study, flood inventory datasets and existing machine learning models. Researchers have implemented different algorithms over different areas of study such as K-Nearest Neighbors (K-NN) (Toth, Brath, and Montanari 2000), Artificial Neural Networks (ANNs) (Falah et al. 2019b), Support Vector Machines (SVM) (Tehrany, Pradhan, and Jebur 2014; Tehrany et al. 2015; Tehrany, Pradhan, and Jebur 2015), random forest (R.F.) (S. Lee et al. 2017; Zhao et al. 2018), Genetic Programming (GP) (Khu et al. 2001), Frequency Ratio (FR) (Tehrany, Pradhan, and Jebur 2015; Rahmati, Pourghasemi, and Zeinivand 2016; Khosravi et al. 2016) and Logistic Regression (LR) (Pradhan 2010; Tehrany, Pradhan, and Jebur 2013). Remote sensing and Geographic Information System (GIS), along with machine learning models,

have made significant contributions in flood modelling (Haq et al. 2012; Pradhan et al. 2014). As a case study, Esfandiari et al (Esfandiari et al. 2020) tested different conditioning factors for flood mapping and reported that five conditioning factors, namely Altitude, Slope, Aspect, Distance from River and Land use/cover, produce the best results once used with a R.F. classifier in the chosen study area. However, these studies map an existing flood as they take training samples from existing satellite images and cannot predict floods in a future event.

One of the major problems of machine learning models is that they require numerous ground truth samples for training. Lee (D.-H. Lee 2013) proposed to use pseudo labels with maximum predicted probability to true labels for neural networks. This is a promising approach that can help with reducing the number of training samples but to the best of our knowledge has never been used in real-time flood mapping. On the other hand, the flood susceptibility maps produced by machine learning provide a general perspective of flood prone areas based on previous flood events; therefore, they can fail in predicting possible future flood events.

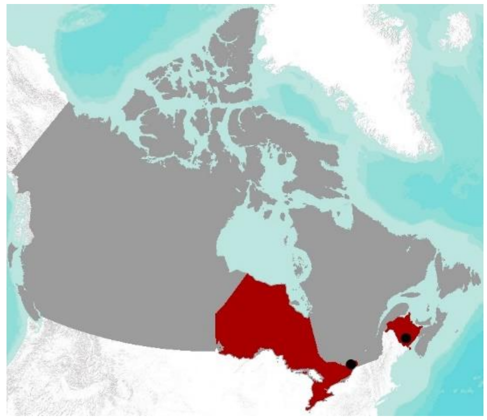
To overcome the problems of machine learning mentioned above (requirement of abundant training samples and low accuracy in flood prediction), we proposed in our paper a new flood prediction model called Pseudo Supervised Random Forest (PS-RF). PS-RF uses the HAND model, a R.F. classifier and Random Sample Consensus (RANSAC) paradigm (Fischler and Bolles 1981) to dynamically estimate the flood extent. In this model, the flood predictions from the HAND model are used as pseudo labels to train a R.F classifier. Through a series of random data selection and outlier detection using RANSAC paradigm, the most reliable predictions are selected for training R.F. To help eliminate erroneous samples, the five best conditioning factors reported by (Esfandiari et al. 2020) are used. Then, the R.F. with the best selected pseudo training subset goes through a cross-validation procedure for hyperparameter optimization. The optimized R.F. classifier, is used to produce the final flood map. In this model, rather than using ground truth samples to train the R.F. classifier, the training samples are selected from a thresholded HAND model with their associated uncertainties. However, since we include other conditioning factors, the resulting trained R.F classifier produces higher accuracies compared to a thresholded HAND model alone.

The flood prediction accuracy of PS-RF was tested using five different flood events in Fredericton, NB, Canada in the years 2014, 2016, 2018, 2019. To prove that the proposed model can work in other areas we also tested it using the 2019 flood in Gatineau, QC, Canada. For accuracy assessment, multiple ground truth points were collected from Landsat 8 OLI and Sentinel-2 satellite imageries. PS-RF was able to produce flood maps with accuracies over 89% which surpasses flood extent estimations produced by the HAND model in all the flood events. The results proved that, by PS-RF, we produced a flood extent prediction model by which we can estimate the inundated areas as a function of the depth of water in the nearby stream. Thus, the relevant authorities can prepare for the rise of water to prevent further damage each year.

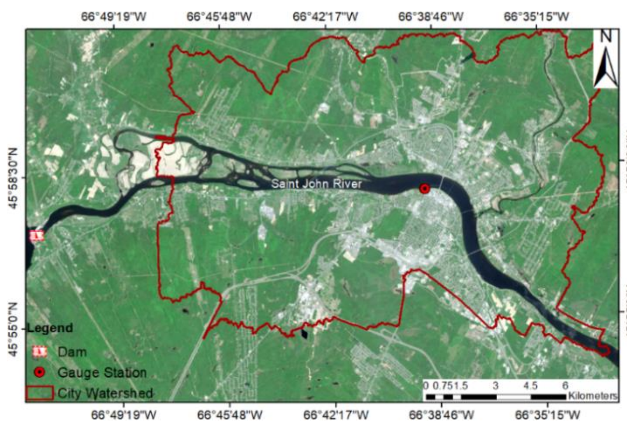
3.2 Materials and Methods

3.2.1 Study Areas

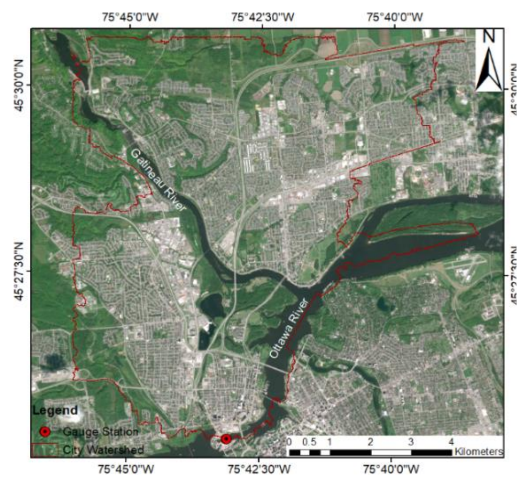
The first study area in this study is in Fredericton (Figure 3-1(b)), the capital of the province of New Brunswick, in Atlantic Canada, which covers an area of approximately 155 square kilometers. The city with its nearly 94,000 population and approximately 22,000 households, has experienced several flood events in history (McGrath, Stefanakis, and Nastev 2016). The Saint John River, which flows from west to east through Fredericton, splits the city into northern and southern sides. Due to cold winters in this region, the surface of the Saint John River freezes every year during winter and it starts to melt as the weather gets warmer in April and May. Extreme snowmelt along with heavy rain can cause a significant rise in the river water level. Also, the Mactaquac Dam, shown in Figure 3-1(b), which is located nearly 19 kilometers upstream from the city, often cannot hold all the melted snow and ice in its headpond, so water may be released to the river downstream. The river, in this study area, is mainly surrounded by relatively flat flood plain which gets inundated by the water level rise.



(a)



(b)

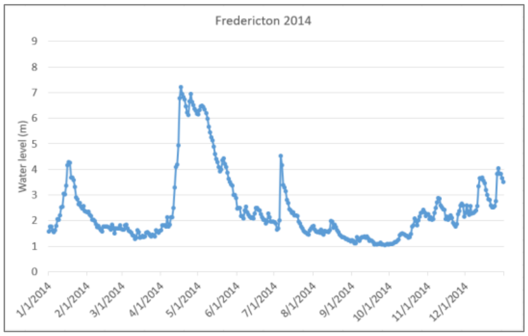


(c)

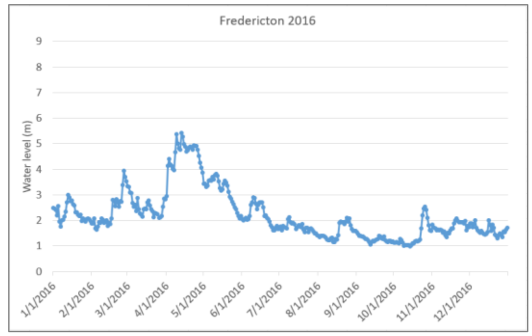
Figure 3-1 Study areas of this paper: (a) map of Canada with the provinces of study highlighted in red, and Fredericton and Ottawa cities highlighted with black dots. More details are provided by Stinel-2 satellite images of (b) Fredericton and (c) and Gatineau study. The red borders in (b) and (c) show the extent of the study areas.

The City of Gatineau (Figure 3-1(c)) is located on the northern bank of the Ottawa River. The city with its 332,000 inhabitants and a total population of 1.3 million is the fourth largest city in the province of Quebec. The Gatineau River goes through the center of the city, flowing from the lakes in the north of the Baskatong Reservoir in western Quebec and joining the Ottawa River in the city. The Ottawa River is regulated with more than 50 major dams and 13 principal reservoirs ("Ottawa Riverkeeper's River Report" 2006). The area of interest contains the relatively flat area starting from Chaudière Bridge, where the hydrometric gauge station of 02LA028 is located, continuing for nearly 6 kilometers along the Ottawa River and Gatineau flood plain starting from Ottawa River continuing upstream for around 9 kilometers. In May 2019, the city and adjacent regions experienced a severe flood causing 111 homes to evacuate, 923 damaged and cost more than 3.4 million for government and insurance companies (CBC News, May 04 2019). The flooding was a result of extreme snowmelt accompanied by heavy rain.

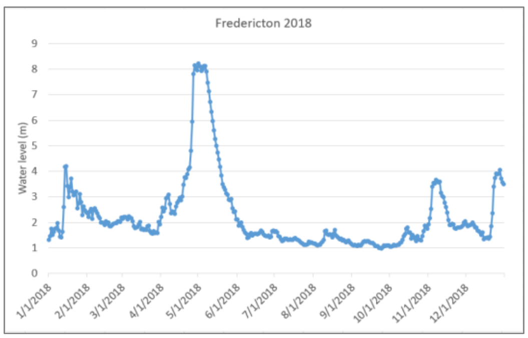
The water level in the rivers is observed with several hydrometric stations along the river every five minutes and the observations are recorded on the Canadian Water Office website (wateroffice.ec.gc.ca). For this research, we used the 01AK003 gauge station located in downtown Fredericton and 02LA028 station located in downtown Ottawa. Figure 3-2 demonstrates the water level in five years of interest in the Saint John River and the Ottawa River.



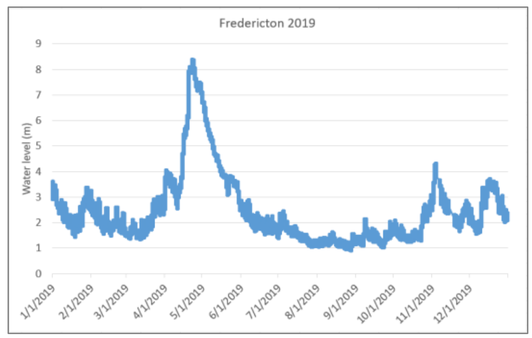
(a)



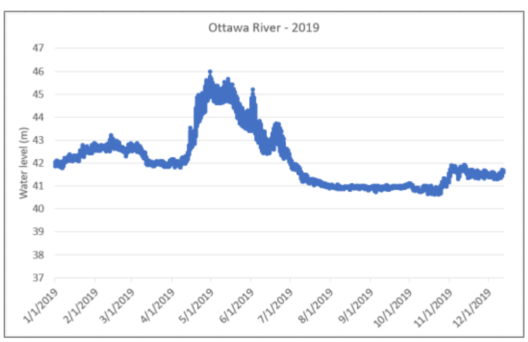
(b)



(c)



(d)



(e)

(a): 1.05 < water level < 7.21 (meters)

(b): 0.99 < water level < 5.42 (meters)

(c): 0.97 < water level < 8.22 (meters)

(d): 0.92 < water level 8.37 (meters)

(e): 40.64 < water level 45.99 (meters)

Figure 3-2 Gauge station readings representing water level for Saint John River in (a) 2014, (b) 2016, (c) 2018, (d) 2019, and in Ottawa River in (e) 2019.

The spring snowmelt and water level rise happens yearly with different degrees of intensity depending on the volume and rate of precipitation and how rapidly the snowpack melts and the depth of the snowpack. In the two consecutive years, 2018 and 2019, the Saint John River reached notable water levels of 8.216 m and 8.356 m respectively, which caused severe damage including the closure of about 50 streets in Fredericton and a total of 1,000 evacuations across the province, with more than $23 million damage to the provincial infrastructure (CBC News, Apr 30, 2018). In the Gatineau area, the water level in 2019 reached 45.994 m which is nearly 3 meters above the normal water level in the Ottawa River. This intense water level rise left 1,212 flood victims and around 575 damaged homes (CBC News, May 04 2019).

3.2.2 Datasets

For this research, we used four different flood events of the Saint John River in Fredericton that occurred in the years 2014, 2016, 2018, 2019 and one flood event in Gatineau city in 2019 (Table 3-1). For accuracy assessment, we used optical satellite images of the area of interest.

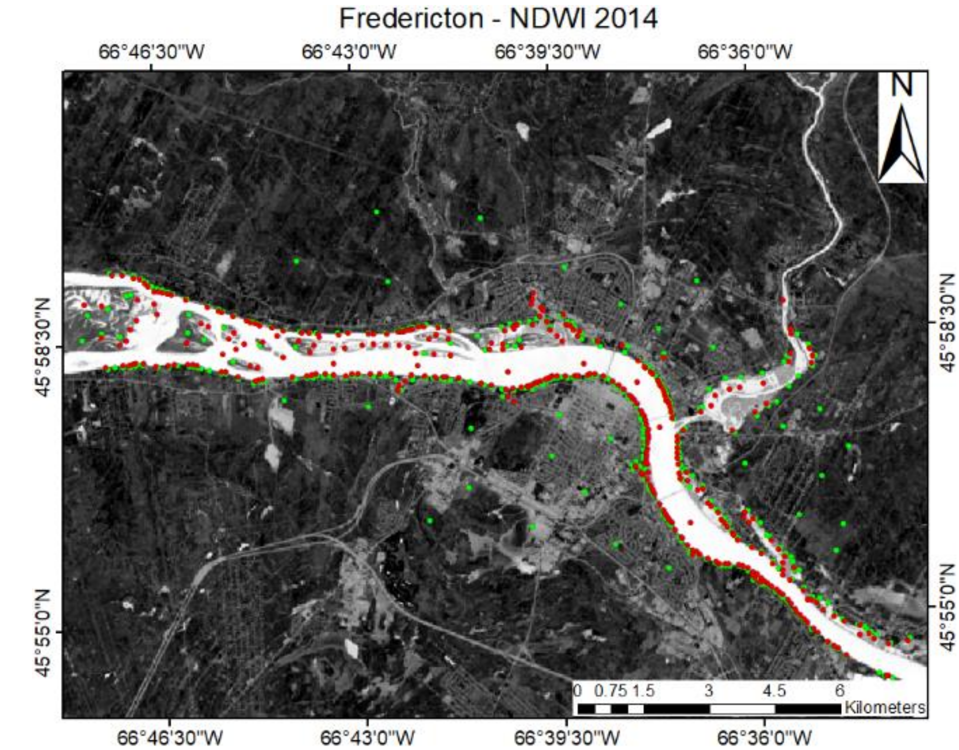
Pan-sharpened Landsat and Multispectral Sentinel-2 images for the five study areas were collected from satellite archives dated either at their peak or close to the peak of the flood events. The Landsat images are pan-sharpened through the fusion of the panchromatic band (15 m) and multispectral bands (30 m) using University of New Brunswick (UNB) pan-sharpening method available in PCI Geomatica software. UNB pan-sharpening addresses significant image fusion challenges including color distortion and dataset dependency and is selected for the fusion of the panchromatic and multispectral bands of the Landsat 8 OLI images, as it properly preserves the spectral signature of the objects as reported by Fathollahi and Zhang (2020). In other words, the pan-sharpening uses the method of least squares to reduce color distortion and applies statistical techniques to remove dataset dependency (Jabari et al. 2017). More details can found in (Zhang 2004).

For each flood event, the author selected a number of sample points whose label is set to either flooded or not-flooded. These points were named depending on the source of labeling. Points interpreted from satellite images labeled as “flooded” or “not flooded” were called “ground truth points”. Satellite images capture one instance of floods and clearly show the flooded pixels. Therefore, these labels are considered as ground truth. For the Fredericton and Gatineau study areas, 700 (350 flooded and 350 not flooded) and 330 (165 flooded; 165 not flooded) ground truth points were collected from satellite images, respectively. The points were well-distributed and spread along the floodplains as shown in Figure 3-3. The ground truth points are only used for accuracy assessment in this study. On the other hand, we labeled the same points using a thresholded HAND model. We call these points “pseudo ground truth points”. Among the pseudo ground truth points, we selected 10% of them as pseudo training points and 90% of them as validation points.

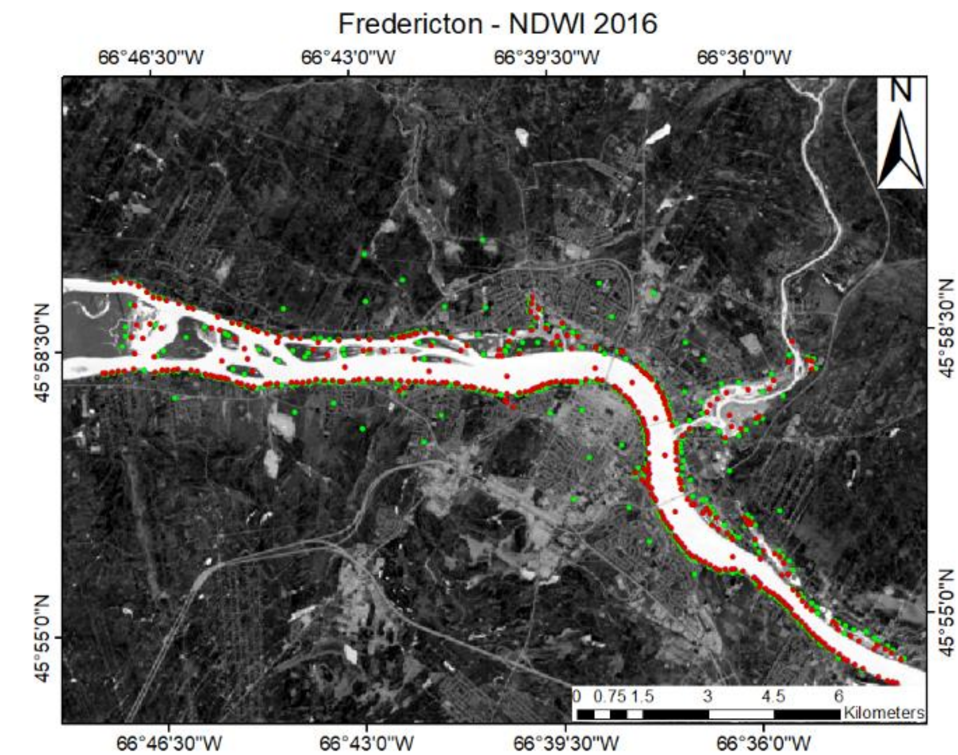
Table 3-1 Flood events and the satellites used for accuracy assessment in each dataset.

Flood Event	Satellite	Spatial Resolution	Time of Imagery	Water depth
Fredericton 2014	Landsat-8	15-m (Pansharpened)	May 07th, 2014 – 15:12:54	6.191m
Fredericton 2016	Landsat-8	15-m (Pansharpened)	April 10th, 2016 – 15:12:54	5.019m
Fredericton 2018	Sentinel-2	10-m	May 02nd, 2018 – 15:26:00	8.036m
Fredericton 2019	Sentinel-2	10-m	May 02nd, 2019 - 15:26:39	6.365m
Gatineau 2019	Sentinel-2	10-m	May 06th, 2019 – 15:49:11	44.990m

For accuracy assessment, the ground truth points were selected using a newly developed Normalized Water Difference Index (NDWI) (Amer, Kolker, and Muscietta 2017) produced from the satellite images (Figure 3-3). Prior to NDWI calculation, top of the atmosphere correction (TOA) is applied to the images to remove the atmospheric effects from the reflectance values. TOA is applied in QGIS software using the Semi-Automatic Classification Plugin (Congedo 2016). Equation (1) demonstrates how NDWI is calculated using blue (Blue) and short-wave infrared (SWIR) bands (bands 2 and 6 for Landsat 8 OLI and bands 2 and 11 for Sentinel-2 satellites, respectively). Flooded and not-flooded ground truth points extracted from the NDWIs were mainly selected within the flood plain to assess the capability of the HAND model and PS-RF in generating accurate flood maps.

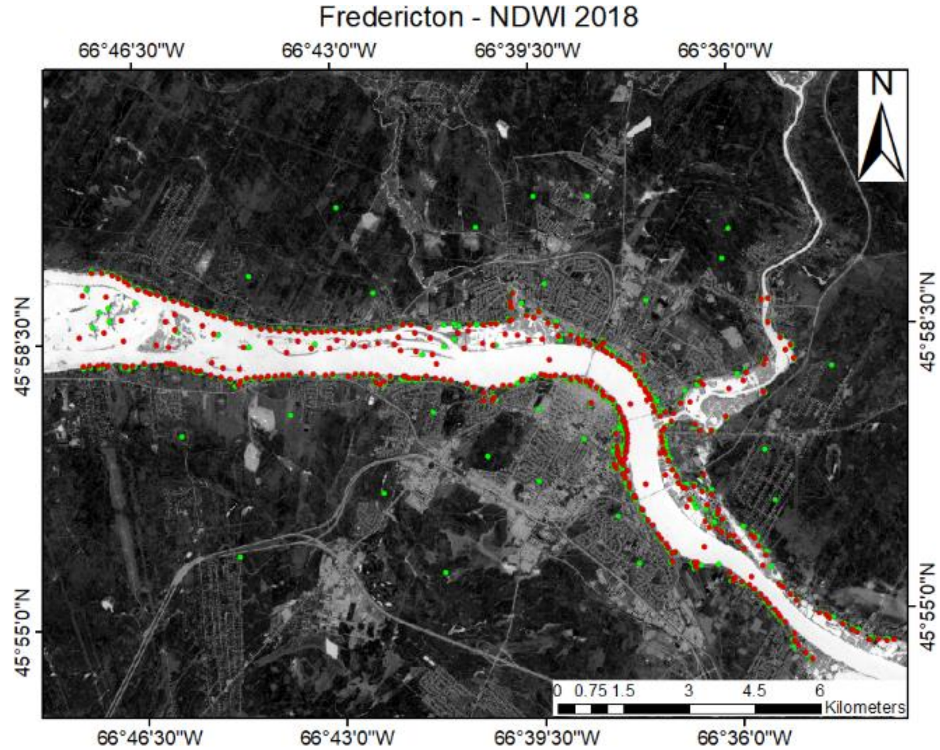


(a)

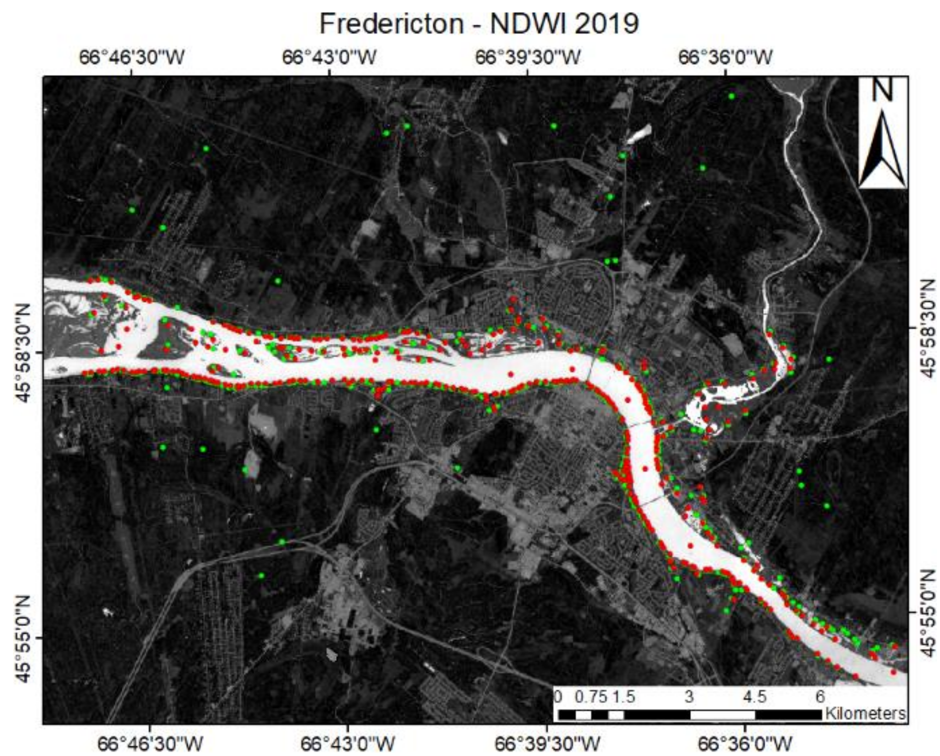


(b)

Figure 3-3 Cont

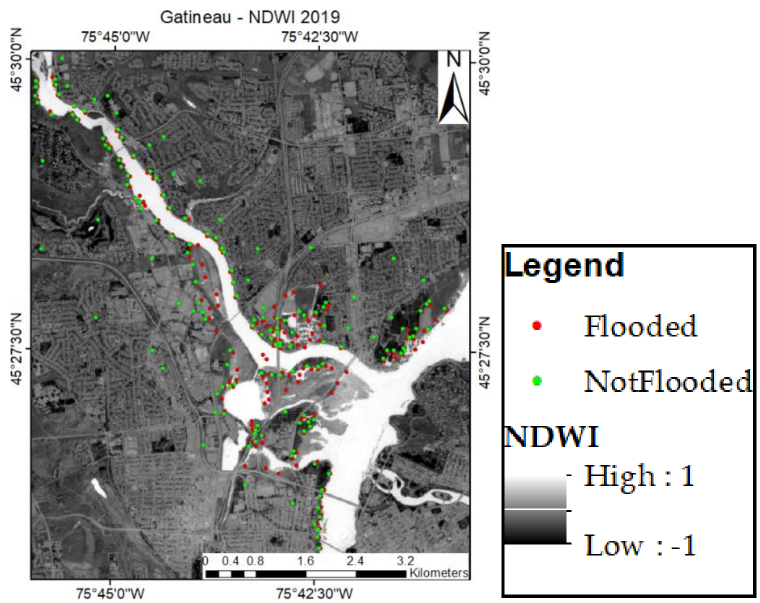


(c)



(d)

Figure 3-3 Cont



(e)

Figure 3-3 NDWI layers produced for (a) 2014, (b) 2016, (c) 2018, (d) 2019 Fredericton and (e) 2019 Gatineau datasets. The coloured dots in the images each indicate a particular location confirmed to be either "not-flooded" (green) or "flooded" (red) at the time of ground truthing.

$$NDWI = \frac{(Green - NIR)}{(Green + NIR)} \quad (1)$$

Light Detection and Ranging (LiDAR) data were used to generate the HAND model and three other conditioning factors including Altitude, Slope and Aspect. For the Fredericton datasets, the LiDAR data were obtained from the geographic data catalog website of the Province of New Brunswick (GeoNB) ("ERD 2015 Lidar" n.d.). The LiDAR data were collected in a period of nearly two months, starting on August 2nd, 2015 and continuing to September 28th, 2015, with an average point density of 6 points per square meter. The LiDAR data has a horizontal positional accuracy of ~ 0.257 meter and vertical positional accuracy of 0.275 meter which were tested using 216 GPS RTK Survey points ("ERD 2015 Lidar" n.d.). Using available high-resolution LiDAR data, a Digital Terrain Model (DTM) with 1-meter resolution was generated for Fredericton area. For the Gatineau dataset, a bare earth High Resolution Digital Elevation Model (HRDEM) (Secretariat Treasury Board n.d.) was downloaded from a Government of Canada website with 1-meter resolution. HRDEM is a national product of high resolution DTM generated from LiDAR across Canada. The downloaded HRDEM was used to generate the HAND model and Altitude, Slope and Aspect conditioning factors for Gatineau city as was done in the Fredericton study area.

3.2.3 Conditioning Factors

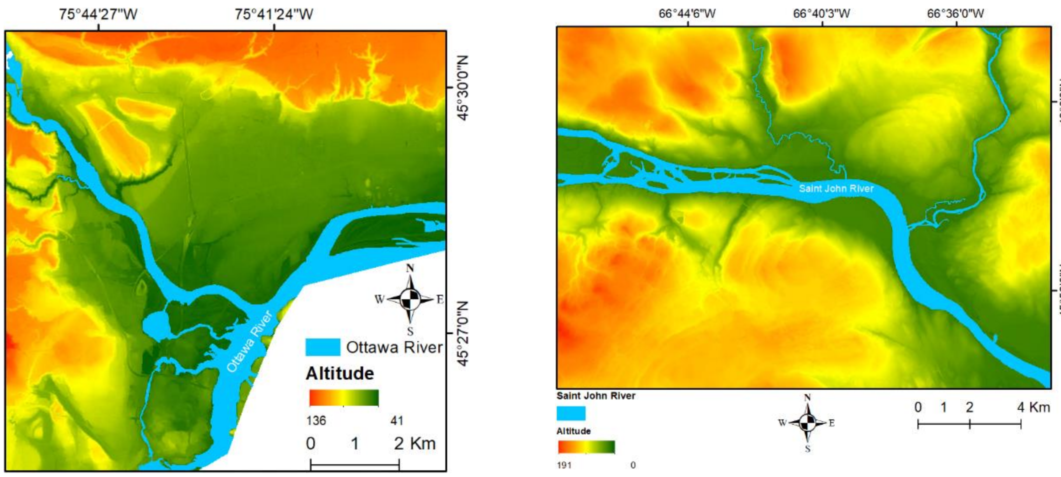
In flood mapping using machine learning, it is crucial to use the appropriate conditioning factors (Kia et al. 2012). In a previous case study (Esfandiari et al. 2020) the best conditioning factors for flood mapping in the Fredericton area were identified. In that study, the effectiveness of 12 different 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 HAND model were investigated. To find the best combination of conditioning factors, different combinations of the 12 layers were tested as input features of a R.F. classifier. The results confirmed that the combination of five layers, namely Altitude, Slope, Aspect, Distance from River and Land use/cover, provided the highest accuracy. It was found that the Altitude layer is one of the most important conditioning factors for flood mapping (Esfandiari et al. 2020). The results also demonstrated that using correlated conditioning factors reduces the accuracy of the R.F. model and a higher number of conditioning factors does not necessarily improve the predictions (Esfandiari et al. 2020;

Tehrany, Jones, and Shabani 2019). Thus, for this research, the 5 best conditioning factors namely Altitude, Slope, Aspect, Distance from River and Land use/cover (Figure 3-4) are used to help with the training of the R.F. model.

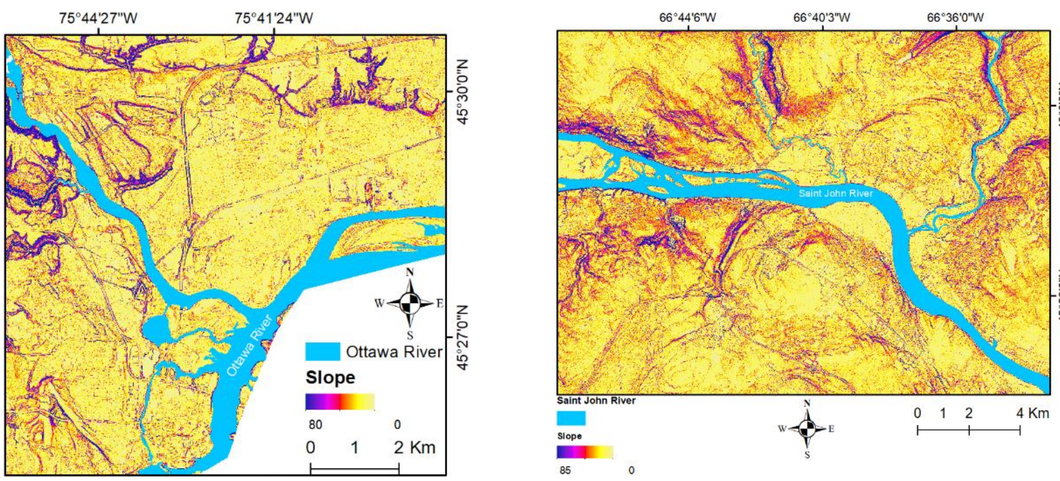
For the Fredericton dataset, the Altitude layer, with 1-meter spatial resolution, was generated from LiDAR data. Slope and Aspect conditioning factors were derived from the Altitude layer in ArcGIS Desktop 10.6.1 using the Spatial Analyst tool. Distance from River was generated using the Euclidean distance tool within the ArcGIS software and generated from the river boundary shapefile which was obtained from the GeoNB website. The Land use/cover layer was generated by overlaying available polygon shapefiles provided on the GeoNB website. The polygons provided seven classes of Urban, Forest, Grass Land, Bare Land, Roads, Water, Wetlands and were obtained through aerial photography.

The generated conditioning factors for the Gatineau city were derived from HRDEM obtained from the federal governmental website. Conditioning factors were generated using the same tools mentioned earlier within ArcGIS software. The land use/cover layer for the Gatineau dataset was downloaded from Agriculture and Agrifood Canada[1].

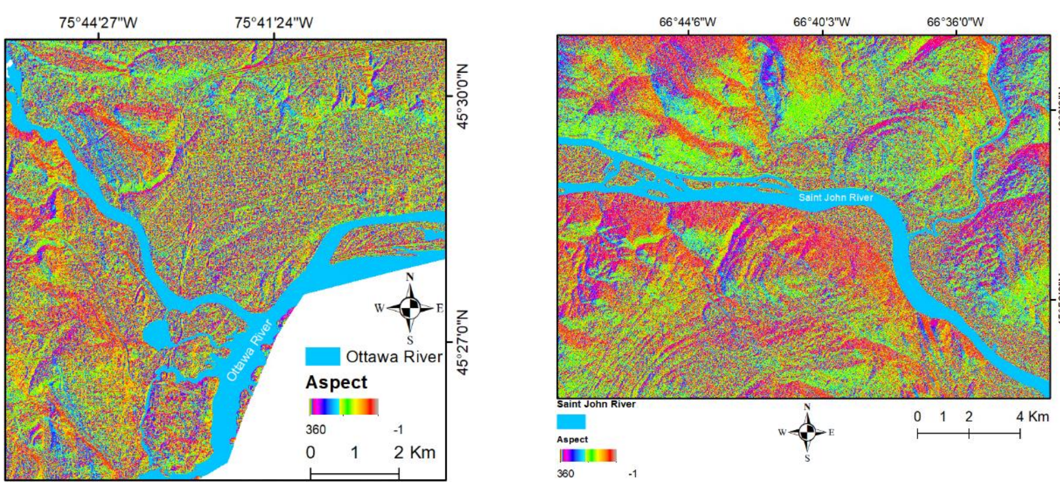
Previously mentioned conditioning factors, Altitude, Slope, Aspect and Distance from River had ordinal values that were normalized from 0 to 1 for better consistency in R.F. (Ihsan, Idris, and Abdullah 2013b). All the conditioning factors were covering the same area extent constructing grids of size $1m^2$, which resulted in raster sizes of 22448 columns and 11533 rows (~ 258 square kilometers) for the Fredericton region. The same conditions for the Gatineau city were applied and the grids in the Gatineau raster files contained 10000 columns and 10000 rows (~100 square kilometers).



(a)

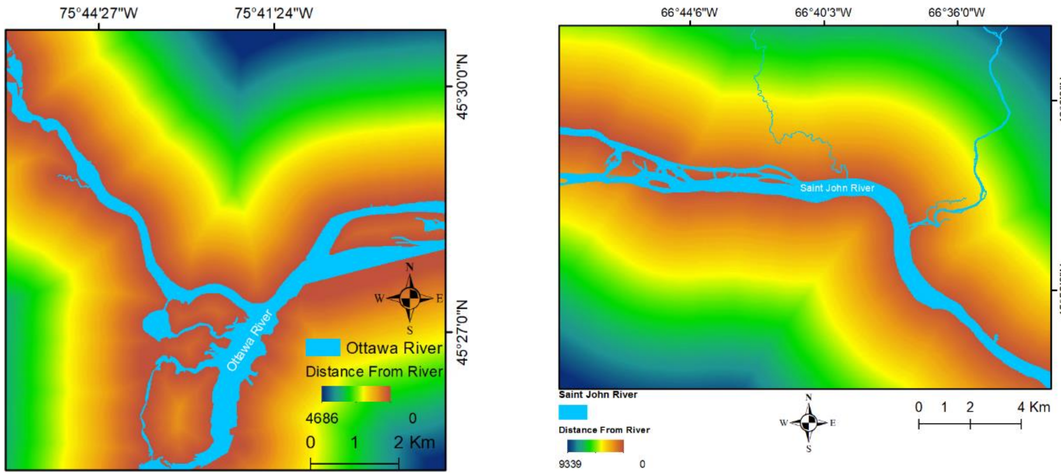


(b)

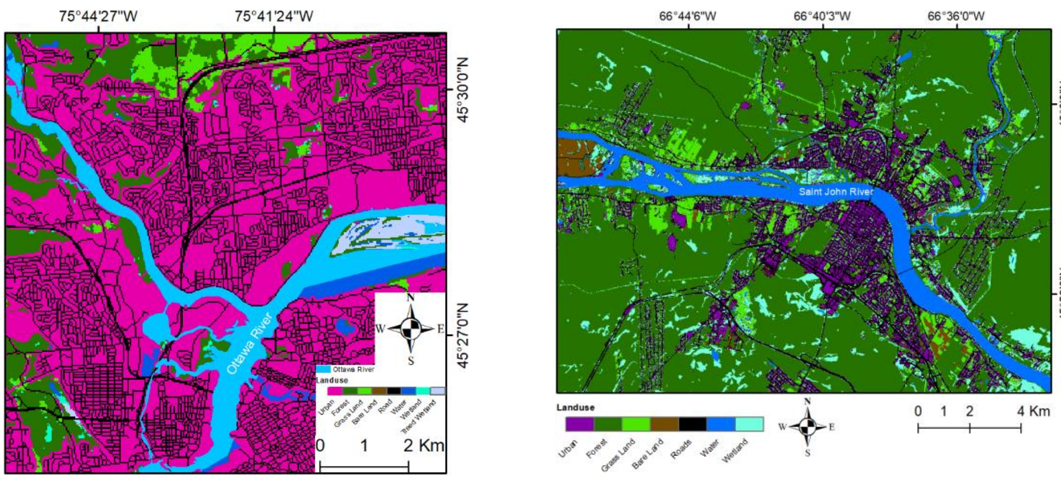


(c)

Figure 3-4 Cont.



(d)



(e)

Figure 3-4 Conditioning factors: (a) Altitude, (b) Slope, (c) Aspect, (d) Distance from River, (e) Land use/cover for (left) Gatineau, (right) Fredericton.

3.2.4 HAND Model

The HAND model is used for several applications such as flood hazard mapping (McGrath et al. 2018b), (Speckhann et al. 2018), landform classification (Nobre et al. 2011) and remote sensing (Chow, Twele, and Martinis 2016). As shown in Figure 3-5, the HAND model is generated through a series of steps. One of the most important requirements of having an accurate HAND model is to use a hydrologically conditioned DTM (Nobre et al. 2011). To have a proper hydrologically conditioned DTM, a total of 34 culverts and bridges were identified through aerial imagery and OpenStreetMap and burnt using Zonal Statistics Tool and Raster Calculator within ArcMap (Step – 1 in Figure 3-5). Then, the pit holes and depressions in the area were filled using "Pit Remove" function from the Terrain Analysis Using Digital Elevation Models (TauDEM) suite within ArcGIS (Tarboton, Sazib, and Dash 2015) (Step – 2 in Figure 3-5). By burning culverts and bridges and filling depressions we ensured that our DTM does not contain erroneous flow directions. Then, the flow direction raster was computed using D-Infinity Flow Directions function from TauDEM, in which the flow is calculated through a triangular facet (Tarboton 1997b) (Step – 3 in Figure 3-5). The last required component of the HAND model is the stream raster which will be considered as the elevation reference of the model and elevation of all pixels will be calculated from this layer, hence, it is necessary to ensure that the stream raster is accurate. The stream raster can be obtained through D-8 Flow Direction and thresholding the major stream based on Strahler stream order (McGrath et al. 2018b; Liu et al. 2016) but, since the stream shapefiles representing all the streams of the province were available in the GeoNB and Ottawa Open Data[2], geospatial databases for Fredericton and Gatineau respectively, we rasterized the shapefile and used it as our stream raster. The "D-Infinity Distance Down" function available in the TauDEM software is the last step which generates the HAND model. Different distance and statistical methods of calculation are available within this tool but for the HAND model generation Vertical method for Distance and Average for Statistical are selected (Step – 4 in Figure 3-5). Once the HAND model is generated through the steps, the model can be used to demonstrate inundated areas based on the water level rise. To do that, the water level difference should be calculated based on the observations from gauge station and the HAND model. For more details on the HAND model, readers can refer to Nobre et al. (2011) and Liu et al. (2018).

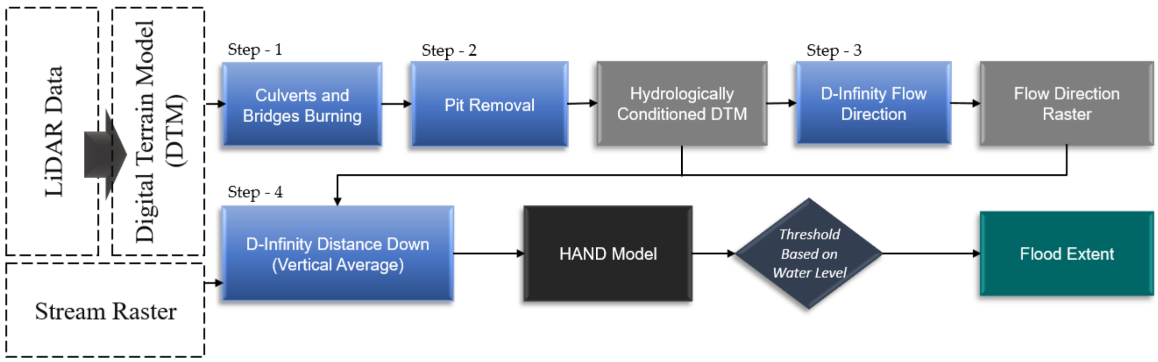
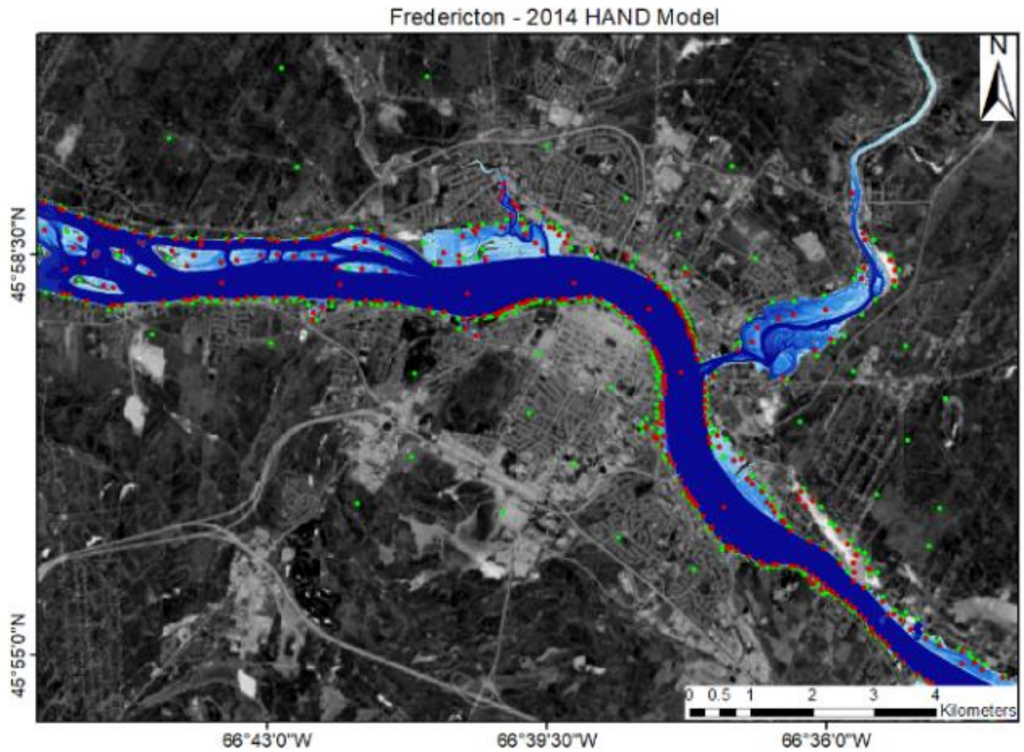


Figure 3-5 Implemented workflow for generating flood extents from the HAND model.

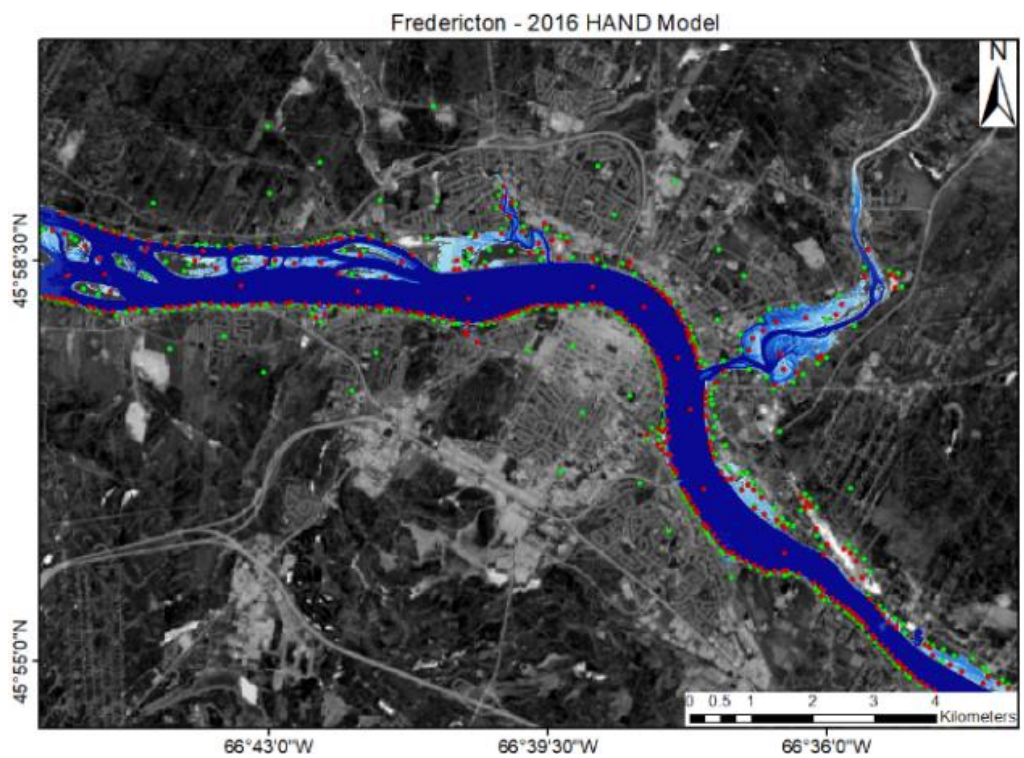
To generate an early estimation of the flood extent, the HAND model is thresholded based on the water level as recorded by the local river gauge. To properly threshold the HAND model, the difference of water level at the time of flood and at the time of LiDAR data collection should be calculated and the model should be thresholded based on that difference (Nobre et al. 2016a). As previously mentioned, in the Fredericton dataset, the LiDAR data were collected in a period of approximately two months from August 2nd, 2015 continuing to September 28th, 2015 and during that time the river water level fluctuated from 1.121 meters to 1.854 meters. We have used the average water level over that period as the river water level at the time of LiDAR data collection which is equal to 1.46 meters (considering the minimum water level of 1.121 meters lead to underestimation in flood maps and the maximum water level causes overestimation). So, the threshold for all the Fredericton datasets were calculated by subtracting the average 1.46 meters from the water level at the time of the flood event. The reference water level for the Ottawa River was considered to be 42.04 meters which was obtained through gauge readings at the time of data collection. Figure 3-6 shows the HAND models generated for the five flood events discussed in this study.

3.2.5 Random Forest

R.F. is a supervised machine learning model and one of the most robust classifiers that makes decisions based on the average of the results of a multitude of decision trees (Breiman 2001). The classifier fits an arbitrary number of trees on different features of the datasets and by averaging multiple predictions controls the over-fitting issue. R.F. was selected because of its capability to handle noisy data, its efficiency to take features with different natures and its ability to rank the features based on their degree of importance (Breiman 2001; Rahmati and Pourghasemi 2017). The feature with the highest degree of importance is the major splitter in all the trees.

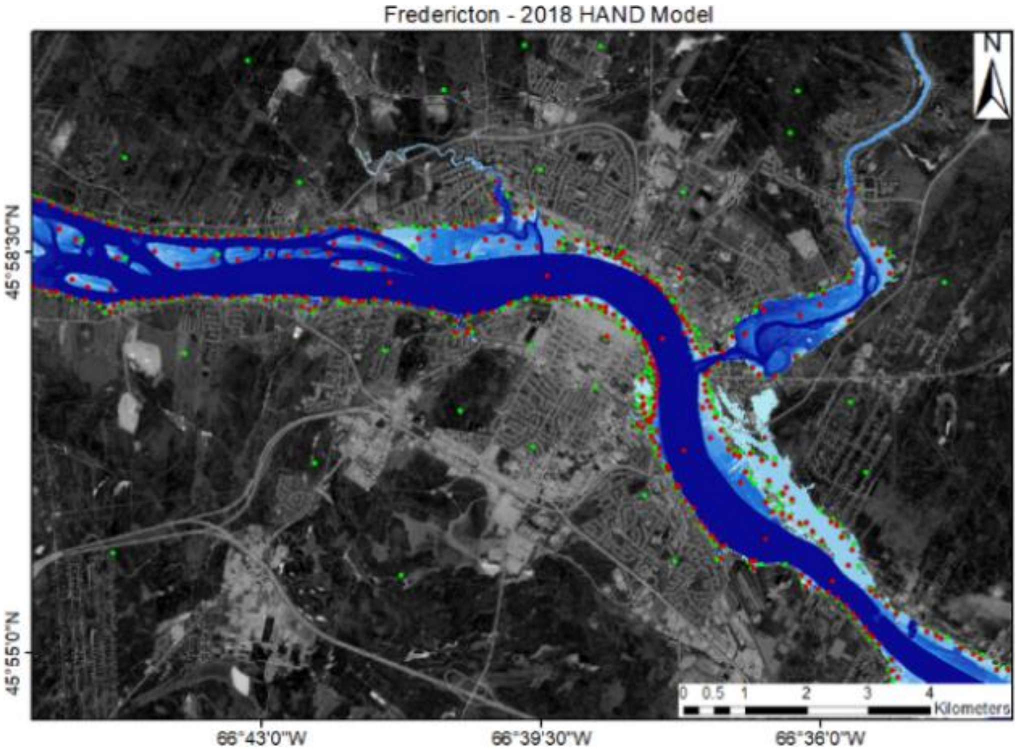


(a)

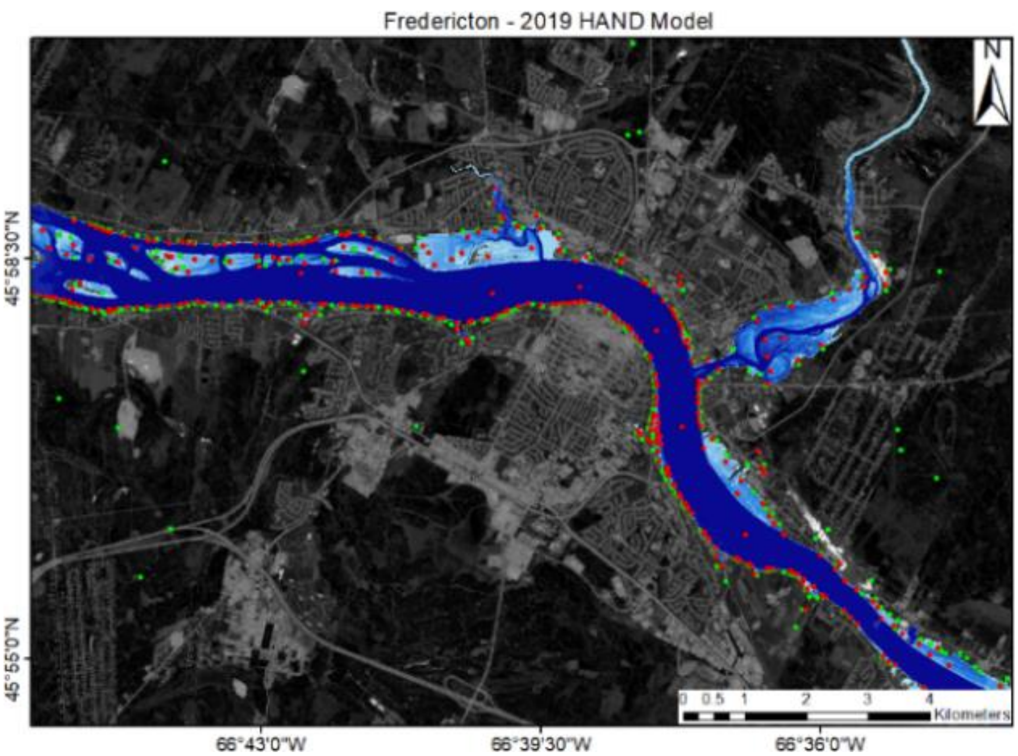


(b)

Figure 3-6 Cont

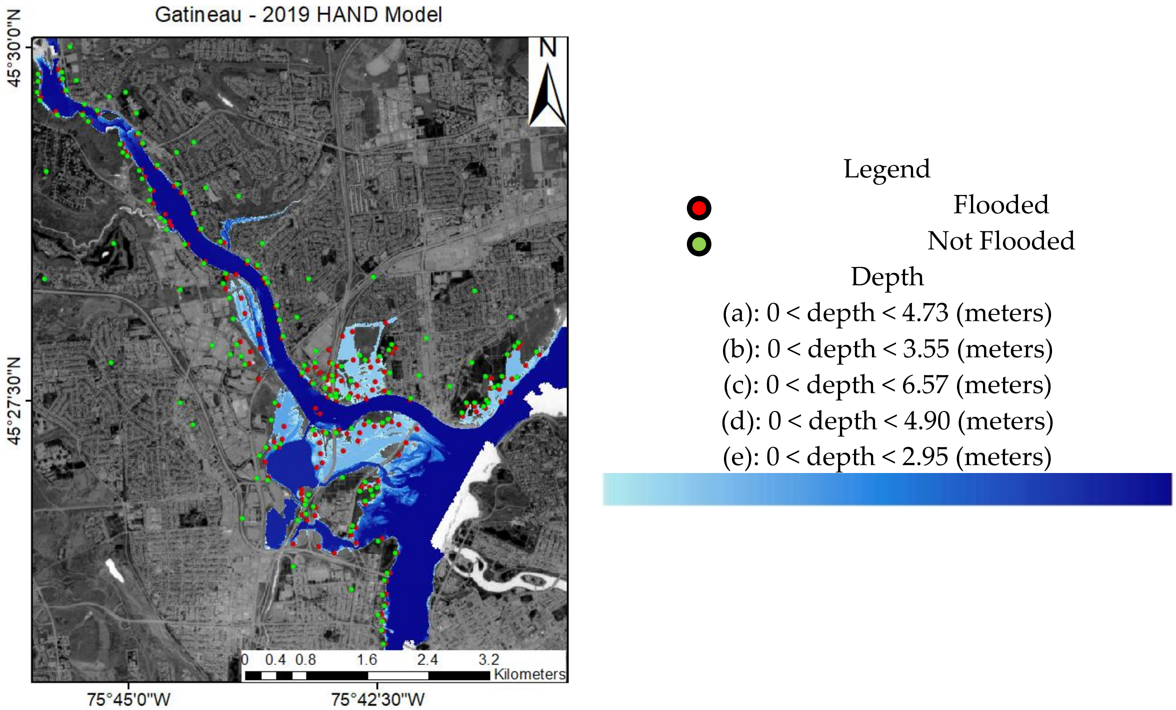


(c)



(d)

Figure 3-6 Cont



(e)

Figure 3-6 Flood extent for five flood events including Fredericton in 2014 (a), Fredericton in 2016 (b), Fredericton in 2018 (c), Fredericton in 2019 (d), and Gatineau in 2019 (e) all generated from the HAND model corresponding to the dates of the satellite imagery.

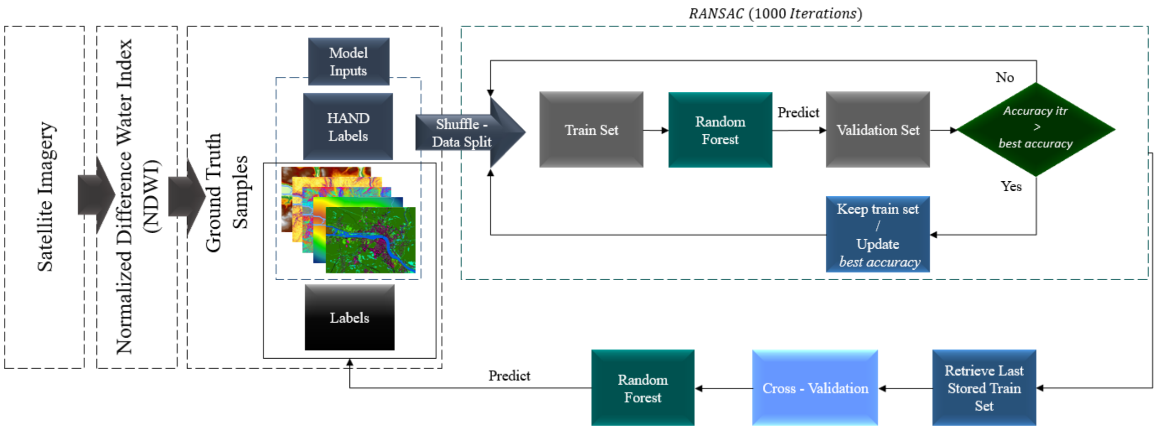


Figure 3-7 PS-RF workflow

3.2.6 Pseudo Supervised Random Forest (PS-RF)

Figure 3-7 demonstrates the workflow of the proposed PS-RF model. In PS-RF, we used the HAND model predictions as pseudo labels to train a R.F. classifier in Python within Scikit-Learn library (Pedregosa et al. 2011). However, due to the deficiencies of the HAND model discussed previously, the pseudo labels obtained from the HAND model are not 100% accurate. Therefore, we need to detect the outliers and remove them so that the R.F. algorithm is trained with the most reliable pseudo training points.

For selecting the best pseudo training points and removing the outliers, the RANSAC paradigm was used. Through the RANSAC process, the HAND model is used as a cost function for R.F. optimization. Through iterations and training R.F. by using randomly selected pseudo training points, the best training subset is selected and, as is evident from Figure 3-8, different training subsets achieve various accuracies. So, the outliers involved in training data provided by the HAND model will be identified.

First, the dataset, including 700 pseudo ground truth points for Fredericton and 330 pseudo ground truth points for Gatineau, with values from conditioning factors and pseudo labels (flooded and not-flooded labels from the HAND model) are randomly split into two subsets of 10% training (in fact pseudo training) and 90% validation. The R.F. algorithm is then trained by the pseudo training data and used to predict the validation subset. The overall accuracy of the validation subset is checked against the HAND model. This accuracy number is stored and the process is repeated 1000 times to make sure different combinations of pseudo training points from the dataset are used to train R.F. Over the iteration process, the pseudo training subset is updated each time the process reaches a higher accuracy than previous predictions over the validation subset.

Once 1000 iterations were completed, the best pseudo training subset was used to train the R.F. model. Also, the model hyperparameters including the number of trees (n_estimators), the minimum number of samples (min_sample_split), minimum number of samples at a leaf (min_samples_leaf) and maximum depth of trees (max_depth) are optimized through 5-fold Cross-Validation. Then, the R.F. model which was trained by the best-achieved pseudo training subset is used to predict the ground truth points of the whole dataset.

3.3 Results

In this study, different measures for accuracy assessment were applied to compare the performance of our results from PS-RF to that of a thresholded HAND model. To assess the accuracy of PS-RF we used accuracy measures such as true positive rate (TPR), true negative rate (TNR), false positive rate (FPR), false negative rate (FNR), overall accuracy, Cohen’s Kappa Coefficient (Campbell and Wynne 2011) and Matthews Correlation Coefficient (MCC) (Matthews 1975) which are depicted in Table 3-2.

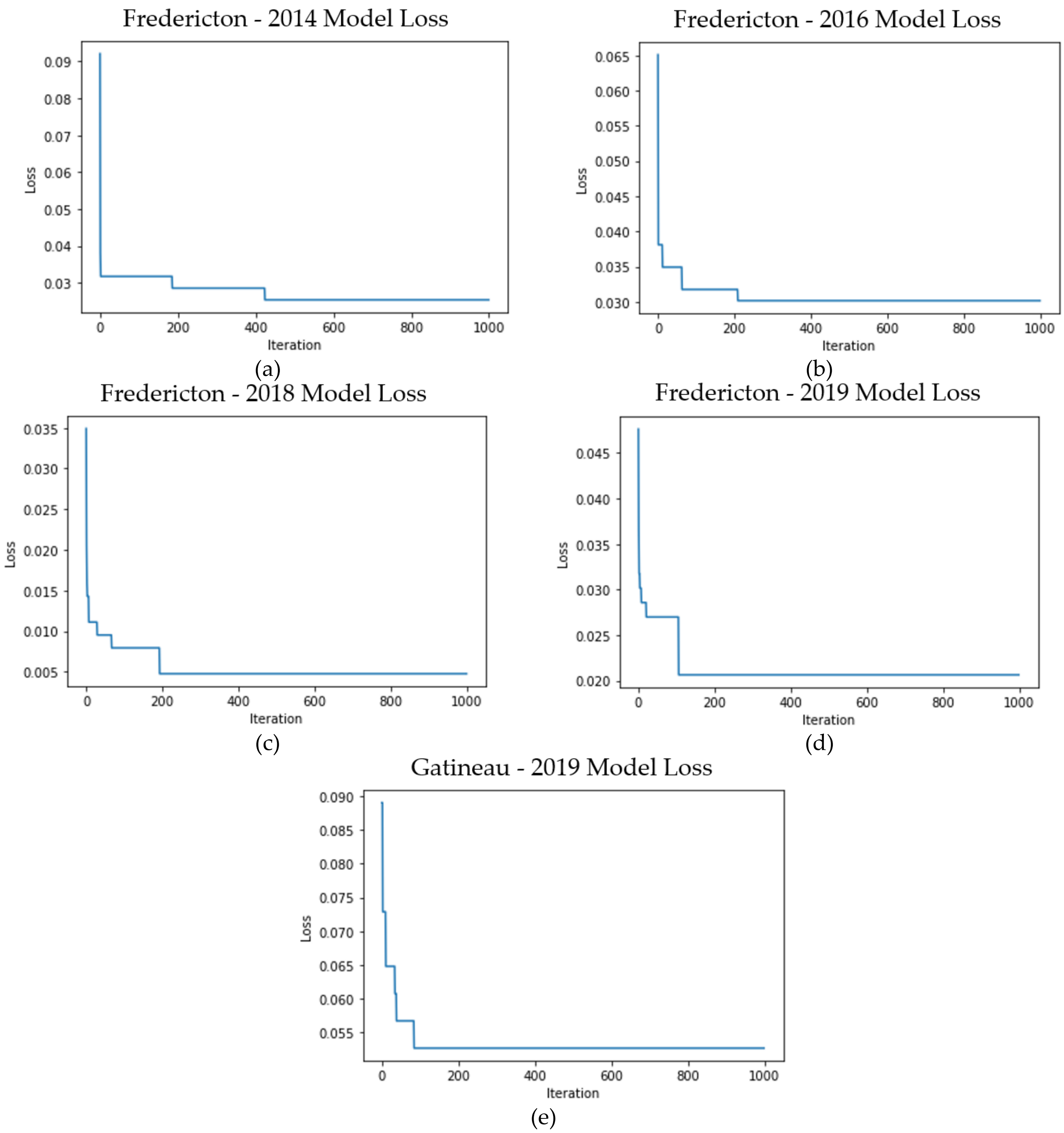


Figure 3-8 Model loss for R.F. through RANSAC process demonstrated for five flood datasets of (a) Fredericton 2014, (b) Fredericton 2016, (c) Fredericton 2018, (d) Fredericton 2019, and (e) Gatineau 2019.

TPR (Sensitivity) and TNR (Specificity) indices show the probability of correctly predicted flooded and not-flooded points by the HAND model and the proposed model. FPR and

FNR indices show the percentage of flooded and not-flooded points which were misclassified. Cohen's Kappa Coefficient represents the rate of likeliness between ground truth and predicted data and its values are interpreted as: None 0 – 0.20, Minimal 0.21 – 0.39, Weak 0.40 – 0.59, Moderate 0.60 – 0.79, Strong 0.80 – 0.90 and Almost Perfect likeliness for values of above 0.90 (McHugh 2012). Due to recent concerns about Cohen's Kappa Coefficient assessments and its undesired behavior discussed in (Delgado and Tibau 2019), the MCC metric was also calculated to ensure the validity of our evaluation. MCC values close to 1 represent perfect agreement between predicted and observed values, a value of 0 is interpreted as random prediction and a value of -1 is interpreted as complete opposite (i.e., observed and predicted values were in complete disagreement). Calculated indices are all based on four parameters of true positive (T.P.), true negative (T.N.), false positive (F.P.) and false negative (F.N.). In Cohen's Kappa Coefficient, P_0 and P_e are relative observed agreement and the hypothetical probability of chance agreement, respectively.

Table 3-2 Formulas related to accuracy assessment section.

Parameter name	Formula	
Sensitivity (TPR)	$\frac{TP}{TP + FN}$	(2)
Specificity (TNR)	$\frac{TN}{TN + FP}$	(3)
Fall-out (FPR)	$\frac{TP}{FP + TN}$	(4)
Miss-rate (FNR)	$\frac{FP}{FN + TP}$	(5)
Overall Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$	(6)
Cohen's Kappa Coefficient (K)	$\frac{p_0 - p_e}{1 - p_e}$	(7)

Matthews Correlation Coefficient (MCC)	$\frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$	(8)

Figure 3-8 represents the Model Loss for the five periods. The vertical axis of the figures demonstrates how the prediction accuracy over validation data has improved and higher accuracy was obtained by selecting diverse subsets of pseudo ground truth points provided by the HAND model. The horizontal axis is the implemented iteration number. In the 2014 dataset, the R.F. model accuracy over validation data were improved by 0.06 through the iteration process. In 2016, it can be seen that an improvement of nearly 0.04 was achieved through iteration. In 2018, the accuracy for R.F. model was enhanced by 0.03 over validation data through shuffling and selecting training data randomly among our pseudo ground truth points. R.F. predictions for 2019 were also enhanced over validation data by 0.025 through randomly selecting pseudo training data. In the Gatineau flood event that occurred in 2019, there was an improvement of around 0.04 through the iteration process.

The achieved overall accuracy, which represents the closeness of predictions to their true classes, shows that PS-RF has improved the accuracy of the HAND model in all the five flood events (Figure 3-9). In the 2014 Fredericton flood event, PS-RF reached an overall accuracy of 94.23% compared to 88.86% from the HAND model. Also, PS-RF was ~11% higher than HAND in both Cohen's and MCC respectively. In the 2016 Fredericton flood event, the PS-RF model has improved the accuracy of the HAND model by 5% with 94.14% accuracy for PS-RF and 89.14% accuracy for the HAND model. Cohen's and MCC metrics were 10% and ~9.5% higher for PS-RF than HAND in 2016 as well. 2018 flood event predictions were nearly similar but still, PS-RF obtained a higher accuracy with 88.43% overall accuracy for PS-RF and 85.57% for the HAND model. In 2018, PS-RF had Cohen's of ~77% which is ~6% higher than HAND and MCC score of ~79% which is ~7%

higher than HAND. In the 2019 Fredericton flood event, PS-RF was able to reach an overall accuracy of ~95% compared to ~90% for the HAND model with ~11% higher Cohen's and MCC score. In Gatineau 2019 flood event, the overall accuracy of PS-RF was 94.85% while the HAND model reached 93.03%. PS-RF's Cohen's and MCC score for Gatineau 2019 flood event were ~3.5% and ~3% higher than HAND, respectively.

The Cohen's and MCC results show that PS-RF has outperformed the HAND model in all the five years of interest and the degree of likelihood in the prediction of flooded and not-flooded points is superior to the HAND model according to Cohen's Kappa Coefficient values. Also, higher values of the MCC index obtained by PS-RF show better classification quality than the HAND model. TPR, FPR, TNR and FNR indices are discussed next and demonstrated in Figure 3-10 for the HAND model and PS-RF.

To analyze the HAND model and PS-RF accuracy in detecting flood extent, TPR, FPR, TNR and FNR indices were calculated. The accuracy indexes are presented in Figure 3-10 and discussed further in the discussion section.

Figure 3-9 Overall accuracy, Cohen's Kappa Coefficient and Matthews Correlation Coefficient (MCC) Score for the HAND model and PS-RF in five flood events of (a) Fredericton 2014, (b) Fredericton 2016, (c) Fredericton 2018, (d) Fredericton 2019, and (e) Gatineau 2019.

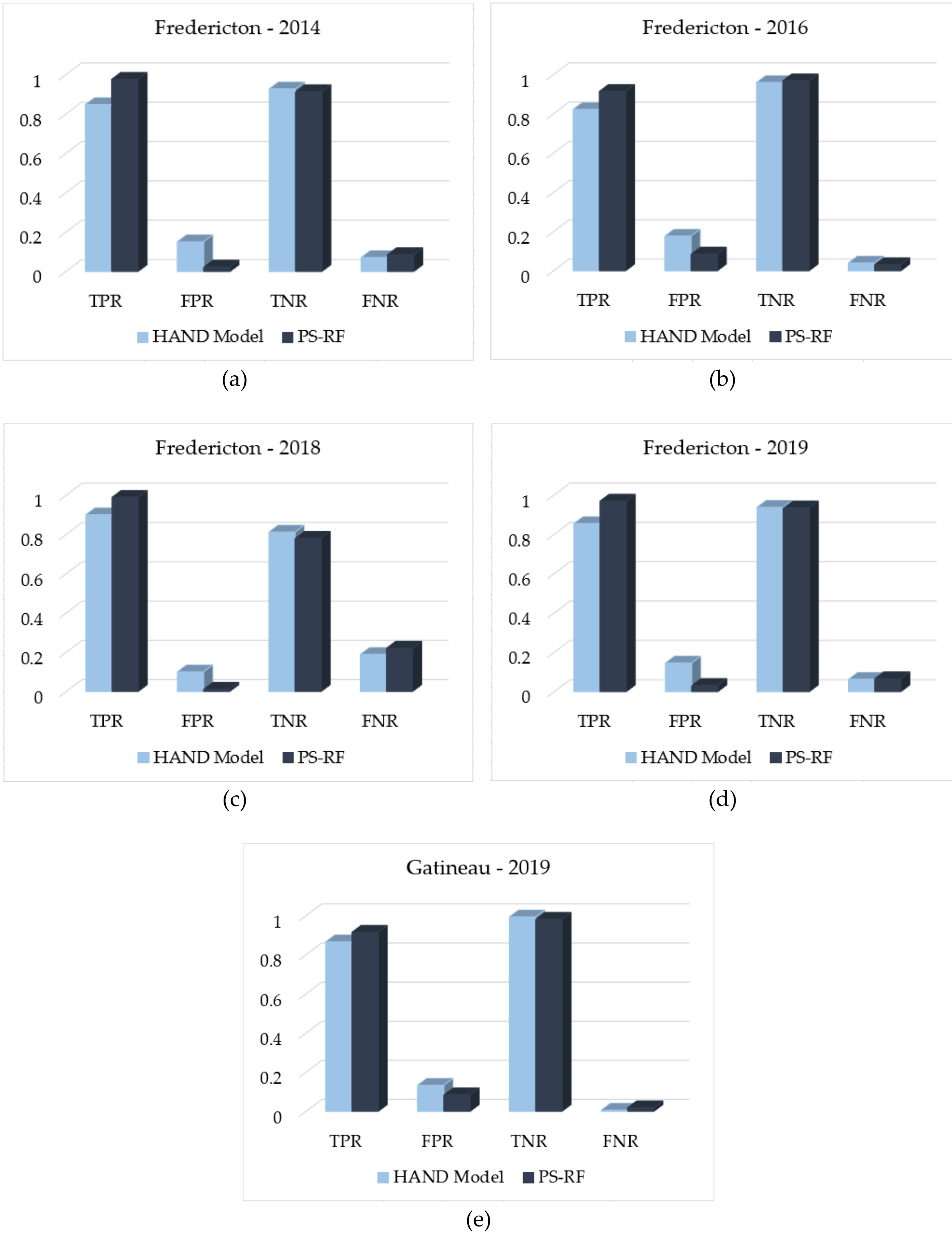


Figure 3-10 True positive rate (TPR), true negative rate (TNR), false positive rate (FPR) and false negative rate (FNR) for the five flood events (a) Fredericton 2014, (b) Fredericton 2016, (c) Fredericton 2018, (d) Fredericton 2019 and (e) Gatineau 2019.

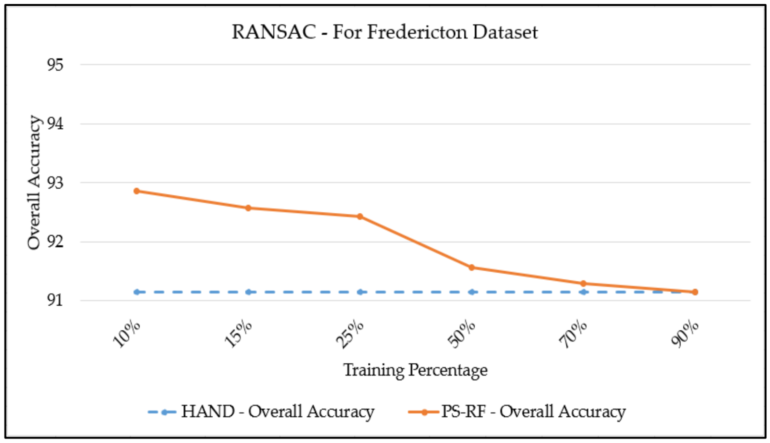
3.4 Discussion

PS-RF was tested in 5 different flood events and resulted in higher accuracy than the HAND model in all. In flood mapping, it is very important to provide precise estimations and the model should not over or underestimate the flood extent. To analyze the performance the HAND model in estimating the extent of 5 tested flood events, different accuracy metrics were calculated and are discussed below.

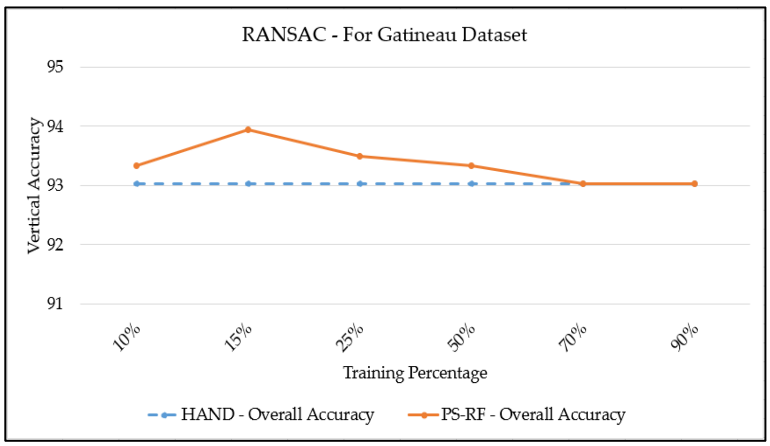
Fore the Fredericton 2014 flood event, the HAND model has correctly predicted ~85% of flooded ground truth points while PS-RF achieved ~98%. In the case of not-flooded points, the HAND model performed slightly better with TNR value of ~93% compared to ~91% for PS-RF. This means PS-RF was able to perform better in detecting flooded points but the HAND model had a better performance in detecting not-flooded points (Figure 3-10 (a)). In the Fredericton 2016 flood event, PS-RF performed better in predicting both flooded and not-flooded points with ~9% and ~1% higher TPR and TNR values (Figure 3-10 (b)), respectively. The Fredericton 2018 flood event contained nearly similar results for PS-RF and HAND model with PS-RF achieving ~9% higher TPR and ~3% lower TNR compared to the HAND model (Figure 3-10 (c)). In the Fredericton 2019 flood event, PS-RF obtained ~11% higher TPR and ~0.3% lower TNR (Figure 3-10 (d)). In the Fredericton 2016, 2018 and 2019 floods, PS-RF performed better in detecting both flooded and not-flooded ground truth points than the HAND model. However, in the Gatineau 2019 flood event, PS-RF was able to reach a TPR value of ~93% compared to ~87% from the HAND model. However, the HAND model performed better in detecting not-flooded points with ~2% higher TNR value (Figure 3-10 (e)). The HAND model can estimate the inundated areas given the depth of water. Therefore, it can produce a dynamic flood mapping model.

However, since it is a simplified model, it suffers from having a limited accuracy as reported in different studies (Nobre et al. 2016a). In this study, we proposed a new model that achieved better results compared to a HAND model. We trained an R.F. model to benefit from different conditioning factors, that is, altitude, slope, aspect, distance from river and land use/cover, to adjust the flood borders produced by the HAND model. Thus, this model combines the HAND results with different conditioning factors to improve the accuracy of the final flood maps.

In PS-RF, the predictions of the HAND model are used as the presumptive labels of ground truth points to train R.F. model with five conditioning factors of Altitude, Slope, Aspect, Distance from River and Land use/cover. However, as concluded previously, the HAND model estimations contain errors that need to be eliminated from the pseudo training subset. To remove the HAND model errors, the RANSAC paradigm was employed. The philosophy of RANSAC is to use as little data as possible to train a model while attempting to remove the errors (Fischler and Bolles 1981). Thus, to find the optimum percentage of the pseudo training points, different percentage numbers including 10%, 15%, 25%, 50%, 70% and 90% were tested over the 2016 Fredericton and 2019 Gatineau datasets. These pseudo training points are used to train the R.F. model. Figure 3-11 demonstrates the overall accuracy of the R.F. with different training percentages over the test subset. For comparison only, we displayed the HAND model accuracy in the Figure as well.



(a)



(b)

Figure 3-11 The overall accuracy of different training percentages over validation dataset tested for Fredericton (a) and Gatineau (b) datasets.

Since the datasets for Fredericton flood events include the same number of ground truth points, different pseudo training subsets were tested over Fredericton 2016 flood event. As is evident in Figure 3-11 (a), the highest overall accuracy, 92.86%, was obtained using 10% pseudo training points and as the training subset was increased, the overall accuracy has become more similar to that of the HAND model. In the Fredericton dataset, where the pseudo training subset was 70%, an overall accuracy of 91.29% was obtained and where it was 90% an overall accuracy of 91.14% was obtained which is similar to the HAND model overall accuracy. In the Gatineau dataset, the highest overall accuracy, 93.94%, was obtained by 15% training data. By increasing the pseudo training percentage to 25%, the overall accuracy reaches 93.48% which is slightly higher than the HAND with 93.03% overall accuracy. As shown in Figure 3-11 (b), other training amounts achieved overall accuracies fairly similar to that of the HAND model while the 15% one attained the highest rate. The results of this study showed how the HAND model can provide pseudo training data for R.F. to model flood events. Through this integration, the required training data for R.F. was obtained and the flood events were modeled with a higher accuracy than by the HAND model.

In this study, we proposed to improve the performance of the HAND model in flood mapping by integrating it with one of the most robust machine learning models, R.F. The R.F. model is used in different studies including (Rahmati, Pourghasemi, and Melesse 2016; S. Lee et al. 2017) for generating flood susceptibility maps by using the geospatial information of abundant available inundated locations as training data. In these studies, the R.F. model makes predictions based on the data related to previous flood events; therefore, the model provides a general assessment of areas that are at risk of flooding by using datasets from multiple flood events. However, the required training data for R.F. might not be always available for all the locations, which makes the use of R.F. very strict in this context. Hence, we proposed to use a simplified conceptual model to provide the training data for R.F. to dynamically identify the inundated areas.

3.5 Conclusion

Near-real time flood mapping requires efficient and accurate estimations. Previous techniques for flood mapping either lack the necessary precision or are very costly and intricate for rapid implementation. In this research, a new model, named PS-RF, was developed for flood mapping through the integration of a simplified conceptual model, HAND and a robust machine learning model, R.F. This model was tested on 5 different flood events occurred in Fredericton and Gatineau achieving higher overall accuracy in all the events compared to the HAND model. First, the HAND model was used to simulate the extent of the flood events based on the river water level. To improve the accuracy of flood maps obtained from the HAND model, the RANSAC methodology was used to filter out the outliers from pseudo ground truth points and maintain the most established ones. Then, those most certain pseudo ground truth points were used as pseudo training points for the R.F. model. Next, the trained R.F. model was used to estimate the overall flood extent in the area. The proposed model obtained higher accuracy than the HAND model in all five years of interest. Also, by integrating the HAND model and R.F., the required training data for the R.F. model were provided through the HAND. This approach can be implemented in areas with different climate and topography where a high resolution DTM is available; however, varied conditioning factors might be selected for other regions depending on the characteristics of a region. In our future work, we will consider how the feature selection can be further automated using different optimization algorithms which will result in a more complete flood mapping procedure.

3.6 References

Amer, R., Kolker, A. S., & Muscietta, A. (2017). Propensity for erosion and deposition in a deltaic wetland complex: Implications for river management and coastal restoration. Remote Sensing of Environment, 199, 39-50.

CBC News · Posted: 2018 5:41 PM AT | Last Updated: April 30, and 2018. 2018. "Worst Floods in New Brunswick History: How 2018 Compares | CBC News." CBC. April 30, 2018. https://www.cbc.ca/news/canada/new-brunswick/st-john-river-flooding-history-1.4641969.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Brunner. 2016. "HEC-RES River Analysis System - User's Manual Version 5.0. US Army Corps of Engineers. Institute ForWater Resources, Hydrologic Engineering Center (HEC), p. 962." 2016.

Brunner, G. W., Piper, S. S., Jensen, M. R., & Chacon, B. (2015). Combined 1D and 2D hydraulic modeling within HEC-RAS. In World Environmental and Water Resources Congress 2015 (pp. 1432-1443).

Campbell, J. B., & Wynne, R. H. (2011). Introduction to remote sensing. Guilford Press.

Chow, C., Twele, A., & Martinis, S. (2016, October). An assessment of the Height Above Nearest Drainage terrain descriptor for the thematic enhancement of automatic SAR-based flood monitoring services. In Remote Sensing for Agriculture, Ecosystems, and Hydrology XVIII (Vol. 9998, p. 999808). International Society for Optics and Photonics.

Congedo, L. (2016). Semi-automatic classification plugin documentation. Release, 4(0.1), 29.

Delgado, R., & Tibau, X. A. (2019). Why Cohen’s Kappa should be avoided as performance measure in classification. PloS one, 14(9), e0222916.

DHI, D. (2003). Mike-11: a modelling system for rivers and channels, reference manual. DHI–Water and Development, Horsholm, Denmark.

MIKE, D. (2012). 21-2D Modelling of Coast and Sea. DHI Water & Environment Pty Ltd.: Hørsholm, Denmark.

ERD 2015 Lidar.” n.d. Accessed June 24, 2020. https://geonb.snb.ca/downloads2/lidar/2015/erd/meta/erd2015.html.

Esfandiari, M., Jabari, S., McGrath, H., & Coleman, D. (2020). FLOOD MAPPING USING RANDOM FOREST AND IDENTIFYING THE ESSENTIAL CONDITIONING FACTORS; A CASE STUDY IN FREDERICTON, NEW BRUNSWICK, CANADA. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences, 5(3).

Falah, F., Rahmati, O., Rostami, M., Ahmadisharaf, E., Daliakopoulos, I. N., & Pourghasemi, H. R. (2019). Artificial neural networks for flood susceptibility mapping in data-scarce urban areas. In Spatial modeling in GIS and R for Earth and Environmental Sciences (pp. 323-336). Elsevier.

Fathollahi, F., & Zhang, Y. (2020). Adaptive band selection for pan-sharpening of hyperspectral images. International Journal of Remote Sensing, 41(10), 3924-3947.

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 24(6), 381-395.

Gaur, A., Gaur, A., & Simonovic, S. P. (2018). Future changes in flood hazards across Canada under a changing climate. Water, 10(10), 1441.

Ghoneim, E., & Foody, G. M. (2013). Assessing flash flood hazard in an arid mountainous region. Arabian journal of Geosciences, 6(4), 1191-1202.

Haq, M., Akhtar, M., Muhammad, S., Paras, S., & Rahmatullah, J. (2012). Techniques of remote sensing and GIS for flood monitoring and damage assessment: a case study of Sindh province, Pakistan. The Egyptian Journal of Remote Sensing and Space Science, 15(2), 135-141.

Ihsan, Z., Idris, M. Y., & Abdullah, A. H. (2013). Attribute normalization techniques and performance of intrusion classifiers: A comparative analysis. Life Science Journal, 10(4), 2568-2576.

Jabari, S., Fathollahi, F., Roshan, A., & Zhang, Y. (2017). Improving UAV imaging quality by optical sensor fusion: an initial study. International Journal of Remote Sensing, 38(17), 4931-4953.

Khosravi, K., Nohani, E., Maroufinia, E., & Pourghasemi, H. R. (2016). A GIS-based flood susceptibility assessment and its mapping in Iran: a comparison between frequency ratio and weights-of-evidence bivariate statistical models with multi-criteria decision-making technique. Natural Hazards, 83(2), 947-987.

Khu, S. T., Liong, S. Y., Babovic, V., Madsen, H., & Muttil, N. (2001). Genetic programming and its application in real-time runoff forecasting 1. JAWRA Journal of the American Water Resources Association, 37(2), 439-451.

Kia, M. B., Pirasteh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A., & Moradi, A. (2012). An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. Environmental Earth Sciences, 67(1), 251-264.

Lee, D. H. (2013, June). Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In Workshop on challenges in representation learning, ICML (Vol. 3, No. 2).

Lee, S., Kim, J. C., Jung, H. S., Lee, M. J., & Lee, S. (2017). Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. Geomatics, Natural Hazards and Risk, 8(2), 1185-1203.

Lhomme, J., Sayers, P., Gouldby, B. P., Samuels, P. G., Wills, M., & Mulet-Marti, J. (2008). Recent development and application of a rapid flood spreading method.

Liu, Y. Y., Maidment, D. R., Tarboton, D. G., Zheng, X., & Wang, S. (2018). A CyberGIS integration and computation framework for high-resolution continental-scale flood inundation mapping. JAWRA Journal of the American Water Resources Association, 54(4), 770-784.

Liu, Y. Y., Maidment, D. R., Tarboton, D. G., Zheng, X., Yildirim, A., Sazib, N. S., & Wang, S. (2016). A CyberGIS approach to generating high-resolution height above nearest drainage (HAND) raster for national flood mapping.

Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochimica et Biophysica Acta (BBA)-Protein Structure, 405(2), 442-451.

CBC News · Posted: 2019 4:00 AM ET | Last Updated: May 4, and 2019. 2019. "Under Water, Again | CBC News." CBC. May 4, 2019. https://www.cbc.ca/news/canada/ottawa/ottawa-river-flooding-2019-recap-1.5119980.

McGrath, H., Stefanakis, E., & Nastev, M. (2016). Rapid risk evaluation (ER 2) using MS excel spreadsheet: A case study of Fredericton (New Brunswick, Canada). ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 3, 27.

McGrath, H., Bourgon, J. F., Proulx-Bourque, J. S., Nastev, M., & El Ezz, A. A. (2018). A comparison of simplified conceptual models for rapid web-based flood inundation mapping. Natural Hazards, 93(2), 905-920.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. Biochemia medica, 22(3), 276-282.

Messner, F., & Meyer, V. (2006). Flood damage, vulnerability and risk perception–challenges for flood damage research. In Flood risk management: hazards, vulnerability and mitigation measures (pp. 149-167). Springer, Dordrecht.

Momo, M. R., Pinheiro, A., Severo, D. L., Cuartas, L. A., & Nobre, A. D. (2016). Desempenho do modelo HAND no mapeamento de áreas suscetíveis à inundação usando dados de alta resolução espacial. RBRH, 21(1), 200-208.

Monaghan, J. J. (1994). Simulating free surface flows with SPH. Journal of computational physics, 110(2), 399-406.

Moulinec, C., Denis, C., Pham, C. T., Rougé, D., Hervouet, J. M., Razafindrakoto, E., ... & Gu, X. J. (2011). TELEMAC: An efficient hydrodynamics suite for massively parallel architectures. Computers & Fluids, 51(1), 30-34.

Nicholls, R., Zanuttigh, B., Vanderlinden, J. P., Weisse, R., Silva, R., Hanson, S., ... & Koundouri, P. (2015). Developing a holistic approach to assessing and managing coastal flood risk. In Coastal Risk Management in a Changing Climate (pp. 9-53). Butterworth-Heinemann.

Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A., & Saleska, S. (2011). Height Above the Nearest Drainage–a hydrologically relevant new terrain model. Journal of Hydrology, 404(1-2), 13-29.

Nobre, A. D., Cuartas, L. A., Momo, M. R., Severo, D. L., Pinheiro, A., & Nobre, C. A. (2016). HAND contour: a new proxy predictor of inundation extent. Hydrological Processes, 30(2), 320-333.

Ottawa Riverkeeper's River Report. 2006. Ecology and Impacts.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.

Pourghasemi, H. R., Kariminejad, N., Amiri, M., Edalat, M., Zarafshar, M., Blaschke, T., & Cerda, A. (2020). Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. Scientific reports, 10(1), 1-11.

Pradhan, B. (2010). Flood susceptible mapping and risk area delineation using logistic regression, GIS and remote sensing. Journal of Spatial Hydrology, 9(2).

Pradhan, B., Hagemann, U., Tehrany, M. S., & Prechtel, N. (2014). An easy to use ArcMap based texture analysis program for extraction of flooded areas from TerraSAR-X satellite image. Computers & geosciences, 63, 34-43.

Prakash, M., Rothauge, K., & Cleary, P. W. (2014). Modelling the impact of dam failure scenarios on flood inundation using SPH. Applied Mathematical Modelling, 38(23), 5515-5534.

Rahmati, O., & Pourghasemi, H. R. (2017). Identification of critical flood prone areas in data-scarce and ungauged regions: a comparison of three data mining models. Water resources management, 31(5), 1473-1487.

Rahmati, O., Pourghasemi, H. R., & Melesse, A. M. (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.

Rahmati, O., Pourghasemi, H. R., & Zeinivand, H. (2016). Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. Geocarto International, 31(1), 42-70.

Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., & Tomasella, J. (2008). HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. Remote Sensing of Environment, 112(9), 3469-3481.

Schiermeier, Q. (2011). Increased flood risk linked to global warming.

Schumann, G., Bates, P. D., Horritt, M. S., Matgen, P., & Pappenberger, F. (2009). Progress in integration of remote sensing–derived flood extent and stage data and hydraulic models. Reviews of Geophysics, 47(4).

Secretariat, Treasury Board of Canada, and Treasury Board of Canada Secretariat. n.d. “High Resolution Digital Elevation Model (HRDEM) - CanElevation Series - Open

Government Portal." Accessed June 11, 2020. https://open.canada.ca/data/en/dataset/957782bf-847c-4644-a757-e383c0057995.

Sidrane, C., Fitzpatrick, D. J., Annex, A., O'Donoghue, D., Gal, Y., & Biliński, P. (2019). Machine learning for generalizable prediction of flood susceptibility. arXiv preprint arXiv:1910.06521.

Smith, L. C. (1997). Satellite remote sensing of river inundation area, stage, and discharge: A review. Hydrological processes, 11(10), 1427-1439.

Speckhann, G. A., Borges Chaffe, P. L., Fabris Goerl, R., Abreu, J. J. D., & Altamirano Flores, J. A. (2018). Flood hazard mapping in Southern Brazil: a combination of flow frequency analysis and the HAND model. Hydrological Sciences Journal, 63(1), 87-100.

Tarboton, D. 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.

Tarboton, D. G., & Mohammed, I. N. (2013). TauDEM 5.1 quick start guide to using the TauDEM ArcGIS toolbox. Utah State University (USU), United State.

Tehrany, M. S., Jones, S., & Shabani, F. (2019). Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. Catena, 175, 174-192.

Tehrany, M. S., Lee, M. J., Pradhan, B., Jebur, M. N., & Lee, S. (2014). Flood susceptibility mapping using integrated bivariate and multivariate statistical models. Environmental earth sciences, 72(10), 4001-4015.

Tehrany, M. S., Pradhan, B., & Jebur, M. N. (2013). Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. Journal of Hydrology, 504, 69-79.

Tehrany, M. S., Pradhan, B., & Jebur, M. N. (2014). Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. Journal of hydrology, 512, 332-343.

Tehrany, M. S., Pradhan, B., & Jebur, M. N. (2015). Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. Stochastic environmental research and risk assessment, 29(4), 1149-1165.

Tehrany, M. S., Pradhan, B., Mansor, S., & Ahmad, N. (2015). Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. Catena, 125, 91-101.

Teng, J., Jakeman, A. J., Vaze, J., Croke, B. F., Dutta, D., & Kim, S. (2017). Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. Environmental Modelling & Software, 90, 201-216.

Thistlethwaite, J., Henstra, D., Brown, C., & Scott, D. (2018). How flood experience and risk perception influences protective actions and behaviours among Canadian homeowners. Environmental management, 61(2), 197-208.

Toth, E., Brath, A., & Montanari, A. (2000). Comparison of short-term rainfall prediction models for real-time flood forecasting. Journal of hydrology, 239(1-4), 132-147.

Vacondio, R. B. S. P. M. P., Rogers, B. D., Stansby, P. K., & Mignosa, P. (2012). SPH modeling of shallow flow with open boundaries for practical flood simulation. Journal of Hydraulic Engineering, 138(6), 530-541.

Ye, J., & McCorquodale, J. A. (1998). Simulation of curved open channel flows by 3D hydrodynamic model. Journal of Hydraulic Engineering, 124(7), 687-698.

Zhang, Y. (2004). Understanding image fusion. Photogramm. Eng. Remote Sens, 70(6), 657-661.

Zhao, G., Pang, B., Xu, Z., Peng, D., & Xu, L. (2019). Assessment of urban flood susceptibility using semi-supervised machine learning model. Science of The Total Environment, 659, 940-949.

Zhao, G., Pang, B., Xu, Z., Yue, J., & Tu, T. (2018). Mapping flood susceptibility in mountainous areas on a national scale in China. Science of the Total Environment, 615, 1133-1142.

---

1. https://www.agr.gc.ca/atlas/aci/

2. https://ottawa.ca/en/city-hall/get-know-your-city/open-data

Chapter 4: PS-RF and GeoNB Flood Maps

4.1 Introduction

After finalizing the results of this research, GeoNB released the flood maps related to the 2018 flood event on August 14th, 2020, which were obtained using the latest version of Hydrologic Engineering Center's River Analysis System (HEC-RAS) model. As reported by GeoNB[i], the flood map was generated by dividing the entire Saint John River into 83 sections and applying maximum water levels to each of those sections. The gauge stations along Saint John River and the data from Hydrology Center's hydraulic model were used for flood forecasting and analysis. Also, the extent of the 2018 flood map was validated by survey teams and aerial imagery. GeoNB did not disclose the date for which this flood map is generated. However, it is mentioned that the flood extent shown represents approximately the highest water level reached during the 2018 flood event from the Mactaquac Dam to the Saint John Harbour.

4.2 Results and Discussion

To assess the performance of the proposed method in this thesis, we ran two tests. In Test 1, we compared the accuracy of PS-RF and the flood maps from GeoNB against the ground truth data obtained from satellite imagery. As reported in Section 3.2.1, the ground truth data used in this study was obtained during the flood peak, and it is also reported by GeoNB that their model was generated during flood peak. Therefore, we assume that they reflect the same flooded area at approximately the same time.

The results showed that the GeoNB flood map has a better performance with 93.3% overall accuracy compared to PS-RF with an overall accuracy of 89.6%. Both models have significant accuracy in detecting Flooded points with TPR of 98.6% and 99.1% for PS-RF and GeoNB flood extents, respectively. However, the method applied by GeoNB has performed better in identifying Not-Flooded points with a TNR of 87.4% compared to 80.6% for PS-RF. As can be seen in Figure 4-1, PS-RF is overestimating mostly in flat areas and detecting Not-Flooded points as Flooded.

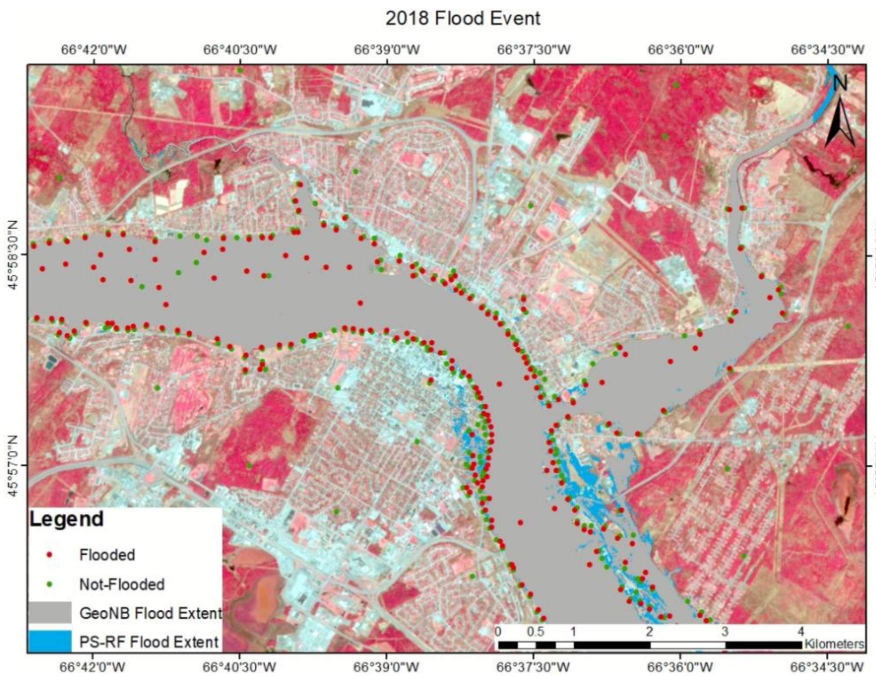


Figure 4-1 Flood maps obtained from PS-RF (in blue) and GeoNB (in grey) representing the flood extent in the 2018 flood event during the flood peak. The red dots represent flooded, and green dots represent not-flooded ground truth data.

Accuracy assessment metrics are also provided for this comparison and can be found in Figure 4-2.

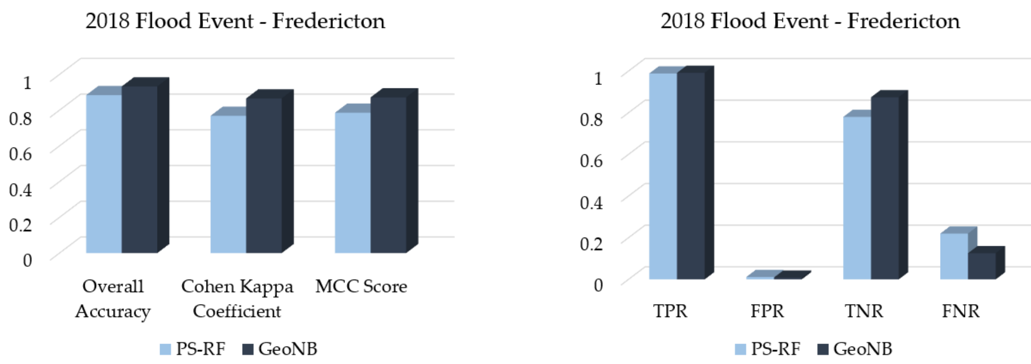


Figure 4-2 Assessment metrics for GeoNB and PS-RF flood extents for 2018 flood event.

In Test 2, we considered the GeoNB flood map as the ground truth and compared the PS-RF flood extent against it. The results showed that the flood map provided by GeoNB in our study area had an area of 31.7 square kilometers from which more than 99% was correctly identified as flooded by PS-RF (GeoNB and PS-RF flood maps shared an overlapping area of 31.4 square kilometers). As can be seen in Figure 4-1, less than 1% of the flooded area reported by GeoNB is missing in PS-RF predictions. In addition, the PS-RF model has identified an area of 2.3 square kilometers as flooded, which are not identified as flooded by GeoNB and can be due to minor water level difference between two flood maps and limited accuracy of sample points selected from satellite imageries with limited spatial resolution. As reported in section 3.2.2, the spatial resolution of satellite images for the 2018 flood event is 10 meters which makes it difficult to precisely distinguish water boundary.

The flood maps generated by GeoNB are simulated by the HEC-RAS model, which simulates the water flow through the open channels and computes water surface profiles. The process is computationally expensive, which requires multiple parameters, including a description of the stream network containing data relevant to all sections and a description of the stream network containing reach locations and connectivity. Also, a description of the model cross-sections containing the location and geometric data as well as additional HEC-RAS modeling information is required. Moreover, it is considerably site-specific, and utilizing it in different locations might complicate the simulation procedure depending on the complexity of the topography of a region. PS-RF model was validated by testing it in five different flood events in 2014, 2016, 2018, and 2019 in Fredericton and in 2019 in Gatineau.

On the other hand, the proposed method in this study has a straightforward workflow with specific input parameters, which were identified and reported in the second chapter of this thesis. Also, the necessity of training data for the R.F. model was handled through the utilization of a hydrological model called HAND and a mathematical model called RANSAC, which was reported in the third chapter. As reported in section 3.4, the proposed model can make dynamic predictions.

Chapter 5: Further Discussion and Conclusion

This chapter summarizes the research conducted for this thesis. First, the summary of the research from chapters 2 and 3 is presented in this section. At the end, some suggestions for future work are provided.

4.1 Summary of Research

This research was conducted to provide an accurate, efficient, and dynamic flood modeling approach through the integration of different conditioning factors, a simplified conceptual flood model called HAND and machine learning. For this study, we used one of the most robust machine learning algorithms, Random Forest (R.F.). In the course of this study, as the first step, the flood conditioning factors that are cited most frequently in the literature, namely, altitude, slope, aspect, distance from river, land use/cover, TWI, TRI, SPI, curvature, plan curvature, profile curvature, and the HAND model are generated. However, in any successful machine learning project, feature selection is essential, which can be a function of the specific classifier used. We focused on selecting features, which are called conditioning factors in this study, that best suit the R.F. classifier.

The remainder of this section explains how, using a R.F. classifier, the essential conditioning factors reflecting inundated areas were selected. The conditioning factors selection was performed through the feature-ranking capability of R.F. For this purpose, we first collected training points from Sentinel-2 satellite imagery acquired around the peak flood date of the 2018 Fredericton flood event. Then, different combinations of conditioning factors were used for generating flood maps for that specific time. The conditioning factors were combined in two different scenarios. In the first scenario, the five most frequently used conditioning factors of altitude, slope, aspect, distance from river, and land use/cover were selected to generate the flood map. Then, the remaining conditioning factors were added one by one until all the conditioning factors were being used. Afterwards, the least important conditioning factors were removed one by one until there were only four conditioning factors left. The accuracies of the produced flood maps were observed and stored. In the first scenario, a total of 16 flood maps were generated. Figure 2-6 reports the accuracy of those flood maps. In the second scenario, different combinations of uncorrelated conditioning factors were used for generating flood maps. The accuracies of the produced flood maps were observed and stored as well. In the second scenario a total of 13 flood maps were generated.

The results of this research, which was published in XXIV ISPRS Congress (2020 edition) showed that the effectiveness of the HAND model in flood mapping using R.F., is the same as using the altitude layer. However, the highest accuracy is achieved if the altitude is used instead of the HAND model. In addition, using correlated layers has a negative impact on the estimations of flood extent and it must be avoided. This can be seen in Figure 2-7 where it is demonstrated that adding extra conditioning factors gradually decreases accuracy.

This method can only represent the most effective conditioning factors fitting best to the ground truth data selected from satellite imagery. In other words, if we did not have optical images of the area for extracting training points, we could not have mapped flooded areas. However, with this study, we identified the best combination of conditioning factors using R.F. Classifier, which are slope, aspect, distance from river, land use/cover, and either altitude or HAND model.

In the second step of this research, it was aimed to provide an accurate flood prediction by integrating the best flood conditioning factors, i.e., altitude, slope aspect, distance from river, land use/cover, and the R.F. algorithm. The objective of this part of the research was to use machine learning to estimate the inundated areas using the water depth of the river. However, machine learning methods require abundant training data, which is either dependent on the availability of flood event optical satellite images, or must be collected through other approaches including aerial photography, site visit, surveying, etc. Still, the training data provided from the satellite images can generate flood extent at one instance of time. i.e., image acquisition time, only. To overcome this challenge, the HAND model was used to provide training data for the R.F. classifier. The HAND model uses the river water level obtained from gauge stations and estimates the inundated areas dynamically; therefore, it is possible to both predict and simulate flood events automatically. The label of the HAND-based training data is selected as flooded/non-flooded using the HAND model flood maps. However, since the accuracy of the HAND model is limited, we call these training samples as *pseudo training data* which needed to be filtered to improve the accuracy of the final classifier.

In this process, RANSAC was used to detect the outliers of the *pseudo training data* samples. This method is called PS-RF. To test this model, a total of 6 flood events were selected: 5 flood events in Fredericton in 2014, 2016, 2018, 2019, and one flood event in Gatineau in 2019. The *pseudo labels* used in this process are obtained from the thresholded HAND model, which is obtained by subtracting the river water level at the time of flood and at the time of LiDAR data collection. Thus, in the first step, all the control points were assigned flooded and not-flooded labels based on the predictions of the HAND model. As demonstrated by the results of Cohen's Kappa Coefficient and Matthews Correlation Coefficient, the PS-RF method achieved higher accuracy and performed better than the HAND model in all five flood events. In the 2014 flood event in Fredericton, PS-RF obtained an overall accuracy of more than 94%, which is nearly 6% higher than what the HAND model could predict. Similar results were obtained in the flood events that occurred in Fredericton in the years 2016, 2018, and 2019 with 5%, 3%, and 5% higher overall accuracy, respectively. To further verify the validity of the proposed method, the PS-RF was tested in a different location, the Gatineau 2019 flood event, and an overall accuracy of nearly 95% was obtained for flood prediction from PS-RF compared to 93% from the HAND model.

In this study, we proposed an efficient flood mapping procedure by first identifying the essential conditioning factors and presented a machine learning-based approach in which a hydrological model called HAND is used to provide the required training data. The results demonstrated that the proposed workflow was able to provide highly satisfactory flood maps efficiently and effectively.

4.3 Suggestions for Future Work

Based on this research, the following recommendations for further enhancing the results are suggested:

- The process of feature selection in machine learning can be further automated using methods like genetic or particle swarm optimization algorithms.

- With the current progress in the field of deep learning, it is possible to eliminate the manual tasks from this procedure and apply deep learning methods for flood mapping.
- The hydrologic model that was used in this study and many of the conditioning factors are generated from DTM. However, such high resolution DTMs are not always available for different places around the world. So, by employing deep learning and satellite imagery, first, a DTM is generated then the process of flood mapping can be carried out through the previously proposed approach.

---

[1]. https://www2.gnb.ca/content/gnb/en/news/news_release.2020.08.0446.html

Curriculum Vitae

Candidate's full name: Morteza Esfandiari

Universities attended:

2018 – Present, MSc Candidate, University of New Brunswick, NB, Canada
2015 – 2017, B.Sc. Geomatics Engineering, Islamic Azad University of South Tehran Branch, Tehran, Iran
2013 – 2015, A.Eng. Surveying Engineering, Darolfonoon Higher Education Center, Qazvin, Iran

Publications:

Esfandiari, M., Jabari, S., McGrath, H. and Coleman, D., 2020. FLOOD MAPPING USING RANDOM FOREST AND IDENTIFYING THE ESSENTIAL CONDITIONING FACTORS; A CASE STUDY IN FREDERICTON, NEW BRUNSWICK, CANADA. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 5(3).

Esfandiari, M., Abdi, G., Jabari, S., McGrath, H. and Coleman, D., 2020. Flood Hazard Risk Mapping Using a Pseudo Supervised Random Forest. *Remote Sensing*, 12(19), p.3206.

Esfandiari, M., Abdi, G., Jabari, S., and Lolla, Vasuki., "Building change detection in off-nadir images using deep learning." In *IGARSS 2021 - 2021 IEEE International Geoscience and Remote Sensing Symposium* (ACCEPTED).

Abdi, G., **Esfandiari, M.**, and Jabari, S., "Building damage detection in post-event high-resolution imagery using deep transfer learning." In *IGARSS 2021 - 2021 IEEE International Geoscience and Remote Sensing Symposium* (ACCEPTED).

Tettey-Asare, M., **Esfandiari, M.**, Jabari, S., Flood Damage Assessment Using 3D City Models - A Case Study in Fredericton, New Brunswick, Canada. In *CSRS 2021 - 2021 Canadian Symposium of Remote Sensing* (ACCEPTED).

Conference Presentations:

M. Esfandiari, H. McGrath, S. Jabari, (2020) **"**Urban Flood Modeling Using Height Above Nearest Drainage (HAND) Model and Random Forest; A Case Study in Fredericton, New Brunswick, Canada", Jan 30th, 2020, Saint John, NB: NRCAN Flood Mapping Workshop.