# WEB-BASED FLOOD RISK ASSESSMENT – RAPID, USER-FRIENDLY TOOLS LEVERAGING OPEN DATA

## HEATHER DANA MCGRATH

**April 2017**

# WEB-BASED FLOOD RISK ASSESSMENT – RAPID, USER-FRIENDLY TOOLS LEVERAGING OPEN DATA

Heather Dana McGrath

Department of Geodesy and Geomatics Engineering
University of New Brunswick
P.O. Box 4400
Fredericton, N.B.
Canada
E3B 5A3

April 2017

**PREFACE**

This technical report is a reproduction of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Geodesy and Geomatics Engineering, April 2017. The research was co-supervised by Dr. Emmanuel Stefanakis and Dr. Miroslav Nastev, and funding was provided by the Canadian Safety and Security Program (CSSP), which is led by Defence Research and Development Canada's Centre for Security Science, in partnership with Public Safety Canada, and additionally supported and funded by the University of New Brunswick and the New Brunswick Innovation Foundation (NBIF).

As with any copyrighted material, permission to reprint or quote extensively from this report must be received from the author. The citation to this work should appear as follows:

# ABSTRACT

Timely and accurate prediction of flood inundation extent and potential negative impacts and consequences is fundamental for the sustainable development of a given region and allows decision makers and the local community to understand their exposure and vulnerability. Complex computer models exist for flood risk assessment and while technologically sophisticated, these programs are intended, first of all, for use by a small number of technical and scientific experts and require considerable processing time and extensive inputs. These existing solutions are generally not well suited for flood prediction in near real-time and often exceed the data available for any given community. This research developed standardized methods, adapted into user-friendly tools which accept limited user input, are based on hydrologic principles and processes, widely accepted risk computation methods and leverage open data. The developed flood mapping approaches access, and through a novel data fusion method, create a better quality digital elevation model (DEM) from multiple open source elevation datasets. This fused DEM is combined with other open source data (e.g., IDF curves, river flow data, watershed boundaries, etc.) to generate a flood inundation surface through two methods: (i) a 0D bathtub model and (ii) a hybrid 1D/2D raster cell storage approach. The 0D model ignores flow rates and changes over time, producing a grid of the maximum spatial extent and depth, calculated as the difference between the terrain elevation and the computed water surface. The hybrid model solves 1D kinematic wave approximation of shallow water equations in the channel and treats the floodplain as 2D flooding storage cells. Water depths from the flood grid are combined with local inventory data (e.g., building structural type, occupancy, valuation, height of the first floor, etc.) to compute

exposure and damage estimates in either a user friendly MS Office application or a web-based API. The developed methods and user-friendly tools allow non-experts the ability to rapidly generate their own flood inundation scenario on demand and assess risk, thus minimizing the gap between the existing sophisticated tools, designed for scientists and engineers, and community needs in order to support informed emergency response and mitigation planning.

# DEDICATION

To my family for their love and support, and to Viaduct.

I'd like to express my gratitude to my supervisors for their encouragement, support,

guidance and mentorship throughout this project.

# ACKNOWLEDGEMENTS

# Table of Contents

# List of Figures

# List of Tables

# 1. Introduction

Every year, disastrous climatological and geologic hazards take place in Canada and across the globe [Nastev and Todorov, 2013]. In recent years, there is a trend of increased damages resulting from natural disasters, specifically floods. In New Brunswick, over 70 floods have been recorded since the 1700s, with a single event in 2008 causing an estimated $23 million dollars in damages [Public Safety Canada, 2014]. The costliest natural disaster on record in Canada is the June 2013 flood in southern Alberta, with reported damages exceeding $6 billion dollars [Environment Canada, 2013]. A number of factors have been identified which contribute to increasing flood damages, including: population growth, increased urbanization in flood-prone areas and the changing climate [Jongman et al., 2012; de Moel and Aerts, 2011; UNISDR, 2011]. As such, government officials, GIS specialists, emergency managers, and first responders look for tools to assess risk, identify vulnerable communities, and develop mitigation strategies and emergency response plans [Neighbors et al., 2013].

Through the use of computer models which simulate hazards and compute exposure one can evaluate the cost effectiveness of mitigation measures, optimize investments, and enable insurance companies, municipalities and residents to prepare for disasters [Apel et al., 2009; de Moel and Aerts, 2011]. A limitation of existing computer models is the requirement of highly trained personnel to prepare the necessary input (hazard, inventory of the built environment, and vulnerabilities) and analyze model outputs. Of the watershed modelling and risk assessment applications available today, few are capable of non-expert implementation [Al-Sabhan et al., 2003]. In addition, the data requirements and data manipulation required to run these complex models often exceed the available

data as well as the technical capabilities of the broader non-expert safety community [Nastev et al., 2015]. These existing models therefore leave a gap between what is needed and when by decision makers [Leskens et al., 2014].

This PhD dissertation presents developed standardized methods and tools for risk assessment which fill in the current gap between rapid, user-friendly tools and sophisticated tools designed for use by scientists and engineers, with respect to flood risk analysis. This research commenced with an identification of major influencing parameters and data used to compute flood hazard and assess community vulnerability. From these results, the key components were identified and applications which compute flood hazard and estimate exposure and potential damages were developed, with a primary impetus on leveraging open source data and minimal user input while incorporating established equations and processes. This is an article-based PhD dissertation, which is presented and supported through the following chapters:

Paper 1 (Peer Reviewed)
**McGrath, H.,** Stefanakis, E., and Nastev, M. (2015). "Sensitivity analysis of flood damage estimates: A case study in Fredericton, New Brunswick." *International Journal of Disaster Risk Reduction,* 14, 379-387. doi:10.1016/j.ijdrr.2015.09.003.

Paper 2 (Peer Reviewed)
**McGrath, H.,** Stefanakis, E., and Nastev, M. (2016). "Rapid Risk Evaluation (ER2) Using MS Excel Spreadsheet: a Case Study of Fredericton (New Brunswick, Canada)." *International Society for Photogrammetry and Remote Sensing Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume III-8, 27-34.

Paper 3 (Peer Reviewed)
**McGrath, H**., Stefanakis, E., and Nastev, M. (2016). "DEM Fusion of Elevation REST API Data in Support of Rapid Flood Modelling." *Geomatica,* Vol. 70, No. 4

Paper 4 (Under Review)

**McGrath, H**., Stefanakis, E., and Nastev, M. (2017). Online Reduced Complexity Flood Modelling: Leveraging Open Data and Limited User Input

## 1.1 Dissertation Structure

This research is presented as a six chapter, article-based dissertation, Figure 1.1. Chapter 1 provides an introduction to and description of the motivation for this research. The next four chapters (Chapter 2 to Chapter 5) present peer reviewed or under review, at the moment of drafting of the thesis, journal papers, while Chapter 6 provides a summary and conclusion of the presented research and contribution. In Chapters 2 through 5, the first author conducted the primary research while the co-authors provided auxiliary advice on content and structure.
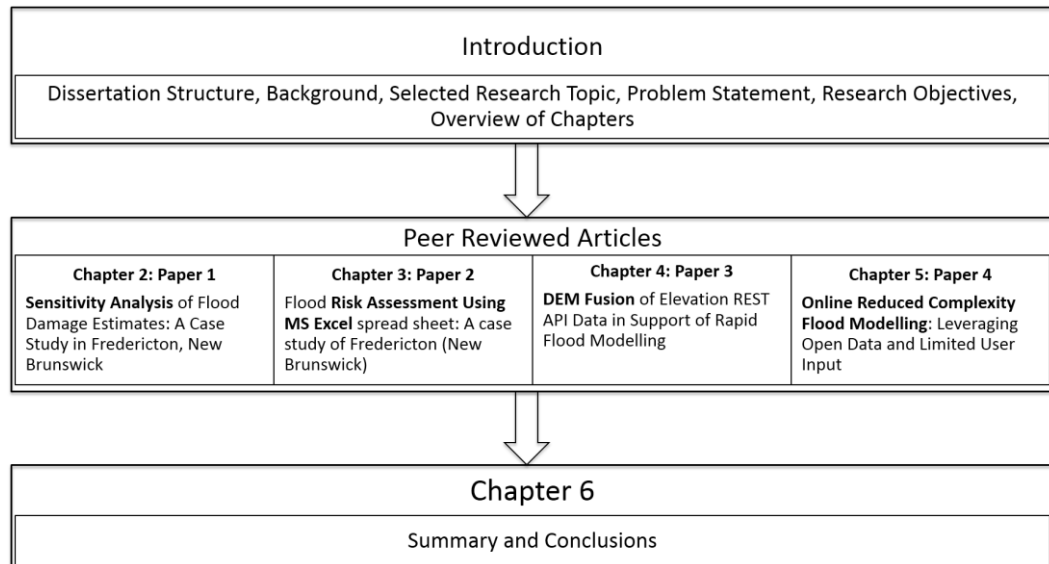
**Figure 1.1** Dissertation structure

## 1.2 Background

Riverine flood risk analysis is the process of measuring the likelihood of the negative impacts and involves the combination of: flood hazard model, inventory model

of the exposed built environment, and a selection of respective vulnerability functions [Nastev and Todorov, 2013], Figure 1.2. While sophisticated software solutions exist, they are typically based on desktop solutions, requiring commercial programs, extensive processing time, sizable inputs, and expert knowledge to run and interpret results [Leskens et al., 2014]. As such, they are not well adapted to respond to the needs of the non-expert public safety community to fully understand their own exposure (value of assets at risk) and vulnerability to inundations. On the other hand, timely and accurate prediction of inundation extent and potential impacts and consequences is fundamental for the sustainable development of a given region and provides valuable information necessary for understanding respective exposure and vulnerability [Scawthorn et al., 2006]. Currently, no application is suitable or available specifically for interventions where flooding is imminent or in progress [Poulin et al., 2012].
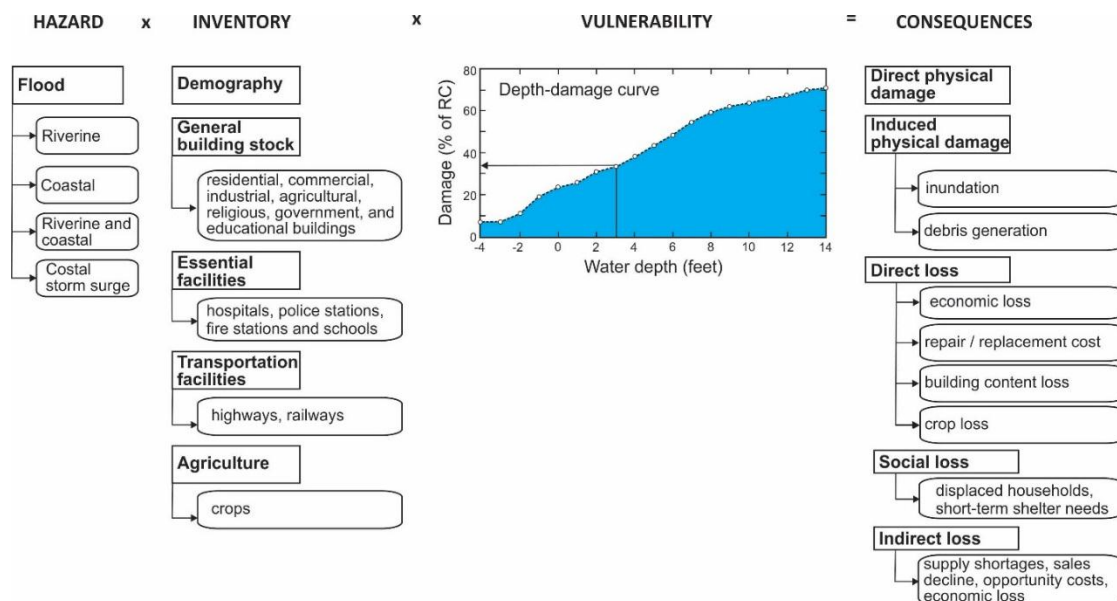


**Figure 1.2** Schematic representation of hazards, inventory, vulnerability and risks.
Source: [Nastev and Todorov, 2013]

### 1.2.1   Flood Modelling

Modelling urban flooding is a complex process and a vast amount of research and literature exist addressing different modelling and simulation aspects [Kulkarni et al., 2014]. Most modelling estimations of design flow are based on methods designed more than 20 years ago and there is ambiguity in model output depending on the input data, type of model selected, and description of river geometry [Cook and Merwade, 2009]. Depth of submersion, flow velocity, sediment load, and duration of flooding are all parameters which contribute to flood hazard, however, access to this data is often limited or non-existent [Poulin et al., 2012].

Hydraulic models are classified through different forms or numerical methods, or by their dimensionality. There are six common approaches: 0D, 1D, 1D+, 2D-, 2D, 2D+, and 3D [Pechlivanidis et al., 2011]. The 0D model involves no physical laws or processes, the flood surface is simply calculated as the difference between the terrain and the modeled water surface, where the water surface may be flat or an inclined plane. 1D models, as found in HEC-RAS (U.S. Army Corps of Engineers (USACE), http://www.hec.usace.army.mil/software/hec-ras/) and Mike 11 (DHI, http://www.mikebydhi.com/products/mike-11), are based on 1D solutions of Saint Venant equations and simulate floodplain flow as part of the 1D channel with an assumption that flow is in the same direction of the main channel [U.S. Army Corps of Engineers, 2010]. The 1D+ approach abandons the assumption of floodplain flow in the same direction and models the floodplain as storage reservoirs, thereby taking a storage cell approach in the floodplain flow simulation. While there are obvious limitations to the 1D approach, case studies in narrow floodplains (where the width of the floodplain is less

than three times the width of main channel) with no separation factors (e.g.: embankments or levees) of HEC-RAS and Mike11 have been validated and show reasonable results [Pechlivanidis et al., 2011].

2D models of flood inundation are derived from the dynamic 2D water flow equations (e.g.: HEC-RAS-2D, Mike 21, FLO-2D (www.flo-2d.com/), and ANUGA (open source, AU)) which describe flowing water in both longitudinal and lateral directions while assuming a hydrostatic pressure distribution. These fully dynamic models require considerable input and computation time and generally yield reliable results in urban environment [Pechlivanidis et al., 2011].

3D models require significant computational power and specialized scientific and engineering expertise primarily found at universities and larger hydraulic laboratories [Haestad Methods, 2003]. Inputs to the 3D model include: velocities in the $x$, $y$, and $z$ direction collected at nodes across the river, water quality and sediment samples. As well, these models can accommodate 3D hydrodynamics, salinity, and sediment transport conditions.

Despite the advantages of sophisticated flood modelling applications, experience shows many of the existing models are of limited use in flood disaster management [Leskens et al., 2014]. A primary disadvantage of these applications is that necessary input data requirements often exceed available data [Al-Sabhan et al., 2003].

The best model is often the one which provides the end user the information required whilst using proxies or reasonably fitting the available data, as the processes necessary to include the best approximate historic/future events are still subject to considerable uncertainties [Bates and De Roo, 2000]. A popular alternative to save

computation time without losing accuracy takes advantage of both the 1D river model

and the 2D model for floodplain simulation. Increasing consensus in the literature

indicates that the channel flow below bankfull depth can be adequately described by a

simplified form of 1D Saint Venant equations [Hunter et al., 2007]; however, the 1D

models have difficulties simulating field conditions when transferred to the floodplain.

Coupling 1D channel flow with a 2D raster storage cell approximation for the floodplain

has produced models which are computationally efficient and suited to adequately

reproduce the hydrograph and inundation measurements simultaneously [Bates and De

Roo, 2000; McMillan and Brasington, 2007]. Advantages of this solution are numerous,

including: reliance on regular gridded digital elevation models to parameterize flows,

quality of spatial predictions which are comparable to similar finite element codes, with

much shorter runtimes. [Bates and De Roo, 2000; McMillan and Brasington, 2007].

### 1.2.2 Vulnerability

In the risk assessment process, vulnerability indicates the susceptibility to sustain a

certain level of damage or loss [Nastev and Todorov, 2013]. Knowledge of vulnerability

remains one of the biggest hurdles in flood risk assessment [Koks et al., 2015].

Traditionally, risk assessment studies include the physical vulnerability of structures to a

certain flood hazard or assessing the risk to life through the assumption of homogenous

vulnerability across the study region.

The most common and internationally accepted method for assessment of urban

flood damage is through the use of depth-damage curves or stage-damage curves [Plazak,

1984; Prettenthaler et al., 2010]. Structure and contents damage resulting from flood

hazard are influenced by many factors; however, usually only building use and inundation depth are considered as damage-causing factors and included in the formulation of depth-damage curves [Merz et al., 2004]. Depth-damage curves relate water depth to estimates of damage to various types of infrastructure to estimate potential damage [Scawthorn et al., 2006]. Depth-damage curves, at a minimum, require two inputs, namely the occupancy classification of the structure and the depth of flooding (Figure 1.3). The output is an estimate of the damage, expressed generally as a percentage of the replacement cost. Depth-damage functions are developed separately for structural or load-bearing components; for contents (e.g.: interior furniture, art, appliances, etc.,); and for inventory in place (e.g.: commercial stock) [FEMA, 2010].

The suite of damage functions referenced in this research are extracted from Hazus, are termed 'credibly weighted', and considered regionally applicable [Scawthorn et al., 2006]. Hazus is a GIS based quantitative risk assessment and decision support tool for natural hazard risk mitigation and emergency management [FEMA, 2010], developed in the U.S by the Federal Emergency Management Agency (FEMA) and adopted for use in Canada (Nastev and Todorov, 2013). These depth-damage curves are based on the best available damage data from floods in the U.S. and represent more than 20 years of losses They were derived from data collected and analysed by the Federal Insurance Administration (FIA) and surveys completed by the U.S. Army Corps of Engineers (USACE) on U.S. infrastructure [FEMA, 2012].
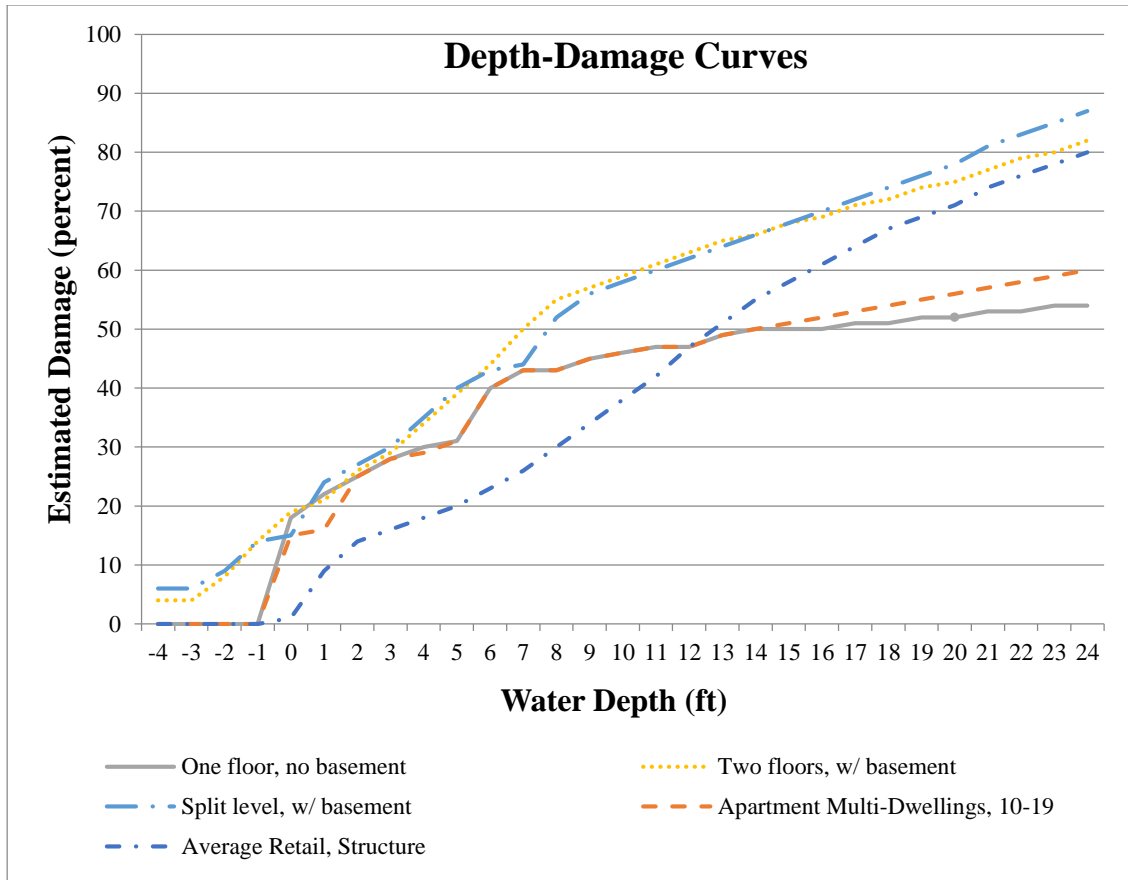
**Figure 1.3** Depth-damage curves for select building occupancy types

### 1.2.3 Consequences

Negative consequences are computed by combining the flood hazard and the vulnerability of the inventory data of the exposed assets. Inventory data includes detailed information regarding, for example, infrastructure, buildings, and population. There is a wide variety of risk assessment models in use internationally, differing substantially in their approaches and estimates of economic costs [Jongman et al., 2012].

Direct losses occur as result of direct physical contact of the flood water with humans, properties, or other objects, while indirect losses represent those which are induced by flood impact and may occur (in time or space) outside of the flood event [Merz et al., 2004]. Direct economic losses include calculations of repair and

9

construction costs resulting from the flood event, whereas indirect economic losses can be related to lost jobs, business interruption, increased transportation costs, etc. [FEMA, 2010]. For this study, it was the calculation of the direct economic losses which were of primary interest in the development of the risk assessment tool referred to as Rapid Risk Evaluation ($ER^2$) (Chapter 3).

$ER^2$ is designed and implemented as described in Chapter 3 in an intuitive and easy to use MS Excel worksheet. To widen the scope and use of $ER^2$ and allow easier access to local inventory data, the underlying calculations, tables and equations were re-used and implemented in a web-based API (hmcgrat1.ext.unb.ca/ER2_Online/Index.html). This API can be used as a web-page where a user individually adds buildings; however, the primary benefit of the API is the ability to programmatically formulate an "unclean URL", for example from a web-mapping application. The request is posted to the host server and the returned result is exposure and estimated damages in a well formatted XML or JSON document. Variables in the *GetDamages* request for building-by-building processing include: building count (BC), building occupancy (BO), foundation type (FT), year built (YB), number of stories (ST), basement (BA), garage (GA), building quality (BQ) and water depth (WD). An example request and response is shown in Figure 1.4.

http://hmcgrat1.ext.unb.ca/ER2_Online/FloodRiskEvaluation?Request=GetDamage&amp;Format=XML&amp;ID=1&amp;BC=1&amp;BO=RES1&amp;FT=0&amp;YB=1987&amp;ST=1&amp;BA=0&amp;GA=0&amp;WD=2&amp;BQ=Average

```xml
<xml>
  <Part_1>
    <Initial_Parameters>
      <ID>1</ID>
      <BA>0</BA>
      <ST>1</ST>
      <WP>100</WP>
      <BO>RES1</BO>
      <DC>105</DC>
      <BQ>Average</BQ>
      <YB>1987</YB>
      <FT>0</FT>
      <GA>0</GA>
      <WD>2</WD>
    </Initial_Parameters>
    <Computed_Variables>
      <BuildingQualityNumber>2</BuildingQualityNumber>
      <CurveDamageID>R11N</CurveDamageID>
      <CurveDamageNumber>105</CurveDamageNumber>
      <AgeSubtraction>29</AgeSubtraction>
      <AgeDeprecation>1.00290952</AgeDeprecation>
      <SquareFootage>2000</SquareFootage>
      <GarageCost>0</GarageCost>
      <AverageBasementCost/>
      <BasementExtraCost/>
      <BoIndex>23</BoIndex>
      <BuildingValuation>186220</BuildingValuation>
      <WaterFoundationSubtraction>2</WaterFoundationSubtraction>
      <StructureDmgPercentage>25</StructureDmgPercentage>
      <ContentDmgPercentage>35</ContentDmgPercentage>
      <ContentCurvePercentage>50</ContentCurvePercentage>
    </Computed_Variables>
  </Part_1>
  <totalBuildingValuation>186220</totalBuildingValuation>
</xml>
```

**Figure 1.4** Example XML request and response from ER$^2$ API

### 1.2.4   Existing solutions

Flood risk assessment tools have been around for decades. However given

technological advances in data collection, processing and modeling, these tools are

undergoing modernization [Messner and Meyer, 2006]. These tools, which historically

have been directed at planning for natural disaster response or developing flood insurance

rate maps [FEMA, 2010], are increasingly being used in a broader range. Existing

software applications include licensed, closed source, open access and open source

solutions, where open source refers to software in which the source code is available to the end user.

In 2011, after researching existing applications which assess geohazard risks, the Public Safety Geoscience Program (PSG) of Natural Resources Canada (NRCan) made the decision that the U.S. Federal Emergency Management Agency (FEMA) Hazus MultiHazard program was one of the best practice methods and decided to adopt the program for the Canadian environment [Nastev and Todorov, 2013]. Both the earthquake and flood modules have been adapted for Canada. A limitation, however, of the Hazus-Canada flood model is that the hydrology and hydraulic computation components have not been fully enabled. Instead, users must upload pre-computed flood hazard grids into the program in order to assess and visualize respective negative consequences. A further concern with the Canadian adaptation is that extensive re-development of Hazus was required. This began in 2011 when Hazus was compatible with Esri ArcMap 10.0. Since that time, FEMA has implemented major changes to the U.S. version, with respect to the underlying data structure, programming and ArcGIS compatibility. The Canadian version is now out of sync with the US version and compatible with an outdated version of ArcMap. Given that this solution only assesses risk and does not compute the flood hazard, there was uncertainty that additional funding for Hazus Canada flood model would be prudent. In the meantime, the support for the Canadian version has been phased out as considerable expense and time would be required to modernize it.

In 2014, the World Bank reviewed open source and open access software packages from around the globe to understand the strengths and challenges of each [World Bank, 2014]. A total of 82 programs covering all natural hazards were initially examined, with

the number being reduced to 31 programs after initial testing and review. A range of

criteria was assessed with respect to these existing programs which offer combined flood

hazard computation and risk assessment:

- Access: open source, open access or licensed
- Flood modeling capabilities
- Data Inputs and outputs
- Technical skill of user/ease of use

Table 1.1 provides a summary of key software programs which made it to the top of their

list, including two licensed programs, Cadyri and Mike11, not considered by the study. In

terms of coding language, Python or Fortran are considered best practice; however C++

is also popular, as is Java for its fast computational abilities [World Bank, 2014]. A well-

presented and easy to interpret graphical user interface (GUI) is important to allow ease

of use, which was found in InaSAFE. The World Bank [2014] found programs which

require inventory data for the risk module should provide sample datasets available for

download and release clear documentation noting structure and details of the required

data. Many of the tested programs from Table 1.1 are extensions of, or plug into, popular

GIS packages, such as: Esri's ArcGIS (Cadyri, Hazus, and Mike11) or Quantum GIS

(InaSAFE), while others, such as RiskScape and Kalypso have designed their own GIS

platform and are therefore not reliant or limited by external software. Most of the

applications, with respect to flood modeling, solve some version of the 1D Saint Venant

equations, while the Cadyri application instead requires a minimum of two existing flood

grids and uses linear regression to compute flood grids at user defined flow rates [Poulin

et al., 2012].

**Table 1.1** Summary of select existing programs, partially adapted from World Bank [2014] report

| Name / Country | License/ Language/ Dependency | Flood Model | Inputs/Outputs | Risk Module | Ease of use |
|---|---|---|---|---|---|
| **Cadyri** (Canada) | Licensed<br><br>Python<br><br>Requires: ArcGIS | *Yes.* Computes new flood grid at user input discharge rate using linear regression | Inputs: Minimum of two flood grids | *ArcGIS*<br><br>python tools built into ArcGIS toolbox | Moderate |
| **Delft 3D Flow** (Worldwide) | Open Source<br><br>(C++) | *Yes.* 3d flow takes all boundary phenomena into account unsteady flow using meteorological and tidal forcing | Outputs: inundation depth, flow, and other hydrodynamic characteristics, including turbulence quantities | *External*<br><br>delft3d-wave or HIS-SSM to view results | Risk assessment difficult to use, require extensive data inputs |
| **Hazus** (U.S.) | Open access<br><br>Requires: ArcGIS | *Yes.* (U.S. version) *No.* (Canadian version)<br><br>Combination of hydraulic & hydrologic modelling confined to floodplain | Inputs: included in program files or link to web to download DEM | *ArcGIS*<br><br>Depth-damage curves, a lognormal pdf vs. inundation depth is used. Social losses calculated via simple function | Training available, well documented manuals. Designed for U.S. |
| **HEC-RAS** (U.S.) | Open Access<br><br>Fortran<br><br>HEC-2 | *Yes.* 1D steady flow, 1, 2D unsteady flow, sediment transport/ mobile bed, water temperature/ water quality | Outputs: depth or duration grids, plots, rating curves, hydrographs, animations | *External*<br><br>Flood Impact Analysis (HEC-FIA) module | Many tutorials, well documented manuals, only support for USACE |
| **InaSAFE** (Indonesia) | Open Source<br><br>Python<br><br>Requires: QGIS | *No.*<br><br>Hazard computed outside program. | User input of exposure from shapefile or OpenStreetMaps | *QGIS*<br><br>Simple vulnerability functions to calculate output | Well documented, simple GUI |
| **Kalypso** (Germany) | Open Source<br><br>Java | *Yes.* Multi-module, hydrology, water level analysis, 1D/2D, flood risk determination | Outputs: hazard inundation maps (2 – 100years) | *External module*<br><br>Risk computed as damage function vs inundation depth or duration/frequency | Wiki-style system and manual |
| **MIKE11** (Denmark) | Licensed.<br><br>AcrGIS for damage estimate | *Yes.* 1D & 2D. Simulate: flow and water level, quality, sediment transport | Outputs include: water level and discharge, comparison maps, plan graphics, animated results | *ArcGIS*<br><br>Toolbox available for flood damage assessment integrates with ArcGIS | Well documented, large number of features |
| **RiskScape** (New Zealand) | Open Access<br><br>Java | *Yes.* Computes flood hazard, easily import and analyze historic flood | Outputs: velocity, inundation depth, ponding and inundation duration | *Integrated*<br><br>Empirical method to create loss | Easy to use, well designed GUI. Software well documented. Designed for NZ |
| **Sobek 1D/2D** (Netherlands) | Open Source<br><br>C++ | *Yes.* 1d/2d hydrodynamic modeling by solving flow equations on 1D network system & 2D horizontal grids | Outputs: Flow velocity and inundation depth at different times | *External module*<br><br>Connect to other (e.g. HIS-SSM for casualties & damage) | Simple to run and install Computationally efficient |

While there are a number of presented applications in Table 1.1 they are designed for use, primarily, in the country of origin and may not easily translate to other regions. Many of the presented solutions require data inputs to solve (some version of) the Saint Venant equations which exceed those available for many Canadian watersheds. Most solutions require interfacing with $3^{rd}$ party software or external module thus another software package is necessary to install, learn and maintain. No single solution offers full capabilities for near-real-time assessment of flood risk within a single easy to use program for easy adoption and use in Canada which may be run by non-expert users.

## 1.3 Research Topic

The primary goal of this research was to help end users from the public safety community run their own flood scenarios and prepare informed emergency response and long term mitigation plans. To this end, different methods were developed and programmed into user friendly tools which allow communities the ability to simulate their own flood hazard scenarios and assess respective risk in support of planning of emergency response and long term mitigation activities. A primary aim was leveraging the abundance of public domain and open source data as primary inputs for these tools.

## 1.4 Problem Statement

Timely and accurate prediction of flood inundation extent and potential negative impacts and consequences is fundamental for the sustainable development of a given region and provides valuable information necessary for understanding respective exposure and vulnerability. Existing solutions require considerable processing time,

extensive inputs, are not well suited for flood prediction in near real-time and/or user-friendly, and often exceed the data available for any given community.

## 1.5 Research Objectives

The objective and original contribution of this research fills in the current gap between rapid, user-friendly tools and sophisticated tools, designed for scientists and engineers, in regards to flood risk assessment. The specific research objectives are as follows:

- Determine the sensitivity of input data and/or parameters to the results of flood risk assessment

- Implement a method of simply calculating loss, through leveraging existing building inventory data, depth-damage curves and published replacement costs

- Develop a new method of DEM Fusion which requires no user input, to produce a better quality DEM from multiple REST API elevation data services

- Develop an application capable of computing flood hazard, 0D and hybrid 1D/2D, requiring minimal user input, which can be run by a non-expert user and provides results accurate enough for mitigation planning and emergency response

The developed standardized methods, based on these specific objectives, is comprised of rapid and easy to use tools based on limited user input, hydrologic principles and processes and accepted risk computation methods which leverage open source datasets. These modular tools form the basis of a flood risk assessment framework which will allow access to otherwise complex flood hazard scenarios and in-depth

knowledge of the community exposure and vulnerability to flood events to the non-expert public safety community.

### 1.5.1    Future Combined Application

By combining the developed tools into a single application, such as a web map, a user with limited expertise can simulate and visualize flood risk based on input via a series of intuitive drop-down menus. The proposed future web application illustrated in Figure 1.5, will first access the DEM fusion tool which extracts elevation data from multiple REST API services and fuses them together into a single, better quality DEM. Next, a flood inundation map is computed based on the user supplied details for a simple flood (0D) or rain event (hybrid 1D/2D). Finally, the potential consequences from the simulated event is computed and visualized by combining the embedded inventory data, the computed flood grid depth and the $ER^2$ API, with the resulting data themed and overlaid on the map.



**Figure 1.5** Example of potential future combined framework web mapping application

## 1.6 Data and metrics

The data resources which were accessed to test the specific research objectives include both open source data and public domain datasets, Table 1.2. Public domain data are data which are freely available to use without restrictions, sharing or modification limitations, while open data refers to those datasets which may include restrictions on use, licenses and/or copyright.

The study area for all research was Fredericton, NB and/or Bathurst, NB. The flood scenarios referenced are the historic 2008 and 2005 floods in Fredericton and a potential 4.5 m sea level rise in Bathurst, NB.

**Table 1.2** Summary of Datasets

| Data Type | Source | Location | Metric | Chapter |
|---|---|---|---|---|
| Inventory (building type, occupancy, etc) Damage parameters Flood Grid | Hazus Canada<br><br>Hazus (U.S.)<br><br>Government of New Brunswick, Dept. Environment and Local Government | Fredericton | Range and $\sigma^2$ of full and depreciated economic losses and Influence Factor in comparison to *base case* | **Chapter 2** McGrath, H., Stefanakis, E., Nastev, M. [2015] |
| Depth-Damage curves | Hazus Canada | Fredericton | Statistical comparison to Hazus Canada results | **Chapter 3** McGrath, H., Stefanakis, E., Nastev, M. [2016] |
| Elevation Data (REST API Elevation Services)<br><br>LiDAR | Canadian Digital Elevation Model (CDEM), Google, and Bing LiDAR, City of Fredericton, City of Bathurst | Fredericton, Bathurst, NB | Statistical comparison to LiDAR ($\sigma,\mu$, $R^2$), FIT Measure | **Chapter 4** McGrath, H., Stefanakis, E., Nastev, M. [2016] |
| Elevation Data Hydrometric Data National Hydro Network (NHN) | (as above) Government of Canada, Water Office GeoGratis, Natural Resources Canada | Fredericton | FIT Measure, profile comparison | **Chapter 5** **McGrath, H**., Stefanakis, E., Nastev, M. [2017] |

## 1.7 Chapter Summaries

In Chapter 1, the motivation, objectives and structure of the research have been presented. Additionally, background information pertaining to how flood risk analysis is being computed by existing solutions have been introduced and their limitations described.

Chapter 2 examines the sensitivity of parameters and data inputs to a flood risk analysis scenario. A baseline solution was generated by running Hazus Canada using the default values and the 2008 flood in Fredericton, NB. These default values were selected by the software provider on the presumption of regional similarities between Canadian provinces and U.S. states south of the border. To illustrate the sensitivity that can be associated to the selection of depth-damage function, flood level, and restoration duration and to identify their relative impacts on the resulting losses, the respective values were modified and the analysis re-run. Each modified parameter was isolated in the re-analysis to determine the relative impact on estimates of exposure and risk.

Chapter 3 presents the developed Rapid Risk Evaluation ($ER^2$) which runs loss assessment analyses in a MS Excel spreadsheet. User input is limited to a handful of intuitive drop-down menus utilized to describe the building type, age, occupancy and the expected water level. The application computes exposure and estimated economic losses related to the structure and the content of the building(s).

DEMs are the primary input to flood inundation mapping. Chapter 4 tests the accuracy of open source REST Elevation API services to evaluate their accuracy and a novel fusion technique is developed which combines multiple DEMs to generate a better quality elevation dataset. The proposed fusion technique incorporates concepts of

clustering and inverse distance weighting (IDW). Three Canada wide elevation sources are included: Canadian Digital Elevation Model (CDEM), Google and Bing.

Chapter 5 illustrates how public domain data can be leveraged and combined with physically based flood inundation models to compute flood hazard and how to optimize computations and generate flood model outputs in near-real-time by users with limited knowledge. Two flood models are explored: 0D bathtub model and a hybrid 1D/2D raster cell storage approach. In the former, user input is limited to a number of points and an associated water depth or data may be extracted from nearby river gauges. In the hybrid model user input includes four categories: (i) geographic location, (ii) rain event, (iii) local conditions, and (iv) average water depth which accept user input from a series of pre-populated drop-down menus.

Chapter 6 presents the conclusions of this research. Included in this chapter are details pertaining to the assumptions and limitations of the proposed framework as well as suggestions for future research and exploration.

By combining the tools developed in Chapter 3 through 5 a standardized method for flood risk assessment has been developed which minimizes the gap between what is needed (and when) by decision makers and the sophisticated tools designed for scientists and engineers currently available for flood risk assessment. Thus these tools help end users from the public safety community run their own flood scenarios and prepare informed emergency response and long term mitigation plans

# REFERENCES

Al-Sabhan, W., Mulligan, M., and Blackburn, G. A. (2003). "A real-time hydrological model for flood prediction using GIS and the WWW". *Computers, Environment and Urban Systems*, 27(1), 9-32.

Apel, H., Aronica, G. T., Kreibich, H., and Thieken, A. H. (2009). "Flood risk analyses— how detailed do we need to be?" *Natural Hazards*, 49(1), 79-98.

Bates, P. D., and De Roo, A. P. J. (2000). "A simple raster-based model for flood inundation simulation." *Journal of Hydrology*, 236(1), 54-77.

Cook, A., and Merwade, V. (2009). "Effect of topographic data, geometric configuration and modeling approach on flood inundation mapping." *Journal of Hydrology*, 377(1), 131-142.

de Moel, H., and Aerts, J. C. J. H. (2011). "Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates." *Natural Hazards*, 58(1), 407-425.

Environment Canada. (2013). Canada's top ten weather stories for 2013. Retrieved from http://www.ec.gc.ca/meteo-weather/default.asp?lang=En&n=5BA5EAFC-1&offset=2&toc=show

FEMA. (2010). HAZUS®MH MR3 Flood Model User Manual. http://www.fema.gov/media-library/

FEMA. (2012). Hazus 2.1 Flood Technical Manual. (Technical Manual). USACE.

Haestad Methods, I. (2003). Floodplain modeling using HEC-RAS. Waterbury, CT: Haestad Press.

Hazus Canada. (2014). Inventory: Exposure by occupancy, natural resources canada, Hazus Canada metadata, hzExposure OccupB_md.rtf . Unpublished manuscript.

Hunter, N. M., Bates, P. D., Horritt, M. S., and Wilson, M. D. (2007). Simple spatially-distributed models for predicting flood inundation: a review. Geomorphology, 90(3), 208-225.

Jongman, B., Kreibich, H., Apel, H., Barredo, J. I., Bates, P. D., Feyen, L., and Ward, P. J. (2012). "Comparative flood damage model assessment: towards a European approach." *Natural Hazards and Earth System Sciences*, 12(12), 3733.

Koks, E. E., Jongman, B., Husby, T. G., and Botzen, W. J. (2015). "Combining hazard, exposure and social vulnerability to provide lessons for flood risk management." *Environmental Science & Policy*, 47, 42-52.

Kulkarni, A. T., Mohanty, J., Eldho, T. I., Rao, E. P., and Mohan, B. K. (2014). "A web GIS based integrated flood assessment modeling tool for coastal urban watersheds." *Computers & Geosciences*, 64, 7-14.

Leskens, J. G., Brugnach, M., Hoekstra, A. Y., and Schuurmans, W. (2014). "Why are decisions in flood disaster management so poorly supported by information from flood models?" *Environmental Modelling & Software*, 53, 53-61.

McMillan, H. K., and Brasington, J. (2007). "Reduced complexity strategies for modelling urban floodplain inundation." *Geomorphology*, 90(3), 226-243.

Merz, B., Kreibich, H., Thieken, A., and Schmidtke, R. (2004). "Estimation uncertainty of direct monetary flood damage to buildings." *Natural Hazards and Earth System Science*, 4(1), 153-163.

Messner, F., and Meyer, V. (2006). "Flood damage, vulnerability and risk perception– challenges for flood damage research." *Flood Risk Management: Hazards, Vulnerability and Mitigation Measures* (pp. 149-167). Springer Netherlands.

Nastev, M., and Todorov, N. (2013). Hazus: "A standardized methodology for flood risk assessment in Canada." *Canadian Water Resources Journal*, 38(3), 223-231.

Nastev, M., Nollet, M. J., Abo El Ezz, A., Smirnoff, A., Ploeger, S. K., McGrath, H., and Parent, M. (2015). "Methods and tools for natural hazard risk analysis in eastern Canada: Using knowledge to understand vulnerability and implement mitigation measures." *Natural Hazards Review*, B4015002.

Neighbors, C. J., Cochran, E. S., Caras, Y., and Noriega, G. R. (2012). "Sensitivity analysis of FEMA HAZUS earthquake model: case study from King County, Washington." *Natural Hazards Review*, 14(2), 134-146.

Pechlivanidis, I. G., Jackson, B. M., McIntyre, N. R., and Wheater, H. S. (2011). "Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications." *Global NEST Journal*, 13(3), 193-214.

Plazak, D. (1984). "A Critical Assessment of Methodologies for Estimating Urban Flood Damages-Prevented Benefits." Colorado Water Resources Research Institute, Colorado State University.

Poulin, J., Chokmani, K., Tanguy, M., & Bernier, M. (2012). "Cartographie dynamique du risque d'inondations en milieu urbain (No. R1428)." INRS, Centre Eau, Terre et Environnement.

Prettenthaler, F., Amrusch, P., and Habsburg-Lothringen, C. (2010). "Estimation of an absolute flood damage curve based on an Austrian case study under a dam breach scenario." *Natural Hazards and Earth System Sciences*, 10(4), 881.

Public Safety Canada, (2014). Canada's National Disaster Mitigation Strategy. Retrieved 3/12, 2014, from http://www .publicsafety.gc.ca/cnt/rsrcs/pblctns/mtgtn-strtgy/index-eng.aspx

Scawthorn, C., Blais, N., Seligson, H., Tate, E., Mifflin, E., Thomas, W. and Jones, C. (2006). "HAZUS-MH flood loss estimation methodology. I: Overview and flood hazard characterization." *Natural Hazards Review*, 7(2), 60-71.

Scawthorn, C., Flores, P., Blais, N., Seligson, H., Tate, E., Chang, S. and Lawrence, M. (2006). "HAZUS-MH flood loss estimation methodology. II. Damage and loss assessment." *Natural Hazards Review*, 7(2), 72-81.

U.S. Army Corps of Engineers. (2010). HEC-RAS river analysis system application guide (Version 4.1 ed.). Davis, CA: U.S. Army Corps of Engineers (USACE).

World Bank. (2014). Understanding risk : Review of open source and open access software packages available to quantify risk from natural hazards. Washington, DC: World Bank Group.

UNISDR (2011): Global Assessment Report on Disaster Risk Reduction – Revealing risk, redefining development, United Nations, Geneva

## 2. Sensitivity Analysis of Flood Damage Estimates: A Case Study in Fredericton, New Brunswick[1]

**Abstract**

Recently, the U.S. FEMA's standardized best-practice methodology Hazus for estimating potential losses from common natural hazards, including earthquakes, flood, and hurricanes has been adopted for use in Canada. Flood loss estimation relies on the combination of three components: flood level, inventory of the built environment, and pre-selected vulnerability parameters such as depth-damage functions, all of which have large associated uncertainties. Some of these parameters, such as occupancy schemes and vulnerabilities, have been carried over from the U.S. version on the presumption of regional similarities between Canadian provinces and states south of the border. Many of the uncertainties can be reduced by acquiring additional data or by improving the understanding of the physical processes. This paper presents results from a series of flood risk analyses to illustrate the sensitivity that can be associated to the depth-damage function, flood level, and restoration duration and to identify their relative impacts on the resulting losses. The city of Fredericton is chosen as the test case as it was subjected in 2008 to flood water levels breaching 1.86 m above flood stage resulting in more than 680 residents evacuated from their homes, and economic costs of more than $23 million. The loss results are expressed by the number of flooded residential buildings which varied between 579 and 623 and the range of replacement cost is $21 million. These results

---

highlight the importance of proper selection of input parameters customized to the study area under consideration.

## 2.1. Introduction

Every year disastrous climatological and geological hazards take place in Canada and across the globe [Nastev and Todorov, 2013]. Of these natural disasters, flooding of river systems is the most frequent and costly natural disaster, affecting the majority of the worlds' countries on a regular basis, and accounts for approximately one-third of total natural disasters related economic losses in Europe [de Moel and Aerts 2011; Jongman et al., 2012]. The costliest natural disaster in Canadian history, the southern Alberta flood in June of 2013, exceeds $6 billion Canadian dollars [Environment Canada, 2013]. In recent decades, the trend of increased damages resulting from flood events may be attributed to a number of factors including: population growth, increased urbanization in flood-prone areas and the changing climate [Jongman et al., 2012; de Moel and Aerts, 2011 Aerts et al., 2015; UNISDR, 2011].

Government officials, GIS specialists, emergency managers, and first responders look for tools to develop mitigation and recovery plans as well as preparedness and response procedures in anticipation of these natural disasters [Neighbors et al., 2013]. Timely and accurate prediction of potential losses is fundamental for the sustainable development of a given region and provides valuable information necessary for understanding of risks and creation and implementation of mitigation measures and post-disaster emergency planning [Scawthorn et al., 2006]. Through the use of computer models which simulate hazards and compute risk we can evaluate the cost effectiveness

of mitigation measures, optimize investments, and enable insurance companies, municipalities and residents to prepare for disasters [Apel et al., 2009; de Moel and Aerts, 2011].

Flood risk analysis involves the combination of three components: a probabilistic or deterministic flood hazard model, an inventory model of the built environment defining the characteristics of the exposed elements (structural type, occupancy category, content), and a selection of respective depth-damage functions [Apel et al., 2009; de Moel and Aerts 2011; Merz and Thieken, 2004]. Loss estimations include physical damage and direct and indirect social and economic losses. A direct loss occurs as a result of direct physical contact of the flood water with humans, property, or other objects, while indirect losses represent those that are induced by the direct impact, and may occur (in time or space) outside of the flood event [Merz et al., 2004]. Physical damage to buildings and certain transportation and essential facilities is estimated based on depth-damage functions which represent the relationship between inundation depth and percent damage [Plazak, 1984 and others]. For buildings, depth-damage functions are developed for structural or load-bearing components; for contents (e.g.: interior furniture, art, appliances, etc.,); and for inventory (e.g.: commercial stock and inventory) [FEMA, 2009]. These three types of damage functions are unique for a given building structural type and occupancy classification (e.g.: residential, commercial, industrial, etc.) The reason behind this is that the underlying structure, for example a single family residence has a different damage response to a given water level than would a multi-family apartment complex. Direct economic losses include calculations of repair and construction costs resulting from the flood event, whereas indirect economic losses are

27

related to lost jobs and business interruption [FEMA, 2009]. The analysis may also include estimates of volume of debris and removal costs. Social impact of the flood event is estimated based on population demographics, flood extent and inundation depths, and is usually expressed by the number of displaced households or people which may require shelter, time needed for re-building (or restoration), recovery needs, etc. Risk analysis can be run on aggregated data, e.g. at the census block level, where the percentage of each census block is determined for a given water level. For more accurate analyses, one can perform a micro-scale analysis where individual structures are introduced with proper parameters and physical damage and direct economic loss estimations are derived on a per structure basis.

Regardless of the applied method, one of the most important aspects of constructing a flood loss model is to identify, quantify, and incorporate uncertainties owing to approximations of the input parameters and simplifications in simulating the physical processes [Merz and Thieken, 2004; Neighbors et al., 2013; FEMA, 2009]. These uncertainties may be linked to the hazard model used (from simple interpolation to sophisticated equations solving the shallow water equations), the choice of vulnerability models and parameters, scale of the study region (micro, meso, or macro), inventory data, or any combination of these [Apel et al., 2009]. In addition, uncertainties propagate through the calculation and accumulate in the resultant damage estimate [de Moel and Aerts, 2011]. Studies acknowledge that flood damage estimates feature a degree of uncertainty, with most efforts focusing on the influence of the hydrological component [de Moel and Aerts, 2011]. Examples of such research include: Dutch FLORIS study using different inundation scenarios, flood frequency statistics and levee breach scenarios

[Apel et al., 2006], boundary effects [Hall et al., 2005], and 1D and 2D numerical models [Horritt and Bates, 2002]. Beyond the hydrologic component, Merz et al., [2004] presented research which found considerable uncertainty in the internationally accepted damage functions – which describe the relationship between the inundation level and damage. Adjusting the value of elements at risk as performed by Egorova et al., [2008] has also shown to affect the loss estimates from a given flood scenario and de Moel and Aerts [2011] computed the influence of four components (inundation depth, land use, value of elements at risk, and depth-damage curves) on the outcome of flood risk analysis.

In this study epistemic uncertainties resulting from incomplete knowledge are considered as they can potentially be reduced by acquiring additional customized data representative of the study region under consideration. The well-known U.S. FEMA's Hazus software, recently adapted for use in Canada, is used to conduct this sensitivity analysis. Hazus is one of the most comprehensive and standardized methodologies presently available for the assessment of potential losses from natural hazards [FEMA, 2009; Neighbors et al., 2013; Nastev 2014].

A number of parameters required for loss estimation including damage functions (e.g.: building, contents, and inventory), restoration functions, and economic replacement values provided with the Canadian version of Hazus are based on U.S. data. Thus the default damage functions suggested to Canadian users were derived from data collected and analyzed by the Federal Insurance Administration (FIA) and surveys completed by the U.S. Army Corps of Engineers (USACE) on U.S. infrastructure. The damage

functions have been regionally adopted into the Canadian model and replacement costs per square foot have been adapted from R.S. Means Co., Inc. [Hazus Canada, 2014].

The parameters reviewed and varied in this sensitivity analysis include (i) structural and contents depth-damage functions for single family residences, RES1, (ii) changes to flood depth, and (iii) changes to restoration duration. The Hazus model was first run using the suggested default values for the considered flood scenario. Additional scenarios were completed with the parameters varied in the respective anticipated ranges to create a range of possible outcomes. Each analysis parameter was isolated so that the influence of each could be determined independently. Parametric analysis was conducted to determine the sensitivity of the final results to each parameter.

The remainder of the paper is structured as follows, in Section 2.3; the study area of Fredericton, New Brunswick (NB) is introduced. In Section 2.4 the methodology of Hazus is briefly described. In Section 2.5 the methodology and parameterization tests are outlined and results are presented. Section 2.6 contains results and conclusions and recommendations for further research are outlined in Section 2.7.

## 2.2. Study Region

The study area selected in this sensitivity analysis is Fredericton, New Brunswick, Canada. Fredericton is located in the west-central portion of this Atlantic province and is bisected by the St. John River (Figure 2.1), a major waterway which runs throughout the province. Its watershed drains an area of approximately 55,000km$^2$, and encompasses much of New Brunswick and parts of Quebec, Canada, and Maine, U.S.A. [McGrath et al., 2014].

Cardy [1976], in a report for the Saint John River Basin Board reviewed flood records and spending in New Brunswick between 1887 and 1971 and found total damages in excess of one million dollars. Between 1971 and 1976 (when their report was published) they report an additional 17 million dollars spent in the province on recovery from flood related damages. A comprehensive database of flood events dating back to the 1600s is available on the Government of New Brunswick web site (www.elgegl.gnb.ca/0001/en/Home/Main). The largest of these are shown in Table 2.1. The second largest flood, used in this sensitivity study, occurred in 2008, with water levels 1.86 m over flood stage. Estimated expenditures across the province for the 2008 flood exceed $23million dollars.

Fredericton was chosen as the study location due to its long flood history, the mix of government and private infrastructure, and the open data policy. Fredericton is the capital of New Brunswick, and, as a result, there are a significant number of government offices and service locations across the city. The community of Fredericton is the third largest in the province with a population of 94,000 [Statistics Canada, 2011] and approximately 22,000 households (Chang et al., 2010). A mix of public (municipal, governmental) and private infrastructure is therefore potentially at risk of flood hazard. The downtown core of Fredericton (along the southern shore of the Saint John River) contains a number of historic buildings, with those in the eastern section of downtown having been built in the late 1700s. In 2011 the City of Fredericton announced an open data policy. The data collected by the city is available to the public via the City of Fredericton (fredericton.ca/en/open-data). This open data policy facilitates locating

31

appropriate datasets for Hazus inventory, including essential facilities, transportation, and

utility networks.
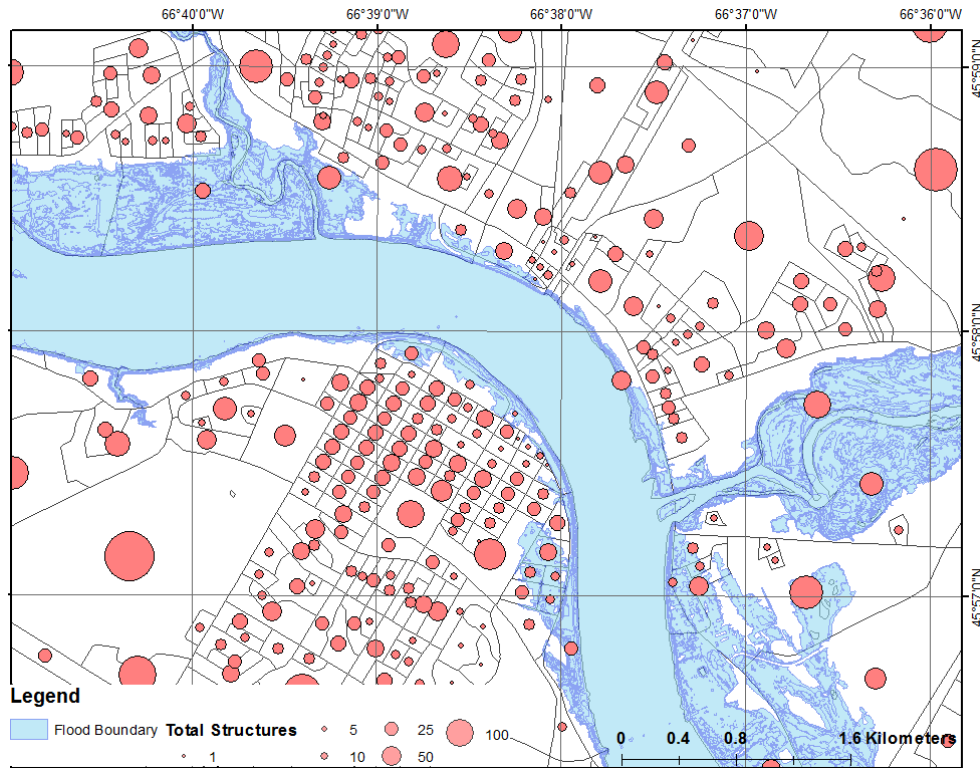


**Figure 2.1** City of Fredericton study area, bisected by the Saint John River. Displaying total structures (per census block) and 2008 flood boundary

**Table 2.1** Flood history database records for Fredericton. Cause key (R = heavy rain, IJ = Ice jam, W = mild weather, T = high tides, SM = snowmelt, F= freshet, SF = snowfall, U = unknown) Damage amount in thousand dollars spent across all communities affected. Source: Government of New Brunswick [2014]

| Date | River level (masl)* | Damage claims | Cause |
|---|---|---|---|
| April 30, 1973 | 8.608 m | $11,877 | R, SM |
| April 30, 1979 | 8.062 m | $2,113 | R,SM |
| April 14, 1993 | 6.60 m | $12,738 | R, F, IJ |
| April 18, 1994 | 7.87 m | $4,130 | IJ, R |
| April 30, 2005 | 7.893 m | $5,600 | IJ, R, W |
| May 1, 2008 | 8.36 m | $23,288 | R,F,SM,W, SF |
| Dec 13, 2010 | 7.73 m | $13,830 | R |
| **\*Flood Stage is 6.50 metres above sea level (masl)** | | | |

The study region created in Hazus contains 617 census dissemination blocks, over 27,000 households and 66,050 residents. Essential facilities in the area include: two hospitals, 30 schools, 4 fire stations and 3 police stations. There are 19,178 buildings within the study area with an estimated total building replacement value (excluding contents) of $5,400 million dollars. Of the buildings in the study area, approximately 90.3% (and 68% of building value) are residential housing.

Within the residential occupancy classification, there are 11 sub-classes: single family residences (RES1), manufactured homes (RES2), apartment buildings (RES3A thru 3F), temporary lodging (RES4), institutional dormitory (RES5) and nursing homes (RES6). To determine if one occupancy classification was more prominent than others in the flood plain, the building stock inventory was reviewed. The distribution of occupancy classifications within the Fredericton project study, based on the nationally supplied inventory layer is shown in Table 2.2. Within the study area over 75% of the residential structures are classified as single family homes, RES1. This fact narrows down the number of potentially representative damage curves which will be varied to examine the sensitivity of the loss estimates. Note that there are over 900 damage curves pre-defined in Hazus [Scawthorn et al., 2006].

**Table 2.2** General building stock count of Residential structures at risk of flood hazard (within flooded census blocks), by specific occupancy classification in the study area

|  | RES 1 | RES 2 | RES 3A | RES 3B | RES 3C | RES 3D | RES 3E | RES 3F | RES 4 | RES 5 | RES 6 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Building count** | 3193 | 561 | 299 | 7 | 8 | 60 | 8 | 0 | 16 | 0 | 2 |

## 2.3. Flood Damage Estimation

The most common and internationally accepted method of estimation of urban flood damage is through the use of depth-damage functions [Plazak, 1984; Smith 1994]. Structure and contents damage resulting from flood hazard are influenced by many factors, however, usually only building use and inundation depth are considered as damage-causing factors and included in the formulation of depth-damage functions [Merz, 2004]. The building age, foundation type, and elevation of the first floor can be included as factors which contribute to the estimated damage of a structure, which are external to the depth-damage functions [FEMA, 2009].

### 2.3.1  Flood Hazard

The flood hazard is based on a depth grid, an ESRI grid file which contains the flooding extent and the water depth. The flood hazard file may be supplied by local government or calculated using, for example HEC-RAS [FEMA, 2009], or CADYRi [McGrath et al., 2014] software.

In this study the flood depth grid was acquired from the New Brunswick Department of Environment. Digital elevation data for the study area was downloaded from GeoBase (http://www.geobase.ca/) and U.S. Geologic Survey Earth Resources Observation and Science (EROS) data Center (http://eros.usgs.gov/find-data).

### 2.3.2  Inventory

The next input parameter is the inventory of assets as risk, such as population and infrastructure. To facilitate this, a number of Hazus datasets come pre-populated. For this study, aggregated building data and population demographics were used. The inventory of the residential buildings and demographics were derived from the 2011 Census data

[Statistics Canada, 2011], while the commercial and industrial structure data was acquired from Dun and Bradstreet [Hazus Canada, 2014].

From the results of the first analysis in Hazus the predominant building type found at risk in this study area is "Wood", followed distantly by "Manufactured Housing" (Table 2.3). There are a total 375 wood structures which are expected to incur damage as a result of the 2008 flood scenario, and 36 manufactured housing buildings.

**Table 2.3** Expected building damage by building type in study area, based on flood hazard 1.86 m above flood stage

| Building Type | Expected Damage | | | | |
|---|---|---|---|---|---|
| | *11-20%* | *21-30%* | *31-40%* | *41-50%* | *Substantial (>50%)* |
| **Concrete** | 0% | 0% | 0% | 0% | 0% |
| **Manufactured Housing** | 0% | 0% | 0% | 0% | 100% |
| **Masonry** | 20% | 0% | 0% | 0% | 80% |
| **Steel** | 100% | 0% | 0% | 0% | 0% |
| **Wood** | 2.13% | 3.2% | 15.20% | 18.93% | 60.53% |

### 2.3.3   Damage Functions

Depth-damage curves represent a fundamental concept regarding assessment of damage resulting from a flood and are internationally accepted as the standard approach to accessing urban flood damage [Plazak, 1984; Prettenthaler et al., 2010]. Depth-damage functions, at a minimum, require two inputs, namely the occupancy classification of the structure and the depth of flooding. The output is an estimate of the damage, expressed generally as a percentage of the replacement cost.

The suite of damage functions which come with Hazus are termed 'credibly weighted' and considered regionally applicable [Scawthorn et al., 2006]. They are based on the best available damage data and represent more than 20 years of losses. For each specific occupancy classification, default damage functions have been regionally selected

for structures, contents and inventory. Due to the similar construction practices, these standardized building types were judged more than sufficient to reflect the characteristics of the building stock in Fredericton. For this study, we have not (presently) derived our own damage functions. Contents loss valuation is calculated as a percentage of the full replacement value per the contents replacement cost ratios [FEMA, 2009]. For residential properties the contents value is typically estimated at 50% of the structure replacement value, while commercial, industrial, and other general occupancy classes range from 50 to 150% of the replacement values.

## 2.4. Methodology

The loss estimation analysis was first run using the default parameters for the Fredericton study region (base case) based on the flood level resulting from the 2008 flood event, where water levels rose 1.86m above flood stage.

The full replacement cost direct economic results for the base case are summarized in Table 2.4. The total estimated losses amount to over $170 million. The largest contributors to this estimate are the building and contents losses with estimates at 52.42% and 46.10% respectively. The occupancy classification with the highest impact on the structural losses is residential (~76%) followed by commercial occupancy (~18%). Given the distribution of results as shown in Table 2.4, the damage curves for residential occupancy were selected as critical parameters and varied in the subsequent loss estimations.

**Table 2.4** Estimated losses based on default settings in Hazus flood model for flood scenario in Fredericton, NB. Results are in thousands of dollars per general occupancy

| | Total | Building | Contents | Inventory | Relocn | Income | Rental Income | Wage Loss | Direct Output |
|---|---|---|---|---|---|---|---|---|---|
| **Education** | $809 | $175 | $625 | $0 | $0 | $2 | $0 | $7 | $74 |
| **Government** | $1,190 | $235 | $912 | $0 | $2 | $1 | $0 | $40 | $8 |
| **Religion** | $3,320 | $785 | $2,513 | $0 | $1 | $6 | $0 | $15 | $99 |
| **Agricultural** | $175 | $50 | $107 | $18 | $0 | $0 | $0 | $0 | $0 |
| **Industry** | $10,660 | $3,337 | $6,254 | $1,068 | $0 | $0 | $0 | $1 | $6 |
| **Commercial** | $48,907 | $16,496 | $31,297 | $735 | $37 | $139 | $23 | $180 | $447 |
| **Residential** | $105,158 | $68,159 | $36,760 | $0 | $110 | $22 | $48 | $59 | $139 |
| **Total** | **$170,219** | **$89,237** | **$78,468** | **$1,821** | **$150** | **$170** | **$71** | **$302** | **$773** |

### 2.4.1   Structure Damage Functions

Two examples of depth-damage functions for single family residences (RES1) are shown in Figures 2.2 and 2.3. The RES1 classification is dissected into sub-curves based on the number of stories of the structure and the presence or absence of a basement. Figure 2.2 represents damage functions for a one-story home with no basement, while Figure 2.3 is for a two-story home with a basement. In both figures the red lines represent the default applied depth-damage function RES1 for the study area; the other lines represent other pre-computed FIA and USACE curves. Negative values indicate water levels affecting only the basement. As expected, for the one-story with no basement, damage remains at 0% until the water depth has reached 0 ft and starts to submerge the first floor. Damages could include items such as flooring, baseboard trim, and drywall, among others. The two-story curves reflect damage to a finished and (possibly) furnished basement.

In total, for the RES1 occupancy classification there are 63 pre-defined depth-damage curves (functions) spread over 8 residence types: one-story, two-story, three-story, and split-level each with and without basements. Of these 63 curves; 24 curves are

defined for one-story structures, 23 for two-story, 3 three-story and 13 split-level depth-damage curves. This allows for a considerable number of potential combinations of damage functions and loss estimates using the pre-defined depth-damage curves. For this study, a sample of n=85 damage curves was chosen to estimate the sensitivity of the loss estimation with respect to depth-damage curves. The damage functions chosen for the sample scenarios represented only those which would be representative of flooding for the Saint John River and the building types found in the study area.
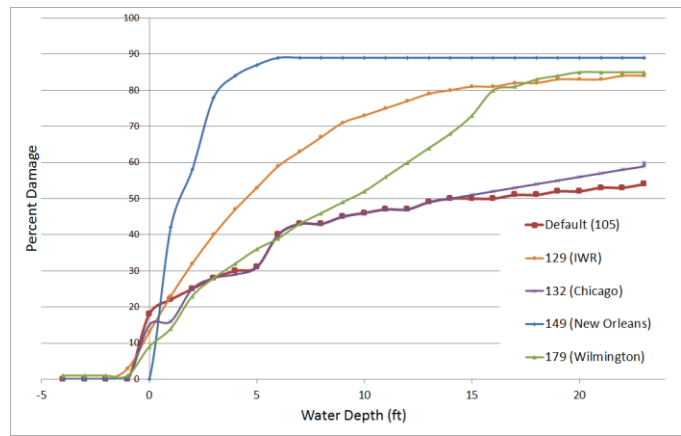


**Figure 2.2** Depth-Damage curves for 1-story structures with no basement
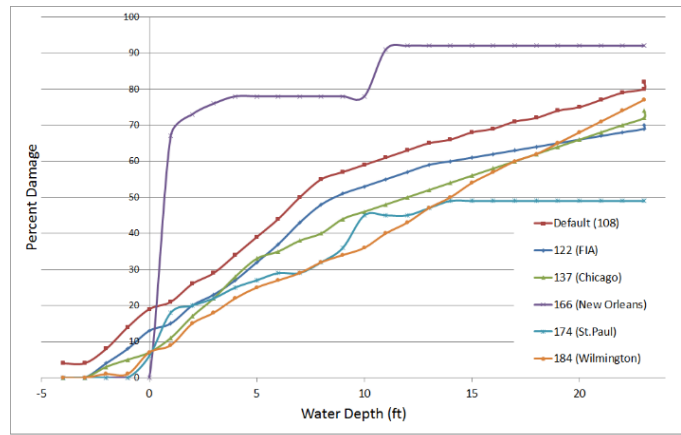


**Figure 2.3** Depth-damage curves for 2-story structures with basement

Functions selected for this sensitivity analysis included those classified from different districts across the U.S (Figures 2.2 and 2.3). The Galveston District damage curves developed using historic flood damage records and are applicable to fresh-water flooding under slow-rise, slow-recession conditions with low velocity on structures without basements [FEMA, 2009]. The Chicago District curves represent generic structure and content functions based on models developed by the Galveston and Baltimore District. New Orleans District functions include those developed based on expert opinion are divided into riverine or rainfall (freshwater) flooding of long or short duration types. The New York District functions include those with and without basements as well as split level structures. The St. Paul District (Minnesota) functions were developed as part of the 1998 flood control project in the Grand Forks area. Two remaining district curves are for Wilmington and the USACE Institute for Water Resources (IWR). The IWR functions are a working project of compiled 'past flood damage surveys' with the objective of becoming the USACE national standard damage functions [FEMA, 2009].

A random sampling of depth-damage curves including both FIA credibly-weighted and USACE District damage functions were selected and analysis was re-run for each of the sample scenarios. Report outputs included general building stock by occupancy (which reports the total square footage per census block of damage), building count (total and damaged), and full and depreciated economic losses. The resulting losses were analyzed in MatLab to determine the range, variance and influence factor.

### 2.4.2 Contents Damage Functions

The contents depth-damage function contains the same 8 sub-classifications of single family residence (RES1) occupancy classes with 43 different pre-defined content depth-damage curves: 15 curves defined for one-story structures, 14 for two-story, 2 for three-story and 12 split-level. A sample of $n = 65$ of content damage curves was selected and the same procedure was repeated as with the depth-damage functions for structural damage. In this case the output from each of these sample analysis were full and depreciated economic losses.

### 2.4.3 Water Depth

As mentioned, the flood depth grid in the base case scenario represents the 2008 flood where water levels were 1.86m above flood stage. To test sensitivity of the loss estimations to water depth, this raster was modified by adding 0.25m (2.11m) and subtracting 0.46m (1.40m). The lower flood level represents the 2005 flood of the Saint John River in Fredericton, while the higher flood level (+0.25m) approximates the largest flood on record in Fredericton which occurred in 1973. The outputs from this analysis include estimation of building loss, building interruption, and debris generated.

### 2.4.4 Restoration Duration

Disruption costs to building owners are considered as restoration time dependent direct economic losses, and include relocation expenses, wage and rental income losses, and capital related income losses. Their calculation includes the following input parameters: restoration time in days, assumed rental costs in dollars per square foot per

day, square footage and water depth per occupancy classification, percent structure

damage, disruption costs in dollars per square foot, and the percent of owner occupied

units in each census block.

The restoration time (measured in days) is computed in four discrete water levels, -

4ft to 0ft, 0ft to 4ft, 4ft to 8ft, and 8ft to 24ft and outputs are given in thousands of

dollars. The number of days to restoration were modified to determine sensitivity of

respective losses based on recovery time. As with previous scenarios, focus was on RES1

buildings. A sample of $n$=4 scenarios were run using the default restoration times (base

case), 30% and 60% longer times, which may represent a remote location with difficulty

getting access to supplies and workers, and a scenario with restoration times 20% shorter

than the base case.

## 2.5. Results

The presented results illustrate the sensitivity associated to the variation of depth-

damage function, flood level, and restoration duration and identify their relative impacts

on the simulated losses.

### 2.5.1   Variation of Depth-Damage Functions

The results based on the variation of the representative depth-damage functions for

both structure and content damage indicate significant variability in the estimated number

of structures affected, total inundated square footage, and replacement costs for both

structures and contents.

**2.5.1.1 Building Count**

In aggregate analyses, losses are calculated based on uniform distribution of

structures throughout the census block and respective damage curves. Using the

percentage of census block inundated, percent damage based on the depth-damage

function, and the numbers of buildings in the census block, an estimate of buildings

affected are calculated for each general occupancy classification. If, for example, 20% of

the census block is considered to have 5ft of water, according to the depth-damage curve

(#105) residential structures sustain ~30% damage, and when damage curve #129 is

applied the damage is ~52% for the same water level. Hence, the estimated number of

flooded structures (building count) changes with respect to the applied depth-damage

function (Figure 2.4). The minimum estimated structures affected from the flood are 579,

while a maximum from the sample scenarios was 623. This range represents the potential

difference in total loss claims submitted by residential households.
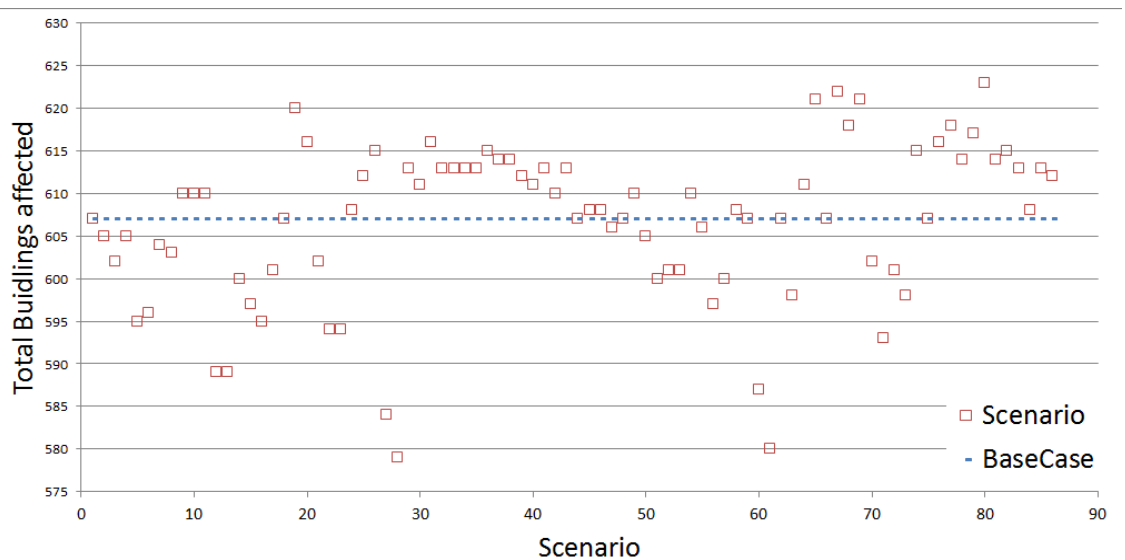


**Figure 2.4** Count of total building affected from the flood per scenario. Solid red line indicates base case with building count of 607

### 2.5.1.2 Damage by Square Foot

Figure 2.5 shows the results of the sample depth-damage scenarios for residential structures based on the percent damage to a structure (damage state). In the $0 - 10\%$ damage range, the variance is less than 5, while the highest variance is found in the 61-70% damage range. In Hazus slight damage is defined within the 1-10% range, moderate damage is between 11-50%, while substantial damage is 51% or greater. Buildings which are categorized as substantially damaged ($>50\%$) represent 33.87% of the base case, with the average of the sample ($n$=85) is very close, 33.27%, however, the distribution within the damage state does vary largely within each category.

### 2.5.1.3 Economic Losses

With respect to full replacement values, the default settings estimate $43,425,000 in structural damages (Figure 2.6). Eighty-one percent of the estimates are found within +/-$5million dollars of the default scenario, while the standard deviation is σ =$4,014,700, Table 2.5. The structural depth-damage function scenarios which estimated losses greater than one sigma include those defined for New Orleans freshwater long flood duration and short duration curves for one and two-story structures, those below one sigma include damage functions as designed for St. Paul District.
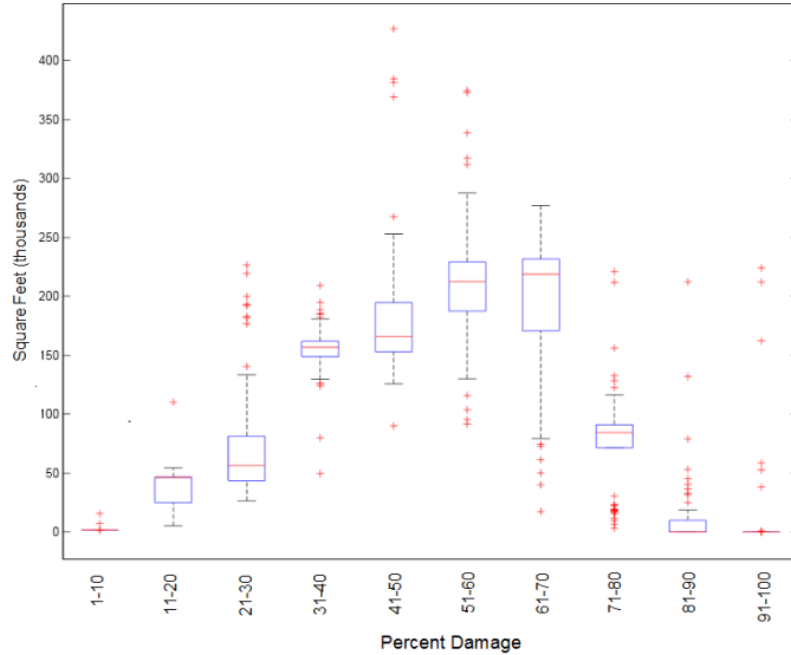
**Figure 2.5** Distribution of simulated RES1 damage by square footage (in thousands) for damage states based on sample of 85 structural damage curves. Boxplots indicate the median, 75th centile, 25th centile, min, and max values
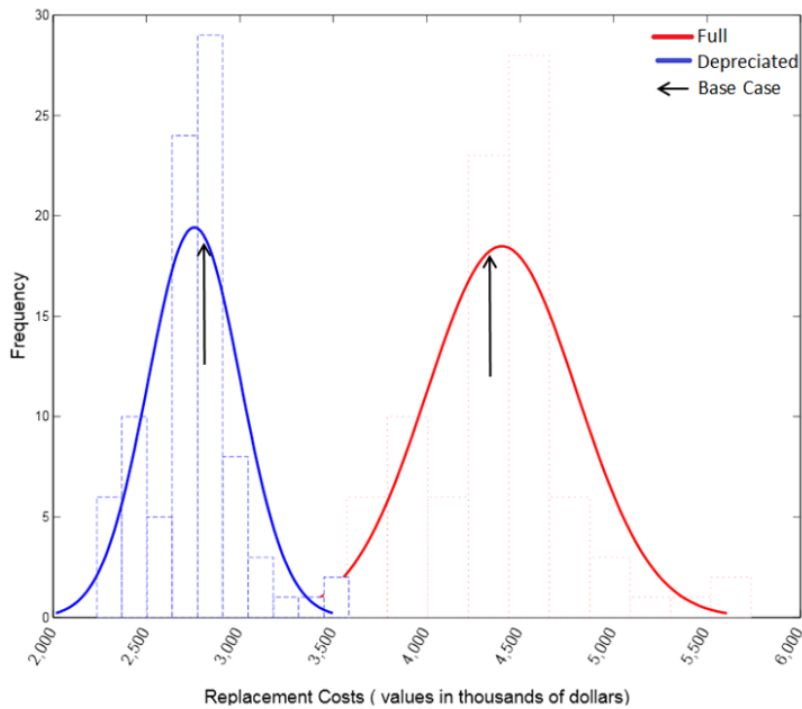


**Figure 2.6** Full and depreciated replacement value (in thousands of dollars) for structures over the sample of 85 depth-damage curve scenarios

**Table 2.5** Economic Losses - full and depreciated losses. (Dollar values in thousands)

| | Full Replacement Value (RES1) | | | | Depreciated Value (RES1) | |
|---|---|---|---|---|---|---|
| | **Building** | **Contents** | **Reloc'n** | **Rental Income** | **Building** | **Contents** |
| *Base case* | *$43,425* | *$21,862* | *$103* | *$18* | *$27,139* | *$13,652* |
| Median | $44,302 | $22,922 | $103 | $18 | $27,729 | $14,328 |
| Standard Deviation | $4,014.7 | $1,864.1 | - | - | $2,471.1 | $1,220.5 |
| Maximum | $57,368 | $30,526 | $103 | $18 | $35,844 | $19,063 |
| **RANGE** | **$21,635** | **$11,463** | **-** | **-** | **$13,518** | **$6,566** |

The contents costs for residential occupancy are considered at 50% of the structural costs. Therefore, if a single-family residence has a valuation of $200,000, the replacement costs for the contents would be $100,000. The damage functions as applied to the contents values display a similar trend to the building structure costs. Contents full replacement values for RES1 base case are $21,862,000 (Figure 2.7). Sixty-eight percent of the scenarios are +/- $1,864,000 of the base case, with a maximum loss estimate ~40% larger. The St. Paul District damage function estimated again the lowest structural replacement costs, however this damage function reported the highest contents replacement cost.

The applied depreciated model is based on industry-standard depreciation method as presented in R.S. Means, Means Square Foot Cost [FEMA, 2010]. The base case scenario estimates $27,139,000 (about 62% of the full replacement costs) in structural replacement costs and $13,652,000 for building contents in the study area. The New Orleans District long and short duration damage functions represent the highest estimates, 15-32% larger for depreciated structural losses, and up to 39% larger contents estimates based on St. Paul District damage functions.
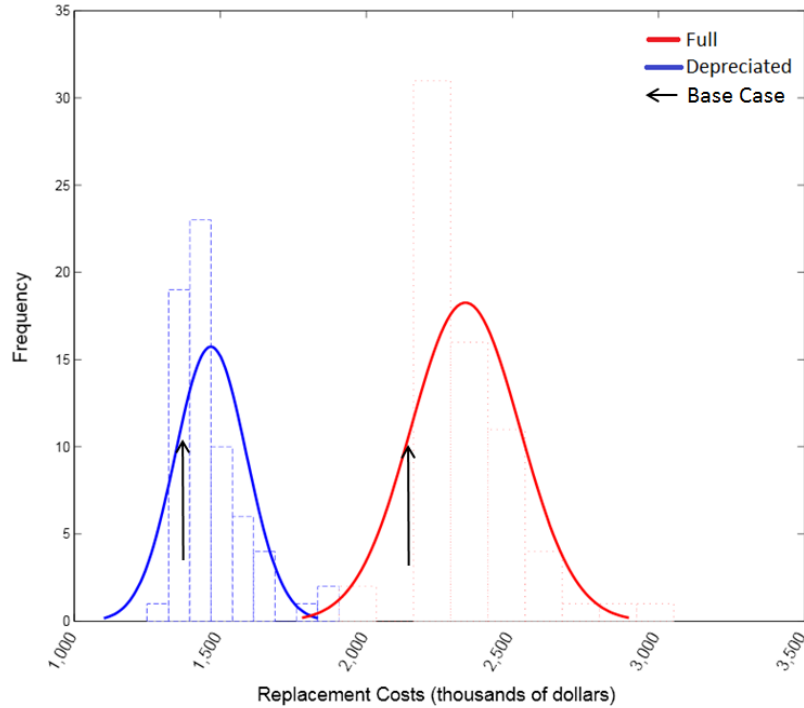
**Figure 2.7** Depreciated replacement value (in thousands of dollars) over the sample of 60 contents depth-damage curve

### 2.5.2 Variation of Water Level

The water level increase was achieved by calculating new cell values by simply adding 0.25m, whereas the scenario of 0.46m below the default value was based on a provincially supplied depth grid. The simulated results indicate that increasing the flood water level, by 0.25m leads to an increase of $93.49 million dollars in total building related loss and business interruption costs across the study area, for a total of $263.73 million dollars (Table 2.6). As the flood boundary was not adjusted with the increased flood level, the loss estimates are likely underestimated. As a result, the estimated shelter needs remain unchanged between this and the base case, i.e., an estimated 1,130 households and 3,390 people are displaced, with 3,096 residents requiring short term shelter.

46

**Table 2.6** Building and Business interruption losses (in millions) over water level scenarios, 1.86m (default), 2.11m (plus 0.25m), and 1.4m (minus 0.46m) above flood stage

| | Base case scenario | | | + 0.25m | | | - 0.46m | | |
|---|---|---|---|---|---|---|---|---|---|
| | Residential | Commercial | Others | Residential | Commercial | Others | Residential | Commercial | Others |
| *Building loss* | | | | | | | | | |
| Structural | 68.16 | 16.5 | 4.59 | 117.37 | 29.03 | 8.13 | 63.10 | 14.36 | 3.8 |
| Content | 36.76 | 31.3 | 10.41 | 53.91 | 38.9 | 13.27 | 34.41 | 28.53 | 9.07 |
| Inventory | 0 | 0.74 | 1.09 | 0 | 0.9 | 1.39 | 0 | 0.62 | 0.9 |
| *Subtotal* | *104.92* | *48.54* | *16.09* | *171.28* | *68.83* | *22.79* | *97.50* | *43.50* | *13.78* |
| *Business loss* | | | | | | | | | |
| Income | 0.02 | 0.14 | 0.01 | 0.03 | 0.17 | 0.01 | 0.02 | 0.13 | 0.01 |
| Relocation | 0.11 | 0.04 | 0 | 0.13 | 0.04 | 0 | 0.09 | 0.03 | 0 |
| Rental Income | 0.05 | 0.02 | 0 | 0.06 | 0.03 | 0 | 0.04 | 0.02 | 0 |
| Wage | 0.06 | 0.18 | 0.06 | 0.07 | 0.21 | 0.08 | 0.05 | 0.17 | 0.05 |
| *Subtotal* | *0.24* | *0.38* | *0.07* | *0.29* | *0.45* | *0.09* | *0.21* | *0.35* | *0.06* |
| **TOTAL** | **105.16** | **48.92** | **16.16** | **171.57** | **69.28** | **22.88** | **97.71** | **43.85** | **13.84** |
| | **170.24** | | | **263.73** | | | **155.4** | | |

The structural and content losses are over 400 times larger than the losses due to business interruption, which considers income, relocation, rental income and wage. In the base case scenario, business interruption amounts to $0.24 million dollars, whereas it is $0.29 million dollars for 25 cm higher water levels (Table 2.6). These values depend much on the applied occupancy scheme, which indicate the percentage of industrial and commercial buildings. At the same time building related losses are estimated at $68.16 million dollars for the base scenario and to $117.37 million dollars when the water levels are raised 25cm. Debris generated from the flood hazard includes finishes (drywall and insulation), structure (wood and brick), as well as foundation (concrete slab, concrete block and rebar). In the base case, it's estimated that 1452 truckloads (at 25 tons per truck) are required to remove the debris. The estimated debris more than doubles, estimating 73,965 tons (2,959 truckloads) if the water level is raised by 25cm.

The second water level considered a flood hazard representing flooding at 1.4 m above flood stage (the 2005 flood event). This scenario puts the water level at 0.46m below the base case.

As expected loss estimations are lower with respect to building and business interruption given the lower flood level. The total estimated economic loss for this scenario is $155.39 million dollars, with residential occupancies comprising 68% of the total loss. Compared to the base case, residential structural and content losses are 92.6%, 93.6% respectively, with a total estimated loss of $97.5 million dollars. Commercial building losses were also lower by 9% to 13%.

Business interruption totals in this scenario indicate less than $10 million dollars from default, with the largest difference being relocation expenses.

The estimated debris generated from this scenario is 30,432 tons or 1,217 truckloads which are ~83% less debris. An estimated 884 households are displaced from this flood event with an estimated 2,322 people requiring temporary shelter. The shelter needs of the 2005 flood scenario are 75=% of those of the base case, and displaced households are 78% of base case.

### 2.5.3   Changes to Restoration Time

The base case suggests 180 days of restoration for RES1 with water between -4ft and 0ft up to nearly 2 years for inundation between 8ft and 24ft. The annualized relocation expenses for the Fredericton study area associated with the base case are $167,000 and rental income losses $74,000 (Table 2.7). Capital related income losses and wage losses in the base case and across all scenarios remained constant, with estimates of

$189,000 and $332,000 respectively. By increasing the estimated restoration days by 30% in each of the water depth classes, there is an estimated ~20% increase in both relocation expenses and rental income losses incurred by residents. A further increase of restoration days to 60% of the base case leads to 48% and 33% increases in relocation expenses and rental income losses respectively. Increases in restoration time continue to increase at 20% or more of the base case in these two scenarios, however the rental income losses increase at a decreasing rate with longer restoration times.

Consideration of more available workers for restoration and increased supply of building materials, for example 20% fewer days than base case, leads to decreases in both relocation expenses and rental income losses. If restoration of a structure with water levels between -4ft and 0ft takes 150 days, instead of 180 in the base case, the relocation expenses are smaller by 15% and 5% for rental income losses.

**Table 2.7** Direct Economic Annualized Income Losses for Buildings

| | Water Depth (in feet) | | | | Income Losses | | | |
|---|---|---|---|---|---|---|---|---|
| | -4 to 0 | 0 to 4 | 4 to 8 | 8 to 24 | Relocation expenses | Capital related income losses | Wage Losses | Rental income losses |
| | *(Days to restoration)* | | | | *(All values are in thousands of dollars)* | | | |
| Less 20% | 144 | 288 | 360 | 576 | $156 | $189 | $332 | $74 |
| *Base Case* | *180* | *360* | *450* | *720* | *$167* | *$189* | *$332* | *$74* |
| Plus 30% | 234 | 468 | 585 | 936 | $203 | $189 | $332 | $91 |
| Plus 60% | 288 | 576 | 720 | 1152 | $248 | $189 | $332 | $99 |

### 2.5.4   Influence Factors

To understand the individual impacts on the results, parameters were varied against a single flood hazard scenario, and multiple flood scenarios are tested using a single set of parameters. In the first set of tests, depth-damage functions were modified to determine their relative influence on the resulting loss estimates in terms of buildings affected and economic impact. The influence factor is determined by dividing the

maximum value by the lowest value in a given test. The influence factor therefore denotes the range, i.e., how far off an estimate may be, which is akin to the method used by de Moel and Aerts et al., [2011]. The results are given in Table 2.8.

- Modification of depth-damage functions: the highest impact is found with respect to the full replacement value (1.605), while the building count is affected the least (1.076) (Table 2.8). This indicates that we can expect to see the largest variance in the results of the total economic losses (building and contents) with respect to full replacement value when the user modifies the depth-damage curve used in the analysis.

- Modification to the water depth: the influence factor is calculated using the total damages to the study area at each water level. The largest influence factor (1.54) is found with the +0.25m (1973 flood) scenario. The flood scenario of -0.46m, has an influence factor of 0.91.

- Modification of restoration time for a single flood scenario (base case): an influence factor of 2 is computed for days to restoration across all water depths in the study area. With respect to income losses, capital and wage losses both have a factor of 1, while rental income influence is slightly larger at 1.3. The largest influence factor when looking at economic losses with respect to restoration time is relocation expenses (1.589).

These results present only a case study in a single region, as recommended by Apel et al., [2009], further test cases in other regions should be undertaken to corroborate the applicability of these conclusions.

**Table 2.8** Influence factor obtained as a ratio between the maximum and the minimum value in a given test

| Varied parameter | Depth-damage Function | | Water Depth | | Restoration | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Building Count | Full Replacement | Depreciated Replacement | +0.25 | -0.46 | Restoration Days | Relocation Expenses | Capital losses | Wage losses | Rental income |
| 1.076 | 1.605 | 1.576 | 1.54 | 0.912 | 2 | 1.589 | 1 | 1 | 1.3 |

## 2.6. Conclusions and Recommendations

In this paper, the sensitivity of flood risk assessment input parameters was investigated using the Canadian version of the U.S. FEMA's loss assessment tool Hazus. The variation of three parameters was considered: depth-damage functions (structure and contents), flood depth and restoration time. First, the default values were used and scenario was considered as the base case scenario. Then, the considered parameters were varied individually. The comparative analysis included: the number of structures which experience damage, economic losses (full and depreciated), and square footage affected.

The selected study area was Fredericton, New Brunswick. The default inventory and demographics from Statistics Canada 2011 census indicate 617 census blocks containing 66,050 residents and 19,178 buildings within the study boundary. Total estimated building replacement value across the study area, excluding contents, is $5,402 million dollars, of which 68% are associated with residential housing. Nationally supplied inventory data were used in this analysis.

Loss estimates indicate that changes in both the structure and associated depth-damage function have significant impact on the final results with respect to building count, total square footage and economic losses. Eighty-one percent of the estimates for full structure replacement are within ±$5million dollars of the default scenario, while the $\sigma = \$4,014.7$. Depreciated replacement costs within one sigma are ± $1,864 million of the base case, with a maximum loss estimate approximately 40% larger.

51

The content range was smaller at (roughly) $10 and $6.5 million for full and depreciated replacement costs respectively. Only the damage curves for single family residence (RES1) were modified in this analysis, and if the project were expanded to include other specific and general occupancy classifications parameters (e.g.: multi-family homes, manufactured homes, commercial, religious structures) the uncertainty in the output is expected to increase.

In this study area, the majority of losses are structural and content building losses, with just a fraction being associated with business interruption, 1.3% in average. As the flood water level with respect to the base-case scenario was raised 0.25m, the total loss estimates to the study region increased by ~55%, conversely, when the flood hazard decreased by 0.49m, the total losses decreased to approximately 91% of the base case. The residential structural classification is estimated to incur the greatest damage in this area, followed by the building contents, and then commercial content.

The above results highlight the importance of proper selection of input parameters customized to the study area under consideration. These should include local inventory of assets at risk (structures, occupancies, and population) and as much as possible consideration of loss assessment on per building basis. Depth-damage damage functions for structures and contents in New Brunswick have not yet been developed. Thus, particular attention should be paid to the adoption of damage curves from other regions where building characteristics are not necessarily the same. These results present only a case study in a single region, further test cases in other regions should be undertaken to corroborate the applicability of these conclusions.

# ACKNOWLEDGEMENTS

# REFERENCES

Aerts, J. C., Botzen, W. W., Emanuel, K., Lin, N., de Moel, H., and Michel-Kerjan, E. O. (2014). "Evaluating flood resilience strategies for coastal megacities." *Science*, 344(6183), 473-475.

Apel, H., Aronica, G. T., Kreibich, H., and Thieken, A. H. (2009). "Flood risk analyses— how detailed do we need to be?" *Natural Hazards*, 49(1), 79-98.

Cardy, W. F. G. (1976). "Flood Management in New Brunswick." *Canadian Water Resources Journal*, 1(1), 40-46.

Chang, W.Y., V.A. Lantz, C.R. Hennigan, and D.A. MacLean (2010). "Benefit-cost analysis of spruce budworm control on Crown land in New Brunswick: incorporating market and non-market values." In: Canadian Resource and Environmental Economics Association Papers and Proceedings 20: 25 pp. (Also In: Proc. SERG International 2011 Workshop. Feb. 8-10, Victoria, BC.)

de MOEL, H., and Aerts, J. C. J. H. (2011). "Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates." *Natural Hazards*, 58(1), 407-425.

Egorova, R., van Noortwijk, J. M., and Holterman, S. R. (2008). "Uncertainty in flood damage estimation." *International Journal of River Basin Management*, 6(2), 139-148.

Environment Canada. (2013). Canada's top ten weather stories for 2013. Retrieved from http://www.ec.gc.ca/meteo-weather/default.asp?lang=En&n=5BA5EAFC-1&offset=2&toc=show

FEMA. (2009). *Hazus 2.1 Flood Technical Manual.* (Technical Manual).USACE.

Government of New Brunwsick (2014). New Brunswick, Environment and Local Government, Flood history database. Retrieved 9/02, 2014, from http://www.elgegl.gnb.ca/0001 /en/Flood/Search?LocationName=Fredericton

Hazus Canada (2014) Inventory: Exposure By Occupancy, Natural Resources Canada, *Hazus Canada Metadata,* hzExposure OccupB_md.rtf,

Hall, J. W., Tarantola, S., Bates, P. D., and Horritt, M. S. (2005). "Distributed sensitivity analysis of flood inundation model calibration." *Journal of Hydraulic Engineering*, 131(2), 117-126.

Horritt, M. S., and Bates, P. D. (2002). "Evaluation of 1D and 2D numerical models for predicting river flood inundation." *Journal of hydrology*, 268(1), 87-99.

Jongman, B., Kreibich, H., Apel, H., Barredo, J. I., Bates, P. D., Feyen, L., and Ward, P. J. (2012). "Comparative flood damage model assessment: towards a European approach." *Natural Hazards and Earth System Sciences*, 12(12), 3733.

McGrath, H., Stefanakis, E., McCarthy, and M., Nastev. M.,. (2014). "Data preparation for validation study of Hazus Canada flood model." International Institute for Infrastructure Resilience and Reconstruction (I3R2) Conference*,* Indiana, 2014

Merz, B., Kreibich, H., Thieken, A., and Schmidtke, R. (2004). "Estimation uncertainty of direct monetary flood damage to buildings." *Natural Hazards and Earth System Science*, 4(1), 153-163.

Merz, B., and Thieken, A. H. (2004). Flood risk analysis: Concepts and challenges. Österreichische Wasser-und Abfallwirtschaft, 56(3-4), 27-34.

Nastev, M., and Todorov, N. (2013). "Hazus: A standardized methodology for flood risk assessment in Canada." *Canadian Water Resources Journal*, 38(3), 223-231.

Nastev, M. (2013). "Adapting Hazus for seismic risk assessment in Canada." *Canadian Geotechnical Journal*, 51(2), 217-222.

Neighbors, C. J., Cochran, E. S., Caras, Y., and Noriega, G. R. (2012). "Sensitivity analysis of FEMA HAZUS earthquake model: case study from King County, Washington." *Natural Hazards Review*, 14(2), 134-146.

Plazak, D. (1984). "A critical assessment of methodologies for estimating urban flood damages-prevented benefits." (*Information Series, n.52 No. 000006953*). Colorado: Fort Collins, Colorado: Colorado Water Resources Research Institute, Colorado State University.

Prettenthaler, F., Amrusch, P., and Habsburg-Lothringen, C. (2010). "Estimation of an absolute flood damage curve based on an Austrian case study under a dam breach scenario." *Natural Hazards and Earth System Sciences*, 10(4), 881.

Scawthorn, C., Blais, N., Seligson, H., Tate, E., Mifflin, E., Thomas, W., and Jones, C. (2006). "HAZUS-MH flood loss estimation methodology. I: Overview and flood hazard characterization." *Natural Hazards Review*, 7(2), 60-71.

Smith, D. I. (1994). "Flood damage estimation- A review of urban stage-damage curves and loss functions." *Water S. A.*, 20(3), 231-238.

Statistics Canada. (2011). Focus on Geography Series, 2011 Census, Retrieved 8/5, 2014, from http://www12.statcan.gc.ca.proxy.hil.unb.ca/census-recensement/2011/as-sa/fogs-spg/Facts-csd-eng.cfm? LANG=Eng&GK=CSD&GC=1310032

UNISDR (2011): Global Assessment Report on Disaster Risk Reduction – Revealing risk, redefining development, United Nations, Geneva

# 3. Rapid Risk Evaluation (ER$^2$) using MS Excel spreadsheet: A case study of Fredericton (New Brunswick)[2]

## 3.1. Abstract

Conventional knowledge of the flood hazard alone (extent and frequency) is not sufficient for informed decision-making. The public safety community needs tools and guidance to adequately undertake flood hazard risk assessment in order to estimate respective damages and social and economic losses. While many complex computer models have been developed for flood risk assessment, they require highly trained personnel to prepare the necessary input (hazard, inventory of the built environment, and vulnerabilities) and analyze model outputs. As such, tools which utilize open-source software or are built within popular desktop software programs are appealing alternatives. The recently developed Rapid Risk Evaluation (ER2) application runs loss assessment analyses in a Microsoft Excel spreadsheet. User input is limited to a handful of intuitive drop-down menus utilized to describe the building type, age, occupancy and the expected water level. In anticipation of local depth-damage curves and other needed vulnerability parameters, those from the U.S. FEMA's Hazus-Flood software have been imported are temporarily accessed in conjunction with user input to display exposure and estimated economic losses related to the structure and the content of the building. Building types and occupancies representative of those most exposed to flooding in Fredericton (New

---

[2] McGrath, H., Stefanakis, E., & Nastev, M. (2016). Rapid Risk Evaluation (ER2) Using MS Excel Spreadsheet: a Case Study of Fredericton (New Brunswick, Canada). *International Society for Photogrammetry and Remote Sensing Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume III-8, 27-34.

Brunswick) were introduced and test flood scenarios were run. The algorithm was successfully validated against results from the Hazus-Flood model for the same building types and flood depths.

## 3.2. Introduction

Every year disastrous climatological and geological hazards take place in Canada and around the globe [Nastev and Todorov, 2013]. In 2010, an estimated 178 million people across the world were affected by flooding and billions of dollars of damage caused [Leskens et al., 2014]. The costliest single disaster on record in Canada is the 2013 flood in Calgary, AB, with an estimated price tag exceeding $6 billion [Environment Canada, 2013]. In New Brunswick (NB), the study area of this project, over $23 million in damages resulted from the Saint John River flooding in 2008, and over 70 floods have been recorded since the 1700s [Public Safety Canada, 2014].

Government officials, GIS specialists, emergency managers, and responders require tools to develop mitigation and recovery plans as well as preparedness and response procedures for natural disasters [Neighbors et al., 2013; McGrath et al., 2015]. Evaluation of risk involves the combination of three components: the potential flood hazard, inventory of the built environment, and representative vulnerability functions – which relate the inundation depth to a percent damage of the asset. Studies of past flood events have shown that the majority of losses arise in urban areas, due to impairment of structures, costs of business shut-down, and failure of infrastructure [Jongman et al., 2012].

Over the past 20 years, considerable progress has been made with respect to flood mitigation strategies by combining strategic planning and risk-based management techniques [Nicholls et al., 2013]. There are a wide variety of flood damage models in use internationally, differing substantially in their approaches to flood computation and estimates of economic costs [Jongman et al., 2012]. There are commercial (AIR Germany Flood Model, AIR UK Flood Model) and open (Basement, Hazus-MH, Kalypso) software solutions available. Many of these solutions have been built specifically to address flood concerns in their country of origin, which, by design, are tailored to meet the conditions, infrastructure, and processes relevant in their geographic region - which often make transferability to another geographic region difficult.

Of the watershed modelling and risk assessment applications available today, few are capable of non-expert implementation [Al-Sabhan et al., 2003]. In addition, the data requirements and data manipulation required for these models to run may exceed the technical capabilities of the broader non-expert safety community [Nastev et al., 2015]. These existing models therefore leave a gap between what is needed (and when) by decision makers and the output a model is able to provide [Leskens et al., 2014].

In this paper, we present the Rapid Risk Evaluation (ER$^2$) tool, which uses a spreadsheet application to compute replacement cost of the building and estimate potential damages resulting from user input flood scenarios. User input is limited to a handful of intuitive drop-down menus utilized to describe the building type, age, occupancy and the expected water level. In anticipation of local depth-damage curves and other needed vulnerability parameters, those from the U.S. FEMA's Hazus-Flood software have been imported temporarily. The computations are done on building-by-

building or aggregate scenario basis. The results of structural and content damage are validated for flood scenarios in Fredericton, NB against Hazus estimates. The paper is structured as follows, In Section 3.3 a brief introduction to flood loss estimation in included and in Section 3.4 the framework of $ER^2$ is described. Section 3.5 introduces the study area. In Section 3.6 $ER^2$ results are presented and comparisons made to Hazus. Section 3.7 summarizes findings and details future research to enhance the application.

### 3.3. Flood Loss Estimation

Direct losses occur as result of direct physical contact of the flood water with humans, properties, or other objects, while indirect losses represent those which are induced by flood impact and may occur (in time or space) outside of the flood event [Merz et al., 2004]. Direct economic losses include calculations of repair and construction costs resulting from the flood event, whereas indirect economic losses are related to lost jobs and business interruption [FEMA, 2010]. It is the calculation of the direct economic losses which were of primary interest in the development of $ER^2$. The most common and internationally accepted method of estimation of urban flood damage is through the use of depth-damage functions [Plazak, 1984]. Structure and contents damage resulting from flood hazard are influenced by many factors, however, usually only building use and inundation depth are considered as damage-causing factors and included in the formulation of depth-damage functions [Merz et al., 2004]. The building age, foundation type, and elevation of the first floor can be included as factors which contribute to the estimated damage of a structure, which are external to the depth-damage functions [FEMA, 2010].

## 3.4. ER$^2$ Methodology

Following the findings of Plazak [1984], Merz et al., [2004] and FEMA [2010],
primary inputs for the computation of direct economic losses focus on depth-damage
functions, inundation depth, building use (occupancy), foundation type, age, and height
of first floor. Using these inputs, estimates of the building value are computed, along with
estimates of damage to the structure, its contents, and the sum of these to show total
losses in Canadian dollars (Figure 3.1). The calculation of building valuation and
estimated damages utilize nationally compiled data tables from the Hazus software,
which have been modified to represent Canadian parameters. The data tables from Hazus
which are used in ER2 include occupancy classification, depth-damage function,
replacement costs per square meter, and floor area (square meter) of a structure,
described in Table 3.1.
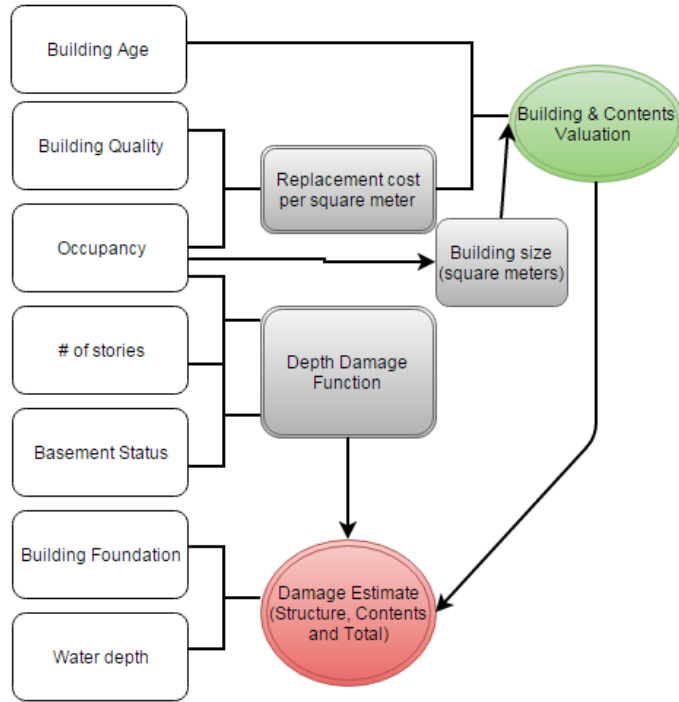
**Figure 3.1** Flow diagram for ER$^2$. White boxes indicate user inputs, grey are computed interim values, and green and red are computed valuation and estimated damages respectively.

**Table 3.1** Hazus tables used

| Table Description | Hazus Table |
|---|---|
| Occupancy Classification | BldgStrDmgFn |
| Depth-Damage Functions | BldgStrDmgFn, BldgContentsDmgFn |
| Replacement costs ($/sf) | hzReplacementCost, hzRes1ReplCost |
| Square footage of building | hzSqftFactors |
| Depreciation factors | DepFunction |
| Exposure contents | hzExposureContentGBldgTypeB |
| Basement status | BldgStrDmgFn, BldgContentsDmgFn, hzRes1ReplCost |
| Garage valuation | hzRes1GarageAdjustment |

The Occupancy Classification table is a primary nominal data category in the worksheet, and is the parent of many other elements in the calculations. Occupancy classifications used in this application include 11 classes of residential occupation, 10 commercial classes, 6 industrial, two for each government and education, and one for

each of agriculture and religious structures. The selection of occupancy is the determiner of the floor area (square meters), replacement costs per square meter, and damage functions. Each of these values is accessed via a lookup table using a common occupancy identifier. The square meters table (hzSqftFactors) is based on distributions of floor area, developed at the dissemination area level for all provinces and territories from the 2011 Canadian Census and Duns & Bradstreet data [Hazus Canada, 2014]. Replacement costs per square meter were derived from the same data, using RSMeans 2006 values for all occupancy codes. The replacement costs per square meter have been averaged over various alternatives for exterior wall construction (e.g.: wood siding over wood frame, brick veneer over wood frame, etc.) [FEMA, 2010].

Presently, the damage functions used in the $ER^2$ algorithm are based on regionally adapted depth-damage functions, based on more than 20 years of claims and measured data in the U.S, however as derived Canadian damage functions become available, they will be imported and set as the default. A table of depreciation factors based on the age of the structure is used to appropriately assign value to the structure. Additionally, tables which account for presence of garage and basement (finished or unfinished) are used in computation of building value and selection of the appropriate damage curve. Using these nationally derived data tables, the $ER^2$ tool results are considered as average for a group of buildings with similar structural and content characteristics. As the data in these tables represent an average, they do not necessarily account for differences and regional variations. To overcome this, an alternative option allowing user input of a known building value is included in $ER^2$ Advanced.

### 3.4.1   ER$^2$ Design Interface

ER$^2$ is comprised of three worksheets which allow user input and present results, and provide a geospatial view via Esri Maps for OfficeTM, while the remaining worksheets provide supporting information for lookup functions and contain computational data. The three worksheets each provide unique scenarios. The first worksheet computes exposure and estimated damages to a single building type with a single user input water level. The second worksheet computes building-by-building losses, allowing up to 300 unique structures, each requiring its own flood level. On the third worksheet users can simulate an aggregated flood risk analysis with up to 50 different building types and four discrete water levels with user input percentages for the dissemination area or block. The tool has been setup for manual user input, or, if data exist in another format, they may be pasted to into ER$^2$ building input section - so long as the fields correspond to expected values.

### 3.4.1.1 ER$^2$ Interface Design

The required user input includes building details and potential water depth, Figure 3.2. The user interface has been configured to accept input to cells via colour coded options. The dark grey cells represent fields for which a user can select an option from pre-determined set of values via a drop down menu. The dropdown menu options have been created using the List option of the Data Validation settings and selecting the defined group name representative of the cells. Light grey cells allow users to input numeric values. Conditional formatting has been used for the manual input cells to restrict user inputs to a given range. Advanced settings which over-ride default

parameters are colored in blue and allow for a user to input a known building value and

manual selection of depth-damage curve or input of a user defined curve.



**Figure 3.2** User input of building details (inventory) and water level (hazard)

### 3.4.1.2. ER$^2$ Result Design

The results available in ER$^2$ are estimates of building exposure and potential direct

economic losses, relevant to the structure and its contents. These data are presented in

tabular format via a pivot table, providing the user the ability to choose the desired level

of detail for reporting and charts which graphically illustrate the damage estimates,

Figure 3.3.
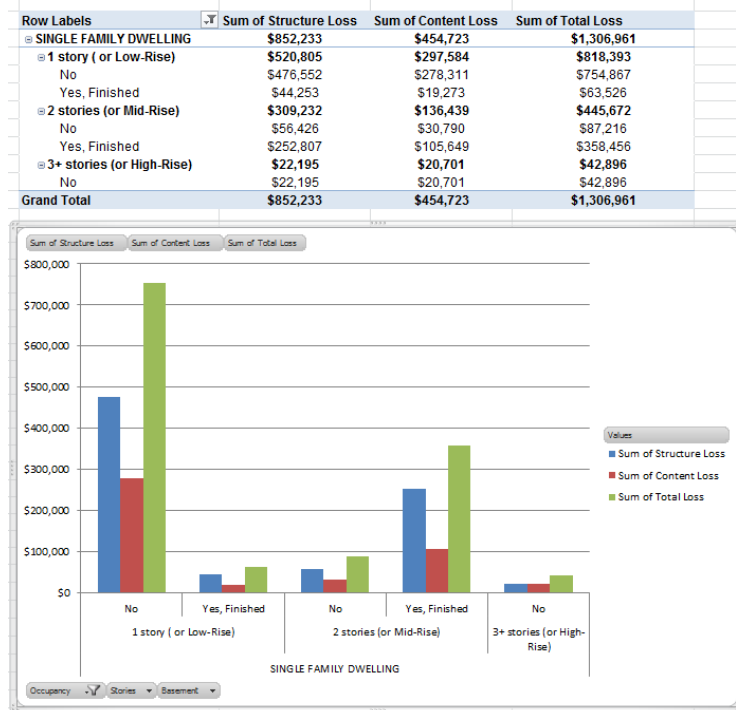
**Figure 3.3** Pivot table and chart results

### 3.4.1.3. ER$^2$ Geospatial Result Design

Included in the building input section of each worksheet is the option to include the spatial location of buildings. The location may be entered in the form of geographic coordinates or the physical address. Through the use of the Esri Maps for MS Office[3], users of the spreadsheet can easily visualize the building details and loss estimations, Figure 3.4. The user may select any field from the spreadsheet to theme the data and select from an abundance of choices regarding colour, symbol and size. To further improve the information portrayed in the map, a user may add any layers from ArcGIS Online or Portal for ArcGIS [ESRI, 2015]. These results may be shared online, sent to a power point presentation or captured as a static image of the map.

---

[3] *Esri Maps for Office* is a registered software product of Esri, Inc., Redlands, CA, USA.
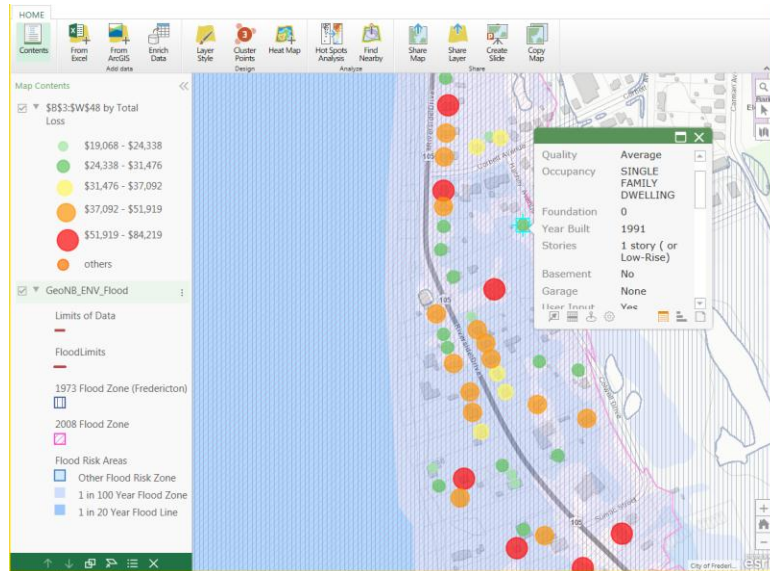
**Figure 3.4** ER$^2$ results, visualized in Esri Maps for Office

## 3.4.2  ER$^2$ Calculations

ER$^2$ computes building valuation (if needed), estimated building losses, structural, contents and total losses, in Canadian dollars.

### 3.4.2.1. Building Exposure Valuation

To avoid confusion with market value or government assessed property, value the cost of a building is computed via a formula reflecting average construction costs. Building value is computed using the age of building, number of stories, size of the structure and the replacement cost per square meter (eq. 3.1). Additional costs are added based on building quality (residential only), basement status and/or garage presence (eq. 3.1).

To compute building valuation, the first input required is the building occupancy classification, for example, consider a single family residence (RES1). Using RES1 occupancy class, the size of the structure ($A_{bld\_k}$) is extracted from the 'hzSqftFactors',

in this example: 185.80m$^2$. RES1 and the number of stories *(s)* are required to look up the replacement cost per square meter $(C_{wi})$, \$709.45 for a one-story average RES1 building. Building value is increased if a basement exists $(b)$ and if it is finished or not $(C_{wb})$. Building value is further increased if a garage is present $(gw_g)$ and depends on the type of garage $(C_{wg})$. The building value is then multiplied by an age depreciation factor $(A_{ge})$ and by the number of structures of the same type $(n)$. Residential structures are a special class, and have multiple RSMeans replacement cost options based on the construction class ($_i$), (e.g.: economy, average, luxury, etc.). These construction classes take into account factors such as the quality of building materials and superiority of craftsmanship [FEMA, 2012].

$$\mathbf{Exp_k} = (\mathbf{A_{bld\_k}}\,(\mathbf{sw_iC_{wi}}) + \mathbf{A_{bld\_k}\,bC_{wb}} + \mathbf{ngw_gC_{wg}})\mathbf{A_{ge}} * \boldsymbol{n} \qquad (3.1)$$

### 3.4.2.2. Contents Exposure

The contents valuation is based on the estimated exposure as described above multiplied by a contents value percentage. The contents value is a fixed percentage, based on the occupancy class of the input structure. For residential structures, contents are set at 50% of the building value, most commercial structures, in part, to account for inventory, compute contents at 100% of the building value, while most industrial buildings, hospitals, and education facilities compute contents value at 150% of the building value. These valuation percentages are based on RSMeans, as read from the 'hzExposureContentGBldgTypeB' table (Table 3.1).

**3.4.2.3. Structure damage estimation**

Of primary influence to the damage estimation is the selection of depth-damage curve, a core of the vulnerability analysis. A damage function is described by an estimated percent of damage at any given water level. There are numerous depth-damage curves available in the literature, describing structural, contents, and inventory damage for each building occupancy type. $ER^2$ (at present) is configured to select a Federal Insurance Administration's (FIA) or modified FIA depth-damage functions for residential structures while the remaining occupancy classes refer to U.S. Army Corps of Engineers (USACE) damage curves. Each building occupancy type has a recommended default damage function associated to it. In addition to the occupancy type, damage functions are unique per basement status and the number of stories. Figure 3.5 illustrates a number of depth-damage functions available in $ER^2$.
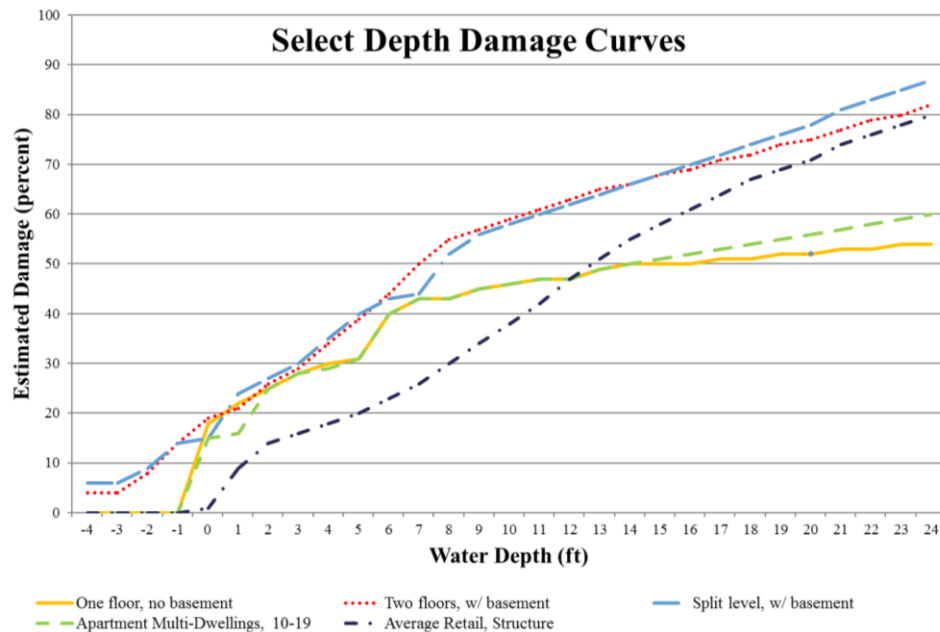


**Figure 3.5** Examples of Depth-Damage curves (FEMA, 2010).

### 3.4.2.4. Contents damage estimation

Estimating the contents damage to a building at a given water level is also based on the relationship of flood depth and percent damage. A separate table of damage curves, designed specifically to represent percent contents damage is referenced. As with the structure depth-damage curves, the contents depth-damage curves are unique per occupancy class, basement status and number of stories. To estimate the contents damages in dollars, the required inputs are: contents valuation ($C_{val}$), and the percent contents damage ($PC$) (eq.3.2). Where $PC$ is equal the intersection of the water depth ($WD$) minus the foundation type ($f$) of the appropriate selected contents depth-damage curve ($CDC$), and $CDC$ is determined based on the occupancy class ($A$), number of stories ($s$), basement status ($b$) (eq. 3.3).

$$CD_k = C_{val} * PC \tag{3.2}$$

$$PC = [CDC_{Asb}]\cap[(WD - f)] \tag{3.3}$$

### 3.5. Study area: Fredericton, NB

The City of Fredericton is the capital of the Canadian province of New Brunswick and located in the west-central portion of the province in Atlantic Canada. The city is split by the St. John River which flows from west to east through the city. The local topography varies considerably with elevations ranging from 2 m to ~175m above sea level, and includes undulating and hilly land [Stobbe, 1940] (Figure 3.6). The basin

immediately surrounding the St. John River and downstream of Fredericton is relatively

flat, with an average mean water level of 2m above sea level [Lantz et al., 2012].

Data from the 2011 Canadian Census indicate approximately 30,000 households in

Fredericton and the surrounding suburbs with a total population of approximately 71,000

[Lantz et al., 2012; Statistics Canada, 2011]. The population of Fredericton is primarily

adults, between 16 and 65 with smaller representation by children (11%) and seniors

(13%). As shown in Figure 3.6 there is a relatively high population density in the flats flat

on the south side of the Saint John River in the Fredericton downtown. Of these

households, approximately 320 buildings have an elevation within 5m of the river bank.

Single family residential buildings are the most common in Fredericton, comprising

73% of the residential structures. Under 10% of residential buildings in the area are

classified as multi-family apartment buildings, and 6.5% of these residences are mobile

homes. Other buildings in the city limits include those classified as commercial (2,100),

industrial (225), Religious (200), government (80) or education (60). The single family

buildings are primarily wood framed, one or two-story buildings, with unfinished

basements or crawlspaces. The mean house value is $212,800.

Flood records, dating back to the 1700s indicate over 70 such events have occurred

in the area, with heavy rainfall, mild weather, snowmelt, and ice jams indicated as

primary causes of this flooding [New Brunswick, Environment and Local Government,

2014]. Flooding in Fredericton occurs when the river level exceeds flood stage, which is

6.5m above sea level. The highest flood levels ever recorded were in 1973, when the

water level was 2.04 m above flood stage, while the second largest flood occurred in

2008, where water level reached 1.86 m above flood stage. The 2008 flood caused

~\$23million in damages to New Brunswick communities situated along the Saint John River [Public Safety Canada, 2014]. For testing of the $ER^2$ algorithms, water levels from two historic flood events were used: the 2008 flood, and the third largest flood on record which occurred in 2005 (Figure 3.7).
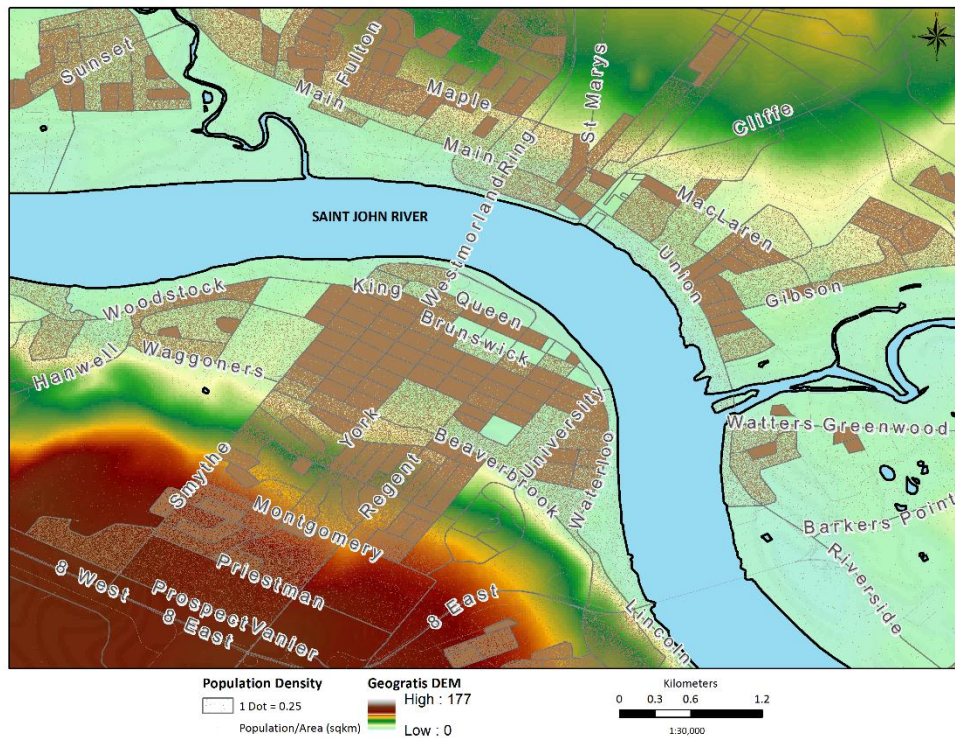


**Figure 3.6** Fredericton, NB. Local topography and population density
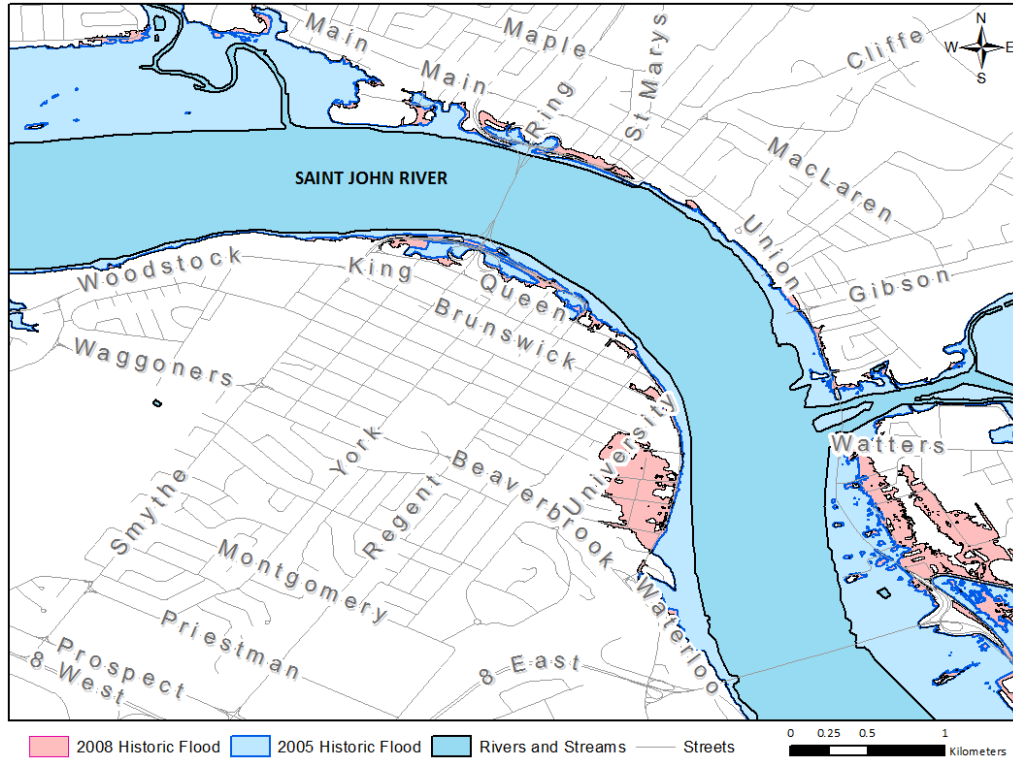
**Figure 3.7** Two historic flood events, Fredericton, NB

## 3.6. Results

Flood depth grid and user defined facilities representing local structures were

imported to Hazus, and analysis was run. The results from Hazus, with (respect to these

user defined facilities), indicate that 356 single family residences experience flooding.

These houses range in value from $41,000 to $495,600, with an average value of

$147,019. The flood depth in these structures, for the 2008 scenario ranges from 1.3cm to

1.67m. The oldest building in the study area has a year of construction dating back to

1901, while the median year of construction 1995. Also damaged in this scenario are 161

manufactured houses, 7 triplexes and 6 multi-family dwellings with 10-19 units. In

addition to residential properties, retail, wholesale trade, churches and entertainment

73

facilities are expected to incur flood damage. Aggregated losses at the dissemination block level were also computed.

At present, local claims data archived by the N.B. Department of Emergency Measures Organization (EMO) are currently unavailable to validate $ER^2$ against field data. Analysis of the results of $ER^2$ are therefore compared only to Hazus in this paper.

As described in Section 3.4, users of $ER^2$ are able input building details and have $ER^2$ compute the estimated building value or they may input a known building value. Each of these options and aggregated results are described in the following sections.

### 3.6.1    Percent Damage

There are over 900 different damage curves defined in $ER^2$. The selection of damage curve for a given structure is based on many factors including: occupancy class, basement, number of stories, type of flooding (riverine or coastal). For a one-story single family residence in Fredericton without a basement, the FIA depth-damage curve with an ID of 105 is the default curve used to predict structure damage (Figure 3.8). The depth-damage definitions are given at foot (0.3048cm) intervals, with percent damage in-between being linearly interpolated. Figure 3.8 illustrates positive correlation of the $ER^2$ damage estimates with the depth-damage curve definitions. However, the Hazus estimates of structure and contents damage do not align as well with the damage curve definitions. Hazus data do follow a similar trend to the depth-damage curve, but tend to be, below 1m water depth: under estimating damage, and over 1m water depth over-estimating damage. The contents damage is underestimated at all water levels in comparison to the curve definition. Plotting percent damage versus water depth for

another flood scenario in Fredericton, NB, the third largest flood (2005) produced a similar figure with the ER$^2$ damages following the damage curve definition while the Hazus results pivot from below to above the curve definition when approximately 1m water depth is recorded.
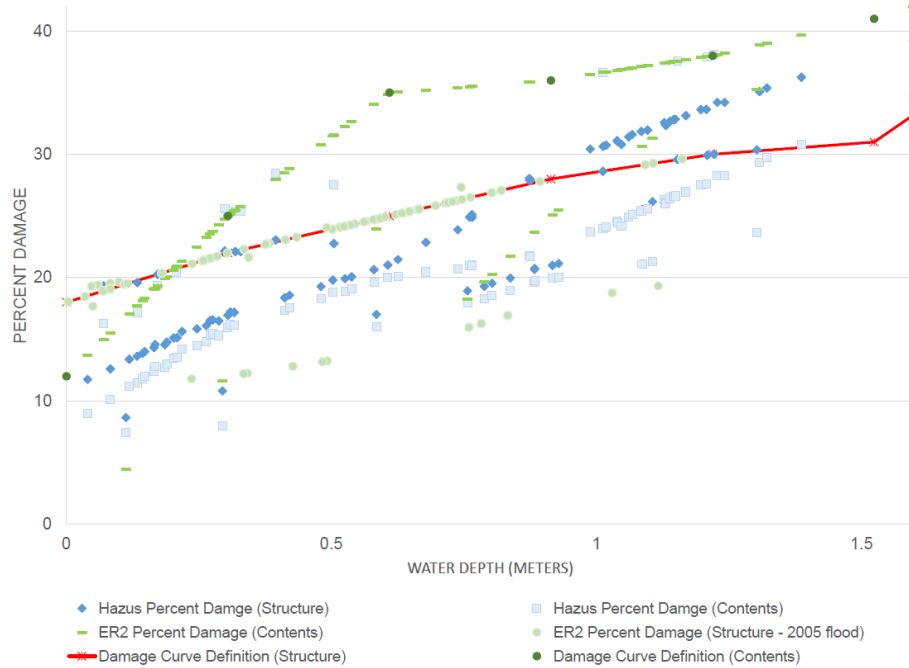


**Figure 3.8** Percent Damage based on water level and depth-damage curve, one-story single family residence with no basement

### 3.6.1    Economic Losses

As discussed in Section 3.4, estimates of damage are based on water depth and percent damage. The percent damage is then converted to currency via the percent of structure value. If a user does not input a known building value, one is computed via average values based on RSMeans. For example, all single family residences are assumed to be 185m$^2$, and replacement cost per square meter of an average quality home is

$709.45. Given this assumptions of generalized costs and size, we can expect the $ER^2$ damage estimates to vary from the Hazus results.

Of the 356 single family residences which, according to Hazus analysis experience flooding, 182 have been identified as single story family residences. Of these 1-story residences, 156 have no basement, while the remainder are considered to have a finished basement. Plotted in Figure 3.9 are estimates of the total losses from each scenario: Hazus, $ER^2$ using computed building value, and $ER^2$ using user input building value. The minimum water depth for which damage is estimated is 1.5cm, with the maximum from the 2008 flood event for one-story single family residences is 1.66 m. There is a trend of increasing uncertainty in the damage estimates as the water depth increases, and the inverse with respect to standard deviation.

When the total damages are plotted against water depth for the three scenarios (Figure 3.9) Hazus, $ER^2$ –user input ($ER^2$-UI) building value, and $ER^2$ – computed building value ($ER^2$-C) we can clearly see a trend to the loss estimates. $ER^2$-C there is a linear increasing trend to the damage estimates with increasing building value. There are two clear trends for $ER^2$-C in Figure 3.9: as water levels increase, from 0 to 0.6m, there is an increase in the total estimated damages across the buildings which have a uniform value. As the water depth exceeds 0.6m there is a shift in damages and the linear trend continues, seeing increased damages with increasing water depth. For the Hazus and $ER^2$-UI the total damage estimates are more randomly distributed, and do not follow the same linear trend, however there is generally an increase of damages with increasing water level. The scattering of the damage estimates for $ER^2$-UI and Hazus represent damages which are not based solely on floor area and a generic value of replacement cost, but an

76

assessed property value which better represents the true value of the building. As $ER^2$-UI and Hazus both use the same input building value, we expect to see a good correlation of total estimated damages.

Comparison of $ER^2$-UI to Hazus structural damage shows, on average an over-estimate of \$1,447, with a standare deviation of 6,058 (

Table 3.2). The difference in the estimated contents damages, between $ER^2$-UI and Hazus also indicate an over estimate by $ER^2$-UI, by an average of \$4,700 across this sample of 156 single family residences with no basement. As $ER^2$-UI uses the same input building value as does Hazus, the difference we see in the total loss estimate can be related back to the differences in percent damage as shown in Figure 3.8. The $ER^2$-C structural and content losses deviate farther from the Hazus and $ER^2$-UI losses, primarily due to the buildings initial value. Since all buildings ($ER^2$-C) started with roughly the same building value, \$186,000, the computed estimated losses have a smaller standard deviation, but are, in magnitude, more different than the Hazus results. On average, the difference between the $ER^2$-C and Hazus loss estimate is 26%, or an over estimate of 26% for structural damage and 45% with respect to contents damage.

**Figure 3.9** Total Damages versus water depth for Hazus, ER$^2$ using user input building value and ER$^2$ using computed building value

**Table 3.2** One-story single family residences, statistics from 156 buildings in Fredericton, NB Estimates from 2008 flood event ER$^2$ user input building value (ER$^2$-UI), ER$^2$ Computed building value (ER$^2$-C )

| | **Hazus** | | **ER$^2$-UI** | | **ER$^2$-C** | |
|---------|-----------|-----------|-----------|-----------|-----------|-----------|
| | **Structure** | **Contents** | **Structure** | **Contents** | **Structure** | **Contents** |
| Average | $27,964 | $12,738 | $29,410 | $17,439 | $35,413 | $18,825 |
| Minimum | $ 8,293 | $ 3,554 | $ 6,637 | $ 2,122 | $0 | 0$ |
| Maximum | $81,451 | $31,942 | $76,364 | $ 48,762 | $54,277 | $34,639 |

In addition to the flooded one-story buildings, there are 80 two-story buildings in the study area which experience flood damage from these flood scenarios. The majority of these residences have a finished basement. The flood depth experienced by these houses ranges from 1.3cm to 1.25m. Given the large percentage of houses with finished basements, total damages are much higher at lower water levels Figure 3.10. Figure 3.10 shows damages computed using the ER$^2$-C (computed building value) which follows a

linear trend which follows the shape of the damage curve, while those which used real property assessment values (Hazus and ER$^2$-UI) are represented by scattered data when plotted against depth.



**Figure 3.10** Total Damages versus water depth for Hazus, ER$^2$ using user input building value and ER$^2$ using computed building value, 2 Story residences, with and without basements

The majority of construction of residential housing in Fredericton began in the 1960s. Plotting the same total loss data from Figure 3.9, using building age instead of water depth along the x-axis, we are able to visualise the estimated damage to structures over four decades (Figure 3.11). From Figure 3.11, it appears that buildings which have a recorded construction date between 1990 and 2005 are those which have the greatest damage incurred based on the 2008 flood scenario. These results seem contradictory to the flood risk mapping and New Brunswick Community Planning Act [New Brunswick Department of Environment, 2013]. A review of the assessment database which lists the

79

building age should be completed to verify structure construction date and thus building age.
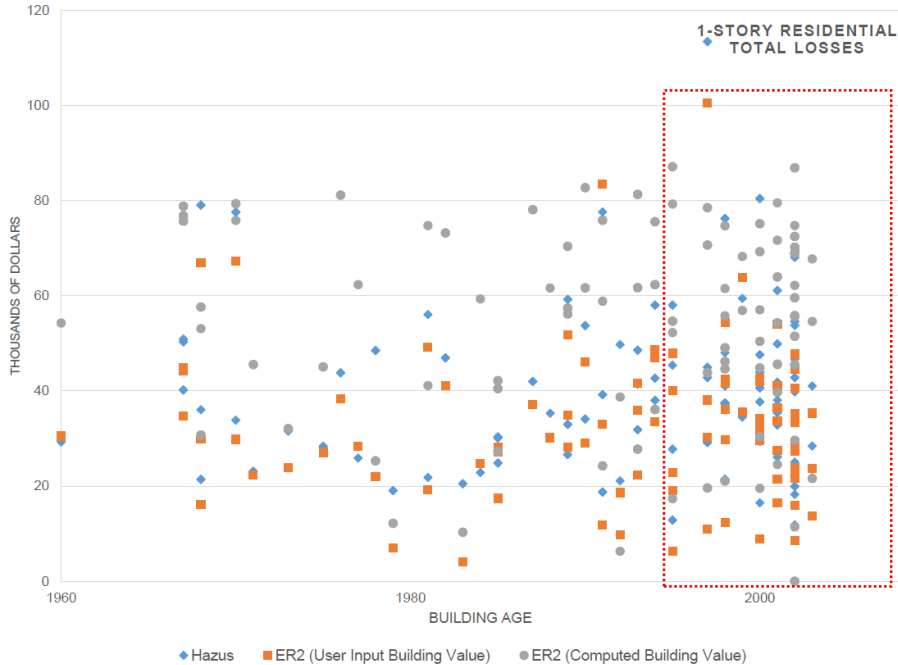


**Figure 3.11** Total Damages versus water depth for Hazus, $ER^2$ using user input building value and $ER^2$ using computed building value by building age. Points within red-dashed polygon represent those with expected damage built between 1990 and 2005

### 3.6.2 Dissemination Area

To simulate an aggregate scenario over a dissemination area similar to Hazus, the buildings are assumed to be equally distributed throughout the area, regardless of their actual location. The water depth is computed as a percent of the dissemination area. Therefore, in the case of the example shown in Figure 3.12, the dissemination area is assumed to have 9% of the area inundated by 3ft of water (5 of 56 cells), 18% with 2ft of water, and 37% with 1ft of water, while 36% of the dissemination area is not flooded. The calculation of damage will assume 4 structures with 3ft of water and multiply this

result by 9% to scale the estimation to represent the inundation area appropriately, and so forth for the remaining water depths.
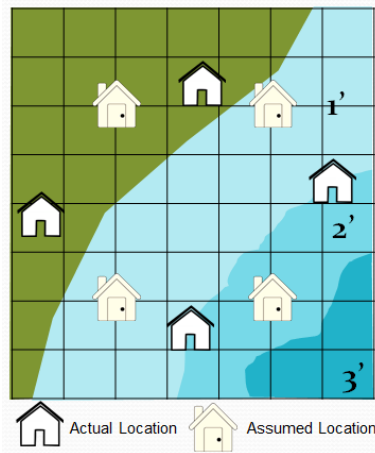


**Figure 3.12** Dissemination area

The sample Fredericton dissemination area contains 52 residential buildings, primarily comprised of 1 story buildings with no basement. Forty percent of the dissemination area was not inundated, while the remainder was flooded with between 1ft and 4ft of water. To simulate four water depths across the study area, the flood levels were queried to find values between 0ft and 1ft, and the average value in this range was selected as the average water depth. This process of selection within an integer range was performed for the remaining ranges and the average value chosen to input to the algorithm. The results of this aggregated style scenario find the total damages to be in agreement within 5.55%, with a larger deviation if looking at just the structure losses individually, Table 3.3.

**Table 3.3** Comparison of ER² to Hazus results for sample dissemination area

|  | **Hazus** | **ER²** | **Difference (%)** |
|---|---|---|---|
| **Structure Damage** | $760,000 | $859,241 | -13.05% |
| **Contents Damage** | $412.000 | $377,874 | 8.28% |
| **Total Damages** | $1,172,000 | $1,237,115 | -5.55% |
| **Exposure** | $11,058,000 | $10,599,561 | 4.15% |

## 3.7. Conclusions

This paper presents the principal characteristics and considerations of the flood risk assessment tool, Rapid Risk Evaluation (ER²) developed within a familiar MS Excel office package, providing capabilities for any user, with (or without) specialized technical knowledge to simulate potential flood risk scenarios and view estimates of exposure and damages.

Results from ER² from two flood scenarios in Fredericton, NB were compared to those results computed by Hazus. We intend to further test these ER² results against field data once access to historic flood claims data is made available.

When ER² User input building values are used, the structural, contents, and total losses computed are in agreement with Hazus results. On average, structure loss differences to be within 1.05% and contents 1.27%. When comparing ER² using computed building values against Hazus loss estimates, the error was higher, an average of 26% difference for structural damage and 45% with respect to contents damage. ER² uses nationally averaged data in the calculation of building area, replacement cost per square meter, and construction costs, it was therefore expected the results from ER² (computed values) to be larger than those computed using assessment data. In both use cases of ER², the damage estimates were found to be higher than those computed using

individual building data. In the aggregated scenario over a dissemination area total damage estimates to be within a ±5.5% range mainly due to the interpolation method.

The comparison is between Hazus and $ER^2$, and these estimates have not been validated against field data. Recommendations for improving this algorithm include expansion of results against insurance claims data and to review results of occupancy classes other than single family residences. Additionally, further testing of $ER^2$ in other communities to see how $ER^2$ compares to Hazus, itself (computed building value and user input building value), and field data are of interest. Furthermore, adding in capabilities which would reflect flow velocity, flood duration, and water contamination are potential enhancements to the application.

Future plans for this algorithm is to offer it up as a web service and to enrich the application with a spatial component which reads building information and details data from local and provincial datasets. This spatial component will provide visualization of results through thematic maps and provide an option alternative to the Esri Maps for Office. $ER^2$ is presently available for download from http://hmcgrat1.ext.unb.ca/ (McGrath, 2015).

## ACKNOWLEDGEMENTS

through the convergence of science and technology with policy, operations and

intelligence. Natural Resources Canada leads the project in partnership with New

Brunswick Emergency Measures Organization, University of New Brunswick, Public

Safety Canada regional office in Fredericton, and New Brunswick Ministry of Transport.

## REFERENCES

Al-Sabhan, W., Mulligan, M., and Blackburn, G. A. (2003). "A real-time hydrological model for flood prediction using GIS and the WWW." *Computers, Environment and Urban Systems*, 27(1), 9-32.

Environment Canada. (2013). Canada's top ten weather stories for 2013. Retrieved 01/06, 2015, from http://www.ec.gc.ca/meteo-weather/default.asp?lang=En&n=5BA5EAFC-1&offset=2&toc=show

Esri. (2015). ArcGIS Maps for Office, Retrieved 02/08, 2015: http://doc.arcgis.com/en/maps-for-office/

FEMA. (2010). Hazus-MH Technical Manual. http://www.fema.gov/resource-document-library

Hazus Canada. (2014). Inventory: Exposure by occupancy, Natural Resources Canada, Hazus Canada Metadata, hzExposure OccupB_md.rtf .Unpublished manuscript.

Jongman, B., Kreibich, H., Apel, H., Barredo, J. I., Bates, P. D., Feyen, L. and Ward, P. J. (2012). "Comparative flood damage model assessment: towards a European approach." *Natural Hazards and Earth System Sciences*, 12(12), 3733.

Koks, E. E., Jongman, B., Husby, T. G., and Botzen, W. J. (2015). "Combining hazard, exposure and social vulnerability to provide lessons for flood risk management." *Environmental Science & Policy*, 47, 42-52.

Lantz, V., Trenholm, R., Wilson, J., and Richards, W. (2012). "Assessing market and non-market costs of freshwater flooding due to climate change in the community of Fredericton, Eastern Canada." *Climatic Change*, 110(1-2), 347-372.

Leskens, J. G., Brugnach, M., Hoekstra, A. Y., and Schuurmans, W. (2014). "Why are decisions in flood disaster management so poorly supported by information from flood models?" Environmental Modelling & Software, 53, 53-61.

McGrath, H., Stefanakis, E., and Nastev, M. (2015). "Sensitivity analysis of flood damage estimates: A case study in Fredericton, New Brunswick." *International Journal of Disaster Risk Reduction*, 14, 379-387.

McGrath H. (2015). Flood Hazard and Risk Assessment Research at UNB, GGE Department, Retrieved from: http://hmcgrat1.ext.unb.ca/

Merz, B., Kreibich, H., Thieken, A., and Schmidtke, R. (2004). "Estimation uncertainty of direct monetary flood damage to buildings." *Natural Hazards and Earth System Science*, 4(1), 153-163.

Nastev, M., and Todorov, N. (2013). "Hazus: A standardized methodology for flood risk assessment in Canada." *Canadian Water Resources Journal*, 38(3), 223-231.

Nastev, M., Abo El Ezz, A., Nollet, M.J., Smirnoff, A., Ploeger, S.K., McGrath, H., Sawada, M. and Stefanakis, E., 2015. "Development of methods and tools for natural hazard risk analysis in Eastern Canada – Use of knowledge to understand vulnerability and implement mitigation measures." *Natural Hazards Review* 18(1) doi:10.1061/(ASCE)NH.1527-6996.0000209

Neighbors, C. J., Cochran, E. S., Caras, Y., and Noriega, G. R. (2012). "Sensitivity analysis of FEMA HAZUS earthquake model: case study from King County, Washington." *Natural Hazards Review*, 14(2), 134-146.

New Brunswick, Environment and Local Government. (2014). Flood history database. Retrieved 9/02, 2014, from http://www.elgegl.gnb.ca/0001/en/Flood/Search?LocationName= Fredericton

Nicholls, R. J., Townend, I. H., Bradbury, A. P., Ramsbottom, D., and Day, S. A. (2013). "Planning for long-term coastal change: Experiences from England and Wales." *Ocean Engineering*, 71, 3-16.

Plazak, D. (1984). A Critical Assessment of Methodologies for Estimating Urban Flood Damages-Prevented Benefits. Colorado Water Resources Research Institute, Colorado State University.

Public Safety Canada, (2014). Canada's National Disaster Mitigation Strategy. Retrieved 3/12, 2014, from http://www .publicsafety.gc.ca/cnt/rsrcs/pblctns/mtgtn-strtgy/index-eng.aspx

Statistics Canada. (2011). Retrieved from http://www12.statcan.gc.ca /census-recensement/2011/as-sa/fogs-spg/Facts-csd-eng.cfm?LANG=Eng&GK=CSD&GC= 1310032

Stobbe, P. C. (1940). Soil Survey of the Fredericton-Gagetown Area, New Brunswick. Experimental Farms Service, Dominion Department of Agriculture.

# 4. DEM Fusion of Elevation REST API Data in Support of Rapid Flood Modelling[4]

## 4.1. Abstract

Digital elevation models (DEM) are an integral part of flood modelling. High resolution DEM data are not always available or affordable for communities, thus other elevation data sources are explored. While the accuracy of some of these sources has been rigorously tested (e.g.: SRTM, ASTER), others - such as Natural Resources Canada's Canadian Digital Elevation Model (CDEM), and Google and Bings' Elevation REST APIs have not yet been properly evaluated. Details pertaining to acquisition source and accuracy are often unreported for APIs. To include these data in geospatial applications and test and reduce uncertainty, data fusion is explored. Thus, this paper introduces a new method of elevation data fusion. The novel method incorporates clustering and inverse distance weighting (IDW) concepts in the computation of a new fusion elevation surface. The results of the individual DEMs and fusion DEMs are compared to a high-resolution Light Detection and Ranging (LiDAR) surface and flood inundation maps for two study areas in New Brunswick. Comparison of individual surfaces to LiDAR find they all meet their posted accuracy specifications, with the Bing data computing the smallest mean bias and the CDEM the smallest RMSE. Fusion of all three surfaces via the proposed method increases the correlation and minimizes both RMSE and mean bias when compared to LiDAR, independent of the terrain, thus producing a more accurate DEM.

---

## 4.2. Introduction

DEMs are an integral part of flood modelling [Cook and Merwade, 2009]. Using the best available elevation data is advised: as DEM resolution and accuracy are the main properties which affect hydraulic and hydrologic modelling results [Vaze et al., 2010]. Flood modelling therefore typically involves high-resolution, e.g. 1m or less, LiDAR elevation data. While LiDAR data is increasingly popular, it is still costly to acquire and computationally expensive [Hummel et al., 2011]. Many larger communities in Canada have budgeted for the expense and invested in a LiDAR dataset; however, this is not feasible for all communities. In lieu of high resolution elevation data coarser and freely available datasets are investigated for rapid flood modelling.

Currently there are several providers of online elevation data with varying levels of coverage, spatial resolution and accuracy. For example, Shuttle Radar Topography Mission (SRTM) (lta.cr.usgs.gov/SRTM), ASTER (asterweb.jpl.nasa.gov/gdem.asp), and Google (developers.google.com) and Bing (msdn.microsoft.com) offer global coverage. Natural Resources Canada (NRCan) offers the CDEM, a Canada wide elevation dataset (geogratis.gc.ca). At the provincial level, several provinces provide access to elevation data, and at a micro level, some municipalities have acquired LiDAR data and enable online access to it. Consequently, multiple datasets covering a single terrain area exist; however, details pertaining to accuracy, acquisition source or technology and collection date vary amongst them and in some cases, are not clearly documented [Buckley and Mitchell, 2004].

While the accuracy of SRTM, ASTER and Cartosat-1 are well documented [Mangoua and Goïta, 2008; Mukherjee et al., 2013; Patel et al., 2016; etc.], the reliability and accuracy of elevation data available from REST APIs, specifically Google and Bing, for uses other than purely recreation, are not. Representational State Transfer (REST) is an architectural style for designing networked applications [Masse, 2011]. An application program interface (API) "exposes a set of data and functions to facilitate interactions between computer programs" and allows for exchange of information [Masse, 2011]. If an API conforms to REST architectural style and guidelines, it is considered a REST API. Though readily available and easy to use, there are certain limitations to these API services, including: source data is often not clearly documented, resolution is not clearly reported, and the interpolation method applied for the calculation is unknown [Stefanakis, 2015]. For example, the Google Elevation API (developers.google.com) is said to be comprised of data from hundreds of providers and is stitched together to provide the best level of coverage available [Thor, 2010]. Recognizing these limitations, there are several advantages to inclusion of these data in the flood modelling process: (i) utilization of REST services is convenient and powerful as there is simple, machine readable data access accessible through a lightweight solution accessed via a URL, (ii) data is processed on the providers' server, and (iii) the data is returned in a well formatted document (JSON or XML). The consistency in the returned data, the ease of access and speed of data acquisition combine to make this a desirable service to include in web projects; and what is probably most important, using this data is computationally inexpensive, resulting in generation of rapid flood models.

Data fusion is considered to improve the trustworthiness of the elevation data of REST API services. This is particularly relevant as resolution and accuracy are the main properties of a DEM which affect hydraulic and hydrologic results [Vaze et al., 2010]. Many researchers, including Costantini et al., [1997] and Buckley and Mitchell [2004] found that fused datasets convey improved knowledge about an area compared to a single dataset. Rationale supporting fusion include the availability of duplicate or partially overlapping datasets, different acquisition techniques, different angles or positions of sensors over different times which may support identification of outliers or inconsistencies [Buckley and Mitchell, 2004]. The benefits of fusion include: (i) surmounted deficiencies associated with one dataset/representation of the area, (ii) increased probability of better surface representation (collection at different angles), (iii) removal of systematic errors, including errors or drifts captured by the inertial navigation system or GPS of the collection sensor, and (iv) variation in resolution, accuracy and scale [Buckley and Mitchell, 2004].

Several data fusion techniques have been proposed, from simple techniques such as data gap filing and weighted averaging [Roth et al., 2002], spatial distribution of error and residual density weighting [Reinartz et al., 2005] and weighting by *a priori* DEM error and derivatives of terrain [Papasaika et al., 2009], to more sophisticated techniques including: sparse representations using segmented patches (unique combinations of terrain shape) along with error weights derived from DEM slope and roughness [Papasaika et al., 2009], frequency domain filtering [Honikel, 1998; Crosetto and Aragues, 2000; Karkee et al., 2008], self-consistency where two DEMS are generated from the same pair of images by switching the reference and target [Schultz et al., 1999],

90

multi-scale kalman filtering [Slatton et al., 2002], and k-means clustering [Fuss, 2013] among others. Many of the fusion techniques are successful in increasing the accuracy and/or the precision of the elevation estimates and typically involve only two DEMs in the analysis. However, despite this synergistic effect, there is no consistent globally applicable solution in place which facilitates merging multiple DEM's [Buckley and Mitchell, 2004; Papasaika et al., 2009]. In addition, most approaches are not data-driven, instead relying on a high quality DEM to improve a less accurate one and on *a priori* knowledge of the accuracy of the input datasets [Fuss, 2013].

Clustering algorithms, typically used in data mining, seem to be one of the most promising methods for merging multiple DEMs. They explore data via an 'unsupervised learning' data driven approach, where *a priori* knowledge of the dataset is not required [Jain and Dubes, 1988]. Data are grouped/clustered based on data similarity, without the need of category labels or described relationships between data. Clusters are considered 'a set of entities which are alike', in that the distance between any points within a cluster is less than the distance of any point outside the cluster [Jain and Dubes, 1988]. Many types of clustering algorithms exist and may be selected based on the data type (e.g.: interval, nominal, binary, ratio, etc.) and how clusters vs individuals will be determined (recursive or iterative, strict or fuzzy membership, etc.). Fuss [2013] analyzed the applicability of k-means, which requires user input of the number of clusters, initial centroid locations, and multiple iterations. On the other hand, the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) method works by finding core samples of high density data and expands clusters from them [Duan et al., 2007].

The development of a framework for simplified flood risk assessment which may be used by any community is explored. The focus is put on DEM use in supporting flood modelling for regions without access to a high resolution DEM. A fused elevation grid from multiple open elevation sources (CDEM, Google and Bing Elevation APIs) is generated using a data-driven approach. The proposed method incorporates the fundamentals of the DBSCAN algorithm [Sander et al., 1998]: neighborhood (eps) and minimum points (MinPts) to form a cluster and identify outliers to exclude from an inverse distance weighted estimate of elevation. To illustrate the effectiveness of the fused DEM for flood modelling, flood grids were created and the resulting grid compared to one derived from a high-accuracy LiDAR surface for two study locations in New Brunswick.

### 4.2.1 Input Data and Study Areas

### 4.2.1.1 Elevation Data

Elevation data used in the fusion process include the Canadian Digital Elevation Model (CDEM) from NRCan (geogratis.gc.ca), Google Maps (developers.google.com) and Bing Maps (msdn.microsoft.com) REST API data. Metadata regarding these data sources are found in Table 4.1. A statistical comparison of the datasets in the two study areas is found in Table 4.2. In Bathurst, the CDEM and Google datasets are very similar, with an RMSE of 1.10 and a mean bias of -0.07m. The highest correlation is found between the CDEM and Google datasets while the poorest involve the Bing DEM, with $R^2$ values of 0.9714 (Google/Bing) and 0.9706 (CDEM/Bing).

The CDEM data were extracted from "hypsographic and hydrographic elements of the National Topographic Data Base or various scaled positional data acquired from the

provinces and territories" [NRCan, 2015]. This dataset stores ground elevations in metres

relative to Mean Sea Level (MSL) based on the North American Datum 1983 (NAD83)

horizontal reference datum and the national Canadian vertical reference Canadian Vertical

Geodetic Datum of 1928 (CGVD28). The measured altimetric accuracy of the CDEM

varies across the country from 0 –70m. In the chosen NB study areas, this data is reported

to be vertically accurate 0 to 10m and verified between 1981 and 1990. Horizontally, the

cell size of the CDEM varies across the country with changing latitude, from 23m x 16m

to 93m x 65m. In these two study areas, the resolution of available data from the Google

and Bing REST APIs is 30m.

 Local LiDAR data, acquired by Leading Edge Geomatics in 2011 (Fredericton) and

2013 (Bathurst) is used for the base elevation data. The average flying height for the

Fredericton, NB project area is 1400m which produces a horizontal accuracy of ~0.35m

@ 1 sigma [Leading Edge Metadata, 2011]. The elevation data was tested against 200

GPS RTK Survey points, 0.13278m at 95% confidence interval, and RMSE of

0.067745m.

**Table 4.1** Elevation service data resources

|  | CDEM (Geogratis) | Google Maps (Google) | Bing Maps (Bing) | Lidar |
|---|---|---|---|---|
| **Resolution/ horizontal vertical** | 23 x 16m to 93 x 65m 0 – 70m *rounded to nearest metre | 10 – 900 m (depending on location) | 10 – 900 m (depending on location) *rounded to nearest metre | 1m horizontal |
| **Source** | NRCan | unknown | unknown | Leading Edge Geomatics |
| **Vertical Reference** | (CGVD28) | WGS84 EGM96 Geoid | EGM2008 2.5' or Height Above Ellipsoid (HAE) WGS84 | (CGVD28) |
| **Date** | 1945- 2011 (study area: 1981 – 1990) | unknown | 2008 | 2011 – 2013 |
| **Coverage Area** | Canada wide | Global | Global: Only latitude in range:  -85 to +85 | Study wide |

**Table 4.2** Statistical Comparison of CDEM, Google and Bing. Shaded boxes with bold characters indicate best fit values, μ =- mean bias, σ = standard deviation, RMSE = root mean square error, $R^2$ = correlation

| Location | DEMs | μ (m) | Avg. μ (m) | σ (m) | RMSE | Avg. RMSE | $R^2$ |
|---|---|---|---|---|---|---|---|
| Fredericton | CDEM/Google | 0.60 | 0.26 | 2.87 | 2.93 | 2.02 | 0.9978 |
| Bathurst | CDEM/Google | -0.07 | | 1.10 | 1.10 | | 0.9969 |
| Fredericton | CDEM/Bing | 0.82 | 0.54 | 3.58 | 3.68 | 3.52 | 0.9964 |
| Bathurst | CDEM/Bing | 0.27 | | 3.34 | 3.35 | | 0.9706 |
| Fredericton | Google/Bing | 0.22 | 0.30 | 4.01 | 4.01 | 3.67 | 0.9956 |
| Bathurst | Google/Bing | 0.37 | | 3.31 | 3.33 | | 0.9714 |

## 4.2.1.2 REST Elevation API

The REST Elevation API requests are constructed as Uniform Resource Locator (URL) strings which contain information about the geographic location of interest, data spacing, and an API key. Once sent to the server, the URL request is processed, and the elevation data is returned to the user in a formatted JSON or XML document. An example of a request to the Google Elevation API service is shown below:

```
https://maps.googleapis.com/maps/api/elevation/xml?path=y1,x1|y2,x2&samples=s&key=API_KEY
```

Variables which are user defined and may change per request include the start and end coordinates of the request: $y_1, x_1, y_2, x_2$ and the number of data points to extract along the path, samples=*s*. Additionally, an API Key is required to access these services.

A Python script was developed which relies on user input in the form of a bounding box and cell resolution. The script takes this input and forms requests of appropriate syntax and length to call to the server. For the Bing API, the request format is in the form of bounding box, including the number of rows and columns - where the number of returned elevation points cannot exceed 1024, while the Google API utilized the path request, as illustrated above. CDEM is also accessible via REST API, taking a similar 'path' form as the Google service. Python libraries including *Math*, *Urllib2*, and *Json*

were utilized to take the user input and compute the appropriate individual requests, such that the number of points requested in a single request did not exceed the stated capabilities. Once the API request url was sliced and formatted, the request was sent to the appropriate server, the returned JSON document was parsed to extract position and elevation data, stitched together to generate a single array, and then converted to GeoTIFF. Geospatial Data Abstraction Library (GDAL) library was utilized for the creation of a georeferenced raster from the elevation data points.

### 4.2.1.3 CDEM Extraction

An alternative to accessing CDEM from the REST API, is via interactive web browser. The CDEM can be downloaded by selecting the geographic extent on the web map from the Geospatial Data Extraction page of Geogratis, (geogratis.cgdi.gc.ca/). The GeoTIFF DEM is emailed to you once the request has been processed by the servers, additionally other DEM derived products can be downloaded: slope, shaded relief, etc. The time this takes depends on current servers' load and complexity of request.

### 4.3  Study Areas

To evaluate the suitability of elevation data from CDEM, Google and Bing Elevation APIs and the proposed fusion technique for rapid flood modelling, two study areas in New Brunswick were selected: Fredericton and Bathurst, NB. These have been selected as there are a mix of urban and rural land use and include shallow and steep sloping terrain with rivers and waterbodies in the vicinity.

### 4.3.1 Fredericton

The City of Fredericton is the capital of New Brunswick and located in the west-central portion of the province in Atlantic Canada, at a latitude of ~45.96°N. The city is split by the St. John River which flows from west to east through the city (Figure 4.1(i)). The local topography varies considerably with elevations ranging from 2m to ~120m above sea level, and includes fairly flat land surrounding the St. John River [Stobbe, 1940]. The Fredericton area is predominately low slope, with 66.7% of the area having slopes < 5°, and less than 2% of the area with slopes > 10°. The city and surrounding region has a population of 56,000 [Statistics Canada, 2014] with approximately 12,000 dwellings. The study area selected for analysis is 5 x 5 km.

### 4.3.2 Bathurst

Bathurst is located on the north east coast of New Brunswick, approximately 47.62°N, on the estuary of Nepisiguit River and the southernmost part of the Chaleur Bay (Figure 4.1(ii)). There are two spits of land, Carron Point and Alston Point, which nearly enclose the entrance to Bathurst Harbour. Additionally, there are two major rivers which flow into Bathurst Harbour, the Tetagouche River and the Nepisiguit River. Although riverine flooding is a concern, the primary susceptibility to flooding within the region is sea level rise and storm surge. Bathurst has about 12,300 residents and approximately 6,300 dwellings [Dietz, 2016]. The city centre is relatively protected by its location well above sea level, however houses and cottages built along the sand pits are directly exposed to high velocity winds and storm events coming off the bay [Dietz, 2016]. The study area is approximately 10 x 15 km, with an elevation range of 0 – 113 m above sea

level. The slope distribution in the Bathurst area is similar to Fredericton: over 72% of the area is low slope (< 5°), 6% in the mid-range, (5 – 10°) and 2% of the area has slopes steeper than 10°.
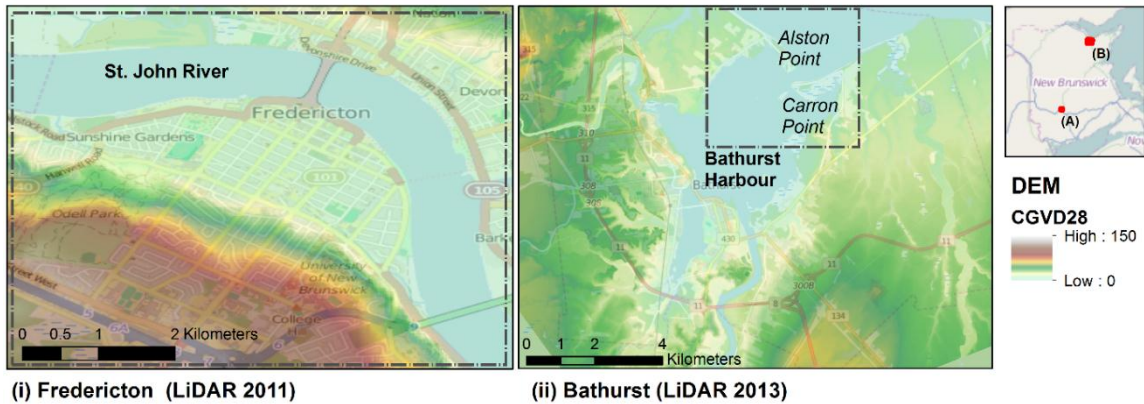


**Figure 4.1** Topography of two New Brunswick study areas, (i) Fredericton, (ii) Bathurst NB, dashed boundaries indicate extent shown in Figures 4.4 and 4.5

## 4.4 Fusion Method

The section describes the proposed method of fusion which incorporates concepts of clustering and inverse distance weighting (IDW). The aim of fusion is to derive a better-quality DEM, where 'better' is defined as minimizing the mean bias and relative height accuracy (RMSE). A multi-step process was applied to generate the fusion DEM: (i) a DEM stack was created first from the input surfaces, (ii) clustering to determine which values fall within each Cluster, (iii) IDW to compute new fusion elevation data, and (iv) the fusion DEM surface compiled.

### 4.4.1 DEM Stack

In this first step, all images were projected to same horizontal projection (NB Double Stereographic) NAD83 CSRS, and vertical reference datum of Canadian

Geodetic Vertical Datum of 1928 (CGVD28). The cell resolution of each of the data

sources differ, varying from 1 m (LiDAR) to 30 m (Bing and Google), with the CDEM

within this range ~23 m. A common cell size of 30m was selected as it was the coarsest

of the available datasets. Point data from the LiDAR and CDEM datasets were used to

generate DEM of common cell size, 30m using the kriging interpolation method. The

LiDAR and CDEM data are both available as CGVD28 reference. For the Bing dataset,

the data was downloaded from the REST API referenced to "height above ellipsoid" and

converted to CGVD28 via the Natural Resources Canada GPS-H tool

(webapp.geod.nrcan.gc.ca/geod/tools-outils/gpsh.php?locale=en). The Google data has an

unknown vertical reference. An assumption was made that, in Canada, the vertical

reference of mean sea level is to CGVD28, and thus data was not vertically adjusted.

Once all the DEMs were in the same horizontal and vertical projection, they were clipped

to the same extents, and the number of rows and columns counted to ensure the cell

assignments were consistent among the DEMs, creating a DEM stack.

## 4.4.2   Clustering

The method proposed utilized the fundamentals of DSCAN clustering. The key

concepts of: (i) neighborhoods (*eps*) of a given radius and (ii) minimum number of points

(*MinPts*) to form a cluster are used, where clusters are considered dense regions in the

data space. The idea is for each data point, the neighborhood (*eps*) of a given radius has

to contain at least a minimum number of points otherwise a cluster is not formed [Xu

1997]. The neighborhood of a point, *p*, is denoted by $N_{Eps}(p)$ and is defined by:

$$N_{Eps}(p) = \{q \in D | dist(p,q) \leq Eps\} \tag{4.1}$$

where $dist(p,q)$ is a distance function for two points, $p, q$. Any data points which do not satisfy the condition of MinPts, are outliers and considered noise.

For each cell in the DEM, the primary cell elevation $(z_{i,j})$ value was extracted along with its eight surrounding neighbors, Figure 4.2(i, ii). In the case of all three DEMS, each cell passes 27 elevation data points through to the DBSCAN algorithm from the *scikit-learn* machine learning library for Python to determine cluster assignment. During cluster assignment, the elevation value was the only considered data, thus the cell location and horizontal distance (e.g. $z_{i+1,j-1}$) was ignored. Based on the selection of eps and MinPts, the elevation values were classified into clusters or noise Figure 4.(i). In Figure 4.(i) there are two clusters which are generated, based on the condition of MinPts = 7 and eps =1.75, and three elevation points are considered noise as they do not meet the minimum density or number of point's required to form a cluster. Of the two derived clusters, the cluster with the greater number of members $(C_{max})$, in this example Cluster 1, is selected and passed to IDW to compute the new fusion elevation $(z_f)$. If zero clusters are formed, elevation data from only cell $z_{i,j}$ are passed to IDW.



(i)

| Source | X | Y | Elevation | Cluster |
|--------|---|---|-----------|---------|
| Bing | i | j | 12.53 | 1 |
| CDEM | i | j | 13 | 1 |
| Google | i | j | 12.53 | 1 |
| Bing | i | j+1 | 13 | 1 |
| CDEM | i | j+1 | 13 | 1 |
| Google | i | j+1 | 13.55 | 1 |
| Bing | i | j-1 | 12.53 | 1 |
| CDEM | i | j-1 | 13 | 1 |
| Google | i | j-1 | 12.1 | 1 |
| Bing | i+1 | j | 13 | 1 |
| CDEM | i+1 | j | 13 | 1 |
| Google | i+1 | j | 15.53 | -1 |
| Bing | i+1 | j+1 | 12.53 | 1 |
| CDEM | i+1 | j+1 | 12 | 1 |

| Source | X | Y | Elevation | Cluster |
|--------|---|---|-----------|---------|
| Google | i+1 | j+1 | 12.05 | 1 |
| Bing | i+1 | j-1 | 10.75 | 0 |
| CDEM | i+1 | j-1 | 11 | 0 |
| Google | i+1 | j-1 | 10.89 | 0 |
| Bing | i-1 | j | 9.71 | -1 |
| CDEM | i-1 | j | 11 | 0 |
| Google | i-1 | j | 10.57 | 0 |
| Bing | i-1 | j+1 | 8.80 | -1 |
| CDEM | i-1 | j+1 | 10 | 0 |
| Google | i-1 | j+1 | 10.56 | 0 |
| Bing | i-1 | j-1 | 12.00 | 1 |
| CDEM | i-1 | j-1 | 12 | 1 |
| Google | i-1 | j-1 | 12 | 1 |

(ii)

**Figure 4.2** (i) Data extraction for fusion of cell i,j and cell size, (ii) Example data from single cell (i,j) and neighbors

The values of eps and Minpts parameters were varied using set values, 0.5 to 10, and values which floated based on the range of the input elevation data to determine the most suitable values. The floating value was computed as the standard deviation of the input $i,j$ cell elevations. MinPts was also varied to determine the best parameter value. MinPts values between 2 and 12 were used as was a floating value of MinPts, based on the range of input elevation values. For example, if the input elevation range was less than 3 m, 10 points were required for a cluster, and if the elevation range was greater than 6 m, 3 points were required for a cluster. The results of the derived fusion surfaces were compared to the LiDAR DEM and the mean difference (bias) and RMSE statistics were considered to determine the most suitable eps and MinPts for each of the fusion combination surfaces,
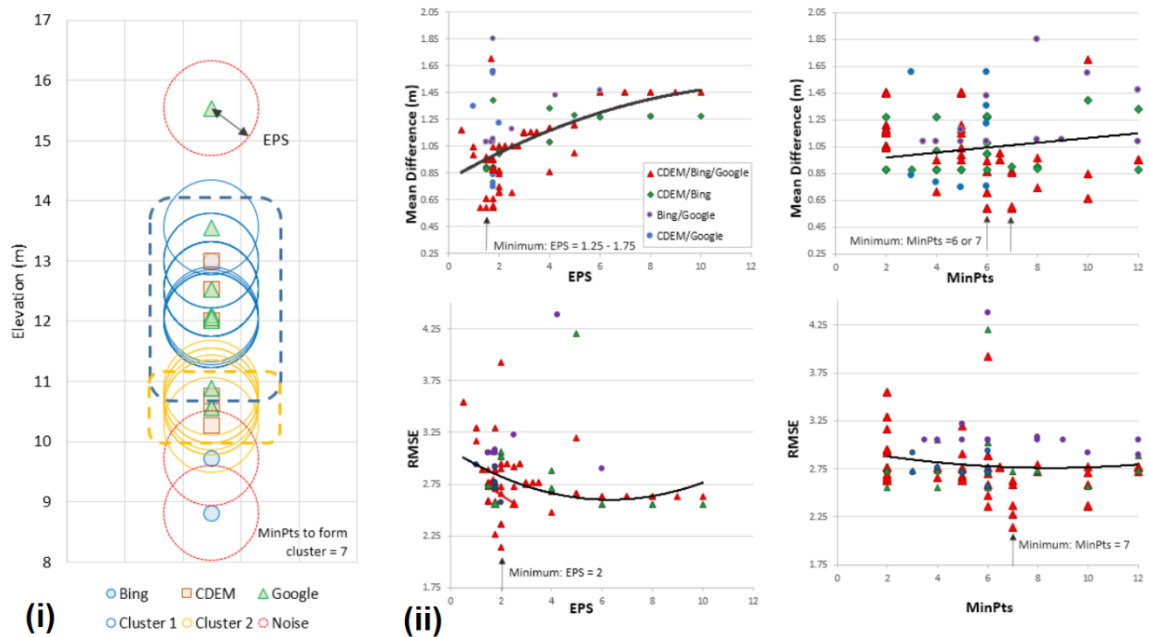


**Figure 4.3** (i) Sample cluster classification based on EPS = 1.75 and MinPts = 7, and (ii) Accuracy results of from varying eps and MinPts parameters. Accuracy results with respect to Mean Difference and RMSE in comparison to LiDAR elevation data.

100

### 4.4.3 Inverse Distance Weight

Elevation data which are part of $C_{max}$ are passed to the inverse distance weight algorithm to compute fusion elevation value $(z_f)$.

To compute $z_f$ for cell $i,j$, data in $C_{max}$ are averaged based on the inverse distance to the cell centre of $i,j$ (eq.2). Data from $i,j$ are given weight of 100%, orthogonal cells (o) have distance of 30m and those diagonal $(d)$ are 42.42m centre-to-centre, these elevation points are thus assigned weights of 30% and 20% respectively.

$$z_f = \frac{\sum_o^n i + \sum_o^n o * w_o + \sum_o^n d * w_d}{n_i * + n_o * w_o + n_d * w_d} \tag{4.2}$$

where i = cell $i,j$ data, o = orthogonal cells, and $d$ = diagonal cells and $w_{ij} = 1, w_o = (1/30) * 10 \ w_d = (1/42) * 10$.

### 4.4.4 Final DEM

The point data was then converted to an elevation raster surface using the GDAL write array to raster function.

## 4.5 Results

The results are divided into two sections: (i) evaluation of the elevation data of the individual and fusion DEMs and (ii) comparison of DEM flood grids to the flood surface derived from the LiDAR surface – including both water depth and spatial coverage of flood extent.

### 4.5.1 DEMs

To facilitate comparison, all DEMs were analyzed with cell size of 30m. While this increase of cell size of the LiDAR data reduces accuracy, when input to hydrologic program, the estimate of volume of runoff is relatively insensitive to this increase [Fellows and Ragan, 1986]. Goulden et al., [2014] found that changes in cell size affect the topography and contribute to variability in watershed attributes, including: basin area, stream length and location. However, given the width of waterbodies in these study areas, the chosen 30m cell resolution is sufficiently able to represent the watershed attributes.

- **Individual DEMS:** The three individual surfaces fit the LiDAR well within their posted accuracy specifications (Tables 4.1 and 4.3). The largest errors are found in the area of an old growth forest in Fredericton (5-20 m) and within the boundaries of the water bodies. In Fredericton DEM errors within the Saint John River, are found to be up to 4 m (CDEM), 2 m (Bing), and less than 1m (Google).

All individual DEMs have a positive bias, on average over-estimating the elevation comparison with the corresponding LiDAR surface. This is likely a result of the elevation data containing surface returns, not just bare-earth measurements as found in the LiDAR dataset. For example, if the elevation data of the Bing or Google REST API is comprised of the SRTM data, elevations will likely be overestimated as SRTM does not produce a bare-earth DEM [Zanderbergen, 2008]. In the SRTM dataset forested areas the elevation values typically fall somewhere between ground and canopy, and in urban areas pixel elevations are affected by any buildings in that pixel [Zanderbergen, 2008], thus producing a positive bias to a bare-earth dataset.

The root mean square error (RMSE) of the CDEM dataset is lower than either the Bing or Google data, making it a better fit, Table 4.3. However, the overall error bias of the Bing data is more closely aligned to the LiDAR in both study areas, with an average bias over the two areas of 0.68m above the LiDAR. Results of the Google dataset find better overall mean bias, but lower RMSE than the CDEM, and the inverse when compared to the Bing results, thus the Google dataset is ranked 2$^{nd}$ in both statistics. The elevation scatter plots in Figure 4.4, 4.5(ii), compare cell by cell elevation values and indicate a strong linear trend across all elevations and fairly high density, indicating good correlation. The Google dataset from Fredericton has the largest $R^2$, indicating the regression model accounts for 0.9951 of the variance, however all individual DEMs are found to have $R^2$ values above 0.97. The histogram of the individual DEM differences from LiDAR show the Google data with the most normally distributed data, with most of the elevation differences within ±2m, while the Bing data is more widely distributed with a large frequency of pixels around -2m and the CDEM data having a right tailed distribution (Figure 4.4 (iii), a-c). The Pearson correlation coefficient (r) of all datasets indicate strong correlation of the individual DEM data to the LiDAR data, with values ranging from 0.9861 to 0.9987.

As the resolution of the CDEM data set varies with latitude, 0.75 arc seconds (south-north) to 3 arc seconds (west-east), and the vintage of the data from CDEM varies across the country from 1945 to 2011, fusion of this dataset with a more current dataset would provide greater confidence in the validity of the data and the representation of present day terrain.

**Table 4.3** Statistical comparison of individual DEMs to LiDAR. Shaded boxes with bold characters indicate best fit values, μ =- mean bias, σ = standard deviation, RMSE = root mean square error, $R^2$ = measure of fit to regression line, r = Pearson's correlation coefficient

| Location | DEM | μ (m) | Avg μ (m) | σ (m) | RMSE | Avg RMSE | $R^2$ | r |
|---|---|---|---|---|---|---|---|---|
| *Fredericton* | Bing | **0.53** | **0.68** | 3.33 | 3.37 | 2.91 | 0.9937 | 0.9969 |
| *Bathurst* | Bing | 0.83 | | 2.30 | 2.44 | | 0.9789 | 0.9894 |
| *Fredericton* | Google | 0.75 | 0.79 | 3.04 | 3.13 | 2.98 | **0.9951** | 0.9976 |
| *Bathurst* | Google | 0.83 | | 2.71 | 2.83 | | 0.9724 | 0.9861 |
| *Fredericton* | CDEM | 1.35 | 1.11 | 2.15 | **2.54** | **2.69** | 0.9974 | **0.9987** |
| *Bathurst* | CDEM | 0.86 | | 2.71 | 2.84 | | 0.9726 | 0.9862 |

- **Fusion DEMs:** Four fusion DEM surfaces were created using unique combinations of the individual DEMs (CDEM/Bing, Bing/Google, CDEM/Google, and CDEM/Google/Bing). The LiDAR DEM is subtracted from the results of the fusion DEMs to create difference DEMs (dDEM).

The fusion dDEMs in Fredericton all have relatively high computed values for Pearson's correlation coefficient (r) and $R^2$ with values of ranging from 0.973 to 0.9974 and 0.9864 to 0.9987 respectively, Table 4.4. The highest correlation coefficient (r) and $R^2$ is found for the fusion surface which uses all three DEMs. This three DEM fusion dataset (CDEM/Google/Bing) also reports the lowest root mean square error in each study area, with an average RMSE = 2.50. However, this fusion DEM has the second smallest mean bias, average +0.01m larger than the Bing/Google dataset. The poorest results were from the CDEM/Bing and CDEM/Google fusion DEMs, however, the average bias and RMSE of these two fusion datasets is equal to or less than that of the individual Google and CDEM DEMs.

**Table 4.4** Statistical comparison of fusion DEMs to LiDAR. F=Fredericton, B = Bathurst, shaded boxes with bold characters indicate best fit values, μ =- mean bias, σ = standard deviation, RMSE = root mean square error, $R^2$ = correlation to best fit line, r = Pearson's correlation coefficient

| Site | Fusion | μ (m) | Avg μ (m) | σ (m) | RMSE | Avg RMSE | $R^2$ | r |
|---|---|---|---|---|---|---|---|---|
| F | *CDEM/Google/Bing* | 0.60 | | 2.80 | **2.27** | | **0.9974** | **0.9987** |
| B | *CDEM/Google/Bing* | 0.84 | 0.72 | 2.60 | 2.74 | 2.50 | 0.9742 | 0.987 |
| F | *CDEM/Bing* | 0.62 | | 2.61 | 3.04 | | 0.9950 | 0.9975 |
| B | *CDEM/Bing* | 0.85 | 0.73 | 2.65 | 2.79 | 2.92 | 0.9730 | 0.9864 |
| F | *CDEM/Google* | 0.75 | | 2.61 | 2.71 | | 0.9964 | 0.9982 |
| B | *CDEM/Google* | 0.85 | 0.80 | 2.68 | 2.81 | 2.76 | 0.9730 | 0.9864 |
| F | *Bing/Google* | **0.59** | | 2.98 | 2.87 | | 0.9958 | 0.9979 |
| B | *Bing/Google* | 0.82 | **0.70** | 2.65 | 2.77 | 2.82 | 0.9730 | 0.9864 |

As discussed above for the individual surfaces, Bing has the lowest mean bias however it does not have the smallest RMSE; CDEM on the other hand, shows the smallest RMSE, but not the smallest mean bias, with the Google dataset ranking second in each of these statistics. In both study areas, comparison of the individual to the fusion DEM data, the CDEM/Google/Bing fusion surface produced the best DEM when compared to LiDAR DEM, with respect to RMSE and $R^2$. The correlation (r) of the data of the CDEM/Google/Bing dataset is equivalent to the best individual result of CDEM in Fredericton. The average mean bias is smallest in the individual Bing datasets, μ = 0.68, however the average mean bias in the Bing/Google and CDEM/Google/Bing fusion data is not much larger, μ = 0.71 and μ = 0.72 respectively. Thus, the fusion process improves DEM accuracy.

## 4.5.2   Flood Inundation Maps

To verify the spatial distribution of the correlation to the LiDAR dataset and the impact of the terrain on the accuracy of the final DEMs, inundation maps were created

for the three individual DEM surfaces and the four final fusion generated DEMs using the

Flood Information Tool (FIT) in ArcMap 10.0 [FEMA, 2010]. The inputs required to

compute the flood depth grid include: the DEM, floodplain boundary and a series of cross

sections with attributes containing height of water and/or discharge. The Fredericton

flood scenario, which is approximately representative of the 2008 flood event, was used

with a discharge volume of 9,120$m^3$/s, in Bathurst, the scenario was based on sea level

rise of 4.5m.

First, the total flooded area is used as a simple measure of the accuracy of the

DEMs. Applying the LiDAR DEM surface, a total of 8.45% and 18.35% of the study

area is flooded in Fredericton and Bathurst respectively Table 4.5. In the Fredericton

scenario, the flood surface obtained with the CDEM/Google/Bing fusion DEM showed

the closest match to the LiDAR surface, indicating 8.79% of the area is flooded, an

overestimation of 0.34% which is equivalent to 0.26$km^2$. In Bathurst, the individual Bing

data was the best match when comparing total area flooded, with a difference of -0.09%,

while the Google and the CDEM/Google/Bing fusion were ranked second, with a

difference of 0.30%. All flood surfaces correlated within ±1.6% of total flooded area or

0.75$km^2$ (Fredericton) and 1.59$km^2$ (Bathurst). In the both study areas, the flood grids

from the fusion surfaces, on average, over-estimated the total flooded area, while the

individual DEMs in Fredericton underestimated total flood area.

However, just comparing the total flooded area does not take into consideration the

overlapping, i.e., which cells are flooded and which are dry in each of the flood grids. A

robust approach to derive areal statistics is achieved using a fit measure, as suggested by

Bates and deRoo [2000], eq.3. In this approach the observed inundated area $(A_{obs})$ is compared to that of the predicted by the model $(A_{mod})$:

$$F = \frac{A_{obs} \cap A_{mod}}{A_{obs} \cup A_{mod}} * 100 \qquad (4.3)$$

when F = 1, the two flood rasters coincide perfectly.

In Fredericton, the CDEM/Google/Bing results most closely matched the LiDAR total flooded area, with a difference of just 0.34% flooded area. However the fit measure, or cell-by-cell comparison finds a relatively low correlation of flooded areas, F = 0.7933, Table 4.5 - indicating that many of the cells flooded in the LiDAR flood grid are not flooded in the CDEM/Google/Bing grid. The Bing dataset has a much higher fit measure, F = 0.9183 in Fredericton and F = 0.8956 in Bathurst. Overall, the fit measures in Bathurst study area are higher than Fredericton, with nearly all being greater than 0.9. In Fredericton the fit measures of the individual surfaces range from 0.7652 (Google) up to 0.9183 (Bing), while the fusion DEMs have fit measures between F = 0.73 and F = 0.80.

- **Fredericton:** In the CDEM/Google flood grid has 49% of the cells indicate water depth ±1 m of the LiDAR flood grid, with 37% of these within ±0.30m, Figure 4.4(i), a-g. The CDEM/Google/Bing fusion DEM created the next best result, with 48% within ±1, with 31% of these within 0.3m, Figure 4.4(d). The individual Bing DEM has the poorest result, with just 5% within ±0.3m and 11% within ±1m, Figure 4.4(c).

**Table 4.5** Difference in flooded area (%) using individual and fusion derived elevation data surfaces and Fit measure (F). Flooded area computed using Flood Information Tool. Results compared to flood surface generated from LiDAR DEM. Bold values represent best values

| ` | Fredericton | | | Bathurst | | |
|---|---|---|---|---|---|---|
| | *% Flooded* | *% change* | *F* | *% Flooded* | *% change* | *F* |
| ***LiDAR*** | 8.45% | - | 1 | 18.35 | | 1 |
| **CDEM ONLY** | 7.46% | -0.99% | 0.8856 | 19.03 | 0.68% | 0.9233 |
| **Google ONLY** | 7.94% | **-0.51%** | 0.7652 | 18.65 | 0.30% | **0.9282** |
| **Bing ONLY** | 7.79% | -0.66% | **0.9183** | 18.26 | **-0.09%** | 0.8956 |
| ***Fusion:*** | | | | | | |
| **CDEM/Google/Bing** | 8.79% | **0.34%** | **0.7933** | 18.65% | **0.30%** | 0.9154 |
| **CDEM/Bing** | 9.22% | 0.77% | 0.7246 | 19.88% | 1.53% | 0.8796 |
| **CDEM/Google** | 9.67% | 1.22% | 0.7866 | 18.81% | 0.46% | 0.9252 |
| **Bing/Google** | 7.79% | -0.66% | 0.7354 | 19.04% | 0.68% | **0.9464** |

In the north-west portion of the map in Figure 4.4(i), north of the St. John River, the entire area is flooded in the base LiDAR case. The Bing, Google and fusion surfaces capture nearly the extent of flooding in this neighborhood, while the CDEM does not. While these surfaces capture flooding in this area, the reported depth of water varies from 0.20m to 5m. Moving east across Figure 4.4(i), along the north bank of the river, the Bing and CDEM/Bing surfaces do not illustrate flooding at all along the north bank, from Ring Road to Gibson St, contrary to the LiDAR flood grid. The remainder of the surfaces do present flooding along the north bank, though not perfectly representing the base case, but the CDEM/Google/Bing and CDEM/Google surfaces are good approximations. On the north shore south of the Nashwaak River the CDEM surface greatly overestimates flooding and the Bing underestimates flooding in this area. The CDEM/Google/ Bing and the CDEM/Google best correspond to the LiDAR in this area. On the southern shore of the St. John River, in the downtown area, all surfaces overestimate flooding, however the Bing data does the worst job at representing flooding, overestimating flooding from the western extent of the study area by up to 6m, all the way along the southern shore to

University Avenue in the east. Overall, based on qualitative analysis, the CDEM/Google/Bing fusion surface, provides the nearest approximation to the LiDAR flood surface.

The elevation scatter plots in Figure 4.4(ii), which compare cell by cell elevation values, indicate a strong linear trend and fairly high density, indicating good correlation in elevation data. The mean and standard deviation are computed from the observed data to the polynomial fit line. The mean for all elevation data, individual and fusion surfaces, are slightly below zero, while the mean for all flood grid data is zero. The individual CDEM data has the lowest standard deviation, with respect to the elevation data, $\sigma = 2.15$, with the CDEM/Google/Bing being slightly larger, $\sigma = 2.16$, and the Bing dataset is the highest, $\sigma = 3.3$. Of the flood grids, the Bing/Google surface has the lowest standard deviation, $\sigma = 1.64$.

Each of the fusion flood grids have smaller computed RMSE than the individual surfaces. The CDEM /Google fusion, in terms of relative mean error, computes the best matching flood grid, with RMSE = 1.17, Figure 4.4(iii). This flood surface also has a relativly small, positive mean bias, $\mu = 0.48$. All of the flood grids produce a positive correlation ($r > 0$). The elevation histograms (Figure 4.4(iii)), for the fusion flood surfaces better approximate a normal distribution, than do the individual DEMs.

Overall, the fusion surface of the CDEM/Google, Google/Bing and CDEM/Google/Bing produce flood grids with strong correlation, r = 0.5758 to 0.6194. The individual CDEM surface is weakly correlated with the LiDAR generated flood grid, with r = 0.2833.

- **Bathurst:** The flood grids generated by the individual and fusion DEMs better approximate the LiDAR flood grid in Bathurst than in the Fredericton area. All DEMs produce flood grids which approximate the flooding extent along both Alston and Carron Point, Figure 4.5(i), however the magnitude of the computed flood differences vary amongst the grids in these areas, with ranges within ±2m of the LiDAR derived flood surface. Along the northwestern shore of Chaleur Bay, west of Alston Point, some inland flooding captured in the LiDAR surface is not represented by any of the individual or fusion DEMs. On the east shore of Chaleur Bay flooding is generally underestimated, as many cells are flooded in the LiDAR surface and not in any of the other grids. However the Google and CDEM/Google/Bing surfaces show a more mixed composition of under and overestimation of flooding in this area.

The elevation scatter plots which plot the LiDAR elevation values against the DEM surface elevation values show lower density than that of the Fredericton study area, especially in the lower elevation ranges, 0 to 40m. Some of this may be explained by differences in elevation computed in water bodies of Chaleur Bay, Bathurst Basin, and the rivers, Figure 4.5(ii). The mean differences from the elevation data to the polyfit line approximate zero. The fusion surfaces derived from combining Bing with CDEM or Google show more outliers, indicating the poorer correlation to the LiDAR data. Overall, the standard deviation of all surfaces are similar, within the range of $\sigma = 2.2$ to 2.6.

The histograms for each surface follow a normal distribution, with peaks near zero. The elevation data from the Bing/Google fusion has the smallest mean bias, $\mu = 0.82$ with the Bing and Google surfaces having the next smallest bias, $\mu = 0.83$. The elevation

histograms are all have two peaks: -1m and 0.8m. Reviewing the spatial distribution of these peaks find these to be primarily within Bathurst Harbour and Chaleur Bay.

Of the individual datasets, the CDEM/Google/Bing surface produces the flood grid which is most statistically similar to the LiDAR flood grid. The CDEM/Google/Bing fusion flood grid computes a large correlation value, r = 0.9945, and a computed relative height accuracy, RMSE = 0.92, Figure 4.(iii).

In the individual datasets, in both locations, the Bing data has the smallest mean bias, the CDEM has the smallest RMSE, while the Google surface is in between in each of these measures. With respect to the fusion surfaces, the CDEM/Google/Bing DEMs compute smaller RMSE and higher correlation when compared to the LiDAR DEM. This fusion surfaces therefore results in flood grids which better approximate the flooding extent and depth than the individual DEMs for both study areas.

**Figure 4.4** Fredericton study area (i) flood depth grids, (ii) correlation scatter plot, and (iii) histogram of individual DEMs: (a) CDEM, (b) Google, and (c) Bing and fusion DEMs (d) CDEM/Google/Bing, (e) CDEM/Bing, (f) CDEM/Google, (g) Bing/Google compared to LiDAR (red) and LiDAR derived flood grid (blue) determined as the baseline surface
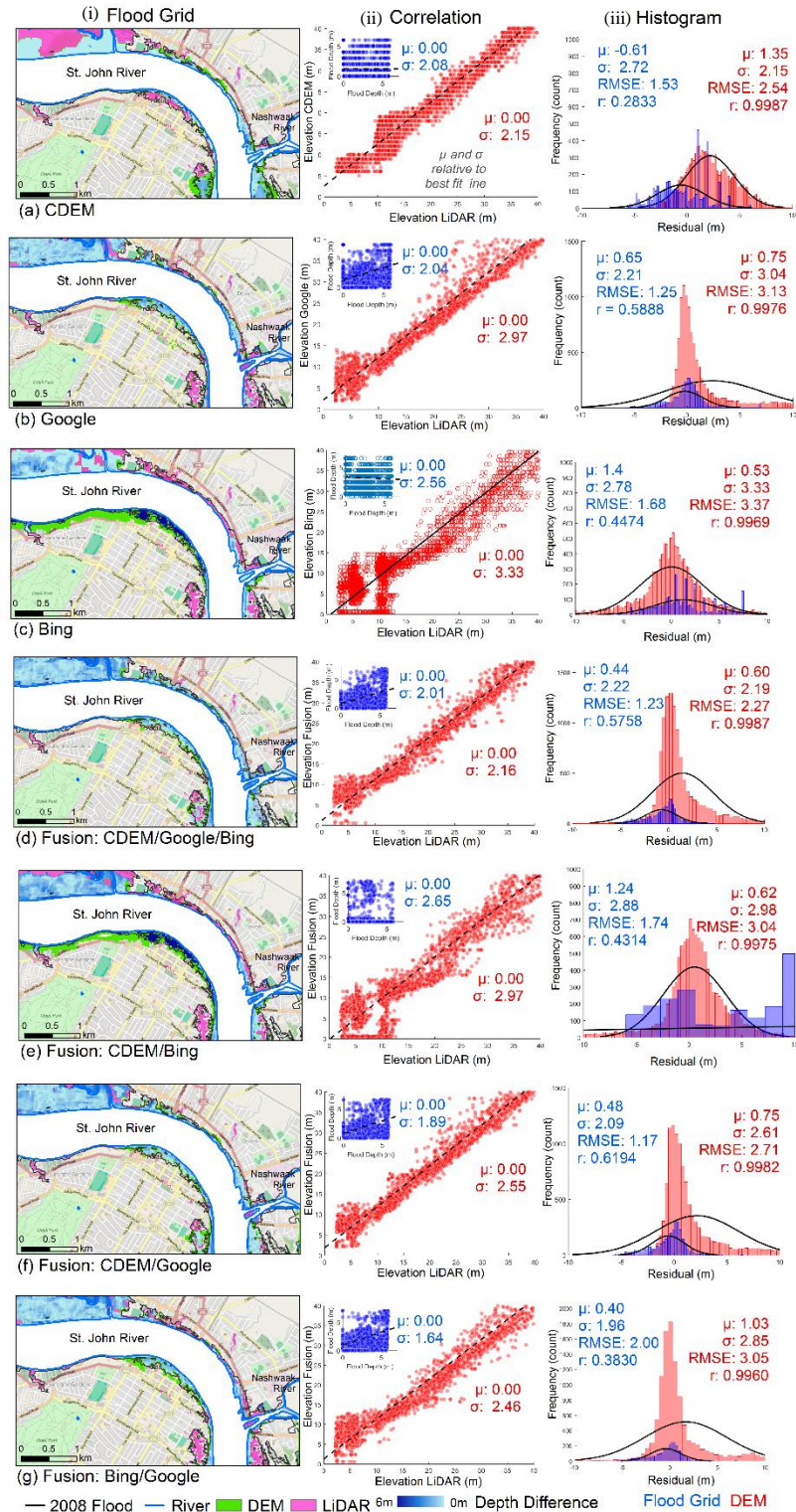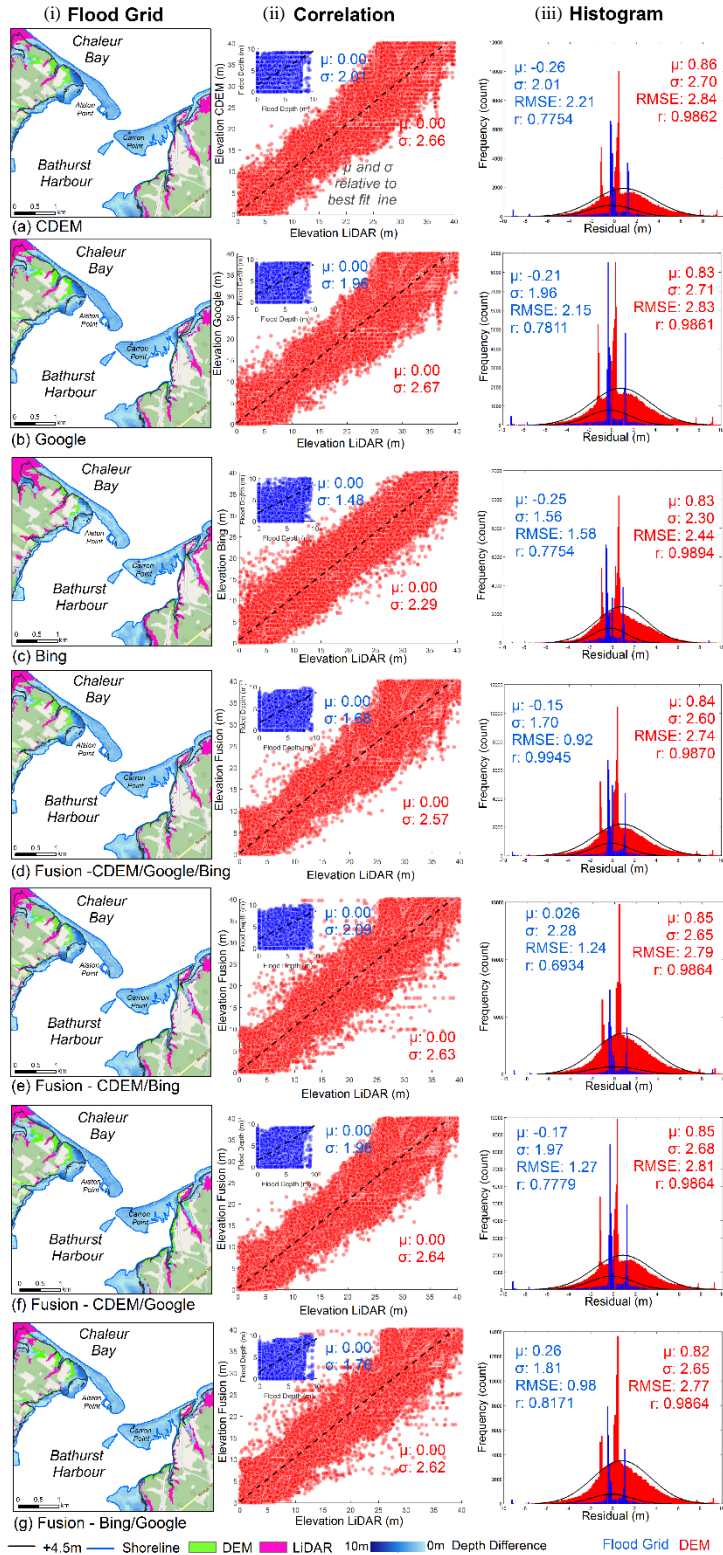
**Figure 4.5** Bathurst study area (i) flood depth grids, (ii) correlation scatter plot, and (iii) histogram of individual DEMs: (a) CDEM, (b) Google, and (c) Bing and fusion DEMs (d) CDEM/Google/Bing, (e) CDEM/Bing, (f) CDEM/Google, (g) Bing/Google compared to LiDAR (red) and LiDAR derived flood grid (blue) determined as the baseline surface

113

## 4.6  Conclusion

The development of a framework for simplified flood risk assessment which may be used by any community is explored, with a focus on how one can support flood estimation for regions without their own high resolution DEM. This is particularly relevant as resolution and accuracy are the main properties of a DEM which affect flood modelling results. Three readily available Canada wide elevation datasets were considered: CDEM, Google Maps and Bing Maps REST APIs. There is limited transparency regarding the source of the data from these providers, thus data fusion is considered to improve their reliability. A fused elevation grid is generated using a data-driven approach to achieve the best-accuracy DEM possible from these inputs. The proposed method used fundamentals of DBSCAN clustering algorithm: eps and MinPts, to determine elevation clusters from noise, and inverse distance weighting to generate a new fusion DEM. To illustrate the effectiveness of the individual and fusion DEMs for flood modelling, they were imported to flood modelling software and the resulting flood depth grid compared to one derived from high-accuracy LiDAR.

Two locations in eastern Canada were used as study areas to test the applicability of open source REST API elevation data over different terrain configurations. Overall, the three individual surfaces fit the LiDAR well within their posted accuracy specification. The CDEM dataset has the lowest RMSE, making it a better choice; however, the error bias of Bing is more closely aligned to the LiDAR in both study areas, with the Google dataset ranking second in each of these statistics.

Fusion was performed via a novel approach which incorporates clustering and inverse distancing weighting to combine the DEMs with the aim of improving DEM

accuracy. In both study areas fusion was able to reduce the RMSE and mean bias while increasing the $R^2$ and correlation. The best results were found when all three DEMs were combined. Flood grids derived from the fusion DEMs generated a better match, in both spatial extent and depth, to the LiDAR derived flood surface than any of the individual DEMs.

Based on these results, the combination of open source CDEM elevation data fused with Bing and Google REST API elevation data provides a better DEM than any of those a forenamed individual surfaces. While the resolution of these data ~30m is significantly coarser than LiDAR, many applications exist for which data of this resolution and accuracy are suitable, including rapid and preliminary flood modelling.

Future recommendations include analysis of other geographic locations to see if they validate these findings and testing these data sources using other documented fusion techniques to determine if this method may provide better results than those previously proposed. Additionally, testing the effect of maintaining 3 dimensions in the clustering algorithm and other clustering methods may be tested to determine if they provide better results.

## ACKNOWLEDGEMENTS

Canada's ability to anticipate, prevent/mitigate, prepare for, respond to, and recover from natural disasters, serious accidents, crime and terrorism through the convergence of science and technology with policy, operations and intelligence.

## REFERENCES

Bates, P., and De Roo, A. P. J. (2000). "A simple raster-based model for flood inundation simulation." *Journal of Hydrology, 236*(1-2), 54-77.

Buckley, S. J., and Mitchell, H. L. (2004). "Integration, validation and point spacing optimisation of digital elevation models." *The Photogrammetric Record*, 19(108), 277-295.

Cook, A., and Merwade, V. (2009). "Effect of topographic data, geometric configuration and modeling approach on flood inundation mapping." *Journal of Hydrology*, 377(1), 131-142.

Costantini, M., Farina, A., and Zirilli, F. (1997). "The fusion of different resolution SAR images." Proceedings of the IEEE, 85(1), 139-146.

Crosetto, M., and Pérez Aragues, F. (2000, March). Radargrammetry and SAR Interferometry for DEM Generation: Validation and Data Fusion. In SAR workshop: CEOS Committee on Earth Observation Satellites (Vol. 450, p. 367).

Dietz, S. (2016). *Climate Change Adaption Plan, coastal erosion and flooding city of Bathurst.* Unpublished manuscript.

Duan, L., Xu, L., Guo, F., Lee, J., and Yan, B. (2007). "A local-density based spatial clustering algorithm with noise." *IS Information Systems, 32*(7), 978-986.

Fellows, J. D. and Ragan, R. M. (1986). "The role of cell size in hydrology oriented geographic information systems." Hydrologic Applications of Space Technology. International Association of Hydrological Sciences (IAHS) Pub, (160), 453-460.

Fuss, C. E. (2013). *Digital elevation model generation and fusion*. Doctoral dissertation, Department of Geography, University of Guelph, ON, Canada

Goulden, T., Hopkinson, C., Jamieson, R., and Sterling, S. (2014). "Sensitivity of watershed attributes to spatial resolution and interpolation method of LiDAR DEMs in three distinct landscapes." *Water Resources Research, 50*(3), 1908-1927.

Hummel, S., Hudak, A. T., Uebler, E. H., Falkowski, M. J., and Megown, K. A. (2011). "A comparison of accuracy and cost of LiDAR versus stand exam data for landscape management on the Malheur National Forest." *Journal of Forestry*, 109(5), 267-273.

Jain, A. K. and Dubes, R. C. (1988). Algorithms for clustering data. Prentice-Hall, Inc.

Karkee, M., Steward, B. L., and Aziz, S. A. (2008). "Improving quality of public domain digital elevation models through data fusion." *Biosystems Engineering*, 101(3), 293-305.

Leading Edge Geomatics. (2011). *LiDAR collection metadata.* Unpublished manuscript.

Mangoua, F. H., and Goïta, K. (2008). "A comparison between Canadian digital elevation data (CDED) and SRTM data of Mount Carleton in New Brunswick (Canada)." International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 1423-1430.

Masse, M. (2011). *REST API Design Rulebook: Designing Consistent RESTful Web Service Interfaces*. O'Reilly Media, Inc.

Mukherjee, S., Joshi, P. K., Mukherjee, S., Ghosh, A., Garg, R. D., and Mukhopadhyay, A. (2013). "Evaluation of vertical accuracy of open source Digital Elevation Model (DEM)." *International Journal of Applied Earth Observation and Geoinformation*, 21, 205-217.

NRCan. (2015). *Canadian Digital Elevation Data (CDED) source.* Unpublished manuscript.

Papasaika, H., Poli, D., and Baltsavias, E. (2009, February). "Fusion of digital elevation models from various data sources." Advanced Geographic Information Systems & Web Services, 2009. GEOWS'09. International Conference on (pp. 117-122). IEEE.

Patel, A., Katiyar, S. K., and Prasad, V. (2016). "Performances evaluation of different open source DEM using Differential Global Positioning System (DGPS)." *The Egyptian Journal of Remote Sensing and Space Science*, 19(1), 7-16.

Reinartz, P., Müller, R., Hoja, D., Lehner, M., and Schroeder, M. (2005, June). "Comparison and fusion of DEM derived from SPOT-5 HRS and SRTM data and estimation of forest heights." In Proc. EARSeL Workshop on 3D-Remote Sensing, Porto (Vol. 1).

Roth, A., Knopfle, W., Strunz, G., Lehner, M., and Reinartz, P. (2002). "Towards a global elevation product: Combination of multi-source digital elevation models." *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(4), 675-679.

Sander, J., Ester, M., Kriegel, H. P., and Xu, X. (1998). "Density-Based Clustering in Spatial Databases: The algorithm GDBSCAN and its Applications." *Data Mining and Knowledge Discovery*, 2(2), 169-194.

Schultz, H., Riseman, E. M., Stolle, F. R., and Woo, D. M. (1999). "Error detection and DEM fusion using self-consistency. In Computer Vision, 1999." The Proceedings of the Seventh IEEE International Conference on (Vol. 2, pp. 1174-1181). IEEE.

Slatton, K. C., Crawford, M., and Teng, L. (2002). "Multiscale fusion of INSAR data for improved topographic mapping." In Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. 2002 IEEE International (Vol. 1, pp. 69-71). IEEE.

Statistics Canada. (2011). Retrieved from http://www12.statcan.gc.ca /census-recensement/2011/as-sa/fogs-spg/Facts-csd-eng.cfm?LANG=Eng&GK=CSD&GC=1310032

Stefanakis, E. (2015, Nov 2). Elevation web services: Limitations and prospects. GoGeomatics Canada, online Magazine

Stobbe, P. C. (1940). *Soil survey of the Fredericton-Gagetown area, New Brunswick.* Ottawa: Experimental Farms Service, Dominion Dept. of Agriculture in co-operation with the New Brunswick Dept. of Agriculture.

Thor. (2010). [Google maps API v3] comparison of google elevation data and SRTM3 elevation data. Retrieved from https://groups.google.com/forum/#!topic/google-maps-js-api-v3/KcjkS-U36dE

Vaze, J., Teng, J., and Spencer, G. (2010). "Impact of DEM accuracy and resolution on topographic indices." *Environmental Modelling & Software*, 25(10), 1086-1098.

Xu, X., Ester, M., Kriegel, H. P., and Sander, J. (1997). "Clustering and knowledge discovery in spatial databases." *Vistas in Astronomy*, 41(3), 397-403.

Zandbergen, P. (2008). "Applications of shuttle radar topography mission elevation data." *Geography Compass*, 2(5), 1404-1431.

Zandbergen, P. A. (2006). "The effect of cell resolution on depressions in digital elevation models." *Applied GIS*, 2(1), 04-1.

# 5 Online Reduced Complexity Flood Modelling: Leveraging Open Data and Limited User Input

## 5.1 Abstract

Modeling flood inundation is a complex process involving numerous variables and factors and extensive literature addressing different modeling and simulation aspects. Existing sophisticated solutions are run on desktop computers requiring considerable processing time, inputs, and user knowledge and are typically not suitable for rapid flood prediction. Hence, important communication gaps between the experts on one side and the needs of the local public safety community and population living in flood prone zones to understand their own exposure and vulnerability on the other exist and are not well addressed. Coupling recent advances in remote sensing techniques and the augmentation of open data with the notion that the best model is the simplest one, a framework for simulating flood inundation via web browser is introduced. Two reduced complexity flood models are explored, (i) 0D flooded water surface and (ii) hybrid 1D/2D cell storage model. The framework considers limited user input, thus facilitating its use by non-expert users. Leveraging freely available open source data, including digital elevation models (DEM), river gauge measurements, intensity duration frequency tables (IDF), and applying basic hydrologic principles and processes, flood scenario modelling capabilities at user specified level or recurrence interval have been created. Initial results reveal significant advantages for the non-expert public safety community: rapidity and ease of use. The presented user-friendly tools allow communities to generate their own

flood inundation scenarios on demand in support of informed emergency response and mitigation planning.

## 5.2 Introduction

Modeling flood inundation is a complex process involving numerous variables and factors and extensive literature exist addressing different modeling and simulation aspects [Kulkarni et al., 2014]. The best model is often the one which provides the user with the information required, whilst reasonably fitting the available data – recognizing that the processes necessary to include to best approximate historic/future events are still subject to considerable uncertainties [Bates and De Roo, 2000]. While sophisticated software packages exist, they are typically based on desktop solutions, requiring extensive processing time, sizable inputs, and knowledge to run and interpret [Leskens et al., 2014]. Timely and accurate prediction of inundation extent and potential impacts and consequences is fundamental for the sustainable development of a given region and provides valuable information necessary for understanding respective exposure and vulnerability [Scawthorn et al., 2006]. Currently, no application is suitable or available specifically for interventions where flooding is imminent or in progress [Poulin et al., 2012]. Kulkarnie et al., [2014] successfully developed a web based GIS and urban flood simulation model, however the run time in their test sites exceeded 6 hours.

Hydraulic models solve unsteady flow along the considered length of an open channel as governed by continuity and momentum equations (eq 5.1, 5.2), together called Saint Venant equations, and are classified through different forms or numerical methods, or by their dimensionality. The components of the momentum equation (eq. 5.2) include: local acceleration, convective acceleration, pressure forces, gravity and friction force

components. One dimensional (1D) models may solve a simplified form of these

equations by assuming shallow water terms have an insignificant impact on the result.

The 1D kinematic wave approximation assumes all the terms, with the exception of

friction and bed gradient, as negligible. On the other hand, the diffusive wave

approximation includes water slope in addition to friction and gravity forces. The

dynamic wave approximation includes all terms of the momentum equation (eq. 5.2).

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = 0 \tag{5.1}$$

$$\underbrace{\frac{1}{A}\frac{\partial Q}{\partial t}}_{\substack{\text{Local} \\ \text{Acceleration} \\ \text{term}}} + \underbrace{\frac{1}{A}\frac{\partial}{\partial x}\left(\frac{Q^2}{A}\right)}_{\substack{\text{Convective} \\ \text{Acceleration} \\ \text{term}}} + \underbrace{g\frac{\partial y}{\partial x}}_{\substack{\text{Pressure} \\ \text{force} \\ \text{term}}} - \underbrace{g(S_0}_{\substack{\text{Gravity} \\ \text{force} \\ \text{term}}} \underbrace{- S_f) = 0}_{\substack{\text{Friction} \\ \text{force} \\ \text{term}}} \tag{5.2}$$

There is increasing consensus in the literature that channel flow below bankfull

depth can be adequately described by a simplified form of 1D equations [Hunter et al.,

2007]. However, these 1D models have difficulties simulating field conditions when

more complex floodplains are considered.  Hybrid models combining the best attributes

of 1D and 2D models have been recently explored with promising results [Fewtrell et al.,

2011; Bates et al., 2010; Hunter et al., 2008; McMillan and Brasington, 2007; Bates and

De Roo, 2000]. Coupling 1D channel flow with 2D raster storage cell approximation for

the floodplain has produced models which are computationally efficient and suited to

adequately reproduce the hydrograph and inundation measurements simultaneously

[Bates and De Roo, 2000; McMillan and Brasington, 2007]. Advantages of this hybrid

solution are numerous, including: regular gridded digital elevation models (DEM) to

parameterize flows, spatial predictions comparable to similar finite element codes and much shorter runtimes [Bates and De Roo, 2000; McMillan and Brasington, 2007]. Additionally, most of the hybrid modelling studies have used cell sizes roughly equivalent to the river width, further simplifying the definition of the river channel as a chain of interconnected cells [McMillan and Brasington, 2007].

Topography represents an integral part and a major source of errors in flood modelling [Cook and Merwade, 2009]. Recent advances in remote sensing techniques have led to widespread availability of sufficiently accurate DEMs which provide nearly globally coverage. The availability of DEMs has improved significantly the ability to parameterize topographic boundaries needed for hydraulic computations [McMillan and Brasington, 2007]; however, the geometric descriptions of river bathymetry is often missing in these data. Input data requirements of existing hydrodynamic software solutions, which model river flow in 1, 2 or 3 dimensions, include: initial flow conditions, change in discharge and side flows, in channel flow characteristics such as velocity, shear stress, water level, etc., land use, and river bathymetry. All these requirements often exceed what is currently available for many Canadian watersheds.

The recognized data limitations coupled with the need for tools which allow sufficiently accurate and rapid computations have prompted this research. The objective of this paper is to illustrate methods developed to: (i) leverage open source and public domain data (e.g., DEM, IDF curves, river flow data, watershed boundaries, etc.) combined with physically based flood inundation models to compute flood hazard, and (ii) optimize computations and generate model output in near real-time. The proposed reduced complexity models are intended to fill a gap in the available applications suitable

specifically for interventions where flooding is imminent or in progress and to allow

access to otherwise complex flood hazard scenarios and in-depth knowledge of the

exposure and vulnerability to flood events to the non-expert public safety community.


## 5.3 Background

Two reduced complexity rapid flood modelling approaches are explored: (i) 0D

bathtub model; and (ii) a hybrid 1D/2D raster cell storage approach.

In the first approach, the flooded area is directly delineated from the DEM and

interpolated water surface using a simple, practical approach for rapid computation. The

flood surface is interpolated by kriging using known water levels at user provided points

or from nearby river gauges. Flood depth in the bathtub model is calculated as the

difference between the terrain elevation and the computed water surface [Yunus et al.,

2016]. Any location where terrain elevation is lower than the water surface is considered

flooded. As elevation data is the primary component, this model may be implemented in

any areas where otherwise detailed hydrological data is unavailable. The computed flood

grid is tested for hydrological connectivity. Usually, the eight-way connectivity rule is

applied: indicating a cell may be flooded if any of its eight neighboring cells is flooded

[Yunus et al., 2016]. In this study, the hydrological connectivity test is based on a cost

raster. The cost raster is initialized with a low friction value (cost unit) for cells which are

flooded and larger friction value for non-flooded cells. For each cell in the study area, the

total cost is computed by summing the cost of each cell it passes to reach the main

channel, based on flow direction computed from the DEM. All cells with total costs

larger than the Euclidean distance from cell to main channel are discarded and set as non-

flooded [Werner, 2001]. In this method flow rates and changes of flood boundary and depth over time are not considered.

The second reduced complexity model implemented is based on the model popularized by Bates and De Roo [2000]. The authors approximated the channel flow by a 1D kinematic wave approximation and treated the floodplain as 2D flooding, where the spreading is simulated using cell storage reservoirs over a raster grid. This physically based raster model solves momentum and continuity equations for the kinematic wave. In this study, a linear scheme which uses the backward-difference method is used to derive the explicit finite-difference equations in terms of space ($j$) and time ($i$) [Chow et al., 1988]:

$$Q_{i+1}^{j+1} = \frac{[\frac{\Delta t}{\Delta x} Q_i^{j+1} + \alpha\beta Q_{i+1}^{j} \left(\frac{Q_{i+1}^{j} + Q_i^{j+1}}{2}\right)^{\beta-1} * \Delta t (\frac{q_{i+1}^{j+1} + q_{i+1}^{j}}{2})]}{[\frac{\Delta t}{\Delta x} + \alpha\beta \left(\frac{Q_{i+1}^{j} + Q_i^{j+1}}{2}\right)^{\beta-1}]}$$

(5.3)

where, $Q$ is volumetric flow rate in Cartesian (x) direction at distance $x$, $\beta$ is a constant, $\alpha$ is computed from Manning's coefficient of friction *(n)*, width of channel *(w)*, and slope of channel, $q$ is lateral flow and *t* stands for time.

First, a linear feature describing the centerline of the main channel is determined from the Canadian National Hydro (NHN) datasets (geogratis.cgdi.gc.ca/). From the main channel the downstream cells are identified. Then for each channel cell the input parameters required to solve the kinematic wave approximation are identified.

Once the bankfull depth is exceeded, and/or to evacuate overland rainfall, water is routed over the floodplain by solving 2D momentum and continuity equations [Bates and De Roo, 2000]. The cell dimensions (length and width) are known from the DEM and

friction coefficient is derived based on land use class from pre-processed LandSAT

imagery (ftp.geogratis.gc.ca/pub/nrcan_rncan/vector/geobase_lcc_csc). For a given cell

in a given time step, the volume variation is computed as the sum of the fluxes of each of

the four neighbouring cells as follows:

$$\frac{dV}{dt} = Q_{up} + Q_{down} + Q_{left} + Q_{right} \tag{5.4}$$

Flow between neighbouring cells is assumed to be a function of the free surface

height difference between the cells, following discretization of continuity and kinematic

momentum, (based on Manning's Law (eq. 5.6)) [Bates and De Roo, 2000; Maugeri,

2012], Figure 5.1:

$$\frac{\Delta h^{i,j}}{\Delta t} = \frac{Q_x^{i-1,j} + Q_x^{i,j} + Q_y^{i,j-1} + Q_y^{i,j}}{\Delta x \Delta y} \tag{5.5}$$

$$Q_x^{i,j} = \pm \frac{h_{flow}^{5/3}}{n} \left( \frac{h^{i-1} - h^{i,j}}{\Delta x} \right) \Delta y \tag{5.6}$$

where, $h^{i,j}$ is the water depth at cell $i,j$, $\Delta x, \Delta y$ are the cell dimensions, $h_{flow}$ is the free

water depth between the two cells and the constant 5/3 is calculated by approximating the

channel shape as a wide, shallow rectangle [Bates and De Roo, 2000; McMillan and
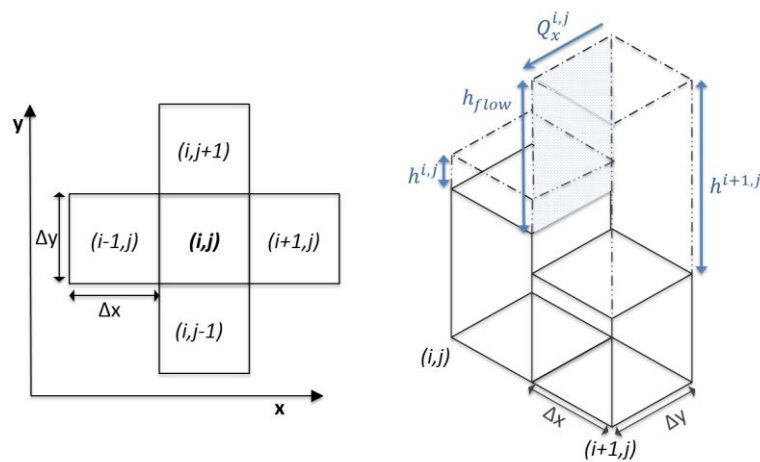
Brasington, 2007].



**Figure 5.1**  (i) discretization of floodplain cells, (ii) flow between cells on floodplain
(eq5.6), after Bates and De Roo, 2000; Maugeri, 2012

A minimum time step for the model is computed based on solution of Courant condition to maintain numerical stability [Chow et al., 1988]:

$$\Delta t \ \leq \frac{\Delta x_i}{c_k} \qquad (5.7)$$

where, $c_k$ is the kinematic wave celerity, and $\Delta x_i$ is distance. The computed time step must be less than the time necessary for a wave to travel a distance of $\Delta x_i$ to satisfy the Courant condition. In general, the explicit method is unstable unless $\Delta t$ is sufficiently small. Although the Courant condition does not guarantee stability, it does provide a guideline for a suitable time step [Chow et al., 1988]. An adaptive time step is used herein in the designed workflow to vary time throughout the duration and minimize runtime.

Assumptions of this hybrid 1D/2D method include: negligible lateral friction, flow velocity and direction controlled purely by friction coefficient and DEM elevation, and no momentum exchange between channel and floodplain, only mass [Bates et al., 2005].

## 5.4 Framework Development

The estimated flood boundary and inundation depths are computed using minimal user input as selected from intuitive drop down menus via a series of developed python scripts which reference and download appropriate online data and process it through application of hydrological processes.

### 5.4.1   0D Bathtub Model

The DEM is the primary consideration for the 0D inundation model. Estimated flood depth at a given number of points, either input by the user or extracted from nearby river gauges are the inputs to ordinary kriging algorithm and an interpolated flood surface

is generated. The DEM data is subtracted from the simulated flood surface to generate the flood inundation extent and water depths and then screened for hydrological connectivity.

**5.4.1.1 Kriging**

In this simulation, observed water level data used for kriging of the flood surface is derived from five river gauge stations in New Brunswick, where water levels were measured at the time of the 2008 flood. These observation data was then treated in a MatLab program: (i) to compute the experimental variogram, which describes the spatial tendency of the data, (ii) select an appropriate variogram model (exponential and isotropic spherical model were tested) and (iii) cross validate the resulting statistics to ensure the best model is selected to generate the flood surface. Regularly gridded (~30m) coordinates within the boundaries of the study area were input as the penultimate step to estimate values of water levels and error variance at each of these unknown cell locations. Lastly, the elevation data of each cell was subtracted from the water level surface to generate the flood grid.

Ordinary point kriging uses weighted averaging of neighboring points to estimate the value at unknown points, using a data driven weighting function, rather than an arbitrary function [Bohling, 2005; Oliver and Webster, 2014]. Kriging estimates have statistically optimal properties, producing the minimum error variances of any linear-estimation method (Snyder, 2008). Additionally, this method maximizes available measured data to help compensate for effects of clustered observations and provides a map of the uncertainty of the estimates.

In ordinary kriging, observed values for the variable $z$ at a number of known locations, $x_1, x_2, ... x_N$ are required, and any new point $x_0$, $Z$ can be predicted by:

$$\hat{Z}(x_0) = \sum_{i=1}^{N} \lambda_i Z(x_i) \tag{5.8}$$

where $\lambda_i$ are weights chosen to minimize the error variance through solution of:

$$\sum_{i=1}^{N} \lambda_i \gamma(x_i - x_j) + \psi(x_0) = \gamma(x_j - x_0) \quad for\ all\ j \tag{5.9}$$

where the sum of all weights $(\lambda_i)$ is equivalent to 1, $\gamma(x_i - x_j)$ is the semivariance

between points $i$ and $j$, $\gamma(x_j - x_0)$ is the semivariance between $j$ and the target point $x_0$,

and $\psi(x_0)$ is a Lagrange multiplier introduced for minimization of the error variance.

The mean-squared error, in terms of the variogram can be computed by Trauth et al.

[2007]:

$$E((\tilde{z}_{xo} - z_{xo})^2) = 2\sum_{i=1}^{N} \lambda_i \gamma(x_i, x_o) - \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j \gamma(x_i, x_j) \tag{5.10}$$

where $E$ is the estimation or kriging variance, $\gamma(x_i, x_o)$ is the variogram between the

observed data point and the 'unknown', $\gamma(x_i, x_j)$ represents the variogram between data

points $x_i$ and $x_j$ and $\lambda_i \lambda_j$ are the weights of the $i^{th}$ and $j^{th}$ data points. In order to

minimize eq. 5.10, a Lagrange multiplier is included, resulting in a linear kriging system

of $N + 1$ equations and $N + 1$ unknowns.


**5.4.1.2 Hydrological Connectivity**

To determine and remove any disconnected flood areas from the flood raster, total

cost of overland flow from each cell to the main channel is computed [Yunus et al.,

2016]. For each cell in the raster, a cost value is assigned. Cells which are non-flooded

receive a large friction value (cost) while cells which are classed as flooded in the

preliminary flood surface are given a low friction value. The total cost of each cell is

computed by summing up the cost of each cell it passes to reach the main river channel,

based on flow direction raster. All cells with total costs less than the Euclidean distance

between the cell and the river channel are kept, while those with higher costs are

discarded, producing the final flood surface.

### 5.4.1.3 Bathtub model flood surface

The final flooded cells are rasterized and three flood surfaces are computed: the

difference between DEM and water level estimate and the difference with the ± variance

per cell from the kriging results.

### 5.4.2   Hybrid 1D/2D cell storage model

The design framework to compute the hybrid 1D/2D flood inundation map is
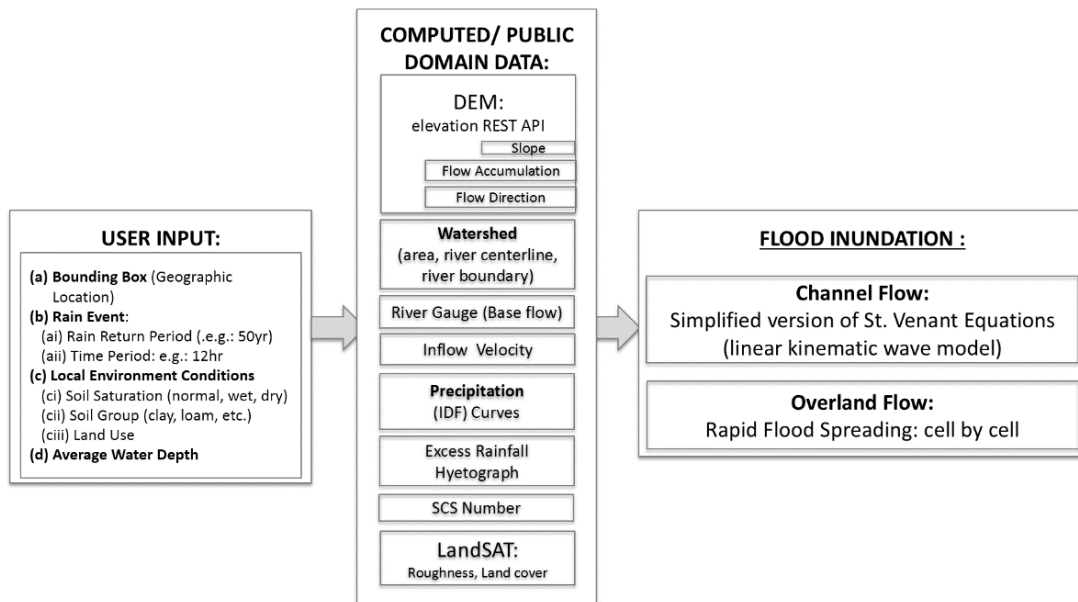
illustrated in Figure 5.2.



**Figure 5.2** Data flow diagram from user input to computed intermediate data, to computation of flood inundation surface

User input is described in four categories: (a) geographic location, (b) rain event, (c)

local environmental conditions, and (d) average river depth. The geographic location is

determined as the watershed limit provided by GeoGratis

(ftp.geogratis.gc.ca/pub/nrcan_rncan/vector/geobase _nhn_rhn/), based on the bounding

box of the map extents or may be manually selected by drawing graphics. The rain event

is selected by the user as a time period of rain, (e.g.: 5m, 1h, etc.) over a specified

recurrence interval from a series of drop-down menus. The local environmental

conditions include: condition (level of saturation), soil group (clays, loam, etc.) and land

use category (e.g.: residential, forest, farmland, etc.), which allow the program to select

an appropriate Soil Conservation Service (SCS) Curve Number (CN). Finally, since the

DEM data do not contain actual river depths, the user is prompted to input an average

depth for the river. This average river depth could be extracted from hydrographic charts,

local knowledge or other data or surveys completed in the area.

The geographic location, bounding box, is used as the extents for which to request

elevation data from Representational State Transfer (REST) Application Program

Interface (API) services, including Google, Bing, and Canadian Digital Elevation Model

(CDEM) from Natural Resources Canada. Additionally, the geographic location is used

to estimate base flow in the river, by extracting the flow level from the nearest station in

the compiled HYDAT (www.ec.gc.ca/rhc-wsc/default.asp?lang=En&n=9018B5EC-1)

database, available through the Water Survey of Canada. If the nearest station does not

contain flow, then another station, on the same river or in the same watershed is selected,

and the flow rate is computed based on the area ratio method [Mann et al., 2004]. Finally,

the geographic location is used to select the nearest station with Intensity-Duration-

Frequency (IDF) data.

### 5.4.2.1 Calculation of Upstream Boundary Conditions

- **Design Rainfall**

Statistics from the IDF files are used to generate a time varying intensity rainfall event over the given time period. Public data from the Government of Canada Engineering Climate datasets ([climate.weather.gc.ca/prods_servs/ engineering_e.html](climate.weather.gc.ca/prods_servs/engineering_e.html)) are the basis for these calculations. The nearest climate station, based on great circle distance is used, and the coefficients $A$ and $B$ are extracted based on user input of return period to compute total design rainfall using the interpolation equation:

$$R = AT^B \tag{5.11}$$

where, $R$ is the interpolated rainfall (mm/h) and $T$ is the rainfall duration (h).

- **Excess Rainfall**

Using the calculated rainfall, excess rainfall is computed using the Soil Conservation Service (SCS) Method of abstraction. The Curve Number (CN) is automatically selected based on the user input characteristics of soil type, land use and degree of saturation. Excess rainfall is computed (eq. 5.12):

$$Q(t) = \frac{(P(t) - I_a)^2}{(P(t) + S - I_a)} \tag{5.12}$$

where, $Q(t)$ is the accumulated depth of effective rainfall (mm) over time (t), $I_a$ is the initial abstraction and $S$ potential storage. The change in discharge per unit width of the river over the length of the rain event is then computed at the upstream boundary of the considered reach as follows:

$$qo = r_i * w_v * \cos(S) \tag{5.13}$$

where, $r_i$ is rainfall, $w_v$ is a constant viscosity of water, set at 20°C, S is slope. The initial depth and velocity at the upstream boundary is computed:

$$y = \frac{n^{0.6}}{s^{0.3}} \, q_o^{3/5} + Q_{gauge} \tag{5.14}$$

$$v = \frac{q_o}{y} \tag{5.15}$$

**5.4.2.2 Channel Flow**

The calculation of channel flow is governed by continuity and momentum equations for the kinematic wave as discussed in section 5.3. The backward difference method is applied to setup the finite difference equations, where $Q_{i+1}^{j+1}$ is found by substituting values of Q from the previous time step ($j$) and space ($i$). The remainder of the river channel flow is computed using this input volume and the water is propagated downstream as per eq. 5.3 [Chow et al., 1988].

**5.4.2.3 Cell storage/Overland flow**

As described in section 5.3, the raster cell storage approximation method is employed when bankfull depth is exceeded for cells in the floodplain. Data required are computed from the input DEM. The DEM was processed in the Python PyGeoProcessing routing package (pythonhosted.org/ pygeoprocessing/index.html) to fill regions in the DEM which don't drain to the edge of the dataset and to create flow accumulation and flow direction rasters. The slope is obtained from the DEM using the Numpy gradient script (www.numpy.org/).

The Python scripts are configured to write output depth and flow rasters at predefined intervals in addition to the final time step. These intermediate rasters allow the end user the ability to visualize the changes in extent and velocity of the inundation area over time.

## 5.5    Model Validation

Both models were applied to a 5km x 5km area in central New Brunswick, centered in Fredericton. This test area is bisected by the Saint John River, which is between 600 and 800m wide in this area. Fredericton is the capital of the province with a population of 94,000, approximately 22,000 households, and a long history of flooding [McGrath et al., 2015]. The primary elevation data used in these models was from the Canadian Digital Elevation Model (CDEM) ([geogratis.gc.ca](geogratis.gc.ca)), with a cell size of approximately 23m x 16m. LiDAR from the City of Fredericton was also used in the 0D model. The study area is located within a sub-watershed of 2,909km$^2$ about 2.5km from the nearest river gauge.

The historic 2008 flood was due to a late spring thaw of a heavier than normal snow pack, 50% above normal, and heavy rains causing runoff flow rates to be 400% greater than normal [Public Safety New Brunswick, 2013]. The historic flood extent from the 2008 flood event is used as the base case in the 0D bathtub model. However, in section 5.5.2, the hybrid 1D/2D model results, the historic boundary has been added only to illustrate historic boundaries and water depth, and is not used for real comparison - as the developed model does not consider snow pack melt.

### 5.5.1    0D Bathtub Results

Kriging estimates for the flood surface based on the known elevations of historic flood levels measured at five nearby gauging stations were computed using two models: isotropic spherical and exponential, through leave-one-out cross-validation [Hyndman, 2010]. Table 5.1 compares the measured and modelled flood surface values to determine the quality of the applied variogram model.

133

**Table 5.1** Cross validation statistics for applied variograms – kriging estimates at observed water depths at considered stations. Values are given in meters

| Station | Measured value | Isotropic spherical | Exponential |
|---|---|---|---|
| | | Estimated / error | |
| 1. Fredericton | 8.286 | 8.286/0.00 | 8.286/0.00 |
| 2. Upper Gagetown | 5.979 | 5.979/0.00 | 5.979/0.00 |
| 3. Gagetown | 5.499 | 5.499/0.00 | 5.499/0.00 |
| 4. Oak Point | 4.634 | 4.634/0.00 | 4.634/0.00 |
| 5. Saint John | 4.433 | -1.494/3.273 | -1.495/2.967 |
| | | | |
| Mean Error | | 0.654 | **0.593** |
| Mean Squared Error | | 2.143 | **1.760** |
| Mean Squared Deviation Ratio | | 1 | 1 |

From the results in Table 5.1, the exponential model appears to produce the lower mean error and mean square error, and was thus chosen as the preferred model for interpolation of the flood surfaces in this study. The elevation data was subtracted from the interpolated water surface for three flood levels: (i) derived from the kriged surface (expected flood level), (ii) minimum boundary computed as the expected flood level minus the error variance at each cell, and maximum boundary computed as expected flood level plus error (Figure 5.3, Figure 5.4).

DEM data obtained from two elevation providers were tested with the kriged water surface to evaluate the accuracy of the derived flood surface, CDEM and LiDAR. The resulting flood surfaces were compared to the observed flood boundary of the 2008 flood event to determine the goodness of fit of the simulated inundation extent. Three profile lines were generated along the northern shore of the Saint John River to compare the water depth for each of the surfaces. In Figure 5.4 are given the mean flood depth along three selected profiles computed using LiDAR and Geogratis elevation data together with the water depth from the historic 2008 flood. It can be observed that the LiDAR water depth data lines up fairly well with the historic flood event, while the CDEM under-

estimates flooding from shoreline inland, with a maximum difference in water depth of

5.6 in profile 2 (Figure 5.4). On average, in all three profiles the difference is 2.38m for

using the CDEM. Over 70% of the mean LiDAR flood surface data points along all three

profile lines are within ~30cm of the water depth from the historic flood grid, indicating a

positive correlation, though there are areas along the profile where there are

discrepancies. A comparison of the CDEM water depth data to the historic flood grid

levels indicates that 10% of the water depths fall within 30cm of historic values and

approximately 42% fall within 2m of those historic values. The LiDAR data illustrates

that this method, when used with accurate elevation data can compute a good

approximation of the flood extent, including areas which may be flooded by backwater

effects. With lower quality elevation data, relative to both cell size and elevation

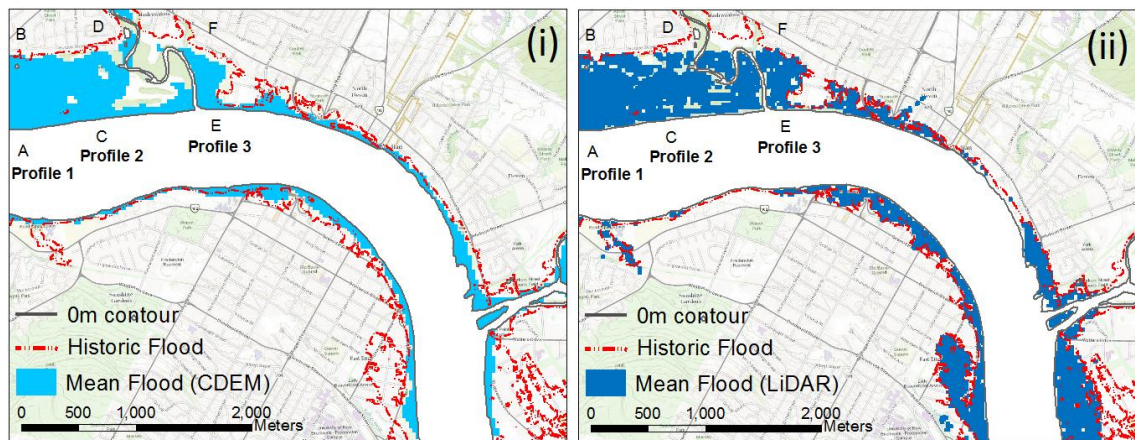accuracy, the quality of the computed flood grid is inferior.



**Figure 5.3** Flood area generated via kriging, mean flood extent using (i) CDEM and (ii) LiDAR DEM surfaces
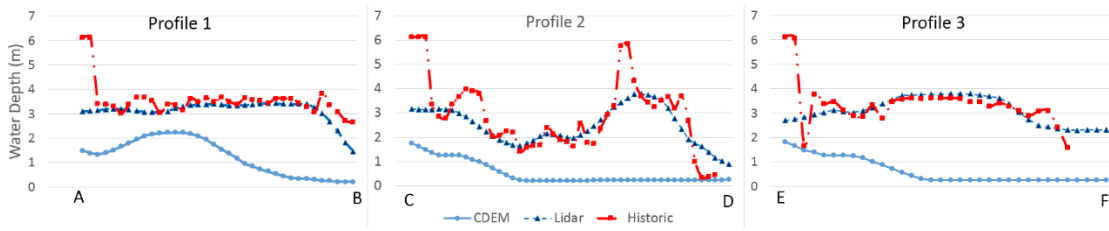
**Figure 5.4** Profile of Flood surface for three x-sections on the northern shore of the Saint John River, profiles reference locations from Figure 5.3.

Spatial statistics were derived using a fit measure, as suggested by Bates and deRoo [2000]. In this approach the observed inundated area from the historic flood ($A_o$) is compared to that of the predicted by the model ($A_p$) and the overlapping area ($A_{op}$):

$$F = \left(\frac{A_{op}}{A_o + A_p - A_{op}}\right)$$
(5.16)

A value of F=1, indicates a perfect match between predicted and observed areas, while lower F values indicates larger discrepancies. In the case where the LiDAR data was the base elevation, the computed flooded area has F value greater than 0.98, indicating near perfect match to the extents of the historic flood, Table 5.2. The flood grid based on CDEM elevation data, has computed F values between 0.77 - 0.80 based on the minimum to maximum flood prediction.

**Table 5.2:** Fit Measure comparing flood surface generated with Bathtub model using (a) CDEM and (b) Lidar data against flood surface from historic 2008 flood in Fredericton

|  | **CDEM** | **LiDAR** |
| --- | --- | --- |
| **Estimate - Error** | 0.7737 | 0.9895 |
| **Estimate** | 0.7853 | 0.9996 |
| **Estimate + Error** | 0.7946 | 1.00 |

Computation time for this method in the tested study area, ~5km x 5km, is under 2 minutes on a 2.70GHz, 4 core Windows computer.

### 5.5.2    Hybrid 1D/2D cell storage model

A 25yr, 1 hour rain event is modeled according to the above described hybrid 1D/2D framework. The results of this simulation are compared to those derived from the input data being reconfigured and input to the LISFLOOD-FP code (bristol.ac.uk/geography/research/ hydrology/models/lisflood/). LISFLOOD-FP is a non-commercial research software developed at the University Of Bristol (UK) based on the simple physically based flood model proposed by Bates and De Roo [2000]. This application is "specifically designed to simulate floodplain inundation in a computationally efficient manner over complex topography" [University of Brisol, 2016]. The primary component of LISFLOOD-FP is a raster DEM which has sufficient accuracy to identify the channel and floodplain topography [Bates and De Roo, 2000]. This method is as described in section 5.2, with respect to 1D channel flow and 2D raster storage cell approximation.

The baseflow, rainfall and change in river discharge are computed as described in Section 5.4.2, Figure 5.5. The time step computed to satisfy the courant condition was 0.96s. The simulation took about 22min to complete on a 2.70GHz, 4 core Windows computer.
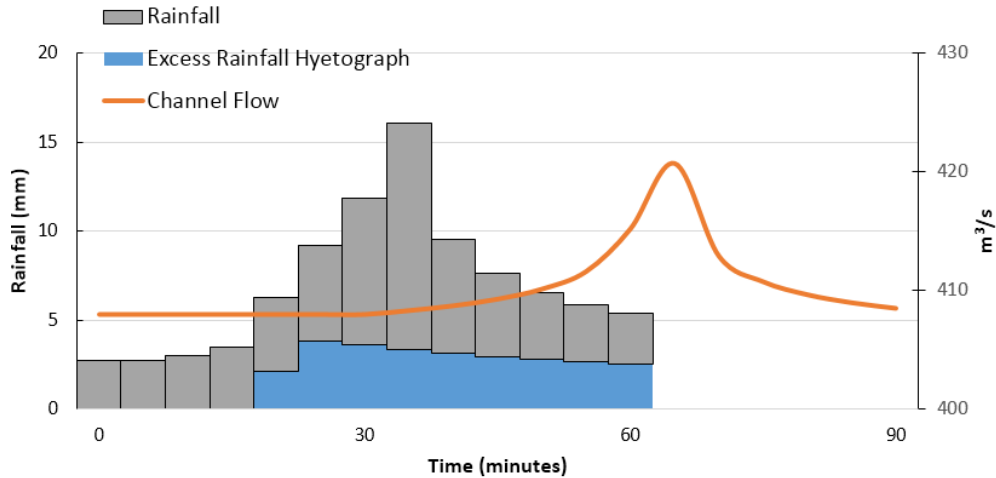
**Figure 5.5** Computed rainfall for 25yr, 1hr event, excess rainfall hyetograph and River flow for study area location

Results from the developed model is compared to that derived using LISFLOOD-FP. The historic flood boundary is shown for illustration purposes, as this historic event included melt of considerable size snow pack which is not incorporated in the tested models.

It's expected, that the simulated event would have lower flow rates and therefore shallower water depths that the historic event. Given this magnitude of rain fall across the study region, the flooded area in the LISFLOOD-FP scenario seems to overestimate the distribution of flooding, on both the northern and southern banks of the river, as is illustrated in Figure 5.6 and the profiles in Figure 5.7. The LISFLOOD-FP solution, at the 1hr time interval, over 45% of the flooded cells have computed water depth exceeding 5 m, additionally, there is considerable water depth in cells on the south side of the river, extending beyond the historic flood boundary. In the LISFLOOD-FP solution, approximately half of the cells along the profile from shoreline to 2008 flood boundary over-estimate the water depth.

The flood boundary and water depths appear more realistically presented by Figure 5.6(i), the designed model, in reference to the historic flood event. Looking at the profiles for the developed model, the first ~100m from the shoreline has computed water depths exceeding the historic flood, while the remainder has lower water depths. Given the steepness of the curve, especially in Profile 1 and 2, Figure 5.7, this may be due to water not properly being transferred to cell storage at the interface between streamflow and floodplain flow. However, the differences may also be due to the accuracy and cell size of the DEM used.
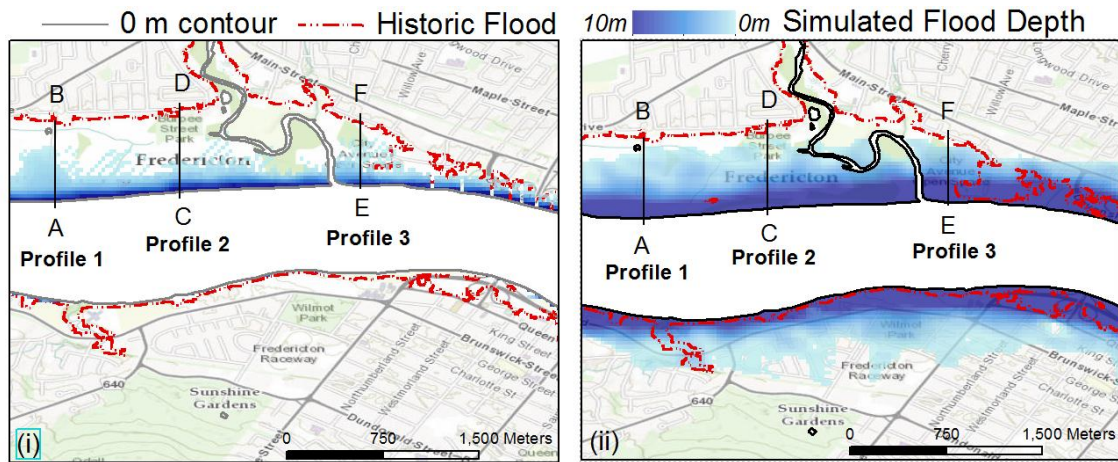


**Figure 5.6** Flood inundation after 1hour, (i) designed algorithm, (ii) LISFLOOD-FP. Red line represents historic flood boundaries and lines 1,2 and 3 profile lines for Figure 5.7.
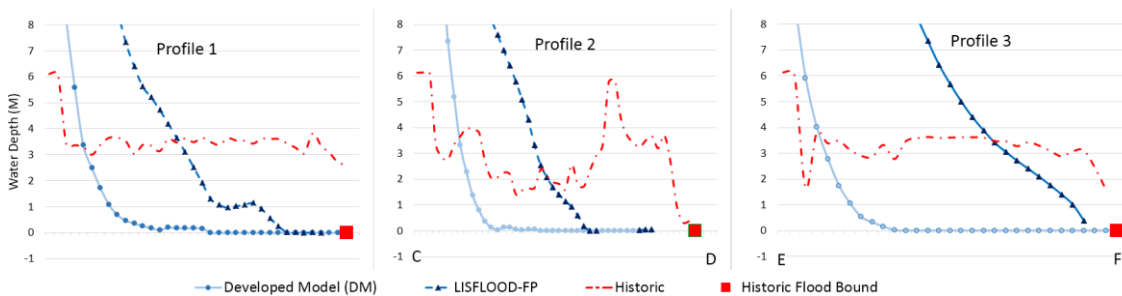


**Figure 5.7.** Profile of Flood surface for three x-sections on the northern shore of the Saint John River, profiles reference locations from Figure 5.6.

**Table 5.3:** Fit measure (F) of proposed algorithm to LISFLOOD generated flood grid.

|  | F |
|---|---|
| **Flooded Area after 1 hr** | 0.2825 |
| **Time Step:** | |
| **15min** | 0.3018 |
| **30 min** | 0.2776 |
| **45 min** | 0.1717 |

The computed Fit measure (F) between the two algorithms is quite poor – which is expected based on visual comparison of Figure 5.6 and the profiles in Figure 5.7, Table 5.3. The fit measure ranges from a maximum F = 0.3018 at 15 minutes into the simulation and lowest F = 0.1717 at 45 minutes.

## 5.6 Conclusions

The objective of this paper was to present ongoing work on the development of methods and tools for rapid and user friendly modelling of flood surfaces and intended for the non-expert public safety community. Reduced complexity models fill a gap in the available applications suitable specifically for interventions where flooding is imminent or in progress. These tools aim to allow access to otherwise complex flood hazard scenarios and in-depth knowledge of the exposure and vulnerability to flood events to the non-expert public safety community. This research focused on limited user input and basic hydrologic principles and processes to compute flood inundation maps. Open source and public domain data was leveraged to provide the necessary input data and combined with physically based flood inundation models.

Two reduced complexity models were tested to generate model outputs in near-real-time by users with limited knowledge. The first reduced complexity method, based on the

'bathtub' model utilizes measured river level data from nearby river gauges or user input data at a number of more locations. An inclined flood surface was computed via ordinary kriging from which the DEM data was subtracted. The resulting flood inundation surface was then screened for hydraulic connectivity and rebuilt. Results from this method indicate Fit values of $0.98 - 1.0$ when using high-resolution LiDAR data and $F \approx 0.8$ with the Canadian national elevation dataset, CDEM, when compared to a historic flood grid. Processing time for this solution was minimal, i.e., under 2 minutes.

The second model, a hybrid 1D/2D solution simulating a probabilistic rain event and simulation of flow in channel using simplified solution of momentum and continuity equations for the kinematic wave and a 2D raster cell overland storage method. This model provides output at numerous time steps through the simulation, thus allowing visualization of changing depth and velocity through time and space. This model is based on an iterative method using small time steps, thus time to run is greater than the former model, ~22min for the 5km x 5km study area. The proposed model did not fit the LISFLOOD-FP solution well at all time steps with greater divergence as time passed. However, the results from this developed method which leverages open data presented a more realistic match when viewed in comparison to a historic flood event.

Suggested improvements to the above framework include integration of the Intensity-Duration-Frequency Curves under Climate Change, available through University of Waterloo (idf-cc-uwo.ca/default.aspx), enhancing the hybrid model to include D-infinity flow directions, and significant refinements to the algorithm to improve speed for the developed hybrid model. Furthermore, adding in capability to

include melting of snow pack is of interest as this is a concern for many Canadian communities.

## ACKNOWLEDGEMENTS

## REFERENCES

Bates, P. D., Horritt, M., Wilson, M., and Hunter, N. (2005). LISFLOOD-FP user manual and technical note. University of Bristol, UK.

Bates, P., and De Roo, A. P. J. (2000). "A simple raster-based model for flood inundation simulation." *Journal of Hydrology, 236*(1-2), 54-77.

Bates, P. D., Horritt, M. S., and Fewtrell, T. J. (2010). "A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling." *Journal of Hydrology, 387*(1–2), 33-45. doi:http://dx.doi.org.proxy.hil.unb.ca/10.1016/j.jhydrol.2010.03.027

Bohling, G. (2005). *KRIGING (C&PE 940).* Unpublished manuscript.

Chow, V. T., Maidment, D. R., and Mays, L. W. (1988). *Applied hydrology*. New York: McGraw-Hill.

Fewtrell, T. J., Duncan, A., Sampson, C. C., Neal, J. C., and Bates, P. D. (2011). "Benchmarking urban flood models of varying complexity and scale using high resolution terrestrial LiDAR data." *Physics and Chemistry of the Earth, 36*(7-8), 281-291.

Hunter, N. M., Bates, P. D., Horritt, M. S., and Wilson, M. D. (2007). "Simple spatially-distributed models for predicting flood inundation: A review." *Geomorphology, 90*(3-4), 208-225.

Hunter, N. M., Bates, P. D., Neelz, S., Pender, G., Villanueva, I., Wright, N. G., Mason, D. C. (2008). "Benchmarking 2D hydraulic models for urban flooding." *Proceedings of the Institution of Civil Engineers: Water Management, 161*(1), 13-30.

Hyndman, R. J. (2010). Why every statistician should know about cross-validation. Retrieved from http://robjhyndman.com/hyndsight/crossvalidation/

Kulkarni, A. T., Mohanty, J., Eldho, T. I., Rao, E. P., and Mohan, B. K. (2014). "A web GIS based integrated flood assessment modeling tool for coastal urban watersheds." *Computers and Geosciences, 64*, 7-14.

Leskens, J. G., Brugnach, M., Hoekstra, A. Y., and Schuurmans, W. (2014). "Why are decisions in flood disaster management so poorly supported by information from flood models?" *Environmental Modelling and Software, 53*, 53-61.

Mann, M. P., Rizzardo, J., and Satkowski, R. (2004). *Evaluation of methods used for estimating selected streamflow statistics, and flood frequency and magnitude, for small basins in north coastal California.* US Department of the Interior, US Geological Survey.

Maugeri, A. (2012). Capabilities of a coupled 1D/2D model for flood inundation simulation. Columbia Water Center.

McGrath, H., Stefanakis, E., and Nastev, M. (2015). "Sensitivity analysis of flood damage estimates: A case study in Fredericton, New Brunswick." *International Journal of Disaster Risk Reduction, 14, Part 4*, 379-387. doi:http://dx.doi.org/10.1016/j.ijdrr.2015.09.003

McMillan, H. K., and Brasington, J. (2007). "Reduced complexity strategies for modelling urban floodplain inundation." *GEOMOR Geomorphology, 90*(3), 226-243.

Oliver, M. A., and Webster, R. (2014). "A tutorial guide to geostatistics: Computing and modelling variograms and kriging." *Catena, 113*, 56-69. doi:http://dx.doi.org.proxy.hil.unb.ca/10.1016/j.catena.2013.09.006

Public Safety New Brunswick. (2013). Public safety. Retrieved from http://www2.gnb.ca/content/gnb/en/departments/public_safety.html

Snyder, D. T. (2008). *Estimated Depth to Ground Water and Configuration of the Water Table in the Portland, Oregon Area.*

University of Brisol. (2016). LISFLOOD-FP. Retrieved from http://www.bristol.ac.uk/geography/research/hydrology/models/lisflood/

Werner, M. G. (2001). "Impact of grid size in GIS based flood extent mapping using a 1D flow model." *Physics and Chemistry of the Earth Part B Hydrology Oceans and AtmosphereE, 26*(7-8), 517-522.

Yunus, A. P., Avtar, R., Kraines, S., Yamamuro, M., Lindberg, F., and Grimmond, C. (2016). "Uncertainties in tidally adjusted estimates of sea level rise flooding (bathtub model) for the greater London." *Remote Sensing, 8*(5), 366.

# 6    Summary and Conclusions

The existing sophisticated tools for flood risk analysis are generally not well suited for rapid flood predictions and they often exceed the data available for any given community; as well, they are ill suited for application by the non-expert public safety community. The purpose of this research was to develop standardized methods and tools which help end users from the public safety community run their own flood scenarios and prepare informed emergency response and long term mitigation plans. The primary focus and contribution of this research is in the development of standardized risk computation methods transformed into a set of user-friendly tools considering limited user input, basic hydrologic principles and processes whilst leveraging open source data.

## 6.1 Summary of Research

The preparatory steps of this research were to: (i) investigate existing software solutions available which combine both computation of the flood hazard and assessment of risk and (ii) to test the sensitivity and relative impacts on the resulting losses due to parameter changes, including: selection of depth-damage function, input flood level, and restoration duration, Chapter 2. Once software options and limitations were reviewed and the relative impact of parametric changes and data were identified, applications were developed which required limited user input to compute damages, Chapter 3, and simulate flood hazard, Chapter 4 and Chapter 5.

Chapter 3 introduced the newly developed tool referred to as Rapid Risk Evaluation ($ER^2$). $ER^2$ was originally programmed as a MS Excel worksheet application,

designed specifically for users to input necessary information on a building-by-building or aggregate scenario basis and respective measured or predicted water levels. The outputs of this application include both exposure and estimates of potential damages and losses - expressed as a percentage of the reconstruction costs. MS Excel was chosen as an initial platform as it is one of the most common spreadsheet applications in use today and therefore has widespread familiarity.

The primary input data considered in this research is in the form of open source and public domain data. As elevation data is an integral, and probably most important, part of flood modelling, a novel method of DEM fusion was developed, Chapter 4, with the objective to create a new, better quality DEM from multiple input DEMs freely available from REST API services. The proposed novel method is based on concepts of clustering and inverse distance weighting and requires no input by the user in selection of clustering criteria or weighting. By using REST API services, the point elevation data is downloaded on-demand and able to be manipulated, as needed – by fusion technique or by users wanting to test out mitigation measures, such as the effect of sandbagging or installing other temporary flood barriers - before the final increased accuracy DEM is generated.

The objective of Chapter 5 was to illustrate how public domain data can be leveraged and combined with physically based flood inundation models to compute flood hazard and how to optimize computations and rapidly generate model outputs by users with limited expertise. Two techniques of reduced complexity flood models were explored: 0D bathtub model and a hybrid 1D/2D raster cell storage approach. In the 'bathtub' model, users select and set expected/historic water levels at several locations.

The flood inundation surface is then computed using kriging. Next, the flooded area is computed by subtracting the terrain elevation data from the computed water surface. To test the flood grid for hydrological connectivity, a cost raster was created and the 'total cost' of travel from each cell to the main channel was computed - keeping only those flooded cells whose total cost was less than the Euclidean distance to the main channel. The second method considers a hybrid 1D/2D flood model developed, assuming channel flow approximated by 1D kinematic wave approximation of shallow water equations, and the floodplain is treated using cell storage reservoirs over a raster grid. The user input is limited to: geographic location, precipitation rate and local runoff conditions. The rain event selected by the user as a time period of a given precipitation rate over a specified recurrence interval and the local runoff conditions include: level of saturation of soils, soil group and land use category, all selected from a series of intuitive drop-down menus.

## 6.2 Achievements of Research

- Development of a relatively simple flood risk assessment method programmed in MS Excel worksheet allowing any user with any level of expertise the capability of estimating building exposure and vulnerability due to a flood event. This worksheet application was subsequently implemented as an API, capable of working with any web based request. The application has been viewed and downloaded about 200 times by users in over 35 countries around the globe, including: Canada, US and Germany. Additionally, this application has been used by Canadian insurance companies, including Deloitte (Sharma, 2015). The website hosting the MS Excel and API applications has

147

been visited more than 2,000 times and is a successful example and a showcase of the University of New Brunswick's expertise in the domain.

- Novel DEM Fusion technique which improved the quality of the elevation data available from REST API services. A multi-step process was applied to generate the fusion DEM which required no user input, other than geographic extent.

- Simulation of flood hazard by non-expert users, leveraging public domain data. Development of a method of 0D and 1D/2D flood modeling which required limited user input and no hydrologic expertise to run. In the hybrid 1D/2D model, users select the study area/watershed, and input details about the rain event and the local environment from dropdown menus. The remainder of the application automatically extracts the appropriate files from pre-defined internet resources and simulates 1D flow in the channel and 2D flow over the floodplain. The program generates spatial variations of depth of water and velocity at multiple time steps, thus allowing users to visualize changes over both space and time. In the 0D model, using LiDAR DEM the model is exemplary in its' ability to reproduce the historic flood event, while the coarser DEM produces ~80% match.

- The DEM fusion and flood mapping tools, written in Python, are considered primary components to the flood risk application framework. The loss-estimation module, originally written in MS Excel is also available as an API. This API, combined with a web mapping application referencing inventory data and the DEM fusion and flood mapping tools allows users with limited knowledge to simulate flood hazard and assess risk.

## 6.3 Limitations and Recommendations for Future Work

While every effort was made to test these developed applications against real-world scenarios and validate the generated results against actual claims and reported damage, access was ultimately not granted as of publication of this dissertation. Instead, to validate the results of the ER$^2$ spreadsheet, it was tested against a solution obtained from Hazus Canada analysis. The need to test these results against actual claims is still an outstanding goal in order to calibrate the model and assign quality or confidence in the derived flood surface and risk estimate. By comparing the results of the developed tools to historic claims, in the tested areas and from other Canadian cities we can begin to establish a confidence interval and accuracy for our results based on different criteria, such as building type or building age for risk. Enhancements to the ER$^2$ application include adding additional depth-damage curves considering flow velocity, flood duration, building foundation type and their incorporation into the damage estimates.

With respect to the DEM fusion method, future recommendations include analysis of other geographic locations to validate the encouraging findings from the two NB communities. Furthermore, additional recommendation for future research of the proposed fusion method is testing the effect of maintaining three dimensions in the clustering algorithm.

In Chapter 5, two reduced complexity flood inundation methods were tested. Enhancements to this portion of the research include optimizing the codes to improve efficiency and further reduce the processing time. Validation in other locations susceptible to riverine flooding is of interest as well as comparison to historic events to better assess the potential accuracy. Additionally, accessing results from the University of

Waterloo online tool which develops IDF curves under climate change would be a great

addition as well as input of snow pack age and depth to better represent spring thaw and

ice jamming.

## 6.4 Conclusion

By combining the tools discussed in Chapters 3 through 5 into a single application,

such as a web map, a user with limited expertise can simulate a flood hazard, based on

limited input via a series of intuitive drop-down menus. The future web application, will

access the DEM fusion tool which extracts elevation data and fuses the multiple

providers datasets into a single, better quality DEM. Next, based on the details input per

the simple flood or rain event a flood inundation map may be computed. Finally, the user

can compute flood risk which uses embedded (or uploaded) inventory data, the computed

flood grid depths and the $ER^2$ API to create data layers which can be themed and overlaid

on the map.

## REFERENCES

Sharma, M. (2016, January 26). Email.

# APPENDIX I

**ELSEVIER LICENSE
TERMS AND CONDITIONS**

Dec 11, 2016

This Agreement between Heather McGrath ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

| | |
|---|---|
| License Number | 4005970007574 |
| License date | Dec 11, 2016 |
| Licensed Content Publisher | Elsevier |
| Licensed Content Publication | International Journal of Disaster Risk Reduction |
| Licensed Content Title | Sensitivity analysis of flood damage estimates: A case study in Fredericton, New Brunswick |
| Licensed Content Author | H. McGrath,E. Stefanakis,M. Nastev |
| Licensed Content Date | December 2015 |
| Licensed Content Volume Number | 14 |
| Licensed Content Issue Number | n/a |
| Licensed Content Pages | 9 |
| Start Page | 379 |
| End Page | 387 |
| Type of Use | reuse in a thesis/dissertation |
| Portion | full article |
| Format | both print and electronic |
| Are you the author of this Elsevier article? | Yes |
| Will you be translating? | No |
| Order reference number | |
| Title of your thesis/dissertation | An Application Framework to Support Web-Based Flood Inundation Mapping and Risk Assessment: Leveraging Open Data |
| Expected completion date | Jan 2017 |
| Estimated size (number of pages) | 135 |
| Elsevier VAT number | GB 494 6272 12 |
| Requestor Location | Heather McGrath 668 Philip St

CANADA Fredericton, NB e3a2y4 Canada Attn: Heather McGrath |
| Total | 0.00 USD |
| Terms and Conditions | |

151

# APPENDIX II

## License and copyright agreement

OPEN ACCESS

The following license and copyright agreement is valid for any article published by Copernicus Publications whose original manuscript was received from **10 December 2007** on.

### Author's certification

In submitting the manuscript, the authors certify that:

They are authorized by their co-authors to enter into these arrangements.

The work described has not been published before (except in the form of an abstract or as part of a published lecture, review or thesis), that it is not under consideration for publication elsewhere, that its publication has been approved by all the author(s) and by the responsible authorities – tacitly or explicitly – of the institutes where the work has been carried out.

They secure the right to reproduce any material that has already been published or copyrighted elsewhere.

They agree to the following license and copyright agreement:

### Copyright

Copyright on any article is retained by the author(s). Regarding copyright transfers please see below.

Authors grant Copernicus Publications a license to publish the article and identify itself as the original publisher.

Authors grant Copernicus Publications commercial rights to produce hardcopy volumes of the journal for sale to libraries and individuals.

Authors grant any third party the right to use the article freely as long as its original authors and citation details are identified.

The article is distributed under the Creative Commons Attribution 3.0 License. Unless otherwise stated, associated published material is distributed under the same licence.

### Creative Commons Attribution 3.0 License

Anyone is free:

**to Share** — to copy, distribute and transmit the work

**to Remix** — to adapt the work

Under the following conditions:

**Attribution** — The original authors must be given credit.

For any reuse or distribution, it must be made clear to others what the license terms of this work are.

Any of these conditions can be waived if the copyright holders give permission.

Nothing in this license impairs or restricts the author's moral rights.

The full legal code of this license.

### Copyright transfers

Many authors have strict regulations in their contract of employment regarding their works. A transfer of copyright to the institution or company, as well as the reservation of specific usage rights, is typical. Please note that in the case of open-access publications in combination with a Creative Commons License, a transfer of the **copyright** to the institution is possible, as it belongs to the author anyway and is not subject to the publisher.

Any **usage rights** are regulated through the Creative Commons License. As Copernicus Publications uses the Creative Commons Attribution 3.0 License, anyone (the author, his/her institution/company, the publisher, as well as the public) is free to copy, distribute, transmit, and adapt the work as long as the original author is given credit (see above). Therefore, specific usage rights cannot be reserved by the author or his/her institution/company, and the publisher cannot include a statement "all rights reserved" in any published paper.

A copyright transfer from the author to his/her institution/company will be expressed in a special "Copyright Statement" at the end of the publication rather than on the first page in the article citation header. Authors are asked to include the following sentence: "The author's copyright for this publication is transferred to *institution/company*".

### Crown copyright

The license and copyright agreement of Copernicus Publications respects the Crown copyright. For works written by authors affiliated with the British Government and its institutions, a copyright statement will be included at the end of the publication.

# GLOSSARY

*Exposure*: the assets and values located in flood prone areas [Koks et al., 2015]

*Vulnerability*: potential for loss of property or life from environmental hazards [Cutter, Mitchell, and Scott, 2000]

*Social Vulnerability:* used to define the susceptibility of social groups to potential losses from hazard events or societies resilience to hazards [Cutter et al., 2000]

*Risk*: likelihood of the event occurring and includes: potential source, impact of the risk, and estimate of its frequency of occurrence

*Mitigation*: suite of efforts to reduce risk or lessen their impacts

# Curriculum Vitae

**Candidate's full name**: Heather Dana McGrath

**Universities attended**:

2014 - 2017: PhD Candidate, Geodesy and Geomatics Engineering, University of New Brunswick, New Brunswick Canada

2012 - 2014: MScE, Geodesy and Geomatics Engineering, University of New Brunswick, New Brunswick, Canada

2000 - 2001: Advanced Certificate (Marine Geomatics), Center of Geographic Sciences, Nova Scotia, Canada

1996 - 2000 B.Sc. Geography and GIS (minor), McMaster University, Ontario, Canada

**Publications:**

**Peer-Reviewed Journal Papers:**

1. **McGrath, H**., Stefanakis, E., Nastev, M. (2016), *A Framework for Reduced Complexity Flood Modelling: Leveraging Open Data and Limited User Input (*under Review*)*

2. **McGrath, H**., Stefanakis, E., Nastev, M. (2016), *DEM Fusion of Elevation REST API Data in Support of Rapid Flood Modelling*. GEOMATICA Vol. 70, No. 4, 2016

3. Nastev, M., Abo El Ezz, A., Nollet, M.J., Smirnoff, A., Ploeger, S.K., **McGrath, H**., Sawada, M. and Stefanakis, E. (2015) *Methods and Tools for Natural Hazard Risk Analysis in Eastern Canada: Using Knowledge to Understand Vulnerability and Implement Mitigation Measures*. ASCE Nat. Hazards Rev., 10.1061/(ASCE)NH.1527-6996.0000209 , B4015002.

4. **McGrath, H.**, Stefanakis, E., & Nastev, M. (2016). *Rapid Risk Evaluation (ER2) Using MS Excel Spreadsheet: a Case Study of Fredericton (New Brunswick, Canada).* ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, 27-34. http://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/III-8/27/2016/, doi:10.5194/isprs-annals-III-8-27-2016

5. **McGrath, H.**, Stefanakis, E., Nastev, M. (2015) *Sensitivity Analysis of Flood Damage Estimates: A Case Study in Fredericton, New Brunswick*, International

**Conference Papers/Presentations**

1. Nastev, M., Nollet, M.J., Abo El Ezz, A., Carrier, M.-A., Smirnoff, A., **McGrath, H**., Stefanakis, E., and Parent, M*., Interactive Web-Based Application for Seismic and Flood Risk Assessment,* World Engineering Conference on Disaster Risk Reduction (WECDRR2016), Dec. 5-6, 2016, Lima-Peru.

2. **McGrath, H**., Stefanakis, M., & Nastev, M. (2016). *Accuracy Assessment of Elevation REST APIs*, Geomatics Atlantic Conference, Sept. 26-28, 2016 Fredericton, NB.

3. **McGrath, H.,** Stefanakis, E., Nastev, M. (2015) *Flood risk assessment using MS Excel spreadsheet: A case study of Fredericton (New Brunswick)*, International Society for Photogrammetry and Remote Sensing, 2016, July 12-19 2016 Prague: http://www.isprs2016-prague.com

4. **McGrath, H**., Stefanakis, M., & Nastev, M. (2015). *Flood Inundation Maps Using Reduced Complexity Models*. 22$^{nd}$ Canadian Hydrotechnical Conference (CHC), April 29-May1, 2015 Montreal QC, extended abstract: http://registration.cgi-pco.com/CSCEhydro/pdf/abstracts/Session3/3D.zip

5. **McGrath, H**., Stefanakis, M., & Nastev, M. (2015). *Flood Risk Assessment Using MS Excel Spreadsheet: A Case Study of Fredericton New Brunswick*. 22$^{nd}$ Canadian Hydrotechnical Conference (CHC), April 29-May1, 2015 Montreal QC, (poster)

6. **McGrath, H**., *Flood Risk Assessment*, ESRI Canada Fredericton User Conference, Fredericton NB, Oct 29, 2014. Presentation online: http://www.esri.ca/sites/default/files/filemanager/Training_and_Events /Conference_Proceedings/2014/Fredericton/HMcGrath_FloodRiskAssessment.pdf

7. **McGrath, H**., Stefanakis, M., & Nastev, M. (2014). *Sensitivity Analysis of Flood Damage Estimates: A Case Study in Fredericton, New Brunswick*. In Proceedings of Canadian Risk and Hazards Network (CRHNet). 22-24 Oct. 2014. Toronto, ON.

8. **McGrath, H.,** Stefanakis, E. (2014). *Development of Data Warehouse for Riverine and Coastal Flood Risk Assessment*. In the Proceedings of the Joint International Conference on Geospatial Theory, Processing, Modeling and Applications (ISPRS/IGU),6-8 Oct 2014. Toronto ON. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-2, 41-48,doi:10.5194/isprsarchives-XL-2-41-2014, 2014

9. **McGrath, H.,** Stefanakis, E., McCarthy, M., & Nastev, M. (2014). *Data Preparation for Validation Study of Hazus Canada Flood Model*. In Randy R. Rapp & William Harland (Eds.), The Proceedings of the 10th International Conference of the International Institute for Infrastructure Resilience and Reconstruction (I3R2) 20-22 May 2014. (14-21). West Lafayette, Indiana: Purdue University.

10. **McGrath**, **H**., Stefanakis (2014) *Flood Mapping and Loss Estimation for Riverine and Coastal Flood Risk Management*. Water Initiative for the Future (WaTIF). 4-6 May 2014. Kingston, ON.

*Other*

1. **McGrath, H.**, Stefanakis, E. , Nastev, M. (2015) *Mapping Flood Risk and Flood Damage Estimates*, Cartouche, #90, Winter/Spring 2015 page 10-11

2. **McGrath, H.,** Stefanakis, E., Nastev, M. (2015) *Flood risk assessment in Fredericton using MS Excel spreadsheet*. Presented online: CANHUG Meeting, Jan 2015.

3. **McGrath, H.,** Stefanakis, E., Nastev, M. (2014) *Hazus-MH for New Brunswick.* Presented online: CANHUG Meeting February 2014