Exploring the Economic Impact of Land Use Changes on Flood Events on the Ottawa River from 2004-2017

Iffat Damji, Shreejan Paudel, Raymond Tran

04/06/2018

Table of Contents

Introduction 3
How the Sustainable Development Goals fit in with the research question 6
Data acquisition 7
Methodology 12
Flood extent mapping: 12
Modelling Soil Runoff: 16
Computing Damages: 26
Results 36
Flood Extent Delineation: 36
Damage outcomes 52
Land Use Classification Results 62
Runoff Modelling Results: 65
Discussion 70
Flood extent delineation: 70
Damage Assessment: 74
Land Use and Runoff Modelling Discussion 79
Conclusion 88
References 92

Introduction

The Ottawa River is one of the most important rivers within North America, stretching over 1200 km and cutting through many major communities in Eastern Canada such as Downtown Ottawa and Gatineau, Quebec. This river has an incredible watershed basin area of 146,300 km^2, flowing through from Lake Timiskaming, to Downtown Ottawa and Montreal, and takes in waters from the Gatineau river and the Rideau river, two other large tributaries. A good majority of the Ottawa river defines the borders between Quebec and Ontario. With a rich history of trade and power generation for Ottawa and Quebec (Britannica, 2018). Furthermore, it is a major tributary of the St. Lawrence River, the largest watershed basin in Canada, which includes in its basin area the five Great Lakes. As such, the Ottawa River plays a large role in the hydrologic processes of Canada.

One of these hydrologic processes that can cause incredible amounts of damages is flooding. The Ottawa River has been very prone to major yearly floods for centuries, which can be linked to major precipitation events. With the increase of global warming, the amount of precipitation seen also increases, due to the water holding capacity of the warmer atmosphere; with every 1C increase, the atmospheres' water holding capacity increases by 7% (Stocker et al., 2013). While these precipitation events happen throughout the year, the most notable flooding events occur in the end of winter. As temperatures increase at the end of winter, the amount of snowmelt that contributes to the river discharge due to surface water runoff increases. This, paired with the increased amount of snowmelt caused by precipitation, as well as its contribution of river flow and increasing soil saturation, can cause incredible amounts of damage to those communities which are living within the major floodplains of the Ottawa River. Furthermore,

increasing temperatures further increase the snowmelt that can be seen, further contributing to the runoff.

Flood management techniques have been implemented, such as flood dams and reservoirs. However, these mitigation techniques are at times incapable of restraining the river stage of major flood events. It is likely the mitigation methods were not designed for the rate of increase of precipitation runoff that was seen, and that these figures were simply underestimated. For example, in May 2017, Downtown Ottawa saw a record breaking amount of rainfall, over 100 mm more rain than the previous year in the same timeframe, causing an incredible amount of damage (Insurance Bureau of Canada, 2017). However, a flood in 2009 was categorized as a 100-year flood event, meaning that in less than ten years apart, a 100-year flood event was seen twice. As such, questions have been raised as to the effectiveness of these mitigation methods, as well as future protection actions.

Due its vast reach, the Ottawa River cuts through many rapidly growing communities. Ottawa has a population growth rate of roughly 8.8% up until 2011, significantly higher than the Canadian average of 5.9% (City of Ottawa, 2011). Furthermore, the Ottawa population growth rate has increased over 10% from 2011 to 2016. These high amounts of population growth speak to the increasing amount of urbanization that will occur, now and in the future. In addition to the increased amount of urbanization due to increased population, it would be expected that industrial land use would increase as well. The city of Ottawa is in a unique position in terms of land use as the city has the Greenbelt, an arc of land which runs through the urban center of the city. This arc is strictly controlled in terms of real estate development, and was developed to reduce the urban sprawl (National Capital Commision, n.d.). The Greenbelt consists of forestry,

wetlands, fields and agriculture, with large amounts of urbanization occurring on either side of the arc. The impact of the Greenbelt on flooding from both the downtown area and the outer urban areas is currently unknown.

While rising temperatures and increased rainfall are major factors in a flood event, another major factor that is at times overlooked is the topography of the watershed basin area. The methods in which rainfall and snowmelt flow across the topology of the study area in itself play a major factor in the flood events that are observed. Different surface materials, roughness, slope and vegetation are just some factors that can play a role in the hydrologic system. There have been studies examining the impact of urbanization and the increased usage of asphalt that both support and rebuff the concept of more urbanization increasing the amount of water runoff (Apolliono et al, 2016). As such, it is important to understand the impacts of land use change that are seen on the hydrological model of a watershed.

The purpose of this project is to examine the applicability of using remote sensing data and techniques in order to determine the extent of the damage the Ottawa River flood can incur on the surrounding landscapes, as well as the economic damage of buildings within the flood extent. Furthermore, this project aims to explore urban area land use change and their potential impacts on not only the hydrology of the area, but also the degree of impact of these changes on flood events. In determining the impact of land use change and its spatial relation to the river in terms of river flooding, as well as determining the economic flood damages that occur due to it, this project aims to not only create a workflow in determining economic damage per land use change, but also examine future techniques of determining flood mitigation

techniques in the city of Ottawa and all other communities within the floodplain of the Ottawa River.

How the Sustainable Development Goals fit in with the research question

The SDGs, also known as the Sustainable Development Goals, were formed by the United Nations in an effort to solve a multitude of global issues by the year 2030. These goals were a precursor to the Millennium Development Goals which were not fulfilled by the year 2015 (Thomson, n.d.). Although our research question is framed in terms of a local context, the extent of our question relates to issues that are occurring worldwide and has wider implications in terms of climate change and action. The SDGs are comprised of 17 goals, three of which can be integrated into the research problem at hand. First off, goal #13, Climate Action, can be integrated into the scope of our research question with regards to the first section of our report in terms of determining the reasoning behind dramatic seasonal changes and preventing the impact to the landscape that these changes have caused.

Goal #9, Industry, Innovation and Infrastructure, also relates to our research question because damages caused by flooding events could lead to governments having to restructure infrastructure that is currently not susceptible to environmental hazards. This could potentially lead to innovative flood damage mitigation techniques being implemented to minimize economic risks that these infrastructural systems can sustained in the future.

Lastly, goal #15, Life on Land, encompasses all of the actors that can be affected by river flooding. If the damages from a flooding event are substantial, citizens, insurance companies and government officials may have more responsibilities in terms of ensuring that the affected region can be restored to its natural condition. Citizens may be forced to relocate temporarily if their houses were damaged during the flood. They may also have to search for alternative modes of transportation if their vehicles were also damaged during a flooding event. Insurance companies will also be affected because they are at the forefront for providing citizens with support to effectively provide their customers with the services and funds to relocate temporarily or to even purchase new vehicles, if that is possible. Federal, provincial and local governments may also incur tremendous costs associated with maintenance of facilities after a flooding event.

Data acquisition

In order to achieve our first aim of the report by mapping flood extent during select peak flood events from 2004-2017, data was obtained and later assessed through the use of remote sensing software. This research set out to cover flooding events on September 10th, 2004; July 4th, 2006; April 5th, 2008; July 28th, 2009; January 25th, 2010 and April 25th-May 13th, 2017. However, we decided to focus more on the 2004, 2010 and 2017 events due to data and time constraints which made it infeasible to focus on analyzing all of the events. It is important to note that we were unable to be assess the flood extent for the flooding extent that occurred in 2010 due to the fact that the ENVI-SAT data that was obtained was unusable as it did not provide any information about the intensity or the amplitude for any of the polarization modes. As such, 2010 data was substituted with 2009 data.

Two types of remote sensing data were considered in order to effectively disseminate the flood extent for the selected dates. This included high resolution optical, multispectral images as well as passive and active microwave radar images. Much of the literature stated synthetic aperture radar data is crucial when it comes to accurately assessing the flood extent of a region during a peak flood event, therefore this type of data was prioritized in terms of delineating the flood extent of peak flood events. This type of remote sensing data takes precedence over optical imagery due to fact that optical sensors are unable to penetrate through cloud cover which is relatively high during these precipitation events.

It should be noted that optical data should not be completely dismissed as it has been used to verify land-use types on SAR images as verified by previous studies. The literature also made note of variations in sensor characteristics (i.e. incidence angle, polarizations and frequency bands) specifically for radar data as those have an effect on the effectiveness of a flood extent map. The literature stated that RADAR-SAT would be preferred since the user can specify a incidence angle that best fits their needs (Bonn et al., 2005). It also mentioned that "ERS-1/2, ENVI-SAT and TerraSAR-X with a 24 degree incidence angle were also considered because of their ability to detect flooding in urban areas" (Mason et al., 2010). The literature also stated that cross-polarization was preferred for flood extent mapping specifically in the HH polarization mode "because it was less sensitive to minor vertical differences on the water surface caused by waves" (Long et al., 2014). We aimed to obtain data that contained HH polarizations, however that was not always possible and therefore may have affected the accuracy of our analysis to some degree.

One of the issues that we encountered while collecting our data was that certain sensors were either discontinued or they were not yet in orbit during our chosen years. We had to work around that issue by utilizing ENVI-SAT data for the chosen years between 2004 & 2009 and utilizing Sentinel-1 data for the flooding event that occurred in Spring 2017. The European Space Agency's scientific hub was used to collect Sentinel-1 data during the April-May 2017 flood period although data was only available for May 12th which was towards the end of the high precipitation period. Therefore this study implemented ENVI-SAT ASAR C-band images with 30 meter resolution and Sentinel-1 Interferometric Wide Swath Level 1- Ground Range Detected image with 5 meter resolution.

Landsat-5 and 7 TM+ (Thematic Mapper) were also considered for the peak flood events but was not used since that data was unavailable for our chosen dates. Since SAR imagery was already obtained and did not need to be atmospherically corrected in order to remove the presence of high cloud cover, it was more beneficial to focus on those images in terms of delineating the flood extent during the peak flooding events. However, pan-sharpened SPOT 4 imagery with a 20 meter resolution obtained from GeoScholarsPortal that was provided by the Ontario Ministry of Natural Resources was used in terms of optical data. This imagery helped us decipher between trees and urban areas that appear to have similar characteristics in SAR imagery during flooding events. It should be noted that although the optical data was from August 30th, 2006 and not on the exact date of the flood event, this image is still within the same season as the 2006 flooding event. This means that not only should the land-use types appear to look the same but the climate variables such temperature and evaporation should be around the same as the day the flooding occurred. Unfortunately due to cost and time restraints, we were

unable to physically collect in-situ measurements during these peak precipitation events in the Ottawa River region which caused some issues in terms of validation. Therefore, we used supplementary shape files for the Ottawa side of the Ottawa River that provided us with thematic and topographic data that was missing from the optical and radar images. These were used as a point of reference as previous studies utilized GIS-based data to further enhance their analysis (Brivio et al., 2002). These files were obtained from the City of Ottawa's open data website which was provided by Carleton University. This included projected road segments, water bodies, tree inventory and parks that were created in 2013 and updated every year since. This data was utilized in order to create the four classes: water, unflooded vegetation, flooded vegetation and urban areas (that are often misclassified as flooded vegetation in SAR images) for the resulting flood extent maps. Although substantial information was obtained, we were unable to obtain LiDAR data which could have further enhanced our analysis but the literature provided other methods were just as effective as using LiDAR data to correct for differences in the terrain.

Landsat imagery was also used for land classification, collected for three dates, one preceding the temporal scale of the study period, in August of 2004, the middle of the study period in July 2009 and one proceeding the April 2017 floods, in August 2017. As this imagery would not be used to delineate the flood extent, rather to classify land, it was important for the time of year was kept as consistent as possible, however all imagery in August 2009 contained significant cloud cover. As such, in the interest of minimizing seasonal differences, imagery in July was selected, rather than September. For both 2004 and 2009 imagery, Landsat 5 mission data was used to keep the consistency between the imagery. While Landsat 7 imagery was considered, it was quickly determined to be unusable due to a failure in a significant component

of the satellite. In 2003, the Scan Line Corrector aboard the Landsat 7 satellites experienced a failure, causing data gaps throughout all imagery (USGS, n.d.). As such, Landsat 5 imagery was used as it was one of few satellites with easily accessible imagery at the time. With a spatial resolution of 30m by 30m, issues regarding resolution quickly arose. Other imageries such as Quickbird, with a panchromatic spatial resolution of 0.65m, and IKONOS, with a panchromatic spatial resolution of 0.82m, were considered. However, major limitations in terms of collecting cloud free, low cost imagery were encountered. With a requirement to receive imagery as far back as 2004, free, high resolution imagery was limited. As such, Landsat 5 was considered as one of few choices.

In terms of the runoff modelling, the model required soil hydrology data. In terms of the Ontario data, this was easily acquired through the Ministry of Natural Resources of Ontario, who compiled the data in January of 2018. This data included a breakdown of the hydrologic soil type as well as a short description and many other miscellaneous information for the soil. These soil types were defined by the USDA and describe the runoff potential of the soil when thoroughly wet (Mockus & Werner, 1994). While the Ontario data was easily collected, the soil information for Quebec was not freely available. While the data had been collected by the Institut de Recherche et de Développement en Agroenvironment (IRDA), the budget of this project did not allow for the steep costs of the soil data. As such, hydrologic soil groups were forced to be interpolated based on the supplementary data. Geological survey data of Canada from the Natural Resources of Canada was used to estimate the hydrologic soil group based on the contents of the surficial materials and the make-up descriptors of the soil groups. Lastly, the USDA Hydrologic for Small Watersheds Technical Release 55 (SCS TR55) report was used as a

simplified estimation of runoff depending on soil type and land use (Croshney et al. 1986). This report is was created by the USDA in order to examine the effects of soil based on urban and urbanizing watersheds, and provides a general run off number to describe soil groups of a particular land use.

The ENVI SAR data were to be used for the flood extent mapping. Optical data were to be used for land classification. SAR and Optical data were both used for flood damage calculation, since the extent calculated from SAR was used for computing total damage. In addition to these data, water depth data was also collected from Department of Environmental and Natural Resources of Canada. However, the water stations did not cover the entire study area, therefore, additional methods were used to extrapolate new data points and interpolate field values based on these data points.

Methodology

Flood extent mapping:

Various studies on flood extent mapping provided us with a guide to effectively assess the flood extent of an area during a peak precipitation or flood event. The majority of the studies that were reviewed for the purpose of this research mentioned that the intensity images were the most crucial for flood extent mapping as they provide us with the backscatter coefficients that allow us to characterize land based versus water entities. The literature also mentioned that pre-processing techniques such as radiometric correction, multi-looking, terrain correction and filtering to remove noise and speckle as that had a significant effect on the overall accuracy of a binary flood extent map. For this reason, using the literature the following methods and

procedures were utilized in order to create flood extent maps for selected years within the chosen time frame of 2004-2017.

Existing tutorials from the University of Waterloo and the European Space Agency provided us with the necessary steps for pre-processing the 2004, 2006, 2008, 2009 ENVI-SAT ASAR C-band images. As mentioned before two remote sensing softwares were utilized for this section of the report. One was Sentinel's Application Platform that is used to calibrate and rectify synthetic aperture radar imagery. ENVI was also utilized for used for filtering, classification and accuracy assessment purposes that will be discussed in further detail later on in this report.

As mentioned before, filtering is a necessary step in ensuring that unnecessary noise and speckle is removed from the images obtained by a radar sensor. This is essential in terms of ensuring that the produced flood extent maps are effective and accurate. Two factors influence the use of the two different filters that were used in this analysis. First, based on the filters that were used in literature, two different filters were used; one was the gamma filter with a 5x5 window and 2 looks (Long et al., 2014), and the other was the median filter with a kernel size that were determined through trial and error and visual interpretation. The second factor was the functionality of the filters as stated by ENVI. In this case it was critical that we used filters that group local pixels together since similar pixels (areas of flooding) may be closer to one another as stated by Tobler's first law of geography, which was mentioned in one of studies that was reviewed (Li et al.,2016). The first filter used was the adaptive gamma filter; the purpose of this filter is to "reduce speckle while preserving edges in radar images as well as replacing pixels with a value calculated based on the local statistics"("Filters", n.d.). The second filter, was the median convolution filter; the purpose of this filter was similar to the gamma filter in which it

also is meant to “reduce speckle while preserving edges larger than the kernel dimensions” while also smoothing the image (“Filters”, n.d.). In this case using a kernel size greater than the default size of 3x3 caused the image to look blurry which would make it even more difficult to interpret.

First, the gamma filter with a 5 by 5 window and 2 looks was utilized and then a median convolution filter with a kernel size of 3 by 3 was used. Both filters effectively removed the noise and speckle distinguished between the bright objects and the dark objects. In terms of classification methods, there were a variety of non-traditional strategies that used automatic classification algorithms and classifications using fuzzy logic. These were however beyond the scope of this research and require an extensive amount of data which we were unable to obtain due to time and cost constraints. Since there was very little variance amongst the final filtered images, only one of them was classified using supervised classification. Based on remote sensing principles about the functionality of radar sensors that allow us to interpret SAR data, we were able to classify the pixels as follows. Since we know that open standing water has a smooth surface, we know that it acts as specular reflector and therefore has a low backscatter return hence why water bodies appear to be dark in SAR images. We also know that urban areas such as buildings and roads have rough surfaces and act as diffuse reflectors that reflect energy back in multiple directions which is why they have such high backscatter returns. Flooded vegetation (canopies and forests) on the other hand have unique characteristics and implement what is called the double bounce scatter effect which makes them appear bright in SAR images as well. This caused issues for our research as we had difficulty differentiating between flooded vegetation and urban areas since they both appear bright/white in the SAR images. For the purpose of this study, the grey pixels were classified as grey pixels were labelled as un-flooded

vegetation since we were primarily concerned with flooded vegetation and urban areas near water bodies.

Since we had prior knowledge of our study area through the use of optical imagery, GIS data and those remote sensing principles, we were then able to create different sets of ROIs (region of interests) to use as our training data. The first set included 14 Water ROI's distributed around all of the water bodies with that extent of our study area that appear dark in both the SAR and optical imageries. It also included 10 flooded area ROI's that were in close proximity to water bodies. These cells appeared bright in the SAR imagery, green in the optical imagery and were within the tree inventory points from the tree inventory shape file that was previously mentioned. Lastly, this set included 8 un-flooded vegetation pixels that were away from water bodies, appeared grey in the SAR imagery and red in the optical imagery. The separability between each of the ROI's were computed and each class had approx. a 1.9 to 2.0 separability meaning that these classes were different enough that they had a minimal chance of being misclassified as one another.

The second set of ROI's included 9 Water ROI's distributed around all of the water bodies with that extent of our study area that appeared dark in both the SAR and optical imageries. It included 7 tree/flooded vegetation ROI's that were in close proximity to water bodies. These cells appeared bright in the SAR imagery, green in the optical imagery and were also within the tree inventory points. It also included 4 urban area ROI's that appeared cyan in the optical imagery and bright/white in the SAR imagery. Lastly, this set included 5 un-flooded vegetation pixels that were away from water bodies, appeared grey in the SAR imagery and red in the optical imagery. The separability between each of the ROI's were computed and each class

had approx. a 1.9 to 2.0 separability meaning that these classes were different enough that they had a minimal chance of being misclassified as one another.

Unsupervised classification was considered with the following parameters: minimum of 2 classes and a maximum of 5 classes, 2 iterations and a minimum of 10 pixels per class. However it did not perform as well as the supervised classification as only two classes were created and it did not account for the differences between land (urban and non-urban) and water bodies.

Modelling Soil Runoff:

Run off modelling was determined using a simple methodology, and was derived from two factors; the soil content and the land use. This methodology was developed by the United States Department of Agriculture (USDA) Natural Resources Conservation Services or previously known as the Soil Conservation Service. As such, the name of resulting factor is the Curve Number Grid, previously known as the SCS curve number, a raster image with values between 30 and 100. To begin, the 2004 imagery was collected and was cut along the study area, which were defined by the following coordinates (**Figure 1**), in order to reduce the amount of wasteful processing that needed to be done. With an original image size of 185km by 185km, classification of this land would require substantial resources. Reducing the study size to simply encompass the Downtown Ottawa area and the Gatineau area substantially reduces the processing time. The study are was defined by the topic of choice; as this topic studies both the land use change and economic building damage of flooding, it was determined that a study area encompassing only the cities would be necessary, as both land use change as well as buildings were expected to be minor within rural areas.

Lat: 45° 31’ 14” N , Lon: 75° 34’ 31” W
Lat: 45° 18’ 45” N , Lon: 75° 34’ 31” W
Lat: 45° 18’ 45” N , Lon: 75° 50’ 11” W
Lat: 45° 31’ 13” N , Lon: 75° 50’ 11” W

Figure 1: Coordinates of Study area

Pan-sharpening of the 2017 Landsat 7 data was considered, however was eventually decided against due to the lack of panchromatic bands in the Landsat 5 data. In an effort to keep consistency between the two different Landsat satellite imageries, a consistent 30m by 30m resolution was reserved. Once completed, all three imageries were visually inspected. As can be seen by the screenshots below, there are minor differences in the imagery below. As such, Landsat calibration was first employed in ENVI to create a top of atmosphere radiance of the imagery. For radiometric calibration, dark pixel subtraction was then employed. Placing the imagery in a false colour composite, dark water pixels were then chosen to as a calibration point. Using this calibration point, atmospheric attenuation is then subtracted from all the imageries. Once completed, these imageries were able to be compared with minimal difference caused by atmospheric attenuation between the two imageries.



Figure 2: Clockwise starting at top left: Study area from 2004, 2009 and 2017

Once the LandSAT imageries were calibrated, one of the first steps was to create a land use raster map for each imagery. To do so, all three imageries were displayed in a false colour composite. To create the training data, the Regions of Interest (ROI) tool in ENVI, portions of different land use were hand digitized. Basic land use categories were used as it was difficult to

determine the differences between specific land types. As such, the land types created were water, residential urban land, industrial urban land, forested land, grassy land and agriculture. A minimum of 40 pixels for each land type were used for classification. See **Figure 3** to see the classification training data. In order to determine the applicability of the training data, a ROI separability report was computed to determine the spectrally distinctness of each feature class, where a value of 2 is maximum. This ROI separability report can be seen in **Figure 4**.

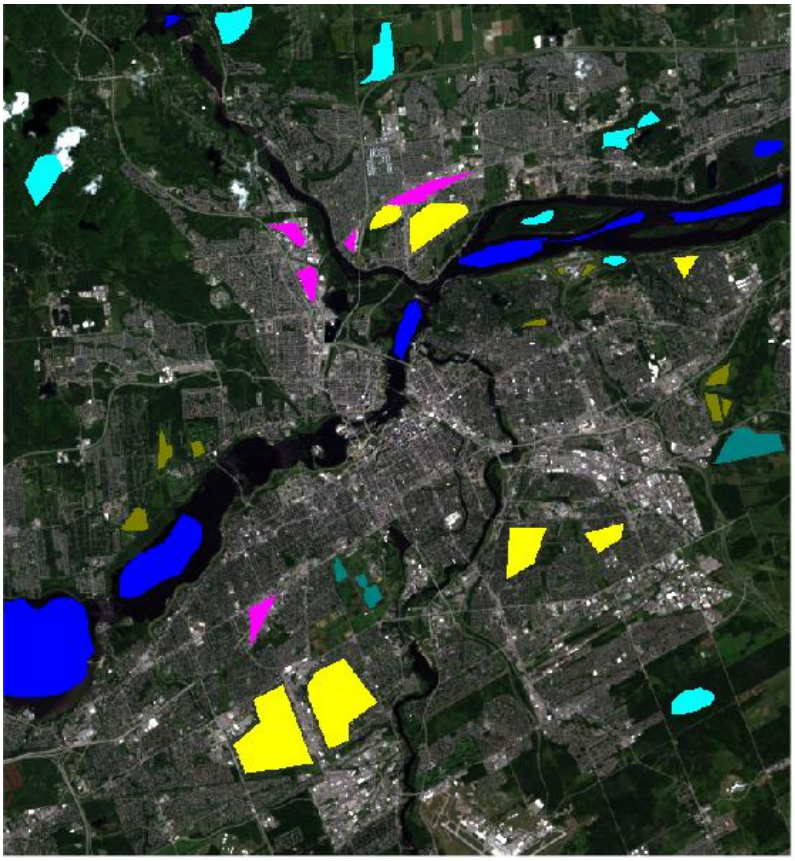


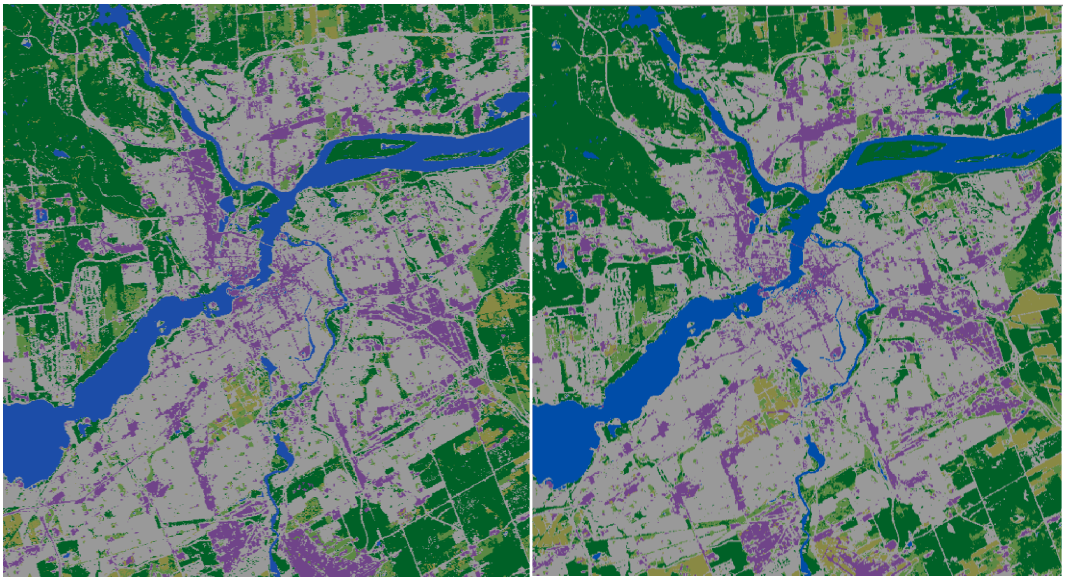
Figure 3: Training data for classification. Yellow = residential, Pink = Commercial, Green = grasslands, Turquoise = Agriculture, Light Blue = Forest, Dark Blue = Water

2004	2009	2017
Agriculture and Grass - 1.45716931 Industrial and Residential - 1.50013486 Grass and Forest - 1.55446928 Agriculture and Forest - 1.72118125 Forest and Residential - 1.82590515	Industrial and Residential - 1.70808649 Grass and Forest - 1.72208830 Grass and Residential - 1.88372922 Agriculture and Industrial - 1.90671797 Forest and Residential - 1.91931390	Agriculture and Grass - 1.60536040 Industrial and Residential - 1.69206544 Agriculture and Residential - 1.76916068 Grass and Forest - 1.80639809 Forest and Residential - 1.84392784

Grass and Residential - 1.85456447 Agriculture and Residential - 1.87333381 Agriculture and Industrial - 1.94560937 Grass and Industrial - 1.94958776 Industrial and Forest - 1.98557944 Industrial and Water - 1.99995637 Forest and Water - 1.99999991 Grass and Water - 1.99999999 Residential and Water - 2.00000000 Agriculture and Water - 2.00000000	Agriculture and Residential - 1.97390465 Agriculture and Grass - 1.99603376 Grass and Industrial - 1.99919117 Industrial and Forest - 1.99947513 Agriculture and Forest - 1.99986305 Industrial and Water - 1.99994930 Forest and Water - 1.99999907 Residential and Water - 2.00000000 Agriculture and Water - 2.00000000 Grass and Water - 2.00000000	Grass and Residential - 1.89617598 Agriculture and Forest - 1.91499434 Agriculture and Industrial - 1.97496586 Grass and Industrial - 1.99657976 Industrial and Forest - 1.99658961 Industrial and Water - 1.99999474 Forest and Water - 1.99999977 Grass and Water - 1.99999997 Residential and Water - 2.00000000 Agriculture and Water - 2.00000000

Figure 4: ROI separability report for each imagery. ROI are kept constant between images

Once the regions of interests were created, the classification of each imagery set could begin. Out of the two major supervised classification methods: maximum likelihood classification (MLC) and support vector machine classification (SVM), previous experiences with the two classifications led to the conclusion that support vector machine classification had a higher degree of accuracy when it comes to supervised classification for land use. Previous literature and studies have also come to this conclusion, showing an overall increase in accuracy of 13.34% when using SVM over MLC (Szuster et al., 2011). As such, the MLC method was forgone in favour of SVM methods. The results of the SVM classification can be seen below, in **Figure 5**.



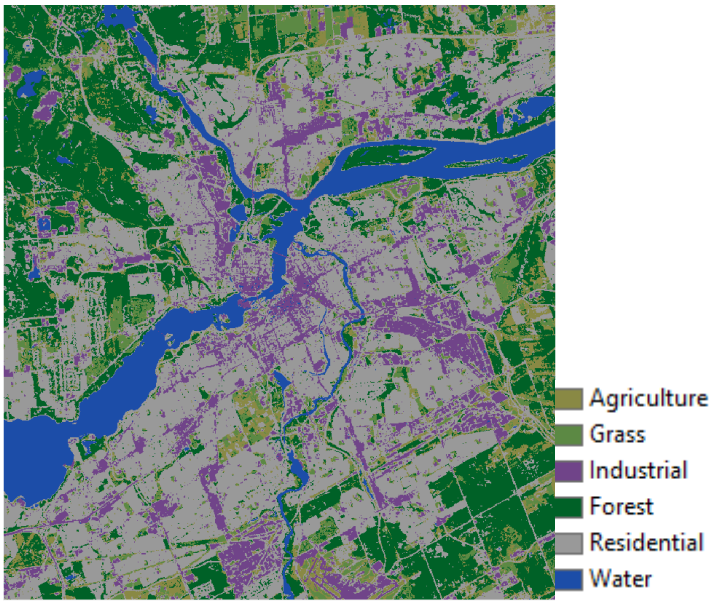


Figure 5: Clockwise starting from top left: 2004 SVM land use classification, 2009 SVM land use classification, land use legend, 2017 SVM land use classification

Once the imagery was classified, a preliminary inspection of the data showed an abundance of salt and peppering within the classification. Because these classified rasters would have to be eventually converted into polygons, the salt and peppering effect would have caused the creation of over 12,000 extraneous polygons, with a majority of these polygons ending up being misclassifications. To combat this, a majority filtering process was used, using a kernel size of 3x3. While a 5x5 kernel size was attempted, it caused the disappearance of too many smaller features and as such, the filtering was redone with a kernel size of 3x3 pixels.

Once the land use classifications were complete and it was possible to inspect the changes in land use over the study period, the next steps were to create the Curve Number raster grid. As the creation of the CN grid is based off two variables, soil and land use, this was split into two sections, beginning with the land use section. To do so, the land use raster was first reclassified into four land use groups; Water, Medium Residential, Forest and Agriculture, all with a land use value of 1, 2, 3 and 4 respectively. As such, the reclassify tool was utilized in ArcGIS to reclassify each original land use into the land use value. Residential and industrial

were reclassified as a land use value of 2, while grass and agriculture were reclassified into a land use value of 4. The reasoning for this methodology is to align with classifications defined by the USGS main land cover types, who are responsible with creating the methodology for the Curve number grid. This was also done to reduce the amount of possible error arising from land use misclassification. Finally, with the land use raster reclassification was completed, the next steps were to convert the raster to a polygon based model. This was done for the possibility to combine the land use classification data with the upcoming soil data. The ArcGIS 'Raster to Polygon' tool was employed for each of the land use rasters to output a vectorized version of the land use classification.

With the land use section of the CN grid creation completed, the next steps were to create the soil data. In order to create the CN grid, specific soil survey data was required. Notably, the hydrologic soil group was required for creation of the soil layer of the CN grid. The hydrologic soil groups, created by the USDA Natural Resource Conservation group, are four distinct groups of soils that are grouped by their ability for water infiltration and infiltration rates (USDA, nd). Based on the runoff potential of the soil group when thoroughly wet, the hydrologic soil groups are defined as such; Soil group A has a high infiltration rate, soil group B has a moderate infiltration rate, soil group C has a slow infiltration rate, and soil group D has a very slow infiltration rates (Mockus & Werner, 1994). To begin, soil data was collected and imported from two sources; the Ontario Soil survey from the Government of Ontario, and Canadian Geological data from the Natural Resources of Canada. As mentioned earlier in the data section of the report Soil survey for the province of Quebec was not freely available. As such, data was required to be interpolated from geological surveys, for the portion of the study areas that laid within the Quebec municipality.

Beginning with the Ontario soil data, after being clipped with the bounds of the study area, the attributes were rearranged to display the proportion of soil groups within the polygon. Shown below in **Figure 6**, it is immediately clear that there is many null values are areas in which data is absent. In some cases, this gap in data is filled in by simply reclassifying the missing values with the most common hydrologic soil group. However, this is not the most accurate method of interpolating soil groups. In order to minimize the amount of error, interpolation of each empty soil group was done using a nearest neighbor methodology. For each soil parcel which has null values, the soil group of all neighbors were collected and the most prevalent soil group was input into the null value. Furthermore, if the soil groups of neighboring polygons were equally split between two soil groups, both soil groups were used, with a percentage based on the proportion of the surrounding soil groups.

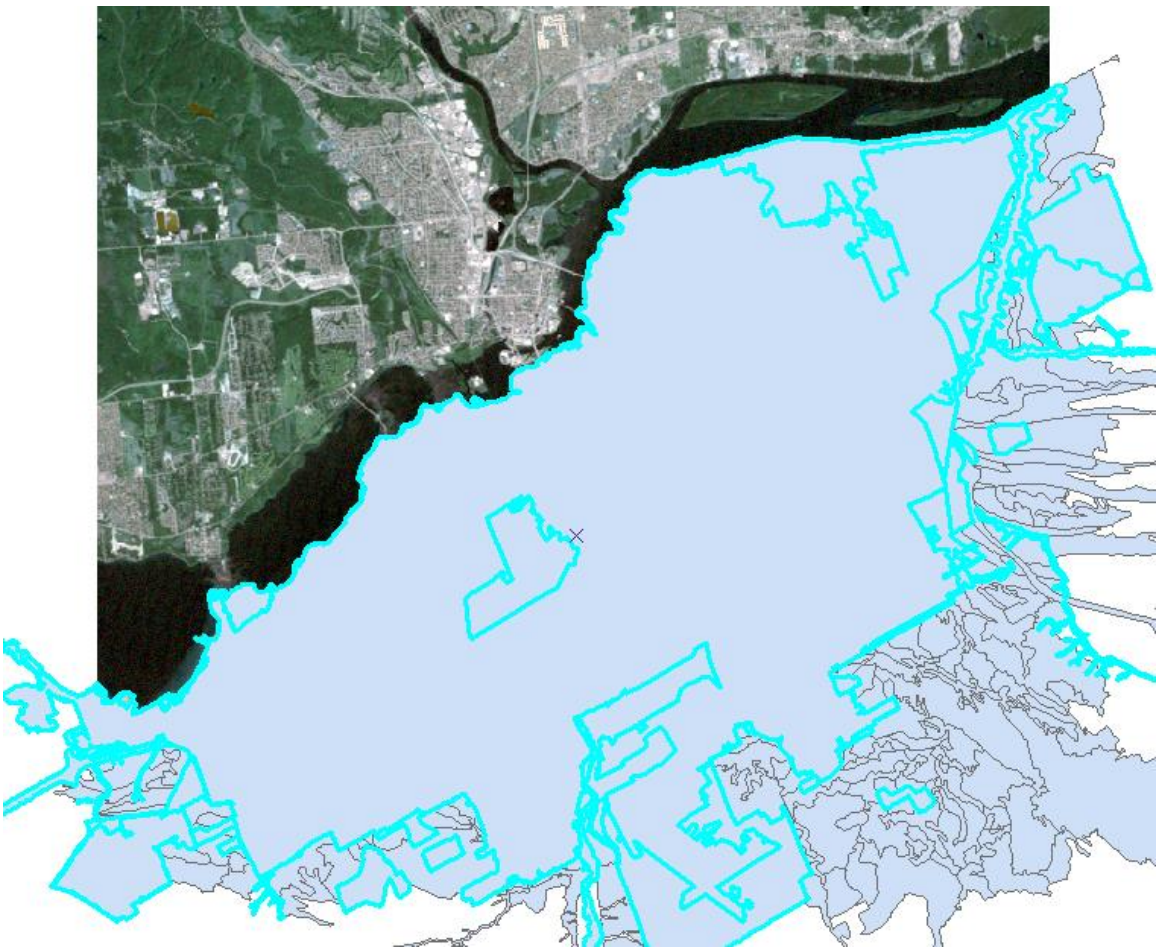


Figure 6: Selected soil polygons with missing values

With the Quebec geology data, the nearest neighbor methods were not applicable as there was no source of data to draw from. As such, geological data was used. Within the surficial geology data, each polygon contained a generic category of the surface materials, texture, thickness and landform (Fulton, 1995). Hydrologic soil groups were estimated using this data. Within the study area, only a small number of geological features were seen; 'fine grained glaciomarine', 'alluvial deposits', 'coarse grained glaciomarine', 'till blanket', and 'till veneer' (Fulton, 1995). Based on the descriptions within the metadata, it was estimated that fine grained glaciomarine would be equivalent to a hydrologic soil group D, due to the fine grained, clay like

features. Alluvial deposits was equated to soil group B, due to the presence of gravel, while coarse grained glaciomarine was equated to soil group C due to the presence of sand and gravel (Fulton, 1995). Till blanket and till veneer were considered soil groups D and C respectively, due to the thickness of the impermeable till materials (Fulton, 1995). Once the soil groups had been created for the Quebec municipalities, the Ottawa soil map and Quebec soil maps were combined to create a soil group map of the study area.

Upon the completion of preparing the soil data, the combination of the soil data and land use data was completed using the Union tool in ArcMap, resulting in polygons with both the land use as well as the soil group composition. Before the CN grid could be created however, a table relating the runoff values to the soil group and land use must be created. A simplistic model is used due to the resolution limitations of the base imagery. As such the relational table can be seen below in table #1. This table was created by the USDA, based off extensive soil sample studies carried out (Cronshey et al., 1986).

Land Use	A	B	C	D
Water	100	100	100	100
Medium residential	57	72	71	86
Forest	30	58	71	78
Agricultural	67	77	83	87

Table 1

The software used to create the CN raster grid is called GEOHMS and is used within ArcGIS. However the CN grid creation is relatively straightforward. The creation of the CN grid is based entirely off the values within the table. For example, a pixel with an agricultural land

use and a total hydrologic soil group composition of group A would result in a pixel CN of 67. This is repeated for each cell, resulting in a raster grid of CN values.

Computing Damages:

For damage calculation, only the 2004 and 2017 data were used. The 2017 and 2004 landsat satellite imagery calibrated as seen in **Figure 2** was used. This raster was then exported as tiff file. To fully calculate damage, the first step was to extract all buildings from the study area. The extraction method that was utilized was provided by Pushparaj and Hegde, but was slightly altered. One of the methodology that was used by Pushparaj and Hegde (2017) for building extraction was: Inputting Image, Removal of Vegetation, Adaptive K-means clustering, Morphological fill operation, Morphological open operation, Extraction of buildings. However, due to limitation of data, the method used in this project for building extraction differs from Pushparaj and Hegde (2017). Pushparaj and Hegde's (2017) data from Quickbird had high resolution of 0.60m x 0.60m to 2m x 2m, whereas the data used in this project is 30m x 30m resolution. Similar to Pushparaj and Hegde (2017) the initial process was to extract vegetation. This was done using Spectral Processing Exploitation and Analysis Resource [**SPEAR**] Vegetation Delineation tool from ENVI. This tool is used to identify the areas of vegetation and determine its quality. However, for this case, it is used for vegetation extraction from the study area. SPEAR Vegetation Delineation is used with auto-correction method of Dark Object Subtraction to remove darkest pixel values from every pixel. Dark Pixel Subtraction corrects haze and glare in the data. Once processed through dark pixel subtraction, it automatically uses NDVI as default to calculate vegetation. The entire range is from -1 to 1, where -1 represents no vegetation and 1 represents total vegetation with good health. From this, four default range was created:

No Veg – values from -1 to 0.25
Sparse Veg – values from 0.25 to 0.5
Moderate Veg – values from 0.5 to 0.75
Dense Veg – values from 0.75 and above.

At the values -1 to 0.25, all the water, and some buildings/roads were included. At 0.25-0.50; a lot of the vegetation were overlapped with buildings due to low-resolution image and lawns being pixelated with homes. At 0.50 and above, it was mostly vegetation, therefore, not included.

The values from 0.25 to 0.50 were masked as 1; all values above and below 0.25 were masked as 0. Areas represented as 1 were extracted as Sparse Vegetation Polygons. The values from -1 to 0.25 were also masked as 1; and all other values above it masked as 0. Areas represented as 1 were extracted as Non-Vegetation Polygons.

Once these vegetation layers were extracted, images used for this project contained water areas and clouds, unlike the extraction method in Pushparaj and Hegde (2017) which had no water areas. Therefore, from these non-vegetation and sparse vegetation layers, water areas, clouds and roads were to be erased for final output of buildings.

Supervised Classification was used for extracting water layers and clouds. There were five different classes: Vegetation, Water, Buildings, Roads and Clouds. The output classifications from these, water and cloud classifications, were then exported with the Classification to Vector tool to conduct vector analysis with Sparse and Non-Vegetation Polygons.

The final extraction left were roads. The roads portrayed in Supervised Classification were mixed with building pixels, therefore, SPEAR Lines of Communication Road with

Supervised Classification was used to extract values for roads. Supervised classification was used with option Spectral Processing (MF, SAM, and MF/SAM Ratio). SPEAR Lines of Communication Road with Supervised Classification enables digitizing, however, the wizard for processing this was closed after the Matched Filter [MF], Spectral Angle Mapper [SAM], and MF/SAM Ratio images were calculated. The MF is used to detect specific materials based on various matches in the image (Harris, n.d.). SAM algorithm is used to determine the spectral similarity between multiple spectra by computing the angles between spectra (Harris, n.d.). The MF/SAM ratio lowers the false positive that may be present in SAM or MF. A pixel that contains road has high value in MF and low value in SAM. If SAM and MF are both high, then when dividing, it lowers the number. Whereas, if MF is high and SAM is low, then it amplifies the road pixel. These pixel values are then exported as .tif to be used further.

All the exported layers in ENVI were imported into ArcGIS for further calculations. However, the MF/SAM ratio layer was still a raster, therefore, it had to be converted into vector before analysis could be done with other vector layers. To convert MF/SAM ratio layer, it was reclassified with old values of 200-255 (brightness) set to 1, and everything below 200 was set to 0. This outputted a raster layer with various road pixels. Then Conditional Evaluation [CON] was used to extract values that were 1. The output from CON was converted into polygon using the Raster to Polygon tool.

Upon completion of the steps above, the result was five different polygon layers: non-vegetation, sparse vegetation, roads, clouds, water. The Erase tool was used to:

1. Remove the polygons of non-vegetation that overlays with water classified polygon.
2. Remove the polygons of non-vegetation that overlays with cloud data
3. Remove the polygons of non-vegetation that overlays with roads data
4. Remove the polygons of sparse vegetation that overlays with water

5. Remove the polygons of sparse vegetation that overlays with cloud
6. Remove the polygons of sparse vegetation that overlays with road data

This ensures that water, cloud and road are also removed from Non-vegetation and Sparse vegetation polygons. Since vegetation, water, roads and clouds are removed, this leaves majority of the remaining polygons as buildings. The resulting outputs consisted of two final layers: final sparse and final non-vegetation. From these vectors, some misclassified polygons were manually deleted. The Extract by Mask tool was used on the original clipped image to extract raster that overlap with the polygons. This gives us the raster layer with values that only overlap with the polygons final sparse and non-vegetation. The CON tool is used on this raster to set sparse vegetation to 1 and non-vegetation to 2.

Raster calculator is then used on the two outputted rasters with the equation:

Con(IsNull("sparse_val"),0,"sparse_val") + Con(IsNull("noval_val"),0,"noval_val")

This sets all Null values in sparse vegetation and non-vegetation to 0 and adds the pixels together. This gave four values for 2017, however only three for 2004. The four values from 2017 are from classification errors, as same pixel should not have been classified as both sparse and non-vegetation. However, the fourth value only contained ten pixels. Thus, the fourth value was then reclassified as sparse vegetation. Afterwards, a new field was added in the attribute table and the area was calculated with the equation: *[count] * 30 * 30* to get the area. The count represents the number of pixels in sparse vegetation and non-vegetation areas. The total building area coverage for 2017 and 2004 can be seen under results, **Figure 24**.

Lastly, one more layer was required for final calculations. The hydro-station data was gathered from the Government of Canada (2018). There were five stations within the study area for 2017 and three for 2004, and these stations did not cover the entire region. Therefore, bordering points which cover the entire study area were needed to be extrapolated before the

study area could be interpolated for depth. However, before these points could be extrapolated, the depth for current station points had to be calculated from river heights. For 2017: The average river height for month of April was calculated for each of the five stations. Then the average height of three non-flooded April months from three different years were calculated; 2016, 2015, 2014. The overall average height of 2016, 2015, and 2014 was subtracted from the 2017 average height to get the difference in height. This value was considered as depth for April. For the calculation of 2004 depth, similar steps were taken except for three stations as opposed to five. Additionally, instead of using the entire September month, only the September 9th value was used, and this was subtracted from the September 9th average height from these years: 2003, 2005 and 2006. Once the depth values for each station was calculated, four border points were extrapolated from known stations. In order to extrapolate this, the station points were selected based on the closeness to one another as well as to the body of water that they were gathering values for. From the stations values that were already known, a 3-dimensional line of two of the closest stations were calculated. This was done by first creating a vector line from two stations, and then expanding it for the entire region. This was calculated as follows:

Given two points of the stations; S1 and S2, the vector is calculated as:

$$\textit{(Coordinates of Station 1) Vector1} = \langle x,y,z \rangle$$

$$\textit{(Coordinates of Station 2) Vector2} = \langle x, y, z \rangle$$

$$\textit{Vector2} - \textit{Vector1} = \langle a\ b\ c \rangle$$

This gives the equation of line as:

$$x = x1 + t(a)$$

$$y = y1 + t(a)$$

$$z = z1 + t(a)$$

From this, z (depth) was calculated by giving t values that moves the x,y coordinates to the edge of the study area. Using this method, four different depth values were calculated for the four boundary points. Once the outer coordinates were determined, ordinary kriging interpolation was computed on all the points to create a depth field. The entire kriging was completed in five major steps, which was all done using the geostatistical wizard in ArcGIS. Ordinary kriging was selected because when using this method, the error variance is minimized (Lefohn, Knudsen, & Shadwick, 2005). In Step three, the model is assumed to be isotropy because the flooding was mostly caused from precipitation therefore, not based on distance. The fields: lagsize, sill, range, and nugget were all unchanged from their default values, which had acceptable semivariogram. In step four, the neighbourhood maximum and minimum is changed from the default 5,2 to 2,1 respectively. This changes the number of closest stations (maximum and minimum) used for each interpolated region. This also ensures that every extrapolated point is used with depth value obtained from the stations, which help to reduce error. Lastly, cross validation was done on the errors and the kriging interpolation is computed. This layer is then extracted as raster for further calculations. However, when computing the kriging, the interpolation caused the vertical and horizontal boundaries to be slanted. This made the kriging raster unaligned with the original raster image causing additional problems. This and other errors from kriging can be seen in the results.

Once the kriging was completed, the next step was to overlay this raster with the rest of the layers to compute areas that are flooded. The flood extent layer computed was then reprojected using ENVI then reclassified using ArcMap. It was reclassified with 1 for flooded area and 0 for non-flooded area. Then raster calculator was used to multiply the flooded region with building layer. This outputted four values for 2017 and three values for 2004, which

represents all the buildings that are within the flooded region. After, a new field was created in the attribute table called Total Area, and the total area of flooded buildings were calculated with the equation: [count] * 30 * 30, where [count] is the number of pixels representing flooded building region. 30 was selected because the pixels are 30 x 30 meters. The total area covered by buildings within the extent can be seen under **Figure 25** in the results.

The next step was to calculate the damage ratio then apply to these buildings for total building damage. Since the only data that was obtained was flood depth, this limited us to very few models that can be used. Four such models were: ICPR, MURL, HYDROTEC, sd-f, and an equation derived from FLEMOps+r RT2 and BT-FLEMO (Kreibich, Botto, Merz, & Schröter, 2016). For this project, the method selected was Hydrotec, because from Eilenburg research, it showed a similar damage ratio to FLEMO+ model (Thieken, Apel, Kreibich, & Aronica, 2007). FLEMO+ is a newer and better model for calculating flooding ratio at meso and micro scales. However, the FLEMO+ model requires additional inputs other than depth, therefore, unable to be used for this project.

Thus, the Hydrotec equation was used with kriging interpolation depth values, in raster calculator, to get raster with damage ratio values. The equation used was:

$$d = (27\ SQRT(h))/100$$

Where d represents the damage ratio and h represents the depth of water.

This loss ratio was then multiplied by the buildings inside the extent to receive damage ratio of each pixel area:

$$Final_Loss = LossRatio * (buildextent/2)$$

LossRatio represents the raster output from hydrotec equation, buildextent represents the buildings that were previously reclassified within the flood extent. The buildings are divided by

2 with the assumption that sparse vegetation areas contain half buildings and half vegetation. Thus, sparse vegetation received the value of 0.5 and no-vegetation areas received the value of 1. This calculated the loss within the building extent, however it does not give the economic value in overall damages.

Therefore, to calculate the economic damages, Rijus's estimated cost per square foot was used. Riju is a custom house building company, which builds houses in southern ontario. This value is based off what they charge to build homes (the average was used). To calculate the cost of damaged buildings, the loss ratio was calculated with the area of building within the flood extent and the cost per square foot. The cost per square-foot is estimated to be 140$ to 160$ for self-built and 170$ to 200$ for custom homes (Rijus, n.d.). With the assumption that the damage will be fixed by the homeowner or hired for repairs, the value 160 per square-foot was selected as reconstruction cost. The overall damages was calculated with the equation:

$$Total\ damage = Final_Loss * 322.917 * 160$$

Where 322.917 represents the area of each pixel in square foot, and 160 represents the dollar cost per square foot.

This final output was reclassified as Natural Breaks (Jenks) with 9 classes, inputting highest value from the given ranges for reclassification. From the reclassified output, the count (number of pixels) was multiplied with the value (cost per square meter) to get total cost.

However, this was not the total cost of building damage for the study area. As previously mentioned, the kriging being slanted caused various pixels to be calculated as zeroes. This can be seen under **Figure 32** where the buildings inside extent have more pixels than the final result. Therefore, this error was corrected by using reverse percentage calculations. The total pixels of buildings damaged after the ratio calculation was 15,919 for 2004 and 9,976 for 2017, instead of,

37,484 for 2004 and 25,771 for 2017 (see **Figure 32** and **Figure 25**). This means that the total flooded buildings represented were only: 38.7% for 2017 and 42% for 2004. Therefore, (1.5833 * 9976) + 9976 = 25771 and (1.3547 * 15919) + 15919 = 37484. Using this method the total cost was calculated with the equations:

*2017 Total Loss = (94390505$ * 1.5833 + 94390505$)*

*2004 Total Loss = (142476447$ * 1.3547 + 142476447$)*

However, this equation assumes that all the pixels that were overlapped with loss ratio and building extent were divided between sparse vegetation and non-vegetation buildings in equal ratio. The Total Loss calculated can be seen under results. Additionally, the flow diagram for the method used for damage calculation can be seen in the figure below (**Figure 43**).

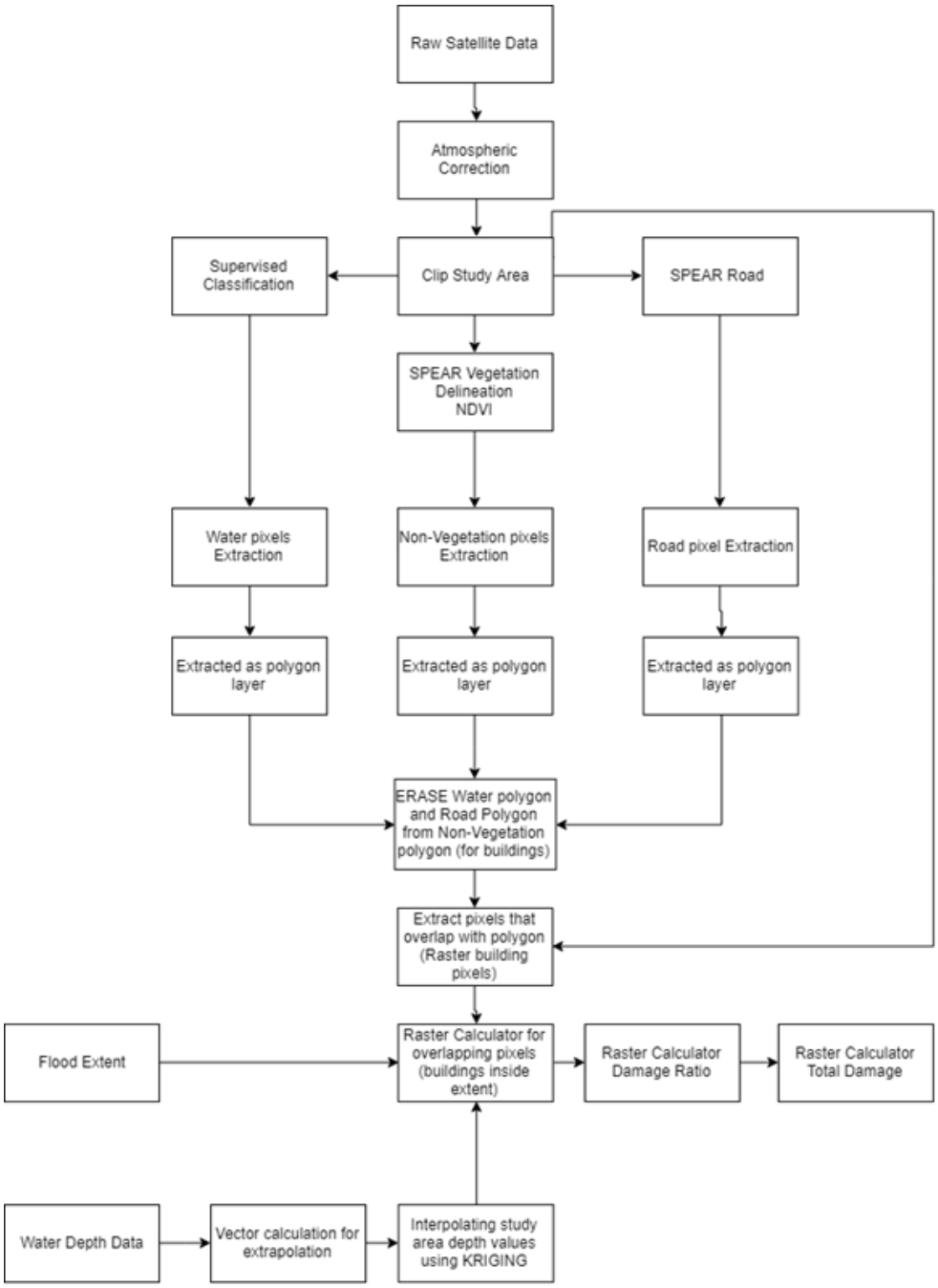


Figure 43 - Damage calculation methodology

Results

Flood Extent Delineation:

For the flooding event that took place on September 10th, 2004, four scenes were collected by the sensor, minutes apart from one another and encompassed all of the Ottawa River region. Using SNAP, the intensity image for the HH band for each of the scenes was radiometrically calibrated which resulted in four Sigma_HH images that provided us with the backscatter coefficients for that polarization mode. The output images were then multi-looked using four looks. The multi-looked images were geometrically corrected using the range-doppler terrain correction function that implemented a predefined STRM 3 second digital elevation model with a NAD83-UTM18N map projection. The resulting images were input into ENVI and were clipped to our study extent boundary. It is important to note that one of the scenes were not within the extent of our study area boundary and therefore were discarded. These images were now ready to be filtered and classified in ENVI. The first set of ROIs were used as training data for the supervised classification and included pixels for regular water bodies (in blue), non-flooded vegetation (in white) and flooded areas in close proximity to water bodies (in green). The image was classified using supervised classification- support vector machine classifier.

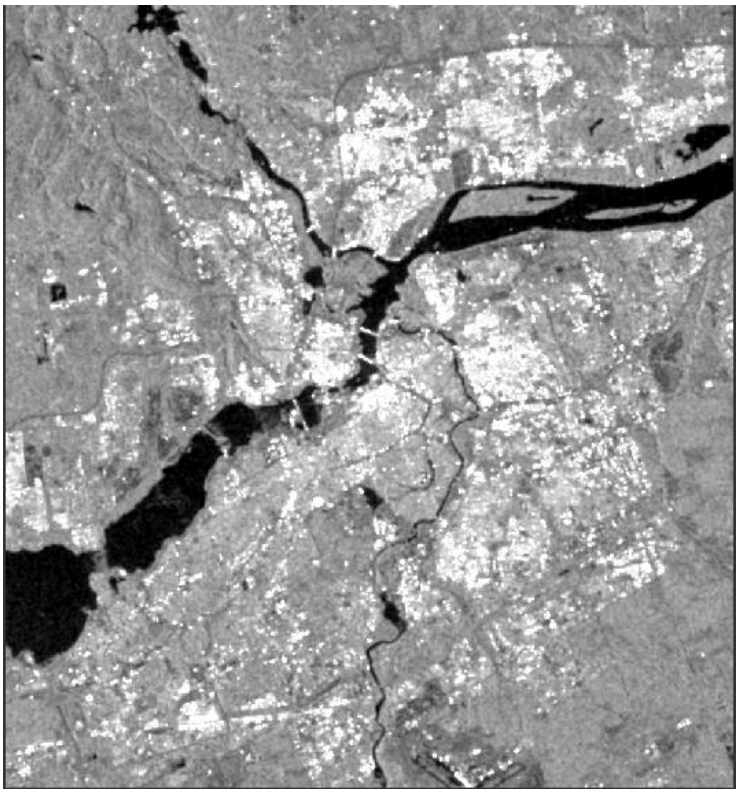


Figure 7..*: September 10th, 2004 intensity HH image with a 5x5 gamma filter with 2 looks*

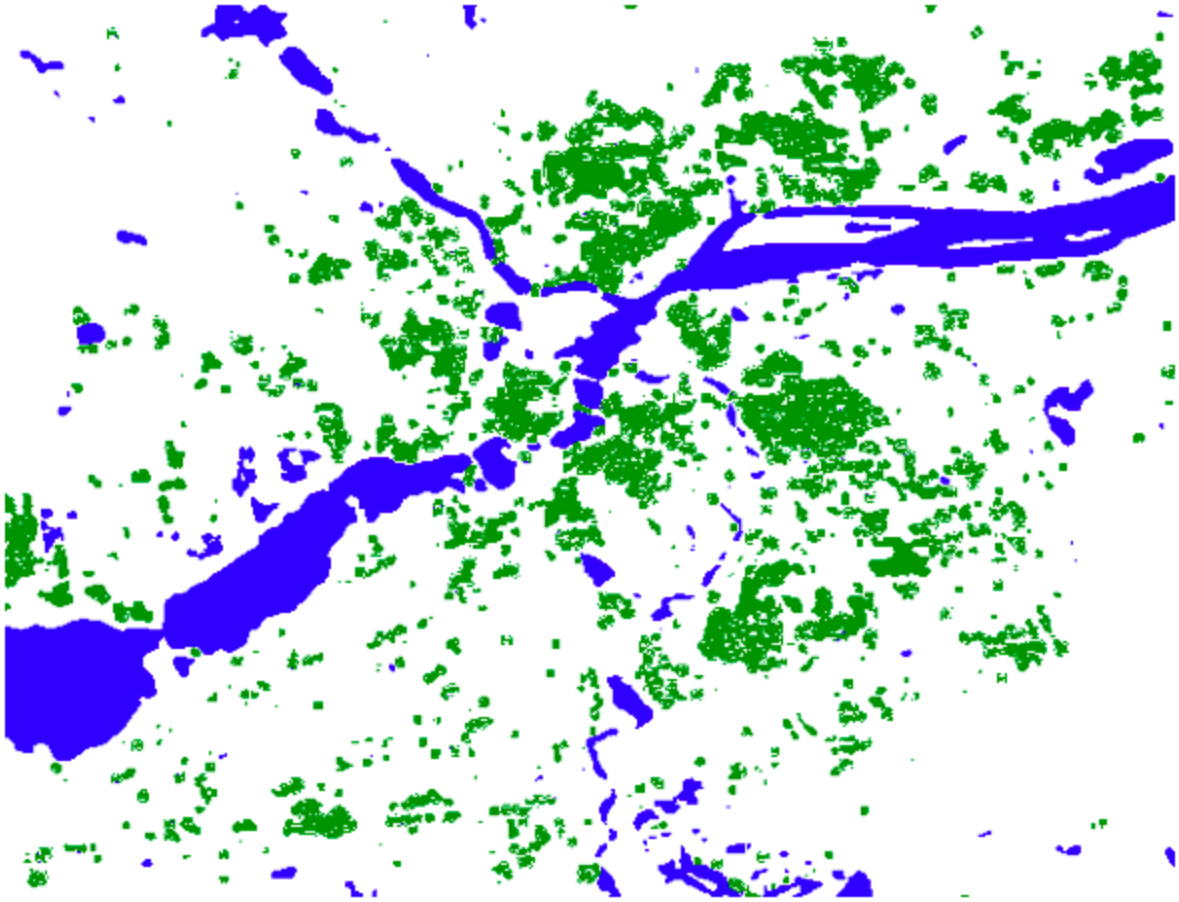


Figure 8..*: September 10th, 2004 final flood extent image created using the support vector machine classifier. Essential parameters: HH band, median filter with a kernel size of 3x3. Flooded areas shown in green; un-flooded areas in white; water bodies in blue*

For the flooding event that took place on July 4th, 2006, five scenes were collected by the sensor, were minutes apart from one another and encompassed a segment of the Ottawa River region. Using SNAP, the intensity images for both the HH & HV bands for each of the scenes were radiometrically calibrated which resulted in five Sigma_HH & Sigma_HV images that provided us with the backscatter coefficients for those polarization modes. The output images were then multi-looked using four looks. The multi-looked images were geometrically corrected using the range-doppler terrain correction function that implemented a pre-defined STRM 3 second digital elevation model with a NAD83 (CSRS98) / MTM zone 8 map projection. The resulting images were input into ENVI and were clipped to our study extent boundary. It is

important to note that not only were two of the scenes not within the extent of our study area boundary but the data also did not fill up that whole area which lead to some issues in terms of misclassification for some pixels in the image. These images were now ready to be filtered and classified in ENVI. The first set of ROIs were used as training data for the supervised classification and included pixels for regular water bodies (in blue), non-flooded vegetation (in white) and flooded areas in close proximity to water bodies (in green). The image was classified using supervised classification- maximum likelihood classifier (which requires two bands to run) and the parallelepiped classifier. In this case, the support vector machine did not perform as well as it did with the 2004 images.

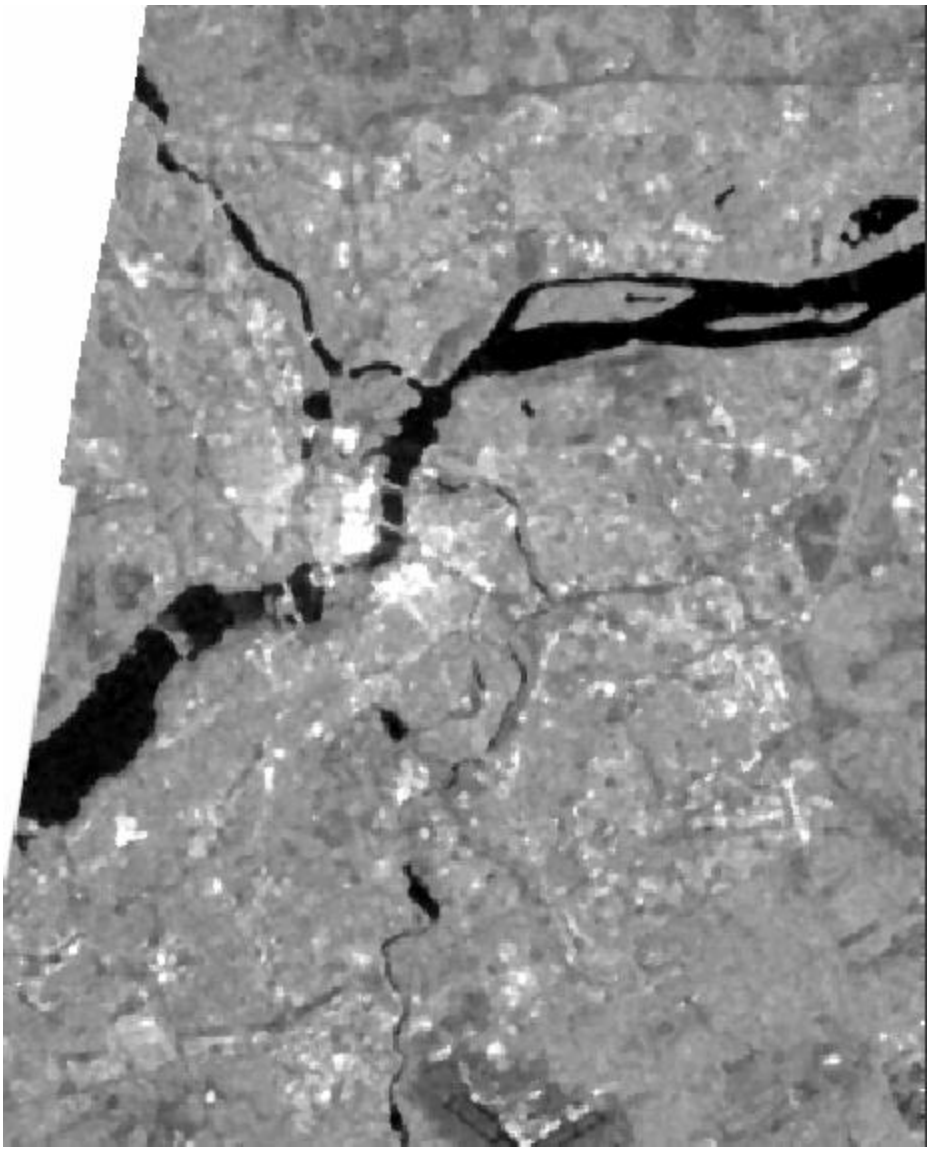


Figure 9..*: July 4th, 2006 intensity HH image with a median 3 by 3 convolution filter*

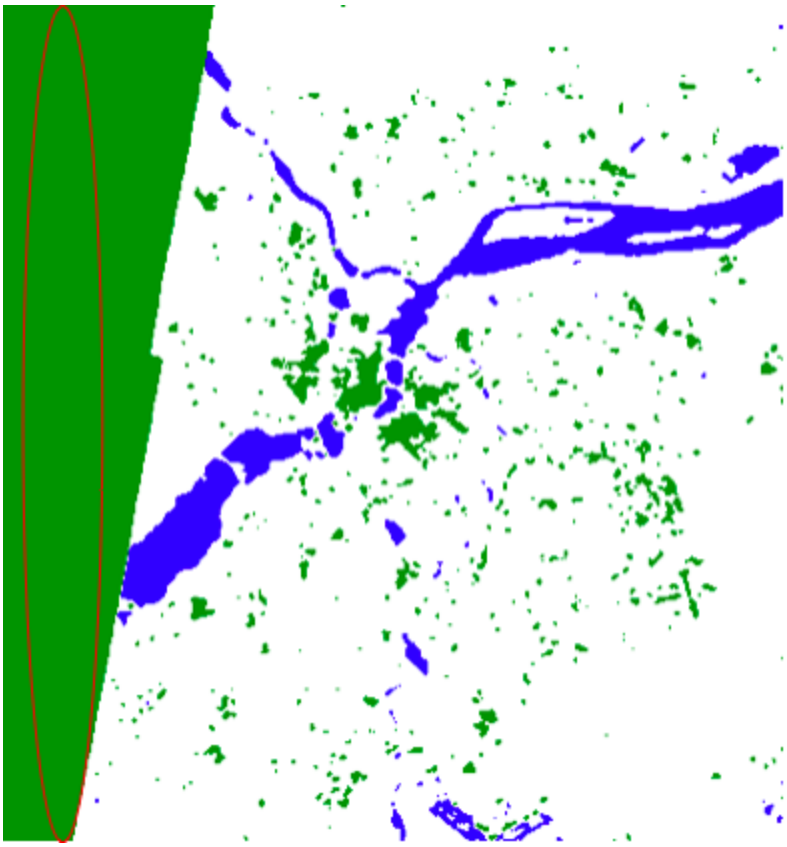


Figure 10..*: July 4th, 2006 final flood extent image created using the maximum likelihood classifier. Essential parameters: HH band. Flooded vegetation shown in green; un-flooded areas in white; water bodies in blue. Note: The area surrounded by the oval outlined in red is the misclassified area.*

For the flooding event that took place on April 5th, 2008, five scenes were collected by the sensor, were minutes apart from one another and encompassed a segment of the Ottawa River region. Using SNAP, the intensity images for both the HH & HV bands for each of the scenes were radiometrically calibrated which resulted in five Sigma_HH & Sigma_HV images that provided us with the backscatter coefficients for those polarization modes. The output images were then multi-looked using four looks. The multi-looked images were geometrically corrected using the range-doppler terrain correction function that implemented a pre-defined STRM 3 second digital elevation model with a NAD83 (CSRS98) / MTM zone 8 map projection. The resulting images were input into ENVI and were clipped to our study extent boundary. It is

important to note that not only were two of the scenes not within the extent of our study area boundary but the data also did not fill up that whole area which lead to some issues in terms of misclassification for some pixels in the image. These images were now ready to be filtered and classified in ENVI. The first set of ROIs were used as training data for the supervised classification and included pixels for regular water bodies (in blue), non-flooded vegetation (in white) and flooded areas in close proximity to water bodies (in green). The image was classified using supervised classification- maximum likelihood classifier (which requires two bands to run) and the parallelepiped classifier.

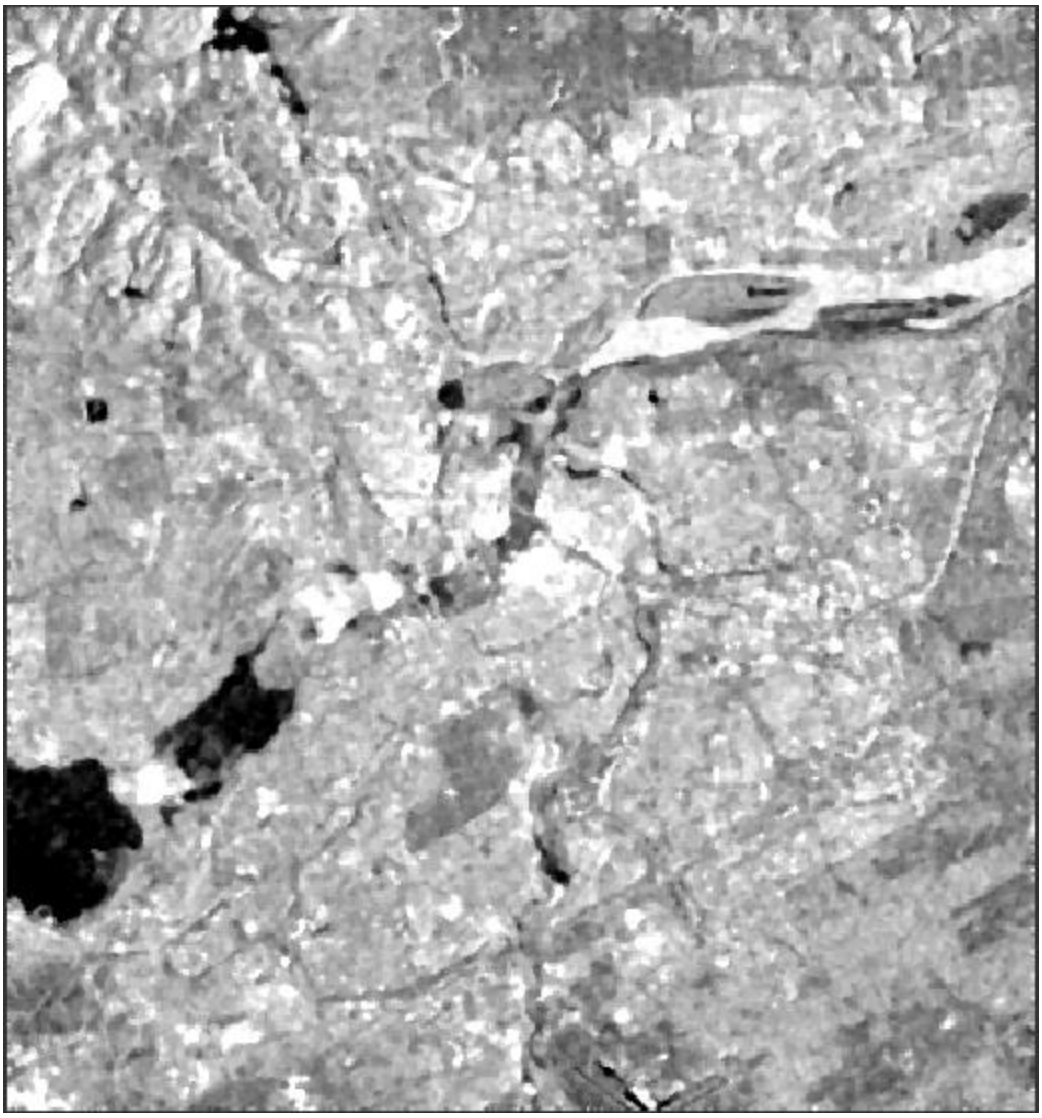


Figure 11..*: April 5th, 2008 intensity HH image with a median 3 by 3 convolution filter*

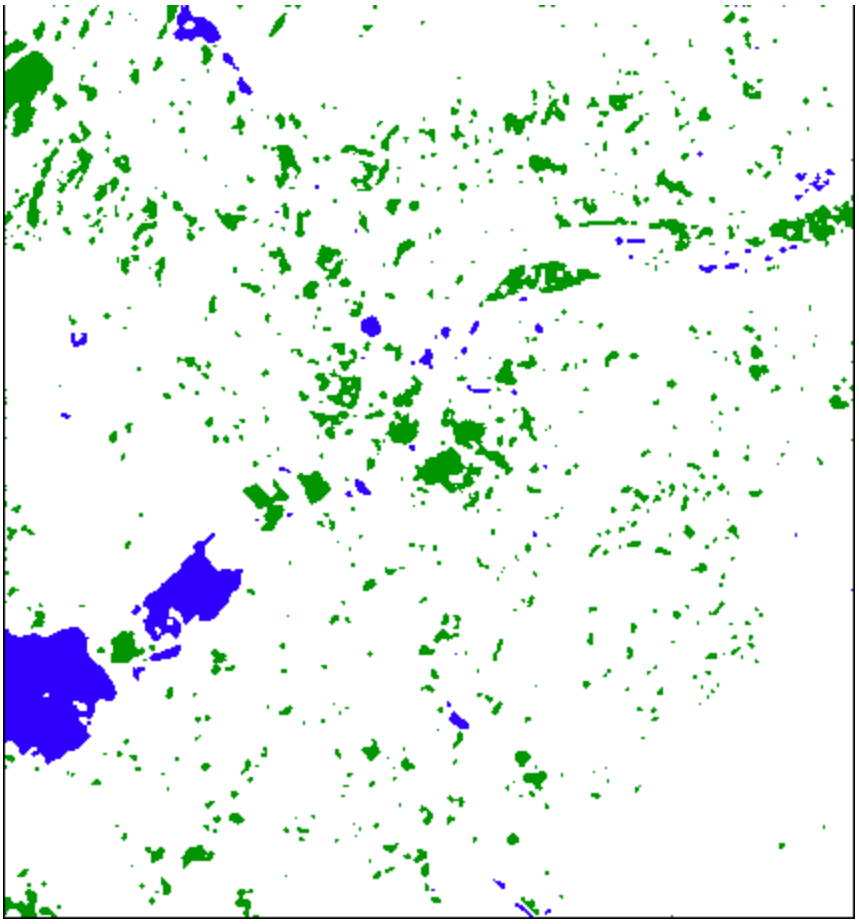


Figure 12..*: April 5th, 2008 final flood extent image created using the maximum likelihood classifier. Essential parameters: HH & HV bands, median filter with a 3x3 kernel size. Flooded vegetation shown in green; un-flooded areas in white; water bodies in blue.*

For the flooding event that took place on, July 28th, 2009, six scenes were collected by the sensor, were minutes apart from one another and encompassed a segment of the Ottawa River region. Using SNAP, the intensity images for both the HH & HV bands for each of the scenes were radiometrically calibrated which resulted in five Sigma_HH & Sigma_HV images that provided us with the backscatter coefficients for those polarization modes. The output images were then multi-looked using four looks. The multi-looked images were geometrically corrected using the range-doppler terrain correction function that implemented a pre-defined STRM 3 second digital elevation model with a NAD83 (CSRS98) / MTM zone 8 map projection. The resulting images were input into ENVI and were clipped to our study extent boundary. It is

important to note that not only was one of the scenes not within the extent of our study area boundary but the data also did not fill up that whole area which lead to some issues in terms of misclassification for some pixels in the overall image. These images were now ready to be filtered and classified in ENVI. The first set of ROIs were used as training data for the supervised classification and included pixels for regular water bodies (in blue), non-flooded vegetation (in white) and flooded areas in close proximity to water bodies (in green). The image was classified using supervised classification- maximum likelihood classifier (which requires two bands to run) and the parallelepiped classifier. Out of the two classifiers, the parallelepiped classifier had the most optimal performance visually as it was the classifier that was recommended in the literature (Bonn et al., 2005).



Figure 13..*: July 28th, 2009 final flood extent image created using the maximum likelihood classifier. Essential parameters: HH & HV bands, median filter with a 3x3 kernel size. Flooded vegetation shown in green; un-flooded areas in white; water bodies in blue. Note: The area surrounded by the oval outlined in red is the misclassified area.*

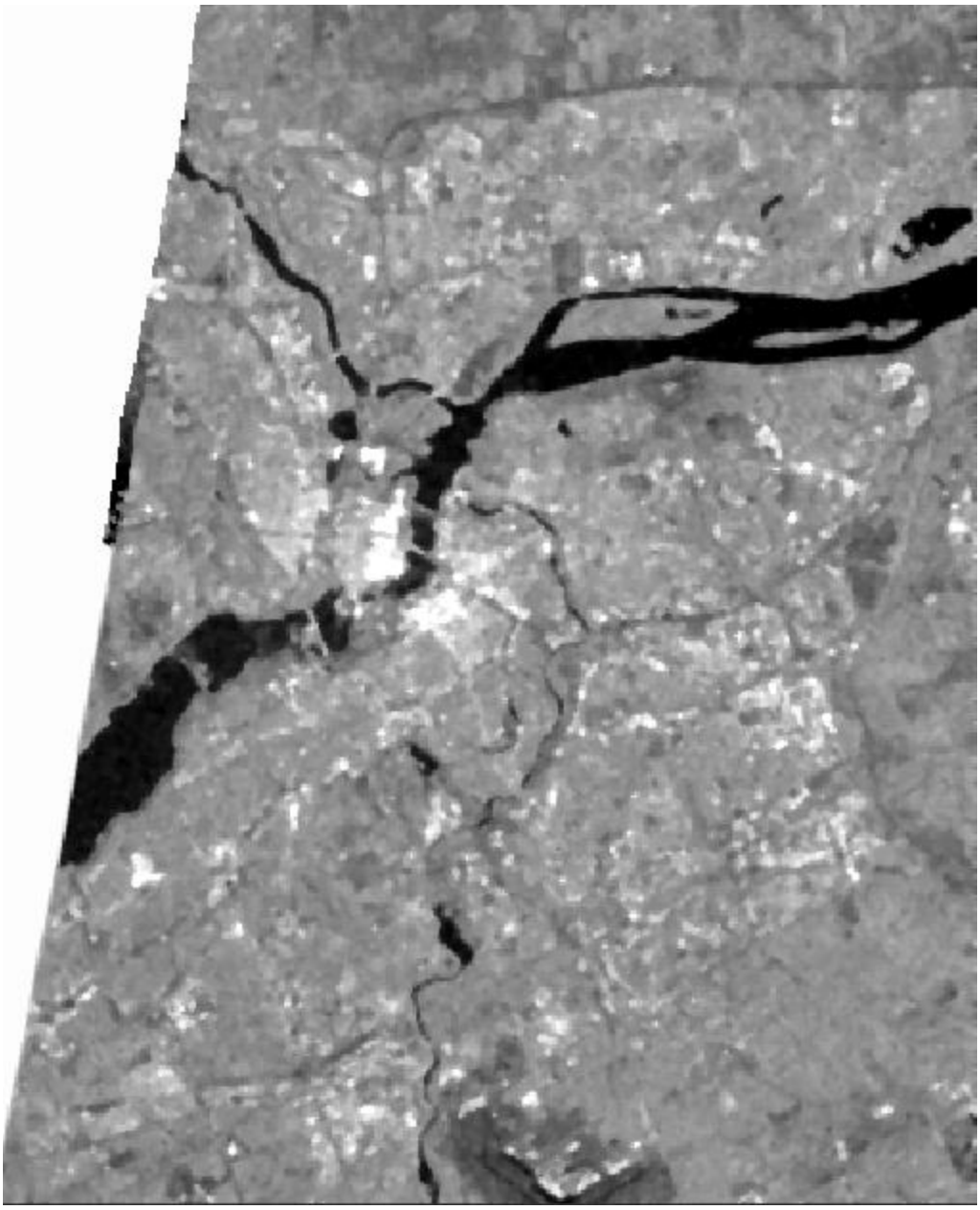


Figure 14..*: July 28th, 2009 intensity HH image with a median 3 by 3 convolution filter*

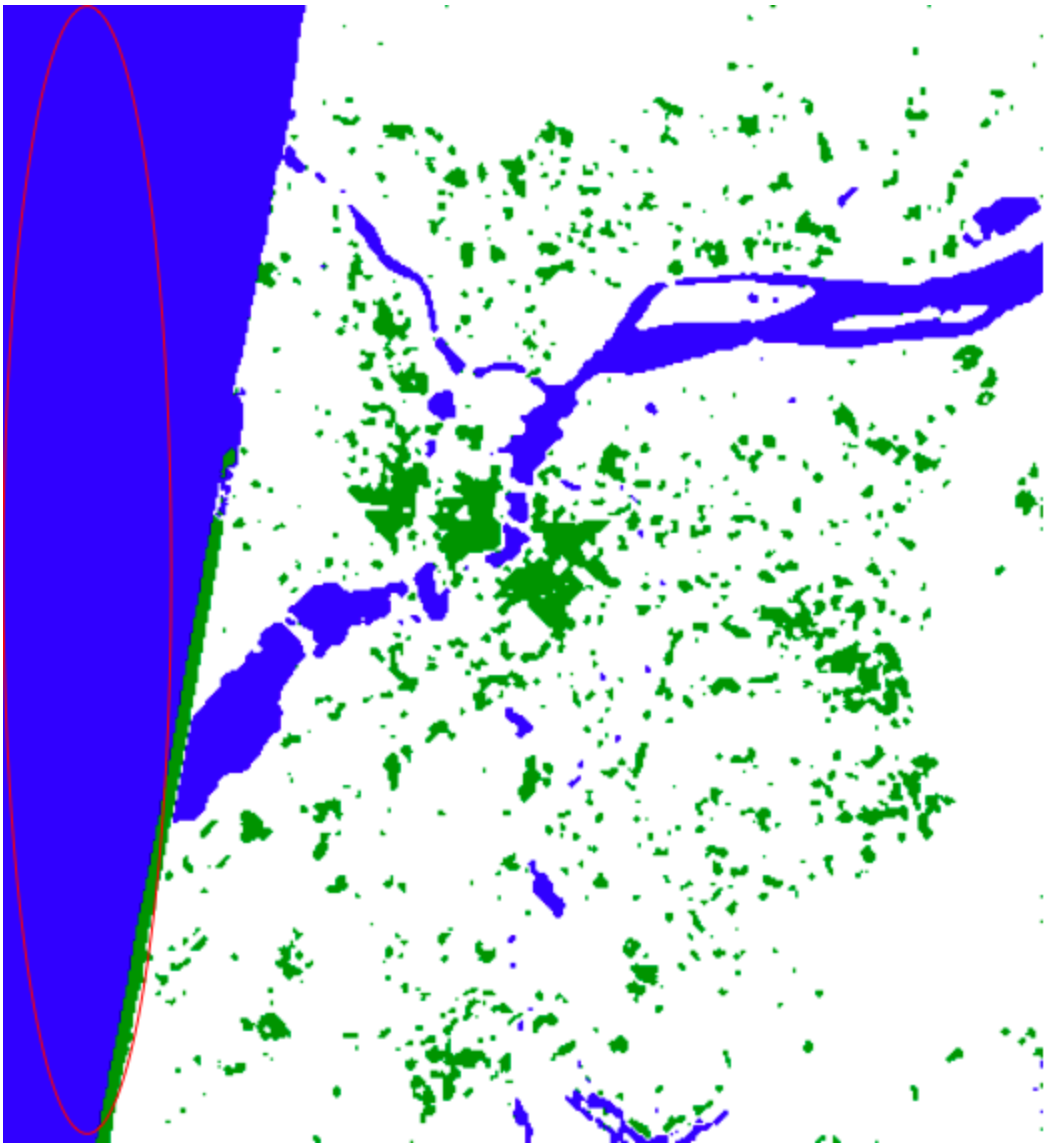


Figure 15..*: July 28th, 2009 final flood extent image created using the support vector machine classifier. Essential parameters: HH & HV bands, median filter with a 3x3 kernel size. Flooded vegetation shown in green; un-flooded areas in white; water bodies in blue. Note: The area surrounded by the oval outlined in red is the misclassified area.*

Unfortunately ENVI-SAT data was unavailable in 2017 as the sensor turned off in 2012 ("Envisat", 2012). The only other available and usable data at cost was Sentinel-1 data which was also provided by the European Space Agency. However, the only data that was available during our specified range for 2017 was collected on May 12th, which was towards the end of the flood period. According to ESA, Sentinel-1 Interferometric Wide Swath Level 1- Ground

Range Detected images are multi-looked and projected to ground range using an Earth ellipsoid model. Since no pre-processing was necessary, the VV and VH polarization images were input into ENVI to be further analyzed. The images were resized to fit the study area boundary and reprojected to NAD83-UTM18N projection. Two different filters were used on the images to see which would help us differentiate. One was the gamma filter with a 5 by 5 window and 2 looks and the other was the median convolution filter with a kernel size of 3 by 3. Both removed speckle and noise effectively but it was still quite difficult to differentiate between the dark and light pixels since we were unable to obtain an HH polarization intensity image. Using knowledge about the backscatter coefficient properties for SAR images and supplementary information about the region, ROIs were created as training data for the supervised classification. The ROIs included pixels for regular water bodies (in blue), urban areas (in yellow), non-flooded vegetation (in white) and flooded vegetation in close proximity to water bodies (in green). Two different supervised classifications were implemented, the parallelepiped classification and the support vector machine classifier. However, the support vector machine classifier out performed the parallelepiped classification for the VH band that was gamma filtered.



Figure 16..*: May 12th, 2017 intensity VV image with a median 3 by 3 convolution filter*



Figure 17..*: May 12th, 2017 intensity VH image with a median 3 by 3 convolution filter*



Figure 18..*: May 12th, 2017 intensity VH image with a gamma 5 by 5 filter with 2 looks*

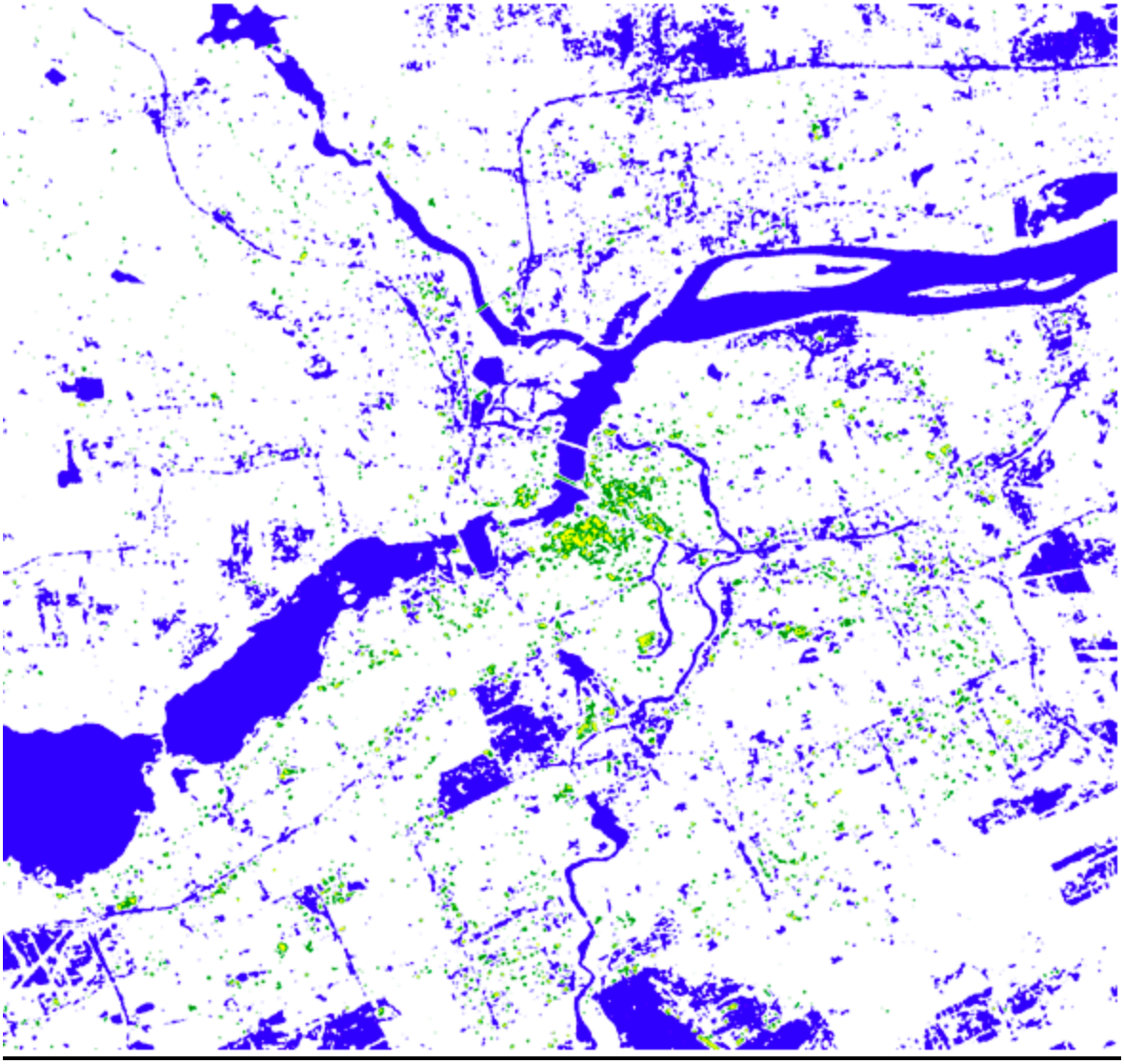


Figure 19..*: May 12th, 2017 final flood extent image created using the support vector machine classifier. Essential parameters: VH band, gamma filter with a 5x5 window & 2 looks. Flooded vegetation shown in green; flooded urban areas shown in yellow, un-flooded areas in white; water bodies in blue.*

Damage outcomes

The SVM separability report for 2017 and 2004 can be seen below. These values are pair separation from least to most. In both of these, roads and buildings have the least separability. This is followed by grass and road for 2017, and bare and road for 2004. All the different classes

have low separability except for water and road, and water and bare, which have complete separability.

For 2017:

```
Buildings and Road – 1.87347137
Grass and Road – 1.98116318
Grass and Buildings – 1.99701730
Bare and Buildings – 1.99843775
Bare and Road – 1.99964356
Bare and Grass – 1.99999785
Water and Buildings – 1.99999960
Grass and Water – 1.99999977
Water and Road – 2.00000000
Bare and Water – 2.00000000
```

For 2004:

```
Buildings and Road – 1.37969941
Bare and Road – 1.93163443
Grass and Road – 1.93515749
Bare and Grass – 1.96217994
Grass and Buildings – 1.96630467
Bare and Buildings – 1.99504792
Water and Buildings – 1.99928343
Grass and Water – 1.99994330
Water and Road – 2.00000000
Bare and Water – 2.00000000
```

Figure 20: Separability Report for 2004 and 2017 ROI's

For 2017:



For 2004:

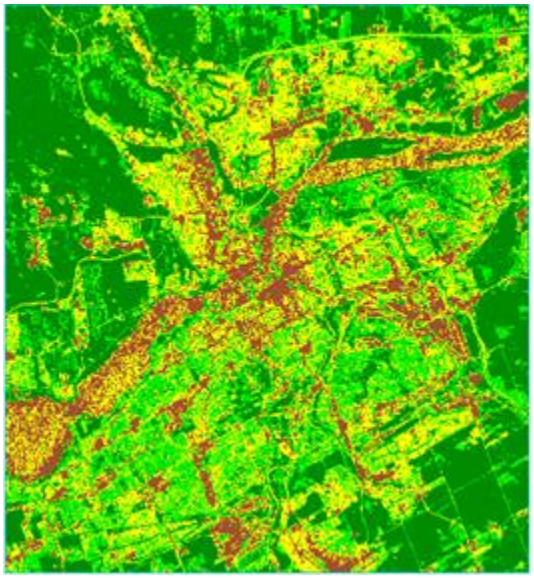


Figure 21: The Sparse and No-Vegetation areas where yellow is sparse and brown is no vegetation

For 2017 For 2004

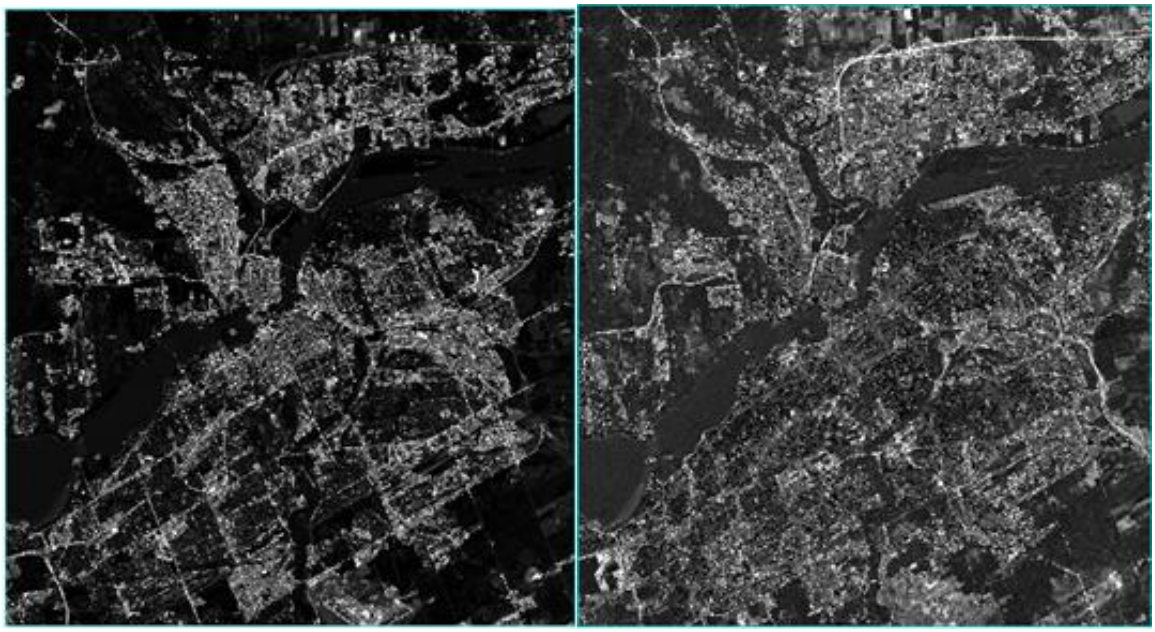


Figure 22: Roads MF/SAM ratio, where the brighter the pixel, the higher chance of it being road.

For 2017: For 2004:

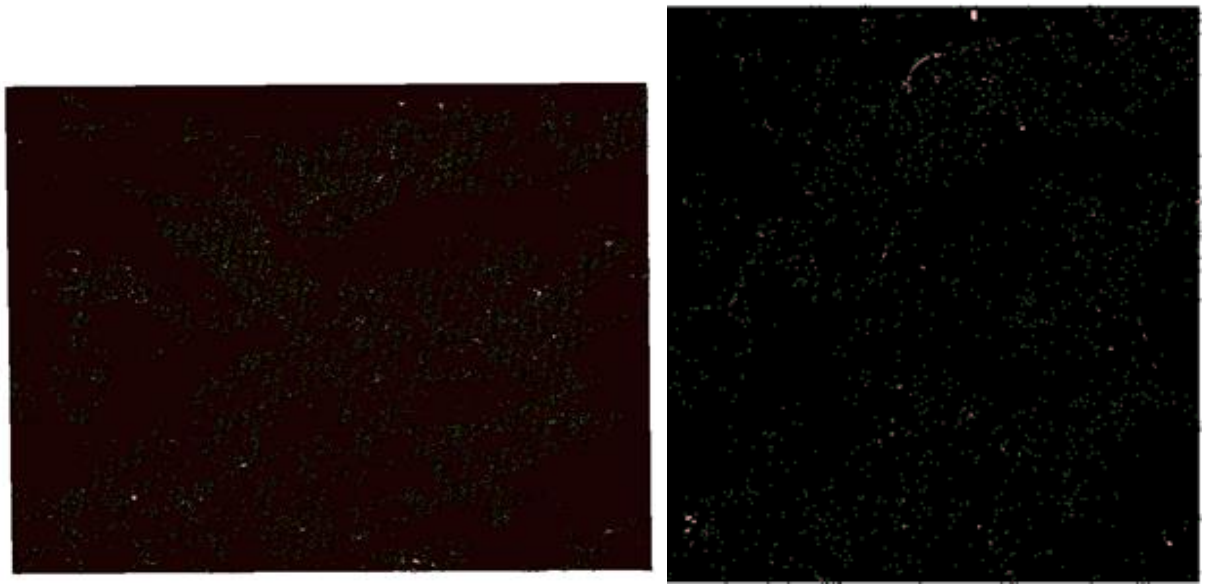


Figure 23: Error from converting raster to polygons

The above pictures shows the error when converting raster to polygons for roads. The green values represents road values on the raster, whereas, the pink represents the polygons that were extracted using raster to polygon too.

Total buildings can be seen below. Green areas represents buildings classified under sparse vegetation, while red areas represents buildings classified under non-vegetation. It was found that there was much more area covered by buildings in 2017, than there was in 2004. The total area covered by buildings in 2017 is: 170,037,000m² and the total area covered by buildings in 2004 is: 132,423,300m².

For 2017: For 2004:

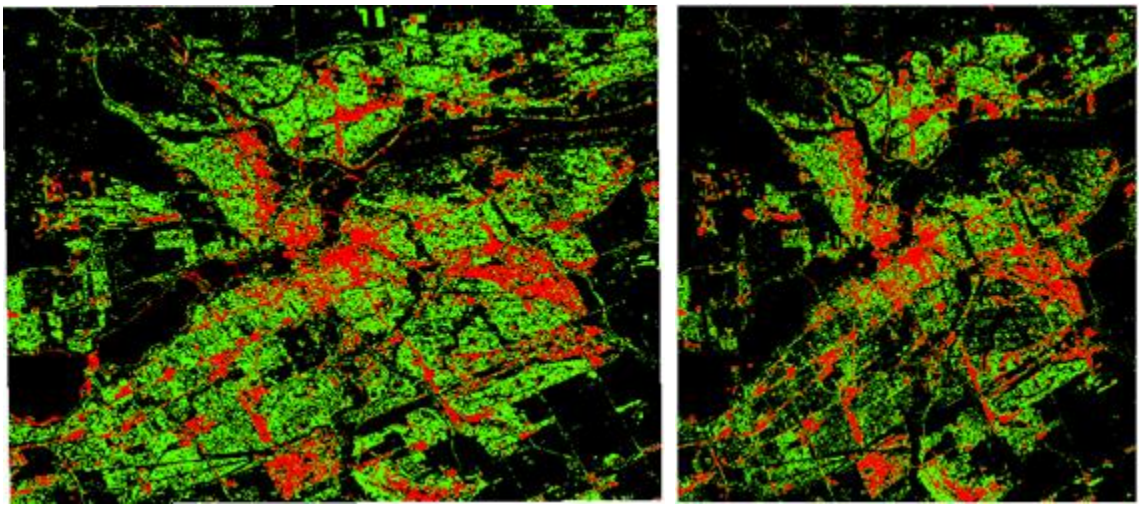


Figure 24: All building coverage within study area

Buildings contained inside the flood extent can be seen below. Green areas represent buildings classified under sparse vegetation, red areas represent buildings classified under non-vegetation. It was found that the area of buildings flooded in 2004 exceed the area of buildings flooded in 2017. The total area of buildings flooded in 2017 is: 23,202,000m², whereas the total area flooded in 2004 is: 33,735,600m². Additionally, the total pixel count for 2017 was 25771 and the total pixel count for 2004 was 37484.

For 2017: For 2004:

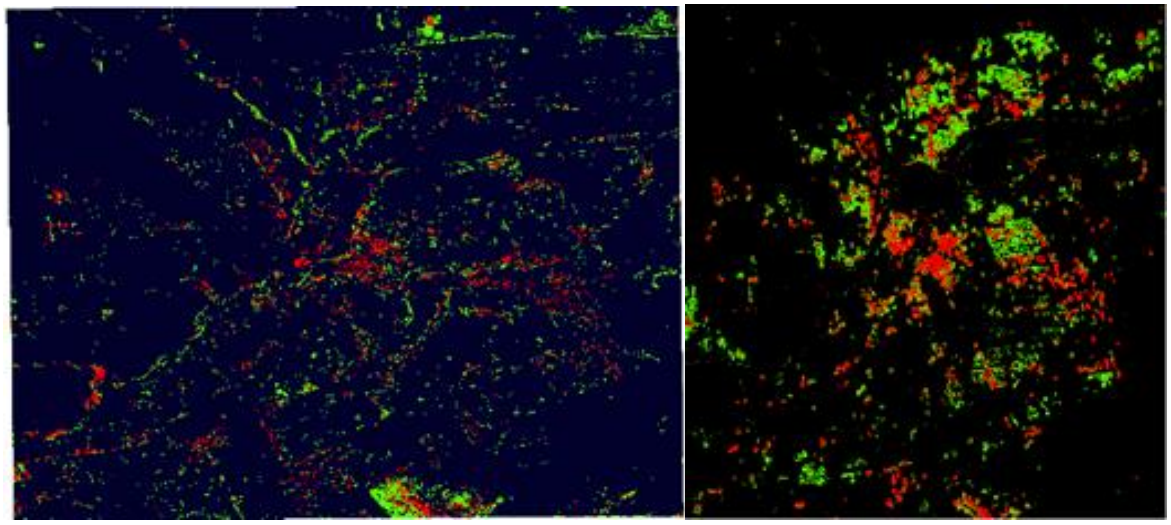


Figure 25: The buildings contained within the flood extent

The stations used for kriging interpolation can be seen below. The green circles with dot in the middle represent the stations with values from Environment Canada, whereas, the circles without the dots are extrapolated points.

For 2017:



Latitude	Longitude	Depth	Shape *	OUTLINE COORDS
45.477107	-75.573126	1.045	Point Z	OUTLINE COORDS RIGHT
45.52016	-75.777477	0.4561	Point Z	OUTLINE COORDS TOP
45.370599	-75.834405	0.3765	Point Z	OUTLINE COORDS LEFT
45.32013	-75.696035	0.4561	Point Z	OUTLINE COORDS BOTTOM
45.343333	-75.808056	0.062	Point Z	02KF015
45.351111	-75.826389	0.38	Point Z	02KF005
45.381389	-75.699722	0.235	Point Z	02LA004
45.439722	-75.692778	0.034	Point Z	02LA027
45.431389	-75.706111	0.73	Point Z	02LA015

For 2004:



	Latitude	Longitude	Depth	Shape *	OUTLINE COORDS
▶	45.477107	-75.573126	0.4092	Point Z	OUTLINE COORDS RIGHT
	45.52016	-75.777477	0.4288	Point Z	OUTLINE COORDS TOP
	45.370599	-75.834405	0.22202	Point Z	OUTLINE COORDS LEFT
	45.32013	-75.696035	0.4288	Point Z	OUTLINE COORDS BOTTOM
	45.351111	-75.826389	0.223	Point Z	02KF005
	45.381389	-75.699722	0.593	Point Z	02LA004
	45.431389	-75.706111	0.321	Point Z	02LA015

Figure 26: The original stations and boundary points with their coordinates and extrapolated depth values

The cross validation errors can be seen below. The five errors: Predicted, Error, Standardized Error and Mean squared Standardized error. The values in 2017 has negatives whereas the ones in 2004 are all positive. This result is re-touched in more details under discussion.

For 2017: For 2004:

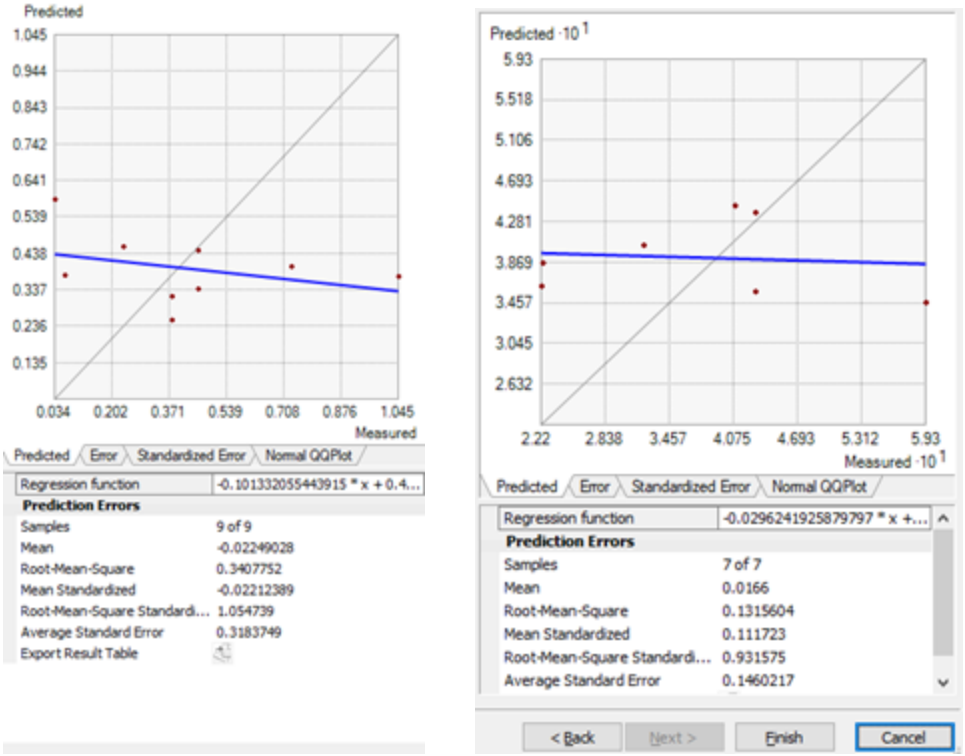


Figure 27: Error cross validation for kriging results

The QQ plot for kriging can be seen below. The QQ plot determines the normality of the data. The 2017 kriging QQ plot shows the points are fairly close to the regression line with small deviation, however, same cannot be said about the 2004 plot. This result is touched in more details under discussion.

For 2017: For 2004:

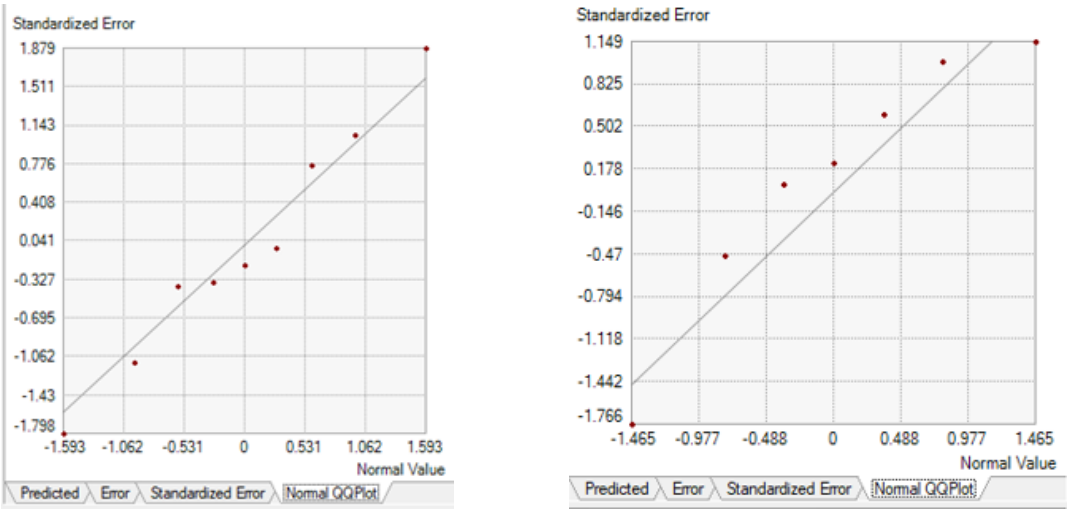


Figure 28: QQ plot for kriging error distribution

The kriging result can be seen below. It displays the different types of parameters selected and results obtained from kriging. As mentioned in the methodology, these reports also display the selected values. It shows the default lag size for both 2004 and 2017, and these vary due to the number of stations used for interpolation purposes.

For 2017

Input datasets	
Dataset	F:\School\geog 471\ARCGIS\coords2.txt_Features
Type	Feature Class
Data field 1	Depth
Records	9
Method	**Kriging**
Type	Ordinary
Output type	Prediction
Dataset #	1
Trend type	None
Searching neighborhood	Standard
Neighbors to include	2
Include at least	1
Sector type	Four and 45 degree
Major semiaxis	0.752137049958
Minor semiaxis	0.752137049958
Angle	0
Variogram	Semivariogram
Number of lags	12
Lag size	0.0626780886747
Nugget	0.0722066703574
Measurement error %	100
Model type	Stable
Parameter	2
Range	0.752137049958
Anisotropy	No
Partial sill	0.282823782671

For 2004:

Input datasets	
Dataset	F:\School\geog 471\ROUGH WORK 2\Calcs\finalStation.txt_Features
Type	Feature Class
Data field 1	Depth
Records	7
Method	**Kriging**
Type	Ordinary
Output type	Prediction
Dataset #	1
Trend type	None
Searching neighborhood	Standard
Neighbors to include	2
Include at least	1
Sector type	Four and 45 degree
Major semiaxis	0.283452276902
Minor semiaxis	0.283452276902
Angle	0
Variogram	Semivariogram
Number of lags	12
Lag size	0.038337963217
Nugget	0.006005944544
Measurement error %	100
Model type	Stable
Parameter	0.3599609375
Range	0.283452276902
Anisotropy	No
Partial sill	0.012674816579

Figure 29: Report for selected kriging parameters

The final kriging output overlapped with the study area can be seen below. It shows that 2017 has a much bigger depth values interpolated than 2004. Additionally, the kriging is slanted from the study area, which would have impact on final damage calculation.

For 2017 For 2004:

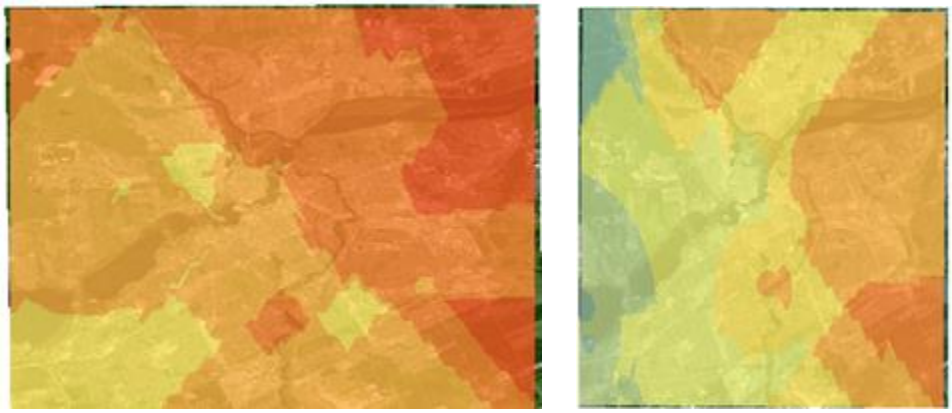


Figure 30: Final kriging output overlaid on top of study area

The loss ratio calculated from Hydrotec model can be seen below: The dark areas are high loss ratio whereas light areas are low loss ratio. The loss ratio for 2017 was approximately within 0.14 and 0.23. Whereas the loss ratio for 2004 was approximately within 0.14 to 0.19. The loss ratio represents the percentage of damage caused to the building. At damage of 1.0, it is total loss; 100% building submersion, complete repair.

For 2017 For 2004

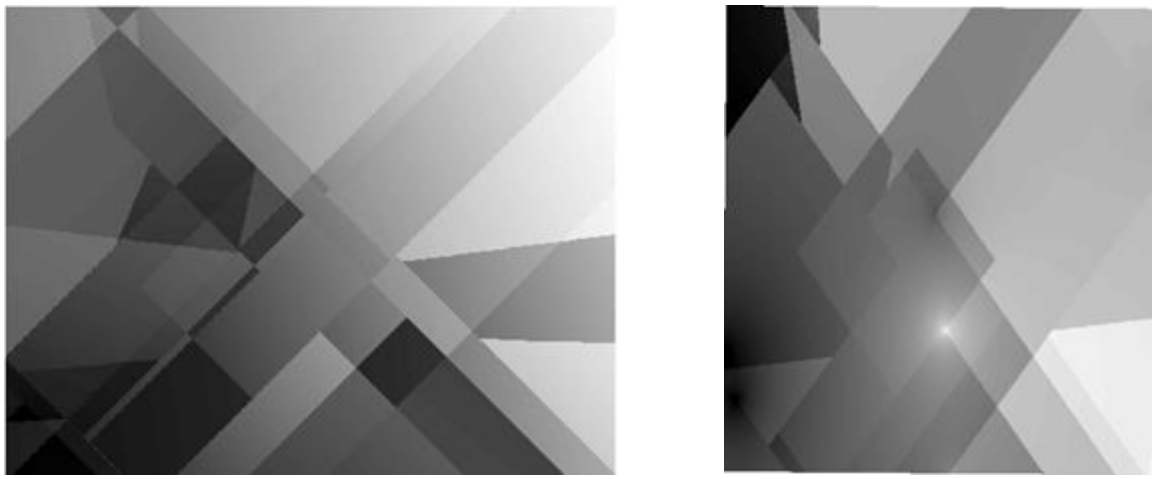


Figure 31: Loss ratio over a surface based on Hydrotec model

The final loss can be seen below. Raster calculator was used to calculate the rasters. The total area coverage by buildings in 2004 was 14,327,100m^2 whereas, the total area coverage by buildings in 2017 was 8,978,400m^2. Therefore, the building coverage pixels for 2004 was 15,919 and for 2017 was 9,976. This area and pixel coverage is substantially less than what was calculated for building area and pixels within flood extent.

For 2017: For 2004:

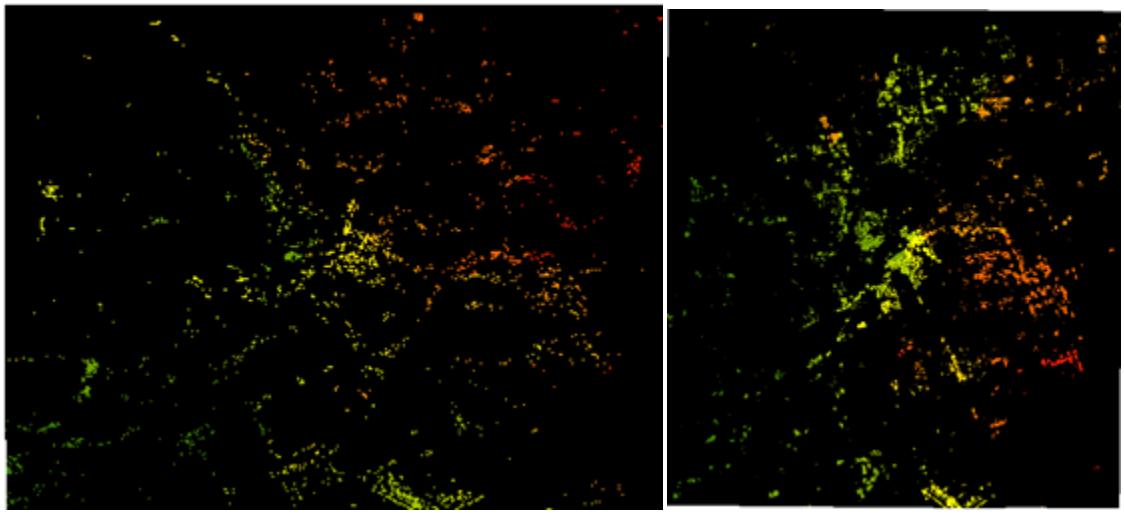


Figure 32: Damage cost per area, green to red representing low to high cost

For 2017: For 2004:



Figure 33: An overlay of building loss raster on top of the study area.

Years	Buildings (Total Area Coverage)	Buildings (Flooded)	Damage (Total)
2017	170,037,000m²	23,202,000m²	$243,838,991
2004	132,423,300m²	33,735,600m²	$335,489,290

Table 2 – Total losses

Land Use Classification Results

When it comes to land use, all three raster images show very similar results, with minimal changes in terms of most classification groups. It also should be noted that the results are based on the majority filtered raster results. Looking at the pixel values, we can see that that the number of pixels classified as water change from 40,877 pixels to 42,402 pixels from 2004 to

2009, a minor increase of roughly 4%, while the results between 2009 to 2017 has a negligible change in the number of classified pixels, from 42,402 to 42,625 pixels being classified as water. This is to be expected as large amounts of alluvial change would be uncharacteristic of a study period of 13 years.

The other land uses however, saw a much larger difference. When examining the forested land use change, the pixel count from 2004 to 2009 show a somewhat modest decrease in forested land, from 145,061 to 137,590, a decrease in forest land of roughly 5.2%. In terms of area of lost forested land, this equates to a drop from 4.35 km^2 in 2004 to 4.12 km^2. From 2009 to 2017, however, this trend seems to be increasing slightly, with a decrease of 8.1%, from 137,590 to 126,494 pixels respectively, or a decrease of total forested land to 3.79 km^2. This equates to a total loss of 13%, or roughly 0.55 km^2 of lost forested land.

Other land uses, however, show a much more surprising level of change. Residential land use in 2004 started off with 7.54 km^2 of land, with a pixel count of 251,321. In 2009, however, this figure decreased by only 0.7%, down to 249,769 pixels, or a land size of 7.49 km^2 of land. Furthermore, the change between 2009 and 2017 is similarly surprisingly small at a total loss of 3%, or a 0.22 km^2 loss, with a total pixel count of 242,415 pixels. With an overall loss of 4.6%, the surprising result of the residential land use is the second most stagnant land use change.

When examining the change in industrial land use, however, the results are much more dramatic. A surprising drop of 9.5% from 2004 to 2009 can be seen, with a pixel count of 45,828, or 1.37 km^2, dropping to 41,482 pixels, or a land size of 1.24 km^2. This is equal to a total decrease of 0.13 km^2 between 2004 and 2009, a counterintuitive result in terms of urbanization. However, the total number of industrially classified pixels in 2017 increase with a jump of

23.4% from 2009, reaching a total count of 51,171 pixels, or an equivalent of 1.54 km^2. This leads to a net increase from 2004 to 2017 of only 11.6%, or a total land use change of 0.16km^2.

Grasslands also had a surprising change in land use between the three years we were studying. Originally, the pixel count values in 2004 had a count of 22,719 pixels, or 0.68 km^2. In 2009 however, a huge increase of 32% can be observed, with a pixel count in 2009 of 29,981 pixels, or 0.90 km^2, equal to a total area increase of over 0.31 km^2. Having a small original area, this large percentage increase may not very strange, however the difference between the years 2009 and 2017 show a very surprising result, with the pixel count of grasslands dropping back down nearly 17%. The pixel count in 2017 was 24,970, or 0.75 km^2. This leads to a net total increase in terms of grasslands of only about 10%. Overall, this increase figure is much more understandable.

Agricultural land use had an equally surprising change over the study period. Starting with a pixel count value of 11,810 in 2004, equating to 0.35 km^2, an increase of 32% or 0.11 km^2 can be seen; in 2009, the land pixel count values of agricultural land use jump to 15,644 pixels, or 0.47 km^2. Similarly to the grasslands increase, a 30% increase for the agricultural land use is not very surprising due to the small original value of only 0.35 km^2, which is small due to the study area. However, from 2009 to 2017, the value of agricultural land use is effectively doubled, with over 29,941 pixels classified as agricultural land use. This equates to a 91% increase over the 2009 values, with a land area growing to 0.90 km^2. With a net increase of 253%, the agricultural land use grows over 0.54 km^2. A summary of the pixel numbers can be seen below.

	2004 Pixel Count	2004 Area (km2)	2009 Pixel Count	2009 Area (km2)	2017 Pixel Count	2017 Area (km2)
Agriculture	11810	0.3543	15644	0.46932	29941	0.89823
Grass	22719	0.68157	29981	0.89943	24970	0.7491
Industrial	45828	1.37484	41482	1.24446	51171	1.53513
Forest	145061	4.35183	137590	4.1277	126494	3.79482
Residential	251321	7.53963	249769	7.49307	242415	7.27245
Water	40877	1.22631	42402	1.27206	42625	1.27875

Table 3: Pixel and Areal Summary of Classification for 2004, 2009 and 2017 imagery.

Runoff Modelling Results:

The results of the CN grid can be seen below in **Figure 34**. As stated earlier, the CN grid is a raster with values ranging between 30 and 100. Values closer to 100 indicate very high run off and low infiltration rates, while values closer to 30 indicate high infiltration rates and low runoff. Visually inspecting the three CN grid rasters for each year, we can see that the study area does generally have high runoff and low infiltration rates. The mean value of runoff for 2004 is 77 while the mean runoff value for 2009 and 2017 are both 78 and 78.3 respectively, showing a miniscule increase in runoff for the study period. All of these values are well over the halfway value between 30 and 100, which is 65.

Firstly, a quick examination of all three CN grids show water bodies having a CN value of 100. This is expected; there is zero expected infiltration for water bodies. In addition, this reflects the table created earlier to assign CN values to particular land uses; parcels classified as water bodies are given a CN value of 100. All three CN grids also show a difference between CN values between Ottawa and Quebec, with the majority of Ottawa parcels having a lighter pixel value in comparison to the Quebec areas. This is most prominent in the downtown areas of both municipalities; the vast majority of downtown areas in Ottawa have a pixel value of 86, while the pixel values located in the downtown Gatineau areas having a pixel value of 71. This is a significant difference in terms of soil permeability and runoff. However, as the inspection continues away from the Ottawa river, both northward and southward, a majority of the pixels from each municipality show similar results, with a pixel value of 78 and 78.5 seen in Ottawa and Gatineau, respectively.Inspecting the histograms of each shows that while the most common pixel value of 71 is fairly regular across all three imageries, the spikes for the other most common CN numbers, 78 and 86, show a significant increase across all three imageries. In fact, all CN numbers greater than 71 are increased slightly from 2004 to 2017. The increase stretches evenly between CN values between 71 and 100, with CN number count increases ranging from 100 to 10,000 per CN number spike.

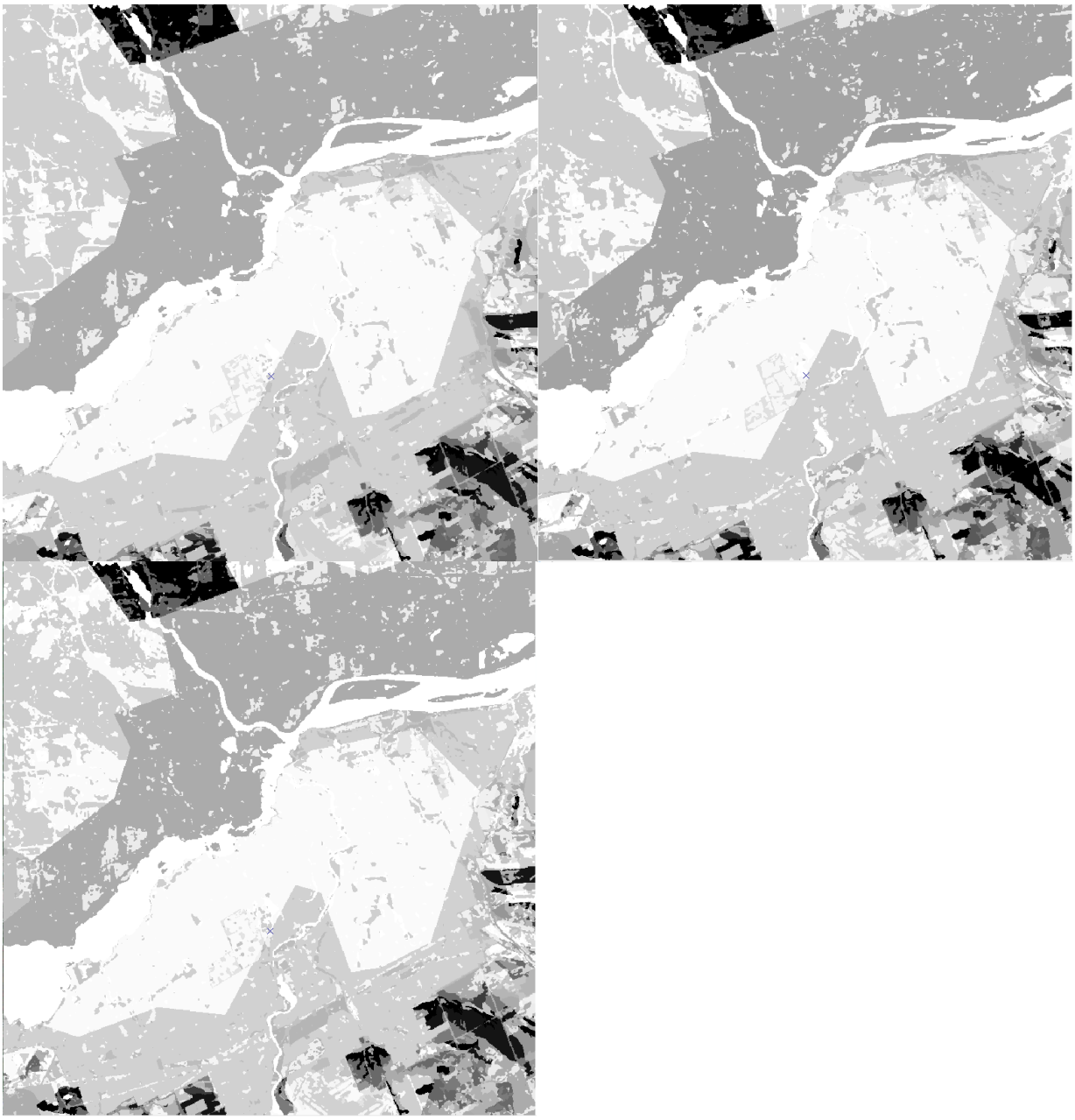


Figure 34: Clockwise from top left: CN Grid result for 2004, 2009 and 2017. Lighter pixels = higher runoff

In order to compare the differences between the CN raster images for each year, simple band algebra was used, with the older imagery's pixel value being divided by the newer imagery's pixel values. The result was a raster in which unchanged CN numbers are given a raster value of 100, while in areas in which CN numbers have increased, and thus the run off, or more specifically, overland flow has increased, have lower values to a minimum value of 30

(**Figure 35**). Similarly, areas in which overland flow has decreased will have higher values to a maximum of 333.33. When examining the change detection imagery, the majority of the landscape sees no change in CN value, however when looking at the change detection from 2004 to 2009, it is apparent that there is a clear majority in terms of increasing overland flow. However, looking at the pixel value data, there are a total of 34,931 pixels which have in increase in overland flow, which equates to 1.05 km^2 of land. Conversely, there are only 20,714 pixels, or 0.62 km^2 of decreased over land flow and increased infiltration rates. This trend of increasing overland flow can also be seen in the 2009 and 2017 comparison imagery. In this comparison, a total of 38,658 pixels have an increased overland flow, equating to 1.15 km^2, while 24,840 pixels, or 0.75 km^2 of land, exhibit an increase in soil infiltration. Lastly, over the period of 2004 to 2017, a total of 46,888 pixels, equating to 1.4 km^2 of land, exhibit an increase in overland flow, while only 19,555 pixels, or 0.58 km^2 of land, show a decrease in overland flow. This net decrease in overland flow is surprising due to the increase in soil infiltration seen in the two previous change analyses.

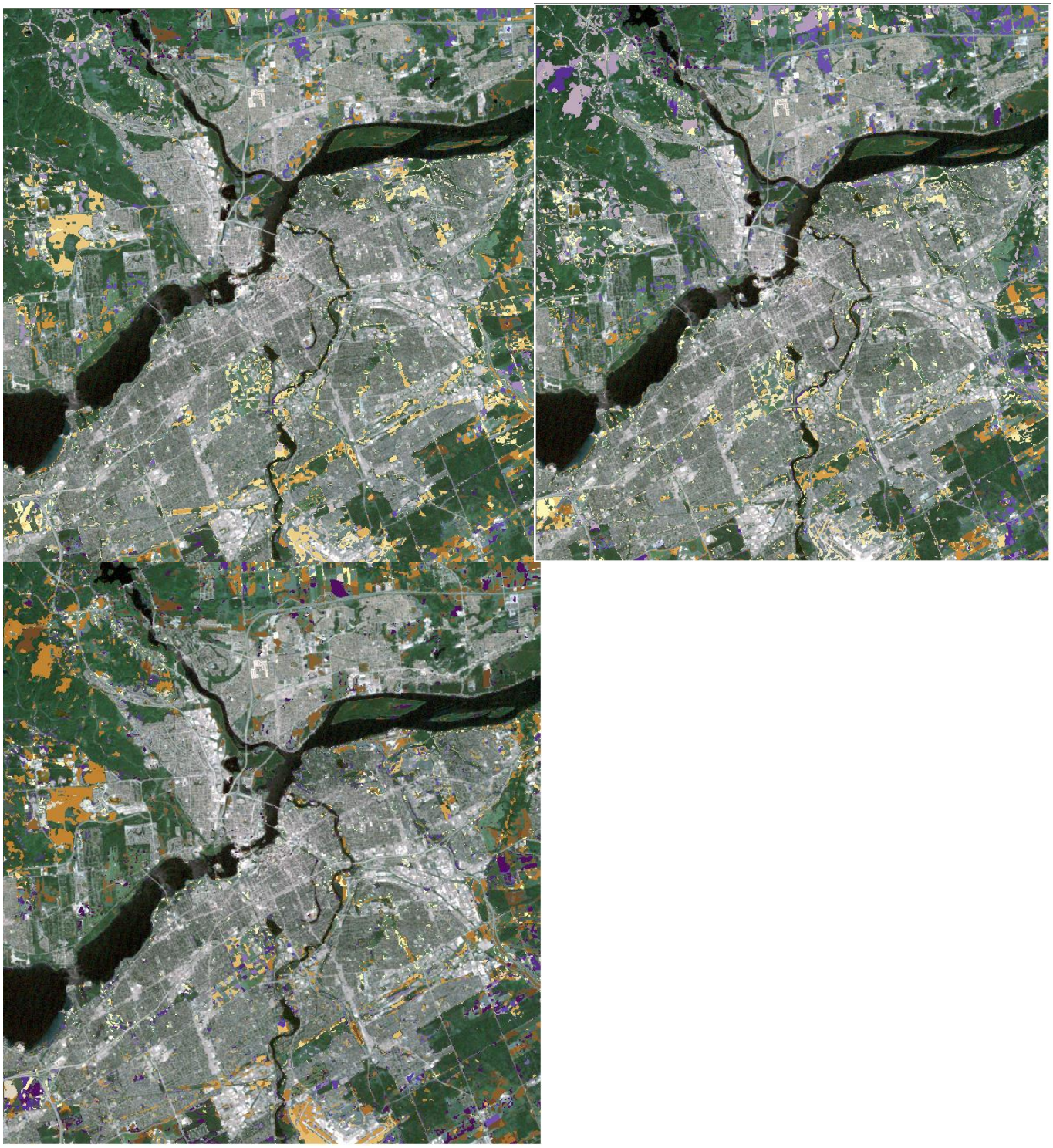


Figure 35: Clockwise from top left: CN change detection from 2004 to 2009, 2009 to 2017 and 2004 to 2017. Purple cells indicate an decrease in runoff, and orange cells indicate a increase in runoff.

Discussion

Flood extent delineation:

As stated before, the literature provided multiple ways of assessing the flood extent of an area during a peak precipitation or flood event. Each step of the analysis was conducted through the use of trial and error in order to determine how to effectively assess the extent of a flood within a urbanized region. There were many limitations that hindered our ability to produce accurate flood extent maps in terms of the temporal, spatial and spectral constraints in the data we had obtained.

In terms of temporal issues, we had issues obtaining active microwave radar data from specific sensors that were mentioned by the literature such as RADAR-SAT, TerraSAR-X and ERS-1 & 2 for the exact peak flooding event dates. Unfortunately, the only data that was available for Spring 2017 was towards the end of the flooding event on May 12th which was mentioned before. This was a significant limitation for our study as imagery that showcased the flood extent during the middle of our time interval would have been able to further verify media reports that came to the conclusion that there was actually a higher flood extent in 2017. Further analysis should be conducted to determine flood extent to a higher degree of accuracy during peak flood times.

We also had issues obtaining in-situ data that could be used a reference points when conducting our classifications and accuracy assessments for the specific land cover/use types. This included land cover/use types such as agricultural areas, trees and urban areas in the case of the flood extent map. Lack of high resolution cloud-free optical data for our specified peak flood

dates also posed as a huge barrier for the accuracy of the flood extent maps. In order to conduct an accuracy assessment, the optical (SPOT 4) imagery that was captured on August 30th, 2006, which was within the range of our study dates was utilized and was classified using a supervised maximum likelihood classifier. Both sets of training data that were created were tested against machine derived training datum that was created from the growth truth image using stratified random sampling with a sample size that includes 90 percent of the pixels in the image. This produced an accuracy of 96-99%; the second set having produced an accuracy of approximately 99 percent. Although we know that the training data created was accurate, we were unable to test the accuracy of the produced flood extent imageries for the selected years.

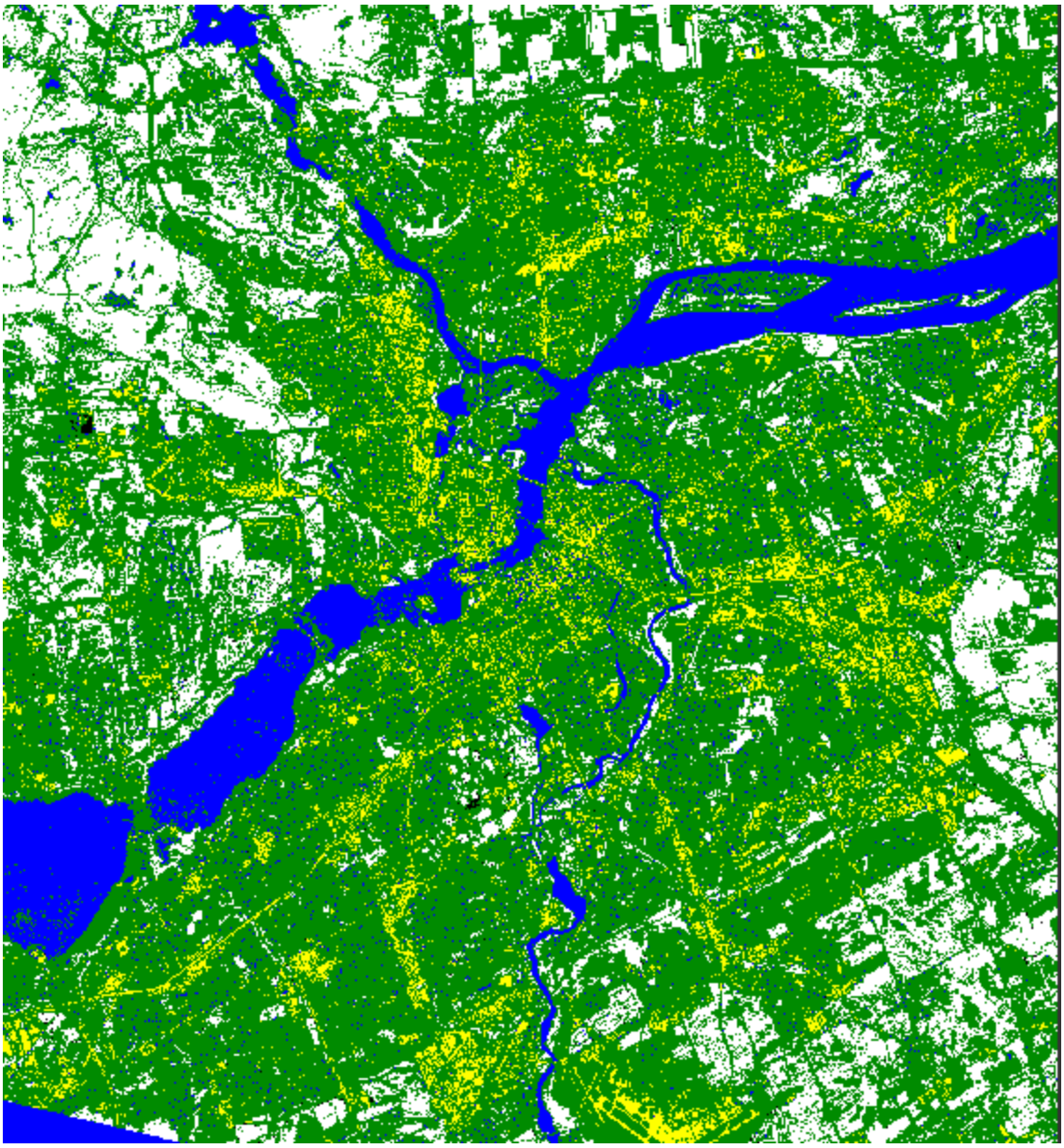


Figure 36..*: SPOT 4 image from August 30th, 2006 classified using maximum likelihood classifier and second set of ROI's. Trees are shown in green; Urban areas are shown in yellow, vegetation shown in white ; water bodies in blue.*

Overall Accuracy = (28753/28985) 99.1996%
Kappa Coefficient = 0.9458

	Ground Truth (Pixels)				
Class	Water	Trees	Unflooded	Urban	Total
Unclassified	0	0	1	0	1
Water	26587	2	0	0	26589
Trees	211	222	10	3	446
Unflooded	0	0	1839	0	1839
Urban	0	5	0	105	110
Total	26798	229	1850	108	28985

	Ground Truth (Percent)				
Class	Water	Trees	Unflooded	Urban	Total
Unclassified	0.00	0.00	0.05	0.00	0.00
Water	99.21	0.87	0.00	0.00	91.73
Trees	0.79	96.94	0.54	2.78	1.54
Unflooded	0.00	0.00	99.41	0.00	6.34
Urban	0.00	2.18	0.00	97.22	0.38
Total	100.00	100.00	100.00	100.00	100.00

Class	Commission (Percent)	Omission (Percent)	Commission (Pixels)	Omission (Pixels)
Water	0.01	0.79	2/26589	211/26798
Trees	50.22	3.06	224/446	7/229
Unflooded	0.00	0.59	0/1839	11/1850
Urban	4.55	2.78	5/110	3/108

Class	Prod. Acc. (Percent)	User Acc. (Percent)	Prod. Acc. (Pixels)	User Acc. (Pixels)
Water	99.21	99.99	26587/26798	26587/26589
Trees	96.94	49.78	222/229	222/446
Unflooded	99.41	100.00	1839/1850	1839/1839
Urban	97.22	95.45	105/108	105/110

Figure 37..: *Accuracy assessment for second set of ROIs against machine derived sample ROIs using 2006 pan-sharpened SPOT 4 imagery*

In terms of limitations with the spatial resolution of the obtained data, it would be ideal to compare the functionality of three supervised classification techniques but that was not always the case since the maximum likelihood classifier required two bands but some of the images only had intensity images for one polarization mode. It would have also been ideal to have at least the HH band included in the Sentinel-1 data that was obtained for the flooding event that took place in 2017, since the bands that were included (VV+VH) were not as effective as the other images that were derived out of the ENVI-SAT ASAR data which included the HH polarization. Water surfaces are susceptible to high wind conditions and heavy amounts of precipitation during peak

flooding events that are better delineated using the HH band primarily as seem especially in the resulting 2004 flood extent image.

In terms of spectral issues, when it came to creating the ROIs which were based on a combination of synthetic aperture radar images, a pan-sharpened optical image as well as supplementary GIS-based data, it was difficult to distinguish trees from urban areas (such as roads and buildings) since they were in close proximity to one another. These two land cover/use types also had very similar backscatter properties in the SAR images which not only posed as a threat for our research but also poses as an issue for future studies on peak flooding events in urban areas with rural attributes as well.

Damage Assessment:

For damage assessment, there are various errors and limitations found throughout the process. Additional improvements could also be made to calculate damages to a higher degree of accuracy. The flood depth was calculated as the average river height of April 2017 instead of using the maximum height from April 2017. Since the flooding occured for the entire month of April and early May, the average was taken to reduce any outlier errors which may be caused by selecting the maximum. Additionally, the DEM was not used for flood depth calculation because the flooding was mostly due to precipitation; since precipitation occurred throughout Ottawa and Gatineau, it was assumed that there would be similar water depth values throughout the study area. This assumes that an elevated area would receive similar precipitation, thus, will have similar depth values. In addition, the soil was assumed to be uniform (for runoff) when calculating ordinary kriging for damage. Different types of ground classifications absorb water

differently, however, when using the kriging interpolation it was assumed that water was not being absorbed.

The ROI separability report can be seen in **Figure 20**. It is shown that, for 2004 and 2017 data, the buildings and roads were similar and harder to distinguish. This affected the result since SPEAR LOC was used with supervised classification and ROI selections for the MF/SAM ratio. However, SPEAR LOC Roads uses additional calculations with the region of interest, therefore the error is not expected to be as large as extracting with SVM.

Figure 21 shows sparse and non vegetated areas for 2004 and 2017. In the images, 2017 shows no sparse vegetation in the river, whereas in 2004 the water also has a large number of sparse vegetation areas. Since the flooding in 2004 was in September, right after summer, the large algae growth from the summer can be seen. Algae cannot grow in winter seasons due to weather conditions, which is seen from 2017. However, since water regions were removed, algae growth has no error problems. However, there are other extractions from this raster which can cause some errors. The sparse vegetation was determined by NDVI, however, NDVI is used for quality of vegetation. Therefore, any bare-lands would most likely be classified as unhealthy vegetation, with the value under 0.50. In addition, some of the vegetation that are low quality would also have lower NDVI values. These non-building and non-road areas classified as non-vegetation will add to additional errors.

The MF/SAM ratio is computed as supervised classification. There does not seem to be large error potential from this section. Regardless, when extracting the road raster into polygons, there is a large discrepancy in the extracted polygons and the raster pixels. This can be seen under **Figure 23** as the pink (extracted polygons) does not fully cover all the green pixels. This

means that the raster to polygon tool did not output polygons that cover the same area that was originally covered by raster. This extraction to polygon was to compute calculations with water areas, sparse vegetation, non-vegetation, and cloud areas which were also extracted as polygons.

Once the polygons of clouds, roads and water are removed from the non-vegetation and sparse vegetation polygons, the outcome should be just the buildings and negligible vegetation mixed with the buildings. However, due to the errors mentioned above, some road networks and non-building types can be seen as classified under sparse and non-vegetation layers. Furthermore, the overall area covered by building is less in 2004 than 2017, as seen under **Figure 24**. This is accurate since the buildings have been increasing due to more urbanization within the area, possibly caused by the growth of mixed use urbanization. However, the area of buildings flooded in 2017 was less than the area flooded in 2004. This error is largely because of the data available for 2017. As mentioned above, The imagery used to calculate extent is near the end of the flood (May 12th), due to lack of data. Therefore, the extent does not cover the entire flooded areas, causing smaller extent, therefore, less buildings being represented as flooded.

There were also errors in the kriging process. The river heights at boundary points that were extrapolated were based on vector equation of a line of two stations which measure river height on the same body of water. This method assumes linear increase/decrease from previous two stations to the boundary point. This is not accurate since the height can be affected by many different factors such as: precipitation, runoff from river islands and river streams. However, without the external points, the entire study area could not be interpolated, therefore, boundary points had to be extrapolated. Additionally, the kriging method with 2 maximum variables was used to reduce these errors, as this will ensure that at least one depth value that is not extrapolated will be used for the interpolation field.

There are few errors that can be seen in the kriging. The cross validation shows the error of predicted regions based on initial points. The values of: mean, root-mean-squared, and mean standardized seen under cross validation results should be as close to 0 as possible for error minimization. The mean value represents the bias from the original values to predicted values **(**ArcGIS, n.d). The root-mean-squared should be close to mean standardized as possible. If the root-mean-squared is greater, this represents an overestimation on the data, whereas, a lower root-mean-squared represents underestimation **(**ArcGIS, n.d). Additionally, the value for root-mean-square standardized must be close to 1. This is fairly close to 1 for 2017 and 2004. Lastly, the average standard error should closely follow the root-mean-square, which can be seen under **Figure 27 (**ArcGIS, n.d). Based on these values, the overall kriging prediction result has some error on interpolated values. Additionally, these errors are not distributed equally throughout the region. QQ plot from the kriging is used to determine the normality of the errors. Since the points do not follow closely to the QQ plot, the errors are not normally distributed throughout the model. Lastly, the final output of kriging has a slant and is not linear vertically or horizontally, as seen under **Figure 30**. This means that the overlap is not aligned perfectly, therefore, any raster that is calculated with this layer will have some errors on the outputting pixels. Although the kriging had error due to alignment problems, it was used with the Hydrotec model to create a damage ratio raster. There are various other methods which could have been used, however, this was method selected due to its simplicity and its close resemblance of damage ratio to the FLEMO+ model (Thieken, Apel, Kreibich, & Aronica, 2007). The error caused by the slant was fixed by calculating the risk based on percentage of pixels lost on the raster calculation. However, this calculation assumes that both sparse and non-vegetation areas were lost in equal

percentage, which may not necessarily be true. Due to the errors stated above, the original hypothesis that 2017 will have higher damage cost than 2004 is inconclusive.

These errors could be largely reduced with a greater amount of higher quality data. The study done by Pushparaj and Hegde (2017) shows a process of building extraction using high resolution data. There was high resolution data from Planet and orthophotography was available. However, since this data only had three bands; red, green and blue, it could not be used for building extraction. However, instead high resolution data with 4 bands can be used for accuracy purposes as it can separate vegetation and roads using object based classification for automatic building extraction. The 30m x 30m Landsat data had too low of a resolution making it difficult to separate different features. In addition to building extraction, the depth calculation could also be performed better with additional data. The more stations there are throughout the Ottawa River and Rideau River, the less assumptions there would be for interpolation purposes. This can be done with in situ data collection. If there were more data points that covered the entire study area, there would have been no need for extrapolation hence reducing some of the errors caused from extrapolation. Additionally, with no reason for extrapolating, the stations selected would extend outwards from the study area, causing no alignment errors.

Additional data would also help with calculating better damage ratio, thus, damage cost. There are no contamination data, building quality data and only some stations have water velocity data. These data could be used to better flood models in terms of providing a more accurate damage ratio. Therefore providing us with a more accurate assessment of the cost associated with those damages (Kreibich, Botto, Merz, & Schröter, 2016).

Land Use and Runoff Modelling Discussion

While many of the land use classification changes were expected, such as the minimal change in the water classification, others were unexpected and required further investigation. An increase in water classification from 2004 to 2009 is likely to have been caused by drought. From 1999 to 2004, Canada had been struck by a serious drought (Bonsal et al., 2011). While this drought had mainly affected the western Prairie Provinces, a significant portion of Quebec had be affected by the droughts, which lasted up until the end of 2004; Sherbrook in Quebec saw a Palmer Drought Severity of -2 in near the end of 2004 (Bonsal et al., 2011). As such, the soil moisture in Quebec in eastern Ontario would be expected to have been drier and capable of holding more moistures, which may have led to abnormal readings in 2004. This would explain the extremely minimal difference between the 2009 and 2017 readings, as well. For the 13% loss of forested land, this is not unusual for an area type such as that of the study area. As the study area became more urbanized over the 13 year period of study, deforestation is one of the most common ways to create more land, whether it be for agricultural purposes or residential purposes. For example, in **Figure 38**, we can see that a large portion of forested area in the north western corner of the study area in Gatineau was removed for the development of residential land use.



Figure 38: Comparison between 2009 urbanization area and 2017 urbanized area

Residential land use saw a loss of 4.6% between 2004 and 2017, a result that seems very peculiar especially considering the loss of forested land to create residential land. If there are examples of increased residential land, how can there be a net loss in residential land? This counter intuitive result is further compounded by the knowledge that Ottawa and Gatineau are both cities with significant population growth, with growth rates higher than that of the average growth rate in Ontario (City of Ottawa, 2011). One explanation of this may be the inclusion of the Greenbelt. The Greenbelt is a section of forested and agricultural land, with a goal to preserve ecological life and stop urban sprawl (National Capital Commission, n.d.). The study area in question perfectly frames the city of Ottawa, with the Greenbelt on the Southern, Western and Eastern borders of the study area. With the Greenbelt, it would be impossible to further urbanize horizontally, as any further urbanization will require encroaching on the Greenbelt land. As such, one possible method for urbanizing within the City of Ottawa is to urbanize vertically; by utilizing mixed used residential and commercial/industrial lands, the City of Ottawa can continue to support its growing population while avoiding the growth of residential sprawl. This can be seen in the industrial classification; the growth from 2004 to 2016 shows an 11.6%

increase in land growth. The relatively small amount of growth can be attributed to the fact that growth in these classifications would take place vertically rather than horizontally.

The 2009 imagery however has an industrial area decrease of 0.13 km^2, another counterintuitive result. After examining the imagery as well as the classification, there does not seem to be any major misclassifications. Comparing the classifications in **Figure 39**, we can see that the 2009 classifications lie within the classification from 2017. However, further inspection shows that the 2009 imagery, especially over the southern half of the imagery, has a significant part of its' imagery occluded to a degree with cloud cover. As such, the brighter returns of the industrial land use are occluded as well, showing difficulties in classification of this land use. For future studies, ensuring imagery with no cloud cover present for all three sections must be regarded with a higher priority.

Another counterintuitive result from the land use classification is the agriculture classification. The incredible jump between 2009 and 2017, with the amount of land nearly doubling from 0.47 km^2 to 0.90 km^2, and nearly tripling from 2004, with an area of 0.35 km^2. The reasoning behind this unintuitive result lies in the difference between the imagery, and likely the sensor itself. Viewing the false colour composite of both Landsat 5 images, it is easy to manually delineate the agricultural land use by eye. However, looking at the 2017 Landsat 7 imagery, the areas of agriculture are very difficult to spectrally differentiate. The other colour composites showed the same difficulties. Inspecting the ROI separability report shows that the spectral difference between agriculture and grasslands are the most difficult for the Support Vector Machine classification to differentiate. Looking at the agriculture classification for the 2017 values, it is clear that the classification is often misclassifying grassy land as agricultural land.

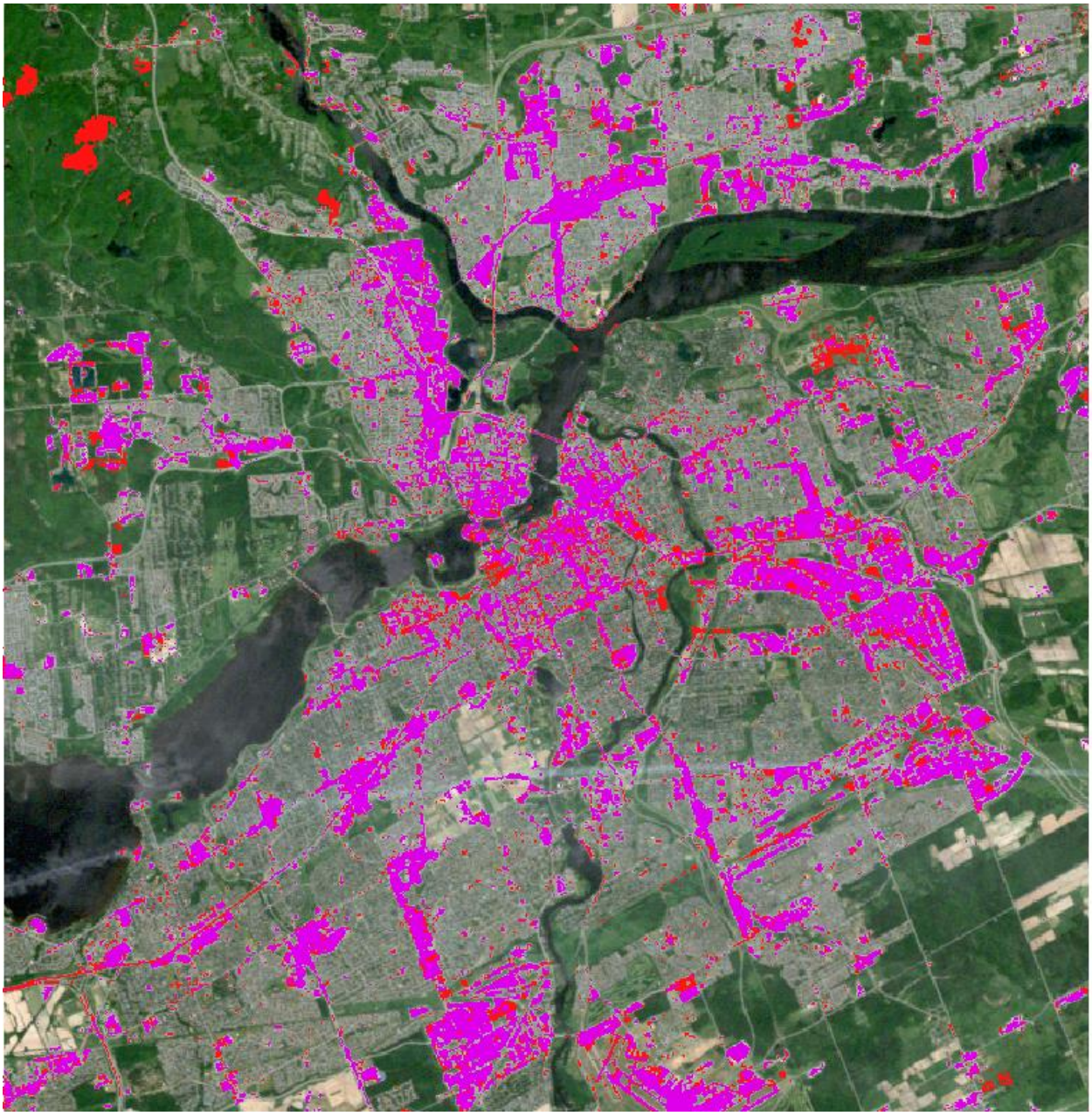


Figure 39: Industrial area classification comparison. Pink areas = 2017 Classification, Red areas = 2009 Classification

While the misclassifications seen may raise concerns about the accuracy of the land classification, it is important to note the reclassification step of the land use classification. When the classification was deemed inaccurate in the agricultural land use, in order to reduce those inaccuracies, agriculture and grass lands were combined. In doing so, the major misclassifications of agriculture were avoided. Furthermore, the classification of residential and

industrial land use were also combined in an effort to increase accuracy. While this leads to a less granular CN grid with only three separate land use groups, the accuracy of these land-use groups would be slightly higher. The land use classification was also limited in its accuracy. Future analyses should consider higher resolution data in order to create more granular classifications. As mentioned earlier, data was limited by the free available data from 2004. Newer and higher resolution data would lead to a higher resolution classification. In terms of accuracy assessments, this project was limited in the amount of data that could be used for assessment. For future projects, a high quality accuracy assessment using in situ measurements should be considered. Higher quality classification methods were also available, as mentioned in the literature, such as using artificial neural networks to spectrally classify the imagery, however these methods were complex and therefore were difficult to implement within the context of this study.

The results of the CN grids show an increase in the amount of possible overland flow for the same conditions on all three time periods. Curve numbers are best demonstrated by the **Figure 40** in the following graph. A curve number of 100 demonstrates perfect amount of runoff, meaning there is no infiltration into the soil, whereas a lower curve number will allow for absorption of the water in the soil until the soil reaches maximum moisture holding capacity (Croshney et al., 1986).

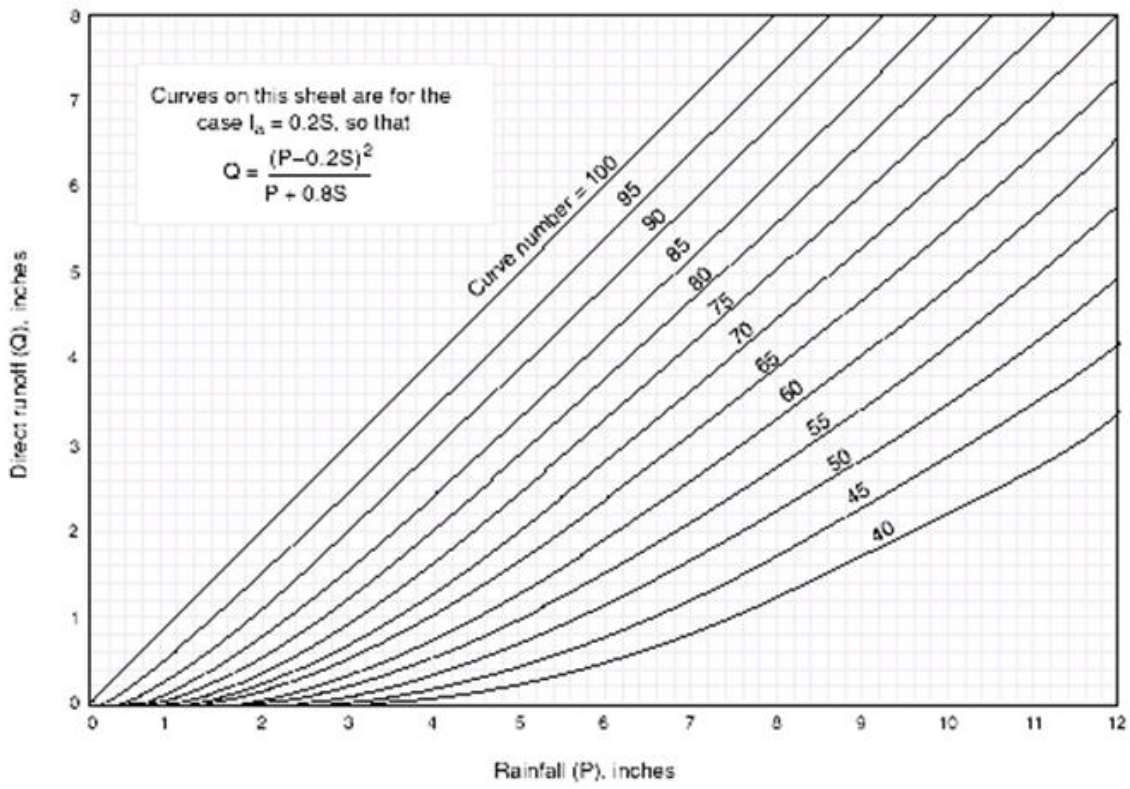


Figure 40: Representation of Curve number grids.

When inspecting the curve number grids, we can see that the curve numbers within Ottawa near the Ottawa River are that of the highest values other than the river itself. This is representative of the flow in terms of precipitation; due to the lack of infiltration from large amounts of urban cement floor and the lack of soil for the water to infiltrate to, the precipitation will flow towards the river. As such, the higher the CN numbers, the more water will be contributed to the river. This is a major cause of flooding in the Ottawa River; due to the large amount of urbanized area, the amount of overland flow that contributes to the river would be massive. Furthermore, with the amount of wetlands located around the City of Ottawa and Gatineau, and with the number of tributaries flowing through both municipalities, areas near the water bodies would be highly saturated, causing a higher amount of overland flow depositing into the water bodies, thereby further raising the water levels of the river. Lastly, floods are most common in the beginning of spring time in the Ottawa River when the added levels of flow caused by rising temperatures, as well as the increased CN numbers due to frozen soils would even further contribute to the overland flow. By inspecting the CN grids, it is clear that the amount of urbanization stretching on the banks of the river have maximized the amount of overland flow on the Ottawa side of the river. Further industrial or commercial urbanization within residential areas would further decrease the amount of exposed soil, thereby increasing the amount of overland flow and causing a more dramatic and damaging flood event.

Further inspection of the CN grids show a discrepancy between the Ottawa and the Gatineau data. This is to be expected due to the aggregation of two data sources, as well as the interpolation of the soil data for the Gatineau region. As such, comparisons between the curve numbers between Ottawa and Gatineau cannot be made at the time. Future studies must contain a consolidated soil survey for both municipalities, it is likely the interpolation of the soil data from

the geological data incurred a large amount of error, as the hydrologic soil groups were simply estimated from the brief descriptions within the metadata. However, basic conclusions can still be made from the comparison of the CN grids within each municipality. Inspecting the change detection rasters for 2004 and 2009, the majority of the changes in the higher curve numbers are within areas of large urbanization. This is mostly taking place within the north western portion of the study area, within Gatineau, as can be seen in **Figure 41**. Interestingly, another portion of land use change which instigates a higher level of overland flow is agriculture. Agriculture provides a higher CN number for all hydrologic soil groups than even medium residential land, perhaps due to the moisture content in the soil. In fact, paired with the large amount of increased land use for agricultural land seen in the classification, it can be argued that the agricultural growth may be a large factor in the increased amounts of flooding within the Ottawa River. Paired with the concept that the Greenbelt is limiting the amount of urban sprawl within the city of Ottawa, agriculture may play a large role in future floods. As such, further studies must be completed to determine a hydrologic model which will conclusively determine the effect of agriculture on flooding.



Figure 41: Comparison between 2004 and 2009 CN change. Compare with Figure #

It is, however, important to note that CN grids do not provide a comprehensive description of hydrologic behaviour. CN grids simply a predictor of runoff that can be seen for average conditions. It does not take into account the directionality of flow, nor does it take into account the water stream network, minimum precipitation or maximum soil retention. CN is simply a variable that can be used to determine these things. Precipitation runoff can be calculated from the below equations. However, curve numbers by themselves cannot predict where and how much areas will be flooded. In the scope of this project, however, it can be used for a general predictor of the contributors to a flood, as the majority of water bodies within the study area lead towards the Ottawa River.

$$\square = \frac{(\square - \square_{\square})^{2}}{\square - \square_{\square} + \square}$$

Where Q = Runoff, □□ is the value of water before runoff event, S is the Soil moisture maximum and P is amount of rainfall

Conclusion

Through an extensive study on the Downtown Ottawa region, it can be understood that while the results of the analyses showed a significant, 43% decrease in terms of flood extent between 2004 and 2017, this decrease can largely be attributed to the failure in data collection. As flood events are quite variable in terms of the water level, thus creating the extent of the flood by capturing the floods during these peak times was not possible. While the 2004 data was able to accurately reflect the amount of flood extent during the peak flood time, the 2017 data was not. The obtained data for 2017 was at least a full week beyond the flood extent time, near the middle of May. This limitation resulted in a much lower water level than what would have been witnessed, thereby showing a much smaller flood extent. From a summary of the 2017 flood created by the Ottawa River Regulation planning board, it is known that the true flood heights at some stations of the 2017 floods rival that of floods since 1915 (ORPB, 2017). In Figure 42, an incredibly sharp decline in water height at the Gatineau station beyond that of May 7th can be seen; this behavior is seen at other measurement stations as well. As such it can be estimated that the flood extent would have been much higher than the extent of the 2004 floods given the correct data.

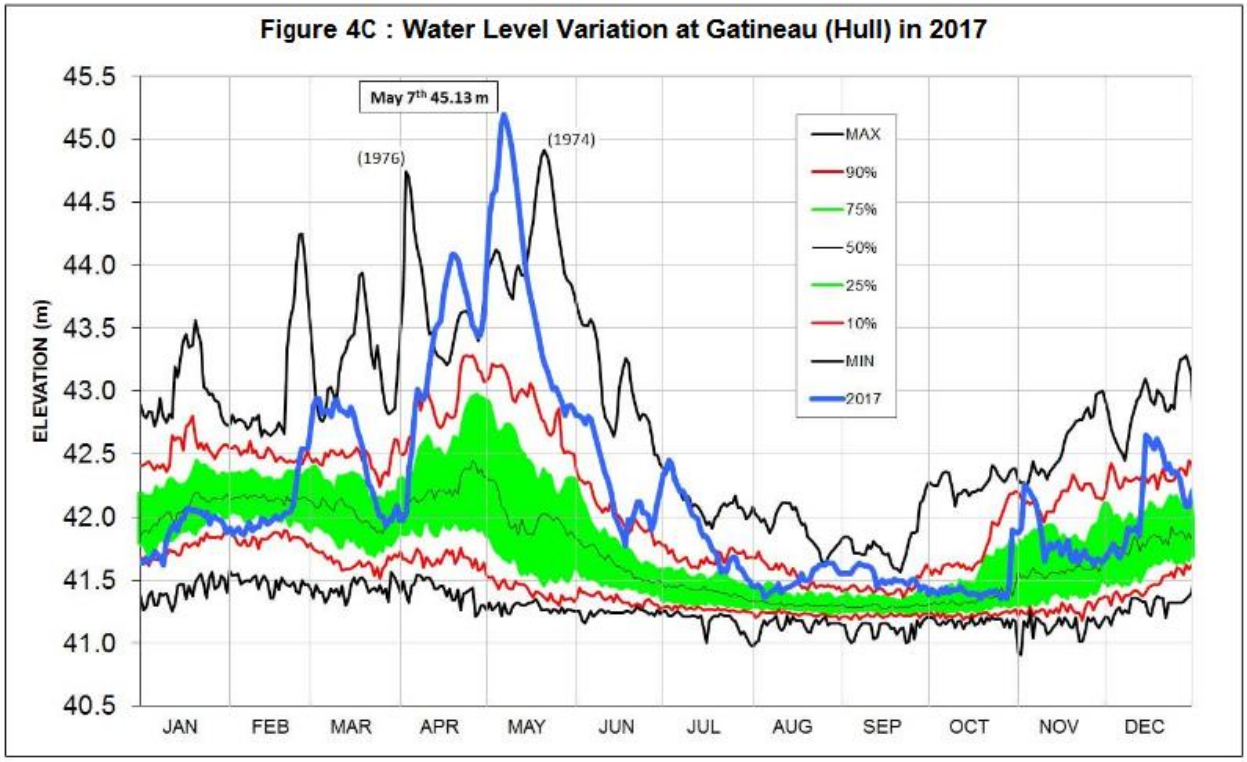


Figure 42: Water Level variation a Gatineau River Station of the 2017 flood.

With an increasing and intensifying trend of such flood extents over the past twenty years, it can be expected that the calculations in terms of economic damages are severely underestimated as well, caused by the underestimation of flood extent. Furthermore, a more accurate assessment in terms of flood depth per area can increase the accuracies in terms of flood depth, as interpolation of flood depths. Future studies should consider alternate methods in order to delineate buildings, such as an artificial neural network. While the potential for error is high for the economic valuation of buildings, the errors are less consequential as the valuation is used for a comparison between the two years and as such should not be used as an empirical economic cost valuation. Furthermore, alternate methods of determining cost evaluation of other land uses such as agricultural loss should be considered. As such, in order for remote sensing to be viable in terms estimating accurate economic costs of flood events, much higher resolution data as well as in situ data must be collected.

When examining the efficacy of determining the impact of flooding in terms of the land use changes in the city of Ottawa, it is obvious that the project is yet again limited by the usable data. In terms of usable data, the Landsat-5 data, with a resolution of 30m, was far too limited to be able to accurately assess the land use. As such, there was a high degree of misclassification, and because of this the amount of land use classes had to be reduced to only four classes. With higher resolution imagery, it would be possible to determine with a higher degree of granularity the amount of land use change, allowing for more conclusions to be drawn in terms of land use change. The difference between agricultural land, grass land and even barren land could yield further insights in terms of urban planning.

Furthermore, this land use classification severely limits the curve number grid results. With only four hydrologic soil groups, four land use classes would yield a total of sixteen possible curve numbers. Furthermore, one of the most important limitations of this project is the lack of soil information collected from Quebec, as well as the omission of certain soil parcels within the Ottawa soil survey. Future studies should consider the availability of available soil data and perhaps perform the soil surveys in house. While the curve numbers simply provide a simplistic view in terms of land use and runoff, it can play an important role in terms of more advanced hydrologic modelling. In fact, while curve numbers can simplistically show how certain land uses may contribute to future flooding, a much more complex hydrological model should be considered in order to conclusively prove land use change and its effects on the entire basin of the Ottawa River. As such, future studies should consider expanding the scope beyond the study area as well as consider a much more detailed hydrologic basin modelling, especially with details such as the City of Ottawa's sewer and drainage infrastructures as well as climatological factors such as soil temperature and snowmelt, alongside hydraulic systems and

networks analysis. This will provide a more comprehensive analysis on planning for mitigation strategies.

While this project has been limited especially in terms of the available data, there are many conclusions that can be made. While the inclusion of the Greenbelt in the City of Ottawa limited the amount of growth in terms of urbanization, it can be concluded that the amount of urbanization, which is known to be a factor in terms of flooding, was not a large contributor to the amount of flooding occurring in the Ottawa floods. It also shows that even though there was a flood extent reduction of 43%, this was seen near the end of the flood event in 2017 and as such the peak floods extent of the 2017 flood would likely have surpassed the 2004 floods. As such, this report shows the need for further research to be done in order to appropriately plan for mitigation strategies for future floods of the Ottawa River.

References

Apollonio, C., Balacco, G., Novelli, A., Tarantino, E., & Piccinni, A. (2016). Land Use Change Impact on Flooding Areas: The Case Study of Cervaro Basin (Italy). Sustainability, 8(12), 996. doi:10.3390/su8100996

ArcGIS. (n.d.). ArcMap. Retrieved April 04, 2018, from http://desktop.arcgis.com/en/arcmap/latest/extensions/geostatistical-analyst/performing-cross-validation-and-validation.htm

Bonn, F., & Dixon, R. (2005). Monitoring flood extent and forecasting excess runoff risk with RADARSAT-1 data. *Natural Hazards*, *35*(3), 377-393.

Bonsal, B. R., Wheaton, E. E., Chipanshi, A. C., Lin, C., Sauchyn, D. J., & Wen, L. (2011). Drought Research in Canada. *Atmosphere-Ocean*, https://www.tandfonline.com/doi/pdf/10.1080/07055900.2011.555103.

Brivio, P. A., Colombo, R., Maggi, M., & Tomasoni, R. (2002). Integration of remote sensing data and GIS for accurate mapping of flooded areas. *International Journal of Remote Sensing*, *23*(3), 429-441.

Cronshey, R., McCuen, R. H., Miller, N., Rawls, W., Robbins, S., & Woodward, D. (1986). Urban Hydrology for Small Watersheds - TR55. *United States Department of Agriculture,* https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1044171.pdf.

“Envisat”. (2012), https://www.esa.int/Our_Activities/Operations/Envisat

“Filters”. (n.d.). *Documentation Center [Harris Geospatial Docs Center]*, www.harrisgeospatial.com/docs/Filters.html.

Fulton, R. J. (1995). Surficial materials of Canada. *Geological Survey of Canada*. doi:10.4095/205040

Government of Canada. (2018, March 14). Real-Time Hydrometric Data Map Search, https://wateroffice.ec.gc.ca/google_map/google_map_e.html?map_type=real_time

Harris. (n.d.). Using ENVI, http://www.harrisgeospatial.com/docs/routines-136.html

Insurance Bureau of Canada. (2017). Canada – Spring Flooding in Ontario and Quebec Caused More Than C$223 Million in Insured Damage, http://floodlist.com/america/canada-spring- floods-ontario- quebec-caused- millions-insured-damage

Infrastructure & Economic Development Dept. (2016). Ottawas Population. Retrieved April 05, 2018, https://ottawa.ca/en/city-hall/get-know-your-city/statistics-and-economic-profile/statistics/ottawas-population

Kreibich, H., Botto, A., Merz, B., & Schröter, K. (2016). Probabilistic, Multivariable Flood Loss Modeling on the Mesoscale with BT-FLEMO. *Risk Analysis*, *37*(4), 774–787. http://doi.org/10.1111/risa.12650

Lefohn, A., Knudsen, P., & Shadwick, D. (2005). Using Ordinary Kriging to Estimate the Seasonal W126, and N100 24-h Concentrations for the Year 2000 and 2003. *United States Department of Agriculture Forest Service*, 1-10, https://webcam.srs.fs.fed.us/ozone/spatial/2000/contractor_2000_2003.pdf.

Li, L., Xu, T., & Chen, Y. (2016). Improved urban flooding mapping from remote sensing images using generalized regression neural network-based super-resolution algorithm. *Remote Sensing*, *8*(8), 625.

Long, S., Fatoyinbo, T. E., & Policelli, F. (2014). Flood extent mapping for Namibia using change detection and thresholding with SAR. *Environmental Research Letters*, *9*(3), 035002.

Mason, D. C., Speck, R., Devereux, B., Schumann, G. J. P., Neal, J. C., & Bates, P. D. (2010). Flood detection in urban areas using TerraSAR-X. *IEEE Transactions on Geoscience and Remote Sensing*, *48*(2), 882-894

.

Mockus, V., & Werner, J. (1994). Hydrologic Soil Groups. *Design Hydrology and Sedimentology for Small Catchments*, 485-495. doi:10.1016/b978-0-08-057164-5.50020-7

National Capital Commission. “Canada's Capital Greenbelt.” *The Ottawa Greenbelt*, ncc-ccn.gc.ca/places-to-visit/greenbelt.

ORPB. (2017). Summary of the 2017 Spring Flood. *Ottawa River Regulation Planning Board*.

Principles of Remote Sensing - Centre for Remote Imaging,Sensing and Processing, CRISP. (2001).

Pushparaj, J., & Hegde, A. V. (2017). A comparative study on extraction of buildings from Quickbird-2 satellite imagery with & without fusion. *Cogent Engineering, 4*(1). doi:10.1080/23311916.2017.1291118

Rijus. (n.d.). HOUSE BUILDING COSTS IN SOUTHERN ONTARIO, http://www.rijus.com/ontario-house-building-costs.php

Stocker, T. F., Dahe, Q., & Plattner, G. K. (2013). Technical Summary. *Climate Change 2013 - The Physical Science Basis,* 31-116. doi:10.1017/cbo9781107415324.005

Szuster, B. W., Chen, Q., & Borger, M. (2011). A comparison of classification techniques to support land cover and land use analysis in tropical coastal zones. *Applied Geography, 31*(2), 525-532. doi:10.1016/j.apgeog.2010.11.007

The Editors of Encyclopædia Britannica. (2018, January 18). Ottawa River, https://www.britannica.com/place/Ottawa-River

Thieken, A., Apel, H., Kreibich, H., & Aronica, G. (2007, May 16). *Comparison of different approaches for flood damage and risk assessment* [Powerpoint Slides]. Germany, Potsdam.

Thomson, S. (n.d.). What are the Sustainable Development Goals?, https://www.weforum.org/agenda/2015/09/what-are-the-sustainable-development-goals/

USDA. (n.d.). Updated Hydrologic Soils Group (HSG) Questions & Answers. *USDA Natural Resources Conservation Service*

USGS (n.d.) "What Is Landsat 7 ETM SLC-off Data?" *What Is Landsat 7 ETM SLC-off Data? | Landsat Missions,* landsat.usgs.gov/what-landsat-7-etm-slc-data.