Impact of Sewershed Characteristics on Rainfall Derived Inflow and Infiltration

Spencer Michael Sebo

Marquette University

Follow this and additional works at: https://epublications.marquette.edu/theses_open

Part of the Engineering Commons

Recommended Citation
Sebo, Spencer Michael, "Impact of Sewershed Characteristics on Rainfall Derived Inflow and Infiltration" (2020). Master's Theses (2009 -). 635.
https://epublications.marquette.edu/theses_open/635
IMPACT OF SEWERSHED CHARACTERISTICS ON RAINFALL DERIVED INFLOW AND INFILTRATION

by

Spencer M. Sebo

A Thesis submitted to the Faculty of the Graduate School, Marquette University, in Partial Fulfillment of the Requirements for the Degree of Master of Science

Milwaukee, Wisconsin

December 2020
ABSTRACT

IMPACT OF SEWERSHED CHARACTERISTICS ON RAINFALL DERIVED INFLOW AND INFILTRATION

Spencer M. Sebo

Marquette University, 2020

Cities rely on sewer systems to transport wastewater and stormwater, but sometimes these systems are overwhelmed, and their capacity is exceeded due to excess water in the system from rainfall derived inflow and infiltration. This can lead to overflows and backups that can be detrimental to health and property. Excess water from rainfall derived inflow and infiltration enters a sewer system through a multitude of ways, including downspout connections, foundation drains, pipe joints, and broken pipes. Identifying these sources individually can be time intensive and expensive if entire service areas need to be addressed. High level screening tools are therefore needed that can apply readily available data to identify areas and sources of rainfall derived inflow and infiltration in a sewer system. This study seeks to address this challenge by using monitoring data from Milwaukee, Wisconsin to derive correlations between known sewershed characteristics and rainfall derived inflow and infiltration. Results show that pipe length per acre, number of parcels, and medium intensity land use are positively correlated to inflow or fast direct flows into the system. In addition, imperviousness, pipe length per acre, low intensity and medium intensity land use are negatively correlated with infiltration or slow inputs from groundwater sources. These findings can be applied by water reclamation managers to narrow the search areas for rainfall derived inflow and infiltration sources within their sanitary sewer systems.
ACKNOWLEDGEMENTS

I would like to thank my advisors, Dr. Walter McDonald, Dr. Anthony Parolari, and Dr. Margaret McNamara for their guidance and contributions to my graduate education and research. Their mentorship was instrumental in my development and training as a researcher and they have greatly broadened my understanding of science and engineering.

I would also like to thank the sponsors of WEF, specifically Milwaukee Metropolitan Sewerage District, Wisconsin Department of Natural Resources, and City of Milwaukee, for funding this research and providing guidance throughout this process.

Finally, I would like to thank all my colleagues with the Hydrology Lab for their help and support with this research. Specifically, Denny Malec, for the help with data collection and formatting. Thank you to everyone in the Hydrology Lab; Joe Naughton, Tamim Sharior, Laine Havens, Liz Regier, Charitha Gunawardana, Isabelle Horvath, Duyen Lam, Kassidy O’Malley, Nathan Hay, and Catherine Sullivan; you have all been incredible colleagues and it was such a privilege to work with such a great group of intelligent, driven, and hardworking individuals.
DEDICATION

I would like to dedicate this work to my parents, Mike and Grace Sebo for their continued support throughout my academic career, and for fostering in me from a young age, a curiosity and love of learning that has driven me to continue my pursuit of a quality education and higher achievement. Without their support I would have never made it this far. I would also like to dedicate this work to Madeleine Behr for her support through the stress, long hours, and late nights it took to get here.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .......................................................................................................................... i
DEDICATION ........................................................................................................................................... ii
TABLE OF CONTENTS ........................................................................................................................... iii
LIST OF TABLES ...................................................................................................................................... v
LIST OF FIGURES ................................................................................................................................... vi

1. INTRODUCTION ............................................................................................................................... 1

2. LITERATURE REVIEW ....................................................................................................................... 4
   2.1 Sewer Overflows ......................................................................................................................... 4
   2.2 Impact of Inflow and Infiltration on Overflows ........................................................................... 4
   2.3 Solution to Sewer Overflows and Basement Backups ................................................................. 6
   2.4 Existing Models of Rainfall Derived Inflow and Infiltration ...................................................... 6
   2.5 Impact of Residential Design on Inflow and Infiltration ............................................................ 9
   2.6 Impacts of Stormwater Management on Inflow and Infiltration ............................................. 9
   2.7 Need for Research ...................................................................................................................... 11

3. METHODOLOGY .............................................................................................................................. 12
   3.1 Data Collection .......................................................................................................................... 12
   3.2 GIS Data Development .............................................................................................................. 15
   3.3 RTK Parameter Development .................................................................................................. 16
   3.4 Statistical Regression Methods .................................................................................................. 23

4. RESULTS ........................................................................................................................................... 25
   4.1 Sewershed Characteristics ......................................................................................................... 25
   4.1.1 Sewershed characteristics .................................................................................................... 25
   4.1.2 Sanitary sewer pipe characteristics ...................................................................................... 27
   4.2 Inflow and infiltration based upon RTK analysis ...................................................................... 29
4.2.1 General inflow and infiltration trends ................................................................. 29
4.2.2 Components of inflow and infiltration ................................................................. 33
4.3 Regression Results ................................................................................................. 34
4.3.1 Relationship between total R and watershed characteristics ......................... 34
4.3.2 Relationship between components of R and watershed characteristics ........ 36
4.3.3 Relationship between R ratios and sewershed characteristics ....................... 38
4.3.4 Multicollinearity among predictor variables ................................................... 42
4.3.5 Multivariable linear regression ........................................................................ 43

5. DISCUSSION .............................................................................................................. 45

6. CONCLUSION ............................................................................................................ 47

6.1 Key Findings ......................................................................................................... 47
6.2 Future Work .......................................................................................................... 48

7. REFERENCES ............................................................................................................ 50

APPENDIX .................................................................................................................... 53

A1. Sewershed and pipe characteristics .................................................................... 53
A2. Linear regression graphs for median R1, R2, R3 ............................................. 56
A3. Linear regression graphs of normalized R1, R2, R3 ........................................ 60
A4. Precipitation and RDII patterns ........................................................................ 64
A5. Multicollinearity .................................................................................................. 66
LIST OF TABLES

Table 1. Typical values of T and K values from Lai et al., 2017 .......................................................... 22

Table 2. Sewershed characteristics and summary statistics for all sewersheds ......................... 26

Table 3. Linear regression results predicting Total R based upon sewershed characteristics ...... 35

Table 4. Linear regression results predicting Total R based upon pipe materials ....................... 36

Table 5. Regression results for R1, R2, and R3 for sewershed characteristics ............................... 37

Table 6. Regression Results for R1, R2, and R3 against pipe characteristics ............................... 38

Table 7. Regression results for R1/Total R, R2/Total R, and R3/Total R for sewershed characteristics ..................................................................................................................... 39

Table 8. Regression results for R1/Total R, R2/Total R, and R3/Total R against pipe characteristics ..................................................................................................................... 42

Table 9. Correlation probability among predictor variables represented as the Pearson productmoment correlation coefficient .................................................................................................................. 43

Table 10. Multivariable linear regression results ............................................................................... 44

Table A 1. Correlations matrix among predictor variables .............................................................. 66
LIST OF FIGURES

Figure 1. Map of all sewersheds utilized in this study and corresponding flow gages and rain gages. All gages are installed and operated by Milwaukee Metropolitan Sewerage District. …… 13

Figure 2. Images from sewershed MS0503 (left) and sewershed MS0454 (right) taken from Google Maps .......................................................... 14

Figure 3. Example of weekday and weekend flows for sewershed DC066E. The pink line represents the weekend flows while the blue line represents the weekday flows. This is reflective of the different water use behaviors of residents on weekdays versus weekends. …… 17

Figure 4. Summation of 3 RTK Synthetic unit hydrographs (Vallabhaneni and Burgess, 2007) .... 21

Figure 5. Percent Imperviousness (left) and medium intensity land use (right) of sewersheds in the case study area ........................................................................................................... 26

Figure 6. Makeup of development within the sewersheds in the study ........................................ 27

Figure 7. Pipe materials as a percentage of the overall pipe length in sewersheds for which data was available ........................................................................................................... 28

Figure 8. Pipe materials in linear feet in sewersheds for which data was available ................. 28

Figure 9. Average total annual R of sewersheds in the study area (left) and distribution of the median annual R across all sewersheds...................................................................................... 30

Figure 10. Example of inflow and infiltration (R) over the course of a year from a selection of eight sites. As illustrated, many sites have greater R values during the winter months……………… 31

Figure 11. Total R and rainfall volume over the course of a year from a selection of eight sites. As illustrated, many sites have greater R values during winter months; however, this does not seem to correlate with rainfall volumes.......................................................... 31

Figure 12. Distribution of total R across all sites for each month of the year (left) and distribution of the median R across all sewersheds considering the entire year in blue and only the months Jun - Oct in red (right) .......................................................... 32

Figure 13. Distribution of total R across all sites for each year of the study ............................. 33

Figure 14. Values of R1, R2, and R3 over the course of a calendar year (a) and a comparison of median annual and Jun - Oct R1, R2, and R3 values across all sewersheds (b)............... 34

Figure 15. Linear regression for standardized R1, R2, and R3 against parcels per acre (a); average imperviousness (b); open space (c); and medium intensity land use (d) ......................... 41
Figure A 1. Spatial representation of the average slope (left) and average R in Jun - Oct (right) for the sewersheds in this study................................................................. 53

Figure A 2. Landcover classification of Milwaukee region and sewershed locations (black outlines) ........................................................................................................ 54

Figure A 3. Map of the sewersheds locations and pipe data. ................................................................. 55

Figure A 4. Linear regression of median R1, R2, and R3 for pipe length per area (a); imperviousness (b); open space land use (c); medium intensity land use (d); mean elevation (e); parcels per acre (f)........................................................................ 56

Figure A 5. Linear regression of median R1, R2, and R3 for pipe length (a); low intensity land use (b); high intensity land use (c); mean slope (d); road length per acre sewershed (e).............................. 57

Figure A 6. Linear regression of median R1, R2, and R3 for cast iron (ft) (a); cast iron (%) (b); clay (ft) (c); clay (%) (d); concrete (ft) (e); concrete (%) (f) ........................................................................ 58

Figure A 7. Linear regression of median R1, R2, and R3 for ductile iron (ft) (a); ductile iron (%) (b); PVC (ft) (c); PVC (%) (d); ABS (ft) (e); ABS (%) (f) ........................................................................ 59

Figure A 8. Linear regression of normalized median R1, R2, and R3 for pipe length (a); area (b); pipe length per area (c); imperviousness(d); number of parcels (e); parcels per acre (f)......... 60

Figure A 9. Linear regression of normalized median R1, R2, and R3 for low intensity land use (a); medium intensity land use (b); high intensity land use (c); mean elevation (d); mean slope (e); road length per acre (f)...................................................................................... 61

Figure A 10. Linear regression of normalized median R1, R2, and R3 ABS (ft) (a); ABS (%) (b); cast iron (ft) (c); cast iron (%) (d); clay (ft) (e); clay (%) (f) ........................................................................ 62

Figure A 11. Linear regression of normalized median R1, R2, and R3 concrete (ft) (a); concrete (%) (b); ductile iron (ft) (c); ductile iron (%) (d); PVC (ft) (e); PVC (%) (f)........................................................................ 63

Figure A 12. Monthly distribution of rainfall and total R (a); median total R, flow duration (hrs) and rainfall (b); distribution of R1, R2, and R3 for each sewershed (c); distribution of t1, t2, and t3 for each sewershed (d) ........................................................................ 64

Figure A 13. Antecedent rainfall as a function of total R (a); R1 (b); R2 (c); R3 (d) ......................... 65
1. INTRODUCTION

Cities across the country rely on sewer systems to transport waste and storm water to prevent flooding and maintain hygienic conditions for residents. These sewer systems generally fall within three categories: storm sewers, sanitary sewers, and combined sewers. Storm sewers collect runoff from storms and transport it to streams and rivers to prevent adverse flooding effects on city infrastructure. Sanitary sewers are connected directly to buildings that discharge wastewater – including homes, businesses, and industries – and transport that wastewater to a water reclamation plant where it is treated and discharged to the environment. Combined sewers are typically older systems that collect all stormwater and wastewater into a single pipe network and transport it to the water reclamation plant where it is treated and discharged to the environment. As a whole, these sewer systems provide critical services to communities, but sometimes they do not always operate as intended.

While sanitary and storm sewer systems have distinct roles, it is not always the case that storm sewers transport only stormwater runoff or that sanitary sewers transport only wastewater. In some cases, there may be illicit connections or cracks in the pipes that allow for inflow and infiltration of unintended sources into sewer pipes. For example, studies have found that storm sewers sometimes have high levels of fecal indicator bacteria that come from human waste intended for sanitary sewers (Sercu et al., 2009; Sauer et al., 2011). Sauer et al. (2011) found that of 4 of 5 storm sewer outfalls were heavily impacted by fecal pollution and human sewage appeared to be the major source. This can present a significant threat to human and environmental health through the discharge of untreated human waste into streams and rivers that are designated as swimmable and fishable water bodies.
In addition, sanitary sewer systems are often subject to increased flows during storm events from inflow and infiltration into stormwater pipes (Zhang et al., 2018). This is because the sanitary sewer pipes are not always a closed system. Pipes can have cracks in them from settling, tree roots, or inadequate joint connections that result in the introduction of stormwater, groundwater, or snow melt. In addition, they may have illicit connections of source water from roof drains, foundation drains, or sump pumps. The result is a significant increase in the volume of water that must be treated at the water reclamation plant.

This is a significant problem for water reclamation plants as treatment processes take time and can be expensive. Treating unnecessary stormwater can drive up the cost of water treatment and in the worst situations, the plant may not be able to handle the entire flow. This can result in overflows of untreated water into the environment, backup of waters into citizens' basements, and potential fines for the water reclamation plant. As such, municipalities invest significant resources into improving the function of their sewer systems to prevent overflows and basement backups. For example, Kansas City has committed to invest $2.5 billion dollars to eliminate overflows of untreated wastewater (Whitley, 2010).

One way that municipalities seek to prevent overflows and basement backups is by monitoring and modeling flows throughout their sanitary sewer system. This helps to provide them with a picture of what is happening within their sewers and understand where inflow and infiltration may be occurring. There is significant literature on modeling inflow and infiltration into these systems that have helped to understand where inflow and infiltration sources may occur within these systems. However, the influence that sewershed scale characteristics have on inflow and infiltration is underexplored.

This thesis seeks to fill this gap through a study that explores the relationship between rainfall-derived inflow and infiltration and sewershed characteristics for sewersheds in the
Milwaukee region. These relationships will allow managers to make informed decisions based upon the spatial characteristics of their sewersheds to reduce the impact of storms on sanitary systems. Current understanding of these impacts is limited and is mainly localized to small modeling studies focused on determining where individual sources of stormwater may be. By taking a broader view of the contributing sewershed and considering overall characteristics, improvements and repairs to sewer systems can be targeted where inflow and infiltration is highest and have the greatest possibility of being reduced.
2. LITERATURE REVIEW

2.1 Sewer Overflows

Sewer overflows occur when large rainfall event contribute vast quantities of water to sanitary and combined sewer systems, and the treatment plant is unable to process the water fast enough. In these situations, treatment plants may need to allow untreated sewage to flow through the system and into a receiving water body. Combined sewer systems serve over 40 million people in 772 communities across the nation and are estimated to overflow up to 850 billion gallons of untreated water per year (EPA, 2011). Sanitary sewer systems are estimated to overflow 23,000-75,000 times per year in the U.S., resulting in the discharge of 3-10 billion gallons of untreated wastewater (EPA, 2004). Basement backups are a related issue where the sewer system becomes so overburdened by the influx of storm water that the sewers end up flowing into people’s basements and flooding them with raw sewage. These events can have profound negative impacts on the environment and human health. Overflows contribute large quantities of polluted waters, containing human pathogens, toxic materials, and heavy metals to surface waters where humans and wildlife may encounter them (EPA, 2011). EPA estimates that 3,448-5,576 annual illnesses can be attributed to overflows impacting the nations beaches (EPA, 2004). Basement backups bring these same contaminants directly into people’s homes. The property damage and risk to human health is significant and therefore regulations have been implemented to force water reclamations plants to better manage flows and eliminate overflows. Failure to comply can result in significant fines if overflows occur.

2.2 Impact of Inflow and Infiltration on Overflows

A significant contributor to overflows is the inflow and infiltration of rainwater into sewers, known as rainfall derived inflow and infiltration (RDII). The inflow portion of RDII is
defined by the EPA as the water, other than sanitary wastewater, that enters sewer systems directly from sources such as downspout connections, foundation drains, manhole covers, cross connections between storm sewers and sanitary sewers, and catch basins (EPA, 2014). This water typically enters the sewer system quickly as a rain event begins and results in a rapid increase in the flow rate in the system. This differs from infiltration which is typically slower and lasts for a longer time after the rain event due to a longer pathway through the soil. The EPA defines infiltration as the water, other than sanitary wastewater, that enters the sewer system from the ground through defective pipes, pipe joints, and connections (EPA, 2014). Inflow and infiltration together make up the entire volume of RDII and during storm events RDII may be a significant portion of the entire flow in the system. By increasing flow, RDII directly leads to overflows that can result in citations and fines from regulators or possibly require expensive upgrades to plants and sewer infrastructure to handle the increased volume.

Solving this problem can come at a significant cost. To prevent overflows, Kansas City has committed to invest $2.5 billion dollars of 25 years to infrastructure improvements (Whitley, 2010). Solutions include costly upgrades to treatment plants, rehabilitation of sewer systems, and an increase in stormwater best management practices such as green infrastructure. Milwaukee Metropolitan Sewerage District (MMSD) increased storage capacity within their system by constructing the deep tunnel system, which consists of 28.5 miles of pipe up to 32 ft in diameter having a storage capacity of 521 million gallons. The deep tunnel system took 30 years to build and cost $1.25 billion (Milwaukee Metropolitan Sewerage District, 2010). MMSD has also committed to capture 740 million gallons of stormwater in green infrastructure by 2035 to prevent overflows at an estimate cost of $1.3 billion (CH2MHILL, 2013). An analysis by the Madison Metropolitan Sewerage District found that just pumping the excess water due to RDII was $235,000 in 2010 (Madison Metropolitan Sewerage District, 2010).
2.3 Solution to Sewer Overflows and Basement Backups

Because of the extensive cost of solving sewer overflows and basement backups, it is critical that municipalities have an accurate understanding of the sources, behavior, and variables that contribute to RDII. Understanding RDII requires accurate sanitary sewer pipe data, flow rate data at critical locations within the system, and models that can elucidate the sources of inflow and infiltration. Doing so can help to determine sources of RDII in order to locate areas of concern such as direct connections and defective pipes needing repair. Once this is understood, infrastructure improvements can be targeted to reduce inflow and infiltration into the system. A study in Columbus, Ohio tested 111 private properties by flooding all possible contributing sources of RDII and found that 98% of measured RDII could be accounted for by inflow from direct connections of downspouts and infiltration through sewer laterals (Pawlowski et. al., 2014). In addition, a study in Germany found a reduction in infiltration of 23.9% and a reduction in inflow of 35.7% with rehabilitation of the sanitary sewer system (Staufer et. al., 2012). The rehabilitation of sanitary sewer systems to address RDII can greatly reduce flows to wastewater treatment plants. Unfortunately, if these measures are not successful, it may lead to expensive upgrades to water reclamation plants that can cost billions of dollars (Sangree, 2014).

2.4 Existing Models of Rainfall Derived Inflow and Infiltration

To determine sources of inflow and infiltration and focus remediation efforts, wastewater treatment plants often rely on models of RDII. There are generally 7 categories of RDII models (Bennett et al., 1999; Crawford et al., 1999):

1. The constant unit rate method. This model uses characteristics of the sewershed to determine a constant rate of inflow and infiltration depending on independent variables. These usually take the form of gallon of RDII per acre of foot of pipe.
These rates can vary from sewershed to sewershed depending on variables such as age of infrastructure or construction practices.

2. The percentage of rainfall volume (R-value) method assumes a constant percentage of rainfall will become RDII and applies this percentage to every storm.

3. The percentage of stream flow method. This model is similar to the rainfall volume model but instead of relating sewer flow to rainfall, it relates sewer flow to streamflow in adjacent streams. This approach assumes that antecedent moisture conditions have an impact on RDII and that the streamflow will account for this. Deriving this relationship is data intensive and requires sewerflow and streamflow data as well as significant streamflow monitoring to be applied to new areas.

4. The synthetic unit hydrograph (SUH) method. This method assumes that RDII responds to rainfall events similar to how stormwater runoff does, with both a volume and duration response. The synthetic unit hydrograph derives a triangular shape from a percentage of the rainfall becoming RDII and a time for that water to flow through the system related to sewershed characteristics determining the time of peak flow.

5. The probabilistic method. This method assumes that similar storms will act in similar ways and relies on frequency analysis to predict RDII. Data of storms and responses are analyzed to determine recurrence intervals. Storms are then classified by their recurrence interval, much like streamflow recurrence intervals, and sewerflow response is based on that classification.

6. The rainfall/sewer flow regression method. This method analyzes historical rainfall and sewer flow monitoring data to calculate a regression equation that allows for
sewerflows to be calculated from rainfall. The regression usually varies throughout the year and different regressions need to be developed for this variance.

7. The synthetic stream flow regression method. This method is rarely applied and requires a calibrated watershed runoff model to calculate RDII from synthetic streamflow. Watershed runoff models are developed using multiple linear regression to relate hydrologic and sewerflow responses to rainfall. According to Bennett et al. (1999) and Vallabhaneni and Burgess (2007), this has been successfully applied in Milwaukee for sewer system improvement planning.

EPA (2008) evaluated all of the models listed above against long term rainfall and sewer flow data to determine their accuracy in predicting variability in events. The data from Bennett et al. (1999) and Crawford et al. (1999) on these RDII models informed the EPA on the development of the EPA Sanitary Sewer Overflow Analysis and Planning (SSSOAP) tool. The EPA SSOAP tool utilizes the synthetic unit hydrograph method for modeling RDII as had the previously developed EPA Storm Water Management Model (SWMM) (EPA, 2008). The synthetic unit hydrograph was selected for this tool because it was found to be the most accurate method at predicting peak flows and event volumes and also required less initial data to develop accurate parameters. This was important to be able to model storms outside of the calibration period so that the EPA SSOAP tool would be useful for simulating future storms and evaluating the impacts on RDII.

In addition to its adoption in several hydrologic models, such as EPA SSOAP and EPA SWMM, the RTK synthetic unit hydrograph method is widely used in practice to study rainfall derived inflow and infiltration. For example, Nasrin et al. (2017) used the RTK unit hydrograph method with the EPA SSOAP tool in Melbourne, Australia to quantify RDII in an analysis of sewersheds to reduce sanitary sewer overflows. Zhang et al. (2018) used the RTK synthetic unit
hydrograph method to model RDII in China related to light, medium, and heavy rainfall events and demonstrated that the method is accurate and easily applied. Siegrist et al. (2016) used the RTK method in EPA SWMM to evaluate RDII in 228 sewersheds in Cincinnati, OH to assist the city in studying and reducing sanitary sewer overflows. Due to the benefits of the RTK approach for estimating inflow and infiltration and its wide use in the literature, this is the approach that is adopted in this thesis for estimating inflow and infiltration into sanitary sewers.

2.5 Impact of Residential Design on Inflow and Infiltration

Much of the literature on inflow and infiltration is focused on the specific sources of inflow and infiltration (Pawlowski et. al., 2014, Gheith et al., 2017), modeling of these systems (Carrico et al., 2017), and remediation efforts (Staufer et. al., 2012). From this, we know that the primary pathways for inflow and infiltration are inflow from direct connections into the sanitary sewer systems, such as from foundation drains or sump pumps, or infiltration into sanitary sewer laterals or mains due to cracks in the pipe or disjointed connections. This knowledge allows for a targeted approach to reducing RDII within sewersheds. Pawlowski et al. demonstrated that a simple logistic model could correctly identify 62% of RDII contributing properties and if applied to the entire sewer district in the study, would result in a savings of $1.5 million over manual testing of each property (Pawlowski et. al., 2014). A greater understanding of the design factors that predict RDII contribution within sewersheds could increase the effectiveness of these models and further increase the benefits of these models.

2.6 Impacts of Stormwater Management on Inflow and Infiltration

Understanding how residential hydrology impacts inflow and infiltration is also important for developing stormwater management solutions in urban areas. In addition to inflow and infiltration, large storm events cause excess runoff from urban areas that have negative impacts on flooding and water quality in receiving water bodies. As a response, many
municipalities are promoting green infrastructure as a way to capture, treat, and infiltrate water at the source. This is an effective way to reduce runoff volumes and improve runoff water quality (Clary et al. 2017). Green infrastructure has also been shown to reduce both peak discharge and total storm flow in storm sewers (Jarden et al., 2016). However, green infrastructure may have unintended negative consequences for sanitary sewer systems. By infiltrating water into the ground, green infrastructure may increase infiltration into sanitary sewer systems by either direct infiltration or raising the groundwater table. This is a concern as green infrastructure is increasingly utilized to address the impacts of climate change and improve resilience in the face of unexpected climate disturbances (Foster et al., 2011, CH2MHILL, 2013). It is important to understand the interactions of sanitary sewers with infrastructure to better plan and place green infrastructure and account for the total costs of these systems.

While some of the primary pathways of inflow and infiltration are known, it is not clear how neighborhood scale hydrologic characteristics impact inflow and infiltration. Within the sewershed of a water reclamation plant, the neighborhood designs can be as diverse as the city itself. There may be areas with dense residential and commercial buildings that leave little room for green space, other medium density residential areas with homes that have roof drains, cemented driveways, and curbs and gutters, as well as low density residential areas with large yards and grass swale stormwater collection systems. Each of these locations may drain to a single treatment plant and have similar wastewater infrastructure; however, the hydrologic characteristics of those locations are vastly different, which could impact the volume, timing, and pathways of inflow and infiltration into a sanitary sewer system.
2.7 Need for Research

While there is significant literature on sources, remediation strategies, and modeling approaches to inflow and infiltration, there has been less focus on the influence that residential hydrology has on inflow and infiltration across sewersheds. This thesis seeks to fill this gap by presenting an empirical study that evaluates the specific hydrological characteristics that impacts RDII. The goal is to determine readily observable spatial characteristics of a sewershed that can serve as indicators of potential RDII so that managers can make informed decisions on where to concentrate efforts to reduce RDII.
3. METHODOLOGY

3.1 Data Collection

The sewersheds studied were located throughout the MMSD service area, which includes much of the Milwaukee metro area and surrounding suburbs. MMSD has two distinct regions within their service area due to the historical nature of the city: combined and separate sewer systems. The combined sewer system is primarily the area closest to the downtown area of Milwaukee and is the oldest. This area has all storm water routed to the sanitary sewer system and then to the wastewater treatment plant. This area was excluded from the study as the inflow of all stormwater during every storm would overwhelm and hide the impact of stormwater infiltration into the sanitary sewer system. The separate sewer system has separate pipes for stormwater and wastewater. Utilizing sewersheds in the separate sewer system areas allowed for a more accurate view of the illicit stormwater entering the system, the specific interest of this study. The separate sewer area is where we selected all 19 of our sewersheds as shown in Figure 1. These gages were selected based upon the availability of the data and their locations in diverse geographic and hydrologic areas.
Figure 1. Map of all sewersheds utilized in this study and corresponding flow gages and rain gages. All gages are installed and operated by Milwaukee Metropolitan Sewerage District.
Since a goal of the research project is to evaluate the inflow and infiltration across sewersheds with various hydrologic characteristics, sewersheds were selected to ensure a broad range in hydrologic characteristics including neighborhood runoff management approach (i.e., curb and gutter versus grass swales), land use, and imperviousness. Figure 2 shows images from two of the sewersheds and provides an example of a sewershed with curb and gutter and significant pavement from driveways and sidewalks (left) and an example of a sewershed with large yards, less pavements, and grass swale ditches to convey runoff (right).

![Figure 2. Images from sewershed MS0503 (left) and sewershed MS0454 (right) taken from Google Maps](image)

This research utilized long term sewer flow data gathered by MMSD with flow meter probes spread throughout the MMSD service area. Specifically, continuous 60-min flow rate data was obtained from 19 gages in separate sanitary sewer systems that spanned over the years 2015 to 2019. All readings from the gages were collected in terms of million gallons per day and downloaded in a Microsoft Excel spreadsheet format, which is compatible for importing directly into the EPA SSOAP Model software.

Rainfall data was also collected by a network of automatic rainfall gauges operated by MMSD with locations shown in Figure 1. Each sewershed was paired with a rain gauge based on geographic proximity to the sewershed centroid. In some cases, a single rain gauge was paired with multiple sewersheds due to the distribution of infrastructure. The rainfall data that was
received from MMSD consisted of cumulative rainfall for every hour of the day, with a reset in the cumulative rainfall each day. This format was incompatible with the EPA SSOAP Model software, which requires rainfall data to be cumulative over only the previous hour. To achieve this format, the entire data set was recalculated by subtracting the prior hours cumulative total from the current hour except for the midnight hour due to the reset. This gave us an hourly cumulative rainfall that could be directly imported into the EPA SSOAP Model.

3.2 GIS Data Development

Once sewersheds were selected, ESRI’s ArcMap was used to obtain spatial parameters for each sewershed. This includes sewershed area, land cover type, imperviousness, parcels, road length, tax assessments, and pipe characteristics. Land cover and impervious data from 2011 – the most up to date national level dataset at the time – was obtained from the National Land Cover Dataset, which is publicly available from the USGS. The national land cover data set consists of a raster of 30m resolution and is divided up into 16 classes of land cover. Summaries of land cover and impervious areas were computed in ArcMap using zonal statistics. Parcel, road, and assessment data was obtained from Milwaukee County (MCLIO, 2020). From this data, parcel counts and road lengths were obtained.

Pipe data proved to be the most difficult to obtain, format, and summarize. Large interceptor pipe data was obtained from MMSD; however, most pipe data is proprietary to each individual municipality. In our case there were 14 municipalities and these had to be individually requested from Bayside, Brown Deer, Cudahy, Glendale, Greendale, Greenfield, Heles Corners, Milwaukee Oak Creek, River Hills, Shorewood, Fox Point, Wauwatosa, and West Allis. Once obtained, all pipe data had to be reformatted to obtain consistent attributes including length, type, material, and slopes. Once combined and formatted, ArcGIS was used to develop
attributes for each sewershed including pipe length and length of specific pipe materials (e.g. linear feet of reinforced concrete pipe).

3.3 RTK Parameter Development

The EPA SSOAP Model was used to quantify the inflow and infiltration into the sanitary sewer system for every storm over 4.5 years at each sewershed. The EPA SSOAP Model allows for rainfall and sewer flow data to be analyzed jointly and broken down into readily comparable attributes. The model does this by first recognizing patterns of flow under different rainfall scenarios and averaging flow to develop baselines. The model then breaks the data down into two main time periods that cover dry weather and wet weather flows.

Dry weather flows are identified as times when the previous 7 days had no reported rainfall. The number of days determined to be dry weather days in this analysis varied between sewersheds but ranged from about 100 to 150 days over the duration of the study. The model averages the flow rate data from these days to generate a baseline dry weather flow pattern. This pattern is further broken down into a weekday and weekend flow to account for the differences in behavior of the system throughout the week. Figure 3 shows an example of the dry weather patterns during the weekday and weekend from one of the 19 sewersheds in this study.
The differences between weekdays and weekends typically arise from people waking later in the morning and staying up later into the evening. As illustrated, the pink line, depicting weekends, begins to rise around 7 am and peaks around 12 noon. This differs from the blue line, depicting weekdays, where the flow begins to rise earlier around 5 am and peaks around 8 am.

During the weekend there is also a shift from industrial wastewater production to more residential sources because people stay home from work. This can lead predominantly residential areas to experience higher flows on weekends because people are at home as opposed to at work. This is illustrated in Figure 3, where the weekend peak flows are larger than the weekday peak flows due to the dominant residential land use in the sewershed. It is

Figure 3. Example of weekday and weekend flows for sewershed DC066E. The pink line represents the weekend flows while the blue line represents the weekday flows. This is reflective of the different water use behaviors of residents on weekdays versus weekends.
therefore necessary to account for this change in flow patterns when trying to determine the volume of inflow and infiltration during rain events.

Determining the average dry weather flow also helps determine the timing and magnitude of the average nightly minimum flow. As illustrated the nightly minimum flow typically occurs around 4 in the morning on the weekday and 5 am on the weekends when the vast majority of people are asleep, and the use of water is greatly diminished. Typically, in residential areas, 90% of nightly minimum flow can be attributed to groundwater inflow while the remaining is wastewater (Lai et al., 2007).

After the determination of these dry weather flow characteristics, the next step is to subtract both daily average dry weather flow and average nightly minimum flow from the raw sewer flow data. This then produces an estimate of the flow attributed to rainfall derived inflow and infiltration. During dry weather conditions, subtracting the average baseflow from the raw data typically produces a flow rate in the system that hovers around zero. However, following rain events, the flow rate can increase to many times the flow rate of average dry weather flow due to inflow and infiltration driven by precipitation. This estimation of inflow and infiltration is the data that we are interested in analyzing in this study.

EPA SSOAP was used to perform the steps above to identify wet weather flow events and determine the magnitude, timing, and percentage of flows that are attributed to different types of inflow and infiltration. Before doing so, there are many issues with data structures and model parameters that had to be addressed. The EPA SSOAP model identifies a storm event as beginning at the point where rainfall derived flow rises above zero due to rainfall of over half an inch and ending at the point where the rainfall derived flow drops back to zero. However, in many cases the time in which the rainfall derived flow reaches zero can extend multiple days or weeks due to the influence of infiltration from a higher water table. In wet times of the year,
such as spring in Wisconsin, another rainfall event could occur before the rainfall derived flow from the previous event reaches zero. Therefore, significant manual adjustments had to be made over the period of record (4 years) for each gage.

Manual identification of storm events was carried out by going through all rainfall derived flow data and rainfall data to adjust when the initial flow begins in relation to the storm and when the slope of the flow approached zero. This slight adjustment to the identification of storms within SSOAP resulted in several more storms being identified and allowed for smaller frequent runoff events to be properly recognized in the model.

The manual identification method also captured seasonal trends in storms that were unaccounted for with the automated method of storm identification. The most significant seasonal trend was the appearance of increased flow events due to the presumed presence of snowfall present on the sewershed that was unmeasured by the rain gage, yet contributed significantly to the inflow and infiltration volume due to melting during rain on snow events. Many inflow and infiltration events were found in January and February without any connection to rainfall. Looking at temperature data during these times we connected these storm-like flows to increased temperature days. We also found that throughout the spring the fraction of water determined to be rainfall derived flow and infiltration was higher in the winter and early spring than during the rest of the year. Again, this increase was believed to be connected to the melting of snow which is not measured by rain gages. Therefore, during the winter and spring months the water balance in the model was inaccurate. To overcome this limitation, we decided to focus on inflow and infiltration during June through October when snowfall is not likely to be present and we can accurately predict the total precipitation that contributes to inflow and infiltration.
The recognition of these seasonal differences differs from a previous unpublished analysis of inflow and infiltration in sewersheds in Milwaukee conducted by an engineering firm (Arcadis, 2019). They previously studied rainfall derived inflow and infiltration and relied on a single set of flow parameters for each sewershed to model that sewershed’s flow, regardless of season. Our analysis indicates that in areas like Milwaukee, where snowfall can remain on the ground and contribute to flow during later storms, it is necessary to determine different parameters for different circumstances present in the sewershed.

An additional outcome from the manual identification of storms was that some of our initial sewersheds had significantly more sewer flow measured than the entire rainfall recorded over the entire sewershed. In these cases, we were able to determine that the mapped area of the sewershed misrepresented the actual drainage area of the sewershed and without full knowledge of every pipes flow direction and elevation it wasn’t possible to remap these sewersheds accurately. This affected 2 of our initial sewersheds and these were subsequently removed from our analysis.

With all storms identified, dry weather parameters computed, and seasonal influences removed, the next step was to use EPA SSOAP to determine the sources of inflow and infiltration into the sanitary sewer system. EPA SSOAP does this with a synthetic unit hydrograph. A synthetic unit hydrograph in its simplest form is a triangular shape that models how a unit of water flows through a system. For example, if an inch of rainfall falls over a sewershed, the water will flow towards the outfall but not all the water will get to the end at exactly the same time. Some of the water will arrive very quickly through direct connections of roof or foundation drains, and other water will get there slowly through infiltration of precipitation into the ground and into the sanitary sewer system through cracks in pipes. The synthetic unit hydrograph typically represents these flows in the sanitary sewer system as a
triangular shape with a quick rise and slow release. However, to properly account for different sources of flows into the system, the EPA SSOAP model further breaks down the hydrograph into three separate synthetic unit hydrographs, representing fast, medium, and slow flow, and adds them together to model the flow as demonstrated in Figure 4. When a calibrated synthetic unit hydrograph is applied to rainfall it can approximate the flow in the system and provide insight into how the system works.

EPA SSOAP utilizes the RTK method to develop each of the three synthetic unit hydrographs. The synthetic unit hydrographs are each defined by three parameters: R, T, and K, which is where the name of this method derives from. R represents the percentage of each unit of rainfall that ends up as flow within the sewer system, effectively the area under the curve. T represents the time to peak flow for that unit of water. Finally, K represents a ratio of time to recession and defines when the flow ends. By utilizing 3 separate synthetic unit hydrographs, each with a different set of R, T and K values, it is possible to approximate the shape of the
actual storm flow response and gain insight into what makes up the flow in that sewer system
due to the breakdown of fast (R1), medium (R2) and slow (R3) flow responses.

Using the EPA SSOAP model, we performed manual RTK curve fitting and defined all RTK
parameters for each storm. In total, 1,632 total storm events were identified and RTK
parameters were fit to each storm. This required manually adjusting 9 variables, R, T, and K for 3
separate synthetic unit hydrographs, to adjust the total flow and timing of flow to best match
the calculated rainfall derived inflow and infiltration. Adjusting these 9 variables required a
visual comparison of RTK curves to measured storm curves. R values were adjusted to match the
general magnitude of the hydrograph. T and K values were adjusted to shape the RTK curve to
the storm in typically ranges shown in Table 1 (Lai et al., 2007).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Range (hours)</th>
<th>Variable</th>
<th>Range (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T_1</td>
<td>0.5 - 2.0</td>
<td>K_1</td>
<td>1.0 - 2.0</td>
</tr>
<tr>
<td>T_2</td>
<td>3.0 - 5.0</td>
<td>K_2</td>
<td>2.0 - 3.0</td>
</tr>
<tr>
<td>T_3</td>
<td>10.0 - 15.0</td>
<td>K_3</td>
<td>3.0 - 7.0</td>
</tr>
</tbody>
</table>

The modeled curve was a summation of these three individual synthetic unit
hydrographs representing the short-, medium- and long-term impact on the sewer system from
the rain event. The short-term impact typically can be thought of as only rainfall derived inflow.
Medium-term is a combination of rainfall derived inflow and infiltration while the long-term
impact generally consists of only infiltration. Next, summary statistics (mean, median, and
standard deviation) of the RTK values across all storms were developed each gage. These
summary statistics of RTK values were what was used as dependent variables in the regression
analysis described in the next section.
There are several limitations to the method that we used here. One limitation is that the fitting of the RTK unit hydrograph requires user input to develop a unit hydrograph that reduces model error by manually changing parameters. This calibration step could be impacted by the skill or preferences of the modeler and may produce inconsistent results between modelers. However, in this case all RTK unit hydrographs were developed by the primary author. Another limitation is that this approach could be considered an overfitting of the model parameters since they are adjusted to each individual storm. An alternative method would be to fit a single parameter for each gage that minimizes the error across all storms, but this would be subject to the same user errors and would not allow an analysis across seasons, which was a goal of this study.

3.4 Statistical Regression Methods

Once hydrologic characteristics and RTK summaries were defined for each gage, simple linear regression was performed using JMP software. JMP was chosen for ease of use in managing the large number of parameters developed and good visualization of data. For the regression, summary statistics of the RTK parameters (i.e. median R1, median R2, median R3) were used as dependent variables, and hydrologic characteristics from GIS analysis (e.g. land over type, pipe characteristics, etc.) were used as the independent variables. Goodness of fit was evaluated using $R^2$. To evaluate whether the regression has statistical significance, the hypothesis test for whether the slope differs from zero was performed (Helsel and Hirsch, 2020).
The multicollinearity of the independent variables was also evaluated using Pearson product-moment correlation.

Finally, multivariable linear regression was performed to develop equations that could predict the inflow and infiltration based upon multiple sewershed characteristics (Equation 1).

\[ y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k \]  (Equation 1)

where \( y \) is the independent variable (i.e. inflow and infiltration), \( \beta \) represents the regression coefficients, and \( x \) represents the dependent variables (i.e., sewershed characteristics). The models were evaluated using goodness of fit metrics include \( R^2 \), adjusted \( R^2 \), and root mean square error (RMSE).
4. RESULTS

4.1 Sewershed Characteristics

4.1.1 Sewershed characteristics

The sewersheds in this study had a wide range of physical characteristics, thereby providing a diversity of sewershed types by which to evaluate. For example, Figure 6 represents the average imperviousness of each sewershed, which ranged between 16-51%, as well as the distribution of medium intensity land uses, identified in the National Land Cover Database, which ranged between 3 – 42%. The National Land Cover Database defines medium intensity land use as areas with a mixture of constructed materials and vegetation, and where impervious surfaces account for 50% to 79% of total area and typically consist of single-family housing units (MRLC, 2011). Table 2 provides all sewershed characteristics along with summary statistics.
Table 2. Sewershed characteristics and summary statistics for all sewersheds

<table>
<thead>
<tr>
<th></th>
<th>Average</th>
<th>Standard Deviation</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area</td>
<td>761.4</td>
<td>465.9</td>
<td>1756.5</td>
<td>106.4</td>
</tr>
<tr>
<td>Mean Elevation</td>
<td>702.61</td>
<td>48.59</td>
<td>801.07</td>
<td>635.82</td>
</tr>
<tr>
<td>Mean Slope</td>
<td>4.5</td>
<td>1.7</td>
<td>8.03</td>
<td>1.08</td>
</tr>
<tr>
<td>Imperviousness</td>
<td>36.18</td>
<td>10.77</td>
<td>51.44</td>
<td>15.54</td>
</tr>
<tr>
<td>Parcels per Acre</td>
<td>2.49</td>
<td>1.37</td>
<td>5.45</td>
<td>0.50</td>
</tr>
<tr>
<td>Pipe length</td>
<td>129453</td>
<td>87258</td>
<td>291734</td>
<td>23665</td>
</tr>
<tr>
<td>Pipe length per acre</td>
<td>177.0</td>
<td>68.4</td>
<td>390.6</td>
<td>96.7</td>
</tr>
<tr>
<td>Number of parcels</td>
<td>1795</td>
<td>1232</td>
<td>4133</td>
<td>53</td>
</tr>
<tr>
<td>Open Space</td>
<td>17.06</td>
<td>12.05</td>
<td>51.92</td>
<td>0.06</td>
</tr>
<tr>
<td>Low Intensity</td>
<td>47.40</td>
<td>12.56</td>
<td>68.79</td>
<td>26.57</td>
</tr>
<tr>
<td>Medium Intensity</td>
<td>20.48</td>
<td>12.28</td>
<td>42.19</td>
<td>2.96</td>
</tr>
<tr>
<td>High Intensity</td>
<td>5.21</td>
<td>4.90</td>
<td>16.11</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Figure 5. Percent Imperviousness (left) and medium intensity land use (right) of sewersheds in the case study area
4.1.2 Sanitary sewer pipe characteristics

In addition to sewershed characteristics, there were some municipalities that had pipe material data that were used to evaluate the pipe material types. In Figures 9 and 10 the pipe materials are shown as a percent of the system and as total linear feet for 19 sewersheds. As illustrated, many of the pipe systems are a composite of PVC, concrete, clay, and other materials. In addition, for several of the sewershed there is a significant portion of the pipes – sometimes well over 50% – that is unknown. While this data provides us with a large sample of
the pipes within each network, it does not capture all of the pipes and therefore there may be some uncertainty within this data.

Figure 7. Pipe materials as a percentage of the overall pipe length in sewersheds for which data was available

Figure 8. Pipe materials in linear feet in sewersheds for which data was available
4.2 Inflow and infiltration based upon RTK analysis

Analyzing the rainfall and sewer flow data with EPA SSOAP resulted in the identification of 1,632 total storm events across the 19 sewersheds. These storms were then modeled with the RTK synthetic unit hydrograph method to develop parameters for each storm.

4.2.1 General inflow and infiltration trends

Inflow and infiltration was evaluated in EPA SSOAP using flow rate data over 4 years (2015-2019), and there were several general trends that were found. The median inflow and infiltration (R) across the sewersheds was 0.135, which indicates that almost 14% of rainfall is being infiltrated into the sanitary sewer system. While this percentage may seem to be low, the volume of inflow and infiltration can still be significant as inflow and infiltration volumes were found to exceed 25 million gallons in some cases. This median R value ranges quite significantly between 0.034 and 0.312 (Figure 9). As illustrated, inflow and infiltration seems to be greatest in the northern portion and southwest portion of the Milwaukee metro area.
There were also seasonal trends in the inflow and infiltration data, with more inflow and infiltration occurring during the winter months. Figure 10 represents the inflow and infiltration for each week of the year. From this figure it is clear that the inflow and infiltration increases during the winter, as represented by the weeks at the end and beginning of the year. The seasonal relationship between total R and rainfall was explored to see if perhaps rainfall volume could explain the percent of rainfall that becomes inflow and infiltration. However, as illustrated in Figure 11, there does not appear to be any relationship between rainfall volume and R for these selected sites.
Figure 10. Example of inflow and infiltration (R) over the course of a year from a selection of eight sites. As illustrated, many sites have greater R values during the winter months.

Figure 11. Total R and rainfall volume over the course of a year from a selection of eight sites. As illustrated, many sites have greater R values during winter months; however, this does not seem to correlate with rainfall volumes.
The seasonal trend in R is further illustrated in Figure 12, which illustrates the distribution of Total R across all sites for each month of the year. As illustrated, the inflow and infiltration is generally highest in early winter and mid-spring. This could be due to rain on snow events in which the snow on the ground is not represented within the precipitation data yet contributes significant volume to the inflow and infiltration volume. Because the rainfall data cannot capture this unknown snowmelt volume, we decided to restrict our analysis to storms that occurred in Jun – Oct, when it is unlikely to have snowpack on the ground that could melt and contribute to the inflow and infiltration volume. Doing so reduces the median total inflow and infiltration as represented in the right image in Figure 12.

We also explored the annual trends in the inflow and infiltration to see if perhaps it was changing over time. As illustrated in Figure 13, the median inflow and infiltration appears to be increasing slightly over time with the lowest median inflow and infiltration occurring in 2015 and highest in 2019. This could be due to several factors including an increase in precipitation patterns, aging of the sewer system, or other variables. While not an objective of this thesis, this could be a potentially important finding to explore further.

Figure 12. Distribution of total R across all sites for each month of the year (left) and distribution of the median R across all sewersheds considering the entire year in blue and only the months Jun - Oct in red (right)
4.2.2 Components of inflow and infiltration

Different components of inflow and infiltration – R1, R2, and R3 – were computed for each storm event and a summary of those components across all watersheds is illustrated in Figure 14. As illustrated, the most significant portion of inflow and infiltration comes from R3 (median of 0.03 in Jun-Oct data) followed by R2 and R1 (medians of 0.01 in Jun – Oct data).

The components of R also had seasonal trends that mirrored R, with greater amounts of inflow and infiltration during the winter months (Figure 14a). Because the winter storms were often rain on snow events, the ratio of inflow and infiltration to rainfall volume is skewed high due to the unaccounted-for snow volume. Therefore, to maintain consistency in the analysis and ensure that the water balance across storms is appropriately captured, only the months of June
– October when snowfall is not present were considered. Similar to the total R, this noticeably affects the distribution of median R1, R2, and R3 values as shown in Figure 14b.

![Figure 14. Values of R1, R2, and R3 over the course of a calendar year (a) and a comparison of median annual and Jun-Oct R1, R2, and R3 values across all sewersheds (b)](image)

### 4.3 Regression Results

#### 4.3.1 Relationship between total R and watershed characteristics

Linear regression was performed to predict the average inflow and infiltration (R) between June and October for each sewershed based upon the sewershed characteristics. Table 1 presents the results from the linear regression with the $R^2$ value, the standardized slope, and the statistical significance of the slope ($\text{prob} > |t|$). Results found that mean elevation ($R^2 = 0.149$) and low intensity development ($R^2 = 0.175$) had the strongest negative correlation; however, neither of these had a statistically significant slope at $p <= 0.05$. Based upon these results, there could be several reasons why a decrease in elevation or low intensity development would impact inflow and infiltration. It could be that the depth to the water table is closely related to the elevation of the land surface (USGS, 2008), and therefore as the elevation goes down the depth to the water table goes down, providing a greater chance for infiltration of groundwater into sanitary sewer systems. In addition, low intensity development...
may not have as many households contributing flows and therefore less chances for inflow into these systems. However, each relationship had relatively low predictability and none had a statistically significant slope at p < 0.05.

Linear regression was also performed to predict total R based upon the pipe materials in the sanitary sewer system, and linear feet of ductile iron was found to have a statistically significant slope (p < 0.05) and an $R^2$ of 0.79; however, due to the low number of data points (5) the regression had a single point with both high influence and leverage that impacted the results (Appendix A2).

While there was some correlation among predictor variables and total R, the correlations were not strong. Total R is a component of multiple sources of flow: inflow and infiltration. Because EPA SSOAP is able to differentiate between these due to the differences in timing of inflow and infiltration, the next section explores the correlations to different components of total R: R1, R2, and R3.

|                     | Total R   | $R^2$ | Slope | $P>|t|\) |
|---------------------|-----------|-------|-------|----------|
| Mean Elevation      | 0.149     | -0.386| 0.103 |
| Mean Slope          | 0.036     | 0.190 | 0.435 |
| Imperviousness      | 0.026     | -0.160| 0.512 |
| Parcels per acre    | 0.025     | -0.172| 0.482 |
| Pipe length         | 0.003     | 0.056 | 0.819 |
| Pipe length per acre| 0.026     | -0.160| 0.512 |
| Number of parcels   | 0.009     | 0.010 | 0.686 |
| Open space          | 0.001     | 0.028 | 0.910 |
| Low intensity       | 0.175     | -0.419| 0.075 |
| Medium intensity    | 0.011     | -0.104| 0.673 |
| High intensity      | 0.064     | 0.253 | 0.295 |
### Table 4. Linear regression results predicting Total R based upon pipe materials

| Material          | n | R²  | Std Slope | P>|t| |
|-------------------|---|-----|-----------|-----|
| Cast Iron (ft)    | 8 | 0.305 | 0.552     | 0.156 |
| Cast Iron (%)     | 7 | 0.128 | 0.358     | 0.431 |
| Clay (ft)         | 6 | 0.333 | -0.577    | 0.230 |
| Clay (%)          | 6 | 0.606 | -0.779    | 0.068 |
| Concrete (ft)     | 15| 0.001 | -0.031    | 0.917 |
| Concrete (%)      | 15| 0.028 | 0.167     | 0.569 |
| Ductile Iron (ft) | 5 | 0.788 | -0.887    | **0.045** |
| Ductile Iron (%)  | 5 | 0.760 | -0.872    | 0.054 |
| PVC (ft)          | 17| 0.007 | -0.084    | 0.748 |
| PVC (%)           | 17| 0.157 | -0.396    | 0.116 |
| ABS (ft)          | 4 | 0.008 | -0.087    | 0.913 |
| ABS (%)           | 4 | 0.165 | -0.406    | 0.594 |

4.3.2 Relationship between components of R and watershed characteristics

Linear regression was performed to predict the median inflow and infiltration in June – October (R1, R2, and R3) based upon sewershed characteristics. For R1 – representing quick inflows into the system – variables with statistically significant slopes at p < 0.05 included positive correlations with pipe length per acre (R² = 0.245) and number of parcels (R² = 0.209). There could be several reasons for these findings. The number of parcels is directly related to the number of homes that have a sanitary connection to the sanitary sewer system. It may be that in these areas, there are also direct connections from foundation drains or downspout that are contributing to inflow. The pipe length per acre represents the density of pipes within the network, which could also be related to the number of homes with connections.

The regression also found that R2 had a negative relationship with low intensity residential land use (R² = 0.149), although no regression equations for R2 had a slope that was significant at p < 0.05. Finally, R3 was negatively correlated with imperviousness (R² = 0.305), pipe length per acre (R² = 0.308) and medium intensity residential land use (R² = 0.267). In all of these cases, high values reflect areas in which residential homes are densely built, pipe density...
is high, and there is high imperviousness. In these cases, the high imperviousness would increase runoff and decrease the amount of water that is infiltrated into the ground and potentially infiltrated into the sanitary sewer system.

Table 5. Regression results for R1, R2, and R3 for sewershed characteristics

<table>
<thead>
<tr>
<th></th>
<th>R1</th>
<th></th>
<th></th>
<th>R2</th>
<th></th>
<th></th>
<th>R3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R²</td>
<td>Slope</td>
<td>P&gt;</td>
<td>t</td>
<td></td>
<td>R²</td>
<td>Slope</td>
<td>P&gt;</td>
</tr>
<tr>
<td>Mean Elevation</td>
<td>0.063</td>
<td>-0.252</td>
<td>0.300</td>
<td>0.076</td>
<td>-0.276</td>
<td>0.253</td>
<td>0.033</td>
<td>-0.180</td>
</tr>
<tr>
<td>Mean Slope</td>
<td>0.000</td>
<td>-0.010</td>
<td>0.970</td>
<td>0.093</td>
<td>0.305</td>
<td>0.204</td>
<td>0.191</td>
<td>0.437</td>
</tr>
<tr>
<td>Imperviousness</td>
<td>0.101</td>
<td>0.317</td>
<td>0.186</td>
<td>0.079</td>
<td>-0.282</td>
<td>0.242</td>
<td>0.305</td>
<td>-0.553</td>
</tr>
<tr>
<td>Parcels per acre</td>
<td>0.093</td>
<td>0.305</td>
<td>0.205</td>
<td>0.014</td>
<td>-0.116</td>
<td>0.635</td>
<td>0.113</td>
<td>-0.336</td>
</tr>
<tr>
<td>Pipe length</td>
<td>0.137</td>
<td>0.370</td>
<td>0.118</td>
<td>0.002</td>
<td>0.048</td>
<td>0.845</td>
<td>0.009</td>
<td>-0.094</td>
</tr>
<tr>
<td>Pipe length per acre</td>
<td>0.245</td>
<td>0.500</td>
<td><strong>0.031</strong></td>
<td>0.049</td>
<td>-0.220</td>
<td>0.363</td>
<td>0.308</td>
<td>-0.555</td>
</tr>
<tr>
<td>Number of parcels</td>
<td>0.209</td>
<td>0.457</td>
<td><strong>0.049</strong></td>
<td>0.012</td>
<td>0.111</td>
<td>0.651</td>
<td>0.006</td>
<td>-0.077</td>
</tr>
<tr>
<td>Open space</td>
<td>0.135</td>
<td>-0.367</td>
<td>0.121</td>
<td>0.029</td>
<td>0.171</td>
<td>0.484</td>
<td>0.138</td>
<td>0.372</td>
</tr>
<tr>
<td>Low intensity</td>
<td>0.017</td>
<td>-0.130</td>
<td>0.595</td>
<td>0.149</td>
<td>-0.385</td>
<td>0.103</td>
<td>0.264</td>
<td>-0.513</td>
</tr>
<tr>
<td>Medium intensity</td>
<td>0.200</td>
<td>0.447</td>
<td>0.055</td>
<td>0.056</td>
<td>-0.236</td>
<td>0.331</td>
<td>0.267</td>
<td>-0.517</td>
</tr>
<tr>
<td>High intensity</td>
<td>0.037</td>
<td>0.192</td>
<td>0.430</td>
<td>0.021</td>
<td>0.143</td>
<td>0.558</td>
<td>0.009</td>
<td>0.095</td>
</tr>
</tbody>
</table>

We also explored the relationship between R1, R2, and R3 and pipe materials in linear feet and as a percentage of the total pipe length in the sewershed. It was found linear feet of concrete pipe and concrete pipe as a percentage of the total pipe length was positively correlated to R1. It is unclear what the reason for these might be, but it could be that these pipe materials are correlated to other explanatory variables, which is explored in Section 4.3.4. In addition to concrete pipe, there were several other materials that had a relatively high R² with R1, R2, and R3, such as ductile iron; however these had small sample sizes (n) and therefore it is tough to draw conclusions as it is unclear if it is representative of the sample as a whole.
### Table 6. Regression Results for R1, R2, and R3 against pipe characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cast Iron (ft)</th>
<th>Cast Iron (%)</th>
<th>Clay (ft)</th>
<th>Clay (%)</th>
<th>Concrete (ft)</th>
<th>Concrete (%)</th>
<th>Ductile Iron</th>
<th>Ductile Iron (%)</th>
<th>PVC (ft)</th>
<th>PVC (%)</th>
<th>ABS (ft)</th>
<th>ABS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1 n</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>15</td>
<td>15</td>
<td>5</td>
<td>5</td>
<td>17</td>
<td>17</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>R1 R²</td>
<td>0.054</td>
<td>0.041</td>
<td>0.001</td>
<td>0.013</td>
<td>0.202</td>
<td>0.190</td>
<td>0.177</td>
<td>0.246</td>
<td>0.005</td>
<td>0.124</td>
<td>0.001</td>
<td>0.305</td>
</tr>
<tr>
<td>R1 Slope</td>
<td>0.232</td>
<td>0.203</td>
<td>-0.750</td>
<td>-0.875</td>
<td>0.539</td>
<td>0.602</td>
<td>-0.421</td>
<td>-0.497</td>
<td>0.071</td>
<td>-0.352</td>
<td>-0.026</td>
<td>0.448</td>
</tr>
<tr>
<td>R1 P&gt;</td>
<td>t</td>
<td></td>
<td>0.581</td>
<td>0.662</td>
<td>0.090</td>
<td>0.004</td>
<td>0.003</td>
<td>0.023</td>
<td>0.481</td>
<td>0.394</td>
<td>0.167</td>
<td>0.974</td>
</tr>
<tr>
<td>R2 R²</td>
<td>0.313</td>
<td>0.099</td>
<td>0.002</td>
<td>-0.553</td>
<td>0.003</td>
<td>0.023</td>
<td>0.883</td>
<td>0.765</td>
<td>0.006</td>
<td>-0.362</td>
<td>0.025</td>
<td>0.034</td>
</tr>
<tr>
<td>R2 Slope</td>
<td>0.317</td>
<td>0.490</td>
<td>-0.400</td>
<td>-0.553</td>
<td>-0.036</td>
<td>0.132</td>
<td>-0.939</td>
<td>-0.875</td>
<td>-0.076</td>
<td>-0.362</td>
<td>0.157</td>
<td>-0.185</td>
</tr>
<tr>
<td>R2 P&gt;</td>
<td>t</td>
<td></td>
<td>0.150</td>
<td>0.188</td>
<td>0.435</td>
<td>0.255</td>
<td>0.903</td>
<td>0.653</td>
<td>0.018</td>
<td>0.052</td>
<td>0.771</td>
<td>0.153</td>
</tr>
<tr>
<td>R3 R²</td>
<td>0.387</td>
<td>0.188</td>
<td>0.014</td>
<td>-0.393</td>
<td>0.025</td>
<td>0.002</td>
<td>0.118</td>
<td>0.048</td>
<td>0.002</td>
<td>0.327</td>
<td>0.004</td>
<td>0.099</td>
</tr>
<tr>
<td>R3 Slope</td>
<td>0.150</td>
<td>0.435</td>
<td>-0.194</td>
<td>-0.393</td>
<td>-0.377</td>
<td>-0.207</td>
<td>-0.343</td>
<td>-0.219</td>
<td>-0.044</td>
<td>-0.327</td>
<td>-0.063</td>
<td>-0.315</td>
</tr>
<tr>
<td>R3 P&gt;</td>
<td>t</td>
<td></td>
<td>0.100</td>
<td>0.330</td>
<td>0.712</td>
<td>0.441</td>
<td>0.184</td>
<td>0.479</td>
<td>0.572</td>
<td>0.724</td>
<td>0.867</td>
<td>0.937</td>
</tr>
</tbody>
</table>

### 4.3.3 Relationship between R ratios and sewershed characteristics

In addition to absolute values, regression was also performed on normalized inflow and infiltration values by dividing R1, R2, and R3 by the total R (Tables 7 and 8). This fraction allows us to compare inflow and infiltration characteristics across watersheds of various scales more directly by normalizing the components of inflow and infiltration values to the total R. As a whole, these ratios have a higher strength of prediction than the absolute values. This could be because the normalization of the components of inflow and infiltration allow for a more appropriate comparison among sewersheds of different scales.

For normalized R1, variables with statistically significant slopes at p < 0.05 included positive correlations with pipe length per acre ($R^2 = 0.48$), medium intensity residential land use ($R^2 = 0.371$), imperviousness ($R^2 = 0.249$), and number of parcels ($R^2 = 0.218$); and a negative correlation with open space land use ($R^2 = 0.237$). While the absolute values of R1 also had statistically significant correlations with pipe length per acre and the number of parcels, the ratio provides new variables correlated with R1: imperviousness, medium intensity residential, and open space land uses. There could be several reasons for these new relationships. Medium
Intensity residential land use is representative of relatively dense single-family homes, and in these locations, there may therefore be more opportunities for direct connections of foundation or roof drains that contribute to inflow. Conversely, in areas of open space land use, there are no structures connected to the sanitary sewer system and therefore less opportunities for inflow. While imperviousness is correlated, it is less clear why this would be directly connected to inflow other than the fact that impervious areas could be correlated with medium intensity land uses that have more homes and therefore more opportunities for direct connections of foundation or roof drains.

Table 7. Regression results for R1/Total R, R2/Total R, and R3/Total R for sewershed characteristics

<table>
<thead>
<tr>
<th></th>
<th>R1 / Total R</th>
<th></th>
<th></th>
<th>R2 / Total R</th>
<th></th>
<th></th>
<th>R3 / Total R</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R²</td>
<td>Slope</td>
<td>P&gt;</td>
<td>t</td>
<td></td>
<td>R²</td>
<td>Slope</td>
<td>P&gt;</td>
<td>t</td>
</tr>
<tr>
<td>Mean Elevation</td>
<td>0.001</td>
<td>0.036</td>
<td>0.884</td>
<td>0.024</td>
<td>0.155</td>
<td>0.526</td>
<td>0.057</td>
<td>0.239</td>
<td>0.324</td>
</tr>
<tr>
<td>Mean Slope</td>
<td>0.015</td>
<td>-0.121</td>
<td>0.620</td>
<td>0.108</td>
<td>0.329</td>
<td>0.169</td>
<td>0.329</td>
<td>0.573</td>
<td>0.010</td>
</tr>
<tr>
<td>Imperviousness</td>
<td>0.249</td>
<td>0.500</td>
<td>0.030</td>
<td>0.109</td>
<td>-0.330</td>
<td>0.167</td>
<td>0.568</td>
<td>-0.754</td>
<td>0.000</td>
</tr>
<tr>
<td>Parcels per acre</td>
<td>0.322</td>
<td>0.567</td>
<td>0.011</td>
<td>0.013</td>
<td>0.113</td>
<td>0.646</td>
<td>0.135</td>
<td>-0.367</td>
<td>0.122</td>
</tr>
<tr>
<td>Pipe length</td>
<td>0.158</td>
<td>0.397</td>
<td>0.092</td>
<td>0.010</td>
<td>-0.100</td>
<td>0.684</td>
<td>0.047</td>
<td>-0.218</td>
<td>0.371</td>
</tr>
<tr>
<td>Pipe length per acre</td>
<td>0.476</td>
<td>0.690</td>
<td>0.001</td>
<td>0.041</td>
<td>-0.202</td>
<td>0.410</td>
<td>0.608</td>
<td>-0.780</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Number of parcels</td>
<td>0.218</td>
<td>0.467</td>
<td>0.044</td>
<td>0.001</td>
<td>-0.032</td>
<td>0.900</td>
<td>0.058</td>
<td>-0.241</td>
<td>0.320</td>
</tr>
<tr>
<td>Open space</td>
<td>0.237</td>
<td>-0.487</td>
<td>0.035</td>
<td>0.066</td>
<td>0.257</td>
<td>0.288</td>
<td>0.394</td>
<td>0.628</td>
<td>0.004</td>
</tr>
<tr>
<td>Low intensity</td>
<td>0.026</td>
<td>0.162</td>
<td>0.507</td>
<td>0.007</td>
<td>0.085</td>
<td>0.731</td>
<td>0.089</td>
<td>-0.298</td>
<td>0.216</td>
</tr>
<tr>
<td>Medium intensity</td>
<td>0.371</td>
<td>0.609</td>
<td>0.006</td>
<td>0.134</td>
<td>-0.366</td>
<td>0.124</td>
<td>0.614</td>
<td>-0.784</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>High intensity</td>
<td>0.003</td>
<td>0.056</td>
<td>0.819</td>
<td>0.069</td>
<td>-0.262</td>
<td>0.279</td>
<td>0.031</td>
<td>-0.176</td>
<td>0.470</td>
</tr>
</tbody>
</table>

Again, there were no statistically significant parameters correlated with normalized R2.

For normalized R3, variables with statistically significant slopes at p < 0.05 included negative correlations with medium intensity residential land use (R² = 0.614), pipe length per acre (R² = 0.608), and imperviousness (R² = 0.568); and positive correlations with open space (R² = 0.394) and mean slope (R² = 0.329). In this case, a decrease in medium intensity land use and imperviousness could mean that there are less connections to the system, less opportunities for
inflow, and more pervious space for infiltration; therefore, the amount of infiltration relative to the whole will be higher. The reason for a negative correlation of pipe length per acre with R3 is not readily apparent, because more pipes in the ground should provide more opportunities for infiltration. However, it could be due to a strong correlation between pipe density and other explanatory variables, as explored further in the next section. Finally, the increase in infiltration with increases in slope runs contrary to other studies that have found that as the slope of the land surface increases, the runoff volume increases and infiltration decreases (Huang et al., 2013). However, these results may indicate that in areas of higher slope – which typically consist of more hills and valleys – rainfall may be pooled into valleys and depressions where it has more time to infiltrate and raise the groundwater table.

The opposite relationships between R1 and R3 when it comes to several predictors may suggest that these predictors can explain where the fraction of inflow and infiltration as a whole will come from. For parcels per acre, average imperviousness and medium intensity land use, there is positive relationship with R1 (i.e., inflow) and a negative relationship with R3 (i.e., infiltration) for both the absolute and normalized values (Tables 5 and 7). This is illustrated graphically in Figure 16. All three of these parameters are representative of the density of development. As the density increases there are more buildings and therefore more chances for direct connections to the sanitary sewer system through foundation or roof drains. Conversely, as density increases, there is less permeable ground for rainfall to infiltrate and therefore in these areas there may a lower relative volume of water in the ground for infiltration.
We also explored the relationship between normalized R values and pipe characteristics, and it was found that concrete had a statistically significant positive relationship to R1 and negative relationship to R3. This would suggest that for sewershed with concrete pipes, there is an increasing amount of quick inflow into the system and a decreasing amount of slower infiltration. However, it could be that concrete pipes are correlated with other explanatory variables as explored in the following section.
**Table 8. Regression results for R1/Total R, R2/Total R, and R3/Total R against pipe characteristics**

| Material   | n  | R²  | Slope | P>|t| | R²  | Slope | P>|t| | R²  | Slope | P>|t| |
|------------|----|-----|-------|-----|-----|-------|-----|-----|-------|-------|-----|
| Cast Iron (ft) | 8  | 0.046 | -0.215 | 0.609 | 0.002 | 0.044 | 0.919 | 0.075 | 0.274 | 0.511 |
| Cast Iron (%)  | 7  | 0.000 | -0.003 | 0.995 | 0.021 | -0.145 | 0.756 | 0.028 | 0.165 | 0.724 |
| Clay (ft)     | 6  | 0.000 | 0.067 | 0.900 | 0.007 | 0.426 | 0.400 | 0.200 | 0.517 | 0.293 |
| Clay (%)      | 6  | 0.028 | 0.174 | 0.741 | 0.006 | 0.516 | 0.295 | 0.164 | 0.417 | 0.411 |
| Concrete (ft) | 15 | 0.223 | 0.714 | **0.004** | 0.007 | -0.026 | 0.930 | 0.099 | -0.727 | **0.003** |
| Concrete (%)  | 15 | 0.160 | 0.623 | **0.017** | 0.005 | -0.119 | 0.684 | 0.033 | -0.690 | **0.006** |
| Ductile Iron (ft) | 5  | 0.036 | 0.189 | 0.760 | 0.175 | 0.418 | 0.483 | 0.020 | 0.140 | 0.822 |
| Ductile Iron (%) | 5  | 0.011 | 0.103 | 0.870 | 0.231 | 0.481 | 0.412 | 0.065 | 0.256 | 0.677 |
| PVC (ft)      | 17 | 0.010 | 0.098 | 0.710 | 0.001 | -0.034 | 0.900 | 0.002 | 0.044 | 0.870 |
| PVC (%)       | 17 | 0.013 | -0.114 | 0.664 | 0.024 | 0.156 | 0.551 | 0.000 | -0.015 | 0.960 |
| ABS (ft)      | 4  | 0.031 | 0.176 | 0.824 | 0.993 | 0.997 | **0.003** | 0.039 | 0.197 | 0.803 |
| ABS (%)       | 4  | 0.113 | -0.337 | 0.664 | 0.697 | 0.834 | 0.166 | 0.017 | -0.132 | 0.869 |

4.3.4 Multicollinearity among predictor variables

The predictor variables were evaluated for multicollinearity to determine the relationship among predictors. One unclear previous finding was the strong relationship between concrete pipes and inflow and infiltration. From the correlation probability matrix below in Table 9, concrete pipe is correlated with the number of parcels and medium intensity land use (a correlation matrix can be found in Appendix A5). Both number of parcels and medium intensity land use were found to correlate to R1 and R3. In this case, although concrete pipes are correlated to an increase in R1 and a decrease in R3, it may be due to the fact that there are more concrete pipes in these areas, rather than a function of the concrete pipes themselves.
Table 9. Correlation probability among predictor variables represented as the Pearson product-moment correlation coefficient

<table>
<thead>
<tr>
<th></th>
<th>Mean Slope</th>
<th>Imp.</th>
<th>Parcels per acre</th>
<th>Pipe length per acre</th>
<th># of Parcels</th>
<th>Open Space</th>
<th>Low Intensity</th>
<th>Medium Intensity</th>
<th>Conc.(ft)</th>
<th>Conc. (%Lft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Slope</td>
<td>0.000</td>
<td>0.007</td>
<td>0.235</td>
<td>0.003</td>
<td>0.911</td>
<td>0.092</td>
<td>0.016</td>
<td>0.800</td>
<td>0.217</td>
<td>0.434</td>
</tr>
<tr>
<td>Imp. Parcels per acre</td>
<td>0.007</td>
<td>0.000</td>
<td>0.032</td>
<td>0.001</td>
<td>0.014</td>
<td>0.001</td>
<td>0.167</td>
<td>0.000</td>
<td>0.029</td>
<td>0.043</td>
</tr>
<tr>
<td>Pipe length per acre</td>
<td>0.235</td>
<td>0.032</td>
<td>0.000</td>
<td>0.008</td>
<td>0.033</td>
<td>0.016</td>
<td>0.021</td>
<td>0.069</td>
<td>0.878</td>
<td>0.350</td>
</tr>
<tr>
<td># Parcels</td>
<td>0.003</td>
<td>0.001</td>
<td>0.008</td>
<td>0.000</td>
<td>0.225</td>
<td>0.016</td>
<td>0.075</td>
<td>0.001</td>
<td>0.067</td>
<td>0.175</td>
</tr>
<tr>
<td>Open Space</td>
<td>0.092</td>
<td>0.001</td>
<td>0.016</td>
<td>0.016</td>
<td>0.091</td>
<td>0.000</td>
<td>0.032</td>
<td>0.000</td>
<td>0.139</td>
<td>0.083</td>
</tr>
<tr>
<td>Low Intensity</td>
<td>0.016</td>
<td>0.167</td>
<td>0.021</td>
<td>0.075</td>
<td>0.794</td>
<td>0.302</td>
<td>0.000</td>
<td>0.683</td>
<td>0.288</td>
<td>0.249</td>
</tr>
<tr>
<td>Medium Intensity</td>
<td>0.080</td>
<td>0.000</td>
<td>0.069</td>
<td>0.001</td>
<td>0.011</td>
<td>0.000</td>
<td>0.683</td>
<td>0.000</td>
<td>0.004</td>
<td>0.016</td>
</tr>
<tr>
<td>Concrete (ft)</td>
<td>0.217</td>
<td>0.029</td>
<td>0.878</td>
<td>0.067</td>
<td>0.025</td>
<td>0.139</td>
<td>0.288</td>
<td>0.004</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Concrete (%Lft)</td>
<td>0.434</td>
<td>0.043</td>
<td>0.305</td>
<td>0.175</td>
<td>0.002</td>
<td>0.083</td>
<td>0.249</td>
<td>0.016</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

4.3.5 Multivariable linear regression

Forward and backwards stepwise regression was performed to develop multivariable linear regression models to predict inflow and infiltration based upon watershed characteristics. Candidate variables were selected as those that were statistically significant in linear regression model discussed previously. Final selected variables had to have significance of each variable at the p < 0.15 level. Table 10 illustrates the final equations that were developed. As illustrated, for both R1 and R1 / total R the pipe length per acre and number of parcels are significant predictors, with each improving the predictive power of the equation. For R3, low intensity and medium intensity development explain 49% of the variance, while for R3 / total R the medium intensity and pipe length per acre explain 72% of the variance.
### Table 10. Multivariable linear regression results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Equation</th>
<th>R²</th>
<th>Adj R²</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total R</td>
<td>$0.342 + 2.9 \times 10^{-4} \times ME - 0.126 \times LI$</td>
<td>0.269</td>
<td>0.178</td>
<td>0.041</td>
</tr>
<tr>
<td>R1</td>
<td>$1.75 \times 10^{-3} + 3.8 \times 10^{-5} \times PLA + 1.83 \times 10^{-6} \times NP$</td>
<td>0.348</td>
<td>0.266</td>
<td>0.006</td>
</tr>
<tr>
<td>R3</td>
<td>$0.103 - 0.1 \times LI - 0.103 \times MI$</td>
<td>0.491</td>
<td>0.427</td>
<td>0.02</td>
</tr>
<tr>
<td>R1 / total R</td>
<td>$0.014 + 6.2 \times 10^{-4} \times PLA + 1.6 \times 10^{-5} \times NP$</td>
<td>0.547</td>
<td>0.491</td>
<td>0.05</td>
</tr>
<tr>
<td>R3 / total R</td>
<td>$0.794 - 0.001 \times PLA - 0.709 \times MI$</td>
<td>0.719</td>
<td>0.684</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Note: ME = mean elevation; LI = low intensity development; PLA = pipe length (ft) per acre; MI = medium intensity development; NP = number of parcels
5. DISCUSSION

This study found that the amount of inflow into sanitary sewers increases with high imperviousness, medium intensity residential land use, and dense pipe networks. These are most likely direct connections to the system through roof or foundation drains. On the other hand, the amount of infiltration into sanitary sewers increase with more open space, more pipe density, less dense development, and less imperviousness. These results have several implications for water reclamation facilities that are considering how to approach reductions in inflow and infiltration.

These results can be used to target infrastructure efforts for reducing inflow and infiltration. From these results, it is clear that if the goal is to reduce inflow, it would be prudent to target infrastructure improvement efforts in sewersheds that have medium density residential land use, higher imperviousness, and less open space. These actions could include disconnections of foundation drains, roof drains, or illicit connections (Staufer et al., 2012). Jiang et al. (2019) found that disconnecting foundation drains reduced flow volume by a minimum of 78%. However, if slower infiltration is a concern, remediation efforts could focus on areas that are less dense, less impervious, and have more open space. These actions could include replacing or relining cracked pipes and fixing improper joints and connections in the sanitary sewer system (Staufer et al., 2012). Robert Jacobsen (2012) demonstrated a method of relining pipes that led to an average 99% decrease in sanitary sewer system exfiltration. In addition, the negative relationship between infiltration and elevation suggests that these efforts could be targeted to sewersheds that have lower elevations.

For municipalities that may be considering whether to introduce green infrastructure as a stormwater management strategy, a primary argument against it is that it may increase inflow
and infiltration. From these results, we see that for many dense sewersheds the primary contributor to RDII is inflow. If green infrastructure is introduced in these watersheds, an increase in infiltration into the soil may contribute to infiltration into sanitary sewers; however, it is not likely to increase inflow and may in fact decrease inflow by removing stormwater from the storm network that may seep into the sanitary sewer system through cracks or improper connections.

The findings of this study could provide valuable information to municipalities across the country; however, there are several factors to consider in generalizing the data. Milwaukee is a post-industrial Midwest city in which much of the development occurred in the early to mid-20th century. Therefore, much of the infrastructure in place is older and subject to deterioration due to aging and many of the homes built may have foundation or roof connections built prior to codes that discouraged them. In municipalities that serve areas that have developed more recently, the function of the sanitary sewer system and design of built environment may be different. In addition, in regions with different precipitation patterns, ground water levels, and tidal influences, among other variables, the dynamics between precipitation and ground water infiltration may behave differently. Therefore, application of these findings outside of the Milwaukee region should consider these factors.
6. CONCLUSION

6.1 Key Findings

This study found that sewershed characteristics can predict rainfall derived inflow and infiltration in sewersheds in Milwaukee, Wisconsin. The key findings from this study were:

- There is significant variability in sewershed characteristics throughout the Milwaukee area, both in surface characteristics and the infrastructure present. Some sewershed characteristics were related to rainfall derived inflow and infiltration and can act as indicators of where inflow and infiltration may be entering sewer systems.

- Inflow (i.e. R1) is positively related to pipe length per acre, number of parcels, and medium intensity land use. All three variables related to the density of development and are correlated with one another. Medium intensity land use contains a high number of parcels and dense pipe infrastructure in the ground. In these areas, more housing and sewer connections increases the potential for direct connections of roof drains, foundation drains, and sump pumps.

- Infiltration (i.e. R3) is negatively correlated with imperviousness, pipe length per acre, and medium intensity land use. Infiltration occurs when rainfall infiltrates into the ground and then into a sanitary sewer pipe through cracks or improper connections. As medium intensity land use increases, so do impervious surfaces, which provide water with less available surface area to infiltrate.

- Infiltration (i.e. R3) is positively correlated with open space land use and mean slope. Areas of higher mean slope typically consist of more hills and valleys. In these areas, rainfall may be pooled into valleys and depressions where it has more time to infiltrate.
and raise the groundwater table. A higher groundwater table has a better likelihood of infiltrating into sewer pipes and becoming RDII.

- No sewershed or pipe characteristics were statistically significant predictors of R2. This may be because R2 is a middle ground between inflow and infiltration and is made up of both late inflow and early infiltration, thus making it difficult to attribute to a single source type.

- Sewershed characteristics were able to predict the normalized R1, R2, and R3 better than the absolute values. This may be because the normalized values allow for more appropriate comparison of sewersheds across various scales and infiltration characteristics because the normalized R1, R2 and R3 directly compare the proportion of RDII attributed to each source.

- Multivariable linear regression found that pipe length per acre and number of parcels explained 55% of the variability in R1 / total R. Since inflow is cause by direct connections to the sanitary sewer from downspouts and foundations, more pipe length and parcels would drive an increase in R1.

- Multivariable linear regression found that pipe length per acre and medium intensity residential land use explained 72% of the variability in R3 / total R. This reinforces the relationships seen with R3 and indicates that inflow is increased, and infiltration is decreased, in areas of higher density.

6.2 Future Work

Some small differences were found between the sewersheds where all laterals were mapped and known and those where they were not. We know in these areas that sewer laterals do exist so not knowing the true length of pipe and possible infiltration surface may skew our data in final analysis. The best case would be to know all this data by going through extensive as
built files which show exactly what the construction crew installed and where. This process would be extremely time intensive and was unreasonable for this study. Some of the communities are slowly addressing this issue by collecting this data as sewer laterals are fixed or upgraded. Obtaining all this data would allow a more robust analysis and more accurate predictions. If this study is repeated it may be beneficial to select sewersheds based on where all of this data is readily available to avoid these confounding issues in analysis.

Furthermore, while this study provides a useful high-level screening tool for municipalities to target inflow and infiltration efforts, it does not provide detailed information on the specific locations within each sewershed that inflow or infiltration may be occurring. To do so would require more detailed monitoring or modeling studies that can provide specific insights into when, where, and how inflow or infiltration is occurring within the sewershed of a sanitary sewer system. Improvements in this type of monitoring and modeling at a smaller spatial scale would be valuable future work to provide water reclamation managers with powerful tools to address inflow and infiltration reduction efforts.
7. REFERENCES


APPENDIX

A1. Sewershed and pipe characteristics

Figure A 1. Spatial representation of the average slope (left) and average R in Jun - Oct (right) for the sewersheds in this study.
Figure A 2. Landcover classification of Milwaukee region and sewershed locations (black outlines)
The unique colors of the pipes in the figure above correspond to unique data sources.
A2. Linear regression graphs for median R1, R2, R3

Figure A 4. Linear regression of median R1, R2, and R3 for pipe length per area (a); imperviousness (b); open space land use (c); medium intensity land use (d); mean elevation (e); parcels per acre (f)
Figure A5. Linear regression of median R1, R2, and R3 for pipe length (a); low intensity land use (b); high intensity land use (c); mean slope (d); road length per acre sewershed (e)
Figure A 6. Linear regression of median R1, R2, and R3 for cast iron (ft) (a); cast iron (%) (b); clay (ft) (c); clay (%) (d); concrete (ft) (e); concrete (%) (f)
Figure A.7. Linear regression of median R1, R2, and R3 for ductile iron (ft) (a); ductile iron (%) (b); PVC (ft) (c); PVC (%) (d); ABS (ft) (e); ABS (%) (f)
A3. Linear regression graphs of normalized R1, R2, R3

Figure A 8. Linear regression of normalized median R1, R2, and R3 for pipe length (a); area (b); pipe length per area (c); imperviousness(d); number of parcels (e); parcels per acre (f)
Figure A 9. Linear regression of normalized median R1, R2, and R3 for low intensity land use (a); medium intensity land use (b); high intensity land use (c); mean elevation (d); mean slope (e); road length per acre (f)
Figure A 10. Linear regression of normalized median R1, R2, and R3 ABS (ft) (a); ABS (%) (b); cast iron (ft) (c); cast iron (%) (d); clay (ft) (e); clay (%) (f)
Figure A.11. Linear regression of normalized median R1, R2, and R3 concrete (ft) (a); concrete (%) (b); ductile iron (ft) (c); ductile iron (%) (d); PVC (ft) (e); PVC (%) (f).
A4. Precipitation and RDII patterns

Figure A.12. Monthly distribution of rainfall and total R (a); median total R, flow duration (hrs) and rainfall (b); distribution of R1, R2, and R3 for each sewershed (c); distribution of t1, t2, and t3 for each sewershed (d)
Figure A 13. Antecedent rainfall as a function of total R (a); R1 (b); R2 (c); R3 (d)
A5. Multicollinearity

The correlation among predictor variables was evaluated using the Pearson product-moment correlation as shown in Equation A1. Results are shown in Table A1.

\[ r = \frac{\sum(x-x\bar{)}(y-y\bar{))}}{\sqrt{\sum(x-x\bar{)}^2}\sqrt{\sum(y-y\bar{)}^2}} \quad \text{Equation A1} \]

<table>
<thead>
<tr>
<th></th>
<th>Mean Slope</th>
<th>Imp.</th>
<th>Parcels per acre</th>
<th>Pipe length per acre</th>
<th># Parcels</th>
<th>Open Space</th>
<th>Low Intensity</th>
<th>Medium Intensity</th>
<th>Conc.(ft)</th>
<th>Conc. (%Lft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Slope</td>
<td>1</td>
<td>-0.595</td>
<td>-0.286</td>
<td>-0.65</td>
<td>-0.028</td>
<td>0.397</td>
<td>-0.547</td>
<td>-0.412</td>
<td>-0.352</td>
<td>-0.228</td>
</tr>
<tr>
<td>Imp. Parcels per acre</td>
<td>-0.595</td>
<td>1</td>
<td>0.493</td>
<td>0.685</td>
<td>0.552</td>
<td>-0.712</td>
<td>0.33</td>
<td>0.922</td>
<td>0.581</td>
<td>0.546</td>
</tr>
<tr>
<td>Pipe length per acre</td>
<td>-0.286</td>
<td>0.493</td>
<td>1</td>
<td>0.588</td>
<td>0.49</td>
<td>-0.546</td>
<td>0.526</td>
<td>0.427</td>
<td>0.045</td>
<td>0.295</td>
</tr>
<tr>
<td># Parcels</td>
<td>-0.028</td>
<td>0.552</td>
<td>0.49</td>
<td>0.292</td>
<td>1</td>
<td>-0.398</td>
<td>-0.064</td>
<td>0.569</td>
<td>0.596</td>
<td>0.756</td>
</tr>
<tr>
<td>Open Space</td>
<td>0.397</td>
<td>-0.712</td>
<td>-0.546</td>
<td>-0.544</td>
<td>-0.398</td>
<td>1</td>
<td>-0.25</td>
<td>-0.748</td>
<td>-0.416</td>
<td>-0.48</td>
</tr>
<tr>
<td>Low Intensity</td>
<td>-0.547</td>
<td>0.33</td>
<td>0.526</td>
<td>0.419</td>
<td>-0.064</td>
<td>-0.25</td>
<td>1</td>
<td>0.1</td>
<td>-0.306</td>
<td>-0.33</td>
</tr>
<tr>
<td>Medium Intensity</td>
<td>-0.412</td>
<td>0.922</td>
<td>0.427</td>
<td>0.693</td>
<td>0.569</td>
<td>-0.748</td>
<td>0.1</td>
<td>1</td>
<td>0.712</td>
<td>0.631</td>
</tr>
<tr>
<td>Concrete (ft)</td>
<td>-0.352</td>
<td>0.581</td>
<td>0.045</td>
<td>0.503</td>
<td>0.596</td>
<td>-0.416</td>
<td>-0.306</td>
<td>0.712</td>
<td>1</td>
<td>0.884</td>
</tr>
<tr>
<td>Concrete (%Lft)</td>
<td>-0.228</td>
<td>0.546</td>
<td>0.295</td>
<td>0.384</td>
<td>0.756</td>
<td>-0.48</td>
<td>-0.33</td>
<td>0.631</td>
<td>0.884</td>
<td>1</td>
</tr>
</tbody>
</table>