

AUTOMATING GIS INPUT FOR DISTRIBUTED URBAN DRAINAGE MODELLING

by

Saalih Mohamed Shamead  
Specialized Honours Bachelor of Environmental Studies (Environmental Management)  
York University  
Toronto, Ontario, Canada  
June, 2012

A thesis

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Master of Applied Science (MASc)

in the Program of

Environmental Applied Science and Management

Toronto, Ontario, Canada, 2014

© Saalih Mohamed Shamead, 2014

## **Author's Declaration**

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

## **Abstract**

### AUTOMATING GIS INPUT FOR DISTRIBUTED URBAN DRAINAGE MODELLING

Master of Applied Science (MASc), 2014

Saalih Mohamed Shamead

Environmental Applied Science and Management, Ryerson University

The development of distributed urban drainage models is becoming more important as cities prepare for the challenges associated with climate change such as more intense precipitation events (McCarthy *et al.* 2010; Allan, 2011; Simões *et al.* 2011; Blumensaat *et al.* 2012; Leitão *et al.* 2012). GIS-based tools were developed to generate input datasets for a 1-D distributed urban drainage model for part of Toronto's combined area, resulting in an efficient model development process compared to those utilizing manual approaches. These automatic GIS-based tools included the delineation of Wet Weather Flow (WWF) subcatchments (stormwater) and Dry Weather Flow (DWF) subcatchments (sanitary). It also included the determination of the intensity of rainfall on a more detailed scale than the coarse coverage provided by the City's rain gauges and the traditional Thiessen polygon interpolation method. Through testing the new tools designed in ModelBuilder, it was determined that 66% and 52% of DWF and WWF subcatchments respectively, were automatically delineated to a degree where they would be "Acceptable" for input into the urban drainage model, InfoWorks CS. Although the rainfall tools were able to continuously interpolate measured rainfall (on a seemingly unprecedented basis), and generate over 700 virtual rain gauges, the validity of the approach remains imperfect due to irresolvable inconsistencies between the City's gauges and those used for validation purposes.

## **Acknowledgements**

I would first like to acknowledge my two supervisors Dr. Doug Banting and Dr. James Li who provided me with this great opportunity and for all their assistance and advice along the way.

I would like to acknowledge the support provided by City staff at Toronto Water especially Grace Lin, Xueying Liu, Jaime Aldana, Ilona Tarvydas, Huu-Thoi Nguyen, Kishan Ramkhelawon, Milan Jekic, Tigist Gebretsadik and Dr. Jian Lei. The support provided by other members of the Ryerson University community including Dr. Celia Fan, Dr. Weihua Cao, Michael MacDonald, Kay Kang, Betsy Zheng, and Jessica Carvalho. The support of Jamie Duncan from the TRCA for the rain gauge monitoring site records provided. I would also like to thank my examiners for their comments.

I would also like to acknowledge the financial support of the Tri-Council Agencies SSHRC and NSERC, the Ontario Graduate Scholarship program, Toronto Water, and Ryerson University through academic scholarships. Finally, would like to acknowledge ESRI Canada for the motivation to pursue research in this field through being a recipient of their annual ESRI Canada GIS Scholarship program.

## **Dedication**

To the Almighty, without whose help, none of this would be possible. I also dedicate this work to my family, whose support I have been able to rely upon since the beginning of this journey.

## Table of Contents

Title Page .....	i
Author's Declaration .....	ii
Abstract .....	iii
Acknowledgements .....	iv
Dedication .....	v
List of Tables .....	ix
List of Figures .....	x
1.0 Introduction .....	1
1.1 Problem Definition .....	1
1.2 Context .....	4
1.3 Research Rationale .....	9
1.4 Purpose and Objective .....	10
1.5 Study Area .....	11
1.6 Organization of Thesis .....	15
2.0 State of Current Methods .....	16
2.1 Wet Weather Flow Subcatchment Delineation Methods .....	16
2.2 Dry Weather Flow Methods .....	23
2.3 Rainfall Interpolation Methods .....	27
3.0 Automated Approach to Dry Weather Flow Subcatchment Delineation .....	30
3.1 Required Datasets .....	31
3.2 Frontage Node Generator Tool .....	33
3.3 Modifying and Running the "Add a Lateral" Tool .....	35
3.4 DWF Subcatchment Generator Tool .....	36
4.0 Automated Approach to Wet Weather Flow Subcatchment Delineation .....	41
4.1 Required Datasets .....	42
4.2 Creating the Boundary Polygon Shapefile .....	44
4.3 Selecting the Appropriate Pipes Based on Catch Basin Leads (Preprocessing and Data Quality Assurance) .....	45
4.4 The WWF Subcatchment Generator Tool .....	51
5.0 Automated Approach to Rainfall Interpolation .....	56

5.1 Datasets Required .....	57
5.2 Determining the Number of Rain Gauges Needed for Interpolation .....	58
5.3 Compiling Rainfall Data and Preparing it for Input into the Tool.....	59
5.4 Determining the Most Suitable Interpolation Method .....	62
5.5 Designing and Running the Continuous IDW Rainfall Interpolation Tool .....	63
5.6 Creating Virtual Rain Gauges.....	64
5.7 Extracting Virtual Rain Gauge Location Data from Interpolated Layers.....	67
5.8 Associating Virtual Rain Gauges to WWF Subcatchments.....	68
6.0 Evaluating Automated DWF Subcatchment Delineation .....	71
6.1 Evaluating Automatically Delineated DWF Subcatchments for Coxwell.....	71
6.2 An Overview of Errors Produced by the Automated DWF Subcatchment Delineation Process .....	79
6.3 Summary of Strengths and Weaknesses of the DWF Tools .....	83
7.0 Evaluating Automated WWF Subcatchment Delineation .....	84
7.1 Checking the Overall Accuracy of the Approach.....	84
7.2 Common Errors Generated by Approach .....	88
7.3 Strengths and Weaknesses of Approach.....	92
8.0 Evaluating Automated Approach to Continuous Rainfall Interpolation.....	93
8.1 TRCA Rain Gauges .....	93
8.2 Comparing Results with TRCA Rain Gauges .....	96
8.3 Utilizing NEXRAD Datasets as an Additional Source for Comparison.....	103
8.4 Summary of Strengths and Weaknesses of Approach .....	108
9.0 Conclusion, Future Research and Recommendations.....	109
9.1 Conclusion .....	109
9.2 Future Research .....	114
9.2.1 Automated DWF Subcatchment Delineation.....	114
9.2.2 Automated WWF Subcatchment Delineation.....	115
9.2.3 Automated Rainfall Interpolation .....	116
9.2.4 General Research .....	116
9.3 Recommendations.....	117
Appendix.....	120

A.1 Source Code for Adjusting XML Configuration File to Create Pseudo Sewer Laterals for Each Parcel.....	120
A.2 Creating a Geometric Network .....	123
A.3 ModelBuilder Diagrams of Selected Tools Developed .....	125
A.4 Results of Interpolation Methods Testing .....	129
A.5 Steps for Creating Virtual Rain Gauges Using the Repeating Shapes Tool .....	131
A.6 Excel Functions Used to Eliminate Non-Numeric Characters from a Field.....	131
A.7 The 93 DWF Subcatchments Visually Inspected .....	132
A.8 The 96 WWF Subcatchments Visually Inspected.....	135
A.9 Raw Data of Rainfall Measurements above Zero Millimetres for one of the two TRCA Rain Gauges Used to Evaluate the Validity of the Rainfall Interpolation Method Developed .....	137
References.....	153

## List of Tables

Table 1. DWF subcatchment assumptions.....	30
Table 2. Layers required to run DWF subcatchment delineation tool.....	32
Table 3. WWF subcatchment assumptions.....	41
Table 4. Layers required to run WWF subcatchment delineation tools.....	42
Table 5. Layers required for virtual rain gauge generator tools.....	57
Table 6. DWF subcatchment criteria.....	71
Table 7. Results of the visual inspection of selected DWF subcatchments.....	77
Table 8. Comparison of number of topology errors identified.....	79
Table 9. WWF subcatchment categorization criteria.....	84
Table 10. WWF subcatchment automatically delineated categorization results.....	87
Table 11. Comparison of number of topology errors identified for WWF subcatchments.....	88
Table 12. TRCA approval code table.....	94
Table 13. TRCA grade code table.....	95
Table 14. Excerpt of a table used for comparison purposes.....	97
Table 15. Evaluation of virtual gauges derived from interpolation.....	102
Table 16. Intensity values recorded by the City's 27 gauges expressed in mm per hour and those of the TRCA's measured in mm at 19:25 on July 7th, 2013.....	105
Table 17. Intensity values recorded by the City's 27 gauges expressed in mm per hour and those of the TRCA's measured in mm on July 8th, 2013 at 16:10.....	106
Table 18. Intensity values recorded by the City's 27 gauges expressed in mm per hour and those of the TRCA's measured in mm on July 5th, 2013 at 15:45.....	107
Table 19. Comparing values from the TRCA HY003 gauge versus virtual gauge.....	137

## List of Figures

Figure 1. The spatial arrangement of the City of Toronto's 36 rain gauges .....	4
Figure 2. Sewer networks where the automated methods were applied.....	7
Figure 3. Only 18 rain profiles being utilized by the model.....	9
Figure 4. Elevation values across the city of Toronto in metres above sea level (mASL).....	12
Figure 5. Landuse types across the City of Toronto .....	13
Figure 6. All the unknown parcels having been visually filled in .....	14
Figure 7. The spatial location of the three boundaries used .....	15
Figure 8. A single manually delineated DWF subcatchment .....	24
Figure 9. DWF subcatchments automatically delineated by InfoWorks CS .....	26
Figure 10. DWF method schema .....	31
Figure 11. Frontage nodes automatically generated by the frontage node tool .....	33
Figure 12. A case where the frontage node tool would not output the frontage node but instead output the centroid .....	34
Figure 13. The user input interface for the Frontage Node Generator Tool .....	35
Figure 14. User input interface for the DWF Subcatchment Generator Tool.....	37
Figure 15. DWF subcatchment created when using an earlier version of the DWF Subcatchment Generator Tool.....	38
Figure 16. How the Multipart to Singlepart Function in ArcGIS converts a multipart feature into separate features.....	39
Figure 17. DWF subcatchment automatically delineated by the tool.....	40
Figure 18. WWF method schema .....	42
Figure 19. WWF subcatchment delineation method broader study area boundary .....	45
Figure 20. Where the areas occupied by both the Coxwell and Interceptor networks that were checked by the modellers in InfoWorks were removed.....	46
Figure 21. The "Target layer(s) features that touch the boundary of the Source layer" reduces the number of erroneous selections .....	47
Figure 22. Merging of the Coxwell and Interceptor pipe networks.....	48
Figure 23. Layer of all the pipes from the unmodified network (from TWAG) selected because they have catch basins attached to them. ....	49
Figure 24. Join Data function used to select the appropriate pipes even if they were moved slightly and therefore no longer snapped to a catch basin lead.....	49
Figure 25. Highlighted blue line from the TCL layer signifies the location of a culvert that was "burnt in" .....	50
Figure 26. User interface for WWF Subcatchment Generator Tool.....	51
Figure 27. Comparison of different raster resolutions from 3-m to 0.5.....	52
Figure 28. Extraneous catchments that were eliminated after pipe information was added.....	53
Figure 29. WWF subcatchments automatically delineated by the tool .....	54
Figure 30. As can be seen, it respects the gradients of the contour lines quite well.....	54
Figure 31. High points flowing downslope towards the pipe's location .....	55

Figure 32. Virtual rain gauge method schema .....	56
Figure 33. RG-020 which is a tipping bucket style rain gauge known as the TR-525USW rain gauge (Cole, 2014a) .....	58
Figure 34. The 27 rain gauges selected.....	59
Figure 35. How rainfall records from each of the 27 rain gauges was selected and downloaded	60
Figure 36. Typical attribute table of a rain gauge's records once all the measurements were spatially linked to their appropriate rain gauge location in a feature class .....	61
Figure 37. User interface for the Split Layer By Attributes Tool .....	62
Figure 38. User interface for Continuous IDW Rainfall Interpolation Tool .....	64
Figure 39. A ModelBuilder diagram of the Continuous IDW Rainfall Interpolation Tool .....	64
Figure 40. User Interface for the Virtual Rain Gauge Generator .....	65
Figure 41. ModelBuilder diagram of the Virtual Rain Gauge Generator .....	66
Figure 42. Virtual rain gauge point locations generated by tools .....	66
Figure 43. CSV input file generated for RG000 with the time being represented in the first column and intensity in mm per hour in the second. ....	68
Figure 44. User interface for Rain Gauge Assigner Tool .....	69
Figure 45. ModelBuilder diagram of Rain Gauge Assigner Tool .....	69
Figure 46. Results generated by rain gauge assigner tool.....	70
Figure 47. Location of Section A and Section B .....	72
Figure 48. Evaluation results for Section A.....	73
Figure 49. Evaluation results for Section B .....	74
Figure 51. Randomly distributed points that were not suitable .....	75
Figure 50. Results of sample size calculator.....	75
Figure 52. The 93 randomly selected subcatchments that were used in the visual check .....	76
Figure 53. Very unusual parcels that are likely inaccuracies.....	78
Figure 54. DWF Error # 1: Incorrect parcel connections caused by frontage node location .....	80
Figure 55. Error # 2 Incomplete subcatchment only containing the road corridor portion with the upstream node contained inside and Error # 3 Incomplete subcatchment only containing parcel portion with no upstream node .....	81
Figure 56. Error # 4 Incorrect parcel connection caused by location of pipe .....	82
Figure 57. Calculation used to determine sample size for WWF subcatchment evaluation.....	85
Figure 58. The 99 WWF subcatchments selected for visual inspection .....	86
Figure 59. Visualization of the result of the WWF subcatchment check .....	87
Figure 60. Error # 1 Incorrect boundary of subcatchment.....	88
Figure 61. Error # 2 Subcatchment contains multiple sets of catch basins belonging to multiple pipes .....	89
Figure 62. Error # 3 Subcatchment containing another smaller subcatchment .....	90
Figure 63. Error # 4 WWF subcatchments consisting of fragmented multipart features .....	91
Figure 64. Spatial arrangement of the TRCA's 9 rain gauges in Toronto .....	94
Figure 65. TRCA rain gauges used as points of comparison.....	96

Figure 66. Comparison of TRCA HY016 Values Versus Interpolated Values Over Time.....	99
Figure 67. Scatter plot of TRCA HY016 rainfall measurements and the corresponding interpolated measurements.....	99
Figure 68. Comparison of TRCA HY003 Values Versus Interpolated Values Over Time.....	100
Figure 69. Scatter plot of TRCA HY003 rainfall measurements and the corresponding interpolated measurements.....	100
Figure 70. Location of the rain gauges used in the correlation coefficient calculations .....	101
Figure 71. NEXRAD radar visualization.....	104
Figure 72. Location of precipitation according to radar (blue rectangles) relative to gauges at 19:25 (23:25 GMT) on July 7th, 2013.....	105
Figure 73. Location of precipitation according to radar (blue rectangles) relative to gauges at 16:10 (20:10 GMT) on July 8th, 2013.....	106
Figure 74. Location of precipitation according to radar (blue rectangles) relative to gauges at 15:45 (19:45 GMT) on July 5th, 2013.....	107
Figure 75. Flowchart of the automated subcatchment delineation tools presented .....	111
Figure 76. The limits of TWAG in accounting for catch basins.....	118

## 1.0 Introduction

### 1.1 Problem Definition

As urban areas become the home to increasingly more of the world's population and the intensity of precipitation events that can potentially lead to flooding disasters increases because of climate change, more focus is being placed upon the ability of cities to prepare for these challenges (McCarthy *et al.* 2010; Allan, 2011; Leitão *et al.* 2012). In response to these emerging issues, great reliance and interest is being placed upon models such as one dimensional (1-D) distributed urban drainage models to provide the answers (Simões *et al.* 2011; Blumensaat *et al.* 2012). Such modelling, typically combines hydraulic modelling, which is the flow of water in the sewer network also known as the minor system, with hydrological modelling, the flow of water over the major system or across the surface of urban environments. Therefore, in some cases it is also referred to as dual-drainage modelling (Leandro, 2008; Chen *et al.* 2009; Leitão *et al.* 2012).

Such a refocus in research has occurred because traditional lumped models do not provide the level of detail that is required to make certain decisions (Rodriguez *et al.* 2008; Hamel *et al.* 2013). In a distributed model, modellers are able to determine whether a particular building is differentially prone to flooding, which would be nearly impossible to do using traditional lumped models because they are too aggregated (Luciani, 2005; Elliott & Trowsdale, 2007). However, distributed models also have a disadvantage when compared to lumped models. In order to generate more comprehensive results, the model itself has to be more detailed. This means the model is so comprehensive that it operates on a pipe-to-pipe basis. Thus, each pipe has its own individual subcatchment or sewershed that represents it, instead of one large polygon representative of multiple pipes, as is the case in a lumped model. Such modelling is referred to as 1-D modelling because it does not specifically identify which directions the floodwaters flow on the surface. Instead merely provides a value for the volume of water that will overflow from the sewer system via individual manholes and or outfalls in the event of a flood (Mark *et al.* 2004, C.W. Baxter, personal communication, July 18, 2013). In addition, such models also require the input of a defined overland flow pathway based on links and nodes (Leandro *et al.* 2009). While there is also 2-D modelling which overcomes these limitations, this approach is much slower in terms of computation time and requires more

detailed and accurate datasets than currently widely available (Leandro, 2008; C.W. Baxter, personal communication, July 18, 2013), and is therefore beyond the scope of the research being undertaken here. Furthermore, at present, it is not being widely used in major cities such as the City of Toronto, although it is one of their future goals (G. Lin, personal communication, July 18, 2013).

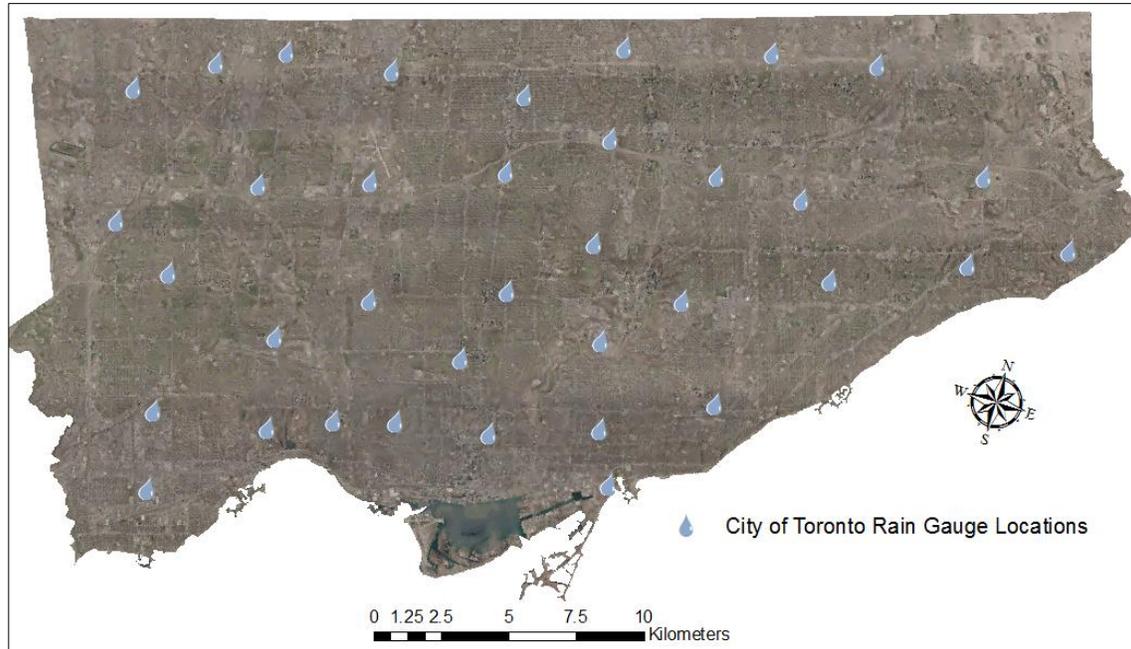
In order to generate the level of detail required for a distributed model, detailed geospatial and temporal data need to be provided. Geographic Information Systems (GIS) are the most appropriate way of developing these datasets because of the ability of this technology to allow the necessary data to be manipulated (Storck 1998; Seth *et al.* 2006; Amaguchi *et al.* 2012) this was the focus of this research. However, the actual hydrologic and hydraulic modelling often takes place in a specialized modelling software environment such as InfoWorks CS, but would not be able to run without reliance on GIS (Edwards *et al.* 2011). Its core components are stored as either shapefiles or an ESRI Geodatabase, with the latter being a proprietary database storage format developed by ESRI to store geospatial data (ESRI, 2010; Edwards *et al.* 2011; Innovyze, 2013).

In the InfoWorks CS modelling software used by the City of Toronto who serve as a case study in this research, modellers need to input subcatchments for both the storm and sanitary systems known as Wet Weather Flow (WWF) subcatchments and Dry Weather Flow (DWF) subcatchments respectively. However, these subcatchments could not be automatically generated to a satisfactory level for City staff by the InfoWorks CS software itself (G. Lin, personal communication, February 13, 2013). Instead, staff must either manually create them in a program specifically designed to handle geospatial files, such as ESRI's ArcGIS 10 with an ArcInfo license, as is being used in this research, or within InfoWorks CS, using the GIS tools offered within it (ESRI, 2010; Innovyze, 2013). This is a challenge faced not only by the City of Toronto but is now recognized as being a major hurdle in constructing urban drainage models rapidly and accurately (Dongquan *et al.* 2009). While other proprietary and open source GIS software options exist such as MapInfo or GRASS, ArcGIS was chosen because it is the software package currently being used by Toronto and other major cities in North America (Johnson, 2009). Furthermore, other modellers in their manual delineation of subcatchments have also used this software (Giron, 2005). Additionally, the City's needs for hydrological modelling also require precipitation measurements recorded by rain gauges to be modelled as

point shapefiles indicating the location of the gauges and a separate comma-separated values (CSV) file.

The CSV contains the actual rainfall measurements, measured as intensity in millimetres per hour, recorded on a consistent time interval basis, such as every 5 minutes, for the duration of the simulation (X. Liu, personal communication, July 10 2013; Innovyze, 2013). Currently, the City of Toronto has approximately 36 rain gauges distributed across the city (Figure 1). However, it was uncertain whether this was dense enough because even using the relatively coarse 4 km by 4 km cells from the NEXRAD radar precipitation data from the Hydrologic Rainfall Analysis Project (HRAP) as a yardstick would require approximately 62 unique rain gauge locations to provide adequate coverage for the City of Toronto (Robayo, 2005). In addition, according to rainfall data resolution requirements set by Schilling in 1991 for urban drainage studies, it is necessary to have one rainfall measurement per square kilometer. Since the entire combined service area is over 160 km<sup>2</sup>, it would require over 160 unique rainfall measurements to meet this standard (Einfalt *et al.* 2004).

Furthermore, although there was some radar-detected rainfall data freely available for Toronto, they do not necessarily exist for other cities across Canada, although, more importantly, it was uncertain whether radar data need to be calibrated and that was deemed beyond the scope of this work. Thus, for this reason, primary rain gauge data would seem to be preferable for Canadian municipalities even though it has been used for flood modelling of American cities (Knebl *et al.* 2005). Additionally, there have already been studies that examined whether rainfall data derived from radar data were more suitable for urban drainage modelling as was conducted in Flanders, Belgium (Goormans & Willems, 2013). Furthermore, studies have already been conducted on combining radar and rain gauge data together (Einfalt *et al.* 2004). However, radar data remains imperfect and can possibly generate large errors in urban drainage simulation values (Berne & Krajewski, 2013; Shafiei *et al.* 2014).



**Figure 1. The spatial arrangement of the City of Toronto's 36 rain gauges**

As previously stated, one of the major reasons why this topic has become so important, is that climate change has been accompanied by the increased risk of exceeding municipal infrastructure capacities. Which is expressed as street and basement flooding, due to surcharged sewer pipes, especially in urbanised areas (Diaz-Nieto *et al.* 2012). This increased risk is caused by the increasing frequency of intense precipitation events associated with climate change that have been identified as responsible for 60% of flood events in urban areas in the UK (Intergovernmental Panel on Climate Change (IPCC), 2007; Leitão *et al.* 2012). In fact, one study that modelled the topography in a GIS environment, for the area around the Arno River at Firenze in Italy, determined that urbanised areas were the regions that were most prone to flooding (Morelli *et al.* 2012). While no such publicly available study was conducted for Toronto, it is likely that such results hold true widely because another earlier, similarly GIS-driven urban drainage model created of the Zhujiang Delta of southern China, also concluded that highly urbanized areas were most prone to flooding (Weng, 2001).

## 1.2 Context

While the previous section introduced the challenge of creating input parameters manually for 1-D distributed urban drainage models and suggested that the density of physical rain gauges is typically less than what is perhaps optimal for modelling purposes, this section

explains how this particular research was initiated. On June 28th, 2012, the Canadian federal government officially released the *Wastewater Systems Effluent Regulations (WSER)* pursuant to the federal *Fisheries Act*. Under these new regulations, a number of new requirements were created that all major wastewater treatment facilities across Canada, excluding anywhere north of the 54th parallel in Quebec or Newfoundland and Labrador, had to comply with (Department of Justice, 2012). As municipalities comprised the majority of operators of such facilities, *WSER* most deeply affected them.

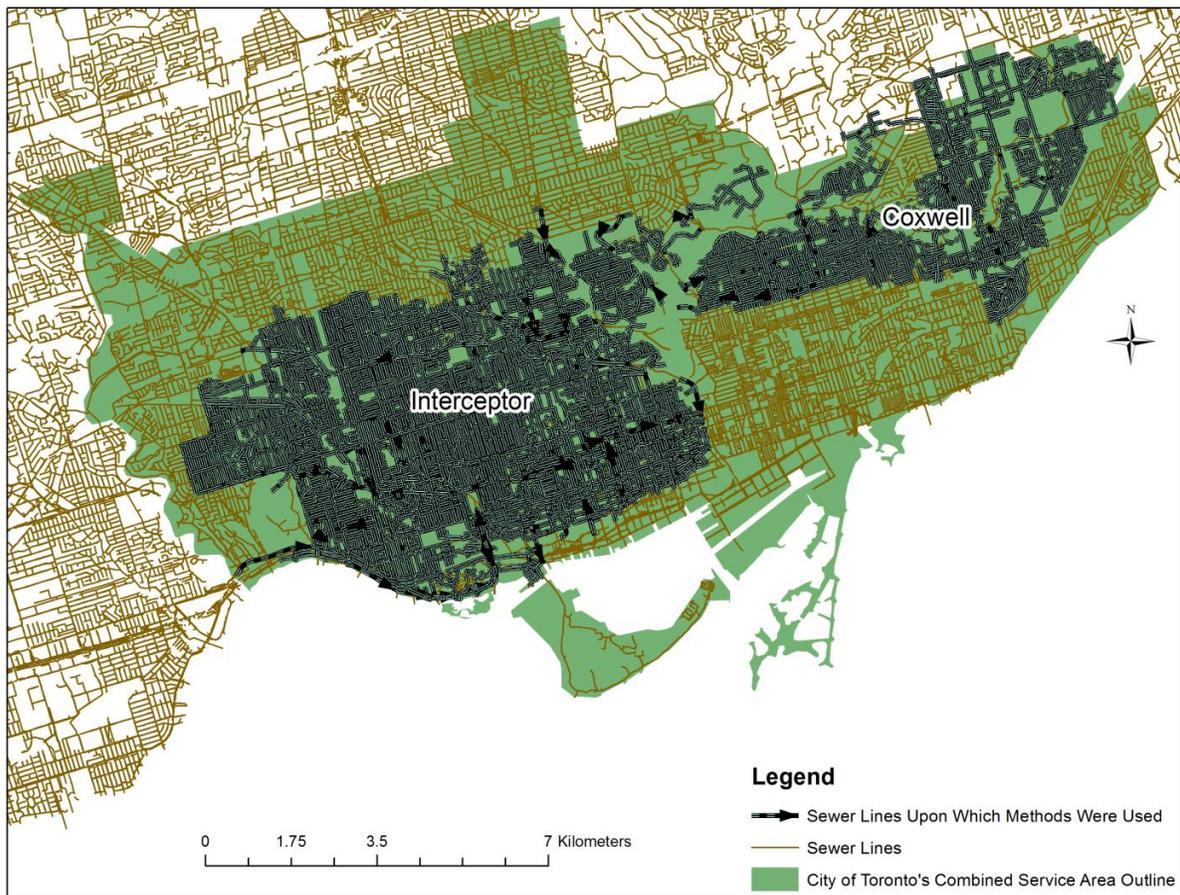
In the fall of 2002, Environment Canada held a number of stakeholder consultation meetings regarding the future management of sewer wastewater effluent in Canada. From these meetings, Environment Canada received feedback, supporting a harmonized approach to wastewater management (Government of Canada, 2012). In respecting the system of federalism that exists in Canada, Environment Canada approached the Canadian Council of Ministers of the Environment (CCME) (Privy Council Office, 2012; Government of Canada, 2012). The CCME is a result of Canadian federalism because, while responsibility for the environment is never explicitly stated in the Constitution Act, the various powers that relate to the environment are shared by both levels of government, making the environment a common responsibility (Greenbaum & Wellington, 2010, pp. 12-3 70). The CCME is comprised of the Federal Minister of the Environment and all of the Environmental Ministers from the ten provinces and the three territories, in an attempt to respect the views of both levels of government (CCME, 2011).

Thus, this intergovernmental organization, being the ideal candidate to help develop this nationwide standards approach, agreed to develop the necessary framework during the following year (Government of Canada, 2012). After several years of consultation with various groups and stakeholders, most of the provincial ministers on the CCME finally endorsed the “Canada-wide Strategy for the Management of Municipal Wastewater” report that was released on February 17, 2009 (CCME, 2009).

Following the official release of these regulations, key reporting requirement dates were identified by the City of Toronto. Among these dates was May 15, 2013 by which date the City of Toronto had to report to Environment Canada through an online portal the spatial location of all its combined sewer overflow (CSOs) locations with coordinates measured in degrees latitude and longitude. The City had to submit a report by February 15, 2014 on the frequency

and volume of overflows from the City's 300 plus CSOs. Combined sewer overflow locations were originally devised to prevent flooding in urban environments by diverting excess stormwater into a nearby water body or stream instead of overloading the local wastewater treatment plant (WWTP). However, as the treatment levels of many WWTP have increased and will be increasing more under the new *WSER* regulations, the water originating from CSOs themselves are now recognized as being significant sources of water pollution to receiving waters (Lau *et al.* 2002). Not only are they major sources of pollution due to their sewage content, but they have also been reported to contain other pollutants such as heavy metals and pesticides, as were identified in studies of CSO discharge in the of Kamloops, British Columbia, Buffalo, New York and Paris, France (McGreer & Belzer, 1999; Irvine *et al.* 2005; Gasperi *et al.* 2008).

There were only two options for the City of Toronto: either placing a monitoring device on every CSO outfall, which was quickly deemed impractical due to the costs, or develop a distributed model to estimate the values required for reporting. The City's existing lumped model of the combined area known as the BPR model could not be relied upon entirely because it did not contain all CSO outfall points that had to be reported on individually; hence new models had to be created. In order to create two of these new models, the City of Toronto commissioned a team of Ryerson University researchers to develop models for the portions of the combined sewer network labelled as Interceptor and Coxwell (Figure 2).

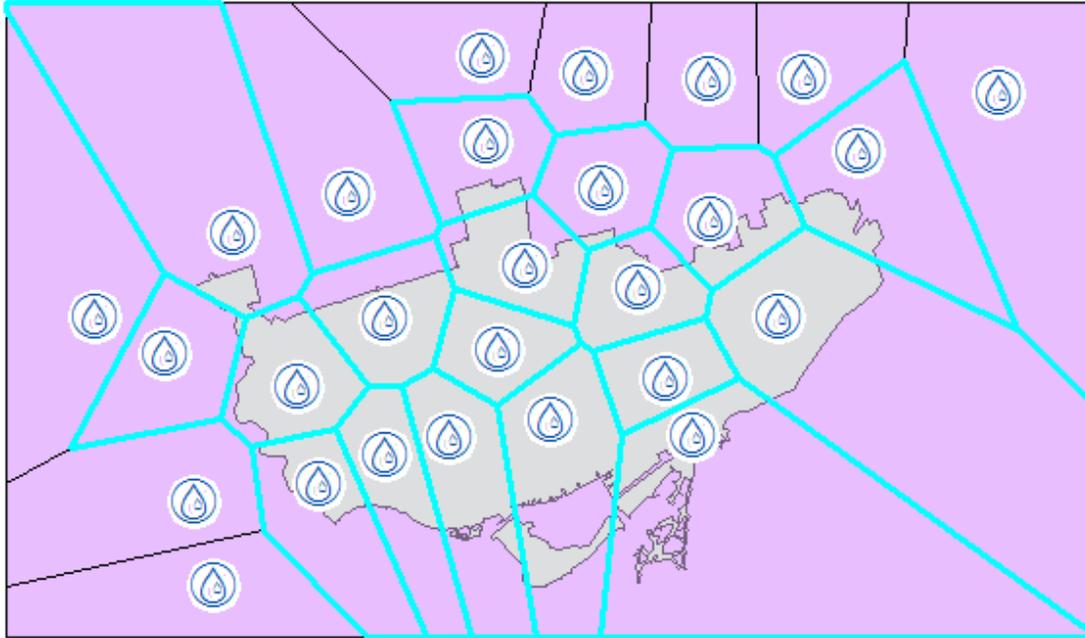


**Figure 2. Sewer networks where the automated methods were applied.**

While the models were not being built entirely from the ground up as most of the pipes had previously been digitized, what was needed was a means to delineate dry weather flow (DWF) subcatchments to simulate normal sanitary baseflow conditions, and wet weather flow (WWF) subcatchments to simulate inflow from precipitation events. It has been the practice at the City to delineate these features manually and is the typical practice elsewhere (Chen *et al.* 2003; Mark *et al.* 2004; Jankowsky *et al.* 2012). However, such an approach was considered untenable as that meant creating as many as 21,572 subcatchments because that was the total number of pipes in the area of the City that Ryerson University was responsible for modelling. This situation provided an ideal opportunity to pursue the primary aim to develop GIS-based automated methods that would reduce the time required for delineating subcatchments in comparison to relying solely on the manual approach. Therefore, it was decided to create these subcatchments by developing new automated GIS-based tools tailored specifically for the City of

Toronto, but which could be potentially applied elsewhere by incorporating existing automation tools.

Moreover, Toronto Water had previously contracted private consultant-engineering firms to create other smaller isolated distributed basement flooding models, one of whom used interpolation methods to generate more rain profiles than were possible with the Thiessen method. As a result, City staff wanted to expand upon the work done by the consultancy (Clarifica, 2010). The consultancy did not have to develop a tool to interpolate rainfall on a continuous basis over a period of seven months because they only simulated two separate precipitation events. Thus, they had likely relied upon the standard interpolation methods available in ArcGIS to generate their results as it is not entirely clear which interpolation method they utilized (Clarifica, 2010; ESRI, 2010). Therefore, City staff also requested consideration of an automated approach using interpolation methods in ArcGIS to generate virtual rain gauges that were based on thousands of continuous interpolations. These interpolations would then be used as input for precipitation into the two InfoWorks CS models that were being developed by Ryerson that covered a portion of the combined system area (Figure 2). Modellers of other areas in the City's combined sewer service area could then potentially use these interpolations as well. This was important because the City had 36 rain gauges, though only 18 of them would be utilized with the traditional Thiessen method when applied to the entire combined service area. However, this does not provide the high spatial and temporal rainfall resolutions that are required for urban drainage modelling purposes (Figure 3) (Chen & Liu, 2012). Thus, a secondary aim was to develop an automated approach relying upon continuous interpolations to generate more rain profiles, than simply relying upon the Thiessen method. This method operates on the basis that any location within each polygon generated is closer to its associated rain gauge than to any other rain gauge point feature (ESRI, 2010).



**Figure 3. Only 18 rain profiles being utilized by the model**

### **1.3 Research Rationale**

Although, in the previous section it is apparent that the impetus for the creation of a new detailed model for a portion of the City of Toronto's combined sewer network region was the new requirement of Canadian federal regulations, the methods developed through this research hold greater potential. The City of Toronto is not going to be using this model to estimate only CSOs' surcharge frequencies, outflows and volumes. Instead, the plan is to use it for more operational tasks such as assessing whether a particular block can handle a new development based on the existing infrastructure as modelled in the virtual environment of InfoWorks CS or whether they need to improve the storm sewer capacity in a certain area to prevent more basement flooding events. As such, these results could potentially be used to assist in prioritizing system upgrades for the City's aging infrastructure and potentially elsewhere (Long *et al.* 2009). These types of applications are why it is now recognised that urban drainage modelling studies are crucial to the operations of sewer systems (Goormans & Willems, 2013).

This research is seen to assist in providing scientific support for the management of CSOs by a municipality through its creation of model input parameters. It is intended that it is useful for creating urban drainage models that can help to answer questions that more and more cities are going to be faced with as they continue to grow and climate change increases the

frequency of intense precipitation events (Marsalek *et al.* 2006; IPCC, 2007; Leitão *et al.* 2012).

#### **1.4 Purpose and Objective**

The purpose was to determine the truthfulness of the primary and secondary hypotheses suggested by the observations presented above. The primary aim was pursued by estimating how much time modellers can potentially save using the automated approach to delineate subcatchments based on the percentage of subcatchments it was able to automatically delineate to an "Acceptable" level. The secondary aim was pursued by attempting to develop an approach that was capable of continuously interpolating existing rain gauge measurements and then determining the accuracy of the interpolated values against an independent reference set of rain gauges.

The objective of this research was to:

Develop GIS-based tools that automatically generate input datasets for 1-D distributed urban drainage modelling.

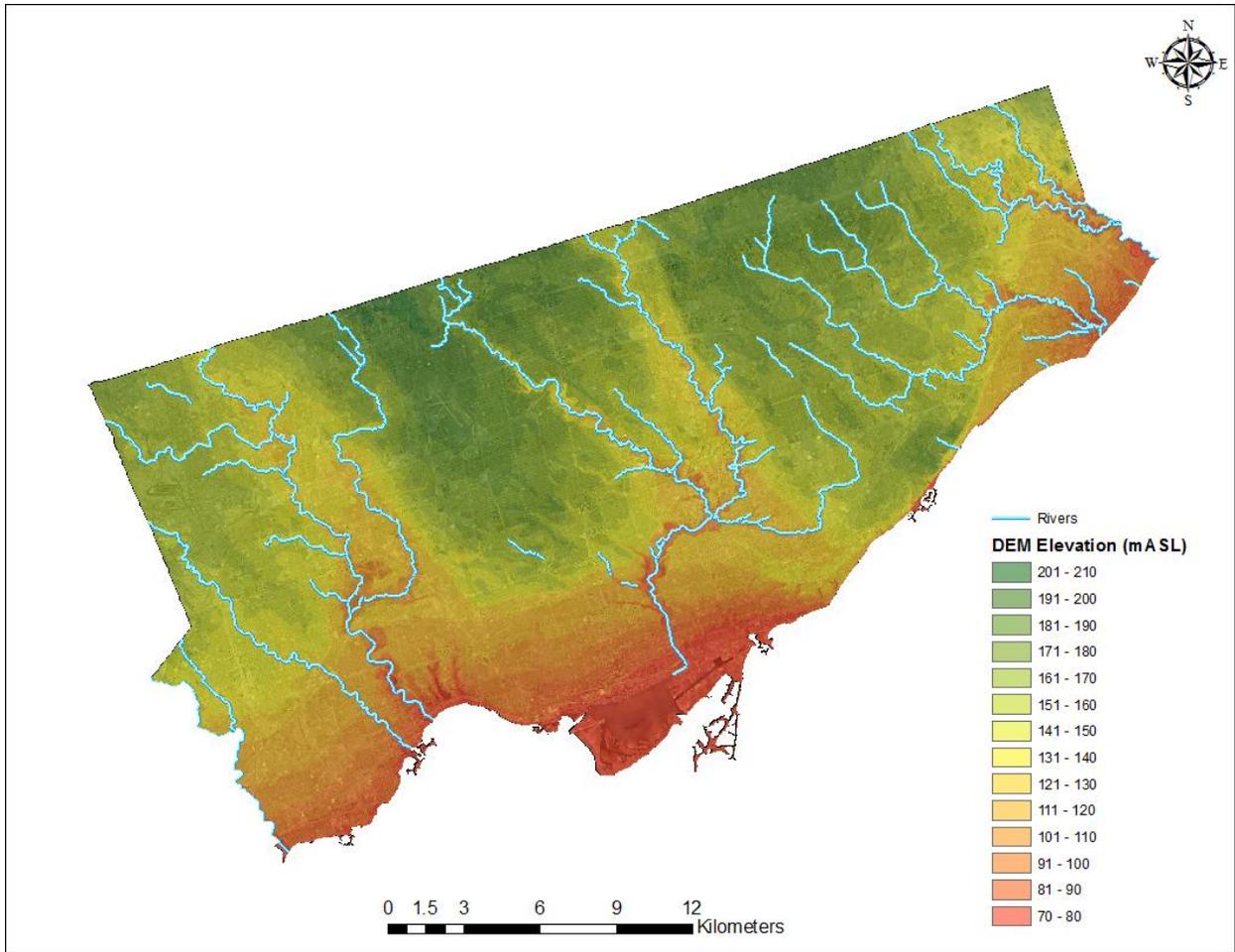
To achieve the objective, the following tasks were completed:

1. Identify the need for the tools being presented here that are more objective and repeatable, an enhancement of the current manual methods.
2. Document the development and implementation of these automated tools in a manner that makes them easy to replicate.
3. Design the tools in such a manner that allows greater future flexibility by developing on a platform that requires minimal programming skills making it accessible to City staff and other potential future users of the methods.
4. Document the effectiveness and limitations of these tools to allow potential users to make an educated decision as to whether or not they want to implement them.
5. Provide recommendations on when and when not to use these tools and identify future directions of research.

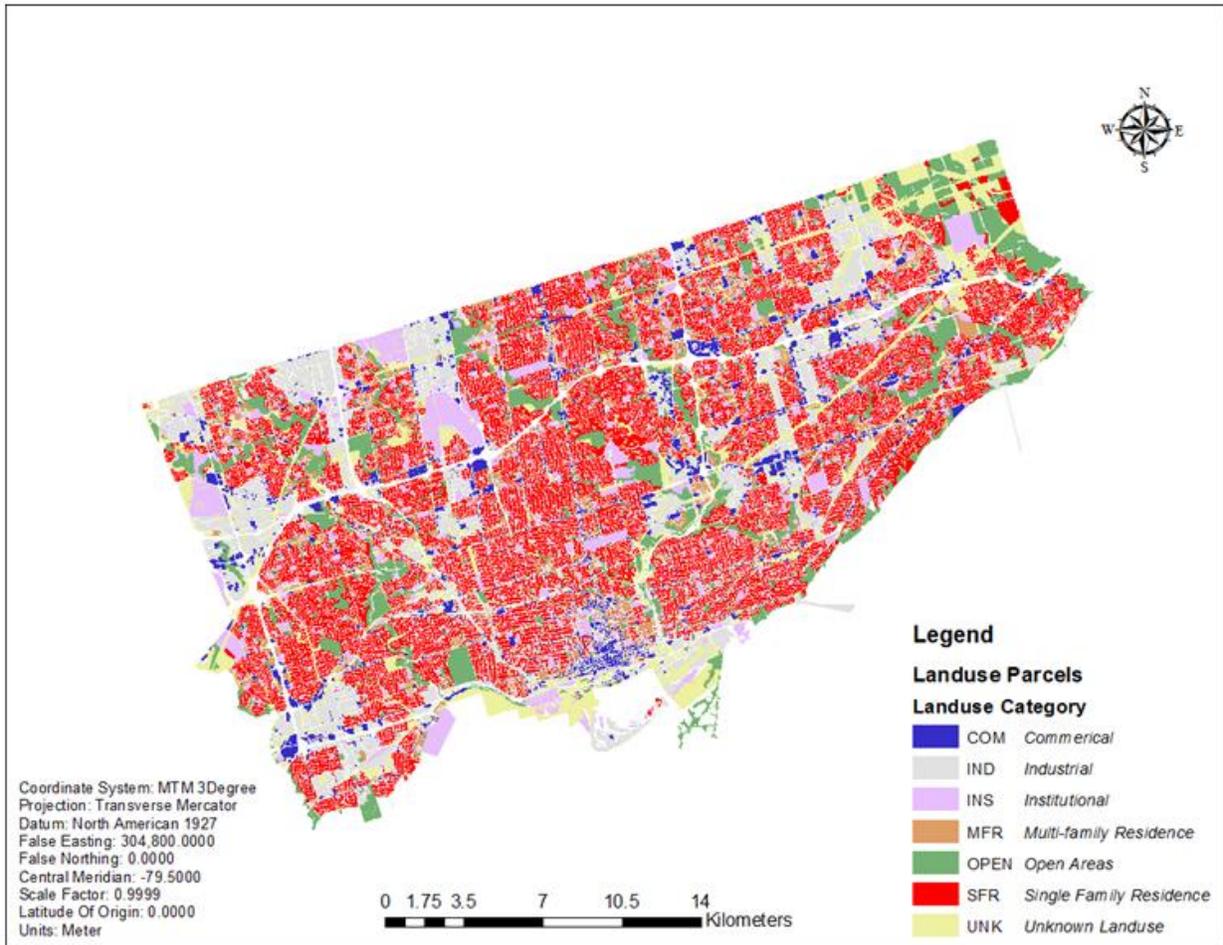
## 1.5 Study Area

The City of Toronto is the most populated city in Canada and spans an area of roughly 630 km<sup>2</sup> (City of Toronto, 2013b). Its population is projected to increase from 2.79 million in 2012 to 3.45 million in 2036, which translates into a greater strain upon its existing infrastructure (Ontario Ministry of Finance, 2013). According to the City's in-house database, known as the Toronto Water Asset Geodatabase (TWAG), the oldest pipe in the system dates back to 1844. While the total length of the city's stormwater and sanitary-sewer lines, according to TWAG, is over 10,532 km and consists of over 166,000 individual pipe segments, City staff have admitted that there are portions of the network that are still missing in TWAG (J. Lei, personal communication, February 6, 2013). This thesis was focused on the combined sewer service area of the city, specifically the two pipe networks identified in Figure 2. A combined sewer is one in which both sanitary and stormwater flow into the same pipe. Therefore, overflows from such a sewer are referred to as combined sewer overflows or CSOs.

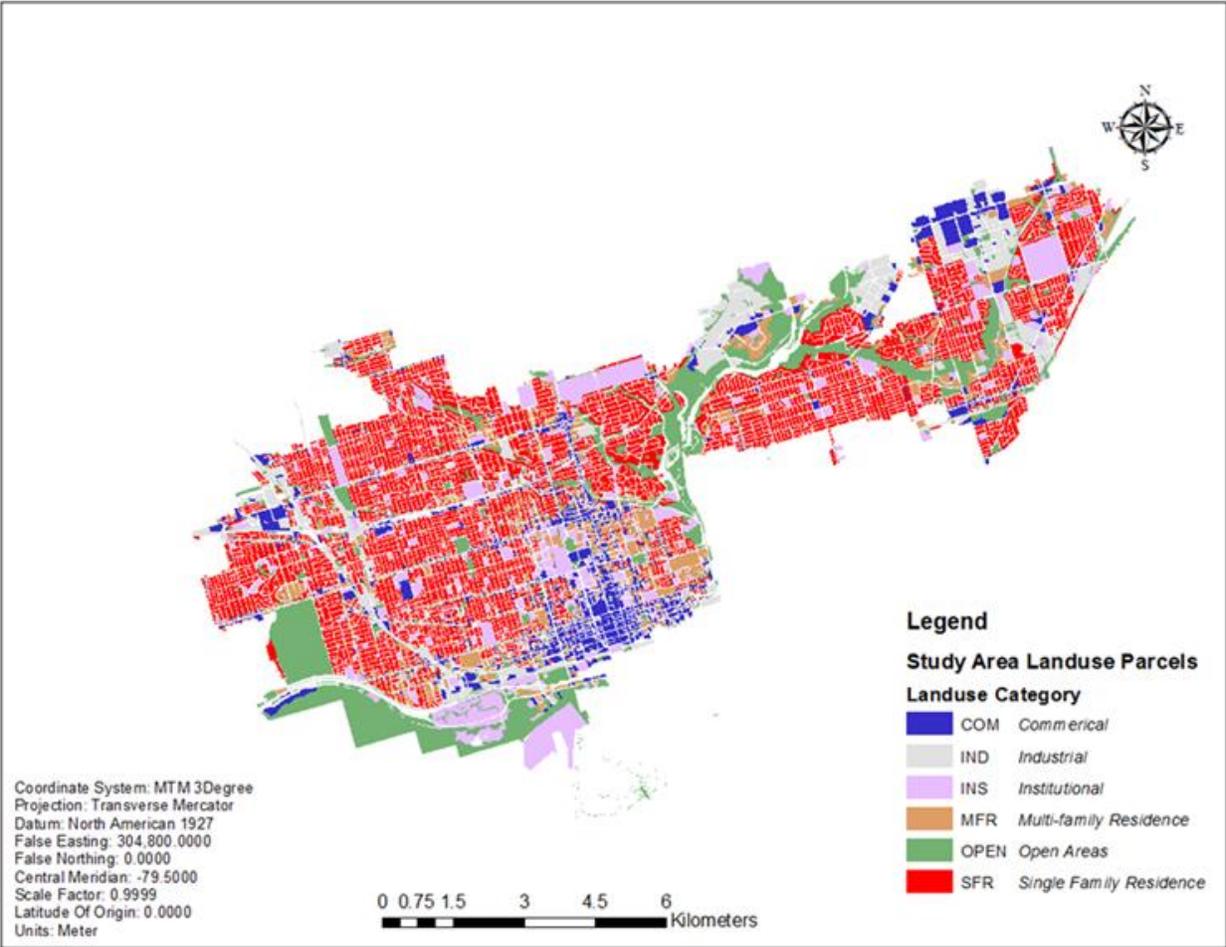
The City of Toronto is relatively flat; the map in Figure 4 was based on a fine network of digital elevation points. Being situated beside Lake Ontario, the land and the sewer networks slope towards this water body where all pipes including the City's CSOs, ultimately discharge. Land use distribution for the City, providing a parameter commonly used in modelling, was available from Toronto Water (City of Toronto, 2006) on a land parcel basis, with some gaps found in the records (Figure 5). To confirm the land use in the missing cases, a visual inspection was undertaken, using a combination of 2011 10-cm high-resolution orthoimagery provided by City staff at Toronto Water (City of Toronto, 2011b) and Google Earth satellite imagery (with the assistance of K. Kang). Several thousand parcels had to be filled in. From this, it is evident that 1.2% were industrial locations, 84.3% were single-family residences, 7.8% multiple family residences such as apartment buildings and condominiums, 1% institutional buildings such as schools and churches, 3.9% were commercial facilities such as malls and offices, and 1.8% was open area (Figure 6).



**Figure 4. Elevation values across the city of Toronto in metres above sea level (mASL)**

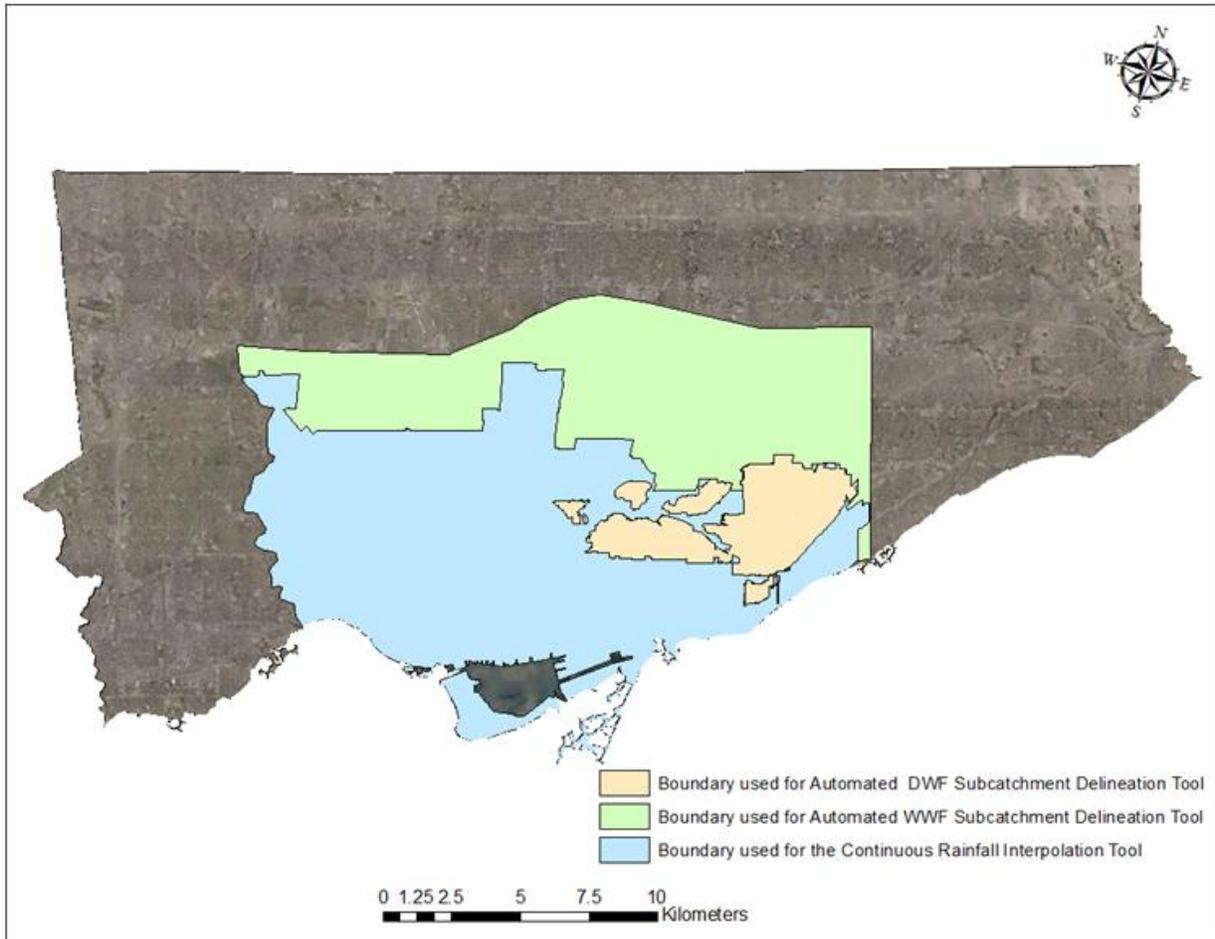


**Figure 5. Landuse types across the City of Toronto**



**Figure 6. All the unknown parcels having been visually filled in**

For the area that is being focused on, there are two main sewer pipe networks, the Coxwell and Interceptor, which are the names of the trunk pipes in their systems. The Coxwell network covers part of Scarborough, while the Interceptor covers most of the pre-amalgamation City of Toronto as seen in Figure 2. Figure 7 provides a clear overview of the extent of the study areas used when developing the final iterations of the automated tools that are covered in future sections. From this, it should be clear that the scope of this thesis strictly remains in the realm of how GIS can be used to assist in the process of developing inputs for 1-D distributed urban drainage modelling by using the City of Toronto as a case study. Thus, concerns beyond this realm such as how these inputs translate into CSO monitoring results from the model are beyond the scope of this thesis and hence not covered.



**Figure 7. The spatial location of the three boundaries used**

## **1.6 Organization of Thesis**

This thesis is structured into nine sections. In this current section, the scope, motivation and need for this research are detailed. Section two provides the necessary background to recognize that while automated methods exist, none is perfect and approaches have been found to be rather lacking in details regarding their implementation. Sections three, four and five present the automated methods developed through this research. Sections six, seven and eight detail analyses conducted to evaluate the methods introduced in the previous sections. In section nine, conclusions are drawn about the methods developed and recommendations for municipalities that may want to implement them are provided.

## 2.0 State of Current Methods

### 2.1 Wet Weather Flow Subcatchment Delineation Methods

Wet weather flow has two major components in which GIS plays a major role. Those two components are the delineation of overland flow paths and the creation of subcatchments which is directly related to the first. This is because both are derived from the same topographic source (Olivera *et al.* 2002). When it comes to topographic sources, there are a number of options available from vector-based triangulated irregular networks (TIN) models to contour lines, which are interpolations that indicate elevation. However, the most commonly used of GIS methods are raster or grid-based because digital elevation models (DEM) are now the most common source of topographic information (Garbrecht & Martz, 2000; Wilson *et al.* 2008). As a result, much of the focus on methods for delineating overland flow paths is directed at utilizing raster layers as the topographic input layer.

However, even DEMs have issues. Many of the DEMs that are easily accessible are not of ideal spatial resolution in terms of both horizontal and vertical accuracy. Horizontal resolution refers to the actual size of each grid cell that comprises the surface of the DEM and the vertical resolution refers to the level of precision by which the actual elevation values are recorded for each cell. Hence, both are very important for accurately modelling overland flow paths in urban environments, which are composed of man-made features such as roads, curb lines, and buildings (Shamsi, 2005; Pan *et al.* 2012). In recent years though, datasets created from a relatively new remote-sensing technique known as Light Detection and Ranging (LiDAR), which utilizes airborne lasers, have been developed that can deliver a horizontal accuracy of 10 cm (Fewtrell *et al.* 2011). While higher spatial resolution is usually considered better, considerable time and effort is required in order to transform the raw LiDAR elevation points into a workable DEM, since there is "noise" present in the raw datasets. There are now some tools available in proprietary software such as ArcGIS to address these issues to a certain degree (ESRI, 2010). In addition, because traditional LiDAR relies upon airborne aircraft, it is unable to detect small overpasses, such as an elevated rail line crossing over a roadway, which allows water to flow underneath them (Diaz-Nieto *et al.* 2008). Traditionally, high-resolution orthoimagery or fieldwork would be necessary, but now there is terrestrial LiDAR technology that overcomes such issues by utilizing lasers attached to a moving vehicle (Sampson *et al.* 2012).

Despite these advancements, LiDAR still has two major drawbacks. One is the cost of the actual data, which means that it is not always readily available. The other is that the more detailed a DEM, the less useful it becomes due to the limitations of computers not being able to effectively process and handle such large datasets (Onafrychuk, 2007; Diaz-Nieto *et al.* 2012). However, despite these limitations, LiDAR is increasingly being used in urban drainage model studies. This is because a comparison of the use of various DEMs with varying horizontal resolutions for urban drainage modelling purposes has found that 5-metre horizontal spatial resolution DEMs were too coarse and that a DEM with a 50 cm horizontal spatial resolution was the ideal resolution (Fewtrell *et al.* 2011). However, in most cases, it is simply not possible to secure such high resolution DEMs, so the best available data source is used, even if it means using a 10-metre horizontal spatial resolution DEM (Gironás *et al.* 2010).

There have been a number of methods that have been developed for delineating overland flow paths. However, before most of those methods can be used, it is essential to recondition the DEM to account for manmade features such as roads and culverts that exist in urban landscapes and alter the natural overland flow paths (Hammond & Han, 2006). In order to do so, one needs to use various methods and tools. This includes filling, which raises the elevation of cells considered sinks. It also includes carving or breaching, which lowers the elevation of a surrounding cell so the flow of water can continue. However, in some cases, maintaining the sinks as found in the DEM is beneficial in identifying how much detention storage capacity these sinks provide, although spurious sinks are sometimes present in the DEM and need to be removed (Lindsay & Creed, 2005; Djokic, 2008). The methods for sink filling vary since they may need to be completely or partially filled in, depending on the discretion of the GIS analyst in consultation with the modelling engineer. There are a few tools already built into ArcGIS' Arc Hydro, such as Sink Evaluation, that can be used to determine the capacity of these sinks or depressions that exist on the surface (ESRI, 2011b).

Another main tool is called stream burning or DEM reconditioning and, as can be inferred from the name, it was initially developed for use in natural landscapes in order to enforce the known location of streams or rivers (Djokic, 2008). However, in urban drainage modelling, this same tool can be modified for use in the burning-in of various features, such as the curb lines, roads or underground sewer pipe locations (Baker *et al.* 2006; Luciani *et al.*

2011). The depths of the aforementioned features are then typically specified to be lowered between 0.1 to 5 metres because this "improves the representation of anthropogenic features in urban DEMs" (Gironás *et al.* 2010, p. 2).

Another tool that is used to create reconditioned, hydrologically correct DEMs is the Arc Hydro Build Walls tool that can be used to raise the elevation of cells around features that act as barriers for water, such as the perimeters of buildings or fences. However, this tool was not designed for buildings because when it constructs the wall, it only raises the outside perimeter of the building, therefore causing a sink or depression to develop inside every building polygon on the reconditioned DEM (Djokic, 2008). As such, this tool needs to be modified before it can be used for this purpose. Alternatively, the Raster Calculator tool, which manipulates raster layers using various operators through map algebra expressions, can be used to artificially raise or lower the elevation of certain cells. This is also known as height correction (ESRI, 2010; Djordjević *et al.* 1999). However, all these tools need to be used with caution, particularly when working with datasets of different scales, as it is illogical to have a low resolution DEM, for instance at a scale of 1:100,000 and then attempt to "burn in" features that were drawn at a scale of for instance 1:5,000. Thus, it is essential to use features of comparable resolution when attempting to "burn in" or to raise cells representing any features (Saunders, 2000).

While there are theoretically more advanced algorithms like the tillage controlled runoff pattern model, the canal enforcement, and the road enforcement algorithm, they are not always applicable to urban environments and in some cases is simply an integration of reconditioning (Choi, 2012). The road enforcement algorithm, for example, takes into account the effect of roads by simply accounting for the location of roadways using ancillary data when processing the DEM and then using a simplistic method to create the overland flow paths (Duke *et al.* 2006). As such, these sorts of methods will not be delved into further.

Instead, the focus is on methods that actually differ in how overland flow water moves from cell to cell across the DEM and that appear most promising, because a complete review of all the methods that exist is beyond the scope of this study. While these methods are usually used to determine the overland flow path of water across a surface, in reality they treat water more like a ball rolling across a surface because they do not account for the infiltration of water (Prodanović *et al.* 2009). The D8 method or eight-direction pour point model developed by

O'Callaghan and Mark (1984) is among the earliest and most widely used methods (Shamsi, 2005). In this method, water flows from the centre of a cell in the middle of a  $3 \times 3$  window into the adjacent cell that has the lowest elevation in one of eight possible directions because it is assumed that water follows the steepest path downhill (Olivera *et al.* 2002). While it works well in some areas, it does not perform as well in flat areas (Tarboton, 1997). However, provided the DEM is assumed to have already been reconditioned to include manmade structures such as roads, this should not be an issue because it has been described to work well in areas with well-developed channels (Garbrecht & Martz, 2000). Furthermore, one of the key advantages of the D8 method is that it is faster than all the rest and in this case speed was beneficial because a 10-cm resolution DEM takes  $10^4$  times longer to process than a 2-m one (Bates, 2012).

The Random eight-node (Rho8) algorithm, introduced by Fairfield and Leymarie, (1991) is a single flow direction (SFD) similar to the D8 method. The difference, however, is that the algorithm introduces some randomness into the direction that the water will flow from one cell to the next. Randomness is incorporated into this algorithm because instead of water flowing out of the cell to the one the steepest descent every time the algorithm is run, it incorporates a weighted randomization so that the water is modelled to flow into the adjacent cell in a  $3 \times 3$  window that has the lowest elevation (Lam, 2004). In this research however, variability in the results is not desirable as one of the objectives was to generate results that are objective and repeatable and thus this was not considered a suitable method for this study.

Instead of assigning the flow of water from the centre of one cell into one of eight possible directions on the surface of a DEM, with the FD8 algorithm, the flow is dispersed among multiple cells. The degree of dispersion is based on a weighted system and because of this it is also known as a multiple flow direction (MFD) algorithm (Quinn *et al.* 1995; Lam, 2004). The FD8 algorithm operates on the logic that water flows more readily in the cardinal directions which are West, East, North, South rather than in the intercardinal directions which are northeast (NE), southeast (SE), southwest (SW), and northwest (NW) because it has to travel a shorter distance in those directions. The likely reasoning for such logic is that in these directions, water has to travel a shorter distance. Hence, these directions are given more weight, 0.5 versus 0.35 (Quinn *et al.* 1991). It also considers the tangent of the slope by determining the difference in elevation between the two points while also factoring in the distance from the

centre of the originating cell to the receiving cell (Quinn *et al.* 1991; Quinn *et al.* 1995). However, the method sometimes over-disperses overland flow into all neighbouring cells with lower elevations, which does not occur in reality (Tarboton, 1997). Nonetheless, the FD8 algorithm is commonly used in order to generate input for the TOPMODEL (a TOPography based hydrological MODEL) (Quinn *et al.* 1995; Beven, 1997; Lindsay, 2009).

The Digital Elevation Model Network Method (DEMON), developed by Costa-Cabral and Burges (1994), is considered a stream-tube method for the characteristic tubes it generates. Thus, the flow of water is generated not at point sources that move from the centre of one cell to the next (the point is considered one-dimensional). Rather it is across a surface in what are called stream tubes or flow tubes (unlike the point, the tube is considered to be two-dimensional because it has both width and length), thereby defining a more-likely path for the flow of water (Costa-Cabral & Burges, 1994). In order to generate these flow tubes, the interpolated pixel corner values of the DEM are used and "a plane surface is fitted for each pixel," (Tarboton, 1997, p. 310). Although the tubes are meant to limit the amount of dispersion, some of the flow paths it can generate are still obviously erroneous (Tarboton, 1997). In a study that compared the computation time of various overland flow algorithms, the DEMON method took the longest, requiring almost 24 hours to run on a 2-metre horizontal resolution DEM covering an area of only 36 km<sup>2</sup>. As it is so complex to implement, it is rarely, if ever, used. This is why it has not been used in this study (Tarboton, 1997; Mouton, 2005).

The D $\infty$  algorithm, developed by Tarboton (1997), was an attempt to improve all of the overland flow algorithms mentioned previously. The reason for this is that the D infinity method, allows water to flow into one cell exclusively or up to two cells from an adjacent cell "based on representing flow direction as a single angle taken as the steepest downward slope on the eight triangular facets centered at each grid point" (Tarboton, 1997, p. 309). Similar to the DEMON Method, it does not limit the flow path into one of 8 possible directions, as water is allowed to travel into an infinite number of possible flow directions (Tarboton, 1997). The flow direction angle is determined as being the direction of the steepest downward slope on the eight triangular facets formed in a 3 by 3 grid cell window, centred on the grid cell of interest. Additionally, unlike the D8 method, if the DEM surface is flat, water moves better across its surface because the flow direction is forced to move into a neighboring cell of equal elevation that flows into a

cell of lower elevation directly adjacent to it (Onafrychuk, 2007). Therefore, the presence of erroneous pits in the DEM would not be as concerning as has been found when using the D8 method (Lam, 2004). The  $D_{\infty}$  algorithm was expected to be simple to use due to the development of a free toolbox for ArcGIS, known as Terrain analysis, using Digital Elevation Models (TauDEM) Tools (Tarboton, 2003). However, it is not being used in this study because, although it is more efficient than the DEMON Method, it was found to be more computationally demanding than the D8 algorithm (Mouton, 2005).

The final major method that is reviewed here is among the most recent developments (Pilesjö & Hasan, 2013) in the field known as the Triangular Form-based Multiple Flow (TFM) algorithm. The TFM algorithm, being an improvement on the  $D_{\infty}$  algorithm, divides single cells into eight triangular sections in order to estimate flow paths on the surface of a DEM in a similar manner to the  $D_{\infty}$  algorithm. Specifically, a grid cell is partitioned into eight triangular centre facets, in which the lines extend from the centre of the cell in question to the centres of the eight neighbouring cells in a  $3 \times 3$  window. Thus, each of the eight facets is created from connecting the centre cell to two other adjacent cells. A constant slope and aspect (i.e. slope direction) are defined for a resulting facet from mathematical formulations. Thus, the overland flow is routed out of the facet and into a neighbouring facet. Estimations can be made in terms of flow direction and divergence/convergence much more intuitively than in other raster-based solutions. However, because it is such a recent innovation, implementing this method required a copy of the source code, which unfortunately could not be secured from the creators of the algorithm. Nonetheless, the reason why it is being featured here is that it appears to be a very promising method. Especially because the creators have indicated that in the source code, they have developed solutions to address flat areas and man-made barriers, both of which are found in urban environments and might reduce the amount of reconditioning that is required (Pilesjö & Hasan, 2013).

Once overland paths have been determined, they can then be used to delineate the subcatchments. While the manual method of WWF subcatchment delineation is viewed as being the most accurate, it is also problematic. Not only is it very time consuming but it is also very subjective which raises questions of its true accuracy. With the manual method, the one delineating flow paths would have to be delineated by simultaneously considering information

contained within various datasets. This includes the location of buildings, parcels, pipes, catch basins, roads, orthoimagery and contour lines in an attempt to delineate subcatchments, as they would not have an overland flow path. Consequently, they would need to use these layers to determine the most likely flow of water, which would define the extent of each WWF subcatchment (Chen *et al.* 2003; Jankowsky *et al.* 2012). In the manual approach, it is typically assumed that water will follow the steepest path down, therefore draining into a pipe if it has catch basins attached to it. Therefore, with the manual method, if there are several parcels that are sloped such that water will flow towards their backyard where there is a catch basin attached to a storm pipe, it is assumed the water will flow to this pipe and so it is delineated accordingly. In most cases however, the overland flow of water is assumed to exit the property through the front yard or driveway and onto the street since most of the properties are graded in the city so that stormwater will flow on to the street. This could be confirmed by checking the contour lines, which can be a very subjective and time-consuming process. A further issue with the manual method is that it does not always reflect the full extent of water's overland flow path because manmade parcels are used as the boundaries of the subcatchments.

Automated methods, while not perfect, offer the advantage of being much more objective because each time the algorithm is run, the same catchments are delineated (Baker *et al.* 2006). Perhaps the most widely used delineation method relies upon a "DEM-based" approach. This method utilizes the algorithms covered to trace the most probable flow paths depending on the information within the DEM (Mark *et al.* 2004). Such an approach can be implemented utilizing the Arc Hydro data model used by the Toronto and Region Conservation Authority (TRCA, 2013) and the State of Florida (Foley, 2008). The Arc Hydro data model extension was designed to support GIS-based hydrologic simulation models of natural landscapes and because it is in the public domain, it is freely available for use. Thus, it provides a framework that conveniently packages most of the ArcGIS tools required to generate subcatchments that go beyond what is provided in the standard Hydrology Toolbox though this too has been used to generate subcatchments (Luciani, 2005; Dongquan *et al.* 2009). For example, a flow direction grid can be generated using the D8 method, which can then be used as the basis for creating subcatchment polygons by following a series of processing steps that are provided as built-in functions (Maidment, 2002; Johnson, 2009). Furthermore, it has been used in places such as Australia in the delineation of urban subcatchments (Barron *et al.* 2011). Nonetheless, one major

disadvantage of this method is that if unmodified or based on default settings, it does not generate subcatchments that can have a predefined number of catch basins (e.g. one catch basin) or manholes in them as it was originally intended for natural landscapes. This is because it is based solely on the DEM, though this can be addressed by modifying the DEM itself as mentioned earlier (Olivera *et al.* 2002). Such an approach appears to have also been used by a consultancy that had previously created a distributed model for a portion of the City in their creation of WWF subcatchments (Clarifica, 2010). However, much uncertainty remains on what was done and how it was accomplished. As a result, this research also sheds some light on what was possibly done as this was of great interest to City staff at Toronto Water (G. Lin, personal communication, February 13, 2013).

The distance-based method overcomes the problem of not being able to control the number of catch basins in subcatchments by using the location of the minor system or sewer system inlets such as the manholes or catch basins, as a basis for delineating subcatchments automatically. Thus, it essentially uses a Thiessen polygon approach to define the drainage area of each catch basin. However the disadvantage of this method is that it ignores the topography of the surface and therefore does not have a direct relation to the overland flow paths that were generated (Mark *et al.* 2004).

## **2.2 Dry Weather Flow Methods**

A manual method for DWF subcatchment delineation is currently employed by the City of Toronto. Other approaches to subcatchment delineation have been presented by a team of researchers using Cincinnati, Ohio as a case study (Chen *et al.* 2003), the Massachusetts Water Resources Authority (MWRA) (Shamsi, 2005) and the developers of InfoWorks (Innovyze, 2013). Shamsi (2005) suggests that other municipalities have been involved in similar automation activities, but details do not seem to be available.

It is the current practice that for simple cases, the City of Toronto's modellers manually delineate a sanitary subcatchment on the basis of linking all parcels to the nearest upstream manhole or node based on the parcel's frontage or where the driveway is located. These locations are derived from the City's high-resolution orthoimagery (10cm-spatial resolution). They then construct a subcatchment polygon that includes all these parcels and encompasses the portion of road between the parcels to which each upstream node or manhole is linked (Figure 8). However,

because the city is so large, such an approach is impractical to be continued in the long-term which is why an automated approach is desirable. Furthermore, the modelling approach is facilitated by the use of ArcGIS Model Builder, since it is not only object-oriented (Allen, 2011), but serves as an effective tool for communicating the logic of an automation procedure. This means that an individual without much programming experience can not only easily design a tool, but can also modify an existing one. All that is required is knowing which specific order is needed for sequencing the GIS functions and whether it is necessary to use any special ModelBuilder functions, because these functions can merely be dragged and dropped from the ArcGIS toolbox. Therefore, this design platform was used in order to meet the objective of designing tools that would be easy to modify in the future.



**Figure 8. A single manually delineated DWF subcatchment**

The Watershed method is innovative in that it utilizes a DEM to create DWF subcatchments. With this method, the user burns in only the sanitary flow portion of the sewer network as though it were a stream while maintaining pipe flow-direction based on the invert values (buried depths) of the ends of the pipe. After that, the modified DEM is treated as though it were being used for delineating overland flow since a flow direction grid is created using the flow direction function in ArcGIS in the standard hydrology toolbox. The resulting output is then used as the input to create the subcatchments as a raster. They are then converted into polygons using the Watershed function in ArcGIS. While this is largely an automated approach, its major

weakness is that it does not follow the boundary of the land parcels (Chen *et al.* 2003). Unlike the WWF, the flow source is household-based and it is appropriate to associate this with individual land parcels.

It should be noted that the Hydrology Toolbox also includes “Basin” and “Watershed” tools. The only difference is that instead of the Watershed function being applied to the modified DEM, the Basin function is used which is normally used to create “a raster delineating all drainage basins” (ESRI, 2010). The Watershed function differs from the Basin function in that it “determines the contributing area above a set of cells in a raster” (ESRI, 2010). Therefore, while this raster-based approach is also largely an automated process, it too fails to follow the boundary of the land parcels and therefore has been suggested to be better suited for delineating WWF subcatchments than DWF subcatchments (Chen *et al.* 2003).

The Euclidean Distance approach developed by the Massachusetts Water Resources Authority involves 5 major steps that are reproduced here in their entirety (Shamsi, 2005, p.102):

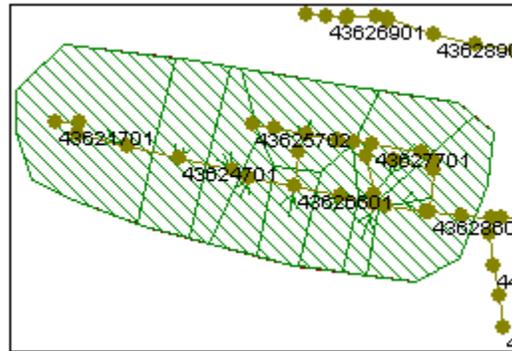
1. All the sewers tributary to a pour point (point where a trunk sewer connects an interceptor) are identified using a custom network-tracing utility (upstream trace). The group of pipes identified like this defines a tributary area.
2. A new attribute with the name of the tributary area is manually added to all the sewers.
3. Vector sewer layers are converted to a grid format.
4. The Euclidean distance allocation function is applied to all the sewer cells. The Euclidean distance is calculated in the same manner as finding the hypotenuse of a triangle. This generates a sewershed boundary around sewers.
5. The raster Euclidean distance boundaries are converted back to vector (polygon) format.

(Shamsi, 2005, p. 102)

While this method is an improvement over the manual method and the two DEM-based approaches, it does not always produce the most logical results and can be potentially quite time-consuming because of the manual additions that have to be made in order for it to run (Shamsi, 2005). Therefore, it is perhaps not the most suitable method for applying on a large scale as is needed here. Furthermore, although all the steps in the implementation of this method have been provided above, it remains somewhat incompatible with the City of Toronto’s approach.

In the InfoWorks CS model, used by the City of Toronto, the automated approach to creating subcatchments relies upon Thiessen polygons. To begin, the modeller is required to manually create or provide a bounding polygon. The modeller can then choose which pipes to create subcatchments for by selecting their upstream nodes while also making sure the bounding polygon is also selected and then using the “Create within Selected Polygon” function to create a

Thiessen polygon-based subcatchment for each selected pipe via its upstream node (Innovyze, 2013). The major issue with this method is that it does not follow the parcel boundary layer as is desirable for cities such as the City of Toronto and produces primitive looking results (Figure 9).



**Figure 9. DWF subcatchments automatically delineated by InfoWorks CS**

Currently, the most promising of existing known methods is the proxy method because according to Chen *et al.* (2003) that developed it, when they used the manual method to verify the accuracy of the results generated by their three automated methods, the results of the proxy method were most similar to those of the manual method. With the proxy method, a scripting program (created in the Avenue programming language) was used to generalize the sewer network that included pipes and loading points where sanitary flow entered the sewer system. Proximity analysis tools were then used to link each parcel to the nearest segment of a generalized sanitary flow pipe and all the parcels that were linked to a common pipe or loading point were dissolved (Chen *et al.* 2003). The dissolve function essentially merges vector features that share a specified common attribute together and in this case, it was all parcels that shared the identical nearest pipe information. Therefore, if five parcels shared the same attribute for the nearest loading point, they would be merged into one large parcel. However, the creators of this method fail to specify how this was done or provide any of the source coding (Chen *et al.* 2003). In addition, the Avenue programming language is outdated and therefore no longer widely used (D. Banting, personal communication, December 16, 2013). Furthermore, this method also fails to address the proper division of the road area among DWF subcatchments as is required by municipalities such as the City of Toronto, for use in their Inflow and Infiltration (I & I) calculations. Nonetheless, this method is still valuable because it utilizes the boundaries of the land parcels layer to delineate the extents of the DWF subcatchments (Chen *et al.* 2003), much as

it is done when they are manually delineated by staff at the City of Toronto or by consultants on their behalf (G. Lin, personal communication, February 6, 2013).

### **2.3 Rainfall Interpolation Methods**

Interpolation involves "deriving a curve, or surface, which passes through a limited number of measured data points in order to predict or determine values of an experimental parameter in between the measured data points" (Spangler, 1992, 305). Hence, interpolation methods extend far beyond their uses in rainfall interpolation to include areas of speech synthesis, geological studies, and automotive engineering just to name a few (Spangler, 1992). However, since this research was focused on how it can be applied to interpolate the rainfall data collected from the City's rain gauges to create virtual rain gauges, only some of those interpolation methods that were currently being used in this field were considered. Before proceeding, it is important to note that two main families of interpolation methods exist. The first are deterministic interpolation methods which are "directly based on the surrounding measured values or on specified mathematical formulas that determine the smoothness of the resulting surface" such as Spline. The second family of interpolation methods consist of geostatistical methods like Kriging (ESRI, 2010).

The interpolation method currently used by the City of Toronto is the Thiessen method. In this method, polygons are generated which are arbitrary vector features that surround each input rain gauge point location, such that any location within this polygon is closer to its associated rain gauge point location than to the rain gauge point location of any other polygon (Lhomme *et al.* 2004). While the Thiessen polygon is very simple, it quantizes space into discrete polygons centred on the point at which its key attribute had been observed. (Ruelland *et al.* 2008). Though these polygons can be converted to a raster, the real issue is the underlying assumption that all the space within the polygon receives the same amount of rainfall, with no way to easily test this assumption. Therefore, the observed reality that rainfall (and other atmospheric) attributes vary gradationally, undermines confidence in the assumption that these generated rainfall zones are in fact valid (D. Banting, personal communication, December 16, 2013). Because of its usage (Figure 3), City staff were only able to generate about 18 possibly unique values at any one time, which according to them and the literature, was likely too coarse

when attempting to develop input for a 1-D distributed urban drainage model (Chen & Liu, 2012, G. Lin, personal communication, February 13, 2013).

The isohyetal method, like the previous method, uses an area-based weighting scheme, used to infer the locations of lines of equal precipitation. Consequently, these isohyets are analogous to contour lines, but instead of being used to indicate areas of equal elevation, they are used to indicate areas of equal precipitation (Johnson, 2009). However, once again, a continuous surface is not created except with further interpolation, so this method would likely produce results that would not be suited for use by the City of Toronto.

Of those interpolation methods that produce a continuous surface or raster by default, which would be more suitable in addressing the needs of the City of Toronto, three are reviewed here. The inverse distance weighted (IDW) method operates on the basis that observed values that are near to one another on the ground, tend to be more similar than those further apart (Piazza *et al.* 2011). Therefore, the amount of rainfall at a non-gauged location can be approximated by a weighted average of observed rainfall values within a radial search neighbourhood. Its radius can be defined by the range of a fixed number of rain gauges that are typically inversely proportional to the square distance between the observed rain gauge values and at non-gauged locations. As a result, it has the tendency of producing a bull's eye pattern around the input point locations (Ruelland *et al.* 2008; Soenario & Sluiter, 2010). Furthermore, the values of the input points are lost, though all the values generated from the interpolation are within the range of the maximum and minimum values that existed in the input dataset. Overall, the IDW method for rainfall interpolation remains an effective approach because as demonstrated by its recent use in Taiwan to interpolate rainfall in its urban areas and it achieved correlation coefficient values of over 0.95 (Chen & Liu, 2012).

Unlike the other methods being reviewed here, Kriging is a geostatistical method that is based on probability models, which means it allows for some measure of the accuracy of the values predicted for the resulting raster. It assumes that the distance or direction between sample points reflects a spatial correlation that can be used to explain variation in the surface, which is represented by a variogram (Hutchinson, 1998; ESRI, 2010). Kriging, while considered the most advanced method in regard to interpolating rainfall values, does have its drawbacks such as

requiring prior calibration of a variogram using three parameters known as the range, nugget and sill and, in some cases, this is difficult to do even for simulated data (Hutchinson, 1998).

Another possibly suitable interpolation method was to use a spline function, which unlike Kriging, maintains the original rain gauge precipitation values in the output raster. Conceptually, it can be thought of as bending a rubber sheet to pass through all the input points, while at the same time minimizing the total curvature of the rubber sheet's surface so it is less computationally intensive than Kriging (Ruelland *et al.* 2008).

In addition, while all these methods already exist as built-in functions in ArcGIS and other GIS software programs, what was lacking was a means of automating the interpolation to run thousands to tens of thousands of times. The reason for this is that the shortest time step at which rainfall is usually interpolated is for 1-hour intervals for a single storm at a time and so running an interpolation thousands of times is rare (Haberlandt, 2007; Ly et al. 2013). However, in this case, the City of Toronto wanted to find a way to rapidly interpolate their rainfall records on 5-minute intervals from April to November, as this is the period of time during which their rain gauges are operational (G. Lin, personal communication, February 13, 2013). While it might not be advisable to attempt to interpolate rainfall on such a short time scale, seeking of first-hand evidence of its feasibility was undertaken.

### 3.0 Automated Approach to Dry Weather Flow Subcatchment

#### Delineation

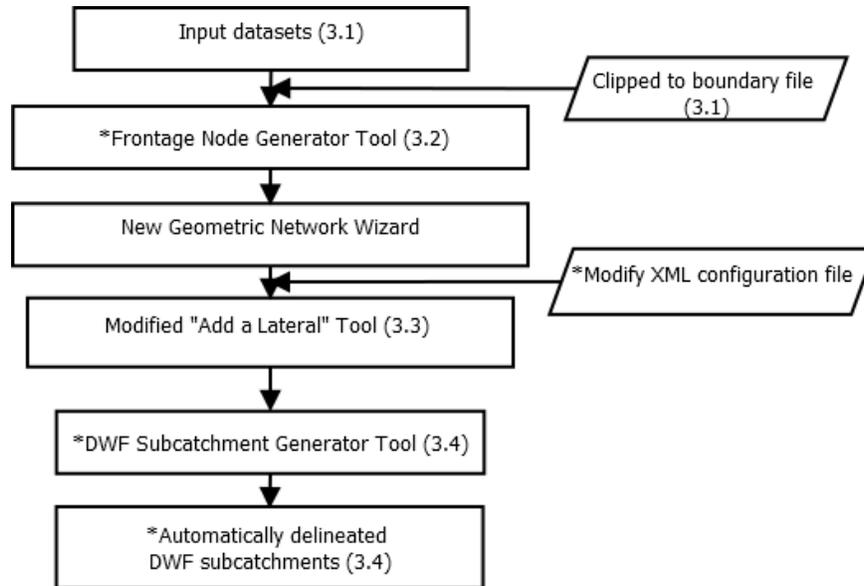
The automated approaches, which were suggested in the published literature (Chen *et al.* 2003; Shamsi, 2005; Innovyze, 2013) were not entirely suitable. Thus, modifications had to be made, and a new approach that relies upon ModelBuilder and the Infrastructure Editing Toolbars in ArcMap (ArcGIS for Local Government team, 2013) was created that conforms to the assumptions laid out in Table 1. While many iterations of this tool were developed, the approach presented here was the most promising and was applied to the Coxwell model pipe network.

**Table 1. DWF subcatchment assumptions**

Assumptions
1. Each parcel connects to the nearest upstream node of a sanitary or combined pipe based on its frontage or driveway because it is assumed that the sanitary lateral is located under it.
2. Each parcel can only connect to one pipe.
3. Each subcatchment conforms to the parcel boundaries.
4. The road corridor proportion containing each upstream node location is part of the subcatchment.
5. There should be no gaps or overlaps between subcatchments.

The major components of the approach are summarized in the flowchart in Figure 10. The numbers enclosed in brackets indicate the section where further details regarding that step or process is located in a manner that attempts to allow for ease of its application in other jurisdictions. Those missing such numbers are located, in the Appendix (A.1-A.2). Upon inspection of Figure 10, it may seem counterintuitive to not integrate the clip function into the tools. However, this was done to minimize any errors. The reason for this is that if the input boundary provided was not perfectly aligned with the other layers, such as the parcels layer, it could generate slivers through the clipping process. Hence, by doing this step separately as seen in Figure 10, it was possible to check for the occurrence of such slivers and therefore minimize their presence early on. Those processes or steps prefixed by an asterisk (\*) indicate the areas of

contribution made by this research and thus cover the tools and or special adjustments made that make the overall approach unique.



**Figure 10. DWF method schema**

### 3.1 Required Datasets

Table 2 identifies the key datasets, their data type and any relevant key attribute fields as developed for application to the City of Toronto, that were used by the tools. It should be noted that the sanitary and combined pipes layer and manhole locations layer used to create and test the tool had been checked by the modeller. This is why they were exported from the InfoWorks CS model and not TWAG, which is maintained by City staff at Toronto Water, as can be seen in Table 2. In addition, the road corridors layer noted in the table below does not only include the area covered by the road itself, but the area beyond the parcels. Hence, it also includes the sidewalk areas, swales and other possibly impervious surfaces. However, the city parcels layer and road corridor layer area were created and maintained by other divisions at the City of Toronto and so limited quality control adjustments could be made to these datasets. As the final iteration of the tool required the input of a defined study area boundary layer, and only a definitive study boundary for the Coxwell network could be supplied, only this network was used in evaluating the final iteration of the tool.

**Table 2. Layers required to run DWF subcatchment delineation tool**

<b>Geospatial File Name or Description</b>	<b>File Type</b>	<b>Key Attribute Fields</b>	<b>Source</b>
1. City parcels layer	Polygon shapefile	1. Land use type field	Grace Lin, Senior Modelling Engineer at Toronto Water, City of Toronto (City of Toronto, 2006)
2. Sanitary and combined pipes layer	Line shapefile	1. UP Asset ID field 2. Down Asset ID field 3. Asset ID field 4. Flow Type Field	Exported from InfoWorks CS Coxwell Network Model
3. Road corridors layer	Polygon shapefile	N/A	Grace Lin, Senior Modelling Engineer at Toronto Water, City of Toronto (City of Toronto, 2010)
4. Manhole locations layer	Point shapefile	Asset ID field which corresponds to the UP Asset ID and Down Asset ID fields of each pipe.	Exported from InfoWorks CS Coxwell Network Model
5. Boundary of study area layer (Coxwell Network)	Polygon shapefile	N/A	Dr. Celia Fan, Modelling Engineer, Member of the Ryerson team responsible for creating the InfoWorks CS model through which this research was made possible.
<p><b>Spatial Reference Shared by all layers to be consistent with the GIS data standards of the City of Toronto:</b>                      Projected Coordinate System: MTM 3Degree                      Projection: Transverse Mercator                      Geographic Coordinate System: GCS North American 1927                      Datum: North American 1927</p>			

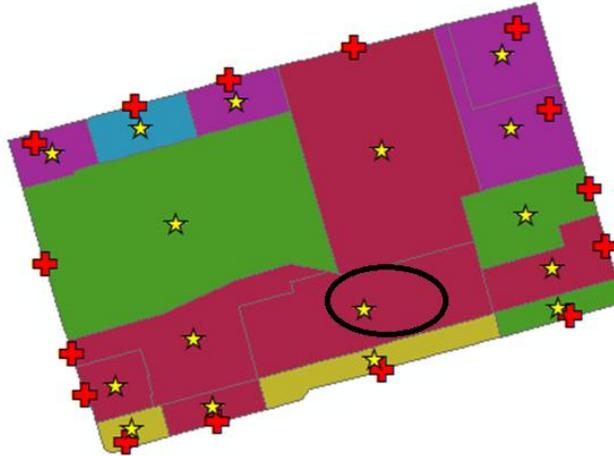
In order to begin the process of automatically generating delineated DWF subcatchments, all datasets were clipped to the boundary area polygon. This was done to minimize processing time and to ensure subcatchments were created where necessary and nowhere else.

### 3.2 Frontage Node Generator Tool

A tool was created using ArcGIS ModelBuilder as the initial step in connecting land parcels and their roadside frontage for DWF network-and catchment-delineation. To program this, a new layer of data elements was needed, such that each lot and its adjacent road segment would be identified at a node. This Frontage Node Generator Tool incorporated the logic that the sanitary outflow pipe of each building in the City of Toronto normally leaves a parcel via its frontage, the side on which the building faces onto the road or the side of the lot that faces towards the road. Other layers, such as the clipped road corridor and parcel layers were put through a Frontage Node Generator Tool created in ModelBuilder. This tool actually creates a 0.5-metre full buffer around the road corridors layer, which in turn creates an area of overlap with each adjacent land parcel. The intersect function is then applied creating a new feature of the overlap of two separate but intersecting layers, from which the centroid or a point within this newly created feature is extracted (Figure 11). Of course, this tool was not successful for all parcels since a small number of lots were surrounded by parcels on all sides. In such situations, the centroid or a point within each parcel was used (Figure 12) and so the tool was designed to automatically merge these points with the rest of the frontage nodes.



**Figure 11. Frontage nodes automatically generated by the frontage node tool**



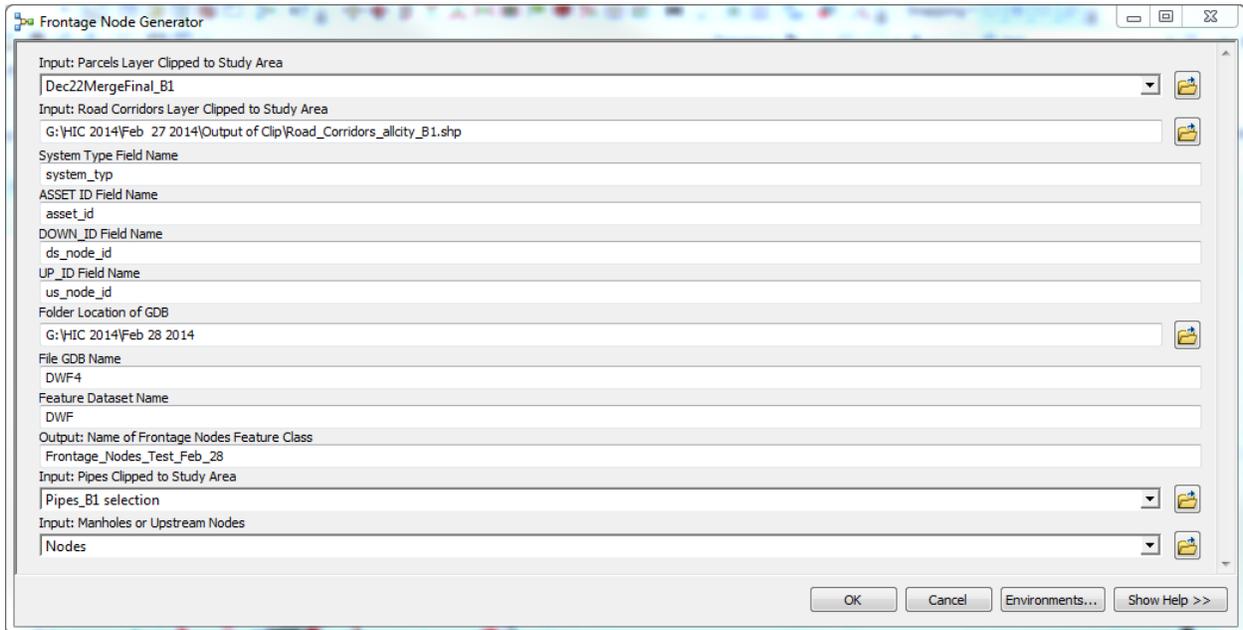
**Figure 12. A case where the frontage node tool would not output the frontage node but instead output the centroid**

In addition, the tool also automatically added new fields to this new frontage node point layer, which was to associate each parcel with the nearest pipe. While it is possible to add or remove fields as necessary, in this case, the tool was designed to automatically add 4 additional fields to the frontage node layer. These layers were as follows:

1. `us_node_id`: Corresponds to a field with this name in the pipe layer that identifies the upstream node id or upstream manholes of each pipe. The reason why this is included is that this field is a required input field for the InfoWorks CS modelling software and is used to identify each subcatchment.
2. `ds_node_id`: Corresponds to a field with this name in the pipe layer that identifies the downstream node id or downstream manholes of each pipe. The reason why this field is included is for reference purposes only and not to be used by the model and therefore could be dropped later.
3. `system_typ`: Corresponds to a field with this name found in the pipe layer that identifies whether it is either a combined or sanitary flow pipe. Field is a required input field for the InfoWorks CS modelling software.
4. `asset_id`: Corresponds to a field with this name found in the pipe layer that identifies the asset id of each pipe. The reason why this field is included is for reference purposes only

and is not used by the model since there maybe two pipes stacked, one on top of the other, underground sharing the same downstream node id and upstream node id. By having this field, it allows the program to distinguish between these two pipes in such cases.

Aside from automatically outputting the frontage node layer into a feature dataset in a new geodatabase, this program also generated an empty Sewer\_Lateral\_Line feature class within this same feature dataset. As well, it automatically imported the pipes and nodes layers into this new feature dataset. It should be noted, that this tool also automatically checks whether there is any layer without any frontage node due to its location, such as seen in Figure 12, and generates a point that is then merged and outputted in the resulting frontage node layer. The interface for this Frontage Node Generator Tool and a ModelBuilder diagram outlining its inner workings that have been explained above can be seen in Figure 13 and Figure A3 respectively. A geometric network was then created based on the output of the Frontage Node Generator Tool with a detailed explanation of its implementation being provided in the Appendix.



**Figure 13. The user input interface for the Frontage Node Generator Tool**

### **3.3 Modifying and Running the "Add a Lateral" Tool**

The Infrastructure Editing and Infrastructure Reporting Toolbars were used to generate the pseudo sewer lateral lines used to create the DWF subcatchments. They are part of the freely

available Water Utility Network Editing package from ESRI (ArcGIS for Local Government team, 2013). While there were perhaps easier ways of fulfilling the same purpose that was provided by the "Add a Lateral" function in this toolbar, this approach was still chosen. This is because the toolbar provides additional features such as the trace upstream and downstream pipe functions that can be used to check pipe connectivity by automatically generating a line that highlights the relevant pipes. These are also useful for data quality confirmation and are not currently being utilized by the City of Toronto.

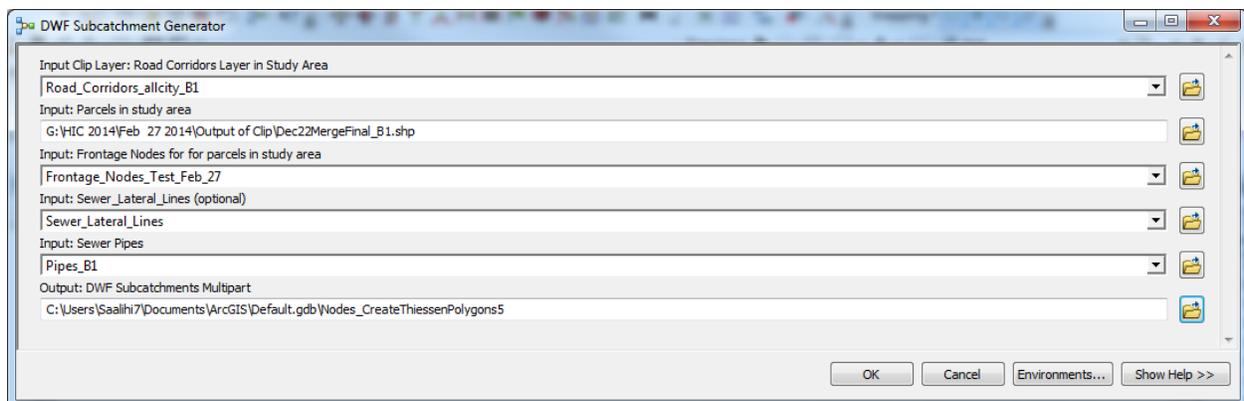
In order to better understand the benefit of this tool, it is necessary to explain how datasets are stored and used at Toronto Water as related to urban drainage modelling. TWAG is a geodatabase with well over 30 unique layers that includes all the City's water related assets such as the locations of pumps, treatment plants, valves and pipes. Its main purpose is to be used as a source of locating these water related assets for maintenance purposes. In order to create the sub-datasets that comprise TWAG, such as pipes, a team of engineering technologists use as-built drawings to manually input each feature within the ArcGIS database, which is being continually updated. These datasets are then optimized for input into InfoNet, which is a "purpose-built Infrastructure Management System (IMS) for water distribution, wastewater collection and stormwater networks" (Innovyze, 2014). This is then normally used for its trace upstream and downstream pipe functions. Since the data from TWAG are not optimized for modelling, they need to be modified by modellers within either ArcGIS or InfoWorks CS, so that it can be used for modelling purposes.

To use the "Add a Lateral" function to create pseudo sewer lateral lines, the configuration file, which is an XML file, had to be adjusted and a geometric network had to be created as seen in the Appendix. The adjustments made allowed an artificial sewer lateral line to be automatically generated from each frontage node to its nearest pipe and also automatically populated the four fields that were added into the frontage node layer by copying the appropriate information from the pipes layer.

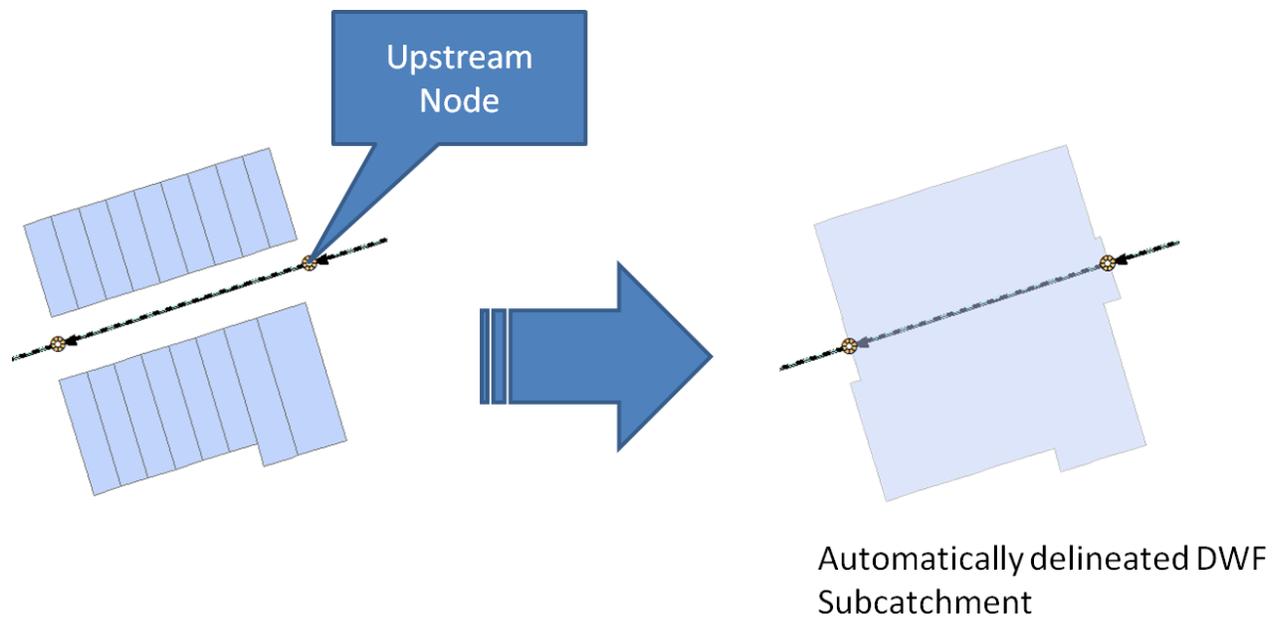
### **3.4 DWF Subcatchment Generator Tool**

The DWF Subcatchment Generator Tool (Figure 14) was created to have two main purposes. First, it merges all parcels that are associated to a common pipe together. Second, it assigns a portion of the road to the pipe associated to each subcatchment and then merges the

road and parcels together to create a subcatchment for each pipe. The program operates on the logic that while the frontage node that is representative of each parcel has had all the pipe information added to it during the generation of the laterals, this information has not yet been transferred to the parcels. Consequently, it was designed to automatically transfer this information to the parcel via a Spatial Join function. The Spatial Join function in ArcGIS essentially combines attributes from one feature to another based on a spatial relationship (ESRI, 2010). In this case, it was based on which parcel each node intersected. Furthermore, the target features and the joined attributes from the join features are written to a new output layer. Since the upstream node defines the region of a road that is crucial to each subcatchment, (in the approach used by the City of Toronto) the tool attempted to ensure that this portion of the road was maintained as part of each subcatchment. Attempts to capture both the up and downstream nodes in all cases only succeeded in some cases (Figure 15).



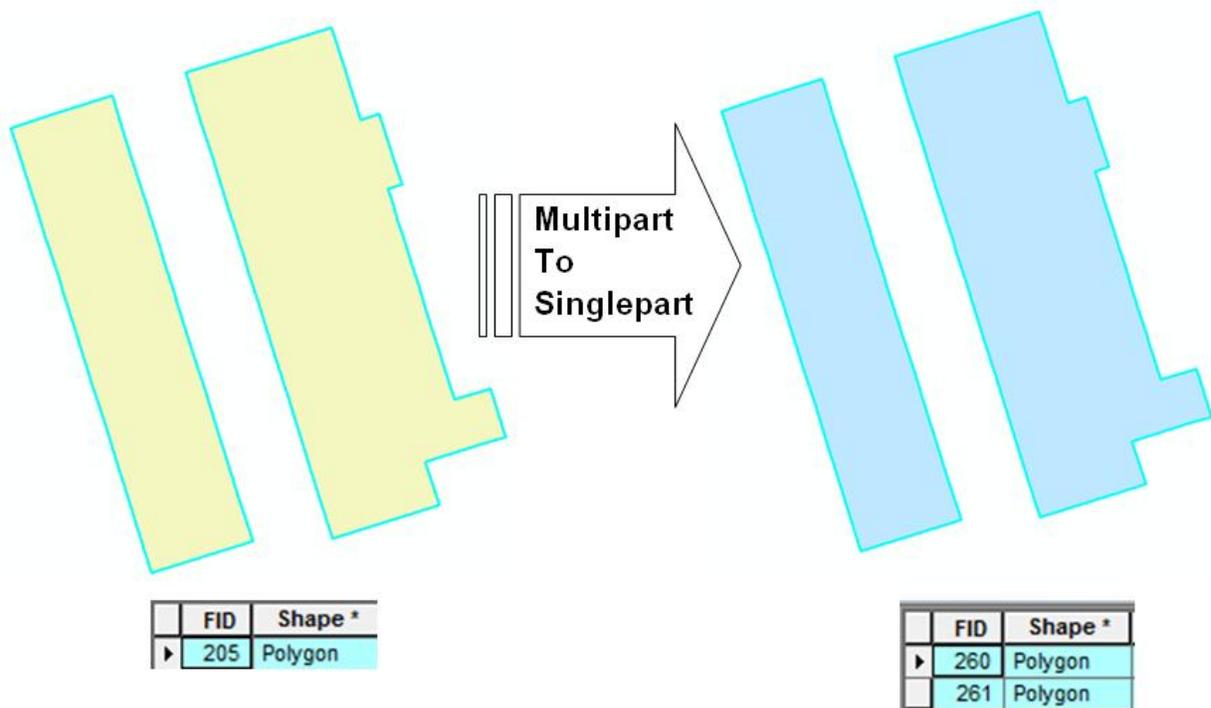
**Figure 14. User input interface for the DWF Subcatchment Generator Tool**



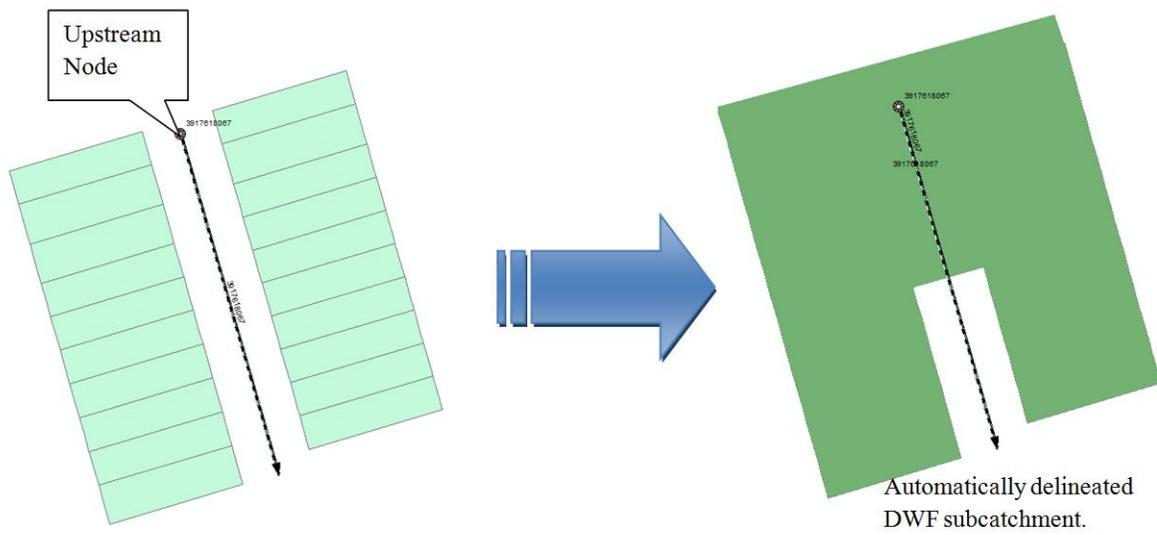
**Figure 15. DWF subcatchment created when using an earlier version of the DWF Subcatchment Generator Tool**

In order to ensure the preservation of the road area, allocated to the upstream node for each subcatchment, the DWF Subcatchment Generator Tool takes the populated Sewer Water Laterals layer generated using the "Add a Lateral" function. It then uses a Select By Location ArcGIS function to identify which pipes were touched by the boundary of the lateral layer, since not all pipes had parcels connected to them. It was for only these selected pipes that pseudo upstream nodes were created and upon which the Create Thiessen Polygons function was applied. The inspiration for using Thiessen polygons came from trying to utilize the best aspects of the automated approach available in InfoWorks CS. Since applying the Thiessen polygon function resulted in a layer that covered the entire study area and not the just roads region, the road corridor layer that was input was used to automatically clip it to only the road portion. Afterwards, the two layers were merged together; these were the Thiessen polygons layer that was clipped using the roads layer and the parcels layer with the pipe information written to each parcel. After the ArcGIS Merge function, they were then combined using the ArcGIS Dissolve function, based on three common attributes, which were fields, added to the frontage nodes initially (`us_node_id`, `asset_id` and `system_typ`). The output of this Dissolve function can be considered as being Multipart DWF subcatchments that were automatically delineated by this automated approach. If there were any gaps, then all that had to be done was to apply the ArcGIS

Erase function on the original boundary polygon. Afterwards, the layer resulting from the Erase was merged to the subcatchment layer and then the Eliminate function was applied. In order to make sure that no multipart features existed or features that share one attribute table entry but were composed of multiple features (Figure 16), the ArcGIS Multipart To Singlepart Function was also applied and then the Eliminate function was used once again to create Singlepart DWF subcatchments (Figure 17). A complete ModelBuilder diagram of the tool can be seen in Figure A4.



**Figure 16. How the Multipart to Singlepart Function in ArcGIS converts a multipart feature into separate features**



**Figure 17. DWF subcatchment automatically delineated by the tool**

## 4.0 Automated Approach to Wet Weather Flow Subcatchment

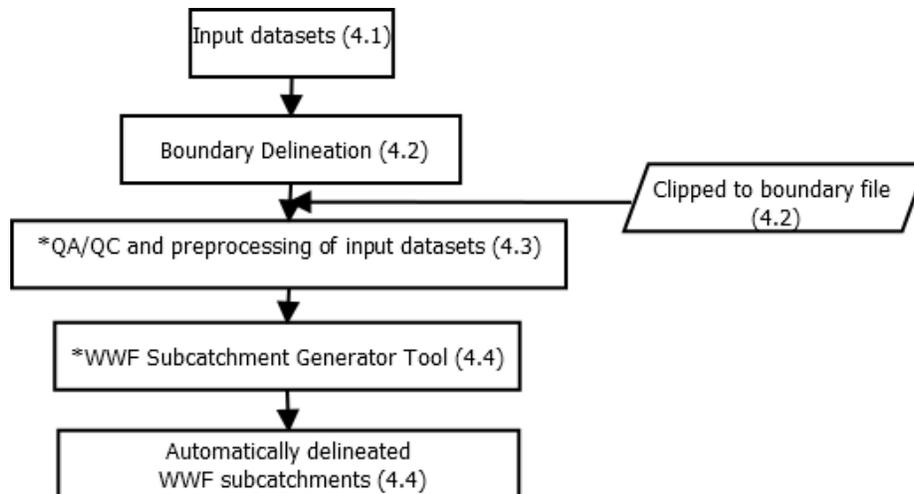
### Delineation

The WWF approaches that were currently available, including those through Arc Hydro version 2.0 for ArcGIS 10, were not suitable without modification to meet the needs of cities, so alterations were made based on the assumptions in Table 3. Hence, the major contribution of this research has been the development of a new automated approach that meets such needs by combining the strengths of the Arc Hydro data model such as its ease of use, with those of the distance-based method. This has been achieved by relying upon both Arc Hydro and ModelBuilder.

**Table 3. WWF subcatchment assumptions**

Assumptions
1. The subcatchments are based on the reconditioned DEM which includes the location of buildings, pipes and roads.
2. Subcatchments are only created for pipes with catch basins correctly associated to them in TWAG.
3. The subcatchments are delineated assuming a rate of 100% roof disconnection.
4. The area outside of the roof area of each subcatchment is assumed to reflect the flow of water on the surface based on the D8 algorithm.

The major components of the approach developed are summarized in the flowchart diagram seen in Figure 18. From this flowchart, it is apparent that this approach relies upon a very comprehensive approach, because all the automated processes are integrated into a single tool, which turned out to be one of the more innovative aspects of this research. It has been able to take a relatively complicated process and simplify it into a straightforward one. In addition, while quality assurance and quality control measures were taken for all of the datasets used in this research, it was especially important for this tool and is why it has been explicitly indicated in Figure 18. It was particularly important because it was in this step that certain measures were taken such as ensuring that the pipes exported from the two models were correctly merged together which helped to minimize potential errors as detailed in section 4.3.



**Figure 18. WWF method schema**

#### 4.1 Required Datasets

Table 4 identifies the key datasets, their data types and relevant key attribute fields required to use the tools. It should be noted from this table that certain layers such as the DEM were supplied by individuals at the City who in turn received them from other staff at the City over the years. Unfortunately, many of these layers lack appropriate metadata or were created especially for this research, which is why no metadata has been provided. From this table, it is also apparent that three discrete sewer pipe networks were utilized. In addition, from this table it is also apparent that two sources of building layers were utilized, each representing the footprint of a building as a polygon. Both were utilized because each contained buildings not contained in the other. Thus, both layers were merged together by first applying the Erase function as was used when merging the various pipe layers as detailed in section 4.3, to minimize any errors such as duplicate buildings.

**Table 4. Layers required to run WWF subcatchment delineation tools**

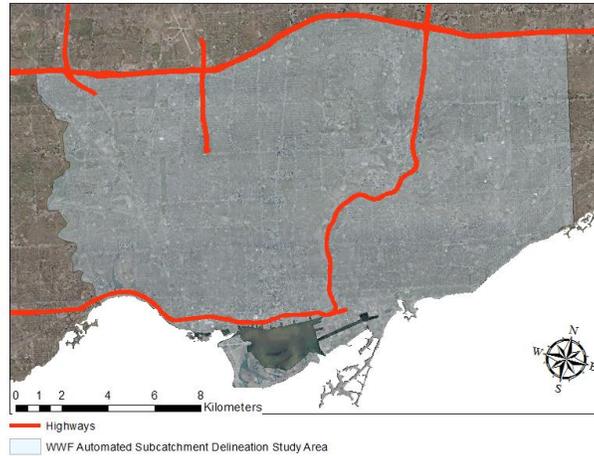
<b>Geospatial File Name or Description</b>	<b>File Type</b>	<b>Key Attribute Fields</b>	<b>Source</b>
3-metre horizontal resolution and 10-centimetre vertical resolution, Digital Elevation Model of Toronto	Raster File	VALUE which indicates the elevation value of each cell measured in millimetres	Jaime Aldana Engineer, Toronto Water, City of Toronto (City of Toronto, 2011a)
Active storm and combined pipes layer # 1 Coxwell Model	Line shapefile	1. UP Asset ID field 2. Down Asset ID field	Exported from Coxwell InfoWorks

Network		3. Asset ID field 4. Flow Type Field	CS Model
Active storm and combined pipes layer # 2 Interceptor Model Network	Line shapefile	5. UP Asset ID field 6. Down Asset ID field 7. Asset ID field 8. Flow Type Field	Exported from Interceptor InfoWorks CS Model
Active storm and combined pipes layer # 3 (these are pipes which are defined as being active within TWAG and not abandoned or not in use)	Line shapefiles	1. UP Asset ID field 2. Down Asset ID field 3. Asset ID field 4. Flow Type Field	TWAG (Toronto Water, 2014)
Catch basin locations layer	Point shapefile	N/A	TWAG (Toronto Water, 2014)
Catch basin leads layer (the pipe that connects the catch basin to either a combined or storm pipe)	Line shapefile	N/A	TWAG(Toronto Water, 2014)
Manhole locations layer # 1 Coxwell Model Network	Point shapefile	Asset ID field which corresponds to the UP Asset ID and Down Asset ID fields of each pipe	Exported from InfoWorks CS Model
Manhole locations layer # 2 Interceptor Model Network	Point shapefile	Asset ID field which corresponds to the UP Asset ID and Down Asset ID fields of each pipe	Exported from InfoWorks CS Model
Boundary of study area layer	Polygon shapefile	N/A	Derived from the combined system area boundary provided by Grace Lin in consultation with Dr. Celia Fan.
Building layer # 1	Polygon shapefile	N/A	Grace Lin, Senior Modelling Engineer, Toronto Water, City of Toronto (City of Toronto, 2007)
Building layer #2	Polygon shapefile	N/A	GCCView: An internal server that hosts many of the City's in-house geodatabases (Geospatial Competency Centre, 2013)

Toronto Centreline (TCL) layer	Line shapefile	FCODE_DESC field (Describes the kind of feature represented i.e. minor arterial road or river	Open Data Toronto Website (Geospatial Competency Centre, 2014)
<p><b>Spatial Reference Shared by all layers to be consistent with the GIS data standards of the City of Toronto:</b>  Projected Coordinate System: MTM 3Degree  Projection: Transverse Mercator  Geographic Coordinate System: GCS North American 1927  Datum: North American 1927</p>			

**4.2 Creating the Boundary Polygon Shapefile**

To ensure that all subcatchments within the area of concern, (the Interceptor and Coxwell networks) would be accurately represented, a new boundary file had to be created. Instead of using the boundary layer of the combined area provided for the DWF subcatchment tool, the boundary was determined through consultation with Dr. Celia Fan, a civil engineer working on the project. Through this consultation it was determined that it would be best to place the Northern boundary at Highway 401, the Western boundary to the Humber River, the Eastern boundary a little beyond the extent of the DWF subcatchments for the Coxwell network, and the Southern boundary at the City's shoreline (Figure 7). This larger boundary was taken only as a precaution, since WWF subcatchments are typically larger than their DWF subcatchment counterparts (C. Fan, personal communication, June 3, 2013). In addition, it appears that a consultant group that previously created automated WWF subcatchments for the City of Toronto implemented a similar practice. However, their report did not disclose their approach (Clarifica, 2010). The boundary layer was created by making a copy of the original combined system service area layer provided by City staff and modifying until it looked like what is seen in Figure 19. The remaining layers, referred to in Table 4, were then clipped using this newly created boundary file that covered an area of approximately 256 Km<sup>2</sup>. This also reduced the overall file sizes, which improved the processing time.



**Figure 19. WWF subcatchment delineation method broader study area boundary**

### **4.3 Selecting the Appropriate Pipes Based on Catch Basin Leads (Preprocessing and Data Quality Assurance)**

To simplify the naming of new geospatial files created during the process, a default geodatabase was created where all these preparation files would be stored automatically. The WWF subcatchments were defined by the delivery of stormwater runoff to the City’s catch basins. Therefore the input shapefile could only include those pipes which were either combined or storm sewers, were in active use as identified in the GIS shapefile from TWAG and also had catch basins attached to them. The catch basin lead layer was crucial because only pipes, that have catch basins connected via leads or pipes, can capture overland flow into them. This occurs during a precipitation event in a system that is fully “disconnected”, in the sense that all downspouts from roofs are no longer tied to the sewer system because this has become a mandatory policy that is being phased in. City staff assume that 50% of disconnection of downspouts necessitates an adjustment to the upstream contributing nodes to be made by the modellers. Since these are applied after this automated process has identified the subcatchments, it is beyond the scope of this study (G. Lin, personal communication, September 19, 2013). Therefore, in this automation of subcatchment delineation, a 100% downspout disconnection rate was assumed.

However, since this research really focused on creating subcatchments for the Coxwell and Interceptor sewer networks, which are two separate models, the modellers checked only pipes part of these two networks within InfoWorks CS. It is important to note that the completeness of these two pipe networks was first examined. It was determined that

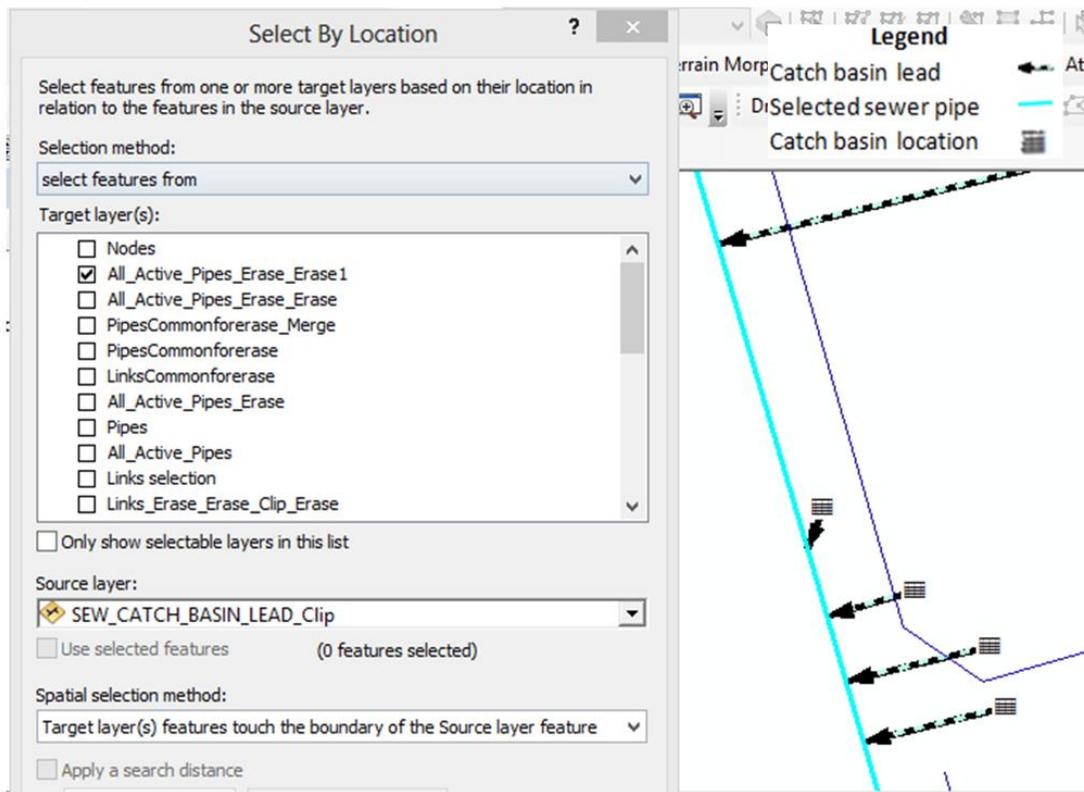
approximately 10% of the pipes were either completely missing or present but missing key fields required for modelling purposes such as the diameter of the pipe. This was established using as-built drawings and consultation with City staff. The modellers then only focused on ensuring the quality of these selected pipes (10%) by examining their attribute accuracy and positional accuracy by ensuring that the existing fields were correctly populated and that the pipes were precisely located using as-built drawings (Haklay, 2010). The pipes checked (pipes belonging to the Coxwell and Interceptor sewer networks) did not cover all the pipes within the larger study area (Figure 19) and so unchecked pipes, meaning pipes, which were not reviewed were also used, these pipes were taken directly from TWAG. To prevent duplication, it was necessary to remove the pipes belonging to the Coxwell and Interceptor sewer networks from the pipe network that was taken from TWAG. To do so, corresponding pipes from both model networks were removed from the from TWAG dataset through a combination of matching the asset id of pipes using the Join Field function and physically erasing pipes with the Erase function (Figure 20).



**Figure 20. Where the areas occupied by both the Coxwell and Interceptor networks that were checked by the modellers in InfoWorks were removed**

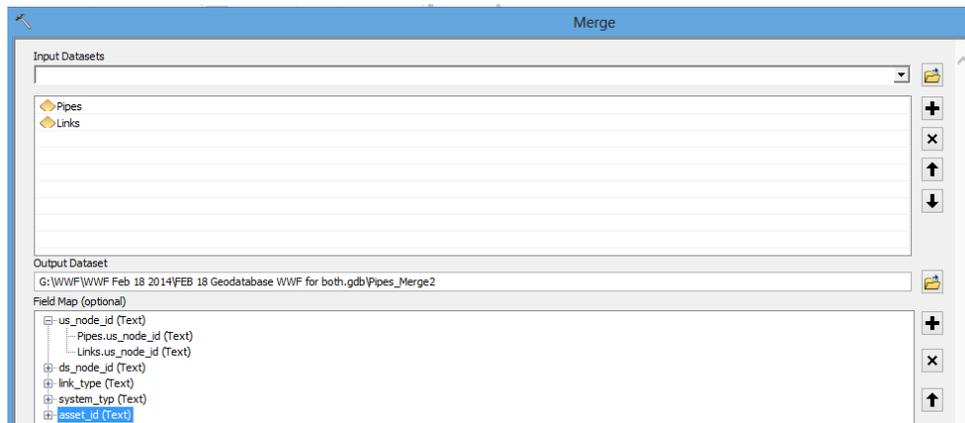
After removing the unmodified pipes from the TWAG pipe network to make room for those modified by the modellers within InfoWorks CS, it was then possible to select those pipes with a catch basin attached to them. As each catch basin node or lead was not assigned to the pipe it was expected to be associated with, as indicated by its attributes, selections were onerous.

In this case, a simple Select by Location-Intersection function would not suffice, since it selected more pipes than was expected. The touch boundary option, which is a permutation of the simple intersection option in ArcGIS, was utilized (Figure 21). It was critical that the quality of the data being inputted into these tools was complete and accurate. By using this selection method, approximately 12,606 pipes of the 32,580 pipes, including those beyond the two actual networks that were of concern (Coxwell and Interceptor sewer networks), were selected from the various layers of pipe networks.



**Figure 21. The "Target layer(s) features that touch the boundary of the Source layer" reduces the number of erroneous selections**

The separate Interceptor and Coxwell model network pipes then had to be merged. While a seemingly straightforward process, many errors occurred when attempting to merge them together without removing any fields from their attribute tables. It was decided to limit the fields that would remain in the output merge layer to those that were considered essential (Figure 22). Additionally, by removing extraneous fields it helped to reduce the file size, which improved processing time when working with such large files.

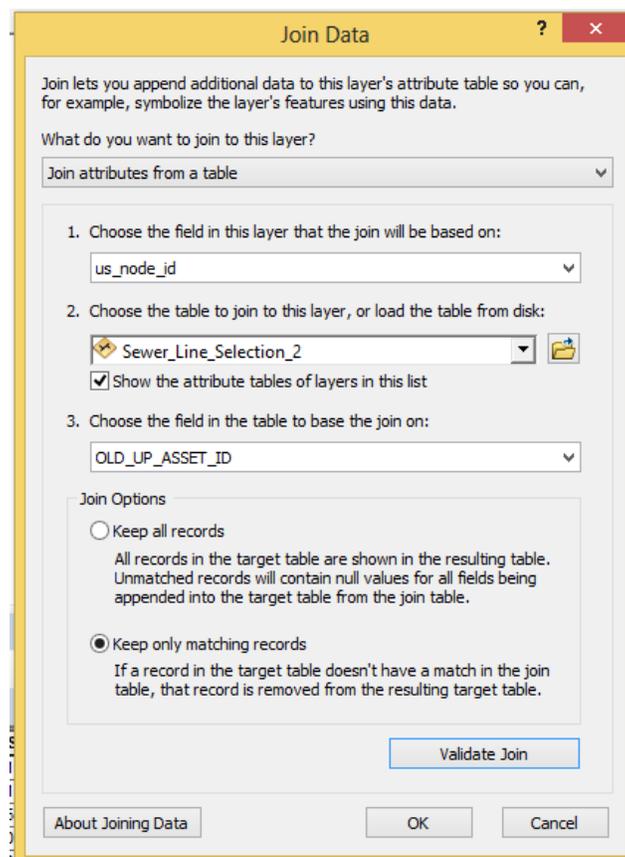


**Figure 22. Merging of the Coxwell and Interceptor pipe networks**

As some of the pipes exported from the InfoWorks CS model (Active storm and combined pipes layers # 1 & #2 from Table 4) had been moved from their original spatial locations, while being edited by the modellers, they no longer connected to their catch basin leads. Therefore, such pipes would not be selected, even though they should have been included in the pipes to be fed into the tool. In an attempt to remedy this issue, all the pipes where catch basin leads were snapped to them in the TWAG network, were selected and exported into a new layer as seen in Figure 23. Afterwards, an ArcGIS Table Join function based on the upstream node field was created. Since this parameter needed to be unique for every pipe in the model, it was the appropriate choice for identifying the correct pipes. The join was with the layer that had resulted from merging together the pipe networks from both the Interceptor and Coxwell models as seen in Figure 24. That in turn, helped to reduce the pipe network to pipes that actually had catch basins attached to them for this region, while still using those pipes that had been exported from the InfoWorks CS model. Those pipes that were selected from the InfoWorks CS models based on their up asset id through this join were then exported into a new layer.



**Figure 23. Layer of all the pipes from the unmodified network (from TWAG) selected because they have catch basins attached to them.**



**Figure 24. Join Data function used to select the appropriate pipes even if they were moved slightly and therefore no longer snapped to a catch basin lead**

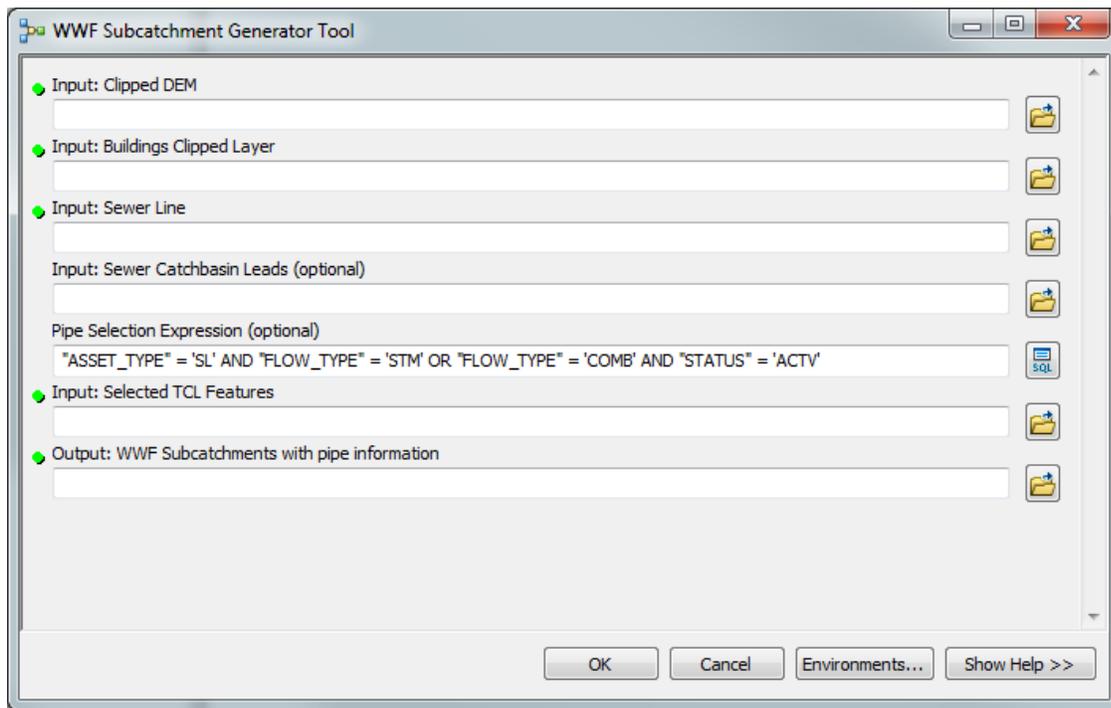
It was then necessary to merge the pipes edited by the modellers in InfoWorks CS to the TWAG pipes maintained by City staff at Toronto Water to create one layer containing all the pipes that would be used as input into the tools developed to create WWF subcatchments. In order for the pipe network to be recognized in Arc Hydro, it was necessary to add "Hydro IDs" to each pipe segment using the "Assign HydroID" function in Arc Hydro.

One final preparatory step had to be completed. It was to address the issues that there was no layer dedicated to identifying culverts or structures that allowed water to flow under manmade structures such as underpasses. The need for considering the locations of culverts and roads were requirements stipulated by City staff at Toronto Water (G. Lin, personal communication, February 1, 2013). The solution involved querying the Toronto Centreline (TCL) layer, which is a line layer that includes major linear features in the City, in order to isolate features that were needed to address these issues (Geospatial Competency Centre, 2014). The categories of features isolated were selected on the basis that they represented features via which water would likely travel such as major arterial and minor arterial roads and rivers or possibly identify the location of culverts (Figure 25).



**Figure 25. Highlighted blue line from the TCL layer signifies the location of a culvert that was "burnt in"**

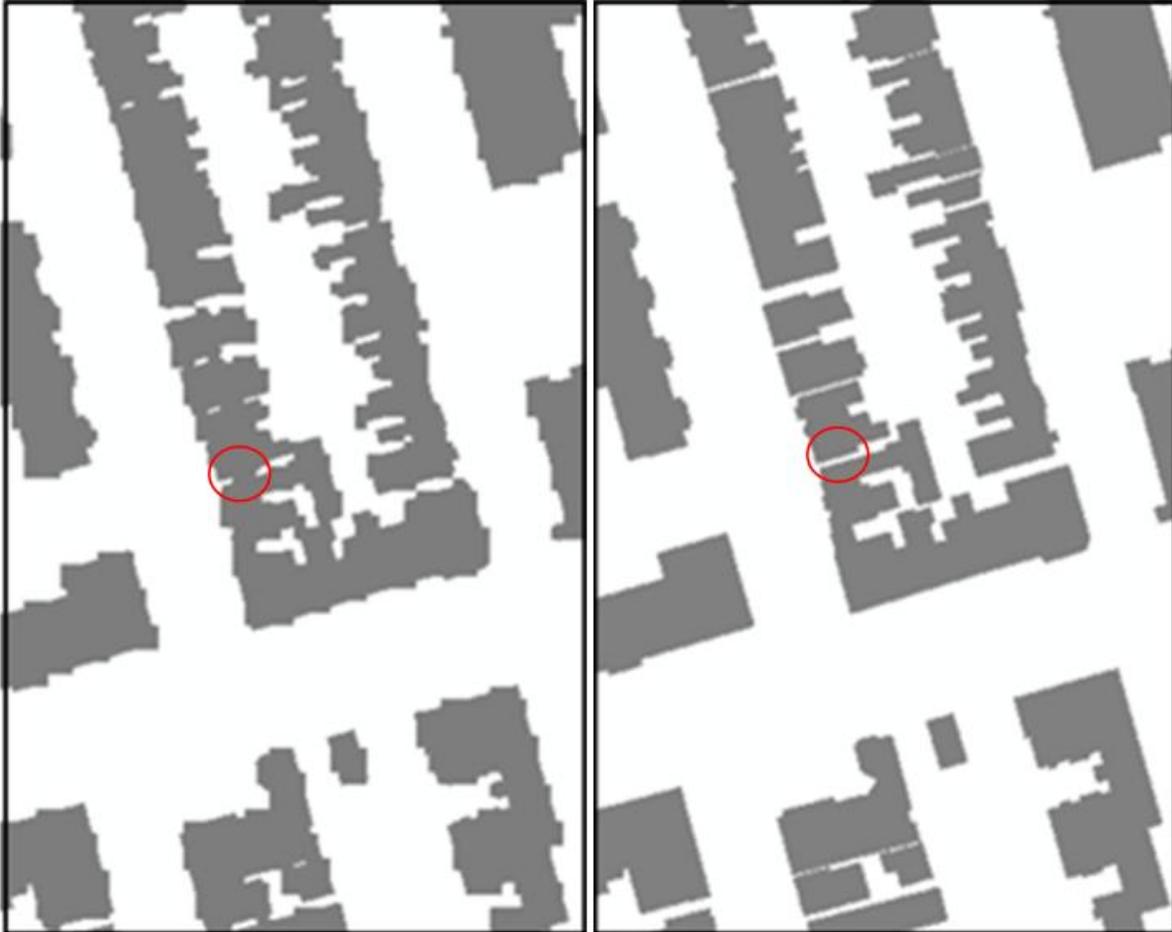
## 4.4 The WWF Subcatchment Generator Tool



**Figure 26. User interface for WWF Subcatchment Generator Tool**

To create this tool, all the steps involved were first conducted manually before being built into a user interface through ModelBuilder (Figure 26 and Figure A5). This tool was designed to be comprehensive, producing WWF subcatchments from the DEM, associating pipe network information with them and generating input for another tool, which generates overland flow paths.

The first step taken by the program made it possible for the tool to correctly "burn in" features later on. Based on initial testing it was found that by not resampling the DEM, it caused issues such as creating artificial barriers between buildings that prevented the flow of water (Figure 27). Such issues reinforce why it is essential to use features of comparable resolution when attempting to "burn in" or to raise cells representing any features in the real world, man-made or otherwise (Garbrecht & Martz, 2000). Fewtrell *et al.* (2011) suggested a horizontal resolution 0.5-metre for addressing such issues while using DEMs in urban drainage modelling studies. ArcGIS provides an opportunity to adjust the resolution of a DEM by the Resample function so the first step taken using the program is to change the spatial resolution to this standard.

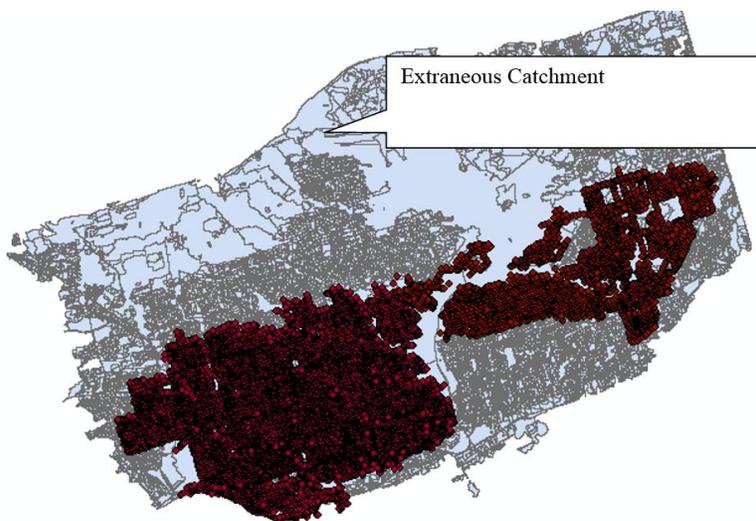


**Figure 27. Comparison of different raster resolutions from 3-m to 0.5.**

Once the DEM was resampled, all the sinks and depressions were identified as a precaution. According to ESRI (2011a) "Depressions are defined by subtracting the input DEM from its filled DEM" whereas sinks are defined by a cell surrounded by higher elevation cells, which traps the water in that cell so it cannot flow onwards. It could be considered a precautionary step since in the next step all these sinks were eliminated with the use of the fill sink function within Arc Hydro. The program then began the process of manipulating the DEM in an attempt to accurately replicate flow in the urban landscape. To do so, it burns in the pipes, roads and other relevant features from the TCL layer through the Arc Hydro DEM Reconditioning function. This function normally lowers the elevation of the DEM at points that can be inferred to coincide with natural streams and rivers. However, in DEMs of urban landscapes, pipes and roads have been "burnt in" place of streams, and the cells are typically lowered by only 0.1 to 5 metres from their original elevation to cause the water to flow along

these flow paths (Baker *et al.* 2006; Gironás *et al.* 2010). Simultaneously, the tool also prevents water from flowing through buildings by raising the location of buildings using a layer that was created by merging the two building layers indicated in Table 4. It does so by using the buildings layer as an extraction mask and then uses the ArcGIS Raster Calculator to add 10 metres to all coinciding cells on this DEM.

Combining the two layers together to create a layer that had roads, rivers and pipes "burnt in" and the location of buildings raised by additional 10 metres, any new sinks were filled in and a flow direction grid was generated using the D8 method (Figure A5) (O'Callaghan & Mark, 1984). To ensure that only one WWF Subcatchment was created for each pipe, the Flow Direction with Streams function in Arc Hydro was applied, utilizing the sewer line layer that only had combined and storm pipes as input. As can be seen from Figure A5, the catchments were first generated as a raster, which were then converted into polygons. After this, the pipe information was added to each catchment through a Spatial Join, which helped remove the extraneous catchments that were not part of the Coxwell or Interceptor networks as seen in Figure 28. The Clip function would not suffice to remove the extraneous catchments because the WWF subcatchment boundary for the Coxwell and Interceptor networks was not known prior to running the WWF Subcatchment Generator Tool and hence delineated by the tool itself. Once this was done, the WWF subcatchments were considered completed (Figure 29, Figure 30 and Figure 31).



**Figure 28. Extraneous catchments that were eliminated after pipe information was added**

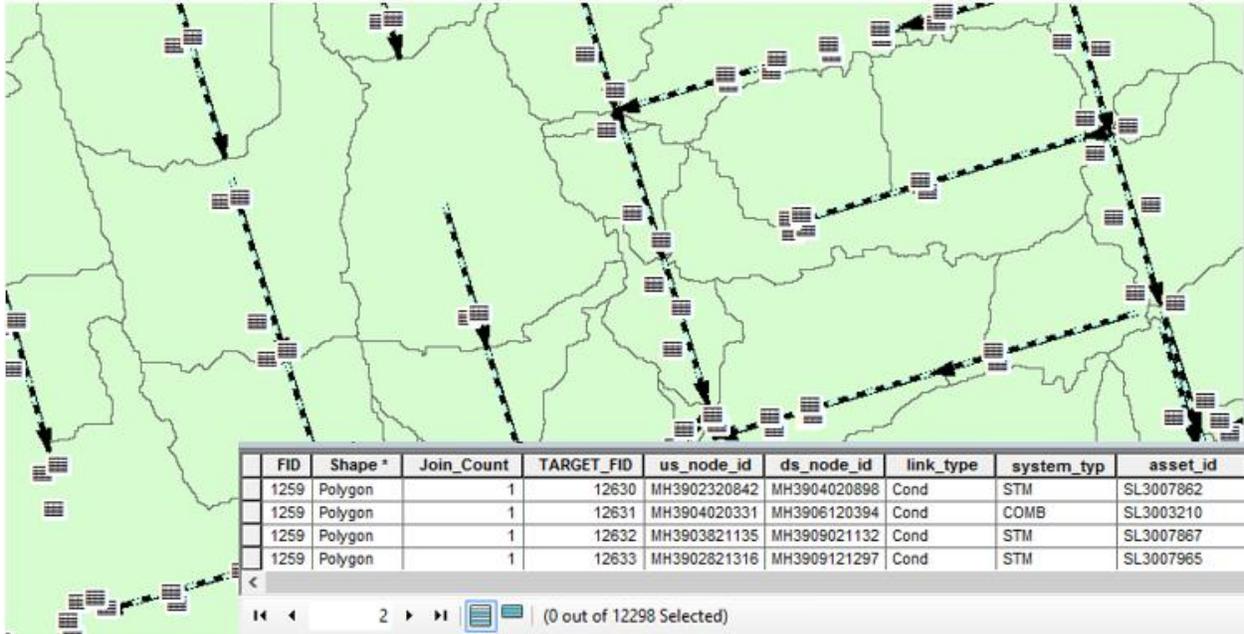


Figure 29. WWF subcatchments automatically delineated by the tool

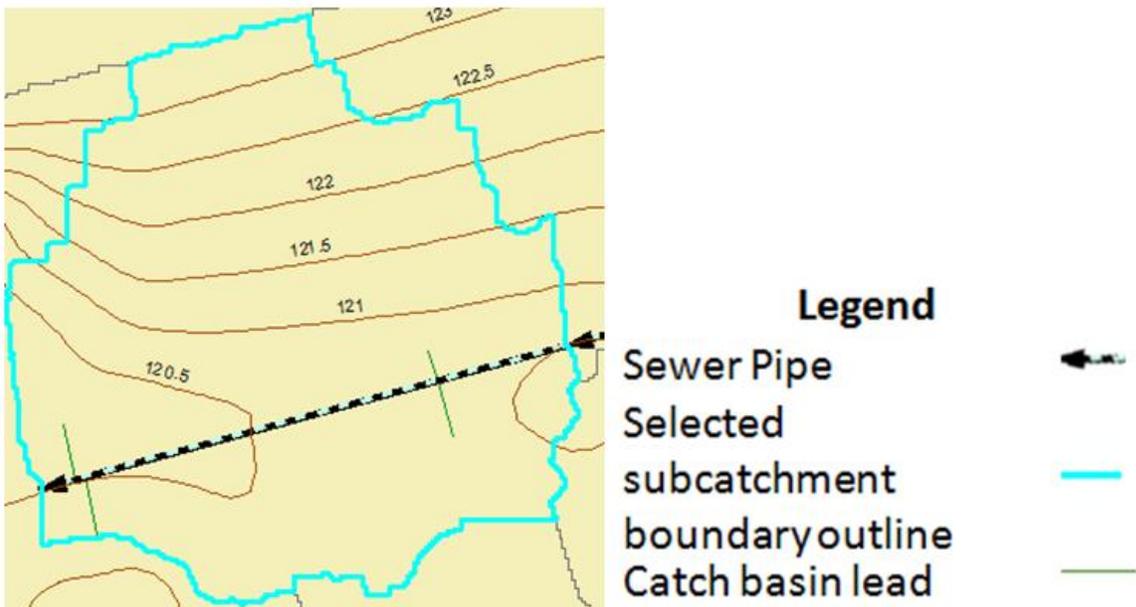
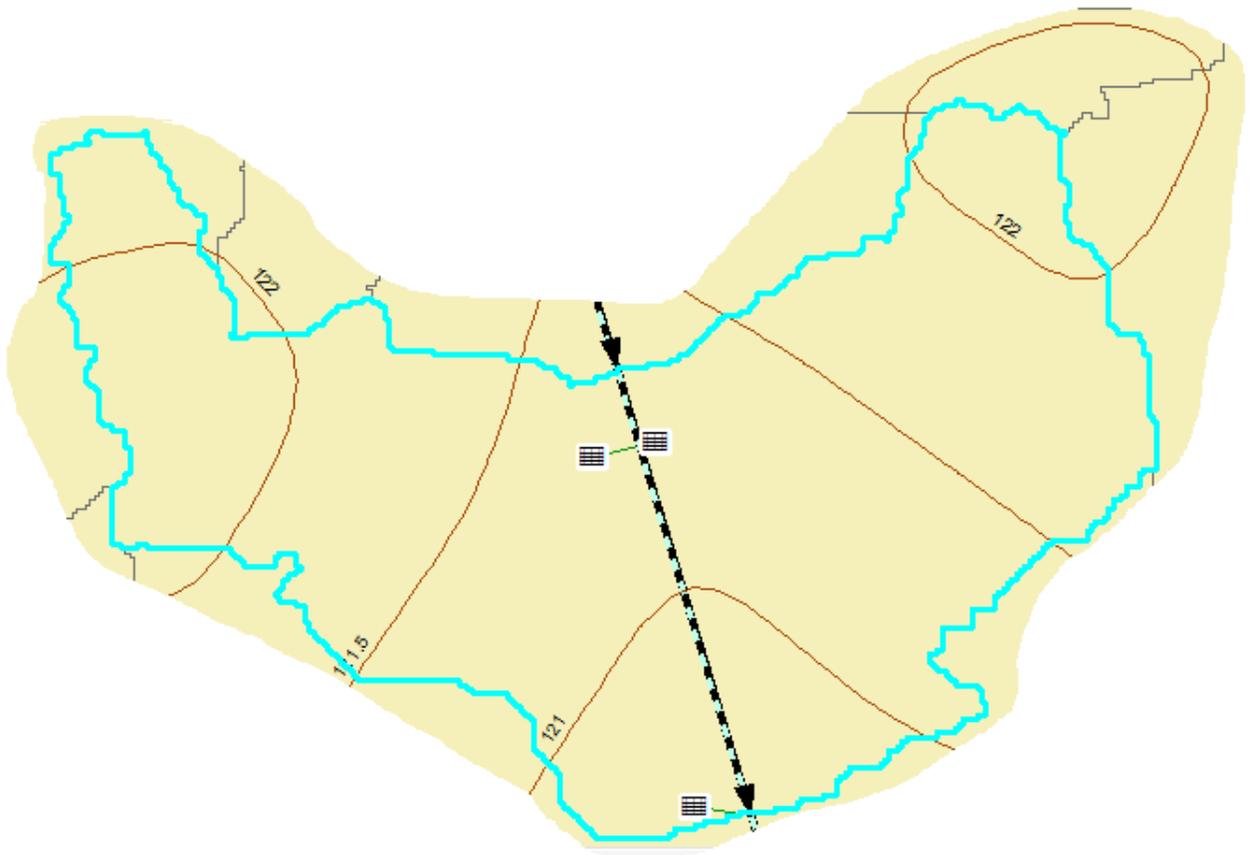


Figure 30. As can be seen, it respects the gradients of the contour lines quite well

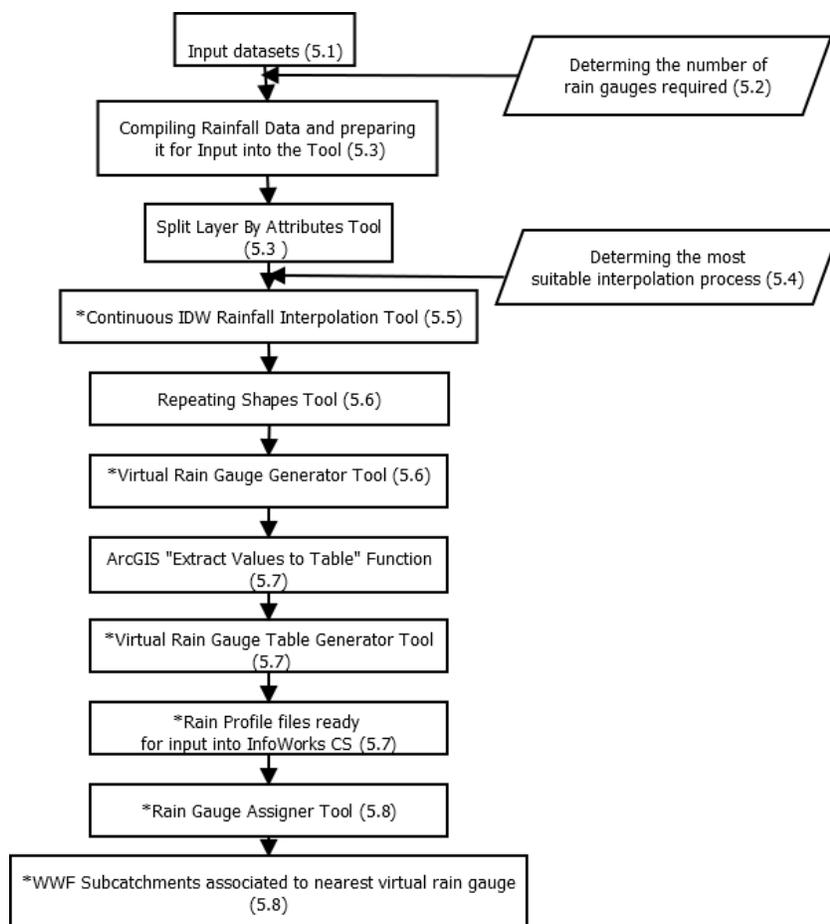


**Figure 31. High points flowing downslope towards the pipe's location**

## 5.0 Automated Approach to Rainfall Interpolation

Virtual rain gauges, as the name implies, are simulated rain gauges situated at points where no real physical rain gauge was present, but instead are populated with values based on the interpolated values of actual rain gauges in the area (Shamsi, 2005; Ahrens, 2006; Haberlandt, 2007). The contribution of this research has been to develop not only a method to continuously interpolate data to generate virtual rain gauges at a seemingly unprecedented rate, but to provide a complete array of tools to facilitate its use as input for urban drainage models such as InfoWorks.

The major components of this approach are provided in the flowchart diagram seen in Figure 32. As can be seen from this diagram, while this approach can still be considered an automated approach, it relies upon a series of tools, both existing and specially designed for this



**Figure 32. Virtual rain gauge method schema**

research. Thus, while some of the tools developed are by no means novel on their own such as the Rain Gauge Assigner Tool covered in section 5.8, the way they have been integrated, as part of a whole package is what makes this research innovative. In addition, certain steps, such as those covered under section 5.6 in the flowchart, may only need to be done once if the same region is going to be interpolated over time. It should also be noted that in this approach, there are also multiple input datasets created for the InfoWorks CS Model.

The first are actual virtual rain gauge locations, the second is the tabular rainfall records associated to each rain gauge and the third are WWF subcatchments pre-associated to their nearest rain gauges which saves the user time.

## 5.1 Datasets Required

The files listed in Table 5 include those that were necessary for generating the virtual rain gauges such as the rainfall measurements recorded on a 5-minute interval in a tabular format. It also includes those that were required for ensuring that the gauges generated could be utilized by the model such as the WWF subcatchments. It can be noted from this table that a key field for the shapefile denoting the location of the City's rain gauge was a field labelled as "COLE\_ID field". This was because Cole Engineering maintains the gauges. Thus, the identification system used by the City to identify their gauges is the same one used by Cole.

**Table 5. Layers required for virtual rain gauge generator tools**

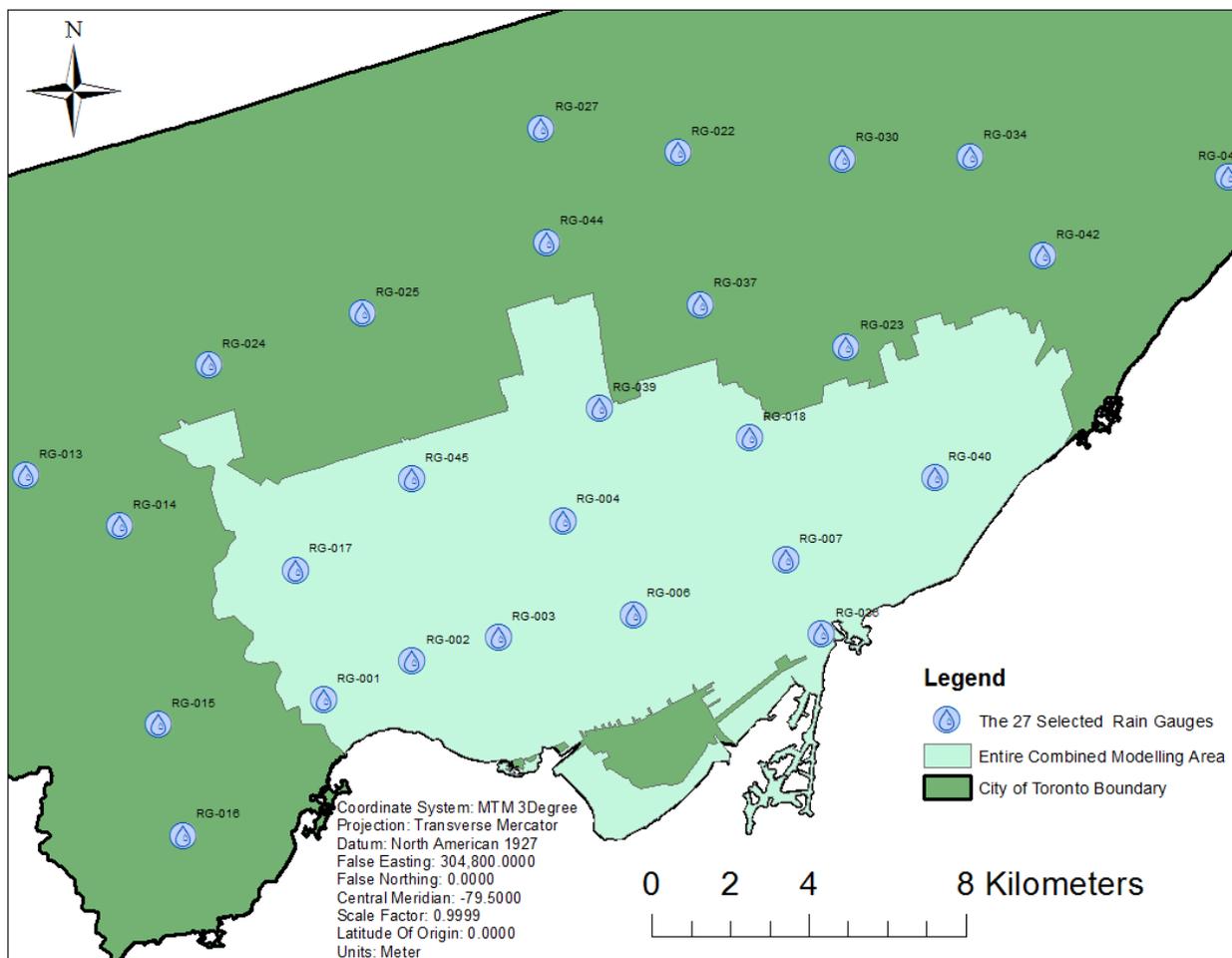
<b>Geospatial File Name or Description</b>	<b>File Type</b>	<b>Key Attribute Fields</b>	<b>Source</b>
WWF subcatchments	Polygon File	Rain Gauge Number field	Delineated by the tool presented in the previous chapter
Location of rain gauges	Point shapefile	COLE_ID field Added X coordinate field Added Y coordinate field	Grace Lin, Senior Modelling Engineer, City of Toronto (City of Toronto, 2013a)
Boundary of study area layer (Entire combined sewer service area)	Polygon shapefile	N/A	Grace Lin, Senior Modelling Engineer, City of Toronto (Lin, 2013)
Rainfall measurements	Tabular file such as CSV	Rain Gauge Number field Rain Gauge (mm) field Time field Added "I" field to signify intensity of rain Added "T" or "TS" field to signify a unique timestamp	Retrieved from Data Current, a secure online portal hosted by Cole (Cole, 2014b)
<b>Spatial reference shared by all geospatial layers to be consistent with the GIS data standards of the City of Toronto:</b> Projected Coordinate System: MTM 3Degree Projection: Transverse Mercator Geographic Coordinate System: GCS North American 1927 Datum: North American 1927			

## 5.2 Determining the Number of Rain Gauges Needed for Interpolation

The study area was the entire combined (CSO) area as defined by the shapefile provided by the City of Toronto and shown in blue in Figure 7. Rainfall was to be interpolated across the entire CSO portion of the City, based on the City's shapefile of municipal rain gauges and recorded monitoring rainfall records for 2013. In total, the City of Toronto has 36 tipping bucket rain gauges. Of these, only gauges in and immediately adjacent to the study area were selected for the interpolation process (Figure 33). However, enough gauges were still required in order to enable interpolated values to cover the study area, which was determined through sensitivity testing. Initially, only those rain gauges in the immediate study area were used. However, this did not generate an interpolation surface or raster that would provide sufficient coverage. Additional gauges were added incrementally, until it was determined that 27 of the 36 rain gauges were necessary to create an interpolation surface that would provide the necessary coverage (Figure 34).



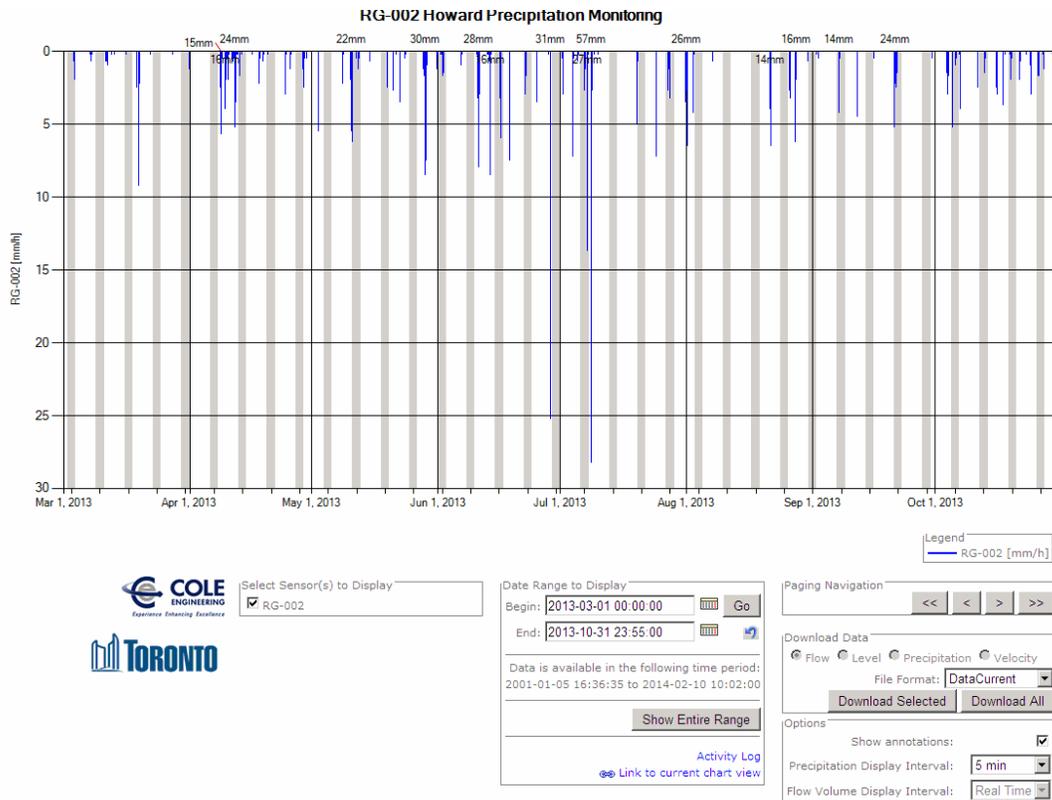
**Figure 33. RG-020 which is a tipping bucket style rain gauge known as the TR-525USW rain gauge (Cole, 2014a)**



**Figure 34. The 27 rain gauges selected**

### **5.3 Compiling Rainfall Data and Preparing it for Input into the Tool**

The City's rain gauges are maintained and operated by Cole Engineering, which is a consultant firm. Cole Engineering in turn provides municipalities access to their rain gauge rainfall records via a secure online portal known as Data Current from which, with the appropriate credentials, authorized personnel are able to download rain gauge measurements from 5-minute intervals to 1 hour intervals in comma-separated-values (CSV) files (Cole, 2014b). The 5-minute rainfall records were acquired for the period between 00:00 hours on April 1<sup>st</sup>, 2013 to 23:55 hours on October 31<sup>st</sup>, 2013 or a total 61,631 measurements for the desired stations (Figure 35). Compilation required linking each CSV file to the shapefile of station locations by adding the spatial coordinate location of each gauge to the CSV files.



**Figure 35. How rainfall records from each of the 27 rain gauges was selected and downloaded**

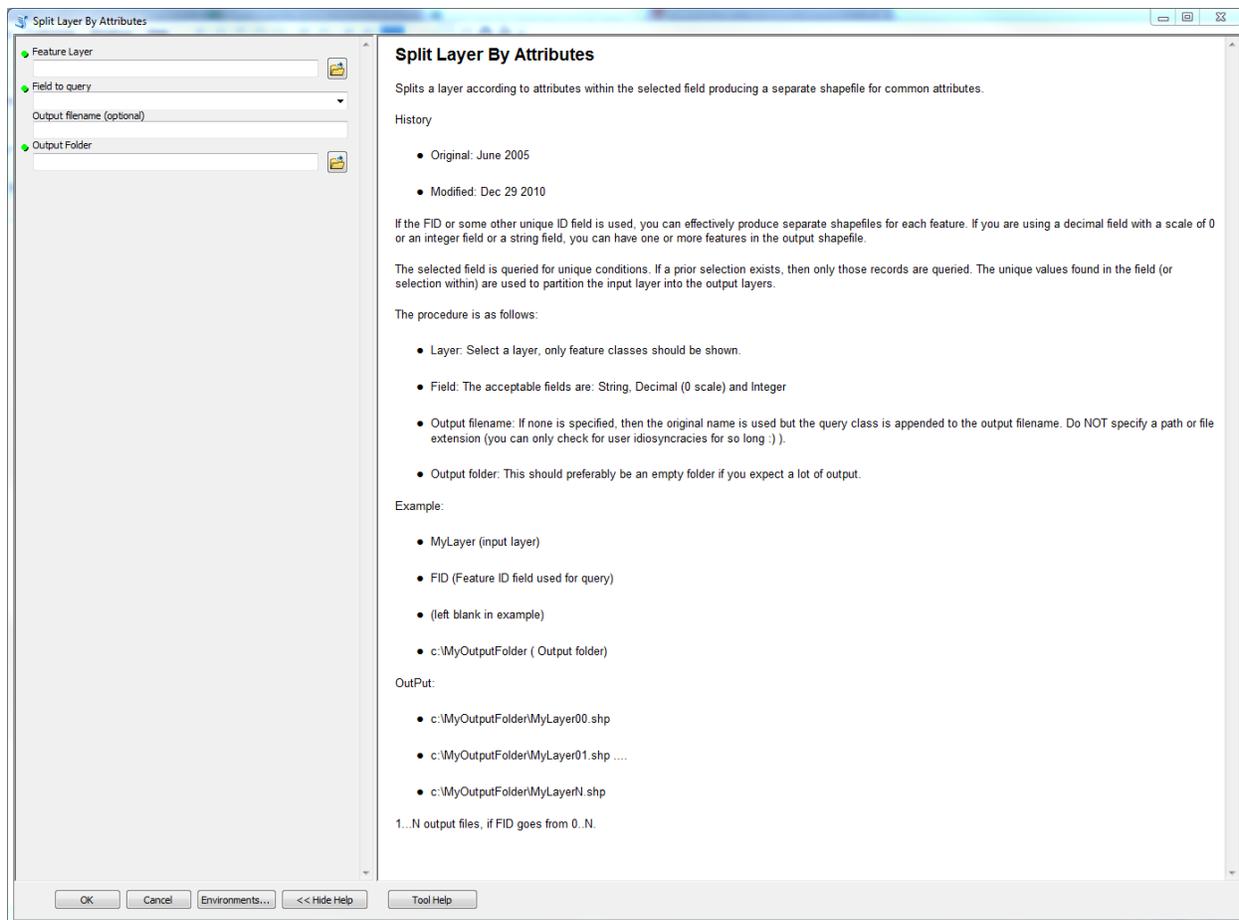
In the conversion process, the attributes were adjusted by adding a new field in which the appropriate sensor id number or COLE ID for that gauge was added (ID). As the time stamp information would be lost once the feature classes were exported into shapefiles, the addition of a field called “T” was added to retain this information and was populated with a value between 0 and 61,630, which corresponded to a unique time and date. In order for the InfoWorks CS model to run, it needed the precipitation expressed as an intensity value. As a result, to determine the intensity on an hourly basis for each five-minute interval, each value had to be multiplied by 12 (Innovyze, 2013). Therefore, the precipitation field was changed so that it was labelled as “P” and a new field called “I” was added to signify the intensity value that was added (Figure 36).

OBJECTID *	Shape *	Time	P	I	ID	T
1	Point	1/1/2013	0	0	RG-01	1
2	Point	1/1/2013 12:05:00 AM	0	0	RG-01	2
3	Point	1/1/2013 12:10:00 AM	0	0	RG-01	3
4	Point	1/1/2013 12:15:00 AM	0	0	RG-01	4
5	Point	1/1/2013 12:20:00 AM	0	0	RG-01	5
6	Point	1/1/2013 12:25:00 AM	0	0	RG-01	6
7	Point	1/1/2013 12:30:00 AM	0	0	RG-01	7
8	Point	1/1/2013 12:35:00 AM	0	0	RG-01	8
9	Point	1/1/2013 12:40:00 AM	0	0	RG-01	9
10	Point	1/1/2013 12:45:00 AM	0	0	RG-01	10
11	Point	1/1/2013 12:50:00 AM	0	0	RG-01	11
12	Point	1/1/2013 12:55:00 AM	0	0	RG-01	12
13	Point	1/1/2013 1:00:00 AM	0	0	RG-01	13
14	Point	1/1/2013 1:05:00 AM	0	0	RG-01	14

**Figure 36. Typical attribute table of a rain gauge's records once all the measurements were spatially linked to their appropriate rain gauge location in a feature class**

The 27 shapefiles were all merged into one large shapefile (27RGMerge). As there were many time steps when none of the gauges registered a value above zero, an attribute query, which only selected those rows when a value above zero was registered for the Intensity (I) field, was used. Isolated, using the ArcGIS Create Layer from Selected Features function, a summary table for the “T” field was created for these time intervals.

The "T" field from this new summary table was used in a join on the large rain-gauge shapefile (27RGMerge) to identify all those time steps when a measurement above zero was registered on at least one of the gauges for each time step. Those rows that were selected were then exported into a new shapefile that only contained information for those time steps when at least one gauge had a value above zero. To simplify the creation of the tool, the freely available Split Layer By Attributes Python-based tool was used (Patterson, 2013). The way in which this script functions and why it was integral to the development of this automated method is that it allows efficient separation of any shapefile or feature class into multiple files on the basis of sharing the same value for a certain attribute (Figure 37) (Patterson, 2013). In this case, the attribute was the field that signified the time step or interval. A separate file was created for each relevant time step so that it would be easier to automatically feed in the information for each interpolation into the tool that would generate all the interpolated layers.



**Figure 37. User interface for the Split Layer By Attributes Tool**

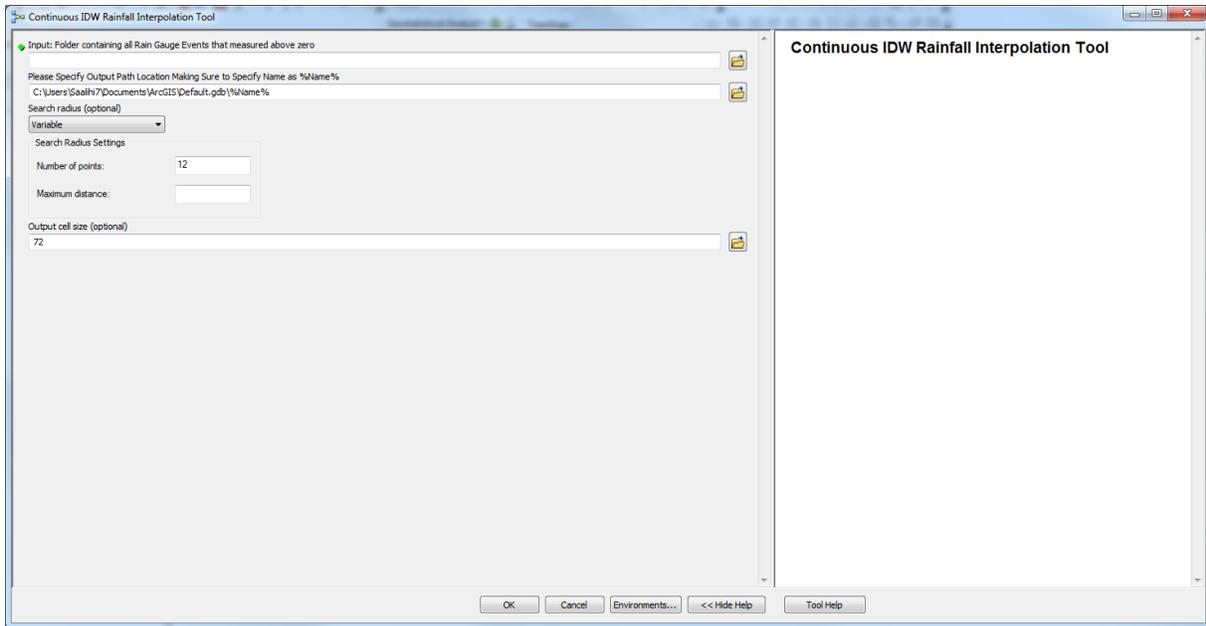
## **5.4 Determining the Most Suitable Interpolation Method**

Since there were thousands of shapefiles, each representing the 27 rain gauge observations of the specified time interval, a sensitivity analysis could be conducted to determine which interpolation method seemed most appropriate. The interpolation methods that were tested were IDW, Kriging and spline as can be seen in Figure A7 and Figure A8. While in the literature (Dirks *et al.* 1998; Ahrens, 2006; Piazza *et al.* 2011; Ly *et al.* 2013) the more computationally demanding spline and Kriging methods might be viewed as being “more accurate”, they did not enhance the results achieved by the much simpler IDW and were therefore concluded to be not appropriate. Even when there were many rain gauges registering values above zero as seen in Figure A7, both spline and Kriging generated irrational values on their output raster. Such results were not acceptable because it is impossible to measure a negative value for rainfall. Thus, such results immediately raised concerns regarding the accuracy of such interpolation methods in this case and which is why they were not pursued. It

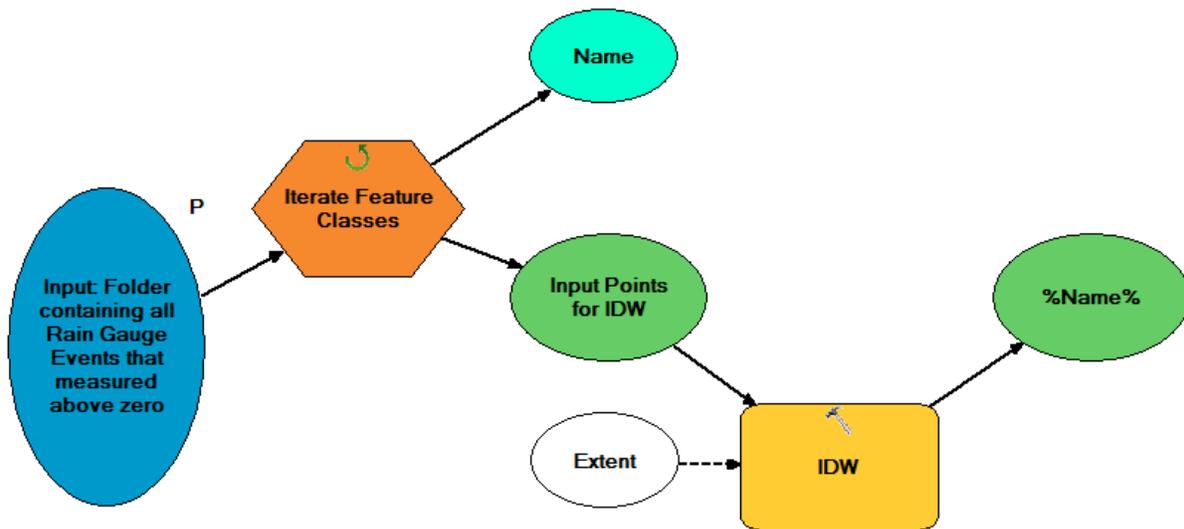
appears when this has happened in similar studies, this issue has been addressed by converting all the negative values into zeroes (J. Liu, personal communication, August 17, 2013).

## **5.5 Designing and Running the Continuous IDW Rainfall Interpolation Tool**

ModelBuilder was used in designing the Continuous IDW Rainfall Interpolation Tool (Figure 38) because it provides great flexibility, allowing modification of the tool in the future without requiring users to know much about programming. These modifications could include adjusting the resolution of the interpolation results, which could be adjusted by changing the output cell size or even the type of interpolation that was employed. The challenge faced, while developing this tool, was figuring out a way to have the tool run interpolations on a continuous basis within ArcGIS as it appeared as though it may not have been done at the scale being undertaken in this research. To overcome this challenge, it was necessary to rely upon all of the individual files that had been generated earlier using the Split Layer By Attributes Python-based tool. Furthermore, the Iterate Feature Class option in ArcGIS ModelBuilder with the IDW interpolation function was used as the basis for the tool (Figure 39). This option allows the same function to be applied to every shapefile or feature class in a specified folder or geodatabase (Allen, 2011). In order for all the interpolations to be completed, it only took a few hours since all steps with zeroes were skipped. The output was thousands of interpolation surfaces in raster format with a cell size resolution of 72 metres.



**Figure 38. User interface for Continuous IDW Rainfall Interpolation Tool**

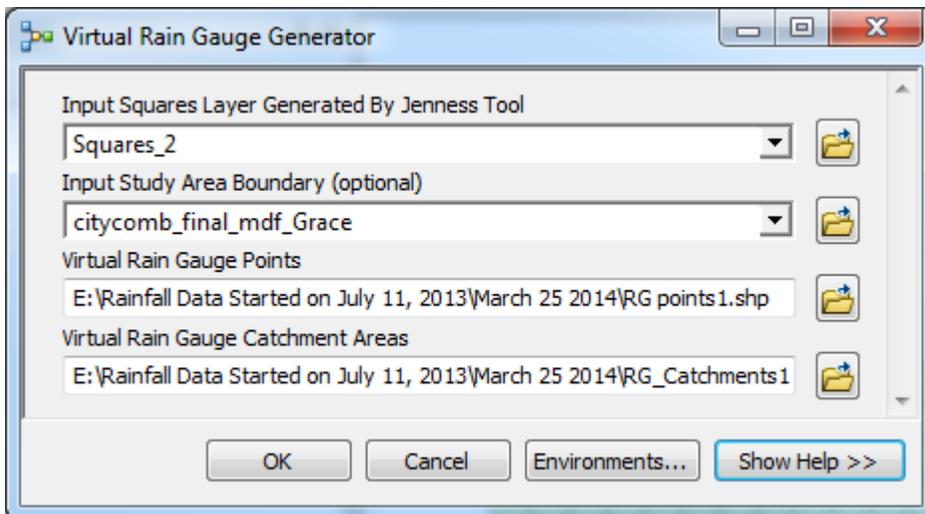


**Figure 39. A ModelBuilder diagram of the Continuous IDW Rainfall Interpolation Tool**

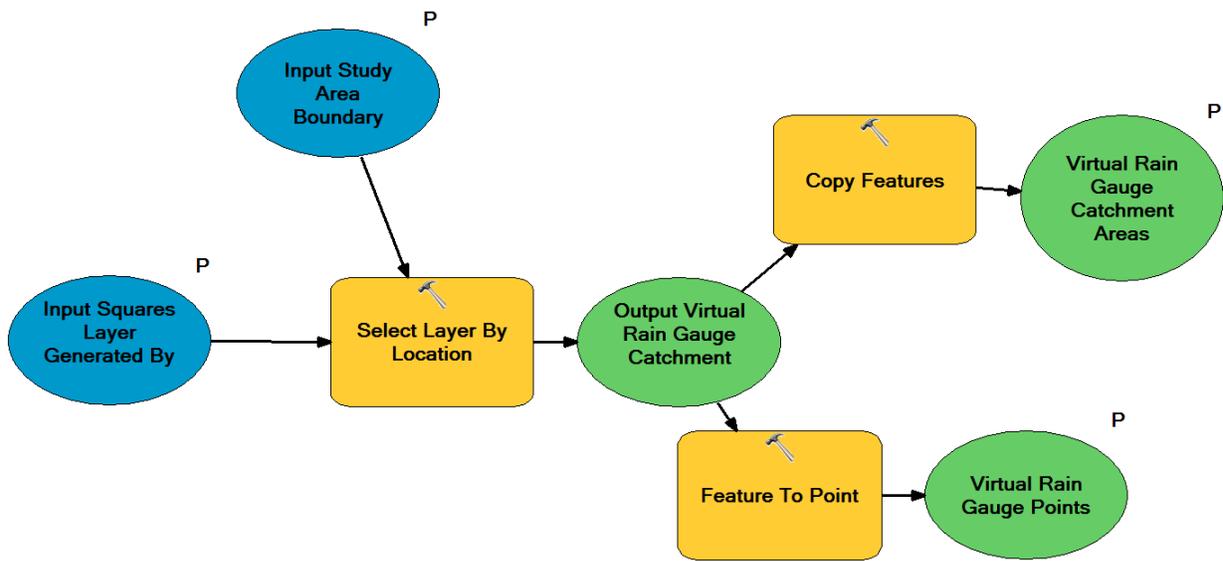
## 5.6 Creating Virtual Rain Gauges

Initially, a manually delineated bounding rectangle around the entire combined (CSO) area or study area shapefile was used as the basis for generating virtual rain gauges. It was then converted into a raster using cell dimensions of 500 metres by 500 metres since this was the size used by the consultant firm that created a model for Toronto, and had interpolated the rain gauge

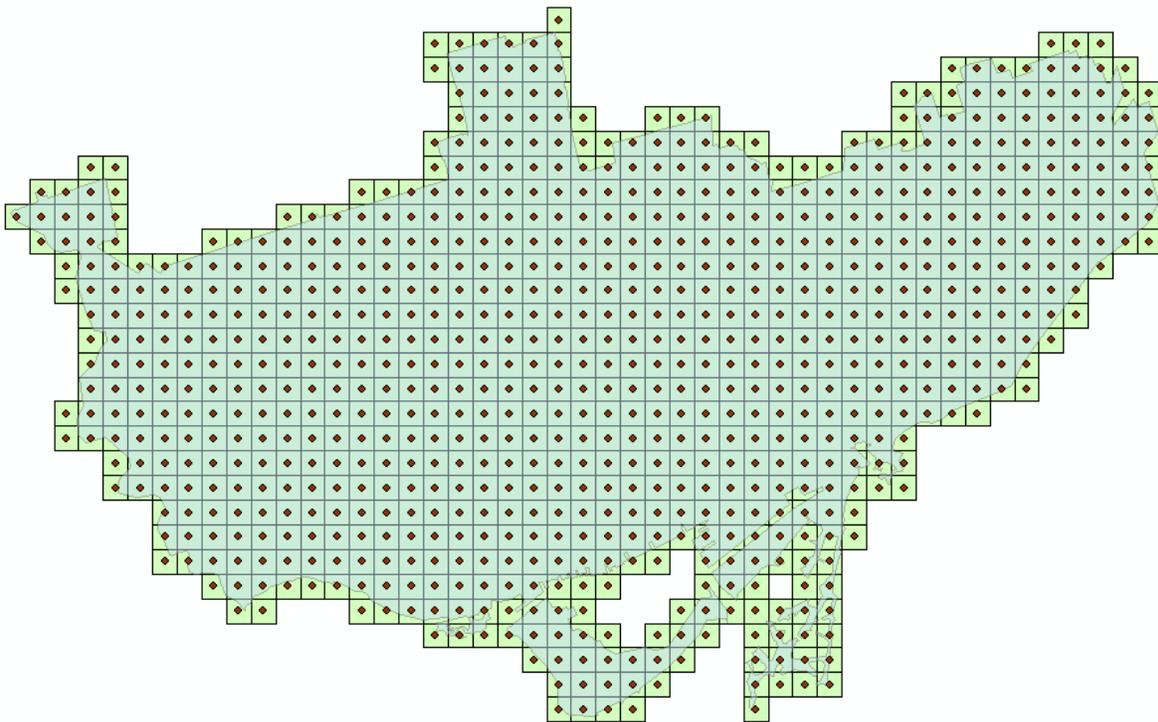
records so that instead of just having a few rain profiles they had over 100 (Clarifica, 2010). The manual creation of the bounding rectangle polygon shapefile was determined to be problematic because it made it difficult to generate consistent results. Though the number of virtual rain gauges generated when the bounding rectangle was first drawn manually was 793, there was no guarantee that this number would always be generated by following such an approach. In an attempt to generate repeatable results, another method was developed. This process was much simpler than the previous one because it relied heavily upon a free tool developed by Jenness (2012) entitled “Repeating Shapes for ArcGIS 10”. Once installed and following the steps (Figure A9) to generate the virtual rain gauge catchment areas (500 by 500 m squares in a polygon shapefile format), it was possible for the actual rain gauge points or locations to be derived. The Virtual Rain Gauge Generator Tool shown in Figure 40 and Figure 41 created the actual virtual rain gauge point locations (Figure 42). Figure 41 also indicates why it was possible to have an odd number of virtual gauges despite using a grid to generate them and that is because only those squares, which intersected the study area boundary, were selected as being virtual rain gauges.



**Figure 40. User Interface for the Virtual Rain Gauge Generator**



**Figure 41. ModelBuilder diagram of the Virtual Rain Gauge Generator**



**Figure 42. Virtual rain gauge point locations generated by tools**

## **5.7 Extracting Virtual Rain Gauge Location Data from Interpolated Layers**

To create a complete time series for each rain gauge that was then ready for input into the InfoWorks CS model (Figure 43), it was necessary to create individual files for each gauge. The virtual rain gauge shapefile was used as the input point layer in the ArcGIS Extract Values to Table function. This function mined all the values from the thousands of precipitation raster layers to generate one large table. It was then possible to associate each interpolated value to its correct time step through joining it to the name of its raster (RasterName) (See A.5). The resulting table was linked to the rain gauges based on the "SrcID\_Feat" field or the field indicating the virtual rain gauge number in order to add their appropriate X and Y coordinates which were then used to create a single shapefile layer. When all the values had been extracted and converted into one large shapefile, the Split Layer By Attributes Tool was used to separate the large shapefile generated into 793 individual shapefiles, one for each virtual rain gauge. All 793 shapefiles were then run through the Virtual Rain Gauge Table Generator Tool designed in ModelBuilder that added back all the time steps when the precipitation and intensity value was zero across all the rain gauges to create a complete time series using the Iterate Feature Classes function and the Join Field function (Figure A6).

	C	D
47002	6/13/2013 4:40:00 AM	0
47003	6/13/2013 4:45:00 AM	0
47004	6/13/2013 4:50:00 AM	0
47005	6/13/2013 4:55:00 AM	0.068019
47006	6/13/2013 5:00:00 AM	0.173343
47007	6/13/2013 5:05:00 AM	0
47008	6/13/2013 5:10:00 AM	0.434264
47009	6/13/2013 5:15:00 AM	0.131548
47010	6/13/2013 5:20:00 AM	2.437718
47011	6/13/2013 5:25:00 AM	2.012591
47012	6/13/2013 5:30:00 AM	2.147233
47013	6/13/2013 5:35:00 AM	2.419241
47014	6/13/2013 5:40:00 AM	0.115871
47015	6/13/2013 5:45:00 AM	0.063529
47016	6/13/2013 5:50:00 AM	2.076037
47017	6/13/2013 5:55:00 AM	0.380509
47018	6/13/2013 6:00:00 AM	0.242846
47019	6/13/2013 6:05:00 AM	2.168808
47020	6/13/2013 6:10:00 AM	0.145813
47021	6/13/2013 6:15:00 AM	2.199435
47022	6/13/2013 6:20:00 AM	0.272088
47023	6/13/2013 6:25:00 AM	0.335534

**Figure 43. CSV input file generated for RG000 with the time being represented in the first column and intensity in mm per hour in the second.**

### **5.8 Associating Virtual Rain Gauges to WWF Subcatchments**

To successfully simulate periods of wet weather in the InfoWorks CS model, each wet weather flow subcatchment needed to be associated to its nearest rain gauge location via a field in its attribute table that clearly identifies this gauge. To associate each rain gauge manually within InfoWorks CS would be untenable because there were over 12,000 individual subcatchments and 793 virtual rain gauges. To save time, a tool that automatically assigns each respective subcatchment to its nearest virtual rain gauge, based on a Spatial Join using the nearest option, was created which is analogous to Thiessen polygon allocation (Figure 44, Figure 45, and Figure 46).

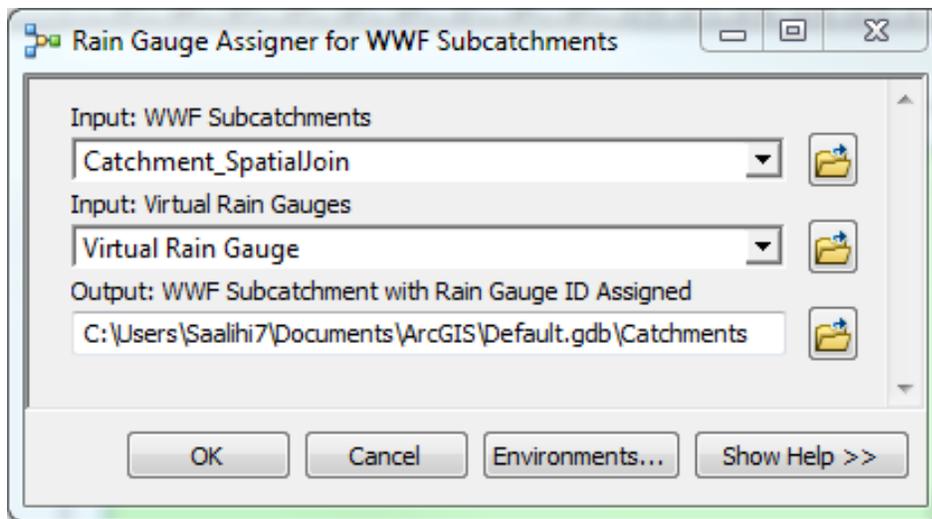


Figure 44. User interface for Rain Gauge Assigner Tool

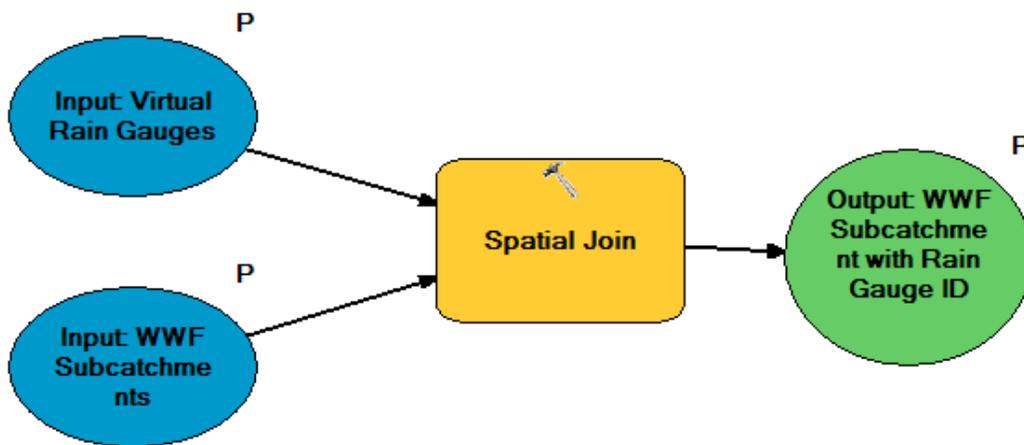


Figure 45. ModelBuilder diagram of Rain Gauge Assigner Tool

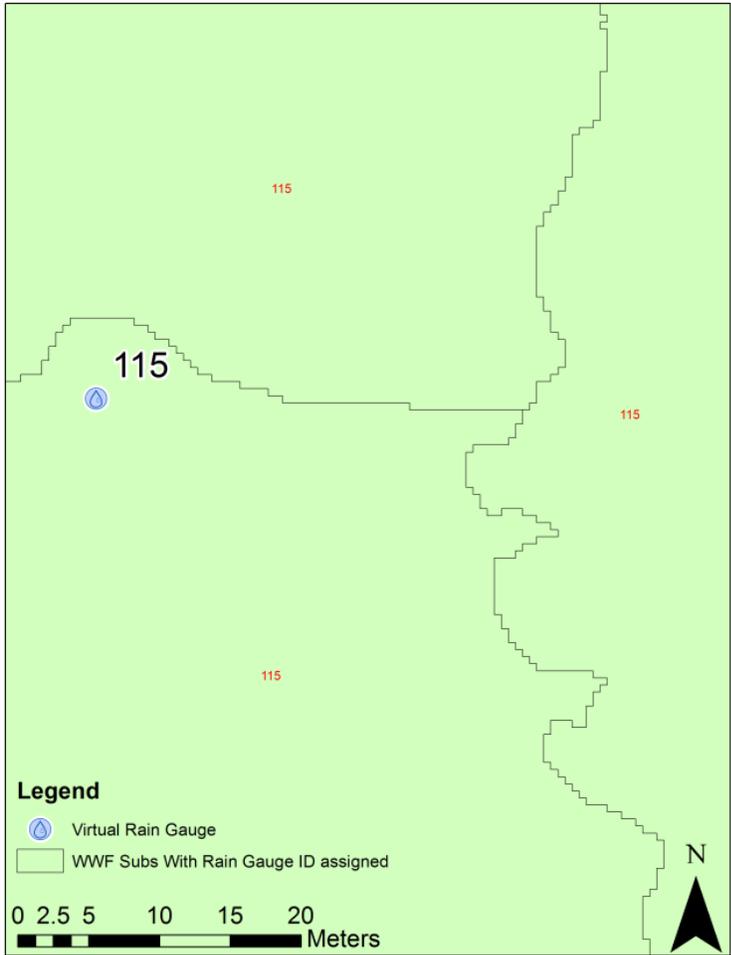


Figure 46. Results generated by rain gauge assigner tool

## 6.0 Evaluating Automated DWF Subcatchment Delineation

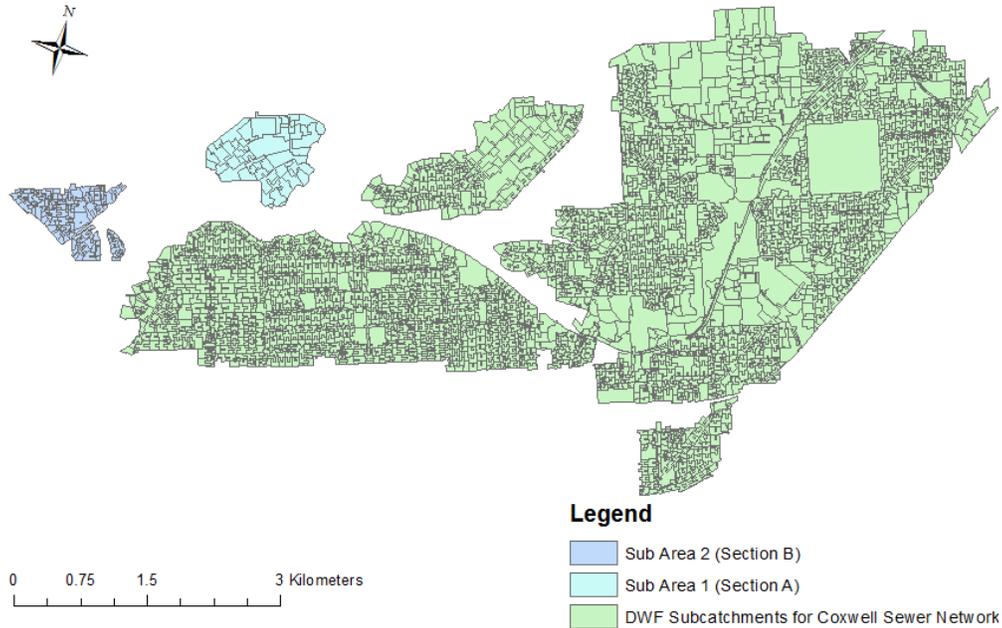
To demonstrate the effectiveness and limitations of the automated approach presented in this document, this section highlights the tools' strengths and weaknesses. In order to determine whether a subcatchment could be categorized as either "Acceptable" or "Manual Intervention Required", the criteria listed in Table 6 were used. These criteria were developed based on the principles used when manually delineating DWF subcatchments by City staff and or consultants. If any of these criteria is not met, then it is categorized as "Manual Intervention Required".

**Table 6. DWF subcatchment criteria**

"Acceptable"
<ul style="list-style-type: none"><li>✓ No gaps between this subcatchment and the surrounding subcatchments.</li><li>✓ No overlap between this subcatchment and the surrounding subcatchments.</li><li>✓ The portion of the road containing the upstream node or manhole of the pipe associated to the subcatchment is part of the subcatchment.</li><li>✓ All connected parcels are downstream of the upstream node of the pipe in the subcatchment.</li><li>✓ The subcatchment has a unique upstream node and asset id combination and also has the field denoting its flow type populated.</li><li>✓ The subcatchment has at least one parcel contained in it.</li><li>✓ The subcatchment has the entirety of each parcel contained within it so that the boundaries of it correspond to the boundaries of the parcels.</li><li>✓ Every parcel downstream of its closest pipe based on the side of its frontage, which is determined by orthoimagery, is part of that pipe's subcatchment.</li><li>✓ The subcatchment is a single part feature because the InfoWorks CS model cannot handle multipart features.</li></ul>

### 6.1 Evaluating Automatically Delineated DWF Subcatchments for Coxwell

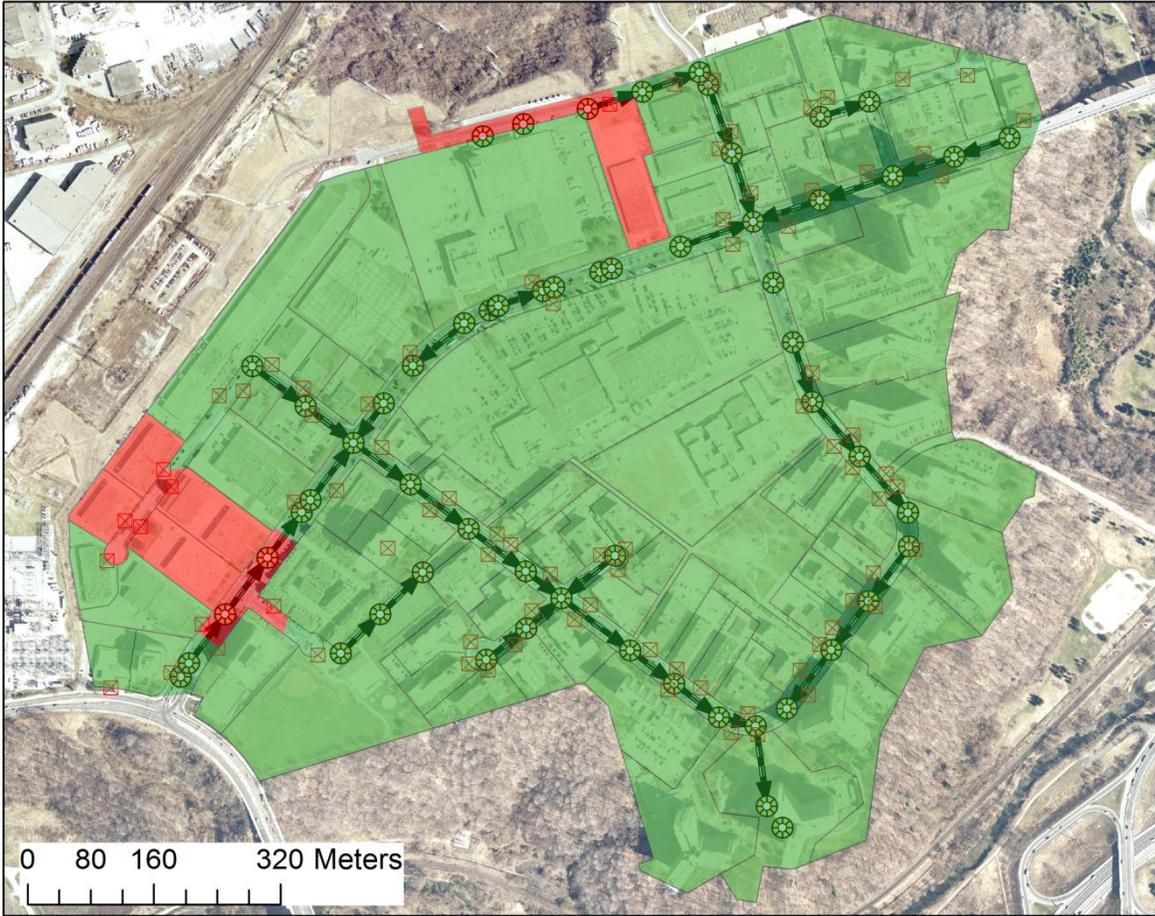
The total number of DWF Subcatchments created with the final iteration of the tool was 2,763 because it was only applied to the Coxwell model pipe network. It was first attempted to evaluate the entirety of sub areas, and or non-random spatially stratified areas in the Coxwell region referred to as Sections A and B (Figure 47). The purpose was to check if location affected whether subcatchments being inspected within these two regions would be categorized as "Acceptable" or "Manual Intervention Required" based on the criteria listed in Table 6.



**Figure 47. Location of Section A and Section B**

The results from sampling Sections A and B of the automatically delineated DWF subcatchments for the Coxwell sewer network, while promising due to the relatively high level of "Acceptable" subcatchments identified, raised questions of whether they were truly representative of the entire dataset (Figure 48 and Figure 49). These areas were isolated pockets, so they could not be relied upon to reliably gauge the overall accuracy of the dataset. Consequently, it was decided to randomly extract subcatchments and check whether they were to be categorized as "Acceptable" or "Manual Intervention Required". To determine which of the two categories each particular subcatchment fell under, the same criteria from Table 6 were applied when visually inspecting each selected subcatchment.

Furthermore, all DWF subcatchments were subjected to topology checks within ArcGIS using the Error Inspection Tool (available only with an ArcInfo Licence) to ensure that errors such as overlaps or gaps among features within the dataset were not present. If there were gaps or overlapping features, it indicated that topology rules were not properly enforced in the datasets. For comparison purposes, during the topology inspection, the automated DWF subcatchments were tested along with previously manually delineated subcatchments from another distributed urban drainage model created by City staff within the combined service area.



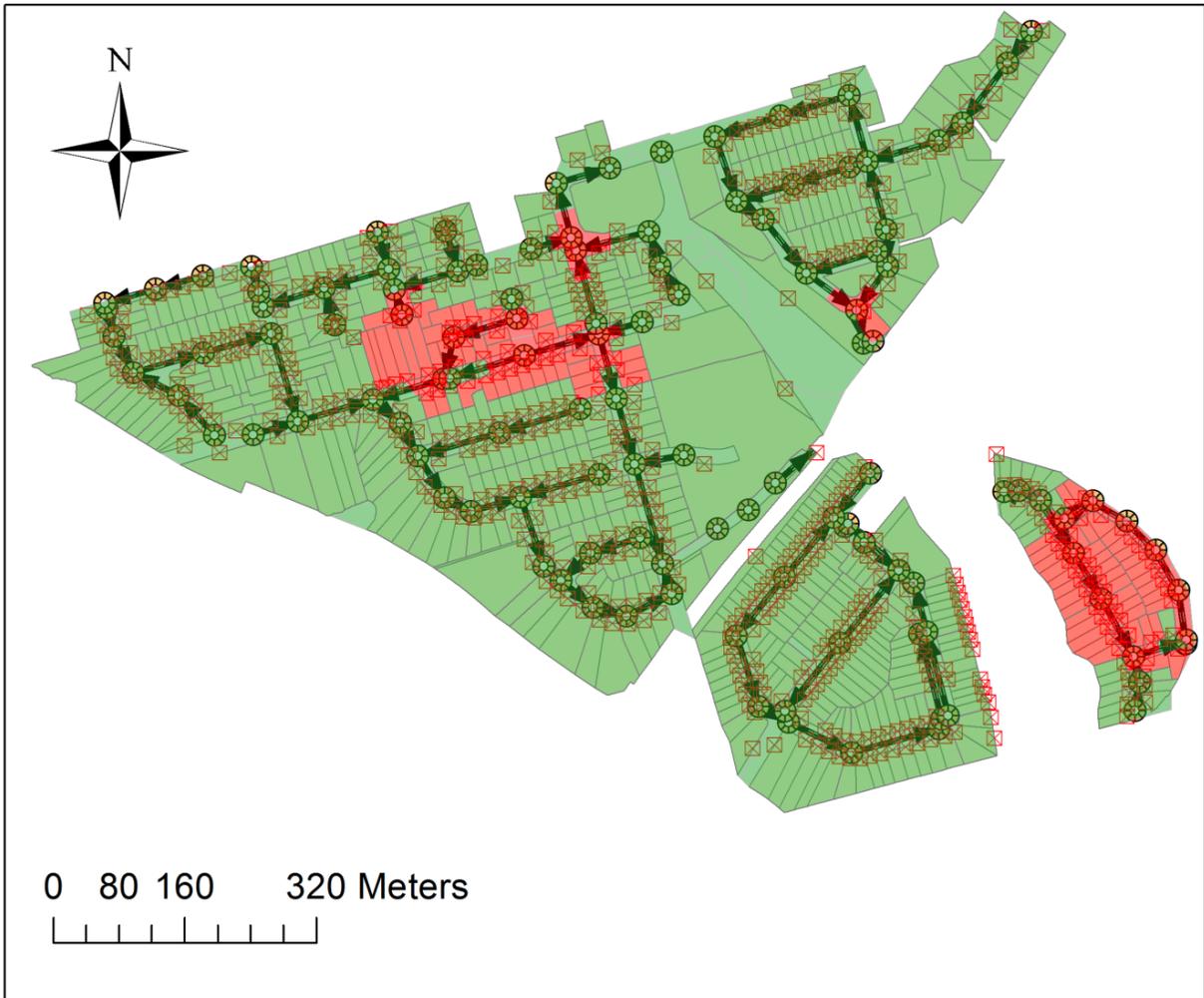
**Legend**

**Result of Evaluation**

- "Manual Intervention Required" 7.1%
- "Acceptable" 92.9%
- Frontage Nodes
- Upstream Nodes
- Sanitary and Combined Pipes
- Parcels for DWF



**Figure 48. Evaluation results for Section A**



**Legend**

**Evaluation Result**

- "Manual Intervention Required" 16.7%
- "Acceptable" 83.3%
- Frontage Nodes
- Upstream Nodes
- Sanitary and Combined Pipes
- Parcels for DWF

**Figure 49. Evaluation results for Section B**

**Determine Sample Size**

Confidence Level: 95%

Population Size: 2763

Proportion: 0.5

Confidence Interval: 0.10

    Upper: 0.60000

    Lower: 0.40000

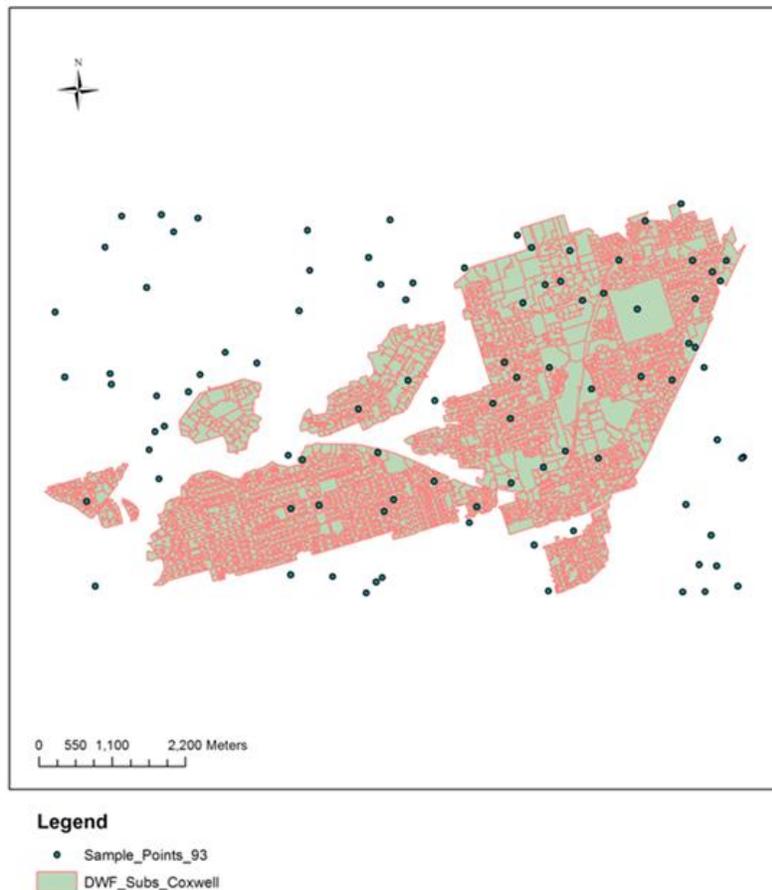
Standard Error: 0.05102

Relative Standard Error: 10.20

Sample Size: 93

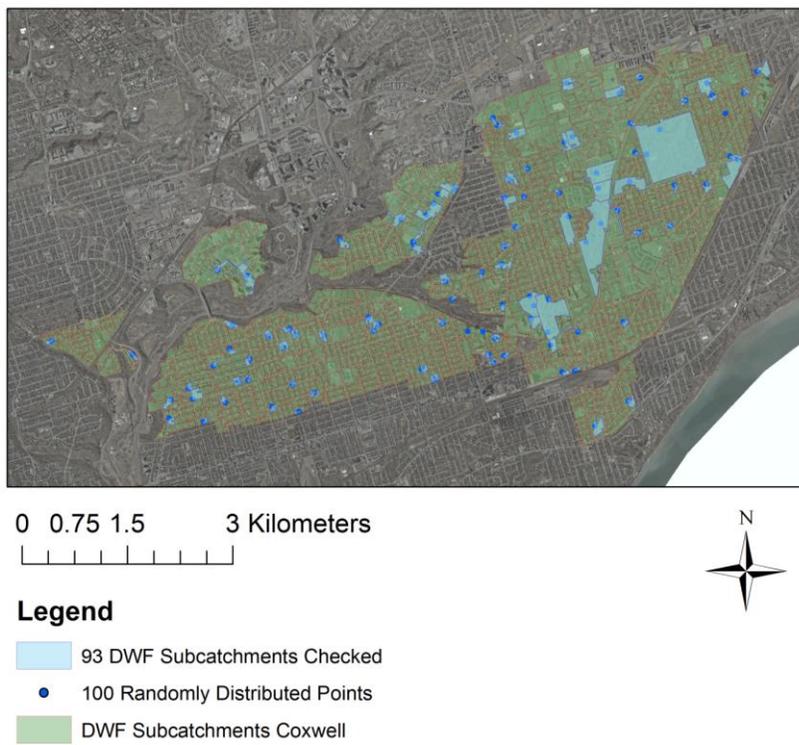
To determine the appropriate sample size for the visual inspection, an online sample calculator was used to perform the appropriate calculations (Figure 50) (Australian Bureau of Statistics, 2013). To attain a 95% confidence level, it was determined that a sample size of 93 subcatchments was required. In an attempt to make the selection of the points as unbiased as possible, the Create Random Points tool in ArcGIS was used using the extent of the DWF subcatchments layer as a constraint. However, this produced an undesirable result, as many of the points were not intersecting any of the DWF subcatchments (Figure 51).

**Figure 50. Results of sample size calculator**



**Figure 51. Randomly distributed points that were not suitable**

Thus, a different approach had to be developed that would only sample within the subcatchment area. While there were a number of ways to achieve this, the approach chosen involved the creation of a new field. With a new field added to the DWF subcatchment's attribute table and used in conjunction with the Dissolve function, one large multipart shapefile was created to act as a confining layer for the point-selection process. In attempting to intersect the randomized “select set” of points with the DWF subcatchments, only 88 were selected because some subcatchments contained multiple points. The number of random points generated was therefore increased to 100 and the number of subcatchments selected increased to 96, which was above the 93 subcatchments necessary to meet the threshold. To reduce this number to 93 subcatchments, while introducing another level of randomization, all 96 of the subcatchments selected by the random points generated were exported into Excel and the random number generator tool was used to create a field comprising random number. After these were generated and ordered, the top 93 rows (93 largest values) were then copied into a new Excel file saved as a Comma-Separated Values (CSV), which is supported by ArcGIS. This file was used in ArcGIS to select the 93 subcatchments randomly selected using the ArcGIS Join Table function based on the "us\_node\_id" field (Figure 52).



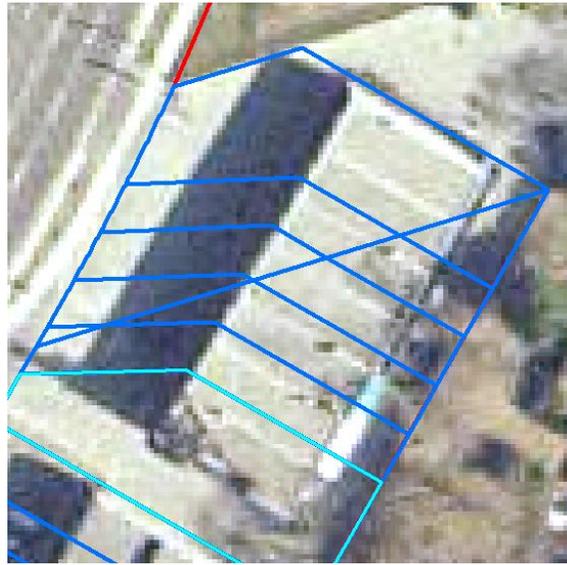
**Figure 52. The 93 randomly selected subcatchments that were used in the visual check**

The results of the visual check are presented in Table 7 with the full results found in Appendix A.6. From this evaluation of randomly selected subcatchments, it was possible to more accurately gauge the true overall accuracy and the kinds of errors of the results in comparison to the initial checks (Section A and Section B). Section A (Figure 48) which only sampled 42 subcatchments, all in the same area, produced more favourable results with an overall accuracy of 93% based on the criteria from Table 6. Therefore, this indicates that errors generated by the automated approach, such as those shown later on, are more numerous in certain areas than others, depending on the spatial complexity, not only of the pipe network (Error # 4 in next section), but also on the complexity of the parcels layer themselves. In some cases, the parcels layer itself would make it rather difficult for even a modeller delineating the parcels manually because some of the shapes present within the dataset are puzzling and appear to be inaccuracies (Figure 53). An interesting observation from the visual inspection was that overall it appears as though subcatchments that require manual intervention tend to be larger than those deemed acceptable. A possible reason for this phenomenon is that these subcatchments have extra parcels connected to them as will be covered in the next section (Error # 1).

Although the percentages in the table indicate the accuracy of the automated method is about 66%, it could also be as high as 76% or as low as 56%, since the margin of error used when calculating the sampling size was plus or minus 10% (see calculation parameters used in Figure 50). Either way, it is a significant improvement from the first iteration of the tool, where the percentage of acceptable DWF subcatchments was approximately 3.3%. While these results are based on the criteria listed in Table 6, they could be further validated through the use of field observations.

**Table 7. Results of the visual inspection of selected DWF subcatchments**

Type of Subcatchment	Number of Subcatchments	Total Area of Subcatchments (Ha)	Percentage
"Acceptable"	61	143.1	65.6%
"Manual Intervention Required"	32	109.3	34.4%
<b>Total</b>	93	252.4	100%



Legend	
Selected Parcel Boundary	<span style="color: cyan;">—</span>
Parcel Boundary	<span style="color: blue;">—</span>
DWF Subcatchment Boundary	<span style="color: red;">—</span>

**Figure 53. Very unusual parcels that are likely inaccuracies**

In terms of the topology check results, the automated subcatchments generated seem to fare much better with no errors being detected. On the other hand, while the manually delineated subcatchments appeared perfect at first glance, the topology check results reveal that there are currently numerous slivers and gaps being created in the manual process as seen in Table 8. Comparison of number of topology errors identified Table 8.

**Table 8. Comparison of number of topology errors identified**

Manual DWF Delineation			Automatic DWF Delineation	
Total # of Subcatchments: 1,726			Total # of Subcatchments: 2,763	
Topology Rule	Number of Errors	Area in Metres Squared	Topology Rule	Number of Errors
Must Not Have Gaps	990	3113528.92	Must Not Have Gaps	0
Must Not Overlap	740	5835.75	Must Not Overlap	0
Total Errors	1,730	3119364.67	Total Errors	0

## **6.2 An Overview of Errors Produced by the Automated DWF Subcatchment Delineation Process**

The four errors covered here represent the range of errors that were identified while reviewing the subcatchments generated against the criteria listed in Table 6. Where possible, an explanation of what caused the error and how to possibly remediate it are provided.

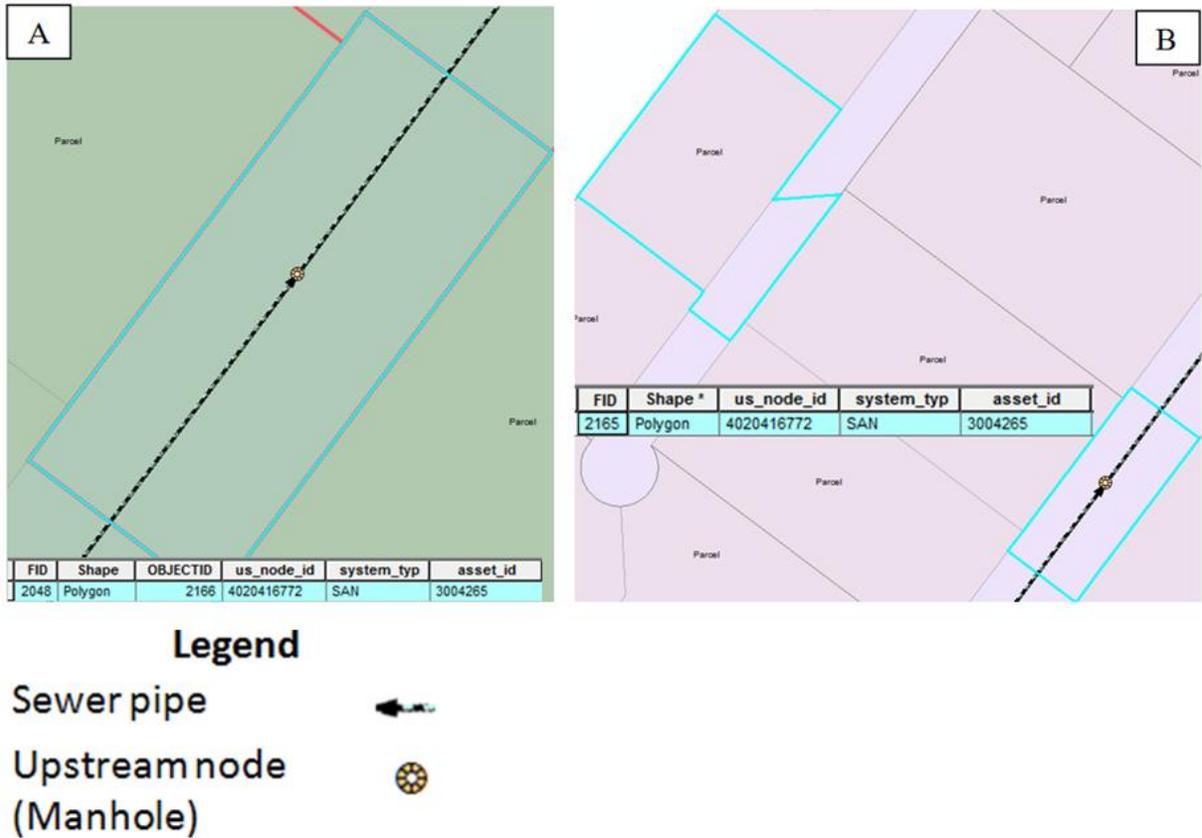


**Legend**

- Sewer pipe 
- Upstream node (Manhole) 
- Frontage Node 
- DWF Subcatchment Boundary 

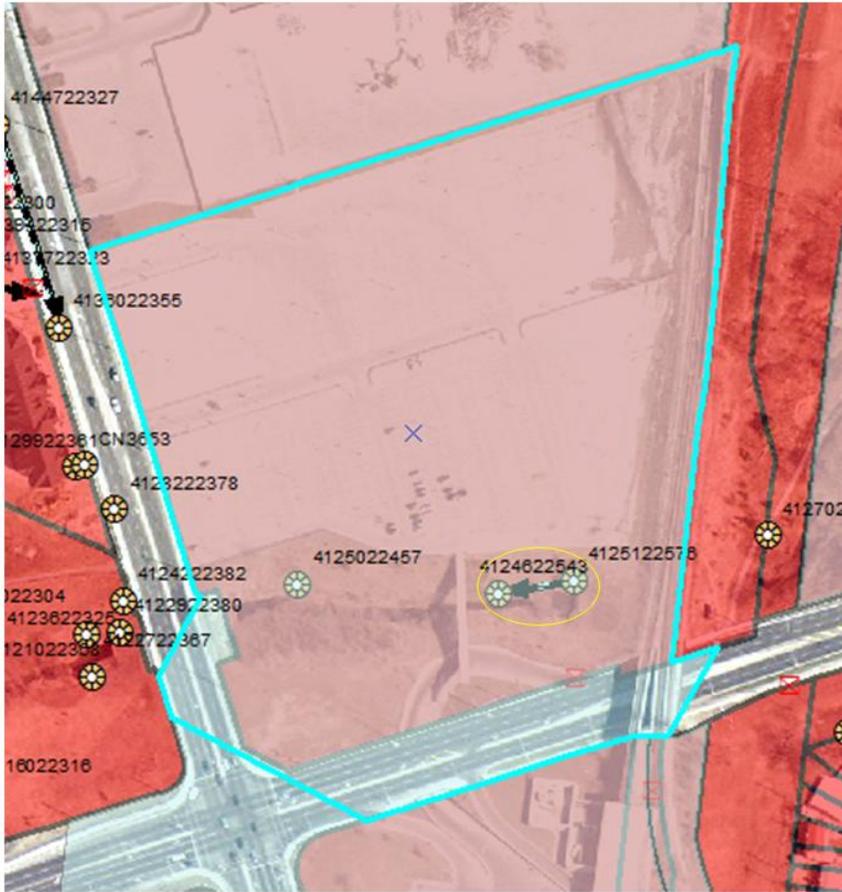
**Figure 54. DWF Error # 1: Incorrect parcel connections caused by frontage node location**

This incorrect parcel connection seen in to a sewer pipe is caused by the location of the frontage nodes. In this case, the two parcels marked by the yellow lines would be connected to the two pipes to their left if the manual approach was applied. The reason why this would occur is that all parcels are supposed to be downstream of the upstream node or the manhole at the head of the sewer pipe they are associated with, though this would not be the case if these parcels were connected to the pipe contained within the highlighted subcatchment. In addition, this error also illustrates the limitation of the tool to be able to effectively assign parcels to the appropriate subcatchment at road intersections.



**Figure 55. Error # 2 Incomplete subcatchment only containing the road corridor portion with the upstream node contained inside and Error # 3 Incomplete subcatchment only containing parcel portion with no upstream node**

The incomplete subcatchment seen above in Figure 55A is generated after converting the DWF subcatchments from multipart features to single-part features. As can be seen from Figure 55B, the subcatchment was initially a multipart feature and because there was no pipe near the parcel highlighted, the program assigned it the closest pipe which happened to be a pipe along another street. While this error is no fault of the tool, as it only attempts to connect each parcel to the nearest pipe while creating subcatchments, it is likely that a pipe that ran along that road was missing. It is probable that this missing pipe belongs to another sewer network, other than the Coxwell network and hence why it was not present within the sewer pipe network provided. If that was the case, the parcel highlighted in blue in Figure 55B should have been excluded from the Coxwell DWF subcatchment boundary provided for input into the tool. By doing so, the error of having no upstream node or pipe contained within this subcatchment's boundary would have been avoided.



**Figure 56. Error # 4 Incorrect parcel connection caused by location of pipe**

While this error (Figure 56) falls under the same type as error # 1, it has a different cause and is why it is being presented here separately. In this case, it occurs when a sanitary or combined sewer pipe is located within a parcel as seen in Figure 56 (pipe enclosed in yellow circle). Thus, while the tool has successfully located the pipe closest to the parcel in creating the subcatchment highlighted in blue in Figure 56, it cannot be correct according to the criteria stipulated in Table 6. According to these criteria, the pipe cannot be the appropriate pipe since its upstream node is not upstream of the parcel. Furthermore, parcels typically do not connect to pipes that are contained within them because these sewer pipes, though they maybe sanitary or combined pipes, are typically trunk sewers. Trunk sewers consist of larger diameter pipes that usually do not connect directly to individual parcels and instead receive flow from other smaller pipes that do have such lateral connections to individual parcels (C. Fan, personal communication, May 15, 2013). One way this error can possibly be minimized in the future is to

a use a filtering step to remove sewer pipes that are contained within parcel boundaries prior to running the tools.

### **6.3 Summary of Strengths and Weaknesses of the DWF Tools**

In order to advise potential future users of these tools, the overall strengths and weakness of the tools have been summarized. The major strengths of the tool can be summarized as follows:

1. No overlapping subcatchments unlike when delineated manually (Table 8).
2. No gaps or slivers between subcatchments unlike when delineated manually (Table 8).
3. The boundaries of the subcatchments perfectly align with the parcels within as long as the input boundary layer is delineated accordingly unlike when delineated manually.
4. Can delineate on average 66% of DWF subcatchments to an "Acceptable" level and therefore potentially reduce the time spent delineating such subcatchments by two-thirds.
5. Has no issues at delineating subcatchments when a single pipe has a connecting pipe at both ends and is along a road.
6. The number of vertices of the resulting subcatchments has been significantly reduced in comparison to earlier versions of the tools.
7. The tool is not subjective and so each time the tool is run, the same results are generated as long as the inputs have not be altered.

The limitations of weaknesses encountered when evaluating the tools were as follows:

1. Unable to always properly allocate parcels to the appropriate subcatchment that are near intersections (Error #1).
2. Unable to properly delineate subcatchments when there is a pipe within a parcel (Error #4).
3. Can only be used if no two pipes share the same upstream node, though this should not be an issue since each pipe must have its own unique upstream node for the InfoWorks CS model to function properly.
4. What is automatically generated as being the frontage node is not always what would be considered the frontage node according to the orthoimagery.
5. If the inputs have errors, these errors are magnified in the input since the tool assumes what is put inside is accurate. However, this can also be viewed as a positive because it alerts the users of possible errors within their input datasets.
6. There is no automated way to know which subcatchments are "Acceptable" after running the tool, which means time still needs to be spent on checking each subcatchment.
7. Designed to work with ArcGIS 10 (ArcInfo licence); there is no guarantee that it works with earlier or later versions.

## 7.0 Evaluating Automated WWF Subcatchment Delineation

To demonstrate the effectiveness and limitations of the automated approach presented for wet weather flow subcatchments this section highlights the tools' strengths and weaknesses. In order to determine whether a subcatchment could be categorized as either "Acceptable" or "Manual Intervention Required" the criteria listed in Table 9 were used. These criteria were developed through consultation with City staff. If any of these criterion is not met, then it is categorized as "Manual Intervention Required".

**Table 9. WWF subcatchment categorization criteria**

"Acceptable"
<ul style="list-style-type: none"> <li>✓ No gaps between this subcatchment and the surrounding subcatchment.</li> <li>✓ No overlap between this subcatchment and the surrounding subcatchment.</li> <li>✓ The portion containing the upstream node or manhole of the pipe associated with it, is part of the subcatchment.</li> <li>✓ The subcatchment has a unique upstream node and asset id combination and also has the field denoting its flow type populated.</li> <li>✓ Single part subcatchments.</li> <li>✓ The subcatchment does not contain smaller subcatchments within it.</li> <li>✓ The flow of the pipe associated to the subcatchment is either combined or storm.</li> <li>✓ Based on the contour lines layer overland flow would flow towards the location of the pipe while also respecting curb line boundaries.</li> <li>✓ The pipe associated to the subcatchment has all of its catch basins within the subcatchment.</li> <li>✓ Does not contain catch basins associated to pipes other than the pipe associated to that subcatchment.</li> <li>✓ Subcatchment does not contain multiple complete sets of pipes and catch basins within it.</li> <li>✓ A subcatchment is associated to a storm or combined pipe that has no catch basins associated to it and does not contain catch basins associated to another pipe*.</li> </ul>

\*While "Acceptable" the tool is designed to minimize its occurrence since these subcatchments would have zero runoff inflow (G. Lin, personal communication, June 25th, 2013).

### 7.1 Checking the Overall Accuracy of the Approach

As the total number of WWF Subcatchments created with the tool for both the Coxwell and Interceptor networks was 12,298 subcatchments, a similar approach employed for evaluating the DWF subcatchments was also used for the WWF subcatchments. Since the previous chapter illustrated how selecting entire subareas or an unstratified random approach was not an effective approach when attempting to evaluate the overall accuracy of a method, it was not pursued in addressing wet weather flow subcatchments. Instead randomly distributed subcatchments were

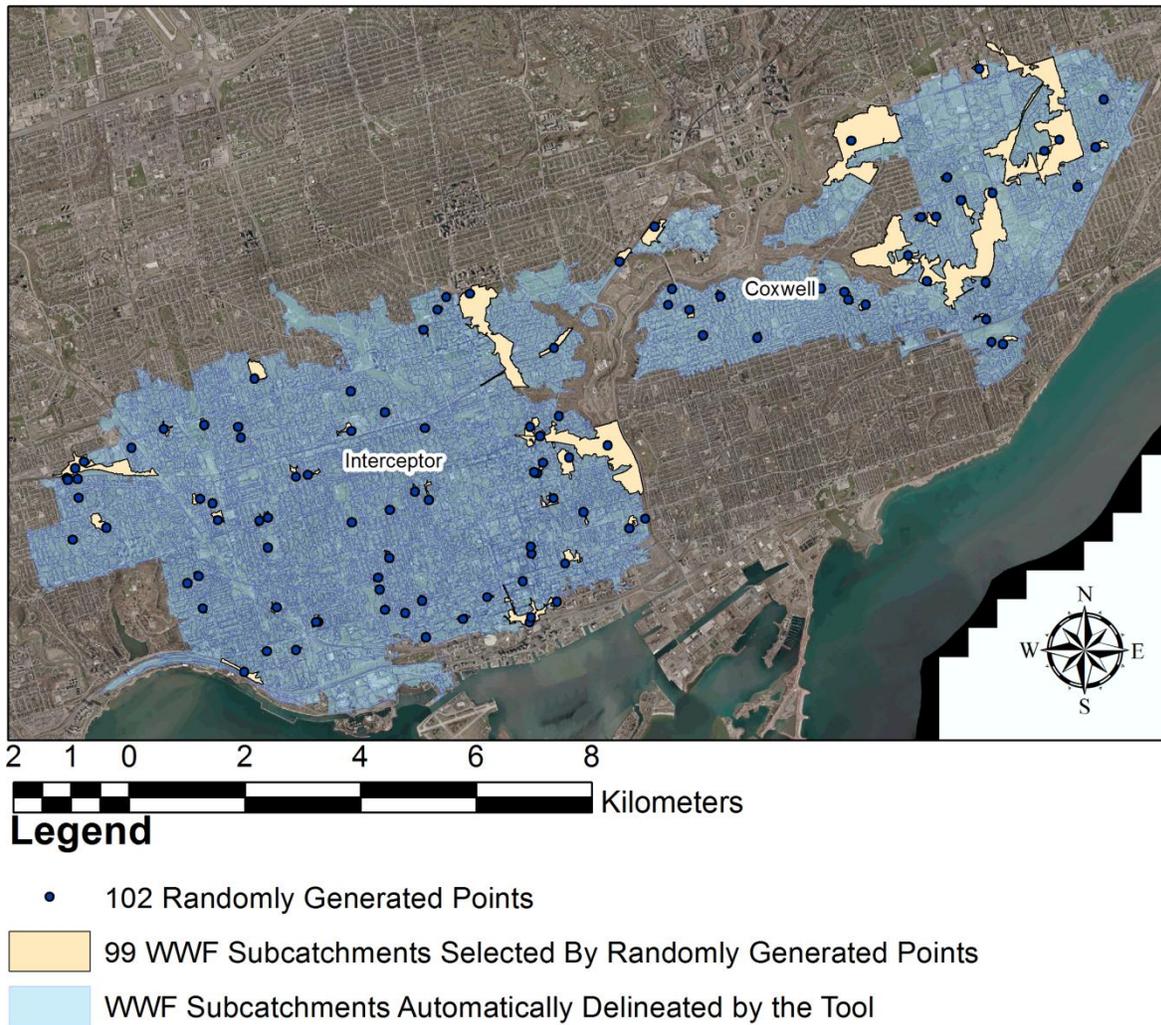
evaluated to determine the overall rate at which the automatically defined subcatchments were "Acceptable" or "Required Manual Intervention". Unlike with the DWF subcatchments, the Sampling Design Tool for ArcGIS 10.0 was used to randomly generate sampling points (NOAA Biogeography Branch, 2013).

Field	Value
Confidence Level	95%
Population Size	12298
Proportion	0.5
Confidence Interval	Selected
Upper	0.60000
Lower	0.40000
Standard Error	0.05102
Relative Standard Error	10.20
Sample Size	96

**Figure 57. Calculation used to determine sample size for WWF subcatchment evaluation.**

The online calculator determined that the required sampling size was 96 for achieving a 95% confidence level as is seen in Figure 57 (Australian Bureau of Statistics, 2013). When the tool was used and 96 points were generated only 88 subcatchments were selected because multiple points were randomly generated within the same WWF subcatchment. With 102 points the number of selected subcatchments rose to 99 (Figure 58). Thus, just as in the evaluation of DWF subcatchments, the approach of using a random number generator table in Excel was employed to randomly select 96 of the 99 selected subcatchments in order to introduce another level of randomization into the selection process. To determine which of the two categories each subcatchment belonged to ("Acceptable" or "Manual Intervention Required"), the criteria

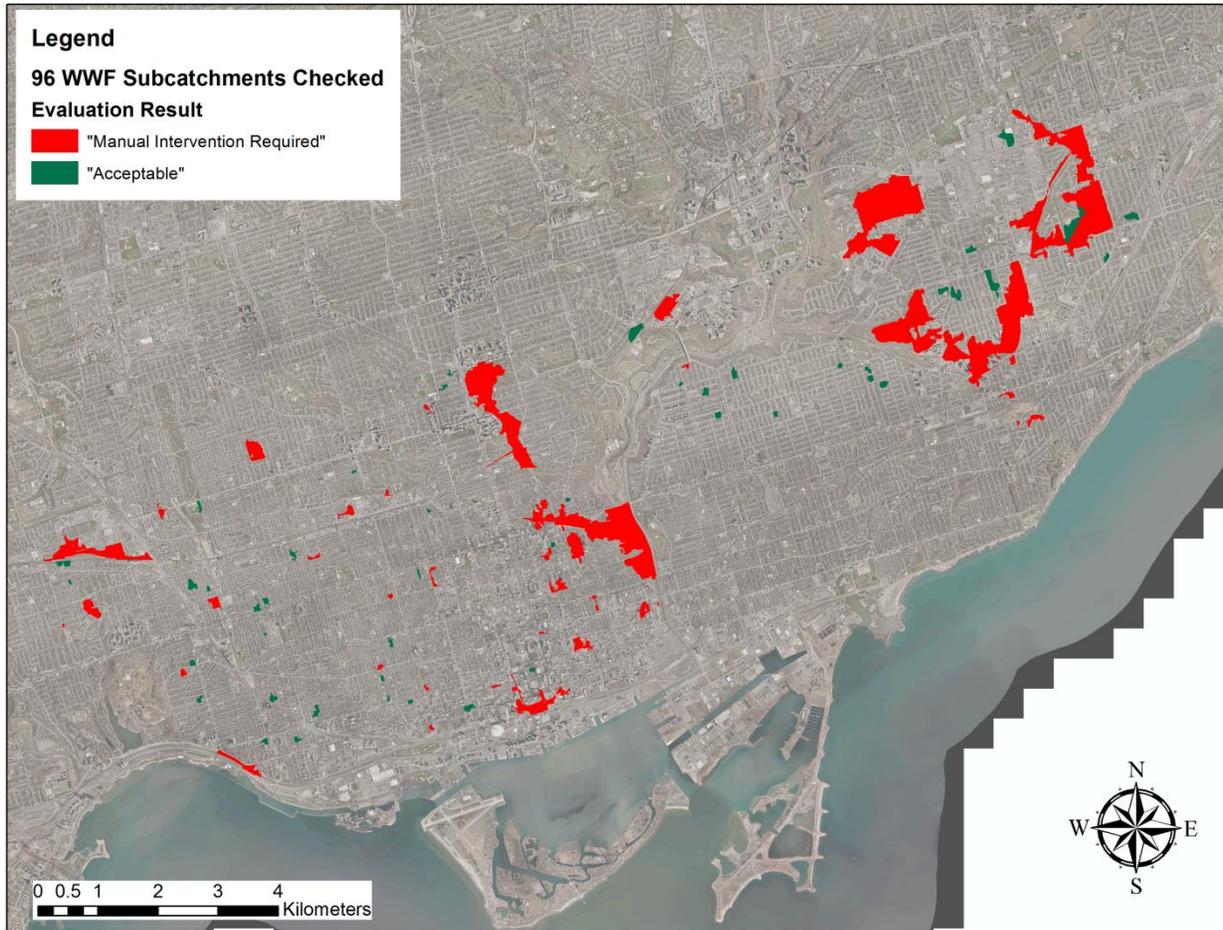
listed in Table 9 were used. In addition to the visual inspection, all the WWF subcatchments, much like the DWF subcatchments were subjected to topology checks. These topology checks were also applied to an existing shapefile of manually delineated WWF subcatchments for another InfoWorks CS model previously delineated by City staff for comparison purposes. These checks were to ensure that no errors such as overlaps or gaps between features in the dataset occurred. If such issues were present it would indicate that topology rules were not being properly enforced in the datasets or that data quality measures were not taken after the subcatchments were delineated.



**Figure 58. The 99 WWF subcatchments selected for visual inspection**

Following the visual inspection of the 96 subcatchments it was determined that the WWF subcatchment delineation tool generated only 52.1% "Acceptable" subcatchments (Appendix A.7). While these results are significant they also reinforce the reality that automatically delineating subcatchments in urban areas is still in the embryonic stages (Figure 59 & Table 10). The results also indicate another useful finding and that is similar to DWF subcatchments that required manual intervention, those WWF subcatchments that were classified into this same category tended to be much larger than their "Acceptable" counterparts. However, the difference in area between the two categories of WWF subcatchments is much more pronounced than was for DWF subcatchments because it was over a 10-tenfold increase. It is possible that these errors are caused by an insufficient number of sewer pipes being inputted into the tool for certain

regions of the two networks. This therefore identifies an area to pursue in future improvements of the tool. Furthermore, while the criteria listed in Table 9 have been used to validate the results, they could be further validated through the use of field observations.



**Figure 59. Visualization of the result of the WWF subcatchment check**

**Table 10. WWF subcatchment automatically delineated categorization results**

Type of Subcatchment	Number of Subcatchments	Total Area of Subcatchments (Ha)	Percentage
"Acceptable"	50	58.4	52.1%
"Manual Intervention Required"	46	594.7	47.9%
<b>Total</b>	96	653.1	100%

In terms of the topology check that was undertaken within ArcGIS, the results in Table 11 indicate that while no errors were detected in those generated by the tools, there were errors

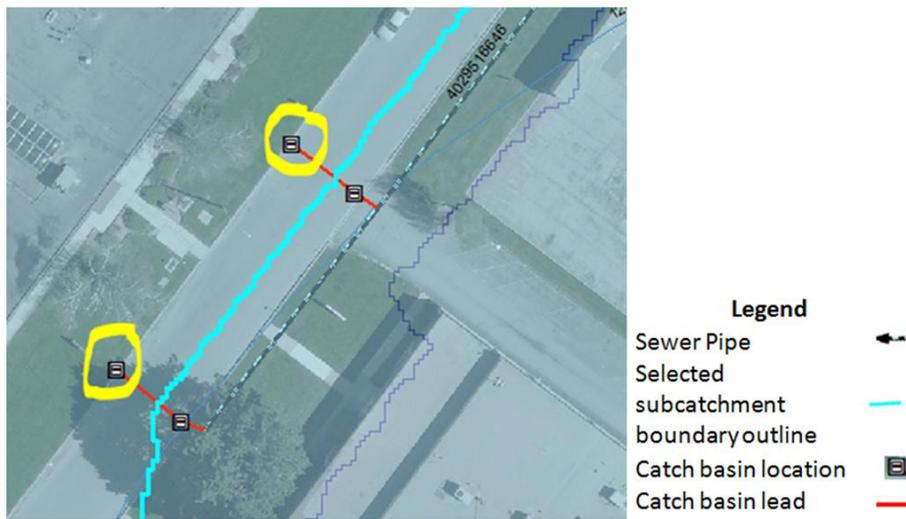
detected in those that have been manually delineated. This highlights the reality that the manual approach currently being used by the City may introduce some topological inconsistencies which should be addressed even if they choose not to employ this automated approach (Table 11).

**Table 11. Comparison of number of topology errors identified for WWF subcatchments**

Manual WWF Delineation			Automatic WWF Delineation	
Total # of Subcatchments: 1,365			Total # of Subcatchments: 12,298	
Topology Rule	Number of Errors	Area in square metres	Topology Rule	Number of Errors
Must Not Have Gaps	885	102,681.15	Must Not Have Gaps	0
Must Not Overlap	1263	955,555.82	Must Not Overlap	0
Total Errors	2,148	1,058,236.97	Total Errors	0

## 7.2 Common Errors Generated by Approach

The four errors covered here represent the range of errors that were identified while reviewing the subcatchments generated against the criteria listed in Table 9. Where possible an explanation of what caused the error and how to possibly remediate it are provided.



**Figure 60. Error # 1 Incorrect boundary of subcatchment**

In Figure 60 the subcatchment's delineated boundary is incorrect because the catch basins circled in yellow should not be part of the subcatchment highlighted in blue. Instead they belong to the subcatchment of the pipe whose upstream node id is 4029516646 as seen in Figure 60 based on the criteria listed in Table 9.



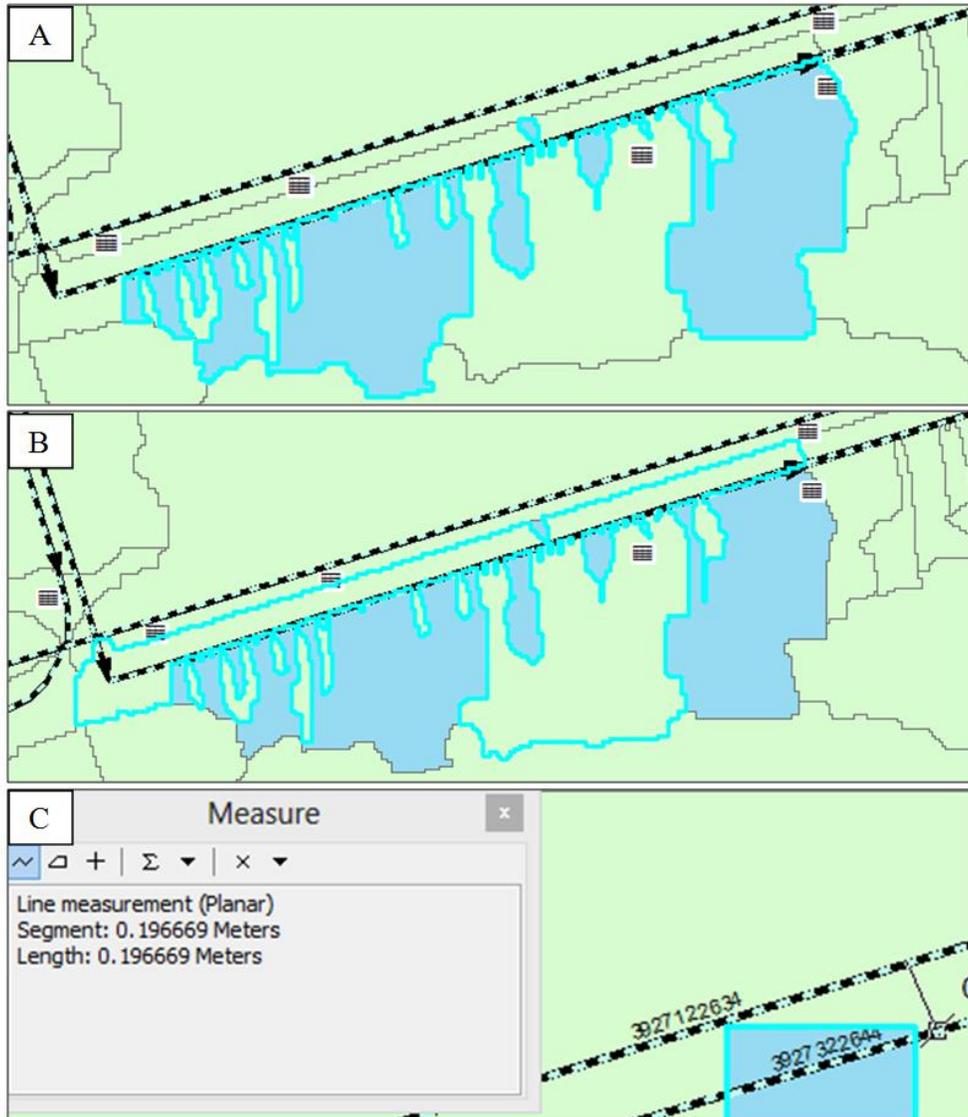
**Figure 61. Error # 2 Subcatchment contains multiple sets of catch basins belonging to multiple pipes**

The subcatchment whose partial outline is highlighted in blue in Figure 61 is much larger than it should be for a distributed model because there are numerous catch basins belonging to a number of sewer pipes. This became apparent after checking TWAG and the pipe network exported from one of the InfoWorks CS models because they both indicated the existence of numerous sewer pipes within this region. Therefore, this subcatchment should not be one large WWF subcatchment but rather multiple smaller subcatchments. However, since the pipes are missing in the figure above it means that these pipes were not properly linked to their catch basin leads within TWAG and hence why they were not selected. Furthermore, this error also highlights another limitation of the tool: if pipes are missing, the tool will fail to respect curb line boundaries in its attempt to locate the nearest pipe that was loaded into the tool when delineating subcatchment boundaries.



**Figure 62. Error # 3 Subcatchment containing another smaller subcatchment**

The subcatchment whose upstream node ID is 3784509614 has another WWF subcatchment contained within it which contradicts the criteria in Table 9. This indicates the need to detect enclosed depressions like this. Furthermore, due to the size of the larger subcatchment highlighted in Figure 62 it is also likely that additional catch basins are missing in this region from TWAG and hence why this subcatchment is so large.



**Figure 63. Error # 4 WWF subcatchments consisting of fragmented multipart features**

The two subcatchments shown in Figure 63 are fragmented multipart features because the tool is unable to effectively delineate two separate subcatchments (Figure 63A and Figure 63B) where the pipes are only 0.2 metres apart (Figure 63C). This is likely caused in part by the fact that the horizontal resolution of the DEM was insufficient because although it was resampled to a resolution where each cell represented 0.5 metres this is larger than the distance between these features. However, the DEM was not resampled further because of computer limitations and this is not what was advised in the literature (Onafrychuk, 2007; Fewtrell *et al.* 2011).

### 7.3 Strengths and Weaknesses of Approach

In order to advise potential future users of the tool, the overall strengths and weaknesses of the tool have been summarized. The major strengths of the tool can be summarized as follows:

1. No overlapping subcatchments unlike when delineated manually (Table 11).
2. No gaps or slivers between subcatchments unlike when delineated manually (Table 11).
3. Can delineate on average 52% of WWF subcatchments to an "Acceptable" level and therefore can potentially reduce the time spent delineating such subcatchments by half.
4. The method is able to automatically take into account elevation, the location of buildings and linear features when delineating subcatchments.
5. Each input pipe has a subcatchment generated for it that contains the upstream node of the pipe.
6. When the tool delineates a subcatchment to an "Acceptable" level, the catch basins associated with that pipe are automatically delineated to be within the subcatchment.
7. The tool is not subjective like an individual and so each time the tool is run the same results are generated as long as the inputs have not been altered.
8. The output from this tool can be used to generate detailed overland flow paths.
9. The Arc Hydro extension required to implement the tool is freely available.

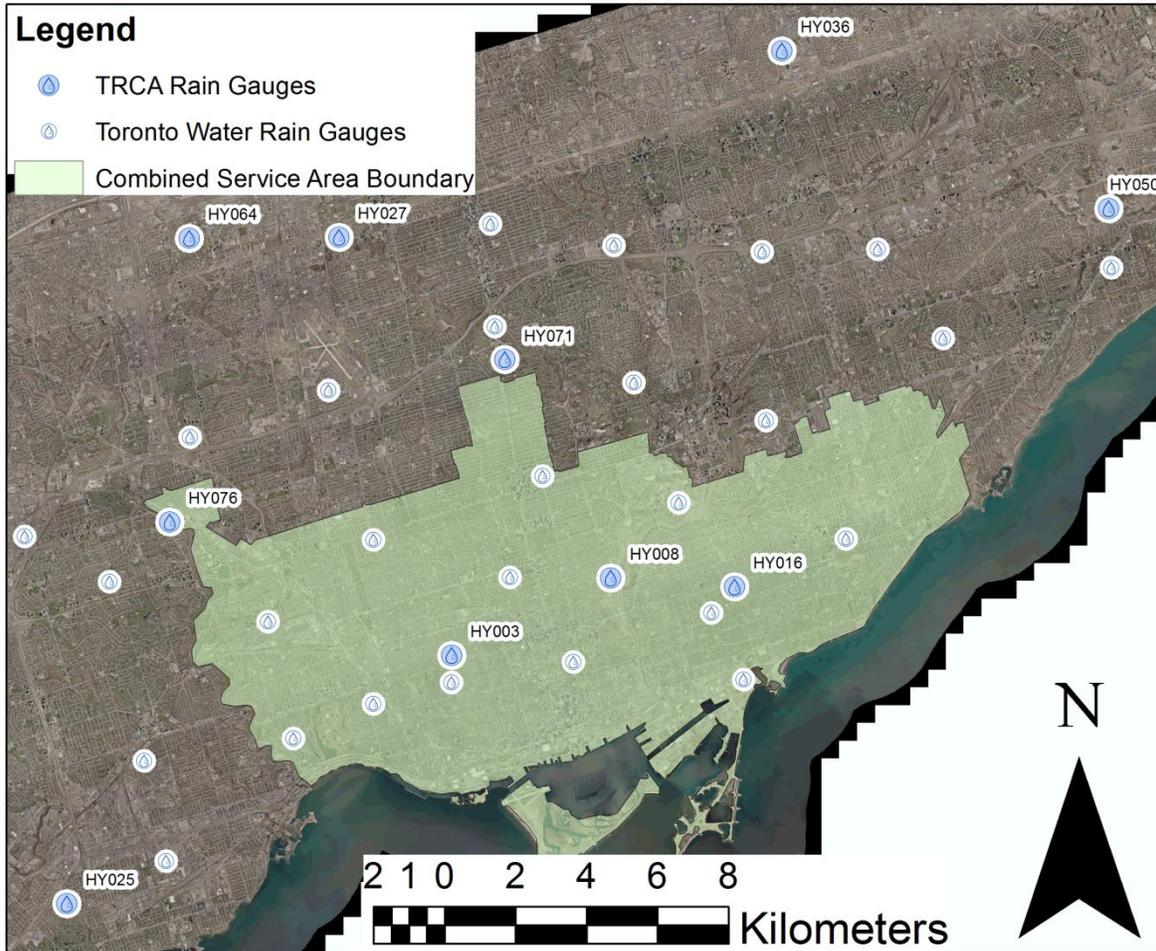
The limitations or weaknesses encountered when evaluating the tool were as follows:

1. Unable to properly delineate separate subcatchments when two pipes are only 20 cm apart. Source data needs to be confirmed to be accurate, complete, and sufficiently precise to enable the automated allocations.
2. Unable to delineate a subcatchment for a pipe if its catch basins are missing or are not linked properly to it.
3. While it can be viewed as a strength that the subcatchments generated do not necessarily follow the parcel boundaries and instead attempt to respect the boundaries of overland flow paths to drainage points, it generates slivers when attempting to split the subcatchments by parcels as specified by the City.
4. Not always able to take into account the locations of catch basins associated with each pipe when delineating subcatchments.
5. An area larger than the area of interest has to be automatically delineated as the peripheral subcatchments normally have issues.
6. There is no tool to automatically identify which subcatchments are "Acceptable" after running the tools, which means time still needs to be spent on checking each subcatchment.
7. Designed to work with ArcGIS 10 no guarantee that it works with earlier or later versions.
8. It requires a spatial analyst licence and ArcInfo level licence for its tools to function.

## **8.0 Evaluating Automated Approach to Continuous Rainfall Interpolation**

### **8.1 TRCA Rain Gauges**

In order to determine whether interpolating rainfall to generate more rain profiles was a valid approach, tests were run to validate the results. As these were interpolated data, the most effective way to test the results was to compare them against a set of independent rain gauges. In this case, the independent rainfall measurements used for validation purposes came from the Toronto and Region Conservation Authority's (TRCA's) monitoring sites (Figure 64). While it would have been ideal to secure rainfall measurements for the entire simulation period which was from April 1, 2013 at midnight to the very end of the day on October 31, 2013, these were not available. Thus, although that was the range for the records requested from the TRCA they did not have information for all of their gauges in Toronto during this period and sometimes when it was available it was not suitable for validation purposes. The reason for this is that the TRCA has a coding system which is part of their internal grading used to determine whether rainfall records are suitable for public consumption (Table 12 and Table 13). Therefore, if any of the records that were received from the TRCA had other than an Approval ID of 4 and a Grade ID of 31, they were deemed unreliable and therefore unusable. In this case that meant that the only available datasets that could be relied upon were records collected between July 1<sup>st</sup>, 2013 and August 1<sup>st</sup>, 2013, so they are the only dates that were used as points of comparison in this check. Furthermore, while this was the only period when approved rainfall measurements were available, not all of the rain gauges shown in Figure 64 met all the approval criteria or had any measurements available. Therefore only records from two of the TRCA's gauges could be used for comparison (Figure 65).



**Figure 64. Spatial arrangement of the TRCA's 9 rain gauges in Toronto**

**Table 12. TRCA approval code table**

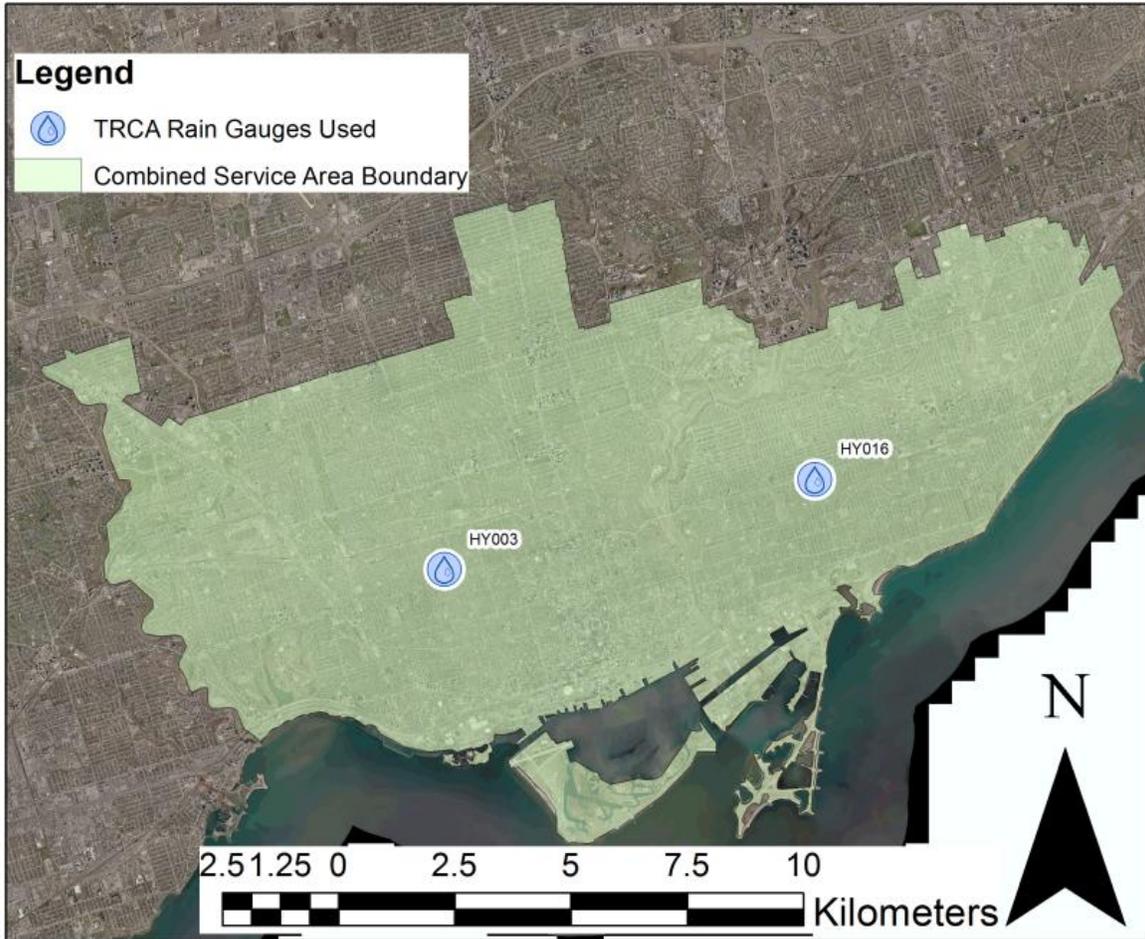
Approval ID	Approval Code	Approval Status
1	<i>WORKING</i>	Raw data
2	<i>CORRECTED</i>	First corrections completed by primary analyst
3	<i>IN REVIEW</i>	Reviewed by secondary analyst and then to be approved by primary analyst once more
4	<i>APPROVED</i>	Finalized data and available for public consumption

(Source: TRCA (J. Duncan, personal communication, February 21, 2014))

**Table 13. TRCA grade code table**

Grade ID	Grade Code	Grade Status
1	<i>UNVERIFIED</i>	Raw uncorrected and unreviewed data. Data may also remain as unverified if there is uncertainty as to the cause of unusual readings in the data. <b>Rating Curves:</b> data which is outside of the Rating Curve and extrapolation will be described as Unverified.
3	<i>ICE</i>	Ice conditions are suspected to be present at the gauge.
6	<i>DAM</i>	Will be applied when a verified physical blockage is located downstream of the gauge and impacts the water level read at the sensor. Most commonly this blockage will be either a beaver or debris jam.
11	<i>POOR</i>	Data will be marked as poor if the data is deemed to be in an unusable state by the analysts. Data marked as poor should only be made available to users with heavy caution that TRCA would not consider the data to be usable. Gaps in data will be marked as poor when it is not reasonably possible to fill those gaps. Data which is faulty due to known sensor malfunctions will also be marked as poor, but will be left in the database so as to maintain the record. Data determined to be faulty during calibration checks will be marked as poor going backwards from the date of the check to the date the analyst can best determine was the start of the failure.
30	<i>ESTIMATED GOOD</i>	<b>Stream:</b> Data which has been modeled with the Aquarius Model Based Correction tool. This data should be of similar quality of data graded Good (31). <b>Rating Curves:</b> This grade will be applied for the extrapolated portion of the curve. <b>Precip/Met:</b> Typically involve filled gaps. Filled gaps will typically be filled using a linear correction.
31	<i>GOOD</i>	Data that has been determined to contain no equipment errors and has been reviewed for errors caused by technician service. This data may include required corrections such as Drift Correction, small Gap Fills (<24 hours) etc.

(Source: TRCA (J. Duncan, personal communication, February 21, 2014))



**Figure 65. TRCA rain gauges used as points of comparison**

## **8.2 Comparing Results with TRCA Rain Gauges**

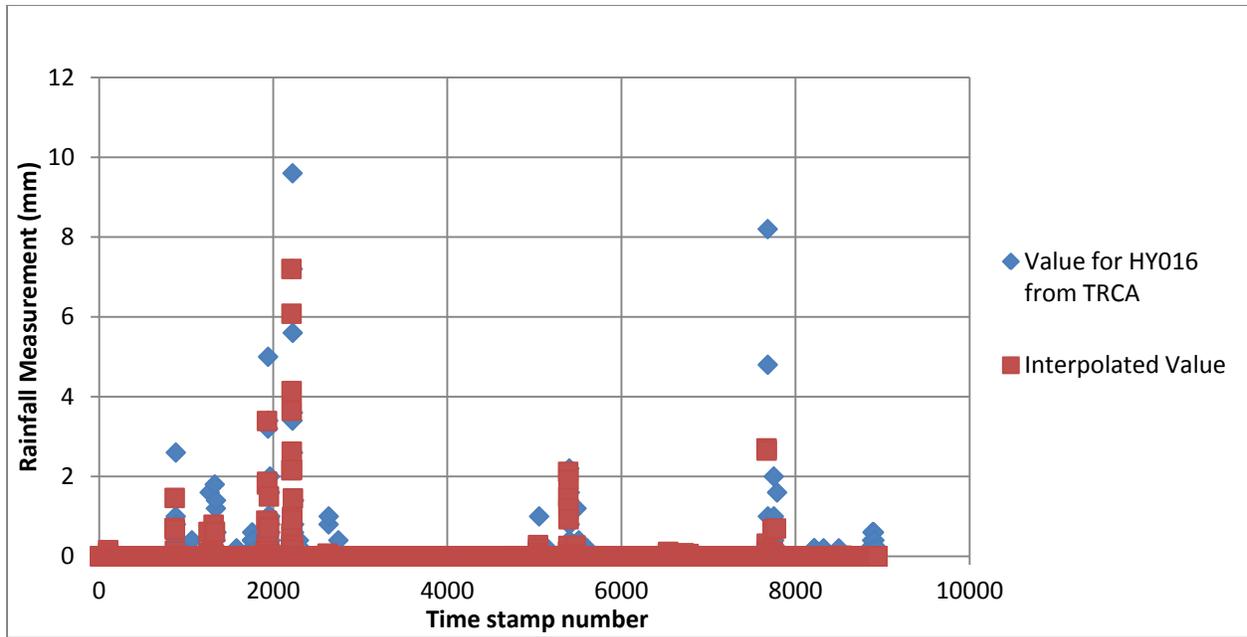
To conduct an appropriate comparison, rainfall data from all 27 rain gauges were isolated for the period between July 1, 2013 at midnight and August 2, 2013 at midnight. These rainfall records were then interpolated using the ArcGIS Continuous IDW Rainfall Interpolation Tool, developed in order to generate raster layers or prediction surfaces. To determine whether what was predicted by the interpolated surface at the exact locations of the two TRCA rain gauges, coincided with what was measured by the TRCA rain gauge, the resulting raster layers were drilled through using the point locations of the two rain gauges from the TRCA in a shapefile. This was done in the same manner that was used to extract values for the virtual rain gauges generated using the Extract Values to Table function.

The actual period of time that was compared was from 1 am on July 1, 2013 to 1 am on August 2, 2013 because this was the full extent of the period when there was usable data from any of the TRCA's gauges. Hence, after the interpolated values were extracted from the raster layers they were loaded into shapefiles and then two tables, the periods of readings from before 1 am on July 1, 2013 and after 1 am on August 2, 2013 were excluded from the analysis. To compare the values each 5-minute interval measurement or individual time step from the TRCA rain gauges was compared to the corresponding time step from those generated through interpolation and therefore representative of a virtual rain gauge at this same location (Table 14). By doing so, it was possible to compare approximately 8,929 unique 5-minute interval periods and determine whether the approach was valid.

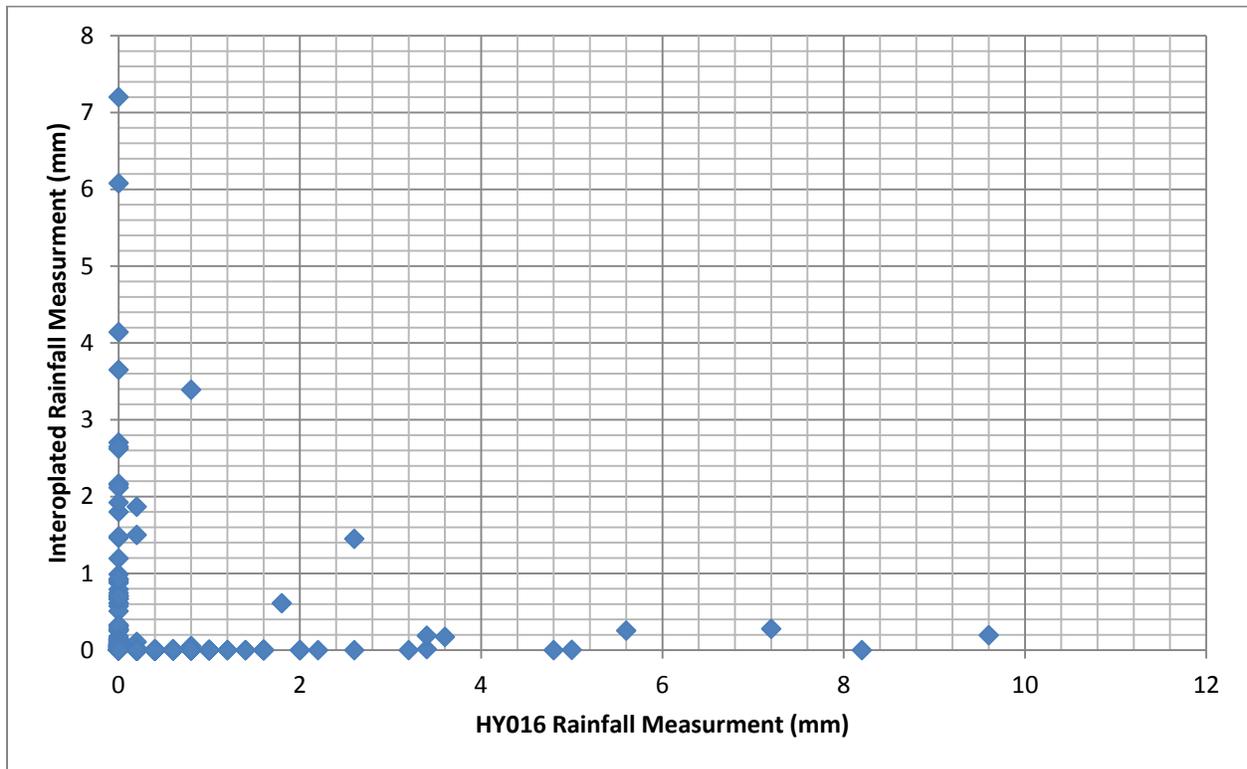
**Table 14. Excerpt of a table used for comparison purposes**

Time	Time Step (TS)	Value for HY016 from TRCA	Interpolated value	Comparison Subtraction: (TRCA-Interpolation)	Comparison Division: (TRCA/Interpolation)	Comparison Division: (Interpolation/TRCA)	(TRCA-Interpolation)/TRCA
8/1/2013 1:00	8940	9.6	0.0	9.6	#DIV/0!	0.0	1.0
8/1/2013 0:55	8939	8.2	0.0	8.2	#DIV/0!	0.0	1.0
8/1/2013 0:50	8938	7.2	0.0	7.2	#DIV/0!	0.0	1.0
8/1/2013 0:45	8937	5.6	0.0	5.6	#DIV/0!	0.0	1.0
8/1/2013 0:40	8936	5	0.0	5.0	#DIV/0!	0.0	1.0
8/1/2013 0:35	8935	4.8	0.0	4.8	#DIV/0!	0.0	1.0
8/1/2013 0:30	8934	3.6	0.0	3.6	#DIV/0!	0.0	1.0
8/1/2013 0:20	8932	3.4	0.0	3.4	#DIV/0!	0.0	1.0
8/1/2013 0:25	8933	3.4	0.0	3.4	#DIV/0!	0.0	1.0
8/1/2013 0:15	8931	3.2	0.0	3.2	#DIV/0!	0.0	1.0
8/1/2013 0:05	8929	2.6	0.0	2.6	#DIV/0!	0.0	1.0
8/1/2013 0:10	8930	2.6	0.0	2.6	#DIV/0!	0.0	1.0
8/1/2013 0:00	8928	2.2	0.0	2.2	#DIV/0!	0.0	1.0
7/31/2013 23:50	8926	2	0.0	2.0	#DIV/0!	0.0	1.0
7/31/2013 23:55	8927	2	0.0	2.0	#DIV/0!	0.0	1.0

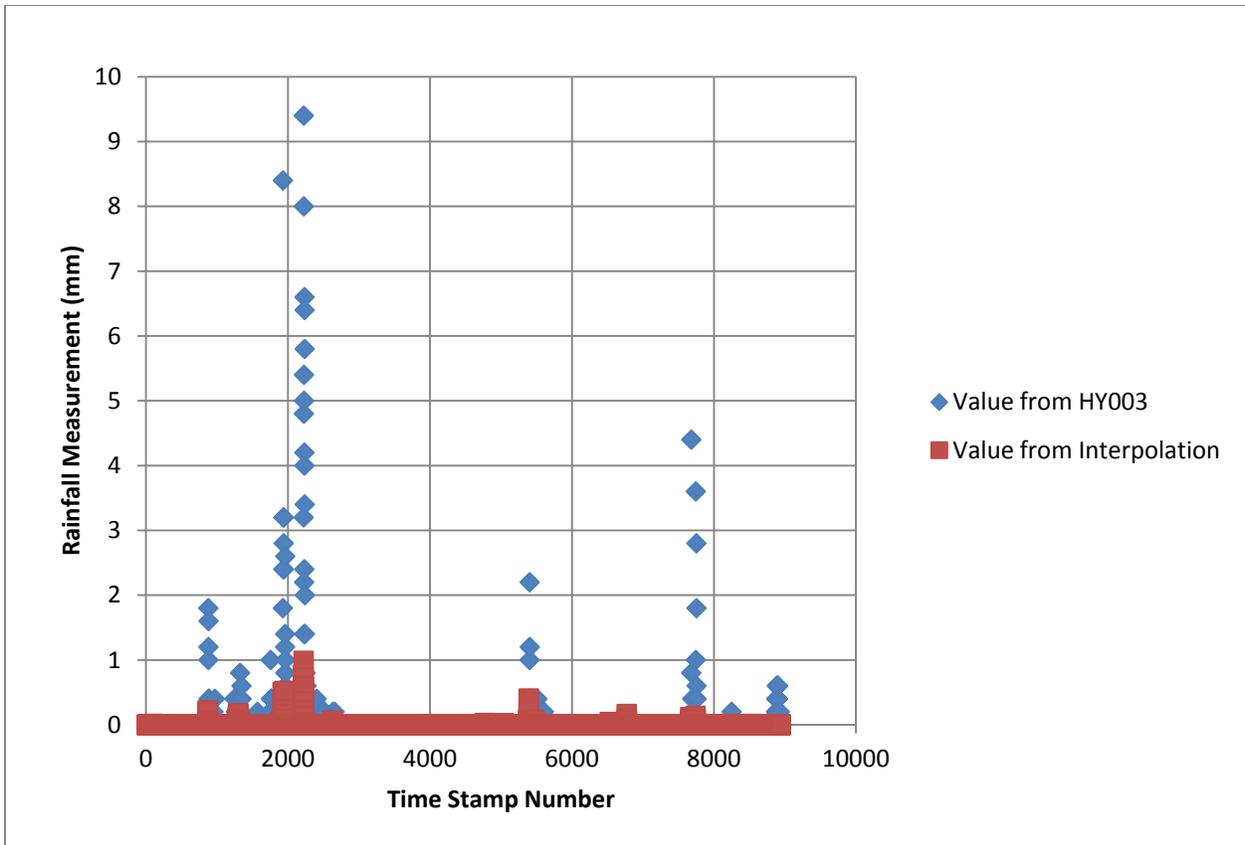
Through comparing each time step with the use of a table like the one shown in Table 14 it was possible to evaluate the validity of the method employed. The results of this evaluation are summarized in Table 15 and as can be seen from these results, at no point was there a time when the measurements were the same, except for when neither of the gauges recorded any rain (Appendix A.8). Instead, the closest the interpolated values or the virtual gauges and the actual rainfall value came to be the same, was on July 7, 2013 at 16:55 (4:55 pm) when the TRCA gauge HY016 recorded a value of 0.20 mm of rain and the virtual gauge predicted that 0.11 mm of rain had fallen. Furthermore, it was also found that while the virtual gauges had the tendency of both underestimating and overestimating actual rainfall, it more often than not underestimated rainfall as can be seen from Figure 66, Figure 67, Figure 68 and Figure 69. However that was expected, since it was based on IDW interpolations and when such interpolations are used, it always has values that are between the maximum and minimum values found in the dataset (ESRI, 2010). In addition, when aggregating the records to calculate total rainfall for the month of July, it was determined that the TRCA gauge (HY003) recorded a total of 178 mm. The nearest physical gauge from the City (RG-003) recorded a total of 162.5 mm while the virtual gauge indicated only a total of 21 mm. Similarly, when aggregating the records for the TRCA gauge (HY016), it was determined that the total rainfall recorded was 147 mm. The nearest City gauge (RG-007) recorded a total of 134.75 mm and the virtual gauge only indicated a total of 74.4 mm. Therefore, the amount of rainfall estimated by the virtual gauges would likely not be more than the total rainfall that is assumed when using the Thiessen polygons currently employed by the City.



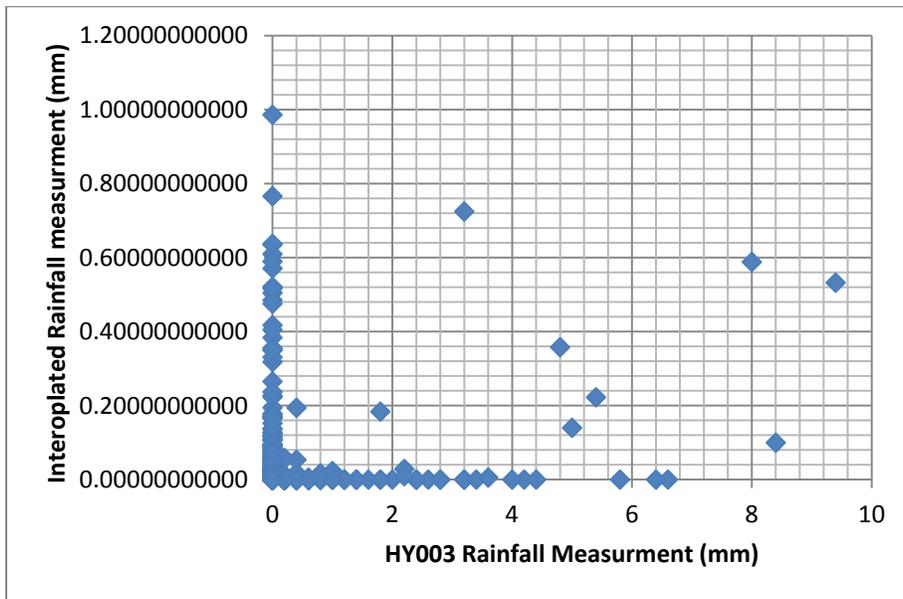
**Figure 66. Comparison of TRCA HY016 Values Versus Interpolated Values Over Time**



**Figure 67. Scatter plot of TRCA HY016 rainfall measurements and the corresponding interpolated measurements**

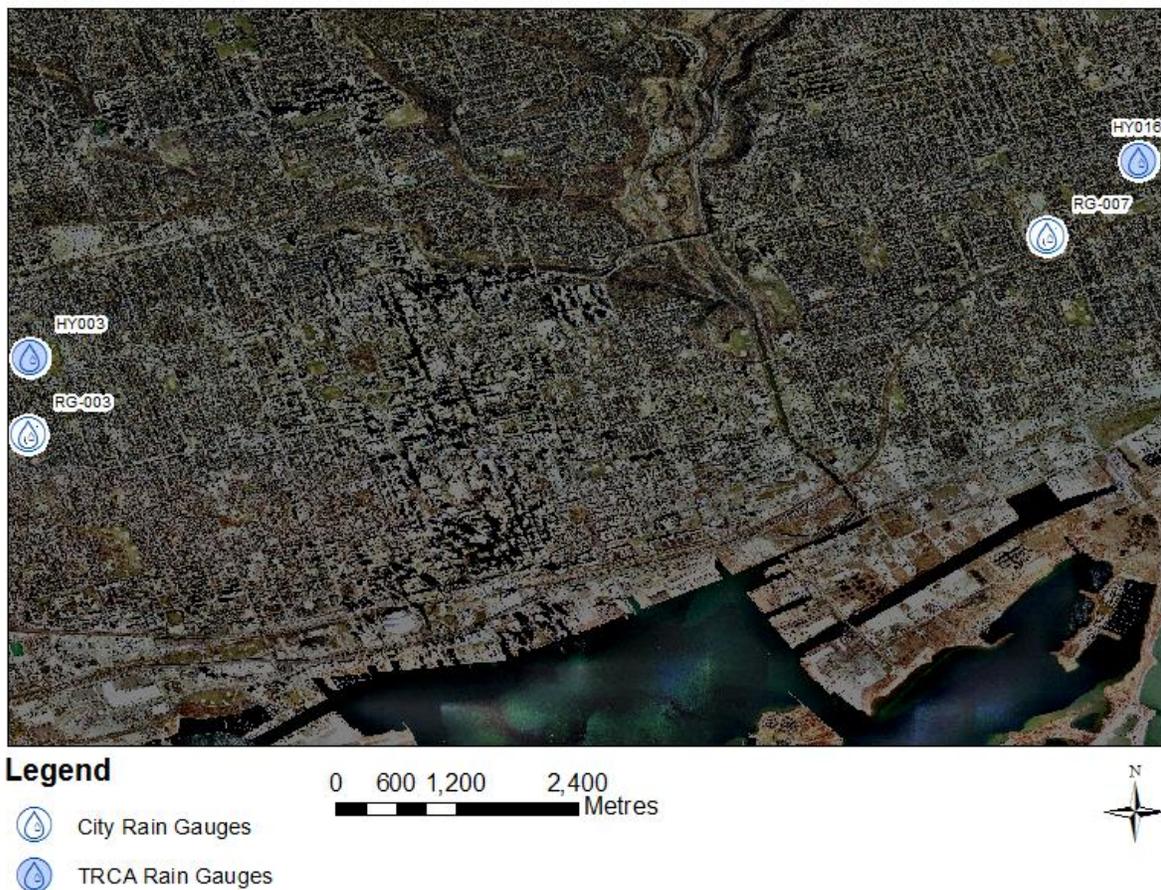


**Figure 68. Comparison of TRCA HY003 Values Versus Interpolated Values Over Time**



**Figure 69. Scatter plot of TRCA HY003 rainfall measurements and the corresponding interpolated measurements**

However, what was most concerning is the degree to which the rainfall was supposed to have occurred according to one dataset and not another and likely one of the main factors why the correlation coefficient values for these datasets calculated within Excel were so low 0.05 and 0.24 (Table 15) and also evident from Figure 67 and Figure 69. The correlation coefficient "is a summary value of a large set of data representing the degree of linear association between two measured variables" (Taylor, 1990, p. 39). Hence, a value of +1 signifies that two datasets share a very strong positive relationship (Taylor, 1990). Even when comparing the correlation coefficient between the closest city rain gauge to each TRCA rain gauge location, the correlation values were only 0.08 and 0.31 (Figure 70). Thus, this discrepancy is likely related to a fundamental difference in the values between the input datasets, the City's rain gauge measurements and the TRCA's rain gauge measurements. However, it may have been that rainfall was inherently inconsistent especially over the episodes and locations examined. Thus, this warranted further investigation due to the limited evidence.



**Figure 70. Location of the rain gauges used in the correlation coefficient calculations**

**Table 15. Evaluation of virtual gauges derived from interpolation**

	<b>HY016</b>	<b>HY003</b>
Frequency of intervals when rainfall was underestimated by the virtual gauge* (When values above zero appear in both datasets)	28 28/8929 = 0.31%	30 30/8929 = 0.34%
Frequency of intervals when rainfall was overestimated by the virtual gauge* (When values above zero appear in both datasets)	3 3/8929 = 0.03%	0 0/8929 = 0%
Frequency of intervals when rainfall occurred according to virtual gauge but not the TRCA gauge	159 159/8929 = 1.78%	167 167/8929 = 1.87%
Frequency of intervals when rainfall was detected by the TRCA gauge but not the virtual gauge	174 174/8929 = 1.95%	155 155/8929 = 1.74%
Largest discrepancy when rainfall occurred according to both datasets (TRCA having the larger value)	9.4 mm	3.6 mm
Smallest discrepancy when rainfall occurred according to both datasets (TRCA having the larger value)	0.09 mm	0.2 mm
Average discrepancy when rainfall occurred according to both datasets (TRCA having the larger value)	1.7 mm	1.9 mm
Maximum multiple a rainfall value was underestimated by the virtual gauge	1328 times	644 times
Minimum multiple a rainfall value was underestimated by the virtual gauge	1.8 times	2 times
Average multiple a rainfall value was underestimated by the virtual gauge	132 times	65 times
Largest discrepancy when rainfall occurred according to both datasets (Virtual Gauge having the larger value)	2.6 mm	N/A
Smallest discrepancy when rainfall occurred according to both datasets (Virtual Gauge having the larger value)	1.7 mm	N/A
Average discrepancy when rainfall occurred according to both datasets (Virtual Gauge having the larger value)	1.8 mm	N/A
Maximum multiple a rainfall value was overestimated by the virtual gauge	9.3 times	N/A
Minimum multiple a rainfall value was overestimated by the virtual gauge	4.2 times	N/A
Average multiple a rainfall value was overestimated by the virtual gauge	7 times	N/A
Correlation Coefficient Value Between the two Datasets	0.05	0.24

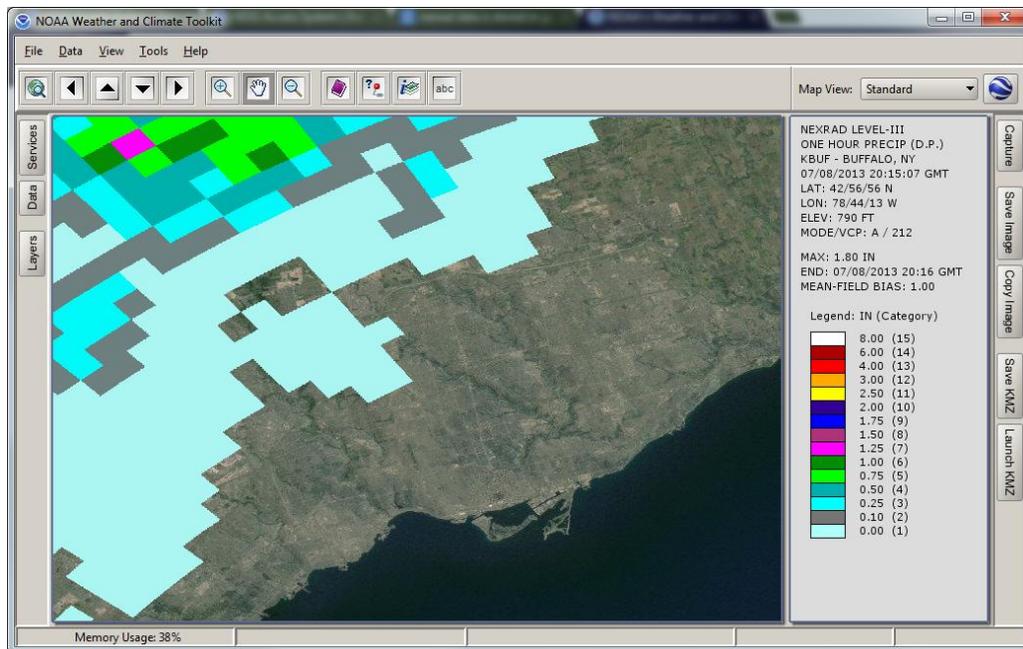
\* - Virtual gauge is synonymous with extracted interpolated values

### **8.3 Utilizing NEXRAD Datasets as an Additional Source for Comparison**

The discrepancy between the two rain datasets raised concerns, which prompted the consideration of radar rainfall records as an additional source of comparison. Radar datasets for the study area from the Buffalo Next-Generation Radar (NEXRAD) station was freely available online (National Climatic Data Center, 2014). Utilizing the visualization capabilities provided by NOAA's Weather and Climate Toolkit (Ansari, 2013) consideration was given to determine whether the City's or the TRCA's gauges more closely coincided with the radar values at times when there were discrepancies between the two gauges during the month of July (Figure 71). Actual comparisons between rainfall measurements could not be conducted, as it was uncertain if or how the datasets had been calibrated.

Based on the limited comparison of several events that was possible, the TRCA gauges were seemingly more accurate. A comparison of all times was not possible because unlike the rain gauge recordings, which were measured on a consistent 5-interval basis, the available NEXRAD radar recordings were not. As such, one radar reading might be for 23:25 and another at 23:28. Furthermore, all the radar readings unlike those from the City or the TRCA were recorded using GMT time as can be seen in Figure 71, so necessary adjustments had be made and are reflected in the comparisons carried out.

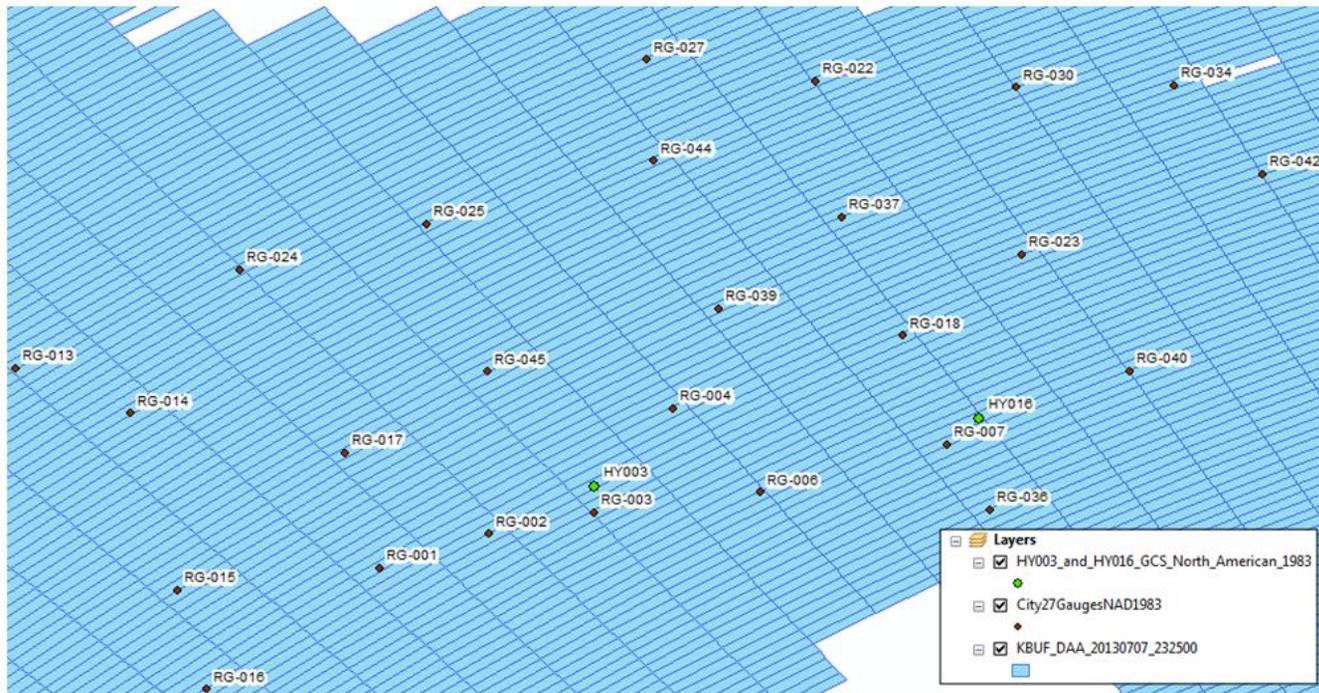
At times when precipitation should have been registered across all the City's gauges, they did not do so whereas those by the TRCA did (Table 16 and Figure 72). There were also times when neither the TRCA's gauge nor NEXRAD detected precipitation but the City's gauges still recorded rainfall (Table 17 and Figure 73). At other times the TRCA's gauges recorded rain that was detected by NEXRAD but not one of the City's gauges recorded any rainfall (Table 18 and Figure 74). However, it remains inconclusive because of the small time period examined and therefore by no means can be considered comprehensive enough to draw any conclusions. Thus, this warrants further investigation due to the limited evidence for the specific time and locations of this study.



**Figure 71. NEXRAD radar visualization**

**Table 16. Intensity values recorded by the City's 27 gauges expressed in mm per hour and those of the TRCA's measured in mm at 19:25 on July 7th, 2013.**

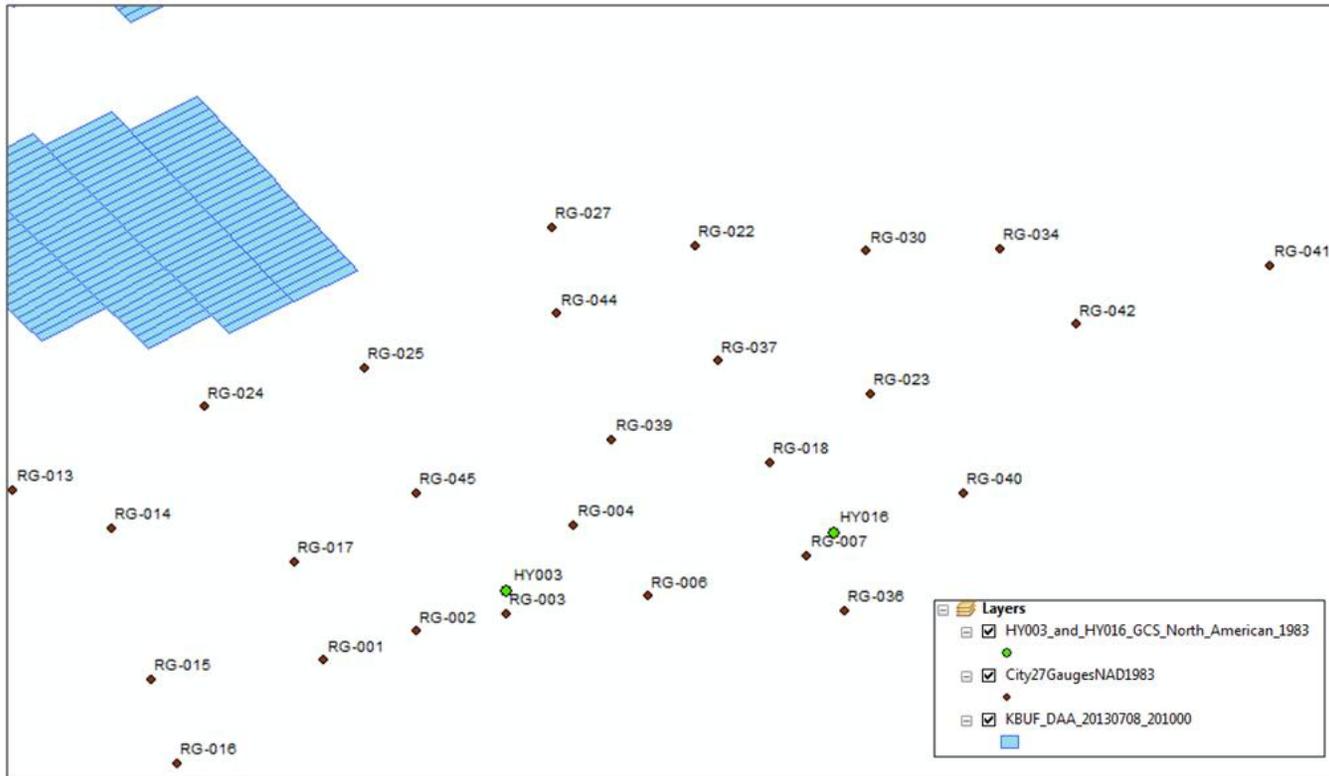
1	Time	RG-001	RG-002	RG-003	RG-004	RG-006	RG-007	RG-013	RG-014	RG-015	RG-016	RG-017	RG-018	RG-022	RG-023	RG-024	RG-025	RG-027	RG-030	RG-034	RG-036	RG-037	RG-039	RG-040	RG-041	RG-042	RG-044	RG-045	HY003	Interpolated HY003	HY016	Interpolated HY016
1951	07/07/2013 19:25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3	3	0	0	2.6	0.0000000000	1.6	0.0000000000



**Figure 72. Location of precipitation according to radar (blue rectangles) relative to gauges at 19:25 (23:25 GMT) on July 7th, 2013.**

**Table 17. Intensity values recorded by the City's 27 gauges expressed in mm per hour and those of the TRCA's measured in mm on July 8th, 2013 at 16:10**

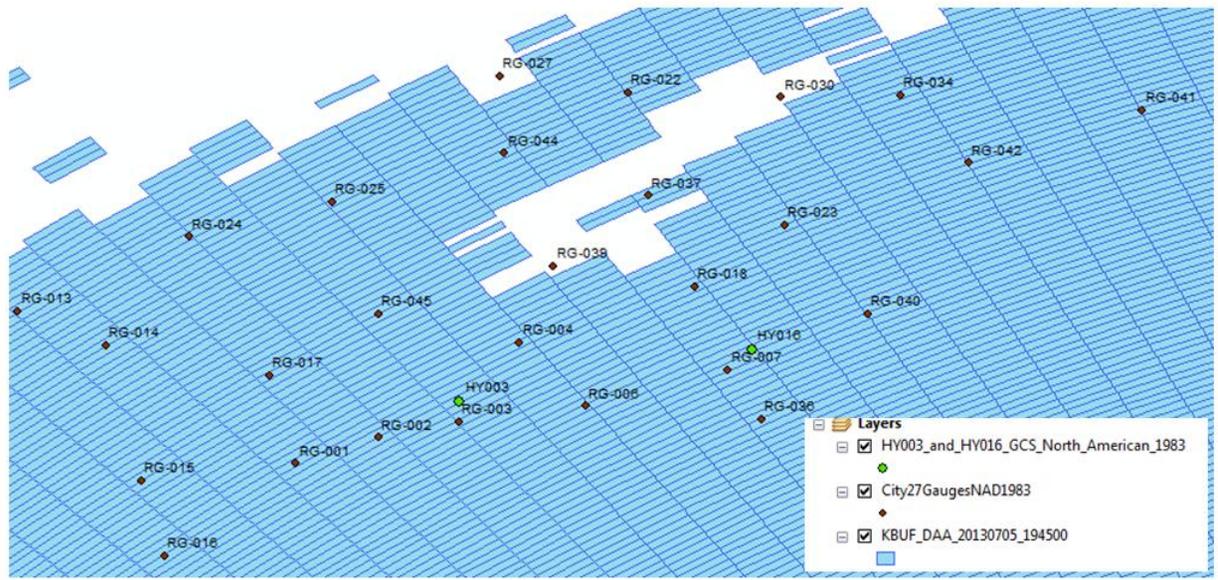
Time	RG-001	RG-002	RG-003	RG-004	RG-006	RG-007	RG-013	RG-014	RG-015	RG-016	RG-017	RG-018	RG-022	RG-023	RG-024	RG-025	RG-027	RG-030	RG-034	RG-036	RG-037	RG-039	RG-040	RG-041	RG-042	RG-044	RG-045	HY003	Interpolated HY003	HY016	Interpolated HY016
07/08/2013 16:10	3	6	78	99	60	90	63	42	27	0	45	87	27	66	51	45	73.152	42	0	51	48	84	93	0	0	78	0	0	0.63514554501	0	6.07869005203



**Figure 73. Location of precipitation according to radar (blue rectangles) relative to gauges at 16:10 (20:10 GMT) on July 8th, 2013.**

**Table 18. Intensity values recorded by the City's 27 gauges expressed in mm per hour and those of the TRCA's measured in mm on July 5th, 2013 at 15:45**

Time	RG-001	RG-002	RG-003	RG-004	RG-006	RG-007	RG-013	RG-014	RG-015	RG-016	RG-017	RG-018	RG-022	RG-023	RG-024	RG-025	RG-027	RG-030	RG-034	RG-036	RG-037	RG-039	RG-040	RG-041	RG-042	RG-044	RG-045	HY003	Interpolated HY003	HY016	Interpolated HY016	
07/05/2013 15:45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.000000000000	1.2	0.000000000000



**Figure 74. Location of precipitation according to radar (blue rectangles) relative to gauges at 15:45 (19:45 GMT) on July 5th, 2013.**

## 8.4 Summary of Strengths and Weaknesses of Approach

In order to advise potential future users of these tools, the overall strengths and weaknesses of the tools have been summarized. The major strengths of the tool can be summarized as follows:

1. Able to continuously interpolate rainfall, this is more efficient than the traditional approach in ArcMap of having to manually initiate each interpolation run.
2. Possible to generate an almost unlimited number of virtual rain gauges this means that more rain profiles than ever before can be loaded in the model.
3. Since it was designed in ModelBuilder, it is quite easy to make adjustments to improve the interpolation results depending on the nature of the input dataset such as switching to spline for the interpolation method. This means that the tool can be used to predict and estimate other atmospheric parameters such as pressure and temperature relatively easily.

The limitations and weaknesses encountered when evaluating the tools were as follows:

1. Results are only as good as the input data.
2. Was not able to predict any rainfall value above zero mm that were identical to those found in the dataset provided by the TRCA for the time increments considered (5 minutes).
3. Tends to greatly underestimate rainfall in its current design.
4. Built and designed to run in ArcGIS 10 so no guarantee that it will work in other versions of ArcGIS and it also requires the spatial analyst extension to execute the actual interpolations.
5. In order to fully benefit from numerous rain profiles that can be generated for modelling purposes requires a high-end workstation.

## 9.0 Conclusion, Future Research and Recommendations

### 9.1 Conclusion

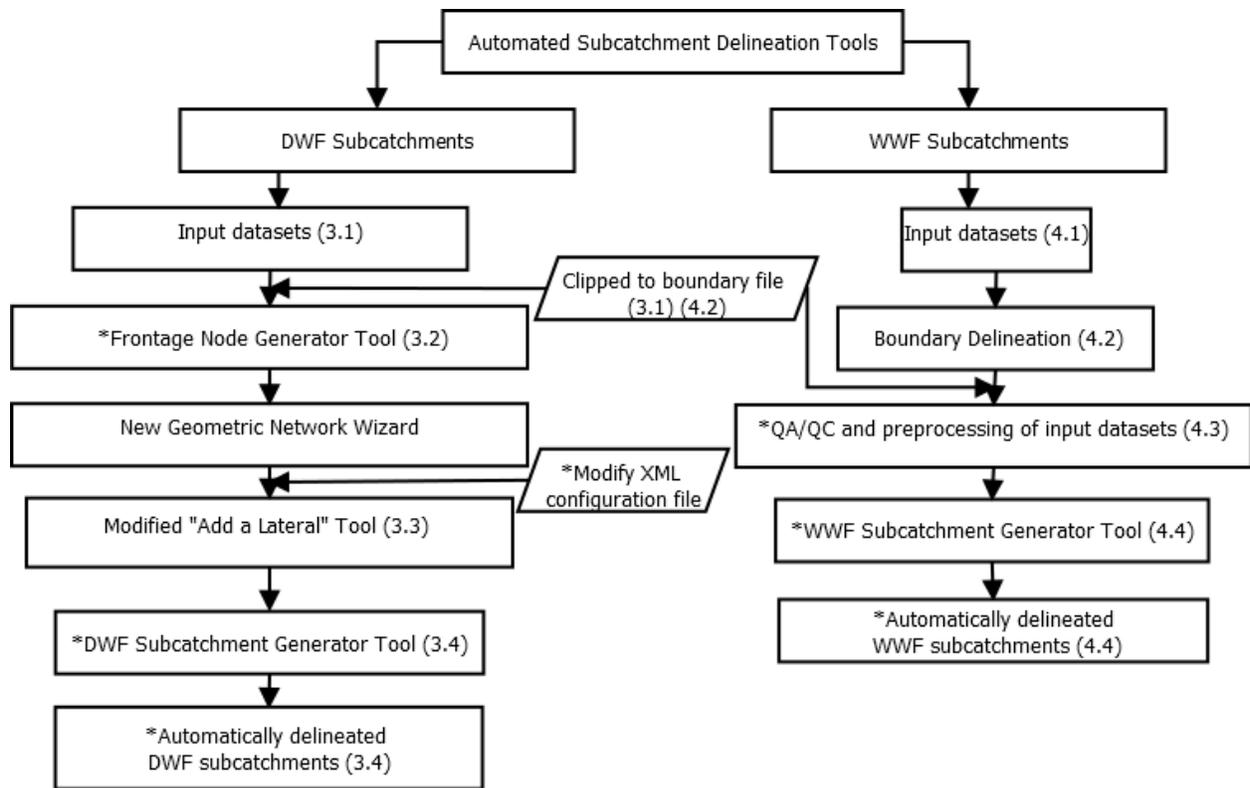
The research undertaken here was initiated by a real-world need for the creation of GIS-based tools to ensure efficient urban drainage model development by the City of Toronto for CSO reporting purposes. It was also set in the larger context that such tools are not only needed by the City of Toronto but by municipalities across the globe. This is because the use of GIS-input-driven distributed urban-drainage modelling has become more widespread as urban centres prepare for the future challenges of climate change coupled with their growing populations (IPCC, 2007; Dongquan *et al.* 2009; McCarthy *et al.* 2010; Allan, 2011; Simões *et al.* 2011; Leitão *et al.* 2012; Blumensaat *et al.* 2012). At the onset, it was uncertain whether this could be achieved. Hence, this was an experiment into whether GIS-based tools could deliver and they did. This is because the central aim, to develop improved automated methods for delineating subcatchments that are more efficient than relying solely upon the manual method, has been achieved to a certain degree. While neither of the automated approaches developed through this research was at a level where it can completely replace the manual method, this was not necessarily the ultimate goal. Instead, City staff stated that they would be satisfied if the tool could reduce the amount of manual labour involved by approximately two-thirds (G. Lin, personal communication, July 18, 2013). In this research, the automated approach developed for delineating DWF subcatchments achieved success, being able to delineate on average 66% of subcatchments to an "Acceptable" level. As such, it is possible that these tools could potentially reduce the time it takes to delineate DWF subcatchments for a model by two-thirds. Although, the automated approach for WWF subcatchment delineation did achieve the same level of success, being able to automatically delineate 52% of subcatchments to an "Acceptable" level, this should not be viewed as a failure but as an opportunity for continued improvement.

If the tools had been designed using a programming language such as Python this would not be as easy. However, since all the tools were designed with its first end-users in mind, City staff at Toronto Water, they were based in the ModelBuilder environment, which provides a platform that is easy to learn and use while still being very powerful (Allen, 2011). Furthermore, there remains scope and flexibility in continually improving the performance of the tool by City staff because of ModelBuilder's modular design. As such, if City staff want to "burn in" the

roads even further down for example, 50 metres instead of the 0.5 metres that was used in this research in an attempt to prevent water from flowing over curb line boundaries, they could easily do so. Thus, the tools could be thought of as being as organic as the models they are being used to create because they could be and should be continuously improved.

This reflects the reality that what may appear to be straightforward processes to automate, are much more complex than they first appear. Therefore, while it was possible to improve by building upon past methods as was done here, room remains for future improvements. In addition, while what was done to create the final tools might seem simplistic in retrospect it should be noted that many designs were tested and it was found that making the tools as simple as possible would improve the chances of them being adopted. Thus, one of the major contributions of this work has been to be able to translate the logic behind the manual approaches of subcatchment delineation into simple tools that could be used by a non-expert efficiently.

Furthermore, the manual method is not without issues. The biggest issue with dry weather flow is that many assumptions made while manually delineating a subcatchment may not necessarily be true. While it was assumed that the sanitary outflow pipe from a building on a parcel is on the side that is considered the parcel's frontage or a property's driveway, there is no easy way to demonstrate this. If the City had a GIS database of all the private sewer lateral connections for every parcel in Toronto, the process of delineating subcatchments could be done with much more confidence. However, no such database exists since these lateral connections are viewed as being private and therefore the City has not developed such a database. In addition, such information is not always present in the as-built drawings they have for their pipes. The only way to be sure that what is manually delineated is accurate, would be to conduct a field check, which is not economically viable as there are over 470,000 parcels across the City. That is why such assumptions have been in place and have been used. In addition, the manual method for both DWF and WWF has issues in the actual delineation processes such as the great number of topology errors that are generated that could perhaps be addressed through the application of some explicit topology rules. Alternatively, the topology inspection tool within ArcMap as was utilized in this research, could also be considered. The subjectivity involved in the manual process could be reduced by the use of the tools such as those that have been presented (Figure 75).



**Figure 75. Flowchart of the automated subcatchment delineation tools presented**

This can also be viewed as an area where the use of GIS in this research has attempted to add a more scientific basis to the delineation process by making it more repeatable. As the manual process is subjective by nature the subcatchments delineated in such a fashion will be slightly different from modeller to modeller with no assurance that the same subcatchments will be delineated each time. Thus, the delineation of subcatchments becomes more of an art rather than science, which is something that using these tools tries to prevent. Furthermore, while WWF subcatchments have been previously delineated through automated means by a consultancy, City staff were not aware of how it was actually determined and when asked could not explain why they were correct (G. Lin, February 13, 2013). With the development of the WWF tool, they now have this same capability in-house and have an understanding of the logic used to create these subcatchments.

The secondary aim to develop an automated approach to generate more rain profiles than available under the Thiessen Polygon Method was also successful with opportunities for enhancement. While the approach developed through this research can be used to successfully

generate far more rain profiles than available via the Thiessen approach, in a manner that is seemingly unprecedented in a municipal environment, it has provided opportunities for further investigation of efficiency in the management of urban rain gauge monitoring. In this case, it has raised concerns about the accuracy of the input rainfall records that the City of Toronto currently uses for modelling purposes which overall does not differ by more than 16 mm for the month of July from the TRCA records examined. Although using interpolations to generate additional rain profiles was originally introduced to the City by a consultant firm they only simulated two events. Therefore, it was not necessary to develop a tool like the one presented in this document that was able to run and generate rainfall for a period of over 7 months, through interpolations requiring only a few hours to complete. This is why the tool can be considered innovative. Furthermore, it is not clear whether the validity of their results (Clarifica, 2010) was evaluated using independent rain gauges as was attempted in this research.

In addition, the results generated by this tool raise an important point: although the capability to generate additional virtual rain gauges has been demonstrated through this research, it does not automatically mean that the results generated are accurate. In this case, it is known that the tool's output would likely underestimate rainfall so this should be considered by any potential user. This can also be viewed as an opportunity to continually improve the tool because while the main challenge of generating interpolations of a continuous basis has been overcome through this research, the tool should be optimized through adjusting various parameters such as changing the sampling size or even the interpolation method. Until then the results from the Thiessen polygon method while seemingly crude, are closer to the measurements from the TRCA gauges than those from the virtual gauges. One possible reason for this is that the period analyzed was during the summertime, when convective thunderstorms occur which are characterized by sharp boundaries (D. Joksimovic, personal communication, August 11, 2014). In addition, the analysis also demonstrated the importance of having independent gauges to validate the results when generating such records, otherwise one cannot have confidence in the results generated nor realize that perhaps the raw datasets were not as accurate as was first thought, recognizing that there is no absolute datum for establishing "reality". Nonetheless, another way the interpolation results can be validated is by using only a subsample of the City's gauges for generating the interpolation raster and then use those gauges excluded from the interpolation as a means of validation.

Overall, the objective and its associated tasks have been met. From the tools presented, covering subcatchment delineation to the generation of input rainfall datasets, it should be clear that the objective has been met. It is hoped that this document has clearly identified the urgent need for the tools presented here, as 1-D distributed urban drainage modelling is applicable to far more than just CSO reporting. It can be used in the planning, design and operation of sewer systems such as in capacity-analysis studies (Blumensaat *et al.* 2012). As such, these tools can be viewed as helping to provide the scientific support necessary for effective management. If it were not for the subcatchments and the input rainfall records, the model could not function and without the model, it would be harder to provide a scientific basis for upgrading a pipe. However, through modelling that is GIS-driven it is possible to simulate the surcharging that will occur if a pipe is not upgraded and hence provide management a more scientific basis as to why they should approve a pipe's upgrading. Furthermore, by attempting to measure CSO surcharge volumes with this same model, City staff can also more effectively attempt to minimize the impact of their CSOs on Lake Ontario by first identifying which locations surcharge the most and then attempting to develop a strategy to address them. Therefore, this illustrates how this research has also attempted to integrate both anchors of the Environmental Applied Science and Management program.

This research, has shown that while there have been approaches similar to the one developed here they all had shortcomings that were addressed. The new automated methods introduced have been intended to explain in a meticulous manner the source codes and ModelBuilder diagrams that facilitate new applications. Furthermore, all these new approaches have been evaluated in a manner that has identified their strengths and weaknesses and therefore allows potential users to decide whether they want to implement them or identify areas which they can improve upon. To improve upon the ModelBuilder tools that were built, reprogramming merely involves dragging and dropping functions from the ArcGIS toolbox and therefore also serves as an educational opportunity for City staff.

It was also determined that rainfall interpolation while perhaps a necessity should not be implemented until the discrepancy between the city's rain gauges and independently available data, such as from the TRCA's rain gauges, has been resolved, as precipitation is the crucial determinant of whether or not CSO overflows or flooding will occur. Thus, if the rainfall

measurements are not accurate it puts into question everything that is based on it, namely the two models' results and the interpolated results (virtual gauges) generated from the City's rainfall measurements.

## **9.2 Future Research**

As this research had an overall focus on GIS-based automation as it relates to input parameters for 1-D distributed urban drainage model, there is much potential for future areas of research. Although this research attempted to be comprehensive by focusing on the main components of an urban drainage model, which are its subcatchments, and rainfall input data, this is such an important field that there is always room for further research. Thus, the suggested areas of future research should not be considered an exhaustive list of all that remains.

### **9.2.1 Automated DWF Subcatchment Delineation**

In terms of DWF subcatchment delineation, while the tool introduced in this research attempts to accurately identify the assumed frontage of each parcel and then locate the nearest applicable pipe to that location, there remains room for further optimization. However, this raises questions as to whether this assumption is always true. Therefore, it would be beneficial research to investigate how often the assumption for the frontage node is truly correct. Such research while not easy, could be undertaken by monitoring locations where maintenance crews were currently working on sanitary pipes and therefore provide an opportunity to do some verification. If this was not possible, a researcher could resort to inspection of technical as-built drawings or CCTV pipe data to determine how often this assumption is correct.

Another area of research that needs to be addressed is how to better position the frontage node especially at parcels located at or near an intersection. As can be seen from Chapter 6, uncertainty in their placement at intersections accounts for a notable proportion of the errors detected. The only resolution of this may depend on the field-evidence and as-built diagrams suggested above.

Yet another important area related to the automated delineation of DWF subcatchments involves the development of a tool that can be used to screen or categorize those subcatchments delineated by the tools presented here. Such a tool would be particularly useful because a modeller would only have to manually adjust those flagged by the tool as "Requiring manual

intervention". This tool could also have a subroutine to walk the modeller through options on how to improve such subcatchments resulting from the DWF tools presented. Consequently, this would significantly improve the time required to produce a polished and complete DWF subcatchment input dataset.

Another goal of future research relating to DWF parcels would be better representation of parcels that are multiuse, because for applications such as this, the land-use classification scheme was simplistic compared to the variations of reality. There are numerous buildings, particularly in downtown areas that are residential or even institutional but are also commercial. However in urban drainage modelling, they can only be described by a single land use. One example that readily comes to mind is the Ted Rogers School of Management located at 55 Dundas St West, Toronto, ON M5G 2C5. Not only is this building the home of Ryerson University's business school but it also houses several diverse commercial outlets. This is important because once it has been determined that the DWF subcatchment boundaries are correct, the next step in the process is assigning wastewater flow values for each parcel that comprises the subcatchment and this is highly dependent upon the type of land use selected for a specific parcel.

### **9.2.2 Automated WWF Subcatchment Delineation**

In terms of future areas of research for WWF as it relates to the City of Toronto it would be ideal if a DEM derived from a finer-resolution (e.g. LIDAR) dataset could be secured and used as input for the tools presented here, in order to compare to the results generated by the current DEM. As computational capabilities increase this dataset could also be used to potentially explore whether higher spatial resolution resolves the current limitation of the tool of not being able to correctly represent subcatchments of pipes less than 20 centimetres apart from one another. In addition, while the D8 algorithm was chosen primarily for its computational efficiency because of the large study area, future studies upon smaller areas could potentially be conducted relying upon more innovative algorithms, such as the Triangular Form-based Multiple Flow (TFM) algorithm suggested by Pilesjö & Hasan (2013). Furthermore, to increase its usefulness, research also needs to be devoted to developing an automated approach to validate automatically delineated WWF subcatchments by the tool as well.

### **9.2.3 Automated Rainfall Interpolation**

Further research needs to be done to properly evaluate the quality of data being recorded or reported as being from the City's rain gauges in order to determine their accuracy. One possible way this could be determined is by creating a smaller model that only included the subcatchments that contributed to one of the City's CSO outflow monitoring locations. In doing so, it would be possible to run the simulation for a single month using rain gauge measurements from the City's gauges and then another simulation run in which the measurements came from the TRCA's gauges or other independent observations. Afterwards, a comparison could be conducted evaluating which of the rain gauge datasets resulted in CSO overflows that most closely coincided with those reported by the City's CSO monitoring devices assuming that the issues with these devices was resolved (H. Nguyen, personal communication, February 25, 2014).

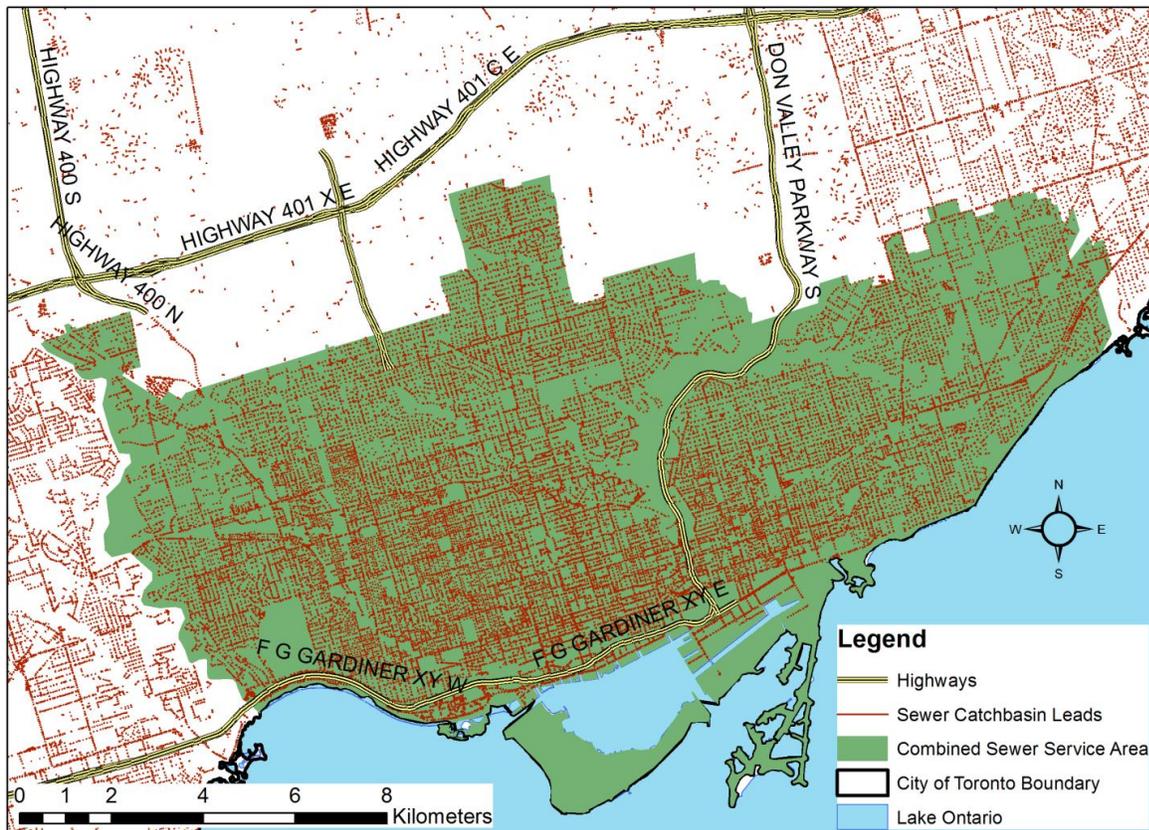
### **9.2.4 General Research**

While the areas attempted to be automated in this research were undoubtedly important there still remains other areas that could not be addressed but are nonetheless important in improving the efficiency with which models are being developed and their overall accuracy. Among these areas is developing automated approaches to determining land cover, which is used in assigning imperviousness values to WWF subcatchments once they have been delineated. Currently, the approach employed by the City through its consultants relies upon a modeller visually inspecting high-resolution orthoimagery from the City or Google Earth satellite imagery and arbitrarily arriving at imperviousness based on what is seen. However, this approach just like the manual delineation methods is very subjective, so in order to maintain consistency and minimize bias in the results only one person from the consultancy is assigned to this task (C. Fan, personal communication, August 6th, 2013). As such, these issues could be eliminated if one were to develop an automated approach to quickly and accurately determine land cover. Fortunately, such research has already been conducted ranging from the use of orthophotos to hyper-spectral satellite imagery (Fankhauser, 1999; Berezowski *et al.* 2012; Demarchi *et al.* 2012). All that remains is to investigate whether any of these options would be viable for the City of Toronto, while bringing it one step closer to being on the cutting edge of urban drainage modelling.

### 9.3 Recommendations

The recommendations provided here are intended for not only the City of Toronto but also others who might be interested in employing any of the methods presented here, either in full or part. Since automated subcatchment delineation is still a relatively new area, it is likely that modellers are still going to resort to manual methods in some cases, so it is important to employ topology rules when manually delineating subcatchments. Furthermore, topology error checks should be completed before inputting geographic data into models, since it does not seem as if this is always practiced. Seemingly insignificant areas of overlap and slivers may not be visible at first glance, but they can add up to inaccurate drainage areas that may also introduce uncertainties to actual field conditions.

It is not advisable for the City of Toronto to attempt to use the WWF tool in the more northern regions of Toronto where there are clearly many catch basins missing from the TWAG data, hence why there are so few pipes being selected (Figure 76). If applications are to be extended to data-sparse areas it is advised to do so only once it has been determined which pipes are appropriate. Then it would not matter whether or not the catch basins and catch basin leads layers in TWAG were up to date. Furthermore, in order to minimize errors in the selection of pipes it is also advisable to not only enforce topology so that each catch basin is snapped to its appropriate pipe, but to also add an attribute field for each catch-basin feature-point that logs the connected pipe. Additionally, such level of detail and accuracy will be necessary in order to successfully run 2-D distributed urban drainage models, which is the long-term goal of many users of such models including the City of Toronto (C.W. Baxter, personal communication, July 18, 2013), as computers become increasingly more powerful.



**Figure 76. The limits of TWAG in accounting for catch basins**

Although the Toronto Centre Line (TCL) layer was used in part to address the lack of a culverts layer, it is recommended that the City create a new layer that contains the location of all such structures. This would ensure that none of the culvert locations were missed during the DEM reconditioning process. In addition, the creation of a polygon layer that focuses exclusively on the road would also be recommended since this would further improve attempts to accurately replicate the geometry of the road within the DEM during the reconditioning process. Furthermore, in order to maintain high standards of data quality it is recommended that all datasets created or acquired by the City in the future come with appropriate metadata.

While the methods presented here are all automated to a certain degree they still require time to initiate and so it best to reserve them for times when one wants to create a model for a relatively large area. In the case of creating a detailed model such as for a single block, it would be quicker to delineate it manually, though being sure to follow the other recommendations made in this regard such as applying topology rules or using ArcGIS' topology inspector tool before loading them into the model.

Another recommendation relates to the type of computer that should be used in order to successfully implement these automated approaches, especially the WWF subcatchment delineation process because it is extremely computer intensive. Hence, the current state of technology suggests a computer with at least an Intel i7 quad core processor, 32 GB of DDR2 RAM, a solid-state drive and a large main hard drive (3 TB). A 3 TB hard drive is recommended because once the ModelBuilder-based tool starts, the computer begins to run out of space if its main hard drive is not large enough and therefore fails to fully execute. As the program is running, it can generate intermediary files that are over 100 gigabytes in size. In this research, the main hard drive was not large enough to accommodate such files so an external drive had to be purchased to allow for the running of the tools though this decreased the processing speed and therefore is not recommended. In addition, the InfoWorks CS modelling results files can also be several hundred gigabytes and so this would ensure that the tools and model could be run on the same machine.

A final recommendation that is applicable and perhaps the most important to all the tools presented here, is ensuring data quality. The data quality of input datasets is especially vital and why it has been repeatedly mentioned throughout this research. This is because all tools will generate results even when their input datasets are not entirely accurate. However, if the input datasets contain numerous potential inaccuracies, the user can have no confidence in the results generated by the tools, diminishing its true value. While it is recognised that no geospatial dataset is a perfect representation of reality, by minimizing the number of potential errors in the input datasets the number of errors generated in the output can be proportionally reduced. In addition, by ensuring the data quality of the input datasets, it makes automation a more viable option as less time will be spent addressing errors that could have been avoided by ensuring the data quality of the input datasets. This way more time can be spent on actually reviewing the results of the model and using this as a scientific basis for making the best management decisions. Thus, if jurisdictions such as Toronto want to truly benefit from automated or semi-automated solutions, addressing the potential inaccuracies in their input datasets should be one of their top priorities.

## Appendix

### A.1 Source Code for Adjusting XML Configuration File to Create Pseudo Sewer Laterals for Each Parcel

In order to allow future readers of this document to easily implement this method on their own the actual configuration changes are conveniently highlighted in yellow below:

```
<!-- Begin Configuration Section for the Add Laterals Tool and Construction Tools, This is an xml array, so you can define any number of AddLateralDetails Entities-->
```

```
<AddLateralsLayers>
  <!-- Start of the Entry, the name is just for reference, not used or presented-->
  <AddLateralDetails Name="Addresses to Mains">
    <!-- The Point layer to connect to the main through a lateral-->
    <Point_LayerName>Frontage_Nodes</Point_LayerName>
    <!-- the layer to look to connect to, the lateral will connect to this from the point-->
    <MainLine_LayerName>Pipes</MainLine_LayerName>
    <!-- The line used to connect the point to the main-->
    <LateralLine_LayerName>Sewer_Lateral_Lines</LateralLine_LayerName>
    <!-- The Template to use to populate the attributes of the lateral line, this can be removed or left blank and the user will be prompt for a template-->
    <LateralLine_EditTemplate>1" Copper Domestic</LateralLine_EditTemplate>
      <FromToFields>
        <!--XML Array of From/To Fields to copy from the Main to the Point-->
        <FromToField>
          <!-- The field in main that contains the attribute you want to apply to the point, this can be left blank or removed-->
          <SourceField>asset_id</SourceField>
          <!-- A field in the point layer that can be populated with a value from the main this can be removed or left blank-->
          <TargetField>asset_id</TargetField>
          <!-- A prefix that can be applied to the value extracted from the main and applied to the point this can be left blank -->
          <Prefix></Prefix>
        </FromToField>
        <!-- Repeat-->
        <FromToField>
          <SourceField>us_node_id</SourceField>
          <TargetField>us_node_id</TargetField>
          <Prefix></Prefix>
        </FromToField>
      </FromToFields>
    </AddLateralDetails>
  </AddLateralsLayers>
```

```

    <TargetField>system_typ</TargetField>
    <Prefix></Prefix>
  </FromToField>
<FromToField>
  <SourceField>ds_node_id</SourceField>
  <TargetField>ds_node_id</TargetField>
  <Prefix></Prefix>
</FromToField>
</FromToFields>
  <!-- Determines the direction to draw the main, flow is set with digitized direction
  this affects the distance set in the point along sections below,
  that is from the start of the main-->
  <LateralLine_StartAtMain>false</LateralLine_StartAtMain>
  <!-- this will check for an existing lateral between the point and the main
  and remove it if one is found, set to false to leave an existing lateral-->
  <DeleteExistingLines>true</DeleteExistingLines>
  <!-- the tolerance to search for the lateral from the point feature-->
  <ToleranceForDelete>.5</ToleranceForDelete>
  <!--restrict searches by layer definition-->
  <SearchOnLayer>true</SearchOnLayer>
  <!--The distance to search for the closest line from the point-- Units is in metres>
  <SearchDistance>500</SearchDistance>
  <!-- This section allows you to create a series of points along the main,
  it can be removed-->
  <PointsAlong>
    <!-- The entry for one point-->
    <PointAlong>
      <!-- The name of the layer to place, this must match the
      layer name in the TOC-->
      <LayerName>Water Curb Stop Valves</LayerName>
      <!-- The distance to place along the lateral, percent or feature units-->
      <Distance>5</Distance>
      <!-- Determines if the distance above is percent or feature units-->
      <DistanceIsPercent>false</DistanceIsPercent>
      <!-- The editor template used to fill in the attributes, this
      can be removed or left blank and the user will be prompted for a template-->
      <EditTemplate>Roundway</EditTemplate>
      <!--Option to intersect a polygon layer and offset the point from the intersection point on
      the polygon boundary, -->
      <PolygonOffsetLayerName>OwnerParcel</PolygonOffsetLayerName>
      <!--To or the From Side of the intersection, the digitized direction of the lateral matters:
      Options - To or From-->
      <PolygonOffsetSide>From</PolygonOffsetSide>
      <!-- ends this layers configuration-->
    </PointAlong>
    <!-- start of next layer, you can copy and repeat these sections for any

```

```

number of point layers-->
<PointAlong>
  <LayerName>Water Fittings</LayerName>
  <EditTemplate>Tee</EditTemplate>
  <Distance>0</Distance>
  <DistanceIsPercent>>true</DistanceIsPercent>
</PointAlong>
<!--Repeat for additional Points-->
</PointsAlong>
<!-- This option will connect two points to the main through a
single lateral if the features are within a tolerance-->
<Dual_When_Two_Selected>>false</Dual_When_Two_Selected>
<!-- This will dual nearby meters even when not selected -->
<Dual_When_Nearby>>false</Dual_When_Nearby>
<!-- the distance to search to create a dual lateral with selected features-->
Dual_Max_Distance_When_Two_Selected>100</Dual_Max_Distance_When_Two_Selected>
<!-- the distance to search to create a dual lateral with nearby features-->
<Dual_Max_Distance_When_Nearby>5</Dual_Max_Distance_When_Nearby>
<!-- Determines how to draw the dual laterals, square or a Y shape-->
<Dual_Option_Make_Square>>false</Dual_Option_Make_Square>
<!-- the distance on the lateral to turn 45 degrees on the main-->
<Hook_DoglegDistance>0</Hook_DoglegDistance>
<!-- Determines if the dogleg is a distance or a percent down the lateral-->
<Hook_DistanceIsPercent>>true</Hook_DistanceIsPercent>
<!-- The angle of the dogleg-->
<Hook_Angle>45</Hook_Angle>
<!-- Option to reset flow after edit - Digitized, Role, None-->
<Reset_Flow>Digitized</Reset_Flow>
<!-- End the config for one point to main with lateral config-->
</AddLateralDetails>

```

### **An Important Note:**

Once the toolbars were installed and the configuration file was adjusted according to the instructions above, artificial or pseudo laterals that would connect each frontage node that is representative of the front side of each parcel to its closest pipe could be created. In order to do so, it was necessary to enter the editing mode and then select up to 5,000 frontage nodes at a time, followed by pressing the "Add a Lateral" from the infrastructure editing toolbar in order for the program to run. Only up to 5,000 nodes were selected at a time, because it was discovered that the laterals were being written to the random-access memory (RAM) of the computer and beyond 5,000 laterals a 32-bit program on even a high end 64-bit system as was used in this study begins to run out of memory and so the process becomes extremely slow.

## A.2 Creating a Geometric Network

In order to create a geometric network it was necessary to right click on the resulting dataset output by the first part of the tool known as the Frontage Node Generator and then click on "New" and then select "Geometric Network" as seen in Figure A1. After that point, it was a simple matter of selecting the right parameters outlined in the flow diagram as seen in Figure A2 in order to successfully create the geometric network. Geometric networks are special networks in ArcGIS that are sets of connected edges or lines and junctions or points that are used to model infrastructure from the real world such as a sewer network in a GIS environment (ESRI, 2010). As such, it was the appropriate type of network dataset to be applied in this case.

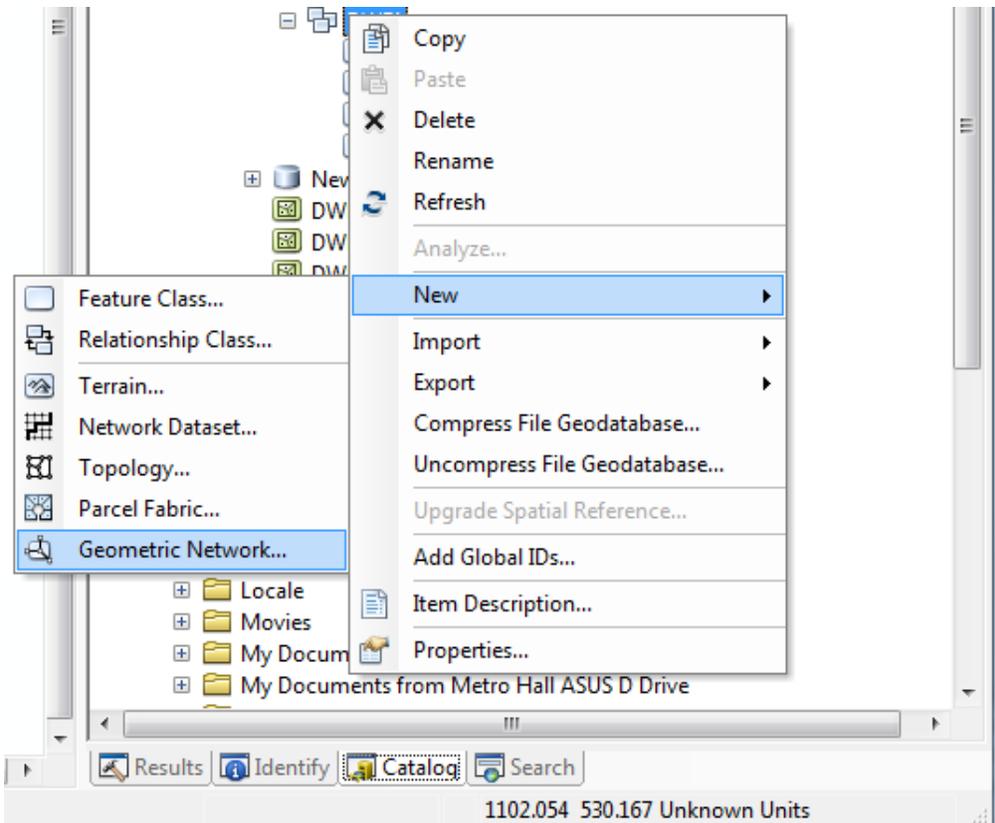
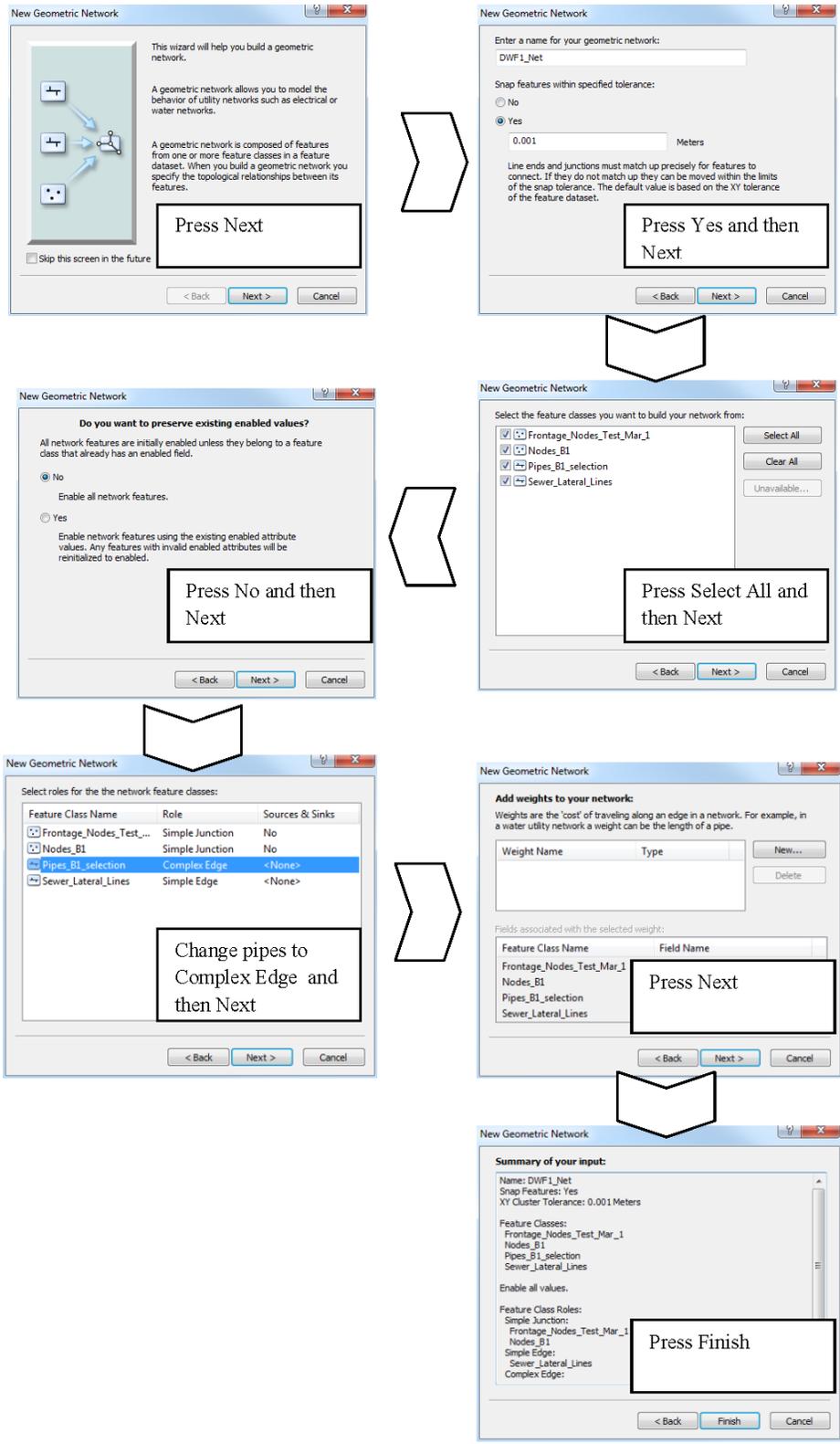


Figure A1. How to initiate the geometric network creation process



**Figure A2. Flow diagram outlining steps required to created geomeric network via screenshots**

### A.3 ModelBuilder Diagrams of Selected Tools Developed

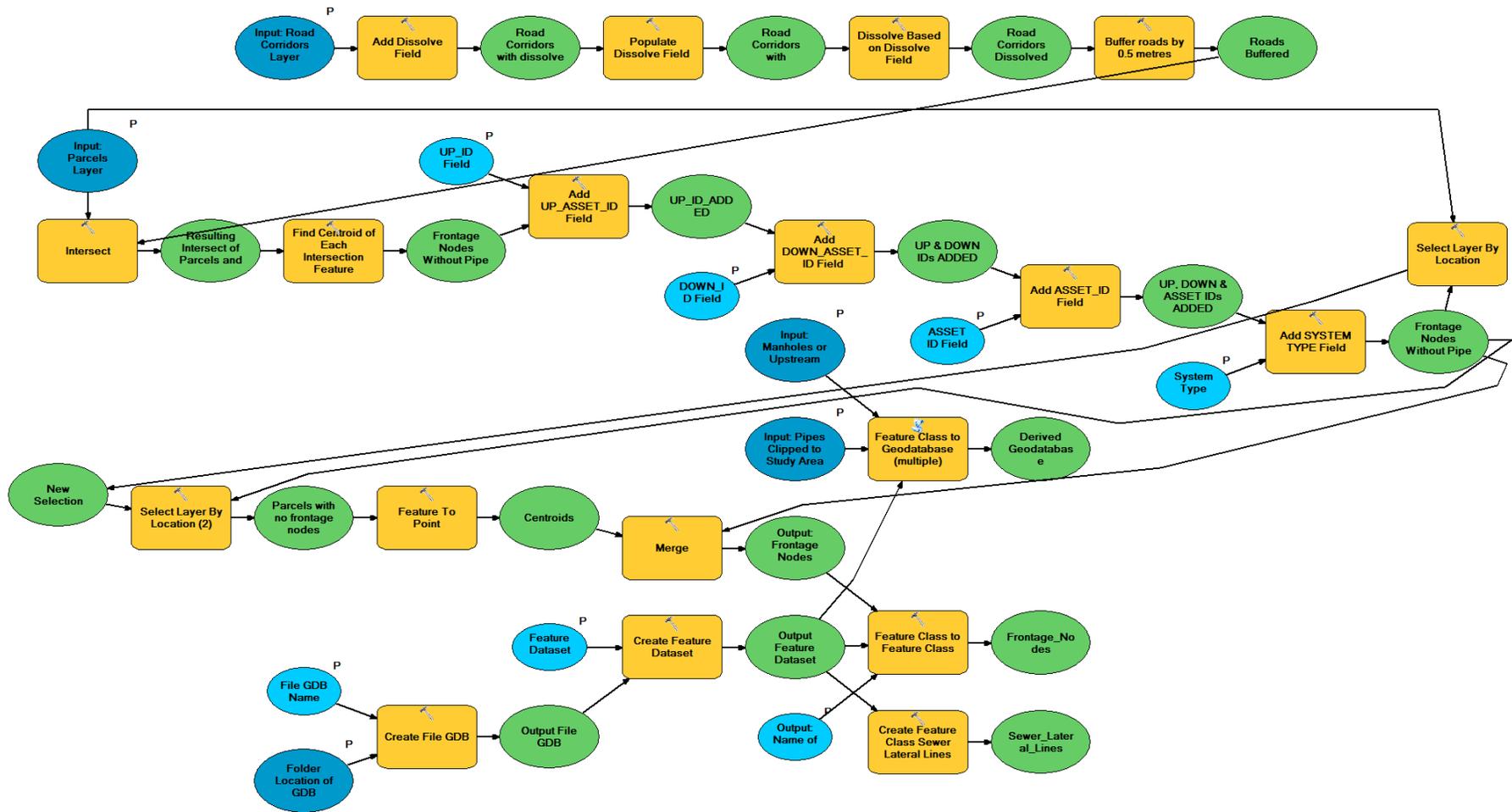


Figure A3. ModelBuilder diagram of the Frontage Node Generator Tool

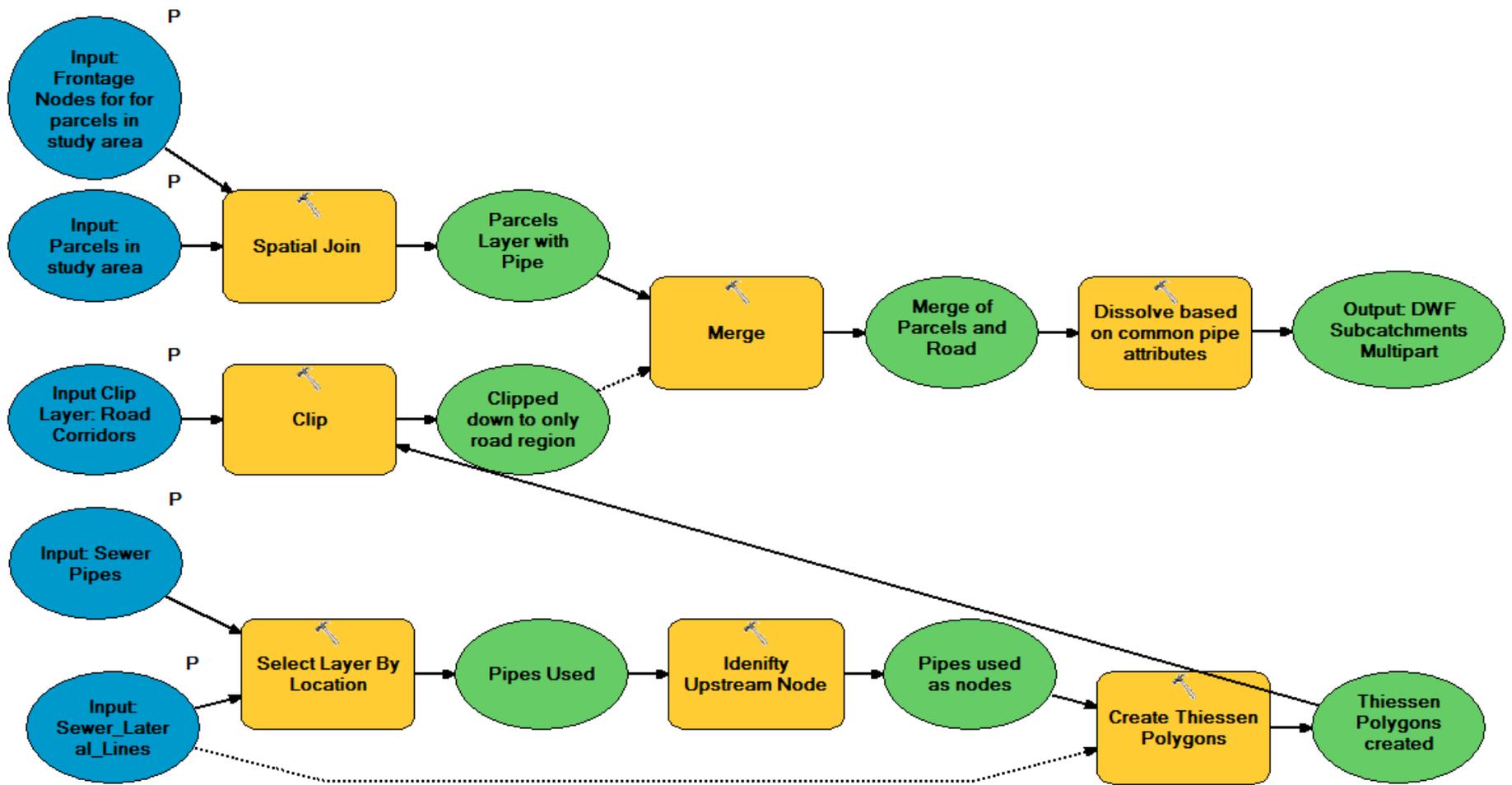


Figure A4. ModelBuilder diagram of the DWF Subcatchment Generator Tool

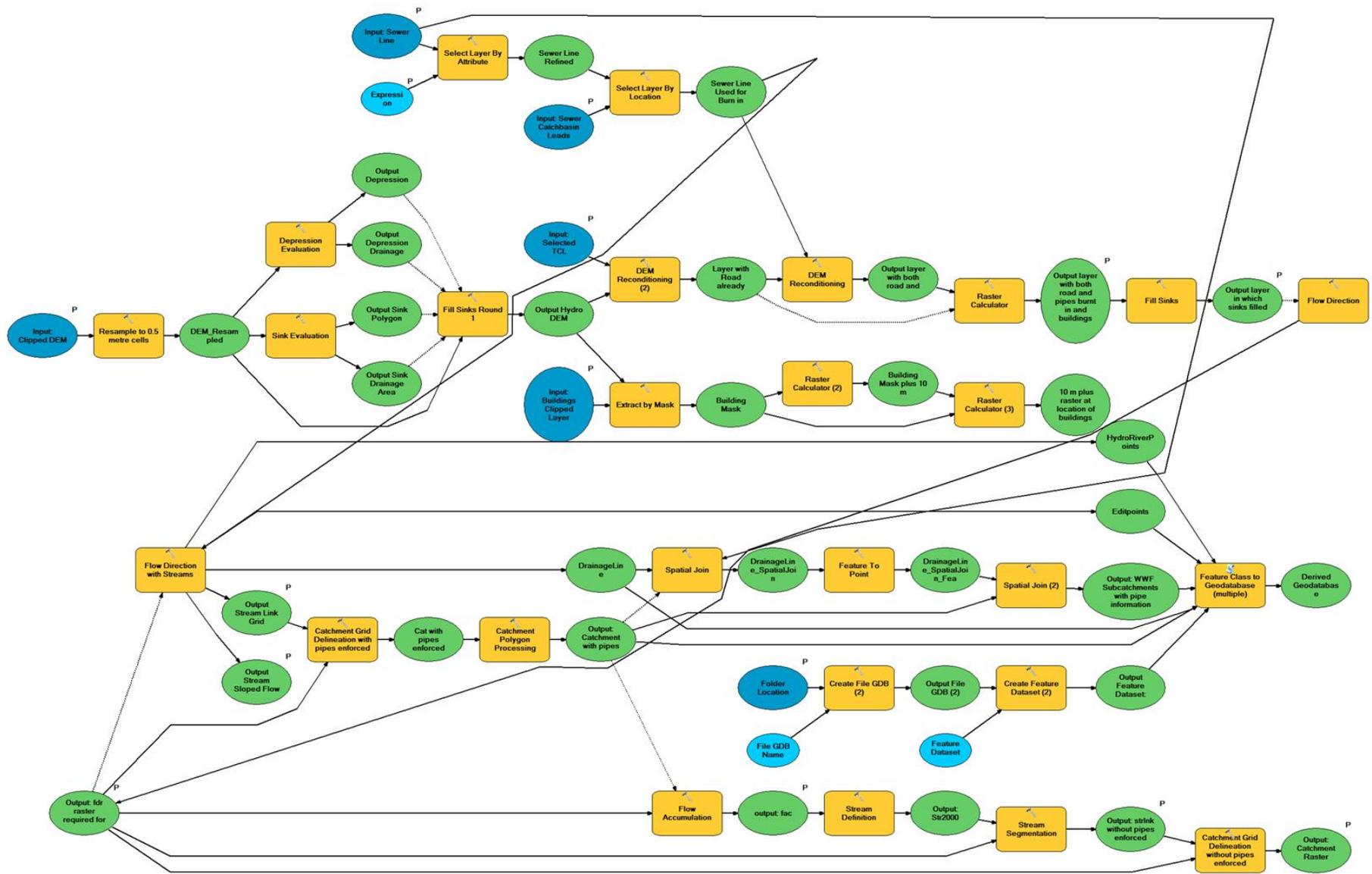


Figure A5. ModelBuilder diagram of the WWF Subcatchment Delineation Tool

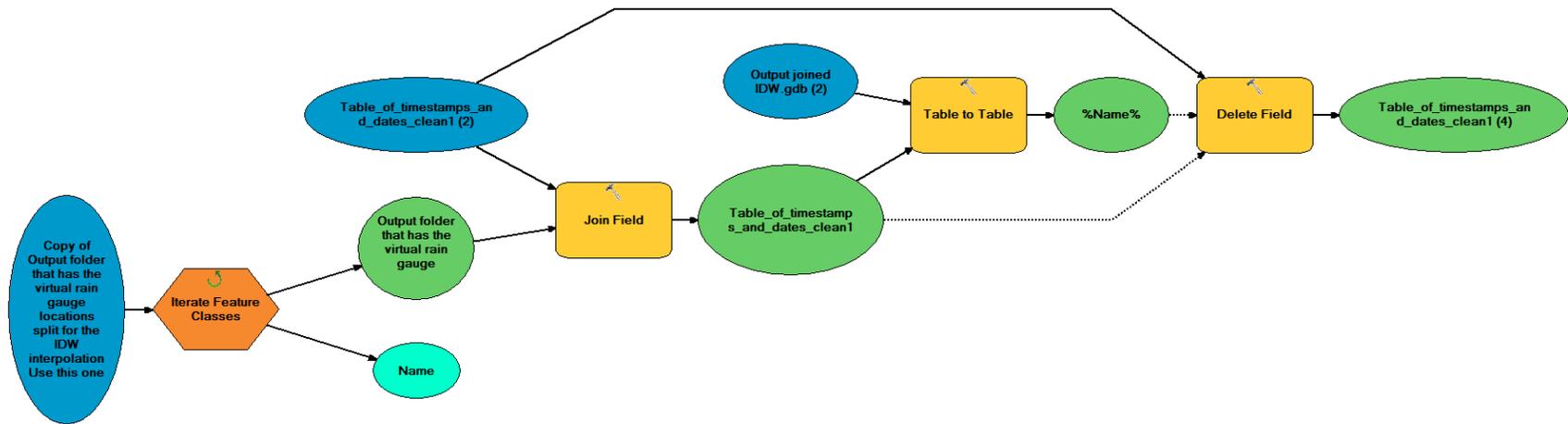


Figure A6. Virtual Rain Gauge Table Generator Tool

## A.4 Results of Interpolation Methods Testing

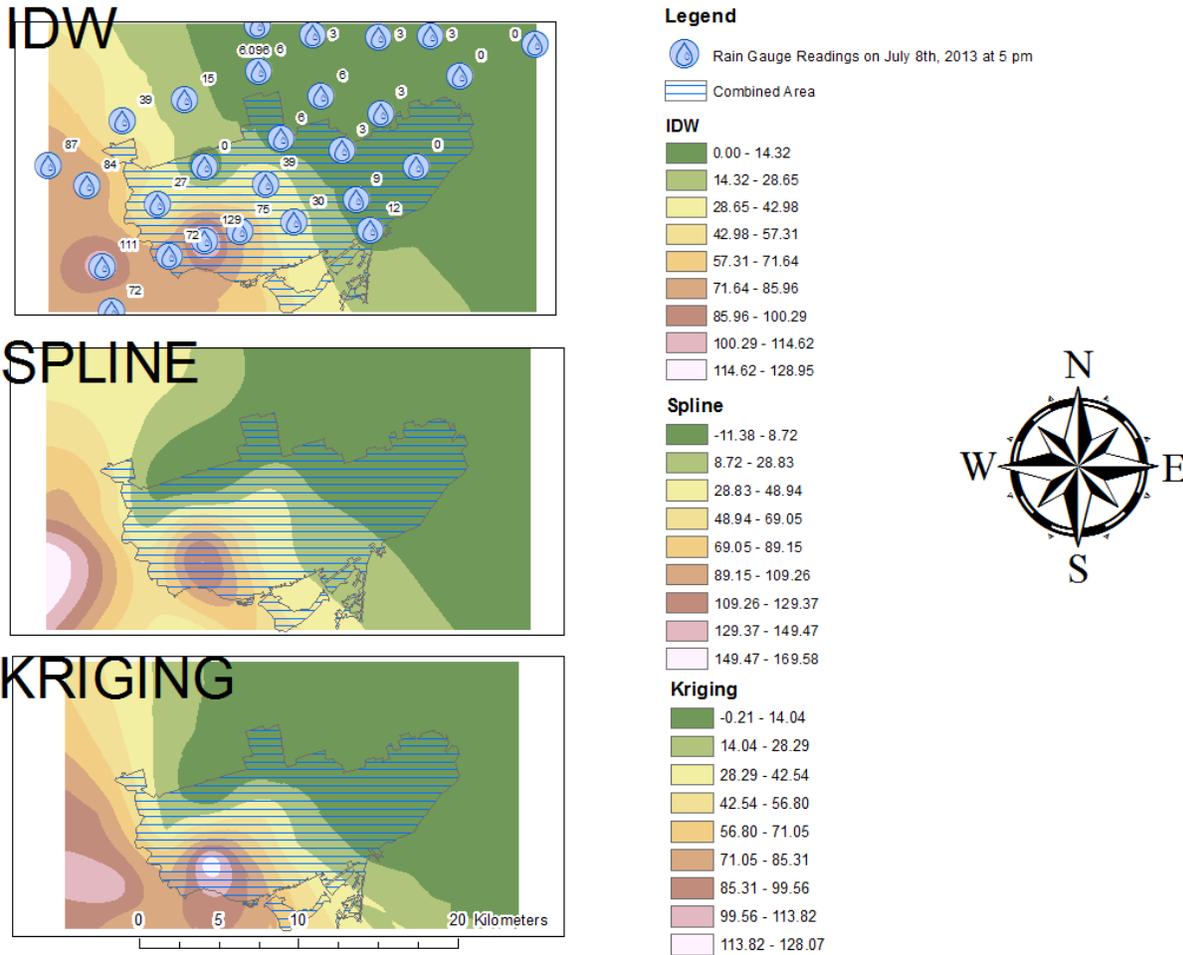
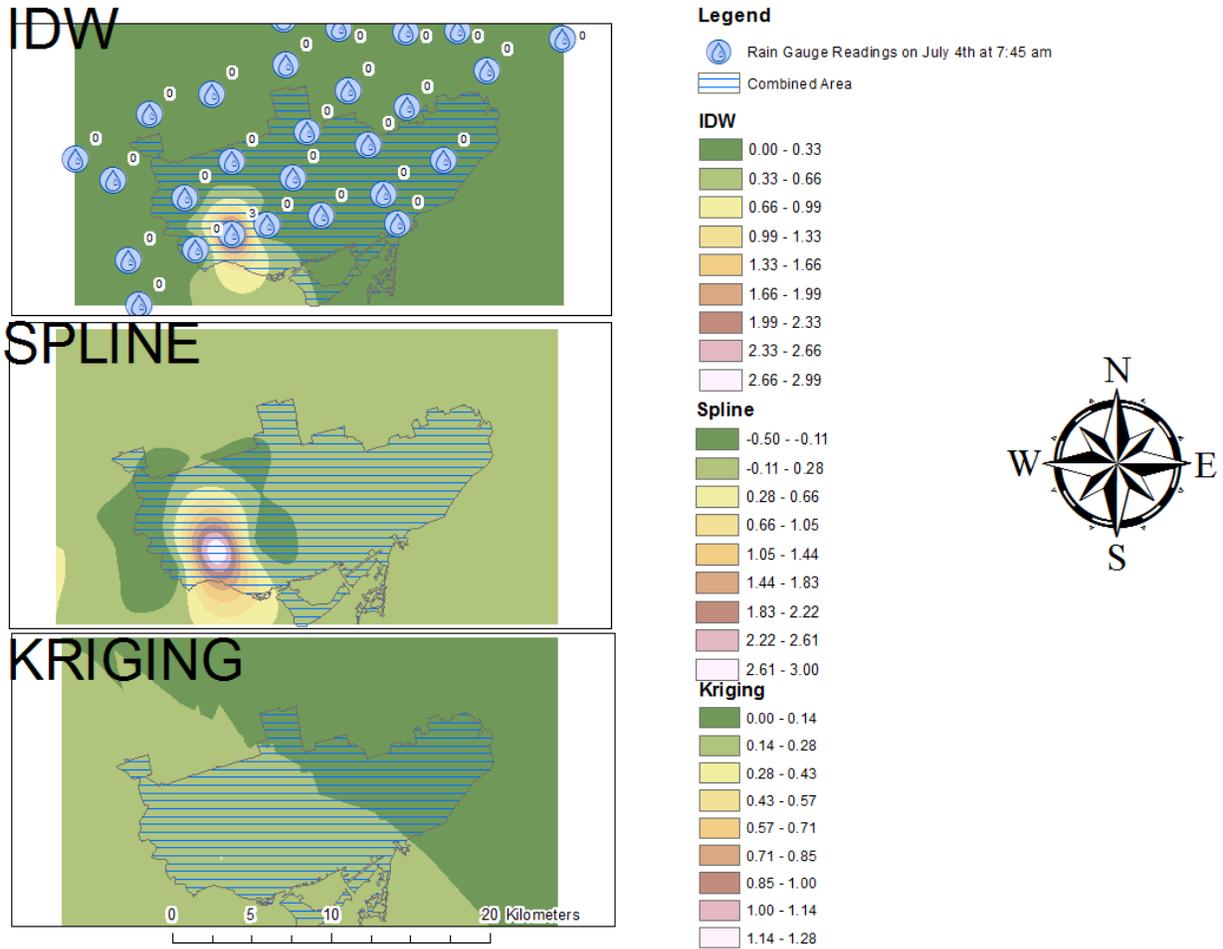
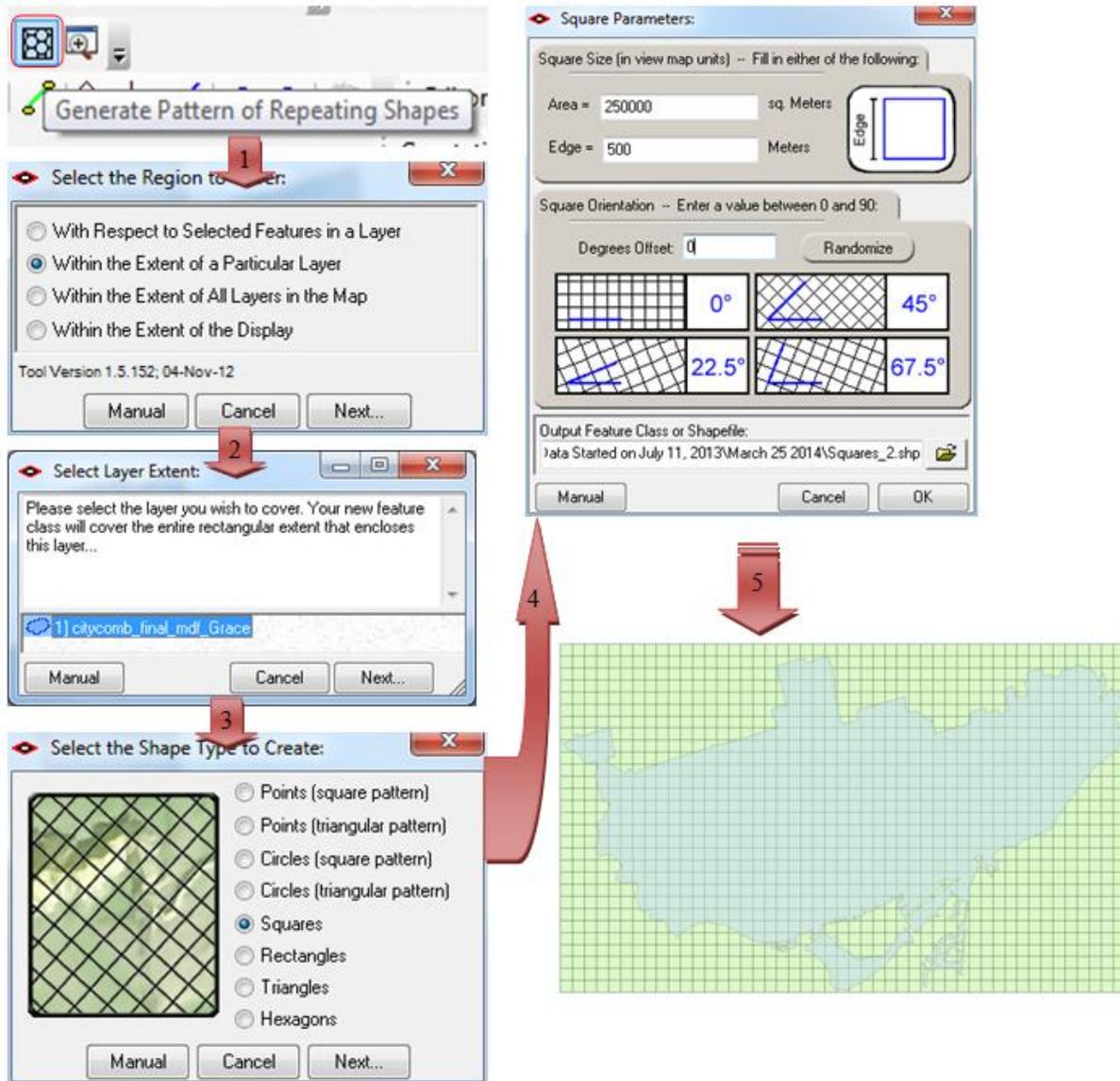


Figure A7. Sensitivity testing results using a time step when most of the rain gauges registered a value above zero



**Figure A8. Sensitivity testing results using a time step when most of the rain gauges registered a value of zero**

## A.5 Steps for Creating Virtual Rain Gauges Using the Repeating Shapes Tool



**Figure A9. Flow chart of parameters that need to be selected at each step in order to generate the virtual rain gauge grid used to derive the virtual rain gauge locations**

## A.6 Excel Functions Used to Eliminate Non-Numeric Characters from a Field

In order to associate each interpolated value to its correct time step through joining it to the name of its raster (RasterName) it is necessary to eliminate the TS before each raster's name. This can be achieved in Excel through a combination of functions. These functions are the

RIGHT function which "Returns the rightmost characters from a text value" and the VALUE function which "Converts a text argument to a number" (Microsoft, 2007). As can be seen from Figure A10 by using these two functions it was possible to eliminate the string portion or file name path portion of the RasterName field to isolate the actual timestep number (TS) so that it was possible to properly associate each interpolation result or raster layer to its correct timestep.

	A	B	C	D
1	OID	RasterName	TS	TS
2	1	E:\Rainfall Data Started on July 11, 2013\April 9 2014\Rasters.gdb\TS0869	0869	869
3	2	E:\Rainfall Data Started on July 11, 2013\April 9 2014\Rasters.gdb\TS0870	0870	870
4	3	E:\Rainfall Data Started on July 11, 2013\April 9 2014\Rasters.gdb\TS0868	0868	868
5	4	E:\Rainfall Data Started on July 11, 2013\April 9 2014\Rasters.gdb\TS0867	0867	867
6	5	E:\Rainfall Data Started on July 11, 2013\April 9 2014\Rasters.gdb\TS0864	0864	864
7	6	E:\Rainfall Data Started on July 11, 2013\April 9 2014\Rasters.gdb\TS0174	0174	174
8	7	E:\Rainfall Data Started on July 11, 2013\April 9 2014\Rasters.gdb\TS0105	0105	105
9	8	E:\Rainfall Data Started on July 11, 2013\April 9 2014\Rasters.gdb\TS0104	0104	104
10	9	E:\Rainfall Data Started on July 11, 2013\April 9 2014\Rasters.gdb\TS0089	0089	89
11	10	E:\Rainfall Data Started on July 11, 2013\April 9 2014\Rasters.gdb\TS0871	0871	871
12	11	E:\Rainfall Data Started on July 11, 2013\April 9 2014\Rasters.gdb\TS0872	0872	872
13	12	E:\Rainfall Data Started on July 11, 2013\April 9 2014\Rasters.gdb\TS1018	1018	1018
14	13	E:\Rainfall Data Started on July 11, 2013\April 9 2014\Rasters.gdb\TS1251	1251	1251

**Figure A10. Editing the names in Excel to isolate the TS number**

### A.7 The 93 DWF Subcatchments Visually Inspected

FID	OBJECTID	us_node_id	system_typ	asset_id	Shape_Area	Acceptable
0	32	4193822028	COMB	0000-035-9101	20734.21497	N
1	33	4079821310	COMB	0000-036-2791	4709.316812	N
3	108	4161621419	COMB	1000-035-8031	2754.509631	N
8	417	3943920252	COMB	3003197	5206.490387	N
9	462	3922018532	COMB	3003341	8478.829562	N
11	477	3950219162	COMB	3003398	6628.015784	N
12	513	3903017213	COMB	3003531	3163.13469	N
14	566	3885817429	COMB	3003624	7454.634078	N
15	572	3927117758	COMB	3003655	8566.114769	N
16	592	3914817975	COMB	3003697	9960.03013	N
17	618	3937617209	COMB	3003730	7790.439412	N
23	833	3795822411	COMB	4000-009-8231	26009.19337	N
25	842	4124521096	COMB	4000-035-7341	22838.25108	N
28	1002	4153921212	COMB	6000-035-7181	8198.368223	N

29	1014	4128521892	COMB	6000-035-8981	4298.493717	N
30	1044	3932422128	COMB	6000-042-2521	1476.446748	N
31	1112	4060222047	COMB	7000-036-8351	61739.88568	N
39	1228	MH3895320867	COMB	SL3003058	3128.770864	N
42	1277	4265324089	SAN	0000-037-1251	6786.838073	N
51	1532	4158423826	SAN	3000-038-9191	663525.7816	N
55	1823	3932019237	SAN	3003396	5054.848341	N
56	1902	3929718032	SAN	3003691	6464.134077	N
58	1954	CN4472	SAN	3003839	6521.990679	N
59	2055	3820417811	SAN	3004072	4624.458771	N
61	2084	3801016334	SAN	3004125	7384.809173	N
65	2150	3803416842	SAN	3004217	4254.550557	N
66	2153	4024017027	SAN	3004227	13010.48246	N
72	2303	4136020390	SAN	3005202	21406.43964	N
81	2509	CN2963	SAN	3011102	9110.742501	N
85	2661	4086223492	SAN	6000-010-8961	5538.401493	N
88	2783	4131723999	SAN	7000-038-9682	10320.61167	N
90	2808	4125122576	SAN	8000-035-6211	63814.06265	N
2	98	3870421993	COMB	1000-010-1351	6255.872871	Y
4	143	3977721770	COMB	1000-039-7361	2754.840802	Y
5	223	4230620953	COMB	2000-038-2971	5175.446756	Y
6	241	3879922187	COMB	3000-009-5251	2790.421653	Y
7	401	3873619988	COMB	3003176	8861.664806	Y
10	474	3944219270	COMB	3003392	8494.201349	Y
13	529	3886017013	COMB	3003547	4938.769706	Y
18	652	3859920131	COMB	3003828	9798.717936	Y
19	700	3833517138	COMB	3003964	6502.393242	Y
20	708	3860717357	COMB	3003991	11285.73531	Y
21	720	3864517498	COMB	3004026	4375.226379	Y
22	734	3855818123	COMB	3004086	5593.948593	Y
24	838	4197722154	COMB	4000-035-5631	32426.43537	Y
26	880	3902321832	COMB	4000-039-7571	13514.19375	Y
27	976	3985021520	COMB	5000-051-2331	61578.33889	Y
32	1146	3829522941	COMB	8000-009-6921	14206.05948	Y
33	1166	4124721416	COMB	8000-035-8191	2179.485716	Y
34	1184	4226421054	COMB	8000-037-0161	5214.105063	Y
35	1185	4209621242	COMB	8000-037-0431	26154.64257	Y
36	1204	3945422838	COMB	8000-042-4521	6328.329133	Y
37	1222	MH3886020916	COMB	SL3003047	5369.558805	Y
38	1226	MH3894121150	COMB	SL3003052	5615.386332	Y
40	1249	MH3923420955	COMB	SL3010243	8756.981134	Y

41	1273	4101622452	SAN	0000-036-8821	262263.2931	Y
43	1291	4226722938	SAN	0000-037-4581	4951.472706	Y
44	1327	4183322977	SAN	1000-035-6951	2590.844341	Y
45	1357	4282021985	SAN	1000-037-3151	25208.09302	Y
46	1388	4305824783	SAN	1000-040-2492	24400.99322	Y
47	1436	4280922781	SAN	2000-037-4241	33452.95918	Y
48	1449	4178524462	SAN	2000-038-6391	44874.17255	Y
49	1486	4292523003	SAN	3000-036-4261	5747.253771	Y
50	1523	4134223078	SAN	3000-038-8021	68123.30942	Y
52	1577	3915914666	SAN	3000806	5624.919509	Y
53	1673	4054621103	SAN	3002630	7983.21676	Y
54	1697	4008717487	SAN	3002830	41882.91281	Y
57	1904	3922818151	SAN	3003702	9742.13282	Y
60	2079	3814418173	SAN	3004106	2455.595438	Y
62	2106	3826116280	SAN	3004150	12101.36336	Y
63	2130	3852016610	SAN	3004190	7300.193327	Y
64	2137	3842416689	SAN	3004202	20321.74293	Y
67	2215	4057918860	SAN	3004699	13973.60213	Y
68	2253	4092719578	SAN	3004890	21572.19987	Y
69	2280	4073219920	SAN	3005066	12178.19921	Y
70	2299	4102020094	SAN	3005163	45809.03217	Y
71	2301	4048619903	SAN	3005180	7406.641989	Y
73	2306	4123820213	SAN	3005205	14361.45737	Y
74	2323	4007220783	SAN	3007181	4725.559106	Y
75	2335	4026221172	SAN	3007242	12295.83205	Y
76	2395	3960321087	SAN	3007664	7817.960066	Y
77	2400	3938521574	SAN	3007780	21204.10127	Y
78	2401	3940821625	SAN	3007783	164566.6554	Y
79	2422	4005620313	SAN	3008474	14345.21706	Y
80	2464	3979520423	SAN	3009726	3305.64656	Y
82	2513	4105822731	SAN	4000-035-6441	13732.73451	Y
83	2616	4274224550	SAN	5000-037-2561	6607.912217	Y
84	2631	4176524394	SAN	5000-038-6332	4412.950567	Y
86	2754	4271724028	SAN	7000-037-0962	37400.00473	Y
87	2778	4128523556	SAN	7000-038-8691	6001.1695	Y
89	2807	4153422301	SAN	8000-035-5941	93738.3019	Y
91	2819	4259823712	SAN	8000-036-5301	6198.816182	Y
92	2827	4070923047	SAN	8000-036-9801	6958.914137	Y

### A.8 The 96 WWF Subcatchments Visually Inspected

FID	us_node_id	system_typ	asset_id	Area_M2	Acceptable
0	MH3950021146	STM	SL3007697	877734.5	N
3	CN7304	storm	1469758	31152.5	N
8	4242923574-D	STM	6000-036-5371	1199904.25	N
9	4121119540	STM	3004925	885203	N
13	4022316583	STM	3001679	96647	N
15	4009320864	STM	3007201	490730.5	N
16	3946822279	COMB	3000-042-2761	11628.5	N
18	3927816877	STM	3005692	6216	N
25	3889922219	COMB	0000-009-5311	13114.5	N
27	3856022682	COMB	8000-009-6201	22691.75	N
29	3844622384	STM	0000-052-6321	3221.25	N
32	3858912640	storm	1412313	7240.75	N
33	3784509614	combined	53081	74982.75	N
34	3767114406	storm	1467268	573440.25	N
36	3721411861	storm	1468437	6425	N
40	3699708104	storm	1439977	14489	N
41	3696914469	storm	1469640	24889.5	N
42	3686511092	storm	1439755	21368.5	N
43	3680909419	storm	1439486	281.25	N
45	3638206666	combined	1438796	10103	N
46	3619614445	storm	1470028	1306	N
47	3621414544	combined	1470256	14543.75	N
48	3627806736	combined	1439986	239849.5	N
49	3615615102	storm	1469402	87295.75	N
51	3619310600	storm	1438348	9476.25	N
55	3661214827	storm	1470277	780598.75	N
59	3570912744	storm	1472187	13602.75	N
60	3560314706	storm	1474531	32099.75	N
61	3556511938	storm	1446390	5575.75	N
63	3533409059	combined	1444530	26431.5	N
64	3543116418	storm	1427535	1369	N
65	3539615388	combined	1471985	9571.5	N
67	3530911323	combined	1445630	289.5	N
68	3535916236	combined	1428447	29494.75	N
69	3519107159	storm	1443912	53099.25	N
70	3506806484	storm	1443725	2118.5	N
71	3488814455	storm	1475798	3201	N
75	3466415254	storm	1477142	40708	N
77	3431811725	storm	1450144	7377.75	N

80	3420808446	combined	1451603	9770.25	N
81	3404513759	storm	1478337	5374.75	N
82	3402414770	combined	1478641	15333.75	N
83	3401412519	storm	1477422	4488.25	N
85	5512532190	storm	4170130	123381.5	N
92	3330412590	combined	1480176	5165.25	N
95	3268809431	storm	1457366	53511	N
1	4189424227-D	STM	3000-038-6222	21764.75	Y
2	4149323183-D	STM	0000-038-8261	70791.5	Y
4	5512538935	STM	3000-038-9462	9351.5	Y
5	4299522161	STM	5000-037-3461	40328	Y
6	4266024360	STM	4000-037-1651	1233.5	Y
7	4135321660	STM	4000-052-9151	14023.75	Y
10	4063921076	STM	3002520	14119.5	Y
11	4060721285	COMB	6000-036-2851	26143	Y
12	4060322044	STM	2000-053-6791	42340.25	Y
14	4004516112	STM	3001583	40101.5	Y
17	3934319301	STM	3009381	5890.75	Y
19	3929519909	COMB	3003118	7366.75	Y
20	3924217681	COMB	3003610	13164.75	Y
21	3912519964	COMB	3003163	9187.5	Y
22	3907816764	STM	3005716	3799.25	Y
23	3905020105	STM	3009336	12717.5	Y
24	3892817257	STM	3004623	12132.25	Y
26	3861218353	STM	3007069	6224.5	Y
28	3853417472	STM	3006549	10858.5	Y
30	3921112930	storm	1412704	2770.5	Y
31	3902712822	storm	1411245	6820.5	Y
35	3756411397	storm	1467906	3477	Y
37	3713014866	storm	1466802	4518.5	Y
38	3692708788	storm	1438879	9297	Y
39	3699112577	storm	1470464	3234	Y
44	3638414694	storm	1470282	5499.25	Y
50	3614610313	combined	1440686	12073.5	Y
52	3606206619	storm	1439580	10225.5	Y
53	3605506451	storm	1439677	7752.25	Y
54	3580812451	combined	1473338	4925	Y
56	3580506600	storm	1443692	3171	Y
57	3574908681	storm	1445376	18574.5	Y
58	3560908936	combined	1447733	4665.5	Y
62	3549909850	storm	1443041	8283.25	Y

66	3536209701	storm	1444621	12195.75	Y
72	3490009888	combined	1451240	1815.5	Y
73	3483411941	combined	1450936	8603.75	Y
74	3468514476	combined	1477200	1889.5	Y
76	3439708716	storm	1448955	8737.75	Y
78	3428414326	storm	1479161	5571.25	Y
79	3420911791	storm	1450816	4641	Y
84	3382309977	storm	1455669	12113	Y
86	3379608794	storm	1453593	7799.25	Y
87	3378012328	storm	1481953	4766.75	Y
88	3377411918	storm	1454235	3154.5	Y
89	3361713310	storm	1480576	15299.75	Y
90	3369910730	combined	1456232	7405.5	Y
91	3362110699	storm	1453658	7001.5	Y
93	3314009854	storm	1453670	8312.5	Y
94	3312410372	storm	1453281	8020.5	Y

**A.9 Raw Data of Rainfall Measurements above Zero Millimetres for one of the two TRCA Rain Gauges Used to Evaluate the Validity of the Rainfall Interpolation Method Developed**

**Table 19. Comparing values from the TRCA HY003 gauge versus virtual gauge**

<b>Time Stamp</b>	<b>TS</b>	<b>Value from HY003 dataset provided by the TRCA ( in mm)</b>	<b>Value from Interpolation (Virtual Rain Gauge) (in mm)</b>
<b>01/07/2013 7:25</b>	89	0	0.00438552396
<b>01/07/2013 8:40</b>	104	0	0.01422704849
<b>01/07/2013 8:45</b>	105	0	0.00711352425
<b>01/07/2013 14:30</b>	174	0	0.00507506868
<b>04/07/2013 0:00</b>	864	0	0.01683998480
<b>04/07/2013 0:15</b>	867	0	0.22299593687
<b>04/07/2013 0:20</b>	868	0	0.19464789331
<b>04/07/2013 0:25</b>	869	0	0.05512521043
<b>04/07/2013 0:30</b>	870	0	0.01933929697
<b>04/07/2013 0:35</b>	871	0	0.01700472645

<b>04/07/2013 0:40</b>	872	0	0.01683998480
<b>04/07/2013 1:05</b>	877	0.2	0.00000000000
<b>04/07/2013 1:10</b>	878	1.8	0.00000000000
<b>04/07/2013 1:15</b>	879	1	0.00000000000
<b>04/07/2013 1:20</b>	880	1.2	0.00000000000
<b>04/07/2013 1:25</b>	881	1.6	0.00000000000
<b>04/07/2013 1:30</b>	882	0.2	0.00000000000
<b>04/07/2013 1:35</b>	883	0.4	0.00000000000
<b>04/07/2013 1:50</b>	886	0.2	0.00000000000
<b>04/07/2013 2:00</b>	888	0.4	0.00000000000
<b>04/07/2013 7:15</b>	951	0.2	0.00000000000
<b>04/07/2013 8:55</b>	971	0.4	0.00000000000
<b>04/07/2013 12:50</b>	1018	0	0.00272907037
<b>05/07/2013 7:45</b>	1245	0.4	0.00000000000
<b>05/07/2013 8:15</b>	1251	0	0.00431635370
<b>05/07/2013 8:40</b>	1256	0	0.00618668133
<b>05/07/2013 8:45</b>	1257	0	0.00766786514
<b>05/07/2013 8:50</b>	1258	0	0.00314550311
<b>05/07/2013 9:30</b>	1266	0.2	0.00000000000
<b>05/07/2013 9:35</b>	1267	0.4	0.00000000000
<b>05/07/2013 9:40</b>	1268	0.2	0.00000000000
<b>05/07/2013 11:25</b>	1289	0	0.00272907037
<b>05/07/2013 11:30</b>	1290	0	0.00711459387
<b>05/07/2013 11:35</b>	1291	0	0.00545814075

<b>05/07/2013 11:40</b>	1292	0	0.03059072420
<b>05/07/2013 11:45</b>	1293	0	0.05299423635
<b>05/07/2013 11:50</b>	1294	0	0.10873520374
<b>05/07/2013 11:55</b>	1295	0	0.09535672516
<b>05/07/2013 12:00</b>	1296	0	0.15231050551
<b>05/07/2013 12:05</b>	1297	0	0.17831355333
<b>05/07/2013 12:10</b>	1298	0	0.16781146824
<b>05/07/2013 12:15</b>	1299	0	0.17090378702
<b>05/07/2013 12:20</b>	1300	0	0.08362885565
<b>05/07/2013 12:25</b>	1301	0	0.02454626188
<b>05/07/2013 12:30</b>	1302	0	0.03789867088
<b>05/07/2013 12:35</b>	1303	0	0.06679838151
<b>05/07/2013 12:40</b>	1304	0	0.03470056131
<b>05/07/2013 12:45</b>	1305	0	0.01398098096
<b>05/07/2013 12:50</b>	1306	0	0.00831852946
<b>05/07/2013 12:55</b>	1307	0	0.00284575764
<b>05/07/2013 13:00</b>	1308	0	0.00462347083
<b>05/07/2013 13:10</b>	1310	0	0.03288712353
<b>05/07/2013 13:15</b>	1311	0	0.04455227777
<b>05/07/2013 13:25</b>	1313	0.4	0.05352083966
<b>05/07/2013 13:35</b>	1315	0	0.02535929717
<b>05/07/2013 13:40</b>	1316	0	0.00346438354
<b>05/07/2013 13:50</b>	1318	0.2	0.00000000000
<b>05/07/2013 14:10</b>	1322	0.4	0.00000000000

<b>05/07/2013 14:15</b>	1323	0.4	0.000000000000
<b>05/07/2013 14:20</b>	1324	0.4	0.000000000000
<b>05/07/2013 14:25</b>	1325	0.8	0.000000000000
<b>05/07/2013 14:30</b>	1326	0.2	0.000000000000
<b>05/07/2013 14:35</b>	1327	0.4	0.000000000000
<b>05/07/2013 14:40</b>	1328	0	0.00766786514
<b>05/07/2013 14:45</b>	1329	0	0.01198421884
<b>05/07/2013 15:30</b>	1338	0.2	0.000000000000
<b>05/07/2013 15:45</b>	1341	0.4	0.000000000000
<b>05/07/2013 15:50</b>	1342	0.2	0.000000000000
<b>05/07/2013 15:55</b>	1343	0.2	0.000000000000
<b>05/07/2013 16:00</b>	1344	0.4	0.000000000000
<b>05/07/2013 16:05</b>	1345	0.6	0.000000000000
<b>06/07/2013 10:45</b>	1569	0.2	0.000000000000
<b>07/07/2013 1:35</b>	1747	0	0.00205288990
<b>07/07/2013 1:50</b>	1750	0.2	0.000000000000
<b>07/07/2013 2:00</b>	1752	0.2	0.000000000000
<b>07/07/2013 2:10</b>	1754	0.2	0.000000000000
<b>07/07/2013 2:20</b>	1756	1	0.000000000000
<b>07/07/2013 2:25</b>	1757	0.2	0.000000000000
<b>07/07/2013 2:30</b>	1758	0.2	0.000000000000
<b>07/07/2013 2:40</b>	1760	0.4	0.000000000000
<b>07/07/2013 2:45</b>	1761	0.2	0.000000000000
<b>07/07/2013 2:50</b>	1762	0.2	0.000000000000

<b>07/07/2013 2:55</b>	1763	0.4	0.000000000000
<b>07/07/2013 3:00</b>	1764	0.2	0.000000000000
<b>07/07/2013 3:05</b>	1765	0.2	0.000000000000
<b>07/07/2013 3:10</b>	1766	0.4	0.000000000000
<b>07/07/2013 3:15</b>	1767	0.2	0.000000000000
<b>07/07/2013 3:20</b>	1768	0.2	0.000000000000
<b>07/07/2013 3:35</b>	1771	0.4	0.000000000000
<b>07/07/2013 11:35</b>	1867	0	0.00462347083
<b>07/07/2013 12:05</b>	1873	0	0.00877104793
<b>07/07/2013 12:15</b>	1875	0	0.01546579879
<b>07/07/2013 14:25</b>	1901	0	0.01546579879
<b>07/07/2013 14:50</b>	1906	0	0.01339165494
<b>07/07/2013 14:55</b>	1907	0.4	0.01528533082
<b>07/07/2013 15:00</b>	1908	0.2	0.00485709310
<b>07/07/2013 15:05</b>	1909	0	0.01850244403
<b>07/07/2013 15:10</b>	1910	0	0.05172131583
<b>07/07/2013 15:15</b>	1911	0.2	0.05543748662
<b>07/07/2013 15:20</b>	1912	0	0.07086832076
<b>07/07/2013 15:25</b>	1913	0	0.11455970258
<b>07/07/2013 15:30</b>	1914	0	0.10783246160
<b>07/07/2013 15:35</b>	1915	0	0.07329131663
<b>07/07/2013 15:40</b>	1916	0	0.08453444391
<b>07/07/2013 15:45</b>	1917	0	0.11656501144
<b>07/07/2013 15:50</b>	1918	0	0.10797354579

<b>07/07/2013 15:55</b>	1919	0	0.05466315150
<b>07/07/2013 16:00</b>	1920	0	0.04526037350
<b>07/07/2013 16:05</b>	1921	0	0.02273924835
<b>07/07/2013 16:10</b>	1922	0	0.06135925278
<b>07/07/2013 16:15</b>	1923	0	0.17471948266
<b>07/07/2013 16:20</b>	1924	0	0.50476533175
<b>07/07/2013 16:25</b>	1925	0	0.31802833080
<b>07/07/2013 16:30</b>	1926	0	0.35572513938
<b>07/07/2013 16:35</b>	1927	0.4	0.19419702888
<b>07/07/2013 16:40</b>	1928	1.8	0.18374982476
<b>07/07/2013 16:45</b>	1929	8.4	0.09984835237
<b>07/07/2013 16:50</b>	1930	0.2	0.05865130946
<b>07/07/2013 16:55</b>	1931	0	0.01256913692
<b>07/07/2013 17:00</b>	1932	0	0.00776897417
<b>07/07/2013 17:05</b>	1933	0.4	0.00000000000
<b>07/07/2013 17:10</b>	1934	0	0.00627813023
<b>07/07/2013 17:20</b>	1936	2.4	0.00000000000
<b>07/07/2013 17:25</b>	1937	3.2	0.00000000000
<b>07/07/2013 17:30</b>	1938	3.2	0.00000000000
<b>07/07/2013 17:35</b>	1939	2.4	0.00000000000
<b>07/07/2013 17:40</b>	1940	2.8	0.00000000000
<b>07/07/2013 17:45</b>	1941	0.2	0.00627813023
<b>07/07/2013 17:50</b>	1942	0	0.02477201447
<b>07/07/2013 17:55</b>	1943	0	0.01197601203

<b>07/07/2013 18:00</b>	1944	0	0.01332534850
<b>07/07/2013 18:05</b>	1945	0	0.03932371363
<b>07/07/2013 18:10</b>	1946	0	0.08822888881
<b>07/07/2013 18:15</b>	1947	0	0.26515242457
<b>07/07/2013 18:20</b>	1948	0	0.52124238014
<b>07/07/2013 18:25</b>	1949	0	0.13660268486
<b>07/07/2013 18:30</b>	1950	0	0.16523861885
<b>07/07/2013 18:35</b>	1951	0	0.09427567571
<b>07/07/2013 18:40</b>	1952	0	0.00933218375
<b>07/07/2013 18:45</b>	1953	0	0.00314550311
<b>07/07/2013 19:10</b>	1958	0.6	0.00000000000
<b>07/07/2013 19:15</b>	1959	1.2	0.00000000000
<b>07/07/2013 19:20</b>	1960	1.4	0.00000000000
<b>07/07/2013 19:25</b>	1961	2.6	0.00000000000
<b>07/07/2013 19:30</b>	1962	0.8	0.00000000000
<b>07/07/2013 19:35</b>	1963	1	0.00000000000
<b>07/07/2013 19:40</b>	1964	0.6	0.00000000000
<b>07/07/2013 19:45</b>	1965	0.4	0.00000000000
<b>07/07/2013 19:50</b>	1966	0.4	0.00000000000
<b>07/07/2013 19:55</b>	1967	0.2	0.00000000000
<b>07/07/2013 21:40</b>	1988	0.4	0.00000000000
<b>08/07/2013 10:50</b>	2146	0.2	0.00000000000
<b>08/07/2013 15:40</b>	2204	0	0.03201349825
<b>08/07/2013 15:45</b>	2205	0	0.12803991139

<b>08/07/2013 15:50</b>	2206	0	0.22655998170
<b>08/07/2013 15:55</b>	2207	0	0.34979164600
<b>08/07/2013 16:00</b>	2208	0	0.48503252864
<b>08/07/2013 16:05</b>	2209	0	0.76617014408
<b>08/07/2013 16:10</b>	2210	0	0.63514554501
<b>08/07/2013 16:15</b>	2211	0	0.60967636108
<b>08/07/2013 16:20</b>	2212	0	0.51688557863
<b>08/07/2013 16:25</b>	2213	0	0.33078551292
<b>08/07/2013 16:30</b>	2214	0	0.47562757134
<b>08/07/2013 16:35</b>	2215	0	0.38391694427
<b>08/07/2013 16:40</b>	2216	0	0.63723814488
<b>08/07/2013 16:45</b>	2217	0	0.57101362944
<b>08/07/2013 16:50</b>	2218	0	0.41748920083
<b>08/07/2013 16:55</b>	2219	0	0.58954864740
<b>08/07/2013 17:00</b>	2220	0	0.98621457815
<b>08/07/2013 17:05</b>	2221	3.2	0.72468852997
<b>08/07/2013 17:10</b>	2222	9.4	0.53230404854
<b>08/07/2013 17:15</b>	2223	8	0.58830398321
<b>08/07/2013 17:20</b>	2224	4.8	0.35797113180
<b>08/07/2013 17:25</b>	2225	5.4	0.22248373926
<b>08/07/2013 17:30</b>	2226	5	0.13998968899
<b>08/07/2013 17:35</b>	2227	2.2	0.02866717801
<b>08/07/2013 17:40</b>	2228	0.8	0.01546579879
<b>08/07/2013 17:45</b>	2229	1.4	0.00000000000

<b>08/07/2013 17:50</b>	2230	2.4	0.000000000000
<b>08/07/2013 17:55</b>	2231	4.2	0.000000000000
<b>08/07/2013 18:00</b>	2232	4	0.000000000000
<b>08/07/2013 18:05</b>	2233	6.6	0.000000000000
<b>08/07/2013 18:10</b>	2234	6.4	0.000000000000
<b>08/07/2013 18:15</b>	2235	5.8	0.000000000000
<b>08/07/2013 18:20</b>	2236	3.4	0.000000000000
<b>08/07/2013 18:25</b>	2237	1.4	0.000000000000
<b>08/07/2013 18:30</b>	2238	2	0.000000000000
<b>08/07/2013 18:35</b>	2239	0.8	0.000000000000
<b>08/07/2013 18:40</b>	2240	0.8	0.000000000000
<b>08/07/2013 18:45</b>	2241	0.2	0.000000000000
<b>08/07/2013 18:50</b>	2242	0.4	0.00272907037
<b>08/07/2013 18:55</b>	2243	0.2	0.00438552396
<b>08/07/2013 19:00</b>	2244	0.2	0.01212049276
<b>08/07/2013 19:05</b>	2245	0.4	0.00704542408
<b>08/07/2013 19:10</b>	2246	0.2	0.000000000000
<b>08/07/2013 19:15</b>	2247	0.8	0.01683998480
<b>08/07/2013 19:20</b>	2248	0.6	0.00545814075
<b>08/07/2013 19:25</b>	2249	0.2	0.00818721112
<b>08/07/2013 19:30</b>	2250	0.2	0.01257273462
<b>08/07/2013 19:35</b>	2251	0	0.01422918774
<b>08/07/2013 19:40</b>	2252	0.4	0.00438552396
<b>08/07/2013 20:10</b>	2258	0.2	0.000000000000

<b>08/07/2013 20:15</b>	2259	0.2	0.000000000000
<b>08/07/2013 20:20</b>	2260	0.4	0.000000000000
<b>08/07/2013 20:25</b>	2261	0.2	0.000000000000
<b>08/07/2013 20:30</b>	2262	0.6	0.000000000000
<b>08/07/2013 20:35</b>	2263	0.2	0.000000000000
<b>08/07/2013 20:40</b>	2264	0.2	0.000000000000
<b>08/07/2013 20:45</b>	2265	0.4	0.000000000000
<b>08/07/2013 20:55</b>	2267	0.4	0.000000000000
<b>08/07/2013 21:00</b>	2268	0.4	0.000000000000
<b>08/07/2013 21:05</b>	2269	0.4	0.000000000000
<b>08/07/2013 21:10</b>	2270	0.4	0.000000000000
<b>08/07/2013 21:15</b>	2271	0.2	0.000000000000
<b>08/07/2013 21:20</b>	2272	0.2	0.000000000000
<b>08/07/2013 21:30</b>	2274	0.4	0.000000000000
<b>08/07/2013 21:45</b>	2277	0.2	0.000000000000
<b>08/07/2013 22:00</b>	2280	0.2	0.000000000000
<b>08/07/2013 22:15</b>	2283	0.4	0.000000000000
<b>08/07/2013 22:30</b>	2286	0.2	0.000000000000
<b>08/07/2013 22:45</b>	2289	0.2	0.000000000000
<b>08/07/2013 23:00</b>	2292	0.4	0.000000000000
<b>08/07/2013 23:10</b>	2294	0.2	0.000000000000
<b>08/07/2013 23:25</b>	2297	0.2	0.000000000000
<b>09/07/2013 8:20</b>	2404	0.4	0.000000000000
<b>09/07/2013 17:40</b>	2516	0.2	0.000000000000

<b>10/07/2013 2:30</b>	2622	0	0.04317809269
<b>10/07/2013 2:35</b>	2623	0	0.06629185379
<b>10/07/2013 2:40</b>	2624	0	0.01816058531
<b>10/07/2013 2:45</b>	2625	0	0.00810470898
<b>10/07/2013 2:50</b>	2626	0	0.00205288990
<b>10/07/2013 2:55</b>	2627	0	0.00153892790
<b>10/07/2013 3:35</b>	2635	0.2	0.00000000000
<b>10/07/2013 5:25</b>	2657	0.2	0.00000000000
<b>17/07/2013 13:40</b>	4772	0	0.02771229483
<b>18/07/2013 12:05</b>	5041	0	0.01422704849
<b>18/07/2013 12:10</b>	5042	0	0.02448607609
<b>18/07/2013 12:15</b>	5043	0	0.01340453047
<b>18/07/2013 12:20</b>	5044	0	0.00629100623
<b>18/07/2013 12:25</b>	5045	0	0.00314550311
<b>19/07/2013 16:40</b>	5384	0	0.00877104793
<b>19/07/2013 16:45</b>	5385	0	0.02323361114
<b>19/07/2013 16:50</b>	5386	0	0.07089038938
<b>19/07/2013 16:55</b>	5387	0	0.08894087374
<b>19/07/2013 17:00</b>	5388	0	0.05517324433
<b>19/07/2013 17:05</b>	5389	0	0.12492776662
<b>19/07/2013 17:10</b>	5390	0	0.23696109653
<b>19/07/2013 17:15</b>	5391	0	0.40488809347
<b>19/07/2013 17:20</b>	5392	0	0.35160511732
<b>19/07/2013 17:25</b>	5393	0	0.03339752182

<b>19/07/2013 17:55</b>	5399	0.4	0.000000000000
<b>19/07/2013 18:00</b>	5400	0.2	0.000000000000
<b>19/07/2013 18:10</b>	5402	0	0.02311735414
<b>19/07/2013 18:15</b>	5403	1	0.02311735414
<b>19/07/2013 18:20</b>	5404	2.2	0.00924694166
<b>19/07/2013 18:25</b>	5405	1.2	0.000000000000
<b>19/07/2013 18:30</b>	5406	0.2	0.000000000000
<b>19/07/2013 18:45</b>	5409	0.2	0.000000000000
<b>19/07/2013 19:55</b>	5423	0.2	0.000000000000
<b>19/07/2013 20:50</b>	5434	0	0.00311668031
<b>19/07/2013 22:00</b>	5448	0	0.02585225366
<b>19/07/2013 22:05</b>	5449	0	0.01814420708
<b>19/07/2013 22:10</b>	5450	0	0.00773346424
<b>19/07/2013 22:15</b>	5451	0	0.00155834015
<b>19/07/2013 23:50</b>	5470	0	0.03139065206
<b>19/07/2013 23:55</b>	5471	0	0.08024308085
<b>20/07/2013 0:00</b>	5472	0	0.02337447554
<b>20/07/2013 0:05</b>	5473	0	0.00314550311
<b>20/07/2013 2:05</b>	5497	0.2	0.000000000000
<b>20/07/2013 2:55</b>	5507	0.2	0.000000000000
<b>20/07/2013 3:00</b>	5508	0.4	0.000000000000
<b>20/07/2013 3:05</b>	5509	0.2	0.000000000000
<b>20/07/2013 3:10</b>	5510	0.4	0.000000000000
<b>20/07/2013 11:05</b>	5605	0.2	0.000000000000

<b>23/07/2013 13:20</b>	6496	0	0.00153892790
<b>23/07/2013 16:20</b>	6532	0	0.00627813023
<b>23/07/2013 16:25</b>	6533	0	0.04252587631
<b>23/07/2013 16:30</b>	6534	0	0.00760079455
<b>23/07/2013 16:35</b>	6535	0	0.02668630704
<b>23/07/2013 16:40</b>	6536	0	0.01513736881
<b>23/07/2013 16:45</b>	6537	0	0.01213624794
<b>23/07/2013 16:50</b>	6538	0	0.00626218319
<b>24/07/2013 7:25</b>	6713	0	0.00314550311
<b>24/07/2013 7:35</b>	6715	0	0.00711352425
<b>24/07/2013 12:15</b>	6771	0	0.16676643491
<b>24/07/2013 12:20</b>	6772	0	0.08998390287
<b>24/07/2013 12:30</b>	6774	0	0.06358114630
<b>27/07/2013 14:50</b>	7666	0	0.00903243851
<b>27/07/2013 14:55</b>	7667	0	0.04923150316
<b>27/07/2013 15:00</b>	7668	0	0.11889864504
<b>27/07/2013 15:05</b>	7669	0	0.10510428250
<b>27/07/2013 15:10</b>	7670	0	0.06336131692
<b>27/07/2013 15:15</b>	7671	0	0.05231860653
<b>27/07/2013 15:20</b>	7672	0	0.02118616924
<b>27/07/2013 15:25</b>	7673	0	0.00274370168
<b>27/07/2013 15:40</b>	7676	0	0.00272907037
<b>27/07/2013 15:45</b>	7677	0	0.00818721112
<b>27/07/2013 16:00</b>	7680	0.8	0.00000000000

<b>27/07/2013 16:05</b>	7681	4.4	0.000000000000
<b>27/07/2013 17:05</b>	7693	0.4	0.000000000000
<b>27/07/2013 20:15</b>	7731	0	0.08042112738
<b>27/07/2013 20:20</b>	7732	0	0.04945622385
<b>27/07/2013 20:25</b>	7733	0	0.09447049350
<b>27/07/2013 20:30</b>	7734	0	0.05996904895
<b>27/07/2013 20:35</b>	7735	0	0.01044660900
<b>27/07/2013 20:40</b>	7736	0	0.01884879544
<b>27/07/2013 20:45</b>	7737	0	0.04923620448
<b>27/07/2013 20:50</b>	7738	0	0.01286625769
<b>27/07/2013 20:55</b>	7739	0	0.02299354970
<b>27/07/2013 21:00</b>	7740	0	0.08042112738
<b>27/07/2013 21:05</b>	7741	0	0.13783702254
<b>27/07/2013 21:10</b>	7742	0	0.07662491500
<b>27/07/2013 21:15</b>	7743	0.4	0.01804734766
<b>27/07/2013 21:20</b>	7744	3.6	0.00558945909
<b>27/07/2013 21:25</b>	7745	1	0.000000000000
<b>27/07/2013 22:05</b>	7753	2.8	0.000000000000
<b>27/07/2013 22:10</b>	7754	1.8	0.000000000000
<b>27/07/2013 22:15</b>	7755	0.6	0.000000000000
<b>27/07/2013 22:20</b>	7756	0.4	0.000000000000
<b>27/07/2013 23:55</b>	7775	0	0.01385454647
<b>29/07/2013 15:25</b>	8249	0.2	0.000000000000
<b>30/07/2013 14:40</b>	8528	0	0.01320138015

<b>31/07/2013 19:35</b>	8875	0.2	0.000000000000
<b>31/07/2013 19:45</b>	8877	0.4	0.000000000000
<b>31/07/2013 20:00</b>	8880	0.2	0.000000000000
<b>31/07/2013 20:05</b>	8881	0.2	0.000000000000
<b>31/07/2013 20:10</b>	8882	0.4	0.000000000000
<b>31/07/2013 20:15</b>	8883	0.2	0.000000000000
<b>31/07/2013 20:20</b>	8884	0.4	0.000000000000
<b>31/07/2013 20:25</b>	8885	0.2	0.000000000000
<b>31/07/2013 20:30</b>	8886	0.2	0.000000000000
<b>31/07/2013 20:35</b>	8887	0.6	0.000000000000
<b>31/07/2013 20:40</b>	8888	0.4	0.000000000000
<b>31/07/2013 20:45</b>	8889	0.2	0.000000000000
<b>31/07/2013 20:50</b>	8890	0.4	0.000000000000
<b>31/07/2013 20:55</b>	8891	0.6	0.000000000000
<b>31/07/2013 21:00</b>	8892	0.2	0.000000000000
<b>31/07/2013 21:05</b>	8893	0.2	0.000000000000
<b>31/07/2013 21:10</b>	8894	0.6	0.000000000000
<b>31/07/2013 21:15</b>	8895	0.2	0.000000000000
<b>31/07/2013 21:20</b>	8896	0.2	0.000000000000
<b>31/07/2013 21:25</b>	8897	0.6	0.000000000000
<b>31/07/2013 21:30</b>	8898	0.2	0.000000000000
<b>31/07/2013 21:35</b>	8899	0.4	0.000000000000
<b>31/07/2013 21:40</b>	8900	0.6	0.000000000000
<b>31/07/2013 21:50</b>	8902	0.4	0.000000000000

<b>31/07/2013 21:55</b>	8903	0.6	0.000000000000
<b>31/07/2013 22:05</b>	8905	0.2	0.000000000000
<b>31/07/2013 22:20</b>	8908	0.2	0.000000000000
<b>31/07/2013 22:35</b>	8911	0.2	0.000000000000
<b>31/07/2013 23:00</b>	8916	0.2	0.000000000000
<b>01/08/2013 0:00</b>	8928	0.2	0.000000000000

## References

- Ahrens, B. (2006). Distance in Spatial Interpolation of Daily Rain Gauge Data. *Hydrology and Earth System Sciences*, 10, 197-208.
- Allan, R. (2011). Human Influence on Rainfall. *Nature*, 470(7334), 344-345.
- Allen, D. (2011). *Getting to Know ArcGIS ModelBuilder* (1st ed.). Redlands: ESRI Press.
- Amaguchi, H., Kawamura, A., Olsson, J., & Takasaki, T. (2012). Development and Testing of a Distributed Urban Storm Runoff Event Model with a Vector-Based Catchment Delineation. *Journal of Hydrology*, 420-421, 205-215.
- Ansari, S. (2013, November 14). *NOAA's Weather and Climate Toolkit*. Retrieved April 14, 2014, from NOAA's National Climatic Data Center Web site:  
<http://www.ncdc.noaa.gov/wct/>
- ArcGIS for Local Government team. (2013, August 19). *Water Utility Network Editing (ArcGIS 10)*. Retrieved October 31, 2013, from ArcGIS Web site:  
<http://www.arcgis.com/home/item.html?id=14bdfb26bc2f4e0388a633fbd0ccca02>
- Australian Bureau of Statistics. (2013, April 24). *Sample Size Calculator*. Retrieved April 1, 2014, from National Statistical Service Web site:  
<http://www.nss.gov.au/nss/home.nsf/pages/Sample+size+calculator>
- Baker, M. E., Weller, D. E., & Jordan, T. E. (2006). Comparison of Automated Watershed Delineations: Effects on Land Cover Areas, Percentages, and Relationships to Nutrient Discharge. *Photogrammetric Engineering & Remote Sensing*, 72(2), 159-168.
- Barron, O., Pollock, D., & Dawes, W. (2011). Evaluation of Catchment Contributing Areas and Storm Runoff in Flat Terrain Subject to Urbanisation. *Hydrology and Earth System Sciences*, 15, 547-559.
- Bates, P. (2012). Integrating Remote Sensing Data with Flood Inundation Models: How Far Have we Got? *Hydrological Processes*, 26(16), 2515–2521.
- Berezowski, T., Chormanski, J., Batelaan, O., Canters, F., & Voorde, T. (2012). Impact of Remotely Sensed Land-Cover Proportions on Urban Runoff Prediction. *International Journal of Applied Earth Observation and Geoinformation*, 16, 54-65.
- Berne, A., & Krajewski, W. (2013). Radar for Hydrology: Unfulfilled Promise or Unrecognized Potential? *Advances in Water Resources*, 51, 357-366.
- Beven, K. (1997). TOPMODEL: A Critique. *Hydrological Processes*, 11, 1069-1085.

- Blumensaat, F., Wolfram, M., & Krebs, P. (2012). Sewer Model Development Under Minimum Data Requirements. *Environmental Earth Sciences*, 65(5), 1427-1437.
- CCME. (2009, February 17). *Canada-Wide Strategy for the Management of Municipal Wastewater Effluent*. Retrieved September 30, 2012, from CCME Web site: [http://www.ccme.ca/assets/pdf/cda\\_wide\\_strategy\\_mwwe\\_final\\_e.pdf](http://www.ccme.ca/assets/pdf/cda_wide_strategy_mwwe_final_e.pdf)
- CCME. (2011, May 10). *About CCME*. Retrieved October 30, 2012, from <http://www.ccme.ca/about/>
- Chen, F., & Liu, C. (2012). Estimation of the Spatial Rainfall Distribution Using Inverse Distance Weighting (IDW) in the Middle of Taiwan. *Paddy and Water Environment*, 10(3), 209-222.
- Chen, J., Hill, A. A., & Urbano, L. D. (2009). A GIS-Based Model for Urban Flood Inundation. *Journal of Hydrology*, 373, 184-192.
- Chen, M., Tucker, C., Vallabhaneni, S., Koran, J., Gatterdam, M., & Wride, D. (2003). Comparing Different Approaches of Catchment Delineation. *2003 Esri International User Conference*. San Diego: ESRI.
- Choi, Y. (2012). A New Algorithm to Calculate Weighted Flow-Accumulation from a DEM by Considering Surface and Underground Stormwater Infrastructure. *Environmental Modelling & Software*, 30, 81-91.
- City of Toronto. (2006, May 12). City of Toronto Parcels (PARCEL3\_Aroll\_group4M). Toronto, Ontario, Canada.
- City of Toronto. (2007, November 29). Buildings2007+4mil. Toronto, Ontario, Canada.
- City of Toronto. (2010, September 9). Road\_Corridors\_allcity. Toronto, Ontario, Canada.
- City of Toronto. (2011a, September 27). DEM\_CITY. Toronto, Ontario, Canada.
- City of Toronto. (2011b). Ortho Imagery 2011 10cm. Toronto, Ontario, Canada.
- City of Toronto. (2013a, May 30). Rain\_Gauges\_2013. Toronto, Ontario, Canada.
- City of Toronto. (2013b, January 30). *Toronto's Geography*. Retrieved January 28, 2014 , from City of Toronto Web site: [http://www.toronto.ca/toronto\\_facts/geography.htm](http://www.toronto.ca/toronto_facts/geography.htm)
- Clarifica. (2010). *Toronto Area 3 Drainage Study City of Toronto*. Technical Memorandum, Markham.
- Cole. (2014a, April 8). *Maintenance Log Details*. Retrieved April 28, 2014, from Data Current Website: <http://cole.datacurrent.ca/editmaintenancelogs.php?action=details&logid=37792>

- Cole. (2014b, April 1). *Precipitation Monitoring*. Retrieved from Data Current Web site: [http://cole.datacurrent.ca/project\\_info.php?pid=161](http://cole.datacurrent.ca/project_info.php?pid=161)
- Costa-Cabral, M., & Burges, S. (1994). Digital Elevation Model Networks (DEMON): A Model of Flow over Hillslopes for Computation of Contributing and Dispersal Areas. *Water Resources Research*, 30(6), 1681–1692.
- Demarchi, L., Canters, F., Chan, J., Ampe, E., & Batelaan, O. (2012). Use of Land-Cover Fractions Derived from MESMA for Urban Water Balance Calculation. *Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1594 -1597). Munich: IEEE.
- Department of Justice. (2012, September 19). *Wastewater Systems Effluent Regulations, SOR/2012-139*. Retrieved September 30, 2012, from Department of Justice Web site: <http://laws-lois.justice.gc.ca/eng/regulations/SOR-2012-139/FullText.html>
- Diaz-Nieto, J., Blanksby, J., Lerner, D., & Saul, A. (2008). A GIS Approach to Explore Urban Flood Risk Management. *Proceedings 11th International Conference on Urban Drainage* (pp. 1-10). Edinburgh, Scotland, UK: IAHR/IWA.
- Diaz-Nieto, J., Lerner, D., Saul, A., & Blanksby, J. (2012). GIS Water-Balance Approach to Support Surface Water Flood-Risk Management. *Journal of Hydrologic Engineering*, 17(1), 55-67.
- Dirks, K., Hay, J., Stow, C., & Harris, D. (1998). High-Resolution Studies of Rainfall on Norfolk Island Part II: Interpolation of Rainfall Data. *Journal of Hydrology*, 208(3-4), 187-193.
- Djokic, D. (2008, December). Comprehensive Terrain Preprocessing Using Arc Hydro Tools. Redlands, California, USA: ESRI.
- Djordjević, S., Prodanović, D., & Maksimović, C. (1999). An Approach to Simulation of Dual Drainage. *Water Science and Technology*, 39(9), 95-103.
- Dongquan, Z., Jining, C., Haozheng, W., Qingyuan, T., Shangbing, C., & Zheng, S. (2009). GIS-Based Urban Rainfall-Runoff Modeling Using an Automatic Catchment-Discretization Approach: A Case Study in Macau. *Environmental Earth Sciences*, 59(2), 465-472.
- Duke, G., Kienzle, S., Johnson, D., & Byrne, J. (2006). Incorporating Ancillary Data to Refine Anthropogenically Modified Overland Flow Paths. *Hydrological Processes*, 20(8), 1827-1843.
- Edwards, J., Koval, E., Lendt, B., Ginther, P., Black, A., Larsen, L., & Hauffen, P. (2011). Benefits of Integrating GIS and Hydraulic Modeling. In L. Armstrong (Ed.), *Hydraulic Modeling and GIS* (pp. 3-22). Redlands, California, USA: ESRI Press.

- Einfalt, T., Arnbjerg-Nielsen, K., Golz, C., Jensen, N., Quirmbach, M., Vaes, G., & Vieux, B. (2004). Towards a Roadmap for use of Radar Rainfall Data in Urban Drainage. *Journal of Hydrology*, 299(3-4), 186-202.
- Elliott, A., & Trowsdale, S. (2007). A Review of Models for Low Impact Urban Stormwater Drainage. *Environmental Modelling & Software*, 22(3), 394-405.
- ESRI. (2010). ArcGIS Desktop Help. Redlands, California.
- ESRI. (2011a). Arc Hydro Toolbox Help. Redlands, California, USA.
- ESRI. (2011b, October). Arc Hydro Tools - Tutorial Version 2.0. Redlands, California, USA: ESRI Press.
- Fankhauser, R. (1999). Automatic Determination of Imperviousness in Urban Areas from Digital Orthophotos. *Water Science and Technology*, 39(9), 81-86.
- Fewtrell, T., Duncan, A., Sampson, C., Neal, J., & Bates, P. (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.
- Foley, A. (2008, September 8). *Using Arc Hydro in Florida*. Retrieved from University of Florida Web site:  
[http://www.clas.ufl.edu/users/mbinford/geo5159\\_gis\\_environmental/seminar\\_presentations/uf\\_archydroinflorida\\_afoley20080908.pdf](http://www.clas.ufl.edu/users/mbinford/geo5159_gis_environmental/seminar_presentations/uf_archydroinflorida_afoley20080908.pdf)
- Garbrecht, J., & Martz, L. (2000). Digital Elevation Model Issues in Water Resources Modeling. In D. Maidment, & D. Djokic (Eds.), *Hydrologic and Hydraulic Modeling Support with Geographic Information Systems* (pp. 1-28). Redlands: ESRI Press.
- Gasperi, J., Garnaud, S., Rocher, V., & Moilleron, R. (2008). Priority Pollutants in Wastewater and Combined Sewer Overflow. *Science of The Total Environment*, 407(1), 263-272.
- Geospatial Competency Centre. (2013, June 13). CITYPRJ\_BUILDING\_A. Toronto, Ontario, Canada.
- Geospatial Competency Centre. (2014, February 7). *Toronto Centreline (TCL)*. Retrieved from City of Toronto Web site:  
<http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=9acb5f9cd70bb210VgnVCM1000003dd60f89RCRD>
- Giron, E. (2005). *Development of a SWMM-GIS Flood Model for New Orleans Drainage Pumping Station No 4 Basin*. University of New Orleans. New Orleans: University of New Orleans.

- Gironás, J., Niemann, J., Roesner, L., Rodriguez, F., & Andrieu, H. (2010). Evaluation of Methods for Representing Urban Terrain in Storm-Water Modeling. *Journal of Hydrologic Engineering*, 15(1), 1-14.
- Goormans, T., & Willems, P. (2013). Using Local Weather Radar Data for Sewer System Modeling: Case Study in Flanders, Belgium. *Journal of Hydrologic Engineering*, 18(2), 269-278.
- Government of Canada. (2012, July 18). *Regulatory Impact Analysis Statement: Wastewater Systems Effluent Regulations*. Retrieved October 31, 2012, from <http://www.gazette.gc.ca/rp-pr/p2/2012/2012-07-18/html/sor-dors139-eng.html>
- Greenbaum, A., & Wellington, A. (2010). *Environmental Law and Policy in the Canadian Context*. Concord: Captus Press Inc.
- Haberlandt, U. (2007). Geostatistical Interpolation of Hourly Precipitation from Rain Gauges and Radar for a Large-Scale Extreme Rainfall Event. *Journal of Hydrology*, 332(1-2), 144-157.
- Haklay, M. (2010). How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environment and Planning B: Planning and Design*, 37(4), 682-703.
- Hamel, P., Daly, E., & Fletcher, T. (2013). Source-Control Stormwater Management for Mitigating the Impacts of Urbanisation on Baseflow: A Review. *Journal of Hydrology*, 485(2), 201-211.
- Hammond, M., & Han, D. (2006). Issues of Using Digital Maps for Catchment Delineation. *Water Management*, 159(1), 45-51.
- Hutchinson, M. (1998). Interpolation of Rainfall Data with Thin Plate Smoothing Splines - Part I: Two Dimensional Smoothing of Data with Short Range Correlation. *Journal of Geographic Information and Decision Analysis*, 2(2), 139-151.
- Innovyze. (2013, February 12). InfoWorks CS Help v13.5. Wallingford, Oxfordshire, United Kingdom.
- Innovyze. (2014). *InfoNet Overview*. Retrieved May 28, 2014, from Innovyze Web site: <http://www.innovyze.com/products/infonet/>
- Intergovernmental Panel on Climate Change (IPCC). (2007). *Climate Change 2007 - The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC*. New York: Cambridge University Press.

- Irvine, K., Caruso, J., & McCorkhill, G. (2005). Consideration of Metals Levels in Identifying CSO Abatement Option. *Urban Water Journal*, 2(3), 193-200.
- Jankowfsky, S., Branger, F., Braud, I., Gironás, J., & Rodriguez, F. (2012). Comparison of Catchment and Network Delineation Approaches in Complex Suburban Environments: Application to the Chaudanne Catchment, France. *Hydrological Processes*, in press. doi:10.1002/hyp.9506
- Jenness, J. (2012). *Repeating Shapes for ArcGIS*. Retrieved from Jenness Enterprises.: [http://www.jennessent.com/arcgis/repeat\\_shapes.htm](http://www.jennessent.com/arcgis/repeat_shapes.htm)
- Johnson, L. (2009). *Geographic Information Systems in Water Resources Engineering*. Boca Raton: CRC Press.
- Knebl, M., Yang, Z., Hutchison, K., & Maidment, D. (2005). Regional Scale Flood Modeling using NEXRAD Rainfall, GIS, and HEC-HMS/RAS: A Case Study for the San Antonio River Basin Summer 2002 Storm Event. *Journal of Environmental Management*, 75(4), 325-336.
- Lam, C. (2004). *Comparison of Flow Routing Algorithms used in Geographic Information Systems*. University of Southern California. Los Angeles: University of Southern California.
- Lau, J., Butler, D., & Schutze, M. (2002). Is Combined Sewer Overflow Spill Frequency/Volume a Good Indicator of Receiving Water Quality Impact? *Urban Water*, 4(2), 181-189.
- Leandro, J. (2008). *Advanced Modelling of Flooding in Urban Areas Integrated 1D/1D and 1D/2D Models*. PhD Thesis, University of Exeter, Exeter.
- Leandro, J., Chen, A., Djordjević, S., & Savić, D. (2009). Comparison of 1D/1D and 1D/2D Coupled (Sewer/Surface) Hydraulic Models for Urban Flood Simulation. *Journal of Hydraulic Engineering*, 135(6), 495-504.
- Leitão, J., Prodanovic, D., Boonya-aroonnet, S., & Maksimovic, C. (2012). Enhanced DEM-based Flow Path Delineation Methods. *Journal of Hydroinformatics, Uncorrected Proof*, 1-12.
- Lhomme, J., Bouvier, C., & Perrin, J. (2004). Applying a GIS-based Geomorphological Routing Model in Urban Catchments. *Journal of Hydrology*, 299(3-4), 203-216.
- Lin, G. (2013, March 3). citycomb\_final\_mdf\_Grace. Toronto, Ontario, Canada.
- Lindsay, J. (2009, November 19). Whitebox Geospatial Analysis Tools Version 2.0.3. Guelph, Ontario, Canada.

- Lindsay, J., & Creed, I. (2005). Removal of Artifact Depressions from Digital Elevation Models: Towards a Minimum Impact Approach. *Hydrological Processes*, 19(16), 3113–3126.
- Long, B., Chang, J., & John, D. (2009). What to Fix First in the Twenty-First Century. In C. Thomas, & N. Humenik-Sappington, *GIS for Decision Support and Public Policy Making* (pp. 24-30). Redlands: ESRI Press.
- Luciani, P. (2005). *Distributed Urban Stormwater Modelling within GIS integrating Analytical Probabilistic Hydrologic Models and Digital Imagery*. MASC Thesis, Ryerson University, Toronto.
- Luciani, P., Li, J., & Banting, D. (2011). Distributed Urban Storm Water Modeling within GIS Integrating Analytical Probabilistic Hydrologic Models and Remote Sensing Image Analyses. *Water Quality Research Journal of Canada*, 46(3), 183-199.
- Ly, S., Charles, C., & Degré, A. (2013). Different Methods for Spatial Interpolation of Rainfall Data for Operational Hydrology and Hydrological Modeling at Watershed Scale. A Review. *Biotechnology, Agronomy, Society and Environment*, 17(2), 392-406.
- Maidment, D. (2002). Why Arc Hydro? In D. Maidment (Ed.), *Arc Hydro: GIS for Water Resources* (pp. 1-12). Redlands: ESRI Press.
- Mark, O., Weesakul, S., Apirumanekul, C., Aroonnet, S., & Djordjevic, S. (2004). Potential and Limitations of 1D Modelling of Urban Flooding. *Journal of Hydrology*, 299(3-4), 284-299.
- Marsalek, J., Jiménez-Cisneros, B., Malmquist, P.-A., Karamouz, M., Goldenfum, J., & Chocat, B. (2006). *Urban Water Cycle Processes and Interactions*. Paris: International Hydrological Programme.
- McCarthy, M., Best, M., & Betts, R. (2010). Climate Change in Cities Due to Global Warming and Urban Effects. *Geophysical Research Letters*, 37(9).
- McGreer, E., & Belzer, W. (1999, November 24). *Contaminant Sources*. Retrieved from Simon Fraser University Web site: [http://research.rem.sfu.ca/frap/S\\_20.pdf](http://research.rem.sfu.ca/frap/S_20.pdf)
- Microsoft. (2007). Excel 2007 Help. Redmond, Washington.
- Morelli, S., Segoni, S., Manzo, G., Ermini, L., & Catani, F. (2012). Urban Planning, Flood Risk and Public Policy: The Case of the Arno River, Firenze, Italy. *Applied Geography*, 34, 205-218.
- Mouton, A. (2005). *Generating Stream Maps Using LiDAR Derived Digital Elevation Models and 10-m USGS DEM*. University of Washington, Seattle.

- National Climatic Data Center. (2014, May 11). *HDSS Access System*. Retrieved from NOAA Web site:  
<http://has.ncdc.noaa.gov/pls/plhas/HAS.FileAppRouter?datasetname=7000&subqueryby=STATION&applname=&outdest=FILE>
- NOAA Biogeography Branch. (2013, September 13). *Sampling Design Tool for ArcGIS Instruction Manual*. Retrieved March 31, 2014, from ArcGIS Web site:  
<http://www.arcgis.com/home/item.html?id=ecbe1fc44f35465f9dea42ef9b63e785>
- O'Callaghan, J., & Mark, D. (1984). The Extraction of Drainage Networks from Digital Elevation Data. *Computer Vision, Graphics, and Image Processing*, 28(3), 323-344.
- Olivera, F., Furnans, J., Maidment, D., Djokic, D., & Ye, Z. (2002). Drainage systems. In *Arc Hydro: GIS for water resources* (pp. 55-86). Redlands: ESRI Press.
- Onafrychuk, C. (2007). *Tracking Saturated Areas in the Landscape: A Topographic Approach*. MSA Thesis, Ryerson University, Toronto.
- Ontario Ministry of Finance. (2013). *Ontario Population Projections Update, 2012–2036*. Toronto: Queen's Printer for Ontario.
- Pan, A., Hou, A., Tian, F., Ni, G., & Hu, H. (2012). Hydrologically Enhanced Distributed Urban Drainage Model and Its Application in Beijing City. *Journal of Hydrologic Engineering*, 17(6), 667-678.
- Patterson, D. (2013, December 10). *Split Layer By Attributes [Python Script Tool]*. Retrieved January 2, 2014, from Geoprocessing Model and Script Tool Gallery Website:  
<http://resources.arcgis.com/gallery/file/geoprocessing/details?entryID=37AEB018-1422-2418-A036-CA6D9920F808>
- Piazza, A., Conti, F., Noto, L., Viola, F., & Loggia, G. (2011). Comparative Analysis of Different Techniques for Spatial Interpolation of Rainfall Data to Create a Serially Complete Monthly Time Series of Precipitation for Sicily, Italy. *International Journal of Applied Earth Observation and Geoinformation*, 13(3), 396-408.
- Pilesjö, P., & Hasan, A. (2013). A Triangular Form-based Multiple Flow Algorithm to Estimate Overland Flow Distribution and Accumulation on a Digital Elevation Model. *Transactions in GIS*, 1-17. doi:10.1111/tgis.12015
- Privy Council Office. (2012, February 8). *Federalism in Canada*. Retrieved October 27, 2012, from Intergovernmental Affairs: <http://www.pco-bcp.gc.ca/aia/index.asp?lang=eng&page=federal&sub=why-pourquoi&doc=why-pourquoi-eng.htm#2>

- Prodanović, D., Stanić, M., Milivojević, V., Simić, Z., & Arsić, M. (2009). DEM-Based GIS Algorithms for Automatic Creation of Hydrological Models Data. *Journal of the Serbian Society for Computational Mechanics*, 3(1), 64-85.
- Quinn, P., Beven, K., & Lamb, R. (1995). The  $\ln(a/\tan\beta)$  Index: How to Calculate it and how to use it Within TOPMODEL Framework. *Hydrological Processes*, 9(2), 161-182.
- Quinn, P., Beven, K., Chevallier, P., & Planchon, O. (1991). The Prediction of Hillslope Flow Paths for Distributed Hydrological Modelling using Digital Terrain Models. *Hydrological Processes*, 5(1), 59-79.
- Robayo, O. (2005). *Map to Map: Converting a NEXRAD Rainfall Map into a Flood Inundation Map*. The University of Texas at Austin. Austin: ProQuest Information and Learning Company.
- Rodriguez, F., Andrieu, H., & Morena, F. (2008). A Distributed Hydrological Model for Urbanized Areas – Model Development and Application to Case Studies. *Journal of Hydrology*, 351(3-4), 268-287.
- Ruelland, D., Ardoin-Bardin, S., Billen, G., & Servat, E. (2008). Sensitivity of a Lumped and Semi-Distributed Hydrological Model to Several Methods of Rainfall Interpolation on a Large Basin in West Africa. *Journal of Hydrology*, 361(1-2), 96-117.
- Sampson, C., Fewtrell, T., Duncan, A., Shaad, K., Horritt, M., & Bates, P. (2012). Use of Terrestrial Laser Scanning Data to Drive Decimetric Resolution Urban Inundation Models. *Advances in Water Resources*, 41, 1-17.
- Saunders, W. (2000). Preparation of DEMs for use in Environmental Modeling Analysis. In D. Maidment, & D. Djokic (Eds.), *Hydrologic and hydraulic modeling support with geographic information systems* (pp. 29-52). Redlands: ESRI Press.
- Seth, I., Soonthornnonda, P., & Christensen, E. (2006). Use of GIS in Urban Storm-Water Modeling. *Journal of Environmental Engineering*, 132(12), 1550–1552.
- Shafiei, M., Ghahraman, B., Saghafian, B., Pande, S., Gharari, S., & Davary, K. (2014). Assessment of Rain-Gauge Networks using a Probabilistic GIS Based Approach. *Hydrology Research*, 45(4-5), 551–562.
- Shamsi, U. (2005). *GIS Applications for Water, Wastewater, and Stormwater Systems*. New York: CRC Press.
- Simões, N., Ochoa, S., Leitão, J., Pina, R., Marques, A., & Maksimović, Č. (2011). Urban Drainage Models for Flood Forecasting: 1D/1D, 1D/2D and Hybrid Models. *12th International Conference on Urban Drainage*, (pp. 1-8). Porto Alegre.

- Soenario, I., & Sluiter, R. (2010). *Optimization of Rainfall Interpolation*. De Bilt: Royal Dutch Meteorological Institute.
- Spangler, C. (1992). Interpolation Methods. In L. Brand, & M. Johnson (Eds.), *Methods in Enzymology: Numerical Computer Methods* (pp. 305-314). London: Academic Press.
- Storck, P., Bowling, L., Wetherbee, P., & Letternmaier, D. (1998). Application of a GIS-Based Distributed Hydrology Model for Prediction of Forest Harvest Effects on Peak Stream Flow in the Pacific Northwest. *Hydrological Processes*, 12(6), 889-904.
- Tarboton, D. (1997). A New Method for the Determination of Flow Directions and Upslope Areas in Grid Digital Elevation Models. *Water Resources Research*, 33(2), 309-319.
- Tarboton, D. (2003). Terrain Analysis using Digital Elevation Models in Hydrology. *23rd Annual Esri International User Conference*. San Diego: ESRI.
- Taylor, R. (1990). Interpretation of the Correlation Coefficient: A Basic Review. *Journal of Diagnostic Medical Sonography*, 6(1), 35-39.
- Toronto Water. (2014, February 11). Toronto Water Asset Geodatabase (TWAG). Toronto, Ontario, Canada.
- TRCA. (2013, July 13). *Restoration Opportunity Planning Process*. Retrieved from Toronto and Region Conservation Authority Web site: <http://trca.on.ca/dotAsset/166351.pdf>
- Weng, Q. (2001). Modeling Urban Growth Effects on Surface Runoff with the Integration of Remote Sensing and GIS. *Environmental Management*, 28(6), 737-748.
- Wilson, J., Aggett, G., Yongxin, D., & Lam, C. (2008). Water in the Landscape: A Review of Contemporary Flow Routing Algorithms. In Q. Zhou, B. Lees, & G.-a. Tang (Eds.), *Advances in Digital Terrain Analysis* (pp. 213-236). Berlin: Springer.