Development of Effective Procedures for Illicit Discharge Risk Mapping in Roanoke, VA and Fairfax County, VA

Paul R. Bender

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Civil Engineering

Randel L. Dymond, Chair

Kevin Young

Kathleen Hancock

May 2nd, 2016

Blacksburg, Virginia

Keywords: Stormwater management, Detection methods, Risk management, Geographic information systems, Nonpoint pollution

Development of Effective Procedures for Illicit Discharge Risk Mapping in Roanoke, VA and Fairfax County, VA

Paul R. Bender

ACADEMIC ABSTRACT

Authorities of municipal separate storm sewer systems (MS4s) are required to address illicit discharges as part of the National Pollutant Discharge Elimination System's (NPDES) stormwater program. Field reconnaissance is an effective measure to detect and identify illicit discharges, but requires substantial staff and financial resources to conduct. While risk analysis techniques and guidelines have been developed to facilitate MS4 prioritization of field operations, neither a standard set of indicators nor a standard operating procedure has been adopted. This study investigates the relationships among indicators of illicit discharge potential (IDP) and the locations of illicit discharges in two Virginia MS4s. Results of the study indicate that certain risk factors are statistically more effective at predicting IDP, suggesting that a core set of factors can be used to map illicit discharge risk. The results also show that risk mapping tools are significantly impacted by uncertainty in model inputs. Recommendations are provided for MS4s interested in pursuing IDP risk mapping as a tool to improve cost-effectiveness and guide illicit discharge program implementation.

Development of Effective Procedures for Illicit Discharge Risk Mapping in Roanoke, VA and Fairfax County, VA

Paul R. Bender

PUBLIC ABSTRACT

Federal and state laws require local authorities to keep stormwater pollution out of streams and other waterbodies. This pollution originates from many different sources, and typically enters local waters through pipes that drain rainfall away from houses, developments, shopping centers, and other developed areas. As the process of inspecting, testing, and eliminating these pollution sources requires significant public funds, it is important that local authorities perform these tasks in an efficient manner. This research utilizes a method known as risk mapping to investigate the relationships between different characteristics, such as land use and age of development, and the locations of pollution sources. The study indicates that certain characteristics may help to determine the location of pollution sources, and that lower quality information describing these characteristics reduces the ability of the risk map to predict where pollution may occur. The study emphasizes the importance of accurate information in managing pollution, and suggests that a risk mapping tool may help authorities to locate areas with a high risk for pollution. These areas can be given top priority for inspection, which provides authorities with a better chance to eliminate more sources of pollution in a quick and efficient manner. In the long term, the benefits to local authorities that make use of the recommendations developed in this study should include improved water quality, improved pollution management, and cost-savings for taxpayers.

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support and guidance of many individuals. First, I would like to thank Dr. Dymond for serving as my committee chair, and for motivating me throughout the research and writing process as I pursued this unique research topic. I also want to thank my other committee members, Kevin Young and Kitty Hancock, for their guidance, insight, and feedback on all aspects of this project.

There are several people outside of Virginia Tech that helped make this research project happen. Special thanks to Christopher Blakeman, Patrick Hogan, and many others at the City of Roanoke for their help obtaining data and information about the City's Illicit Discharge Detection and Elimination program. Special thanks also to James Patteson, Takisha Cannon, Jonathan Okafor, Chad Grupe, Keith Appler, and others at Fairfax County for their help obtaining data and familiarizing me with the County's Illicit Discharge and Improper Disposal program. Thanks also go out to Diana Handy at Arlington County and Lori A. Lilly at Lori A. Lilly Environmental Solutions for their insight and information during the interview process.

I would also like to thank all the members of my research group, and especially Marcus Aguilar, Clay Hodges, Walter McDonald, and Amanda Dritschel, for their candid input and answers to my never-ending barrage of questions and ideas. Last but certainly not least, a special thank you to all of my friends and family for their support and confidence – I couldn't have done this without you all by my side.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

- ArcGIS for Desktop
  - computer software that enables storage and display of data in spatial maps and databases. 27
- BMP
  - Best Management Practice 4
- CCTV
  - Closed Circuit Television 13, 21
- CGP
  - Construction General Permit 128
- COR
  - City of Roanoke 59
- CSO
  - Combined Sewer Overflow 1
- CSS
  - Combined Sewer System 1
- CSV
  - Comma Separated Values 30
- CWA
  - Clean Water Act 3, 20
- CWP
  - Center for Watershed Protection 12
- DEQ
  - Virginia Department of Environmental Quality 4, 20, 26, 128
- DWF
  - Dry Weather Flow 11
- DWS
  - Dry Weather Screening .... *See* ORI
- ECHO
  - Enforcement and Compliance History Online 26, 72, 127
- EPA
  - Environmental Protection Agency 3, 20
- FFX
  - County of Fairfax 59
- FOG
  - Fats, Oils, and Grease 30, 38, 94
- GIS
  - Geographic Information System 14, 22, 27, 53, 123
- HHW
  - Household Hazardous Waste 5
- I/I
  - Inflow and Infiltration 16
- IDDE
  - Illicit Discharge Detection and Elimination . 4, 20
- IDID
  - Illicit Discharge and Improper Disposal .. 4, 20
- IDP
  - Illicit Discharge Potential 15, 26, 40, 52
- IHRR
  - Industrial and High Risk Runoff 5
- MCM
  - Minimum Control Measure 4, 20
- MEP
  - Maximum Extent Practicable 4
- MS4
  - Municipal Separate Storm Sewer System 1, 34, 52, 124
- NHD
  - National Hydrography Dataset 125
- NLCD
  - National Land Cover Dataset 28, 49, 122
- NPDES
  - National Pollutant Discharge Elimination System 2, 3, 19, 54
- OEM
  - Office of Environmental Management 25
- ORI
  - Outfall Reconnaissance Inventory. ... 8, 11, 21, 25
- SCM
  - Stormwater Control Measure 127
- SIC
  - Standard Industrial Classification 14
- SOP
  - Standard Operating Procedure 5, 20
- SSO
  - Sanitary Sewer Overflow 16
- SSS
  - Separate Sewer System 1
- TMDL
  - Total Maximum Daily Load 4, 21
- UK
  - United Kingdom 12
- USD
  - United States Dollar 2, 19

1 INTRODUCTION

In 2008, the National Academy of Engineering identified the need to restore and improve urban infrastructure as one of the fourteen grand engineering challenges facing the United States during the 21st century (Perry et al., 2008). Urbanization has greatly contributed to the need for improved urban infrastructure to manage overburdened water, power, and transportation systems. The magnitude of the effects of urbanization on local ecosystems has risen exponentially with growth in urban centers, and has caused local governments and managers of infrastructure and environmental systems to fall behind on maintenance and mitigation needs, resulting in an aging infrastructure network which contributes to the continued degradation of already impaired ecosystems (Baird, 2009). Furthermore, urban centers tend to be located near bodies of water, including streams, rivers, bays, and estuaries, resulting in more frequent interactions between infrastructure and the water environment.

The emergence of aging and overburdened water infrastructure in the urban environment, including potable water, wastewater, and stormwater management systems, has resulted in several challenges for ecosystem sustainability. The development of forests and grassland into streets and buildings has increased the overall percentages of total and connected impervious surfaces which generate stormwater runoff, impacting fish, benthic macroinvertebrates, and aquatic habitats through increased peak flows and pollutant levels (Booth and Jackson, 1997; Booth et al., 2002; Wang et al., 2001). Urban growth and expansion has also contributed to a mix of wastewater treatment infrastructure in urban areas. Older urban centers that were established before the 1900s often used combined sewer systems (CSS) to transport wastewater and stormwater for treatment. These systems are susceptible to overloading during storm events, when excess runoff causes a combined sewer overflow (CSO) to discharge untreated sewage to a local waterbody (Lilly et al., 2012). Since the 1900s urban design has shifted to using separate sewer systems (SSS) to transport wastewater to treatment facilities, while municipal separate storm sewer systems (MS4s) convey stormwater runoff to local waterbodies. Additionally, rural fringes of urban centers which were previously served by septic systems are now becoming urbanized, resulting in a patchwork of piped sewer service expansions and septic treatment systems (Spirandelli, 2015).

The health effects attributed to water pollution problems are also widespread. Sanitary and combined sewer overflows are often responsible for increased bacteria levels and contamination of potable water supplies. Stormwater pollution contributes high concentrations of heavy metals to fresh water supplies, and these metals must be removed by potable water treatment processes. These additional contaminants, combined with the presence of pathogens such as cryptosporidium and giardia, and the use of certain disinfection protocols, create a host of challenges in potable water treatment, elevating the risk for gastrointestinal illness and other serious health issues originating from the potable water system. In 2003, the annual cost of waterborne illnesses in the US was estimated to be $2.1-13.8 billion (2002 USD equivalent), a sum comparable to the estimated 20-year cost to repair and improve stormwater infrastructure and management practices (Gaffield et al., 2003).

While stormwater contributes significantly to current water pollution problems, it has been argued that sewage and other inputs to the stormwater network that occur as dry weather flows are a more significant contributor than wet weather stormwater pollution. This theory has been supported by studies which show that even after communities have spent billions of dollars on infrastructure repairs to eliminate discharges from combined and separate sewer systems, bacteria and nutrients from sewage remain a primary source of pollution in local waterways (Center for Watershed Protection and Biohabitats, 2011; Kaushal et al., 2011). To fully address the effects of pollution attributed to urbanization, efforts must be made to reduce and eliminate dry weather flows containing illicit discharges with high concentrations of pollutants.

1.1 BACKGROUND

1.1.1 Illicit Discharges Defined

The legal definition of an illicit discharge is set forth by the National Pollutant Discharge Elimination System (NPDES) legislation as any non-stormwater flow entering a municipal separate storm sewer system (MS4), with some exceptions (USEPA, 2010). Some common examples of illicit discharges include:

- sewage from sanitary sewer cross-connections, overflows, and straight pipes,
- sediment from construction sites or landscaping work,
- fats, oils, and grease from food service establishments, and
- household hazardous wastes such as paint, antifreeze, fertilizers, and pesticides.

The list of illicit discharges provided by the NPDES legislation includes a wide variety of pollutants which can enter the stormwater network or a local stream in multiple ways. A common path is via the connection of pipes draining non-stormwater flows to the stormwater network, which can occur when construction plans are incorrectly interpreted, or when poor construction methods are employed. These discharges are termed illicit connections and represent fixed locations where a chronic or constant discharge occurs. Some examples of common illicit connections include sanitary sewers, shop floor drains, or cooling tower discharges. Another way by which illicit discharges enter the stormwater network is through illegal dumping. Illegal dumping refers to incidents in which a substance is improperly disposed of down a storm drain, and is a common pathway for products such as paint, spent motor vehicle fluids, and cooking oils to enter the stormwater network. Spills are another way for pollutants to enter a stormwater network, often caused by vehicular accidents or poor housekeeping practices in businesses or on construction sites. Illegal dumping and spills are not constrained to a particular geographic location, and are often intermittent or one-time occurrences, although businesses or residents that dump materials may be likely to repeat the offense.

In addition to naming specific illicit discharges, the NPDES legislation provides an explanation of the major exceptions to the illicit discharge rule. The two primary exceptions listed are discharges from facilities operating under a NPDES permit, and discharges from firefighting activities (USEPA, 1999). Also described is a broad category of non-stormwater flows which are not considered illicit by the federal government, but which are left to be regulated at the discretion of local governments. Some examples of these flows include groundwater and spring flow, sump pumps and trench drains, residential car washing, and dechlorinated swimming pool water (USEPA, 1999). A locality must decide if these sources contribute to local water quality problems, and should therefore be considered a significant source of pollutants and regulated as an illicit discharge. This provides localities with some flexibility to identify specific pollutants of concern to their communities and local waterways.

1.1.2 Stormwater and Illicit Discharge Regulations

1.1.2.1 Federal and State Stormwater Regulations

Stormwater regulation begins at the federal level with the Environmental Protection Agency (EPA) and the Clean Water Act of 1972 (CWA), which prohibits the discharge of pollutants from point sources into navigable waters of the United States (USEPA, 2016a). The CWA established the National Pollutant Discharge Elimination System (NPDES) which required discharging facilities, such as industrial sites and municipal treatment works, to obtain a permit for point source discharges, defined as pipes or manmade ditches discharging to navigable waters (USEPA, 2010). In response to slow progress in meeting water quality goals nationwide, the NPDES legislation was expanded in 1990 to include municipal separate storm sewer systems (MS4s) discharging to surface waters. MS4 networks constitute the pipes and channels draining stormwater runoff from urban and developed areas into concentrated point source discharges, or outfalls, to a surface water (USEPA, 1999). The 1990 regulations, known as the NPDES Phase I Program, apply to medium and large MS4s, defined as census designated urbanized areas with populations between 100000 and 250000, and greater than 250000, respectively. In 1999, Phase II of the NPDES regulations permitted small MS4s, defined as urbanized areas with populations less than 100000.

Through the NPDES program, the EPA assigns permits allowing Phase I and Phase II MS4s to discharge stormwater into surface waters. Individual permits are generated for Phase I MS4s, while Phase II MS4s are typically covered by a general permit in each state.

States which can demonstrate the capacity and resources to handle NPDES permitting are given authority over NPDES program administration, with oversight from the EPA. The state of Virginia has been granted this authority, and administers NPDES MS4 permits through the Department of Environmental Quality (DEQ). The DEQ releases a permit to an MS4 after the entity submits a Program Plan describing the MS4's detailed approach to addressing stormwater pollution. Upon approval of the Program Plan and receipt of the NPDES MS4 permit, the MS4 is responsible for administering the plan and providing the DEQ with an annual report on progress towards meeting Program Plan goals.

The federal guidelines for Phase I and Phase II permits provide recommendations for best management practices (BMPs) to minimize pollution from stormwater discharges to the maximum extent practicable (MEP). The MEP standard has been the source of several lawsuits in the past decade (EDC and NRDC versus USEPA, 2014; Fowler et al. v. USEPA, 2010) which have resulted in proposed changes to the Phase II rule (USEPA, 2016b) to clarify and expand upon the definition of MEP and ensure that permittees are implementing adequate measures to reduce stormwater pollution. These changes have not yet been approved, but if accepted the proposal may impact state permitting authorities and MS4 permittees by requiring a more thorough MS4 Program Plan.

1.1.2.2 Federal and State Permit Requirements for Illicit Discharges

In the state of Virginia, Phase I individual permits include a Total Maximum Daily Load (TMDL) Action Plan, which describes the steps taken by the MS4 to comply with pollution diets that have been set for local waterbodies impaired by specific pollutants. The individual permit also includes programs and requirements tailored to the specific MS4 entity. All Phase I MS4s are required to address illicit discharges through a section of the permit called Illicit Discharge and Improper Disposal (IDID), and many Virginia Phase I MS4s have additional programs in their permit relating to illicit discharges. Phase II MS4s receive a General Permit from the state which requires a TMDL Action Plan, as well as the implementation of BMPs within six categories of stormwater interventions, known as Minimum Control Measures (MCMs, Table 1-1). Phase II MS4s are required to address illicit discharges under MCM 3 – Illicit Discharge Detection and Elimination (IDDE).

Table 1-1: Minimum Control Measures (MCMs) required as a part of Phase II MS4 permits.

MCM 1	Public Education and Outreach
MCM 2	Public Involvement and Participation
MCM 3	Illicit Discharge Detection and Elimination
MCM 4	Construction Site Stormwater Runoff Control
MCM 5	Post-Construction Stormwater Management
MCM 6	Pollution Prevention/Good Housekeeping

Virginia MS4 permits are generally more stringent and comprehensive than current federal permit requirements, particularly when it comes to illicit discharge programming. Individual permits for Phase I MS4s are tailored to the specific MS4, and programs implemented voluntarily by the MS4 become part of the permit over time. As a result, Phase I permits often contain additional programming and stricter standards than those mandated by federal legislation, and represent a mix of federal and state mandated programs as well as local initiatives (Virginia DEQ, 2013a, 2015a). A list of some of the components of Virginia Phase I MS4 Program Plans related to illicit discharges is provided below. These components may be found in the IDID section of the permit, or as separate programs with objectives addressing illicit discharges:

- Legislation, ordinance, and codes prohibiting illicit discharges
- Dry weather screening of MS4 outfalls
- Public and staff education, outreach, and training
- Complaint investigation and reporting mechanisms
- Industrial and High Risk Runoff (IHRR) facility inspection programs
- Separate sanitary and stormwater infrastructure maintenance and inspection
- Household Hazardous Waste (HHW) collection programs
- Floatables monitoring and reduction programs
- Spill prevention and response programs
- Good housekeeping Standard Operating Procedures (SOPs)

The Virginia Phase II MS4 General Permit is organized similar to the EPA’s general permit in that permittees must include a TMDL Action Plan and address the six MCMs (Table 1-1). However, the Virginia general permit requires more from MS4s than the federal permit to fulfill the requirements set forth in the MCMs and the TMDL Action Plans. MCM 3, Illicit Discharge Detection and Elimination (IDDE), outlines the state's requirements for illicit discharge programming. The NPDES Phase II rule requires four objectives be met by the MS4 (USEPA, 1999):

1. Produce a map of the storm sewer system showing all outfall locations and receiving water bodies.
2. Develop and implement ordinance and enforcement procedures to prohibit illicit discharges.
3. Draft and carry out a plan to detect and eliminate illicit discharges of all forms.
4. Inform businesses and the public about the hazards stemming from illicit discharges and improper disposal.

The federal rule also addresses discussion from the comments period and provides recommendations on additional methods and programs which may be incorporated in MCM 3; however, the legislation does not legally require any further action on these recommendations.

The state of Virginia expands the depth of responsibility of Phase II MS4s by requiring additional work be undertaken to meet these four objectives, and specifying a timeline for completion of the tasks (DEQ, 2013). In particular, the state requires MS4s to provide additional information about each outfall on the map requested in (1), including contributing drainage area and applicable TMDLs, and stresses continued mapping of new discharge points. Regarding (3), the state explicitly requires MS4 operators to draft SOPs for a dry weather screening plan to evaluate outfalls visually and parametrically, investigate any observed flows for illicit discharges, and track the status of an illicit discharge investigation. The state also requires MS4 operators to develop a prioritized list of field activities based on local factors including prior complaints and land use. Furthermore, the state expands on (4) by requiring MS4 operators to develop and publicize a pathway for citizens to report illicit discharges for investigation and follow-up by MS4 personnel. As with Phase I MS4s, the Phase II general permit requires each MS4 to submit an individual Program Plan and annual reports describing any applicable TMDL Action Plans and progress toward compliance with each of the MCMs. Additional or voluntary programs reported by the MS4 are incorporated into the Program Plan, thus expanding further on the minimum federal requirements.

1.1.3 Dry Weather Pollution

In the context of water quality, dry weather pollution refers to pollutants that enter a waterbody via non-stormwater flows, which occur independent of rain events. During precipitation events, pollutants from dry weather flows are diluted by the excess amount of stormwater runoff generated by impervious surfaces in developed areas. However, during periods of no precipitation, dry weather flows will contribute high concentrations of pollutants directly to surface waters, and many of these flows originate from illicit discharges.

In the Baltimore-Washington, D.C. area, a series of investigations were carried out in several local watersheds to determine the impact of dry weather pollution and illicit discharges on stream water quality. The Sligo Creek watershed, located north of Washington D.C., was one of the study watersheds. Field crews sampled dry weather flow from outfalls and base flow from in-stream locations, and tested samples for parameters including bacteria, ammonia, nitrogen, and phosphorus, among others. In Sligo Creek, a comparison of the total nitrogen and phosphorus loads in-stream with those from outfalls revealed that the loadings were nearly equivalent, and that nearly all of these nutrient loads originated from outfalls which exceeded one or more water quality criteria, and therefore are suspected to be illicit discharges (Lilly et al., 2011). In addition, grab samples taken at outfalls and in-stream locations revealed that outfalls exceeding one or more water quality criteria contained significantly higher average bacteria concentrations than ‘clean’ outfalls and the in-stream samples (Lilly et al., 2011). Another study of Western Run, a subwatershed of Jones Falls north of Baltimore, consisted of sampling for nutrients and estimating the annual loads of nitrogen and phosphorus to the stream from outfalls. The results revealed that while 6% of the annual stream discharge originated from outfalls with potential illicit discharges (an exceedance of any water quality criteria), these outfalls contributed 17% and 58% of the annual in-stream nitrogen and phosphorus loads, respectively (Lilly and Sturm, 2010).

The findings in Baltimore and Washington, D.C. support that dry weather flow has a significant impact on stream water quality, and that a large portion of this impact can be attributed to illicit discharges travelling to waterbodies via the MS4 network. These findings have generated interest across the country in determining the effects of dry weather flow, particularly in watersheds draining to impaired waterbodies like the Chesapeake Bay and Puget Sound. In 2009, a study in the City of Federal Way, WA, evaluated the annual impact of residential car washing and found that, when these small individual events are aggregated city-wide, they contributed significant loadings of pollutants, including petroleum hydrocarbons, nutrients, ammonia, and surfactants, to local waterbodies (Smith and Shilley, 2009). And since these pollutants are entering impaired waters, they are also contributing to pollutant loads regulated under TMDL programs.

1.1.4 TMDL Reduction Potential

Section 303 of the CWA requires states to establish designated uses for waterbodies, and set water quality standards (WQS) which support those uses. In the state of Virginia, all waterbodies are designated for recreational use; the population and growth of a balanced, indigenous population of aquatic life; and the production of edible and marketable natural resources (Virginia DEQ, 2014). The CWA also requires states to evaluate waterbodies for compliance with these designated uses, and intervene when water quality standards are not met. TMDLs are the regulatory documents created by state environmental agencies to (1) describe the quantity of a pollutant that an impaired waterbody can assimilate from point and non-point sources without compromising designated uses, and (2) prescribe strategies for reducing pollution inputs to meet the required pollution diet.

Since illicit discharges can contribute to these pollutant loads, MS4s have attempted to calculate the potential load reduction that could be achieved by eliminating these discharges. A study was conducted in the Baltimore-Washington, D.C. area to estimate the magnitude of illicit discharge loads compared to required TMDL reductions for the Sligo Creek and Western Run watersheds. By conducting Outfall Reconnaissance Inventories (ORIs), researchers investigated, identified, and measured potential illicit discharges in the two watersheds, and estimated the fraction of the required TMDL load reductions that could be attributed to potential illicit discharges. For Western Run, the nitrogen, phosphorus, and bacteria contributions from potential illicit discharges represented 63, 21, and 49 percent, respectively, of the required load reductions (Lilly et al., 2012). The fractions for Sligo Creek were slightly lower at 17, 6, and 21 percent of nitrogen, phosphorus, and bacteria load reductions, respectively (Lilly et al., 2012).

These estimates indicate that significant pollutant load reductions could be achieved simply by eliminating illicit discharges from the MS4 network. However, when TMDLs are formed, illicit discharges are typically not factored into TMDL pollution diet calculations. Therefore these discharges may be slipping by unaccounted for in many TMDL programs. Furthermore, state environmental agencies typically create Implementation Plans for TMDLs by recommending various types and quantities of structural and non-structural best management practices (BMPs) which may be employed to improve water quality and meet the TMDL pollution diet requirement. Structural BMPs in particular are often costly for MS4s to implement, and when compared to an Illicit Discharge Detection and Elimination (IDDE) program, structural measures are much less efficient at removing nutrients and bacteria from dry weather flows. As an example, the potential cost to a local government to treat nitrogen stemming from potential illicit discharges in the Sligo Creek watershed is 18 times more expensive when using an end of pipe measure such as a dry swale over an IDDE program (Lilly et al., 2012). This should provide incentive for local governments to expand their IDDE programs; however, these programs do not currently qualify for pollution reduction credits in some TMDL Implementation Plans, including the Chesapeake Bay TMDL (Lilly et al., 2012). The result is a lack of incentive to pursue IDDE and a shift towards other more costly control measures and accounting processes to eliminate nutrients and bacteria using end of pipe solutions instead of source control.

1.1.5 Summary

Illicit discharge detection and elimination is an important topic in stormwater management, and regardless of the status of incentives towards pollution reduction via IDDE programs, the practice is required in MS4 permits and will continue to be a toolset within an MS4's available best management practices. However, while non-structural programs are relatively more cost-effective than structural BMPs for treating pollution from illicit discharges, these programs still require a significant investment of local resources to administer. It is therefore imperative that IDDE programs remain cost-efficient and financially feasible for all MS4s, and especially for Phase II MS4s like those in Virginia, where the state general permit requires significant additional investment beyond federal regulations, particularly with respect to the prioritization of illicit discharge field methods. This challenge presented to smaller Phase II MS4s beginning or operating an IDDE program with limited available resources and experience is a problem only recently addressed in literature, and provides an opportunity for further research to improve the methods used to administer IDDE programs.

1.2 PROBLEM STATEMENT

Cities and localities which operate MS4s are required by the EPA to improve the quality of stormwater entering waters of the United States (WOTUS). This mandate includes a requirement to carry out field reconnaissance and outfall mapping to identify and prevent illicit discharges in the stormwater network. While illicit discharges are a significant source of TMDL pollutants, and field reconnaissance and outfall mapping operations are effective in detecting these sources for elimination, maintaining an effective field program requires significant staff and financial resources, which can present a burden to smaller Phase II MS4s as compared to Phase I entities. Larger Phase I entities often use desktop assessments and risk analysis to prioritize inspections and increase the cost-efficiency of field reconnaissance. Risk analysis can be integrated with mapping requirements to form a valuable tool set for Phase II MS4s. The goals of this research are to investigate the effect of various land, hydraulic, and hydrologic risk factors on the location of illicit discharges within MS4 jurisdictions, and to determine the uncertainty surrounding the quality and accuracy of these factors. Risk factors created using quality data that effectively capture instances of illicit discharges may be utilized by small MS4s to guide field reconnaissance and provide significant, cost-effective water quality improvement through improved detection and elimination of illicit discharges.

1.3 OBJECTIVES

The main objectives of this research study are to:

1. Determine the IDDE methods currently in use or development through review of current literature and discussions with select Virginia Phase I and Phase II MS4s.
2. Investigate potential spatial relationships and interactions among land use, hydraulic, and hydrologic characteristics and the occurrence of illicit discharges.
3. Evaluate the sensitivity of an illicit discharge risk mapping tool to changes in relative weighting.
4. Describe the uncertainty associated with different land use, hydraulic, and hydrologic risk factors used to calculate illicit discharge potential.
5. Develop recommendations for effective illicit discharge risk mapping for Phase II MS4s to utilize this tool during prioritization of outfall reconnaissance and other illicit discharge operations.

2 LITERATURE REVIEW

The following sections provide a review of the relevant literature surrounding illicit discharges. The current field methodologies used in practice to detect and identify illicit discharges are reviewed, and cost comparisons between methods are made where data is available. Additional program components are discussed in the context of the benefits of a holistic approach to illicit discharge detection and elimination, and to describe the limitations presented by staff and financial resource barriers. The current use of risk analysis and risk mapping is also discussed, including the potential to expand and standardize the process. Finally, three different approaches to defining the analysis unit for a risk map are presented, and the benefits and disadvantages of each are compared.

2.1 FIELD RECONNAISSANCE METHODS, UTILIZATION, AND COSTS

2.1.1 The Outfall Reconnaissance Inventory

One of the most widely used and well-known field methodologies is the Outfall Reconnaissance Inventory (ORI), also known as Dry Weather Screening (DWS) in many MS4 Program Plans. The ORI generally consists of visual observations of outfall structures combined with analysis of dry weather flow (DWF), if present, for water chemistry parameters. The ORI has been used by municipalities for many years as an asset management tool and to assess maintenance needs, and it has also been used to establish links between dry weather flow and illicit discharges in several municipalities (Peterson and Grout, 1992). The ORI program was standardized in 2004 for implementation by regulated Phase II MS4s (Brown et al., 2004) and has been the subject of multiple studies comparing the effectiveness of illicit discharge detection and elimination methods (Christian et al., 2004; Irvine et al., 2011; Johnson and Tuomari, 1998; Pitt and Rittenhouse, 2001). While the ORI can serve the dual purposes of asset management and illicit discharge detection, municipalities often find it costly to implement. Field inspection requires significant staff resources, as well as an equipment budget that can vary greatly depending on the screening goals of the municipality. Catalfio et al. (Catalfio et al., 2007) evaluated the cost of the ORI in the Rouge River Watershed in Michigan and determined that it cost approximately $200 per outfall to perform the ORI, or $7,600 per illicit discharge identified. Similarly, Christian et al. (Christian et al., 2004) evaluated the cost effectiveness of the ORI for use by the Michigan Department of Transportation and found that half of their budget was spent on traffic control and safety measures. These studies indicate that outfall reconnaissance may be more effective for traditional versus non-traditional MS4s, and requires strategic planning to ensure the effective use of budgeted resources.

2.1.2 Water Quality Sampling

Another common field method for detecting illicit discharges is to perform water quality testing at outfalls with dry weather flow. A host of different parameters may be analyzed depending on municipal issues and goals, but some common parameters include ammonia, chloride, pH, potassium, and E. coli. By analyzing the water chemistry, it is possible to determine the likely source contributing dry weather flow to the stormwater network. Efforts have been made to develop water quality standards and decision making tools to guide MS4s in the assessment of dry weather flow conditions (Irvine et al., 2011). In particular, the Center for Watershed Protection (CWP) has published a flow chart which uses a subset of common parameters to aid field crews in determining the most likely source of flow (Brown et al., 2004). Water quality testing is often performed on outfalls first, but is also used to track illicit flows upstream and narrow the search area for a source (Johnson and Tuomari, 1998; Tuomari and Thompson, 2004). Depending on the parameters selected, water quality tests may be carried out in the field using rapid test kits, or through sampling laboratory analysis. The costs of lab and field analyses were compared in a study sponsored by the Western New York Stormwater Coalition (WNYSC); the report found that field testing for E. coli using Coliscan Easy Gel® was more than three times cheaper than competitive laboratory rates, while testing for a suite of parameters including phosphorus, nitrate, potassium, chromium, and copper using field reagent kits was more than five times cheaper than laboratory rates (Irvine et al., 2011). Researchers and field technicians have also developed innovative field techniques to measure pollutants of concern without the need to rely on a laboratory. In the UK, testing devices made from cotton tampons were placed in storm sewers to absorb optical brighteners, compounds often found in soaps and detergents that may originate from gray or black water illicit connections. The study showed promising results and warranted additional investigations to further develop the sampler, which cost twenty pence each, or the equivalent of $0.30 US (Chandler and Lerner, 2015).

2.1.3 Dye Tests and Smoke Tests

Dye and smoke testing are two other commonly used field methods in illicit discharge programs. These methods are most applicable for identifying the source of a potential illicit discharge, and they are often used as the final step to identify a discharge isolated to a small geographic region (Chandler and Lerner, 2015). MS4 operators use smoke and dye to test for cross-connections between sanitary and stormwater pipes, and to identify illicit plumbing connections from properties or facilities to the stormwater network (Tuomari and Thompson, 2004). Dye testing has also been used during facility inspections to evaluate the performance of septic tanks or onsite disposal systems (Johnson and Tuomari, 1998). The cost of dye testing is relatively cheaper than smoke testing due to the necessary equipment, and dye testing is used more often in illicit discharge detection programs due to its simplicity. Michigan MS4 operators have used dye testing extensively in their illicit discharge programs, and have determined that the cost of equipment and labor to identify illicit connections averaged $900 per facility inspected (facilities were defined as a residence, commercial, or industrial site) (Catalfio et al., 2007). While this cost is higher than the unit cost quoted for the Outfall Reconnaissance Inventory ($200 per outfall), the cost per illicit discharge discovered is much lower for dye testing, estimated at $2,000 compared to $7,600 for the ORI (Catalfio et al., 2007).

2.1.4 Other Field Detection Methods

There are many other field methods which are less common, but still integral in many MS4 illicit discharge programs. Several studies and reviews of illicit discharge detection methods discuss the use of televised inspections and aerial or thermal imagery. Inspection of stormwater pipes using closed circuit television (CCTV) and robotic camera units can provide helpful information in the case of discharges which are difficult to trace (Johnson and Tuomari, 1998; Tuomari and Thompson, 2004). Thermal imagery has been used in a handful of MS4s to identify potential septage and sewage contributions to local waterbodies (Johnson and Tuomari, 1998; Pitt and Rittenhouse, 2001). While the technique is still under study, municipalities in Arkansas have used thermal imagery to detect differences in temperature between inflow and receiving surface water bodies, which may allow the viewer to pinpoint the locations of failing septic systems or sewer leaks (Eddy, 2000). In-stream water quality monitoring is also a part of many large MS4 program plans, and while the technology at present may not be able to detect individual illicit discharges, monitoring is still useful for identifying large one-time pollution events and constant discharges, and for tracking conditions over time (Chandler and Lerner, 2015; Zielinski and Brown, 2003).

2.1.5 Combining Field and Programmatic Strategies

Given the list of potential strategies for detecting illicit discharges in the field and their benefits and disadvantages, it is not surprising that many MS4s utilize multiple methods in their illicit discharge program. In combination with the development of the Center for Watershed Protection’s IDDE Guidance Manual for Phase II MS4s, a national survey of MS4s (primarily Phase I MS4s) was conducted to determine the degree of acceptance and use of various IDDE methods, and the problems encountered with each method (Zielinski and Brown, 2003). The survey revealed that the average municipality dedicates two-thirds of annual IDDE staff time to field activities, and a large portion of budgeted costs cover staff salaries and education and training programs. Many MS4s implement other programs besides field detection and facility inspections, including education and training for staff and the public, as well as hotline or online reporting systems through which complaints or reports of illicit discharges can be logged for follow-up action by authorities.

Overall, the results of the survey indicate that several elements are common to almost all IDDE programs, namely pollution reporting hotlines, outfall reconnaissance inventories, and receiving water body monitoring (Zielinski and Brown, 2003). Most jurisdictions found that the ORI was helpful in detecting chronic problems on the first survey, but was less useful afterwards for detecting intermittent or one-time discharges. To supplement the ORI, many jurisdictions cross-trained staff from other departments to recognize and report illicit discharges, and relied on public education and hotline reporting to catch sporadic discharges. ORIs were sometimes augmented with dye or smoke tests, water chemistry tests, and inspections of sanitary and stormwater infrastructure to identify additional discharges after addressing chronic problems.

2.2 ILLICIT DISCHARGE RISK ANALYSIS

2.2.1 Use by Current MS4s

In the previously discussed survey by Zielinski and Brown, an effort was also made to gage the mapping capabilities of municipalities and the use of mapping tools in IDDE programs. Almost all jurisdictions reported using paper or digital maps to track outfalls and discharge data, and many of these communities used data to prioritize outfalls and other sites for inspections or dye-testing (Zielinski and Brown, 2003). This was confirmed by a more recent survey of 30 Phase I and Phase II MS4s, where 80% of respondents reported using data to prioritize IDDE actions, and 40% reported using some form of desktop geographic information system (GIS) assessment (Lilly, 2015). Mapping and risk analysis techniques have been reviewed in several publications in the IDDE community, and are recommended to MS4 operators as a tool to aid in program implementation (Brown et al., 2004; Pitt and Rittenhouse, 2001). The processes of risk analysis and risk mapping provide MS4 operators with the ability to assign priority levels to geographic regions more likely to exhibit illicit discharge problems, which allows for a more cost-efficient approach to spending budgeted resources on IDDE programs.

Mapping tools have become a key component of several Phase I MS4 program plans. In particular, Wayne County, Michigan used commercial and industrial facilities with Standard Industrial Classification Codes (SIC Codes) as input data for a GIS tool which ranked drainage areas in the Rouge River Watershed for facility inspections (Tuomari et al., 1995). A later paper by officials in Wayne County describes the use of a GIS method incorporating information on combined sewer systems, land use, businesses, and water quality issues to prioritize areas for illicit discharge evaluations (Johnson and Tuomari, 1998).

Several Virginia Phase I MS4s have developed risk analysis techniques to guide site selection for dry weather screening and wet weather monitoring programs. Fairfax County, Virginia, generates dry weather outfall screening maps which display land use, development age, infrastructure age, impaired waters, complaints, and past offenses to aid in the selection of at least 100 outfalls for screening each year (Fairfax County Stormwater Planning Division, 2014). In addition, the County has developed a more rigorous risk mapping approach to select sites for wet weather monitoring stations. The County generally monitors two high risk areas on a quarterly basis for one year, and then switches to a new pair of sites. The County uses GIS processing tools and information describing MS4 service areas, land use, utility easements, and facilities with SIC codes to select high risk subwatersheds. Statistics generated from the GIS risk analysis are used to rank each subwatershed in terms of the likelihood of illicit discharge problems (Versar Inc., 2014a).

Arlington County, Virginia, is another Phase I MS4 which has developed risk analysis procedures for dry weather and wet weather screening programs. Due to the County’s small size and historical knowledge of commercial and industrial centers, the dry weather screening program focuses on surveying the outfalls from a discrete number of commercial and industrial hot spots within the County (Versar Inc., 2014b). Arlington County’s wet weather monitoring program uses a desktop GIS analysis similar to Fairfax County to score and rank facilities for annual monitoring programs. In Arlington County’s case the method emphasizes sites with industrial and commercial land uses, high percentage of parking lot area, older buildings, and priority SIC classifications (Versar Inc., 2014c).

2.2.2 Factors Affecting Illicit Discharge Potential

The broad definition of illicit discharges as non-stormwater flows that enter an MS4 (USEPA, 2010) leaves MS4s with a multitude of potential pollutants and sources which may contribute to illicit discharges, and which must be narrowed down based on local concerns. As a result, each jurisdiction that chooses to perform some form of illicit discharge risk analysis has created its own set of criteria and factors which contribute to illicit discharge potential (IDP). This widespread use of engineering judgement and differing levels and types of experience has led to a laundry list of potential factors and statistics which are presumed to be linked to IDP (Table 2-1).

Table 2-1: Risk factors used by MS4s or recommended in literature for the evaluation of IDP.

Dataset	Description	Wayne County, MI[a]	WNYSC, Buffalo, NY[b]	Fairfax County, VA[c]	Arlington County, VA[d]	City of Roanoke, VA[e]	Town of Blacksburg, VA[f]	Zielinski and Brown 2003	Brown et al. 2004	Pitt and Rittenhouse 2001
Stormwater Infrastructure	Location, age, condition of outfalls, pipes, and drainage channels	X	X	X	X	X	X	X	X	X
Sanitary Infrastructure	Location, age, condition, of the sanitary sewer network			X		X			X	X
Combined Sewer Infrastructure	Location, age, condition, of any combined sewer service areas	X							X	X
Septic or Onsite Disposal Systems	Existence and performance of septic or onsite disposal systems					X			X	X
Parcel Records, Development Age	Aging infrastructure, cross-connections, out-of-code construction		X	X	X	X	X		X	
Land Use or Zoning	High-risk activities specific to land use or industry	X	X	X	X	X	X	X	X	X
SSO Reports or I/I Studies	Failing or underperforming sanitary infrastructure								X	X
Receiving Waterbodies	Aquatic ecosystems sensitive to or impaired by pollution		X		X			X	X	X
Watersheds	Drainage area defined by topography and/or subsurface drainage	X	X	X		X	X		X	X
Land Cover or Impervious Surface	Indicator of pollution sources, measure of development intensity				X	X	X			
Business and Special Facilities	Standard Industrial Classifications (SIC) or discharge permits	X		X	X	X	X		X	X
Population Density	Intensity of development and urbanization									X
Historic Complaints and Reports	Locations where discharges have occurred in the past	X	X		X	X			X	
Water Quality Data	Past and current sampling efforts and monitoring programs	X							X	
Historical Knowledge or Experience	Seasoned employees, personnel with professional experience	X			X	X			X	X
Easements and Rights-of-Way	Land designated for streets, utilities, or other linear use rights	X		X		X			X	X
Aerial Imagery	Aerial photographs of the land surface	X							X	X

Note: SSO stands for sanitary sewer overflow, and I/I stands for inflow and infiltration.
a (Johnson and Tuomari, 1998; Tuomari and Thompson, 2004; Tuomari et al., 1995)
b (Irvine et al., 2011)
c (Fairfax County Stormwater Planning Division, 2014; Versar Inc., 2014a)
d (Versar Inc., 2014b, 2014c)
e (Dymond et al., 2015a, 2015b)
f (Howard and Schirmer, 2013)

Phase I MS4s with more resources have historically been the champions of risk analysis for illicit discharges. In the face of new legislation requiring the prioritization of IDDE programs (Virginia DEQ, 2013b), Virginia Phase II MS4s, which typically lack sufficient resources to assess the full array of IDP indicators, are left without a sound path forward. Due to the limitations on Phase II MS4 staff and financial resources, MS4 operators must choose a subset of risk factors to evaluate. Phase II MS4s are also limited by a lack of experienced personnel and existing data, whether in paper or digital format (Zielinski and Brown, 2003). In an effort to provide Phase II MS4s with guidance to begin or expand an IDDE program, IDDE methods including the risk mapping process have been publicized in the form of a guidance manual (Brown et al., 2004) and studies have recommended that constrained localities focus on a limited set of indicators to include the stormwater network (outfalls, pipes, and drainage channels), waters of the US, and land use (Zielinski and Brown, 2003). Despite these efforts to provide Phase II MS4s with a straightforward approach to IDDE, the risk analysis process has not been fully standardized, and research remains to be done to determine whether certain indicators have direct links to illicit discharge occurrence.

2.2.3 Risk Mapping Methods

While illicit discharge risk mapping has been implemented by several MS4s, there is still disagreement in the literature about the proper approach to risk mapping, namely the mapping unit upon which the analysis is performed. Generally there are three approaches which have been identified in the literature and practice. The first method uses watersheds as the mapping unit. This approach is detailed in the CWP’s IDDE Guidance Manual (Brown et al., 2004) and specifies that individual watersheds should receive a score for each indicator, allowing for the development of a watershed’s overall score and rank for IDP. This method allows localities to break up their jurisdiction into subunits based on hydrology to prioritize field investigations. The benefit to this method is that watershed boundaries are publicly available data and would not require significant resources to obtain; however, publicly available watershed data is coarse in scale and does not generally incorporate subsurface drainage infrastructure, which may lead to an inaccurate description of the true hydrology of the region.

The second method, proposed by Pitt and Rittenhouse (Pitt and Rittenhouse, 2001) and implemented by several Virginia Phase I MS4s (Virginia DEQ, 2013a, 2015a) uses the MS4 Service Area as the mapping unit for analysis. The MS4 Service Area is delineated by surveying the stormwater network and identifying the drainage areas corresponding to each outfall from the MS4 to a state waterbody. These drainage units are typically smaller than a full watershed, and when combined may not include all land area within a jurisdiction, particularly if that land is not served by the MS4 network. This approach leads to a ‘Swiss cheese’ model where risk is evaluated using only those illicit discharges which occur within the MS4 Service Area, and excluding any potential dumping incidents, spills, or other pollution sources entering a water of the US from outside of the service area but within MS4’s political jurisdiction. In addition, this method requires extensive survey work and resources on the part of the MS4. While knowledge of the full stormwater network is essential to comprehensively address pollution problems and is gradually being required by federal and state permits, as in Virginia (Virginia DEQ, 2013b), this method is still outside of the reach of many Phase II MS4s at present.

The third method, implemented by several Virginia Phase II MS4s to-date (Dymond et al., 2015a; Howard and Schirmer, 2013), utilizes political or jurisdictional boundaries, whether for subsets such as neighborhoods or the full jurisdiction of the MS4. This allows for analysis of all potential pollution sources within the locality’s jurisdiction, and the identification of hot spots which may not be isolated to particular watersheds or services areas. In addition, jurisdictional boundary data is readily available to localities and requires no additional resources to obtain. Political boundaries also do not rely upon accurate hydrologic boundaries. In the end, the decision of the scale used for risk analysis must be made by MS4 operators, as it will depend on local problems and concerns, the type and quality of available data, and the resources at hand to collect new data.

2.3 SUMMARY

Illicit discharge programs have been required by NPDES law for over 25 years in the case of Phase I MS4s (16 years for Phase II MS4s), and implementation over this timespan has enabled regulatory and regulated entities, as well as research and academic institutions, to formulate and evaluate the capabilities of a wide variety of field methods to detect illicit discharges. It is difficult to estimate or predict the costs of these programs due to variations in MS4 size and the case-specific political and pollutant challenges that must be addressed. However, it is clear that MS4s are required to carry out resource intensive programs to eliminate illicit discharges from the stormwater network and to prevent them from occurring in the future. Risk analysis and risk mapping procedures present a significant opportunity for MS4s to prioritize their IDDE efforts, target programs to specific geographic areas, and improve the cost-effectiveness of MS4 operations, providing benefits to the community through reduced pollution and program savings. As with most other IDDE programs, there are distinct staff and financial resource barriers to performing a risk assessment. While guidance documentation has been created for smaller MS4s, more can be done to standardize the risk mapping process, establish a set of indicators which are related to the occurrence of illicit discharges, and alert MS4s to the uncertainty inherent in the data used to perform such analyses.

3 DEVELOPMENT OF EFFECTIVE PROCEDURES FOR ILLICIT DISCHARGE RISK MAPPING

3.1 INTRODUCTION

In 2008, the National Academy of Engineering identified the need to restore and improve urban infrastructure as one of the fourteen grand engineering challenges facing the United States during the 21st century (Perry et al., 2008). Urbanization has greatly contributed to the need for infrastructure to manage overburdened water, power, and transportation systems. The emergence of aging and overburdened water infrastructure in the urban environment has resulted in several challenges for ecosystem sustainability. Aquatic habitats and organisms are impacted by higher peak flows, lower base flow, and higher pollution levels due to increased runoff from new development and impervious surfaces (Booth and Jackson, 1997; Wang et al., 2001). Urban sprawl has contributed to the development of a patchwork of aging wastewater infrastructure, complicating the protection of local waterways (Spirandelli, 2015). Furthermore, the interactions between aging wastewater and stormwater infrastructure systems and the natural water environment have contributed to increased public health hazards including bacteria, pathogens, and heavy metals in drinking water supplies and recreational waters (Gaffield et al., 2003; Swann, 2001). In 2003, the annual cost of waterborne illnesses in the US was estimated to be $2.1-13.8 billion (2002 USD equivalent), a sum comparable to the estimated 20-year cost to repair and improve stormwater infrastructure and management practices on a national scale (Gaffield et al., 2003). While stormwater pollution is a major contributor to current water quality problems, studies have shown that dry weather pollution also has a significant impact on water quality. Even after communities across the nation have collectively spent billions of dollars to repair wastewater infrastructure and to reduce overflows from combined and separate sewer systems during storm events, bacteria and nutrients from dry weather discharges remain a primary source of pollution in local waterways (Center for Watershed Protection and Biohabitats, 2011; Kaushal et al., 2011).

Dry weather discharges comprise a broad array of non-stormwater flows which enter stormwater conveyance systems through natural or anthropogenic pathways. Many dry weather discharges originate from illicit connections, illegal dumping, or spills, which are collectively known as illicit discharges. An illicit discharge is defined in the National Pollutant Discharge Elimination System (NPDES) legislation as any non-stormwater flow entering an MS4, with some exceptions, including discharges from NPDES permitted facilities, firefighting, and those deemed an insignificant source of pollution to the local waterway (USEPA, 1999, 2010). Illicit connections are pipes which are improperly connected to the stormwater network and represent fixed locations where sewage, wash water, industrial wastes, and other pollutants may enter the MS4. Illegal dumping occurs when a substance is improperly disposed of in a storm drain, and is a common pathway for paint, spent motor vehicle fluids, and similar wastes. Spills occur when a discharge reaches the MS4 unintentionally, often during vehicular accidents, or because of poor pollution management at a business or construction site. Unlike illicit connections, illegal dumping and spills are often intermittent or one-time occurrences and are not constrained to a geographic location, although entities that dump materials are more likely to repeat the offense.

Illicit discharges are regulated by stormwater legislation and NPDES MS4 permits under the Clean Water Act (CWA) (USEPA, 2016a). At the federal level, the Environmental Protection Agency (EPA) regulates urban areas through Phase I and Phase II NPDES permits. Phase I permits regulate medium and large urban areas with populations greater than 100,000 and 250,000, respectively, while Phase II permits regulate small urban areas with populations less than 100,000. In the Commonwealth of Virginia, individual Phase I permits and the Phase II General Permit are administered by the state Department of Environmental Quality (DEQ), which reviews all Program Plans submitted by MS4 operators detailing the particular ordinances and programs implemented to manage stormwater quality and quantity.

While federal permit requirements include measures to prohibit, prevent, identify, and eliminate illicit discharges, Virginia permits expand beyond the federal mandates. Phase I MS4s undergo a permit review process where voluntary programs initiated by each MS4 are integrated into the MS4’s individual permit. The resulting MS4 Program Plan is a mix of mandated programs and local initiatives, many of which are associated with the Phase I permit requirement known as Illicit Discharge and Improper Disposal (IDID) (Virginia DEQ, 2013a, 2015a). For Phase II MS4s, the Phase II General Permit regulates illicit discharges through one of six Minimum Control Measures (MCMs). This particular MCM is known as Illicit Discharge Detection and Elimination (IDDE), and the permit explicitly requires additional effort from MS4s towards the fulfillment of IDDE requirements beyond the federal mandates. The expanded responsibilities include (1) more detailed mapping requirements, (2) Standard Operating Procedures (SOPs) and prioritized lists of field activities, and (3) the development and broadcasting of citizen reporting mechanisms (DEQ, 2013).

Most illicit discharges occur independent of precipitation events, therefore illicit discharges are often found by inspecting stormwater infrastructure for dry weather flows. Dry weather flow investigations performed by localities in the Baltimore-Washington, D.C. area show that for Sligo Creek, in-stream loads of nitrogen and phosphorus are nearly equivalent to the combined load from dry weather flowing outfalls that exceed one or more water quality criteria, suggesting that potential illicit discharges contribute substantially to nutrient levels in streams (Lilly et al., 2011). A similar study in the Western Run watershed in Baltimore revealed that while only 6% of the annual stream discharge originated from outfalls with potential illicit discharges, these outfalls accounted for 17% and 58% of the annual in-stream nitrogen and phosphorus loads, respectively (Lilly and Sturm, 2010). Furthermore, a study in the City of Federal Way, WA, revealed that while residential car washing activities individually generate minimal amounts of pollution, the aggregate of these events across an entire city creates a substantial pollution burden on downstream receiving waters (Smith and Shilley, 2009).

The resulting effect of potential illicit discharges is the degradation of water quality in downstream waterways, which limits the designated uses of those waterways. In waters with Total Maximum Daily Load (TMDL) requirements, pollution loading from illicit discharges may account for a significant proportion of the required TMDL reduction. In Sligo Creek, the nitrogen, phosphorus, and bacteria contributions from potential illicit discharges were estimated to be equal to 17, 6, and 21 percent, respectively, of the required TMDL reductions (Lilly et al., 2012). It was also estimated for Western Run that 63, 21, and 49 percent of nitrogen, phosphorus, and bacteria load reductions, respectively, may be achieved if potential illicit discharges were eliminated from the MS4 (Lilly et al., 2012). These pollution loads are not accounted for during the TMDL development process, and a lack of TMDL credits for illicit discharge programs often forces localities to fulfill regulatory requirements by using more costly end-of-pipe measures to treat illicit discharges instead of controlling or eliminating pollution at the source.

As part of the illicit discharge program, MS4 entities utilize field reconnaissance methods to detect and eliminate discharges and make significant improvements to water quality. The Outfall Reconnaissance Inventory (ORI), also known as dry weather screening, is one of the most common field methods for illicit discharge detection. The ORI generally consists of visual observations at outfalls and an analysis of the water chemistry of any dry weather flow discovered. Water quality sampling programs are used to test stormwater for a wide variety of parameters, which helps to identify pollutants and their potential source(s). Dye- and smoke-testing are field methods used to identify the source of a potential illicit discharge, and they are often used as the final step to identify a discharge isolated to a small geographic region (Chandler and Lerner, 2015). Other field methods, such as closed-circuit television (CCTV) inspections, thermal imagery analysis, and in-stream water quality monitoring are also used by many MS4s. While these field methods each provide distinct advantages for illicit discharge detection, they are similar in that they require significant financial and staff resources. The most effective illicit discharge programs often utilize multiple field methods combined with programmatic strategies, such as reporting hotlines or training and education programs, to increase the efficiency of field work (Zielinski and Brown, 2003).

Many MS4s have turned to risk analysis as a tool to direct illicit discharge program implementation. Surveys have shown that MS4s use a variety of data to prioritize outfall or facility inspections (Zielinski and Brown, 2003), and MS4s have often turned to desktop geographic information system (GIS) assessments to handle this task (Lilly, 2015). While recommendations have been published to provide guidelines to MS4s interested in risk analysis (Brown et al., 2004; Pitt and Rittenhouse, 2001; Zielinski and Brown, 2003), a review of current risk analysis techniques indicates that a core set of variables has not been adopted (Table 3-1). In addition, MS4s use different analysis methodologies, and evaluate risk using differing geographic units, such as watersheds, political jurisdictions, or MS4 service areas.

Table 3-1: Datasets for illicit discharge risk analysis used by MS4s or recommended in literature.

Dataset	Description	Wayne County, MI^a	WNYSC, Buffalo, NY^b	Fairfax County, VA^c	Arlington County, VA^d	City of Roanoke, VA^e	Town of Blacksburg, VA^f	Zielinski and Brown (2003)	Brown et al. (2004)	Pitt and Rittenhouse (2001)
Stormwater Infrastructure	Location, age, condition of outfalls, pipes, and drainage channels	X	X	X	X	X	X	X	X	X
Sanitary Infrastructure	Location, age, condition, of the sanitary sewer network			X		X			X	X
Combined Sewer Infrastructure	Location, age, condition, of any combined sewer service areas	X							X	X
Septic or Onsite Disposal Systems	Existence and performance of septic or onsite disposal systems					X			X	X
Parcel Records, Development Age	Aging infrastructure, cross-connections, out-of-code construction		X	X	X	X	X		X	
Land Use or Zoning	High-risk activities specific to land use or industry	X	X	X	X	X	X	X	X	X
SSO Reports or I/I Studies	Failing or underperforming sanitary infrastructure								X	X
Receiving Waterbodies	Aquatic ecosystems sensitive to or impaired by pollution		X		X			X	X	X
Watersheds	Drainage area defined by topography and/or subsurface drainage	X	X	X		X	X		X	X
Land Cover or Impervious Surface	Indicator of pollution sources, measure of development intensity				X	X	X			
Business and Special Facilities	Standard Industrial Classifications (SIC) or discharge permits	X		X	X	X	X		X	X
Population Density	Intensity of development and urbanization									X
Historic Complaints and Reports	Locations where discharges have occurred in the past	X	X		X	X			X	
Water Quality Data	Past and current sampling efforts and monitoring programs	X							X	
Historical Knowledge or Experience	Seasoned employees, personnel with professional experience	X			X	X			X	X
Easements and Rights-of-Way	Land designated for streets, utilities, or other linear use rights	X		X		X			X	X
Aerial Imagery	Aerial photographs of the land surface	X							X	X

Note: SSO stands for sanitary sewer overflow, and I/I stands for inflow and infiltration.

a (Johnson and Tuomari, 1998; Tuomari and Thompson, 2004; Tuomari et al., 1995)
b (Irvine et al., 2011)
c (Fairfax County Stormwater Planning Division, 2014; Versar Inc., 2014a)
d (Versar Inc., 2014b, 2014c)
e (Dymond et al., 2015a, 2015b)
f (Howard and Schirmer, 2013)

The current MS4 regulatory environment requires that field reconnaissance and other non-structural programs be included as part of a comprehensive solution to address illicit discharges in spite of the high resource demand and lack of TMDL incentives. Virginia further requires that Phase II MS4 entities provide prioritized lists of field activities and programs to demonstrate permit compliance. Risk analysis can achieve both of these objectives by increasing the cost-effectiveness of field operations and prioritizing field implementation. The challenge remaining for many smaller Phase II MS4s is to implement risk analysis while operating with limited available resources and experience. The goals of this study are to investigate the effects of various land use, hydraulic, and hydrologic risk factors on the location of illicit discharges, and to determine the uncertainty surrounding the quality and accuracy of these factors. Data from two case study municipalities are used to investigate potential spatial relationships among risk factors and illicit discharges. Composite risk maps are generated and evaluated for sensitivity to risk factor performance, risk factor interactions, and uncertainty. Finally, recommendations are compiled to assist MS4 implementation of risk analysis to prioritize field reconnaissance and improve the cost efficiency of the illicit discharge program.

3.1.1 Case Study Location 1: County of Fairfax, VA

The County of Fairfax, VA, covers 1,039 square kilometers (406 square miles) south and west of Washington, D.C., has a population of 1,081,726 (US Census Bureau, 2015), and has been permitted as a Phase I MS4 since 1997 (Virginia DEQ, 2015b). The County is located within the Chesapeake Bay watershed, which is the largest watershed in the US for which a TMDL study has been undertaken by the EPA. The Chesapeake Bay TMDL requires localities to reduce nitrogen, phosphorus, and sediment pollution into local waterways which drain to the Bay. Several Fairfax County watersheds are under additional TMDLs for other pollutants, increasing the regulatory burden of water quality improvement carried by the County government (Virginia DEQ, 2015a).

The County implemented a real estate tax in July of 2009 covering most of the County, with the exception of Fort Belvoir, to provide a more stable source of income for stormwater programming and capital improvement projects (Fairfax County, 2010). The County is the largest jurisdiction by population in the Baltimore-Washington, D.C. region and in the Commonwealth of Virginia, and is the second wealthiest local jurisdiction in the United States, as measured by median household income, superseded only by neighboring Loudoun County (US Census Bureau, 2015). As such, the real estate tax, which has gradually been increased over time, has provided significant financial resources for the County's stormwater fund.

The socioeconomic position of the County of Fairfax has enabled County officials to develop a robust MS4 Program Plan, including multiple programs dedicated to IDID permit requirements (Virginia DEQ, 2015a). In particular, the County operates dry weather screening and wet weather monitoring programs which utilize risk analysis to identify geographic areas at high risk for illicit discharge. These sites are assigned a rank based on the risk analysis and then prioritized for screening and monitoring.

3.1.2 Case Study Location 2: City of Roanoke, VA

The City of Roanoke covers 110 square kilometers (43 square miles) in southwest Virginia along the Roanoke River, has a population of 97,032 (US Census Bureau, 2015), and has been permitted as a Phase II MS4 since 2003. The City of Roanoke is located within the Roanoke River watershed, which is the second largest TMDL project in the Commonwealth of Virginia behind the Chesapeake Bay TMDL. The Roanoke River is listed as "impaired" under the Clean Water Act Section 303(d) for sediment, bacteria, and polychlorinated biphenyl (PCB) loads in excess of water quality standards, and several City watersheds have additional impairments for bacteria nested under the Roanoke River TMDL. In 2015, a TMDL implementation plan for sediment and bacteria in the Roanoke River watershed was finalized by the DEQ (The Louis Berger Group Inc. and Virginia Department of Environmental Quality, 2015), prompting the City of Roanoke and other local MS4s to draft TMDL Action Plans in partial fulfillment of their MS4 permit obligations.

In anticipation of increasing permit responsibilities and a growing financial burden to improve water quality, the City of Roanoke implemented a stormwater utility fee effective July 2014 and began restructuring the public works department to create a new stormwater division dedicated to addressing water quantity and quality issues (City of Roanoke, 2015). The population of Roanoke is much smaller than Fairfax County; in addition, the median household income in Roanoke is about $30,700 and 16% of the population lives below the poverty line (US Census Bureau, 2015). As a result, the utility generates a smaller pool of resources for stormwater programming and capital improvement projects, when compared to Fairfax County.

While Fairfax County, as a Phase I community, has been permitted longer than the City of Roanoke, Roanoke has many comparable programs and initiatives in place (Virginia DEQ, 2013c). In particular, Roanoke conducts an ORI through the Office of Environmental Management (OEM), which is responsible for performing dry weather screening at a minimum of 50 outfalls annually to detect and eliminate illicit discharges (Virginia DEQ, 2013c). With funding now available from a stormwater utility, the City is interested in expanding its MS4 programming, including the implementation of risk analysis to aid in the prioritization of field activities required in the IDDE section of its MS4 permit.

3.2 METHODS

The methodology for this study was divided into four major parts. First, interviews were held with MS4 operators and illicit discharge personnel to gage the current use of field methods and risk analysis, and to identify and obtain available datasets for risk factor production. Risk factors were then developed in a GIS environment using available data from MS4s and public sources to describe illicit discharge potential (IDP), or the risk of an illicit discharge due to some condition described by the individual risk factor. Risk factor performance was evaluated individually, and then risk factors were compared using statistical methods to develop an overall performance ranking. Finally, a composite risk map combining all risk factors was produced using four sets of coefficients to test sensitivity to risk factor performance, risk factor interactions, and uncertainty.

3.2.1 Data Collection

Interviews were conducted with multiple Phase I and Phase II MS4s in the Commonwealth of Virginia to determine the field methods currently used to detect, identify, and eliminate illicit discharges. Conversations by phone combined with field visits provided insight into the field operations and office-based risk assessment and planning that MS4 operators use in their IDID and IDDE programs, and the data that operators found particularly useful for determining IDP.

The County of Fairfax maintains a large public database of GIS data used to manage County programs and services. Public data such as watershed delineations, parcel, building, and land records, street polylines, and land use and zoning polygons were collected from the Fairfax County GIS and Mapping website (Fairfax County, 2016). Spatial data were downloaded as shapefiles and organized in a geodatabase, while non-spatial data were downloaded as excel spreadsheets and stored outside of the geodatabase. Coordination with County agencies and GIS technicians was required to obtain additional datasets which were maintained offline for privacy and security reasons, including: MS4 and sanitary infrastructure; septic field locations; dog licenses; and detailed records of illicit discharges.

The City of Roanoke also maintains a small public database, from which several datasets were obtained describing streets, streams, and zoning. Coordination with City officials enabled further access to the City's operational geodatabase and information on MS4 infrastructure, dog licenses, business addresses, and building records. Coordination with the Western Virginia Water Authority (WVWA) was required to obtain information on sanitary sewer infrastructure and service records for Roanoke.

Data for both case study jurisdictions was also collected from several publicly accessible databases, including the US Department of Agriculture (USDA) Geospatial Data Gateway (USDA - NRCS, 2016), the EPA Enforcement and Compliance History Online (ECHO) database (EPA, 2016), the DEQ (DEQ, 2016), and the US Census Bureau (US Census Bureau, 2016). Information was collected from these sources regarding land cover, streams, business locations, construction sites, and total population. As it was hypothesized that inconsistent data quality and format would introduce uncertainty in the risk mapping results, metadata was collected for each dataset used in the analysis. Metadata was reviewed to determine the data collection methods, as well as to qualitatively describe the age, coverage, completeness, and accuracy of the dataset.

3.2.2 Risk Factor Production

ArcGIS for Desktop (Environmental Systems Research Institute, 2015) was the primary tool used to transform spatial and non-spatial input data into spatial layers describing the IDP within each jurisdiction. Each spatial layer described the risk of occurrence of an illicit discharge attributed to one of thirteen risk factors chosen for this study (Table 3-2). The thirteen risk factors were grouped into three categories based on the risk factor data type and the unit of analysis in the GIS environment. The fundamental steps of the risk mapping process are summarized here, with detailed procedures described in Appendix A – Risk Mapping Procedures.

Table 3-2: Definitions of risk factors analyzed as indicators of IDP. Risk factors are organized by data type and unit of analysis.

Risk Factor Type	Risk Factor Name	Definition
Categorical	1. Land Cover	Existing land cover according to the National Land Cover Dataset, seven categories
	2. Land Use	Existing land use on a parcel, six categories
Continuous Over Subunit	3. Population Density	Number of residents per square kilometer, measured over census block groups
	4. Percent Impervious	Proportion of watershed area covered by impervious surfaces
	5. Development Age	Median number of years since a major structure was built on land parcels in a watershed
	6. Outfall Density	Average number of outfalls per kilometer of stream in a watershed
	7. Aging Sanitary Infrastructure	Risk of pipe failure in a watershed based on age and material and weighted by pipe length
	8. Drainage Density	Kilometers of pipes and channels per square kilometer of watershed area
Continuous Over Jurisdictional Area	9. Generating Site Density	Number of potential pollution generating facilities per square kilometer
	10. Infrastructure Access Density	Number of access points to MS4 infrastructure per square kilometer
	11. Construction Site Density	Number of construction projects with active VSMP permits per square kilometer
	12. Aging Septic System Density	Number of septic systems older than 30 years per square kilometer
	13. Dog License Density	Number of dogs per square kilometer

Note: VSMP stands for Virginia Stormwater Management Program, which is the regulatory program governing stormwater management for land-disturbing activities in the Commonwealth of Virginia.

Two of the risk factors analyzed, land cover and land use, were categorical in nature. The input data for the land cover and land use layers were reclassified into seven and six predefined types, respectively, to generate the risk factors used for the spatial analysis. In the case of land cover, the reclassified raster data were temporarily converted to vector polygons to carry out the spatial overlay process described in the next section.

Five of the risk factors analyzed were continuous in nature, and described point densities which varied continuously over the study areas. These factors included: infrastructure access density, aging septic system density, generating site density, construction site density, and dog license density. Since each of these risk factors could be modeled as a continuous surface over the study areas, a kernel density function was applied to the point datasets in ArcGIS to generate raster layers with 30m x 30m cells (Figure 3-1). This raster cell size was used for all risk factors as it is the same resolution as the National Land Cover Dataset (NLCD) (Jin et al., 2013), which was the coarsest resolution dataset used in this analysis. Continuous raster layers were reclassified to produce raster layers that qualitatively describe three IDP risk levels (Low, Moderate, and High). Reclassification of IDP scores into risk levels relied upon recommendations from prior studies and manuals (Brown et al., 2004; Swann, 2001); however, in many cases no reference publications were available. In these cases engineering judgement was used to determine the appropriate cutoff for each risk level. The raster layers were then temporarily converted to vector polygons to facilitate the spatial overlay process.

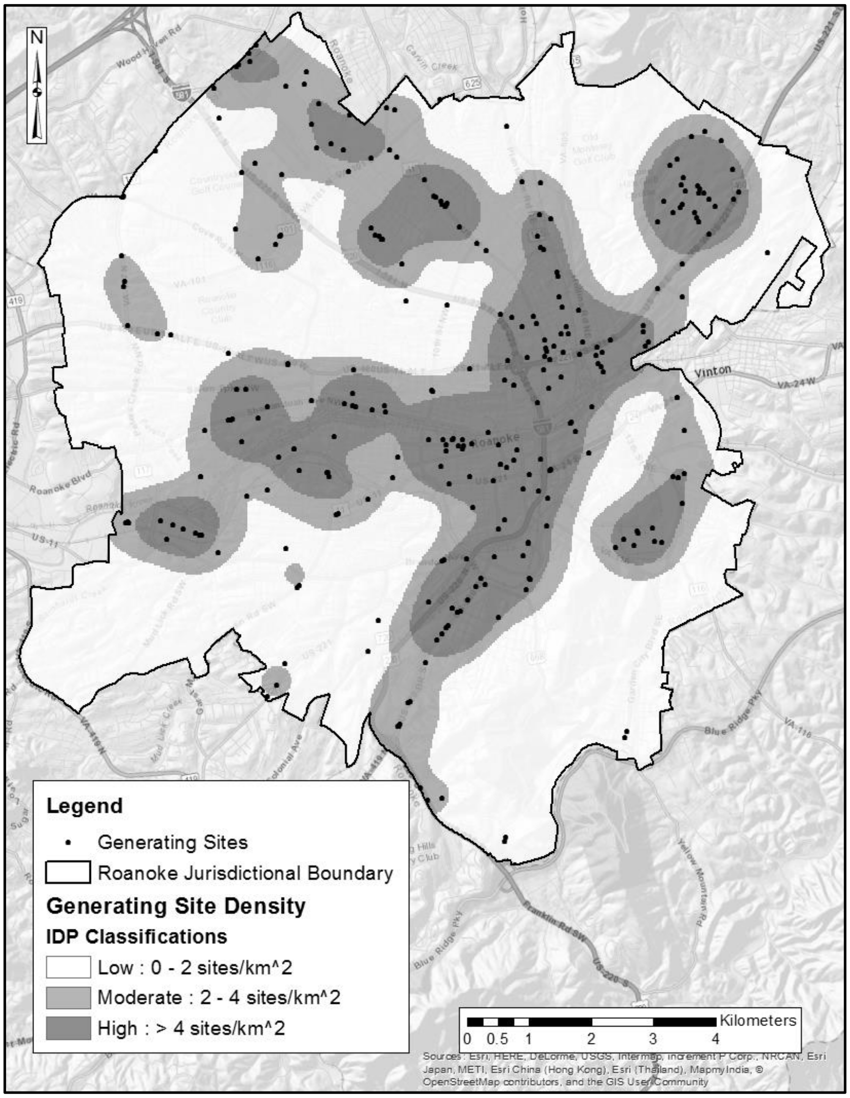


Figure 3-1: IDP risk as described by the generating site density risk factor. Data obtained from the EPA's ECHO database (EPA, 2016).

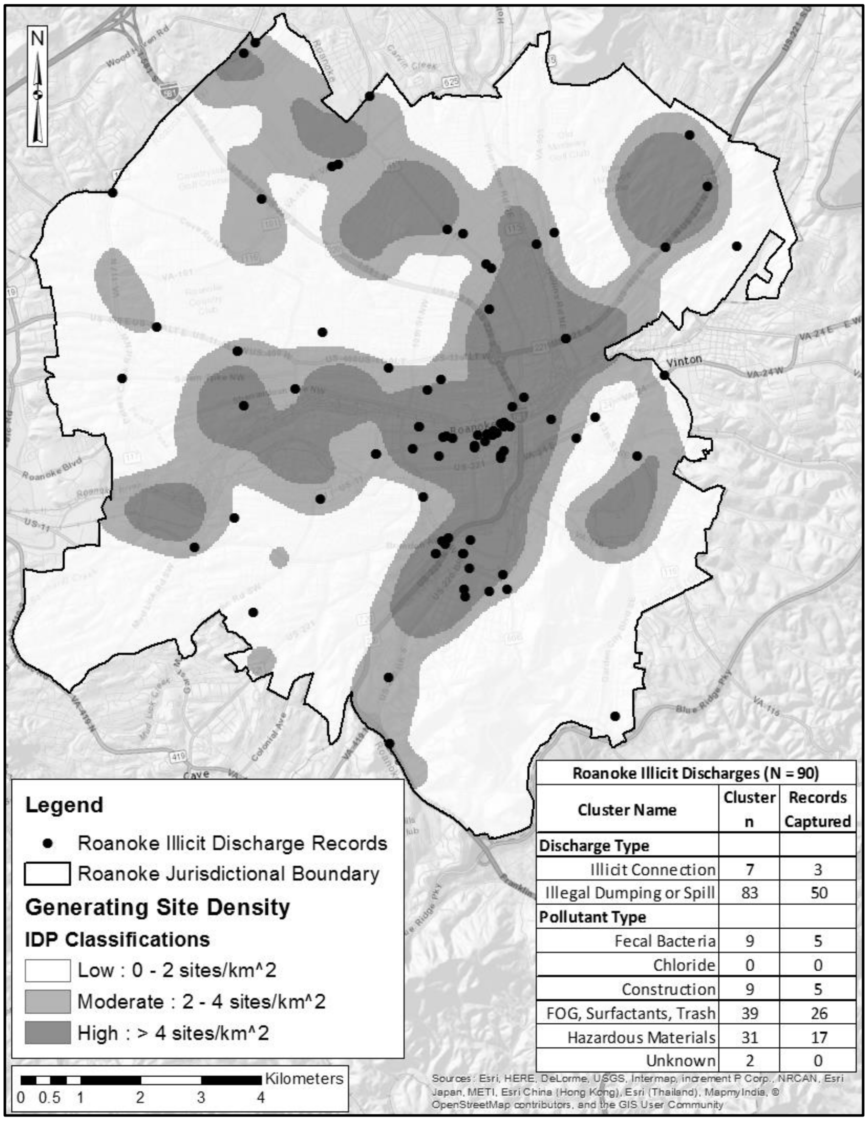
Six of the risk factors analyzed were a mix of densities and other continuous variables which were characteristically linked to subunits within each jurisdiction, and were evaluated within these subunits before creating the risk layer for the full jurisdictional area. These factors included: population density, outfall density, percent impervious surface, development age, drainage density, and aging sanitary sewer infrastructure. Population density was evaluated over census block groups within each jurisdiction, and the rest of these risk factors were evaluated over watersheds within each jurisdictional boundary. In all cases, calculations were made for each subunit to obtain the value of the variable in question, such that risk layers for each jurisdiction consisted of subunits described by discrete numbers from a continuous domain of possible values. These layers were then reclassified to produce vector layers qualitatively describing the IDP risk level (Low, Moderate, and High). Reclassification of IDP scores into risk levels relied upon recommendations from prior studies and manuals (Brown et al., 2004; Swann, 2001); however, in many cases no reference publications were available. In these cases engineering judgement was used to determine the appropriate cutoff for each risk level.

3.2.3 Spatial Analysis

Before the spatial analysis was performed, records of illicit discharges were grouped using two different discharge classification schemes. First, as it was hypothesized that some risk factors may be associated with discharges whose locations are fixed, illicit discharge records were split into clusters based on illicit discharge type - illicit connections (fixed locations where pipes cross-connect) and illegal dumping and spills (discharges assumed to occur independent of location). Second, discharges were classified based on pollutant type to evaluate whether risk factors are associated with specific pollutants. Discharge records were reviewed for similarities in physical and chemical characteristics and discharge sources, and resulted in the following six pollutant clusters:

1. Fecal Bacteria – sanitary sewer overflows, cross-connections, and animal waste dumping
2. Chloride – salt spills, salt storage runoff, and chlorinated swimming pool discharges
3. Construction – sediment, concrete washes, and plant material dumping from utility and roadway work, site development/redevelopment, and landscaping projects
4. Fats/Oils/Grease (FOG), Surfactants, and Trash – similar discharges from restaurants, groceries, food trucks, car washes, fairs, dry cleaners, residences, and institutions
5. Hazardous Materials – industrial, commercial, and residential discharges of hazardous compounds including paints, pesticides, herbicides, process wastes, hydrocarbons, coolant fluids, and metals
6. Unknown – discharges lacking the descriptive evidence or testing to be classified

After each discharge record was assigned to a discharge and a pollutant cluster, discharge records were spatially overlaid on top of individual risk factor layers (Figure 3-2). An illicit discharge was considered to be “captured” by a risk factor if the discharge fell within the high-risk overlay for that factor. A comma separated values (CSV) file was created to store the data for each jurisdiction such that each risk factor occupied a row, and for each discharge or pollutant cluster a column was created and populated with the number of discharge records captured by the corresponding risk factor.



Roanoke Illicit Discharges (N = 90)

Cluster Name	Cluster n	Records Captured
Discharge Type		
Illicit Connection	7	3
Illegal Dumping or Spill	83	50
Pollutant Type		
Fecal Bacteria	9	5
Chloride	0	0
Construction	9	5
FOG, Surfactants, Trash	39	26
Hazardous Materials	31	17
Unknown	2	0

Sources: Esri, HERE, DeLorme, USGS, Intermap, increment P Corp., NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), MapmyIndia, © OpenStreetMap contributors, and the GIS User Community

Figure 3-2: Example of the spatial overlay process performed for the generating site density risk factor.

Statistical methods were employed to compare risk factor performance and determine a relative ranking that describes the ability of individual risk factors to predict overall IDP. All statistical analyses were performed in the R statistical software environment (R Core Development Team, 2015) where the CSV files were used to construct contingency tables pairing all risk factors against each other in each jurisdiction. Fisher's exact test was used to evaluate the contingency tables as it is capable of testing 2x2 tables and is robust against low frequencies in observed data (greater than or equal to 5 data points). The test was originally proposed by Fisher (Fisher, 1970) as a method to determine statistical significance between the proportions of two groups. The method was further developed by Routledge (Routledge,

2005) and has been implemented in software by Mehta and Patel (Mehta and Patel, 1986). A two-tailed Fisher test with α = 0.05 was used to determine if there was a significant difference in performance between any two risk factors, as a two-tailed test is more conservative in estimating significance than a one-tailed test. Performance was measured by the proportion of captured records in a discharge or pollutant cluster. One-tailed Fisher tests with α = 0.05 were then performed to determine the direction of significance. The Fisher analyses and capture proportions were then used to rank risk factors in each jurisdiction, according to the factors' ability to predict the locations of discharges in each discharge and pollutant cluster.

It was postulated that several root factors (percent impervious, land use, development age, and population density) can induce a high-risk overlay in other risk factors if the root factor exhibits high risk. The effect of this interaction would be an exaggeration of the true IDP in high-risk areas shared by multiple interacting risk factors. This hypothesis was tested by comparing the high-risk overlays of root factors to other risk factors and identifying high-risk areas that overlapped. After analyzing high-risk overlays for potential interactions among risk factors, the jurisdictional rankings were further consolidated to provide a shortlist of risk factors which can reasonably estimate IDP for all discharge and pollutant clusters.

3.2.4 Sensitivity Analysis

After the spatial analysis was completed, composite risk maps were produced. First, all vector and raster layers from the spatial analysis were converted to a decimal raster format. For risk factors which had been converted from raster to vector polygons for the spatial analysis, the initial raster file was used for this operation to reduce uncertainty from added processing. Decimal raster layers for each of the thirteen risk factors were then combined in the GIS environment using map algebra to form a single, composite risk map for each jurisdiction describing the overall IDP. The composite IDP scores in Fairfax and Roanoke were then aggregated using watershed boundaries to produce composite maps describing the mean IDP on a watershed basis. All composite maps were reclassified into three qualitative levels of risk (Low, Moderate, and High).

The initial risk maps were created by treating each risk factor as equally important. However, jurisdictions often have knowledge about critical issues which should be included in the risk mapping process. In addition, risk factor performance, interactions, and uncertainties should be considered when determining the weighting or relative importance of risk factors in the risk map. For this study, metadata on the age, coverage, completeness, and accuracy of input data sources were considered as this information points directly to uncertainties which become compounded during processing. Additional sources of uncertainty originating from GIS processing, as well as risk factor performance and interactions observed during the spatial analysis, were considered in determining the relative importance of risk factors.

Three different weight schemes and the equal weight control were used to generate composite maps and assess sensitivity to risk factor weighting (Table 3-3). Scheme 1 considered all available information from metadata, performance, interactions, and uncertainty when assigning risk factor coefficients. Scheme 2 weighted risk factors based on the same information as Scheme 1, but also discounted risk factors presumed to be influenced by other risk factors, theoretically reducing exaggerated risk. Scheme 3 assumed that no interactions exist between risk factors, and only used metadata, performance, and uncertainty to determine the coefficients.

Table 3-3: Summary of risk factors and the coefficients applied to each during the sensitivity analysis.

Risk Factor Type	Risk Factor Name	Control	Scheme 1	Scheme 2	Scheme 3
Categorical	1. Land Cover	1.0	1.0	0.5	2.0
	2. Land Use	1.0	2.0	2.0	2.0
Continuous Over Subunit	3. Population Density	1.0	2.0	2.0	1.5
	4. Percent Impervious	1.0	2.0	2.0	2.0
	5. Development Age	1.0	1.0	1.0	2.0
	6. Outfall Density	1.0	1.0	0.5	1.0
	7. Aging Sanitary Infrastructure	1.0	1.0	1.0	1.5
	8. Drainage Density	1.0	1.0	0.5	1.5
Continuous Over Jurisdictional Area	9. Generating Site Density	1.0	1.0	0.5	2.0
	10. Infrastructure Access Density	1.0	1.0	0.5	1.5
	11. Construction Site Density	1.0	1.0	0.5	1.5
	12. Aging Septic System Density	1.0	1.0	1.0	1.0
	13. Dog License Density	1.0	1.0	0.5	1.0

3.3 RESULTS AND DISCUSSION

3.3.1 Spatial Analysis

After collecting input data, producing the risk factors, and performing the spatial overlay process, summary tables were constructed to show the discharge capture rates for each risk factor overlay. Discharge records were available in the City of Roanoke for a period of 8 years, from October 2007 through April 2015 (Figure 3-3). Each cell in the figure contains the percentage of records in the corresponding discharge or pollutant cluster that were captured by a high-risk factor overlay, also indicated by the cell shading (darker shading indicates a higher percent). The initial results for Roanoke indicate that many illicit discharges are concentrated in regions of medium to high intensity development, and in commercially-zoned areas. Many illicit discharges were located in regions with easy access to MS4 infrastructure, or in regions with higher concentrations of pollution generating facilities or older developments.

City of Roanoke **(n = 90)** 64 Percentage of cluster total captured by high-risk factor overlay	**Illicit Connection**	**Illegal Dumping or Spill**		**Fecal Bacteria**	**Chloride**	**Construction**	**FOG, Surfactants, Trash**	**Hazardous Materials**	**Unknown**
Records in Cluster	7	83		9	0	9	39	31	2
1. Land Cover									
Developed, Open Space	0	0		0	0	0	0	0	0
Developed, Low Intensity	0	16		11	0	11	0	35	0
Developed, Medium Intensity	57	37		67	0	33	36	32	100
Developed, High Intensity	43	47		22	0	56	64	32	0
Barren	0	0		0	0	0	0	0	0
Agricultural	0	0		0	0	0	0	0	0
Undeveloped	0	0		0	0	0	0	0	0
2. Land Use									
Agricultural	0	0		0	0	0	0	0	0
Commercial	14	63		56	0	44	85	29	100
Industrial	43	10		0	0	22	0	29	0
Institutional	0	5		11	0	0	3	6	0
Recreation and Open Space	0	7		11	0	0	3	13	0
Residential	43	16		22	0	33	10	23	0
3. Population Density	0	11		0	0	11	3	19	50
4. Percent Impervious	29	42		33	0	56	46	35	0
5. Development Age	57	61		56	0	67	62	61	50
6. Outfall Density	0	4		0	0	0	0	10	0
7. Aging Sanitary Infrastructure	29	43		33	0	44	54	32	0
8. Drainage Density	29	43		33	0	44	51	35	0
9. Generating Site Density	43	60		56	0	56	67	55	0
10. Infrastructure Access Density	29	46		56	0	44	62	23	0
11. Construction Site Density	14	24		33	0	11	23	23	50
12. Aging Septic System Density	0	2		0	0	0	0	6	0
13. Dog License Density	0	12		0	0	11	5	19	50

Figure 3-3: Discharge capture rates for risk factors in the City of Roanoke.

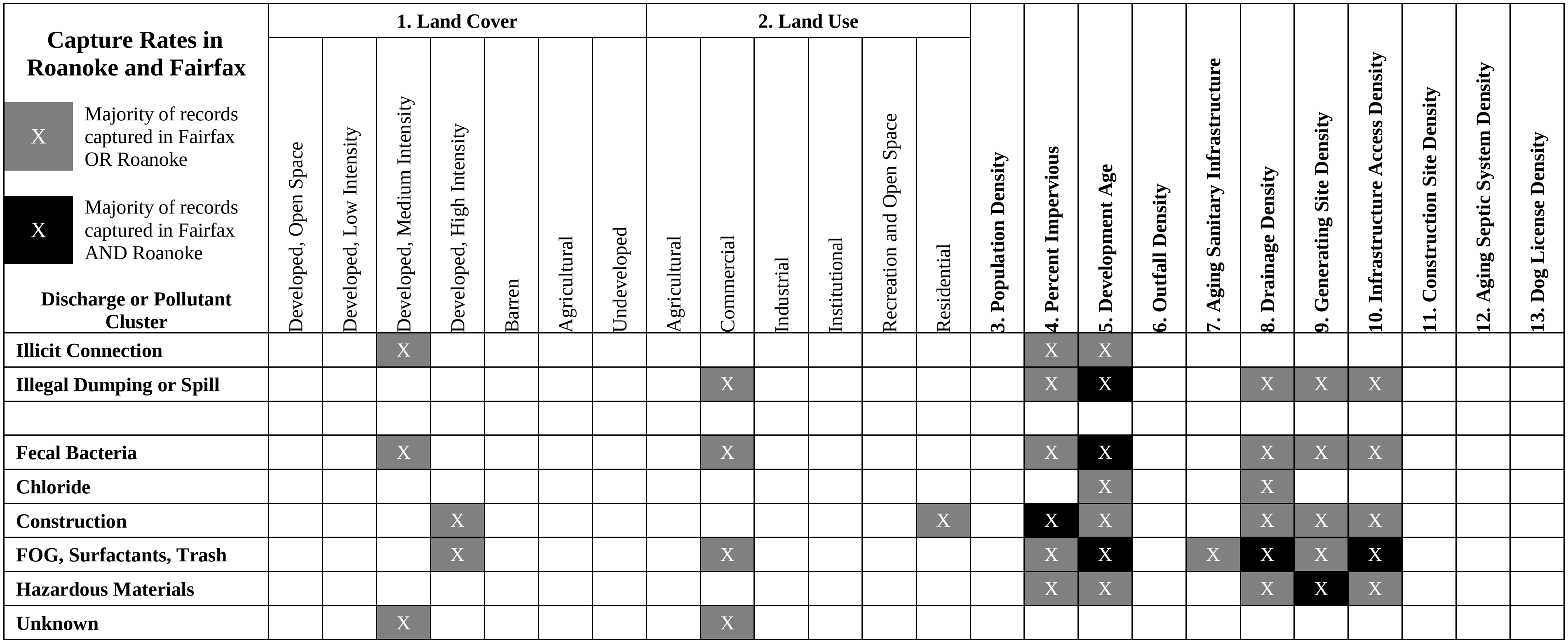
Within the County of Fairfax, discharge records were available for a period of 6 years, from October 2009 through July 2015 (Figure 3-4). The results from Fairfax County suggest that illicit discharges in this jurisdiction are distributed more evenly between multiple land cover types and land uses. However, higher watershed impervious percentages, the presence of more drainage infrastructure and access points, and the presence of older developments all appear to influence the locations of illicit discharges.

County of Fairfax (n = 100) 64 Percentage of cluster total captured by high-risk factor overlay	**Illicit Connection**	**Illegal Dumping or Spill**		**Fecal Bacteria**	**Chloride**	**Construction**	**FOG, Surfactants, Trash**	**Hazardous Materials**	**Unknown**
Records in Cluster	23	77		8	8	12	30	36	6
1. Land Cover									
Developed, Open Space	17	17		25	13	42	13	14	0
Developed, Low Intensity	30	16		38	50	33	7	11	33
Developed, Medium Intensity	30	32		13	13	25	33	42	33
Developed, High Intensity	4	30		13	13	0	40	28	0
Barren	0	0		0	0	0	0	0	0
Agricultural	0	1		0	0	0	0	3	0
Undeveloped	17	4		13	13	0	7	3	33
2. Land Use									
Agricultural	0	0		0	0	0	0	0	0
Commercial	30	34		25	13	8	47	36	33
Industrial	4	21		0	0	8	27	22	0
Institutional	4	4		13	13	8	0	3	0
Recreation and Open Space	30	10		13	38	0	10	19	17
Residential	30	31		50	38	75	17	19	50
3. Population Density	4	0		0	0	0	0	3	0
4. Percent Impervious	61	71		88	50	58	70	75	50
5. Development Age	39	56		75	63	42	57	44	50
6. Outfall Density	26	14		13	38	25	13	17	0
7. Aging Sanitary Infrastructure	48	44		13	38	50	47	50	50
8. Drainage Density	48	60		75	63	58	57	53	50
9. Generating Site Density	43	36		25	13	17	37	53	50
10. Infrastructure Access Density	48	64		38	13	58	73	69	33
11. Construction Site Density	43	30		25	0	33	37	44	0
12. Aging Septic System Density	4	5		0	13	0	7	3	17
13. Dog License Density	4	26		25	0	33	30	11	33

Figure 3-4: Discharge capture rates for risk factors in the County of Fairfax.

Results from the spatial overlay process in both case study jurisdictions were combined to identify similarities among risk factors which might influence capture rates (Figure 3-5). Cells in the figure with an “X” indicate risk factors which captured 50% or more of the records in the corresponding cluster. Lighter shaded cells indicate that a majority of records in the cluster was captured in only one jurisdiction, while darker shaded cells indicate that a majority was captured in both jurisdictions. Figure 3-5 shows that several discharge and pollutant clusters were consistently captured at a higher frequency by five risk factors, namely infrastructure access density, generating site density, percent impervious, development age, and drainage density. The similarities in the performance of these risk factors across a wide array of discharges can be explained by trends in development and urbanization.

The characteristics of developed areas determine the risk present for various types of discharges. For instance, developments with a high percent impervious surface generate more runoff, which requires more subsurface drainage infrastructure. High infrastructure densities provide more opportunities for illicit connections and dumped or spilt materials to enter the stormwater network. In Fairfax and Roanoke, illegal dumping and spills in dense commercial areas often contained FOG, surfactants, and trash, while residential areas experienced more dumping and spills related to construction. Additionally, fecal bacteria discharges were found in commercial and residential land uses, but may have originated from different sources. Commercial areas have more concentrated businesses and infrastructure which facilitates illicit connections, while residential areas have larger dog populations which may result in more dumping or bacterial runoff.



Capture Rates in Roanoke and Fairfax

▨ X Majority of records captured in Fairfax OR Roanoke

■ X Majority of records captured in Fairfax AND Roanoke

	1. Land Cover							2. Land Use						3. Population Density	4. Percent Impervious	5. Development Age	6. Outfall Density	7. Aging Sanitary Infrastructure	8. Drainage Density	9. Generating Site Density	10. Infrastructure Access Density	11. Construction Site Density	12. Aging Septic System Density	13. Dog License Density
Discharge or Pollutant Cluster	Developed, Open Space	Developed, Low Intensity	Developed, Medium Intensity	Developed, High Intensity	Barren	Agricultural	Undeveloped	Agricultural	Commercial	Industrial	Institutional	Recreation and Open Space	Residential											
Illicit Connection			▨ X												▨ X	▨ X								
Illegal Dumping or Spill									▨ X						▨ X	■ X			▨ X	▨ X	▨ X			
Fecal Bacteria			▨ X						▨ X						▨ X	■ X			▨ X	▨ X	▨ X			
Chloride																▨ X			▨ X					
Construction				▨ X									▨ X		■ X	▨ X			▨ X	▨ X	▨ X			
FOG, Surfactants, Trash				▨ X					▨ X						▨ X	■ X		▨ X	■ X	▨ X	■ X			
Hazardous Materials															▨ X	▨ X			▨ X	■ X	▨ X			
Unknown			▨ X						▨ X															

Figure 3-5: Summary of risk factor capture rates for both Roanoke and Fairfax.

The processes by which urbanization and development occur also play a part in determining IDP. Older developments and aging infrastructure were often present in areas of Fairfax and Roanoke where illicit connections produced large numbers of bacteria, chloride, surfactant, or hazardous material discharges. Outdated building codes and poor construction practices may have affected where these discharges occurred, although illicit connections are known to occur in more recent developments when construction is not thoroughly reviewed and inspected. Building codes in Roanoke may also explain why no chloride discharges have been reported during the period of record, although this absence could also be due to differences in screening or reporting programs.

In Fairfax County, a temporal trend due to changes in reporting mechanisms was observed to affect the number and type of illicit discharges reported during a given year. From 2009 to 2014, illicit discharges were discovered through the County’s dry weather screening program, which focused on detection of illicit connections, and found mostly cooling tower and surfactant discharges. The total number of discharges found per year was also very low, since most outfalls that were screened did not show signs of chronic discharges from illicit connections. In 2014, a complaints hotline reporting mechanism was implemented by the County. Data from 2014-2015 show that within two years this additional reporting mechanism increased the total number of illicit discharge records reported since 2009 by 300 percent. In addition to the increased number of records reported each year, records were spread more evenly across discharge and pollutant clusters.

The spatial overlay process revealed the importance of a handful of risk factors, but when combined with the results of the Fisher analysis, a more comprehensive assessment of risk factor performance could be made. The two-tailed Fisher tests identified several risk factors which had multiple significant comparisons ($p < 0.05$). While a few of these risk factors were unique to one jurisdiction or another, seven factors exhibited similar results across both jurisdictions. Those risk factors were infrastructure access density, generating site density, construction site density, percent impervious, development age, drainage density, and aging sanitary infrastructure. However, while the two-tailed Fisher tests provided an idea of which risk factors were statistically different in their performance, the tests did not indicate which factor performed better in each comparison.

The one-tailed Fisher tests determined which risk factors performed better than others, as measured by a statistically significantly larger capture proportion, and allowed for risk factor rankings to be produced. A slightly higher number of significant results were returned in the one-tailed analysis over the two-tailed analysis, which was expected since the one-tailed test is less conservative. After the performance rankings were developed from the capture proportions and Fisher analyses, it was clear that many of the risk factors identified as significant in the spatial overlay analysis and the two-tailed Fisher analysis outperformed most other risk factors. These risk factors consistently ranked in the top five factors for predicting any given discharge or pollutant cluster due to their high capture rates and resulting significant p-values. In addition, another group of risk factors consistently ranked in the top 50% in both jurisdictions and across discharge and pollutant clusters, but occupied different ranks depending on individual performance in each cluster. Likewise, a separate set of risk factors consistently performed poorly across jurisdictions and clusters. These observations suggest that certain risk factors are generally more applicable to predicting IDP in urban and suburban jurisdictions; however, the results of the one-tailed Fisher tests do not account for interactions between risk factors which may affect performance.

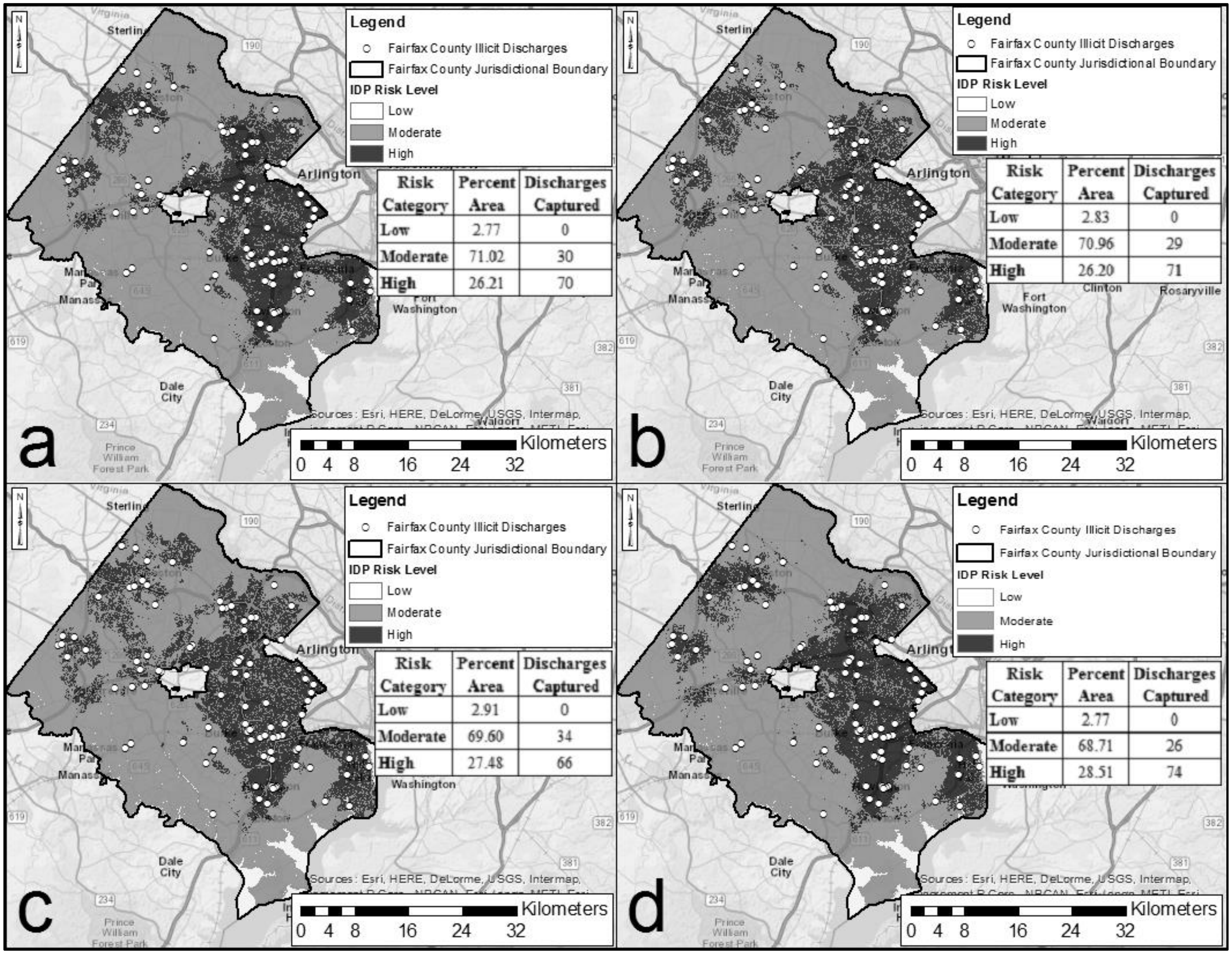
The analysis of potential risk factor interactions found that high-risk overlays for percent impervious, land use, and population density appear to influence the risk associated with several risk factors. The causes of the observed interactions are the same development and urbanization trends that were identified during the spatial overlay analysis. High percent impervious surface leads to increased drainage infrastructure, while commercial and residential land use designations cause increases in generating site density and dog license density, respectively. High-risk land uses also correspond to regions of higher intensity development, which explains the overlap between land use and land cover. Furthermore, densely populated regions are more likely to exhibit higher dog license density due to the increase in potential pet owners. Construction is also more prevalent in more populated areas due to the need for upgraded infrastructure or increased housing demand. An interesting observation was that development age did not show any definitive relationship with aging septic system density and aging sanitary sewer infrastructure, and appeared to be related to percent impervious. The septic and sanitary risk factors also appeared to interact with each other in different manners depending upon the jurisdiction in question.

The rankings produced from the Fisher analyses combined with the results of the risk factor interactions suggest that risk factor performance across urban and suburban jurisdictions is relatively consistent, and can be effectively modeled using a subset of six factors: population density, land use, percent impervious, development age, infrastructure access density, and generating site density. It should be noted that although generating site density and infrastructure access density do appear to interact with the percent impervious surface and land use risk factors, they are still included due to their ability to identify isolated discharges and hot spots. This subset of six risk factors should provide urban and suburban MS4s with a relatively cost-effective risk map to identify high-risk areas for all discharge and pollutant clusters. MS4s would have the option to include additional risk factors based on local concerns and prior knowledge or experience. If an even smaller set of risk factors is desired, as in the case of small MS4s with tight budget constraints, the land use, development age, and infrastructure access density risk factors would be the most effective combination to pursue. These three risk factors provide the best chance to capture the majority of discharges across all clusters analyzed, while highlighting both large high-risk areas and isolated hot spots.

More tests and data from other MS4s would be required to determine if these observations hold true when comparing urban and suburban MS4s to rural jurisdictions. It is important to note that clusters with lower discharge counts produced fewer significant results during the Fisher analyses, indicating that clusters with less than 10 records are still too small for the Fisher exact test to detect a statistically significant difference in risk factor performance.

3.3.2 Sensitivity Analysis

In evaluating the sensitivity of composite risk maps, the weighting process attempts to mitigate data quality and uncertainty issues which originate from (1) the inputs to the risk model, and (2) data processing during model development. Sensitivity analysis also attempts to incorporate risk factor performance and interactions into the model to improve the estimate of IDP in the final composite map. The composite risk maps developed for Fairfax County (excluding the City of Fairfax at the center of the County) (Figure 3-6) and Roanoke (Figure 3-7) show a gradual increase in high-risk area within each jurisdiction, moving from the control up to Scheme 3. However, the different schemes emphasize this high-risk area in different geographic regions within each jurisdiction. In Fairfax County, all maps show higher risk of illicit discharge around Washington, D.C., due to older developments and more impervious surfaces, but Schemes 2 and 3 also emphasize increased risk away from Washington, D.C. This emphasis is due to discounting or weighting of risk factors which bring out aging septic and sanitary infrastructure problems in the western portion of the County (Figure 3-6b and c). In Roanoke, the central downtown region is emphasized across all maps due to its age and impervious surface; however, commercial corridors and residential areas farther out from the city core are emphasized based on the relative impact of weighted risk factors such as land use and aging septic system density (Figure 3-7b and c). There are also differences in the types of land classified as high risk. In Fairfax County, there is more emphasis on land parcels than roads (Figure 3-6b and c) due to increased weight on the land use risk factor. This trait does not show up in Roanoke (Figure 3-7b and c) because in Roanoke, the land use risk factor was produced using zoning polygons which covered parcels and roads, while in Fairfax, the land use risk factor was produced using existing land use data assigned to parcels.



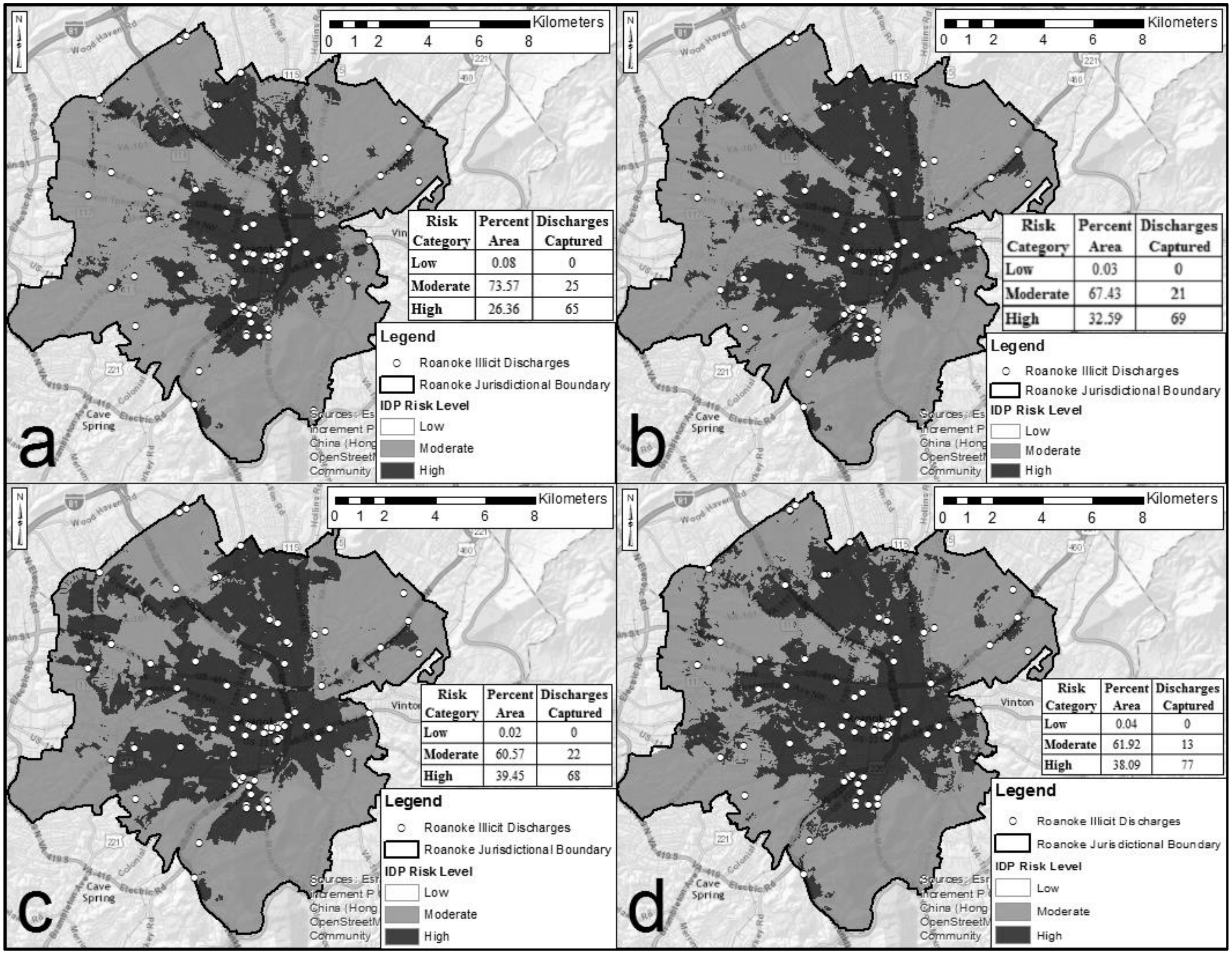
Risk Category	Percent Area	Discharges Captured
Low	2.77	0
Moderate	71.02	30
High	26.21	70

Risk Category	Percent Area	Discharges Captured
Low	2.83	0
Moderate	70.96	29
High	26.20	71

Risk Category	Percent Area	Discharges Captured
Low	2.91	0
Moderate	69.60	34
High	27.48	66

Risk Category	Percent Area	Discharges Captured
Low	2.77	0
Moderate	68.71	26
High	28.51	74

Figure 3-6: County of Fairfax composite IDP risk maps produced using (a) the control weights, (b) Scheme 1, (c) Scheme 2, and (d) Scheme 3.



(a)

Risk Category	Percent Area	Discharges Captured
Low	0.08	0
Moderate	73.57	25
High	26.36	65

(b)

Risk Category	Percent Area	Discharges Captured
Low	0.03	0
Moderate	67.43	21
High	32.59	69

(c)

Risk Category	Percent Area	Discharges Captured
Low	0.02	0
Moderate	60.57	22
High	39.45	68

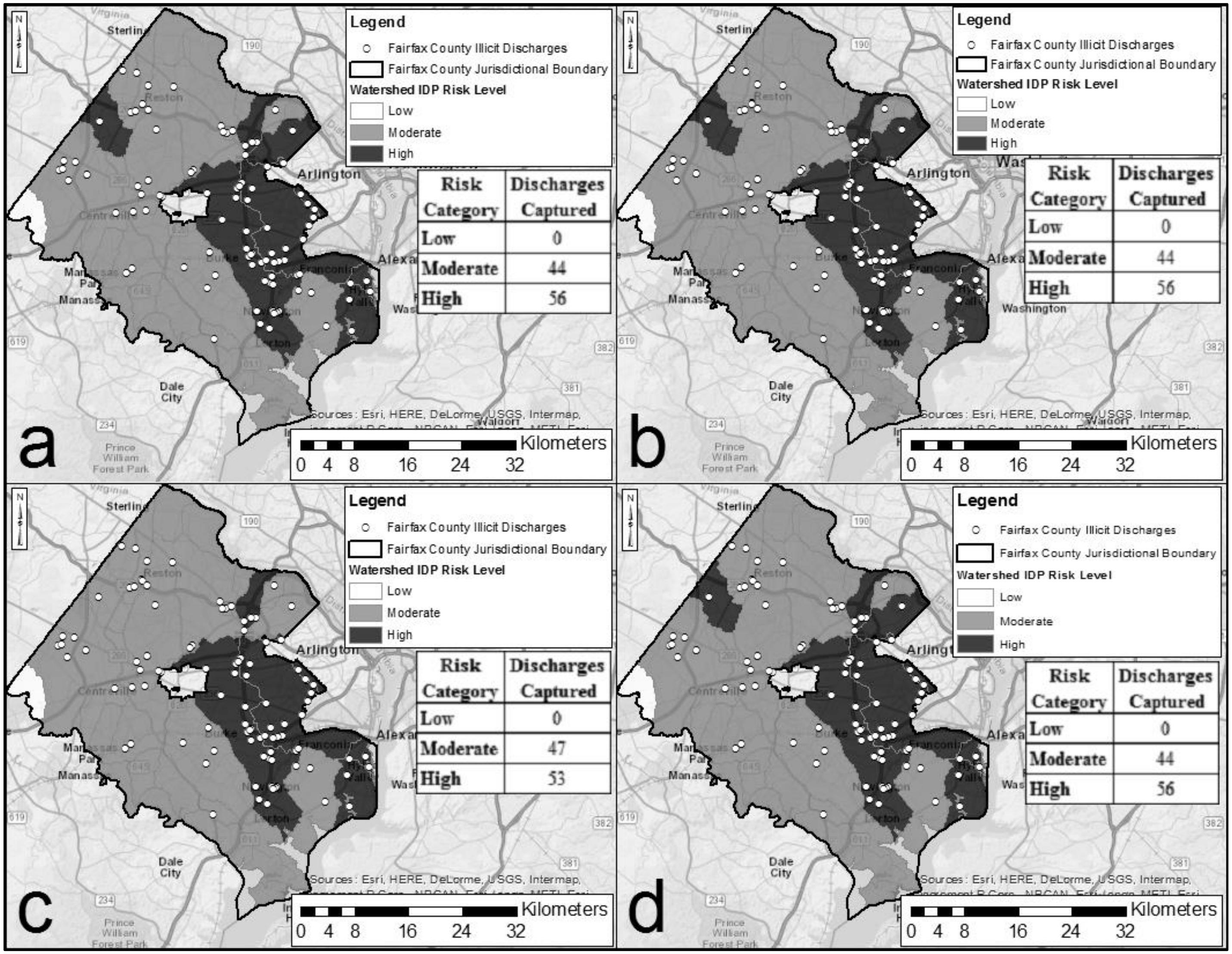
(d)

Risk Category	Percent Area	Discharges Captured
Low	0.04	0
Moderate	61.92	13
High	38.09	77

Figure 3-7: City of Roanoke composite IDP risk maps produced using (a) the control weights, (b) Scheme 1, (c) Scheme 2, and (d) Scheme 3.

The distribution of land area and discharge records over the three risk levels in each composite map provide a numerical measure of the sensitivity of the overall risk map to different weight schemes. While the four schemes provide similar results, Scheme 3 outperformed most others in terms of total high-risk area and proportion of discharges falling in high-risk areas, while Scheme 2 underperformed when compared to the control. Scheme 3 was developed with consideration given to metadata, risk factor performance, and uncertainty, but without consideration to the possibility of interactions among risk factors which could exaggerate risk. Therefore, it is possible that while Scheme 3 captures the most illicit discharge records, it may over-estimate the high-risk area in a jurisdiction leading to the increase in records captured. Similarly, Scheme 2, which takes interactions into account by weighting root factors and discounting related risk factors, underperforms by increasing the high-risk area but decreasing the number of records captured, suggesting that Scheme 2 emphasizes areas which may not actually have high IDP. It is likely that Scheme 1, which considers all of these conditions, including weighting root factors, is a better improvement over the control. Regardless of the scheme used, the composite risk map will not capture all discharges due to uncertainties in the input data, as well as the partially random nature of illicit discharge occurrence.

Aggregating IDP scores using watersheds as subunits of analysis revealed that 25% of the watersheds captured 55% and 43% of the illicit discharge records in Fairfax and Roanoke, respectively. The watersheds labeled as high-risk in Fairfax County were fairly consistent regardless of the coefficients applied to the risk factors (Figure 3-8). This may be caused by the intense development that has occurred near the Washington, D.C. area, which produced high-risk overlays in several risk factors, including multiple factors analyzed over watershed subunits. In the City of Roanoke, two watersheds are indicated to be at high-risk in every sensitivity scenario, while the risk levels of several other watersheds varied based on the coefficients applied to form the composite map (Figure 3-9). This suggests that the two watersheds at the core of the city are inherently high-risk due to multiple risk factors, while other watersheds have specific problems which cause their relative risk to change based on changes in the weights applied to certain risk factors. Roanoke watersheds are also smaller than those in Fairfax County, which allows smaller increases in high-risk area to shift the mean IDP score towards a high-risk value.



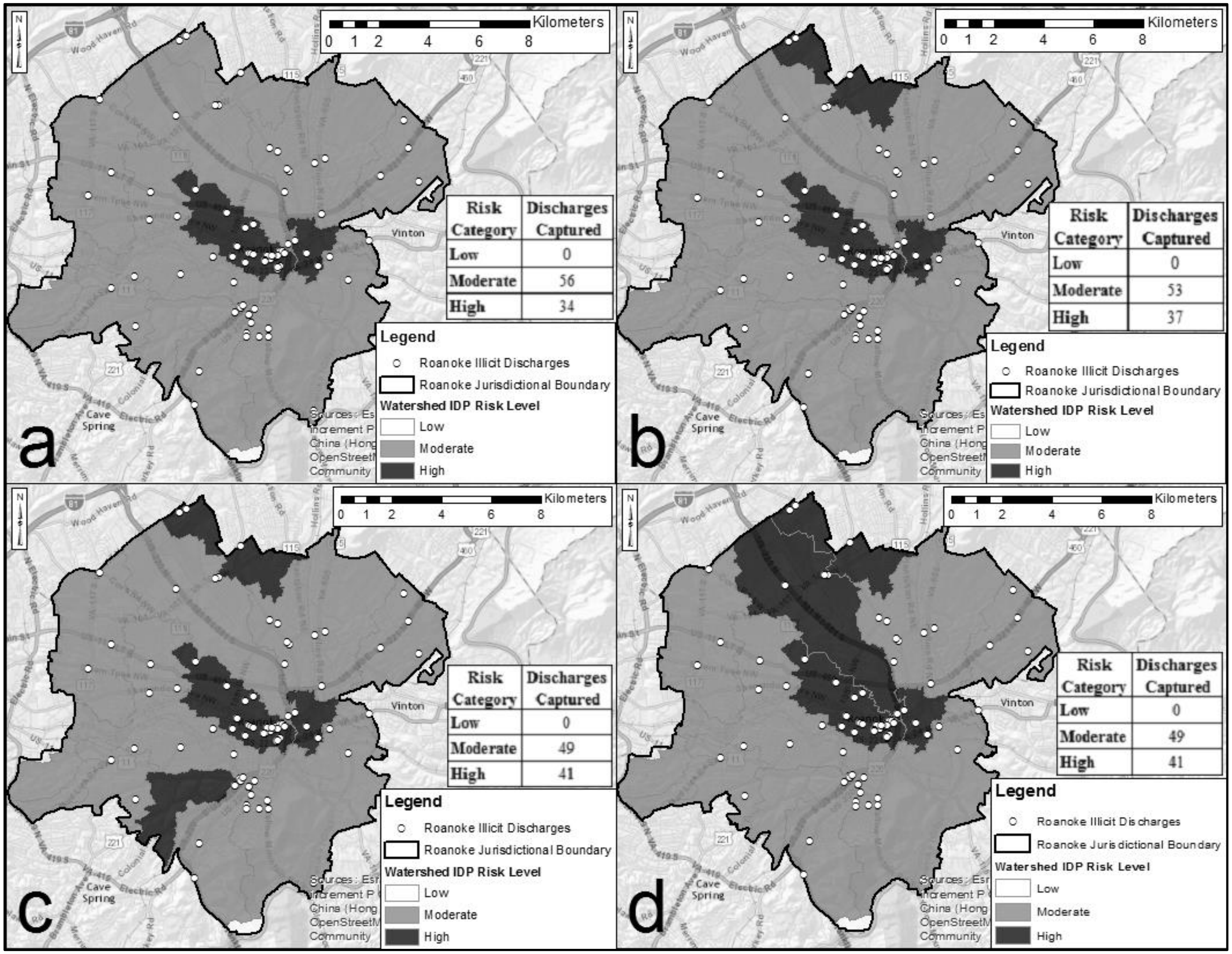
Risk Category	Discharges Captured
Low	0
Moderate	44
High	56

Risk Category	Discharges Captured
Low	0
Moderate	44
High	56

Risk Category	Discharges Captured
Low	0
Moderate	47
High	53

Risk Category	Discharges Captured
Low	0
Moderate	44
High	56

Figure 3-8: IDP scores aggregated over Fairfax County watersheds for (a) the control weights, (b) Scheme 1, (c) Scheme 2, and (d) Scheme 3.



a

Risk Category	Discharges Captured
Low	0
Moderate	56
High	34

b

Risk Category	Discharges Captured
Low	0
Moderate	53
High	37

c

Risk Category	Discharges Captured
Low	0
Moderate	49
High	41

d

Risk Category	Discharges Captured
Low	0
Moderate	49
High	41

Figure 3-9: IDP scores aggregated over Roanoke City watersheds for (a) the control weights, (b) Scheme 1, (c) Scheme 2, and (d) Scheme 3.

While a watershed IDP approach appears to capture a large number of discharge records, the number of records captured by high-risk watersheds is significantly lower than the number of records captured by high-risk raster cells in the original raster-based composite maps. This difference in risk map performance is due to the calculation of the mean IDP score. When the average IDP score is evaluated over a watershed containing high-risk hot spots surrounded by large low- or moderate-risk areas, the hot spot scores are diluted. Therefore, the original raster-based composite maps, which evaluated IDP over 900-square-meter cells, are more precise at determining the location of hot spots and pinpointing problems which are isolated to a smaller region.

3.3.3 Limitations of Risk Mapping

The sensitivity analysis conducted in this study attempts to improve the risk map by accounting for the effects of three variables: (1) risk factor performance, (2) risk factor interactions, and (3) uncertainty. Uncertainty represents the most difficult variable to quantify due to the sheer number of datasets and data sources utilized in the creation of the composite map; therefore, uncertainty significantly limits the conclusions that can be drawn from composite risk maps. During the sensitivity analysis, several sources of uncertainty were observed to affect the overall quality of the risk maps produced in this study. Discussion of major sources of uncertainty is given below, while an extended discussion of these issues is provided in Appendix D – Extended Discussion of Risk Factor Uncertainty.

The most important source of uncertainty in the model originates from datasets which cause the over- or underestimation of IDP. This problem was especially evident in the five risk factors evaluated over watersheds within each jurisdiction: percent impervious, development age, outfall density, aging sanitary infrastructure, and drainage density. The overall unit of analysis was a political boundary independent of hydrology, and watersheds on the fringes of the study area were clipped to this boundary. As the five risk factors analyzed over watersheds were linked to hydrologic conditions, clipping the fringe watersheds to political boundaries produced partial datasets, reflecting an incomplete analysis and leading to significant over- and underestimation of risk.

All risk factors may have been impacted by the political boundary as each dataset used in this analysis was clipped to this boundary. As a result of clipping this data, the effects of conditions in neighboring jurisdictions were not considered, and therefore the risk at the fringes of the case study jurisdictions may have been generally over- or underestimated. Risk also may have been over- or underestimated through the binning of continuous variables into qualitative risk levels. Eleven of the thirteen risk factors analyzed in this study represent continuous variables, and thresholds were assigned using generally recommended thresholds from previous studies or design guidelines. However, where guidelines were unavailable, or where values were significantly different from those observed in other studies, the choice of thresholds between low, moderate, and high risk would have impacted overall IDP.

Another primary source of uncertainty in this model arises from the availability of high-quality data to evaluate risk factors. Multiple risk factors in this study, including aging septic system density, development age, and land use, used surrogate information in place of some desired dataset which did not exist, or was not easily accessible. The quality and relevance of the surrogate data directly contribute to uncertainty in the model. However, any available dataset exhibits a certain level of inaccuracy, and will often have gaps in coverage. Inaccurate data may simply be out of date, as with census population data and NLCD land cover and impervious surface data, which are collected infrequently. Additionally, if the data collection methodology is not thorough enough, specific attributes or large subsections of the dataset may not be collected. Data completeness, age, and accuracy problems were common in many datasets used in this analysis, and underscore the importance of metadata in determining the appropriate weight to assign to a risk factor.

An additional source of uncertainty in this model stemmed from the format of input data and the risk factor production process. Many localities maintain databases in non-spatial formats, such as Microsoft Excel or Access, which had been the standard practice for many years before geographic information systems started to gain popularity. These non-spatial databases are convenient for the mass management of dog licenses, real estate tax information, development records, and illicit discharge records. Since risk mapping involves the manipulation of data in a spatial environment, it was necessary to translate large amounts of non-spatial data into a spatial database. The poor quality of the tabular records and the spatial datasets used to locate those records on a map often resulted in only partial transfer of the dataset, compromising input data quality and increasing uncertainty. Data format was also a problem between spatial datasets because datasets with differing resolutions could not easily be combined. Since the coarsest dataset defines the limiting resolution in spatial operations, fine resolution datasets were distorted or partially lost as the datasets were changed to the coarser resolution. Other tools generated uncertainty during risk factor production by changing polygons into points (affected density calculations) and converting between vector and raster datasets (resulted in a loss of resolution).

While some sources of uncertainty can be mitigated or avoided by carefully considering the available sources of input data and the processing necessary to produce the desired risk factor, much of the uncertainty affecting IDP is inherent in the available data. During the risk mapping process, it is important to identify risk factors which are produced from highly uncertain datasets. The ability of a risk map to identify discharges related to those risk factors will be limited, unless the level of uncertainty can be reduced through additional data collection or quality control.

3.3.4 Implications

The utility of an IDP risk map to MS4s is realized in the prioritization of illicit discharge programs. Field reconnaissance and facility inspections can be prioritized based on the local IDP risk level, such that facilities and outfalls in high-risk areas are inspected before, or are inspected more frequently than those in low-risk areas. Public education and other programs designed to help prevent illicit discharges can be targeted to specific high-risk neighborhoods within a locality. By prioritizing these programs, localities should be in a better position to demonstrate compliance with MS4 permit requirements and achieve a higher cost-efficiency in their illicit discharge programs.

There is an important distinction to be noted between illicit discharge potential and illicit discharge occurrence. A risk map made with high quality data will still only point to areas with high discharge potential, meaning that it is more likely that discharges will occur in high-risk areas, but that discharges will likely still occur in areas with a very low perceived risk. Therefore, risk maps should not be used to eliminate regions for inspections or other programming, but serve as a tool to focus or direct those programs towards larger high-risk areas to improve efficiency. After the major problem areas are addressed, a second round of investigations may be run to focus on smaller high-risk hot spots, and a third round may attend to low-risk communities to identify smaller, more isolated problems that may go unnoticed during risk analysis. The ultimate goals of IDP risk mapping are to integrate the risk mapping process with ongoing and emerging programs and initiatives, and for the risk map to act as a tool to guide program planning.

In addition to producing a composite risk map for overall IDP, the concept can be extended to create risk maps for specific pollutants or local objectives. This would involve selecting a subset of relevant risk factors from the overall IDP map, combining those risk factors with other local knowledge or indicators describing the pollutant of concern, and producing a new composite map to inform relevant programs. In the case of Roanoke, where MS4 operators have identified bacteria as a pollutant of concern, a bacteria IDP risk map could inform programs related to sanitary sewer inspection and maintenance, as well as septic-to-sanitary utility projects for residential property owners.

3.4 CONCLUSIONS

This study has demonstrated that, in the City of Roanoke and the County of Fairfax, many land use, hydraulic, and hydrologic characteristics affect the location of illicit discharges. In particular, three key risk factors, namely land use, percent impervious, and population density, were found to influence the risk associated with other risk factors in the analysis in accordance with trends in development and urbanization. The spatial analysis yielded rankings of risk factor performance which indicated that a handful of risk factors, namely percent impervious, land use, development age, population density, infrastructure access density, and generating site density, are most important in determining IDP in urban and suburban MS4s. MS4 operators should focus their mapping efforts on analyzing these risk factors, especially since many of the datasets required to map them are, or will be required to fulfill MS4 permit requirements in Virginia. Smaller or more cash-strapped MS4s may be able to narrow their risk mapping efforts to just the land use, development age, and infrastructure access density risk factors to achieve comparable, cost-effective results. Additional risk factors may be added to a composite risk map, but MS4 operators should be aware that interactions between the additional and the recommended factors may exaggerate IDP risk if not properly accounted for during production of the composite map.

The sensitivity analysis concluded that it is imperative to consider risk factor performance, risk factor interactions, and uncertainty to develop an effective composite risk map that controls for over- and underestimation of IDP. A functional weight scheme will assign more importance to higher quality, better performing risk factors, while minimizing the exaggerated risk contributed by risk factor interactions. An exact weight scheme will depend on local data and must be chosen based on a thorough evaluation of the quality of available datasets, and an assessment of the performance of and interactions among selected risk factors.

Uncertainty arising from input data and processing techniques will ultimately limit the power of an IDP risk map. This uncertainty can be mitigated or reduced through careful planning before beginning the risk mapping process. MS4 operators should consider the available sources of data within the MS4 organization, as well as publicly available data. While public data sources may be useful, MS4 operators should consider whether the public dataset is sufficient, or if it requires additional data collection or manipulation on the local scale. After thoroughly reviewing available data, MS4s should determine the most appropriate mapping scale and unit of analysis that will fit local needs and available data formats. Risk analysts should develop a GIS procedure which minimizes data processing and data misrepresentation. For internal datasets, MS4s can reduce data uncertainty by storing data in manners that are more conducive to spatial analysis, and by performing routine quality control on local datasets. In many states as in Virginia, existing regulations and permit requirements may provide an incentive to transition to or improve upon GIS as a universal data storage and asset management tool, which will further improve cost efficiency and illicit discharge program management in the long term.

4 CONCLUSION

4.1 IMPLICATIONS

This study has demonstrated that, in the City of Roanoke and the County of Fairfax, many land use, hydraulic, and hydrologic characteristics affect the location of illicit discharges. While many of these factors are interrelated, several factors appear to be effective indicators of illicit discharge potential in urban and suburban areas. These risk factor interactions may be mitigated during the production of a composite map by applying a weight scheme that emphasizes the risk factors that appear to be the root cause of interactions.

This study supports that IDP risk mapping is an effective tool that should be utilized by MS4s across the nation to structure the implementation of illicit discharge programming. The risk map provides the framework for a tiered approach to addressing illicit discharge risk. Large regions identified as high-risk can be assigned first priority for facility inspections, educational mailings, and capital improvement projects. Outfalls draining these areas can also be given first priority during outfall reconnaissance operations. After the major problem areas are addressed, a second round of investigations may be run to focus on smaller high-risk hot spots, and a third round may attend to low-risk communities to identify smaller, more isolated problems that may go unnoticed during risk analysis. This plan of attack allows jurisdictions to apply staff and financial resources first to regions predicted to have the most problems, resulting in program-wide cost savings.

The following recommendations provide guidance that will help MS4s to implement effective risk mapping techniques while accounting for risk factor performance, interactions, and uncertainty.

- MS4 operators should focus on mapping land use, development age, and infrastructure access density, and should add percent impervious surface, population density, and generating site density if resources and data are available. While other risk factors may be added to incorporate local concerns or objectives, these six risk factors performed the best in terms of identifying overall IDP while minimizing interactions.
- During the initial planning process, MS4 operators should gather data from local sources and public databases, and review the quality and relevance of each dataset. MS4 operators should consider whether public and local data should be combined to create a more complete picture of the local conditions.

- MS4 operators should determine whether the available datasets are sufficient, or if local data collection or quality control is necessary to improve input data accuracy. This task may be coordinated with other permit requirements or data maintenance initiatives, such as field infrastructure surveys to develop an outfall map. A GIS organizational structure for data storage is recommended for any data collection undertaken by the MS4.
- Once mapping data is collected and reviewed, MS4 operators should determine the appropriate mapping parameters for the available data. Parameters that should be considered are the unit of analysis (e.g. jurisdictional area, MS4 service area, watersheds, neighborhoods, etc.) and the output data format and scale (e.g. vector or raster maps and raster cell size). Finer mapping scales are recommended due to their ability to detect smaller hot spots.
- MS4 operators may want to coordinate with technology or GIS personnel when possible to develop mapping parameters and a GIS procedure that will produce risk factors efficiently and in a manner which minimizes the potential to distort the input data.
- When producing a composite map from individual risk factors, a weight scheme should be developed that incorporates risk factor performance, interactions, and uncertainty on a jurisdictional basis. These variables can be evaluated during the risk mapping process, and the resulting scheme will further mitigate interactions and uncertainties unique to each jurisdiction.

4.2 FUTURE WORK

While this study has generated some promising results regarding the capabilities of IDP risk mapping, this tool has only recently been implemented, and further study is warranted in several respects. As the case study jurisdictions evaluated here are highly urbanized with little agricultural land, the recommended risk factors may not be generally applicable across all jurisdictions under the MS4 program. Many MS4s contain more rural agricultural land in addition to census-designated urbanized areas. While the MS4 permit only regulates urbanized areas, water quality issues may originate from both regulated and unregulated lands; therefore, an MS4 may benefit from performing an IDP risk assessment over its entire jurisdiction. Different risk factors may need to be evaluated for performance in rural or agricultural regions as compared to urban or suburban MS4s to provide a more accurate description of IDP.

Several opportunities also exist to improve the current risk mapping process for urban and suburban MS4s. Case studies can be performed to evaluate other risk factors not investigated in this study, and to improve risk factor performance by developing more complete datasets for individual risk factors.

Improvements can also be made to the GIS procedures used for risk factor development to reduce the distortion of input data, as well as to include the effects of conditions in neighboring jurisdictions on the IDP risk at the jurisdictional boundary. While risk maps focused on specific pollutants of concern were mentioned as a potential extension of the overall IDP risk map, they were not evaluated in this study. More research to evaluate the effectiveness of risk maps targeting pollutants of concern, and the cost-effectiveness of these risk maps and the overall IDP risk map, may provide more information to support the implementation of this tool in illicit discharge programs.

4.3 FINAL WORDS

As urbanization and development continue, and pollution problems increase in spite of current regulations, illicit discharges will move further into the spotlight of environmental regulators. Local enforcement and public awareness campaigns will continue to be critical components of MS4 Program Plans to affect changes in the human behaviors and practices which lead to many illicit discharges. As a municipality or jurisdiction subject to NPDES and MS4 legislation, it is imperative to improve the effectiveness of illicit discharge programs, not only to demonstrate compliance with permit obligations, but also to move towards expedient water quality improvement.

REFERENCES

Baird, R. C. (2009). Coastal Urbanization: the challenge of mangement lag. *Management of Environmental Quality: An International Journal*, *20*(4), 371–382.

Booth, D. B., Hartley, D., and Jackson, R. (2002). Forest Cover, Impervious-Surface Area, and the Mitigation of Stormwater Impacts. *JAWRA Journal of the American Water Resources Association*, *38*(3), 947–835. doi:10.1111/j.1752-1688.2002.tb01000.x.

Booth, D. B., and Jackson, C. R. (1997). Urbanization of aquatic systems: Degradation thresholds, stormwater detection, and the limits of mitigation. *Journal of the American Water Resources Association*, *33*(5), 1077–1090. doi:10.1111/j.1752-1688.1997.tb04126.x.

Brown, E., Caraco, D., Pitt, R., and Rittenhouse, B. (2004). *Illicit Discharge Detection and Elimination: A Guidance Manual for Program Development and Technical Assessments*. Retrieved from http://www.ncsu.edu/ehs/environ/IDDE Guidance Manual.pdf.

Catalfio, C., Hughes, C., Zorza, G., Mullett, N., and Tuomari, D. (2007). Alternatives for Finding Illicit Discharges to Stormwater Systems. In *WEFTEC* (pp. 2079–2104). Water Environment Federation.

Center for Watershed Protection, and Biohabitats. (2011). *Healthy harbor plan: A plan to make Baltimore Harbor swimmable and fishable by 2020*. Ellicott City, MD; Baltimore, MD.

Chandler, D. M., and Lerner, D. N. (2015). A low cost method to detect polluted surface water outfalls and misconnected drainage. *Water and Environment Journal*, 1–5. doi:10.1111/wej.12112.

Christian, D., Scholl, J. E., Herrmann, J., and Phillips, S. (2004). A transportation authority's perspective on an illicit discharge elimination program. In *Watershed 2004* (pp. 1607–1612). Dearborn, MI: Water Environment Federation. doi:10.1017/S000748530002229X.

City of Roanoke. (2015). Storm Utility Fee and Credits. Retrieved March 14, 2016, from http://www.roanokeva.gov/1843/Storm-Utility-Fee-and-Credits.

DEQ. (2013). *9 VAC 25-890 - General Permit for Discharges of Stormwater from Small Municipal Separate Storm Sewer Systems*. Richmond: Virginia Department of Environmental Quality.

DEQ. (2016). Virginia Department of Environmental Quality. Retrieved December 17, 2015, from http://www.deq.virginia.gov/.

Dymond, R. L., Aguilar, M. F., Bender, P., and Hodges, C. (2015a). *Lick Run Watershed Master Plan*. City of Roanoke, VA.

Dymond, R. L., Aguilar, M. F., Bender, P., and Hodges, C. (2015b). *Roanoke Urban Stormwater Research: Phase II Final Report*. Blacksburg, VA.

EDC and NRDC versus USEPA. (2014). Environmental defense center and natural resources defense council versus USEPA. Petition for a writ of mandamus.

Eddy, N. (2000). Arkansas Sanitarian Uses Infrared Technology to Track Down Sewage. *Small Flows Quarterly*, *1*(2), 22–24. doi:10.1017/S000748530002229X.

Environmental Systems Research Institute. (2015). *ArcGIS for Desktop 10.3.1*. Redlands, CA.

EPA. (2016). Enforcement and Compliance History Online. Retrieved December 17, 2015, from https://echo.epa.gov/.

Fairfax County. (2010). Fairfax County Stormwater Service District. Fairfax County, VA: Fairfax County, VA.

Fairfax County. (2016). Fairfax County GIS and Mapping. Retrieved December 17, 2015, from http://www.fairfaxcounty.gov/maps/.

Fairfax County Stormwater Planning Division. (2014). *Dry Weather Screening Standard Operating Procedures*. Fairfax County, VA.

Fisher, R. A. (1970). The Exact Treatment of 2x2 Tables. In E. A. Cornish (Ed.), *Statistical Methods for Research Workers* (14th ed., pp. 96–97). Darien, CT: Hafner.

Fowler et al. v. USEPA. (2010). C. Bernard Fowler, Chesapeake Bay Foundation et al. v. USEPA. Settlement agreement. *United States District Court for the District of Columbia*. Retrieved from http://www.cbf.org/Document.Doc?id=512.

Gaffield, S. J., Goo, R. L., Richards, L. A., and Jackson, R. J. (2003). Public Health Effects of Inadequately Managed Stormwater Runoff. *American Journal of Public Health, 93*(9), 1527–1533. doi:10.2105/AJPH.93.9.1527.

Howard, K., and Schirmer, A. (2013). *MS4 Program Plan for the Town of Blacksburg Small MS4 General Permit (VPDES Permit No . VAR 040019)*. Blacksburg, Virginia.

Irvine, K., Rossi, M. C., Vermette, S., Bakert, J., and Kleinfelder, K. (2011). Illicit discharge detection and elimination: Low cost options for source identification and trackdown in stormwater systems. *Urban Water Journal*, *8*(6), 379–395. doi:10.1080/1573062X.2011.630095.

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., and Xian, G. (2013). A comprehensive change detection method for updating the National Land Cover Database to circa 2011. *Remote Sensing of Environment*, *132*, 159–175. doi:10.1016/j.rse.2013.01.012.

Johnson, B., and Tuomari, D. (1998). From theory to implementation - Finding illicit connections. In *Water Environment Federation 1998* (pp. 1–10). Denver, Colorado: Water Environment Federation.

Kaushal, S. S., Groffman, P. M., Band, L. E., Elliott, E. M., Shields, C. A., and Kendall, C. (2011). Tracking nonpoint source nitrogen pollution in human-impacted watersheds. *Environmental Science and Technology*, *45*(19), 8225–8232. doi:10.1021/es200779e.

Lilly, L. (2015). Illicit Discharge Detection and Elimination (IDDE) Program Assessment (Responses). Howard County, MD: Lori A Lilly Environmental Solutions.

Lilly, L., Hoffmann, G., Seipp, B., Drescher, S., Lane, C., Caraco, D., Swann, C., and Sturm, P. (2011). *Field Findings: Pollution Detection and Elimination in Sligo Creek, Montgomery County, MD*. Ellicott City, MD.

Lilly, L., Stack, B. P., and Caraco, D. (2012). Pollution loading from illicit sewage discharges in two Mid-Atlantic subwatersheds and implications for nutrient and bacterial total maximum daily loads. *Watershed Science Bulletin*, *3*(1), 7–17.

Lilly, L., and Sturm, P. (2010). *Technical memorandum: Illicit discharge monitoring in Baltimore watersheds*. Ellicott City, MD.

Mehta, C. R., and Patel, N. R. (1986). Algorithm 643: FEXACT: a FORTRAN subroutine for Fisher's exact test on unordered r×c contingency tables. *ACM Transactions on Mathematical Software*, *12*(2), 154–161. doi:10.1145/6497.214326.

Perry, W., Broers, A., El-Baz, F., Harris, W., Healy, B., and Hillis, W. D. (2008). *Grand challenges for engineering*. *National Academy of Engineering* (Vol. 37). Washington, D.C. doi:10.1109/EMR.2009.4804341.

Peterson, H., and Grout, W. R. (1992). Dry weather field screening as an indicator for urban drainage system rehabilitation. In *Water resources planning and management: saving a threatened resource, in search of solutions: proceedings of the water resources sessions at Water Forum '92* (pp. 516–522). American Society of Civil Engineers. doi:10.1017/S000748530002229X.

Pitt, R., and Rittenhouse, B. (2001). *Methods for detection of inappropriate discharges to storm drainage systems - Background literature and summary of findings*. Washington, D.C.

R Core Development Team. (2015). *R Language for Statistical Computing version 3.2.3*. Vienna, Austria: R Core Development Team.

Routledge, R. (2005). Fisher's Exact Test. In *Encyclopedia of Biostatistics*. Wiley.

Smith, D. J., and Shilley, H. (2009). *Residential car washwater monitoring study*. Federal Way, WA.

Spirandelli, D. (2015). Patterns of Wastewater Infrastructure along a Gradient of Coastal Urbanization: A Study of the Puget Sound Region. *Land, 4*(4), 1090–1109. doi:10.3390/land4041090.

Swann, C. (2001). The Influence of Septic Systems at the Watershed Level. *Watershed Protection Techniques, 3*(4), 821–834.

The Louis Berger Group Inc., and Virginia Department of Environmental Quality. (2015). *Roanoke River Bacteria and Sediment TMDL Implementation Plan Part I*.

Tuomari, D. C., and Thompson, S. (2004). Sherlocks of stormwater - Effective investigation techniques for illicit connection and discharge detection. In *Proceedings of the Water Environment Federation* (Vol. 2004, pp. 489–496). Wayne, MI: Water Environment Federation. doi:10.2175/193864704784147098.

Tuomari, D., Foley, J. P., and Taylor, E. S. (1995). Rouge River watershed illicit sewer connection detection program: A GIS application. In *ESRI Conference* (pp. 2–14).

US Census Bureau. (2015). State and County QuickFacts. Retrieved April 26, 2016, from http://www.census.gov/quickfacts/table/PST045215/00.

US Census Bureau. (2016). American Fact Finder. Retrieved December 17, 2015, from http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml.

USDA - NRCS. (2016). Geospatial Data Gateway. Retrieved December 17, 2015, from https://gdg.sc.egov.usda.gov/.

USEPA. National Pollutant Discharge Elimination System - Regulations for revision of the water pollution control program addressing stormwater discharges (1999). USA.

USEPA. PART 122 — EPA ADMINISTERED PERMIT PROGRAMS : THE NATIONAL POLLUTANT DISCHARGE ELIMINATION SYSTEM (2010). United States of America.

USEPA. (2016a). About NPDES. *National Pollutant Discharge Elimination System (NPDES)*. Retrieved February 18, 2016, from http://www.epa.gov/npdes/about-npdes.

USEPA. National Pollutant Discharge Elimination System (NPDES) Municipal Separate Storm Sewer System General Permit Remand (2016). United States of America: United States Government.

Versar Inc. (2014a). *Fairfax County Wet Weather Screening Program Plan 2014*. Fairfax County, VA.

Versar Inc. (2014b). *Arlington County Dry Weather Screening Program: Site Selection and Screening Plan*. Arlington County, VA.

Versar Inc. (2014c). *Arlington County Wet Weather and High Risk Screening Program: Site Selection and Screening Plan*. Arlington County, VA.

Virginia DEQ. (2013a). VA0088579 - Arlington County - Authorization to Discharge Under the Virginia Stormwater Management Program and the Virginia Stormwater Management Act. Arlington County, VA: VADEQ.

Virginia DEQ. (2013b). General Permit for Discharges of Stormwater from Small Municipal Separate Storm Sewer Systems. Richmond.

Virginia DEQ. (2013c). General Permit No. VAR040004 - City of Roanoke. Roanoke, VA.

Virginia DEQ. (2014). Water Quality Standards. Retrieved September 15, 2014, from http://www.deq.virginia.gov/Programs/Water/WaterQualityInformationTMDLs/WaterQualityStandards.aspx.

Virginia DEQ. (2015a). VA0088587 - Fairfax County - Authorization to Discharge Under the Virginia Stormwater Management Program and the Virginia Stormwater Management Act. Fairfax County, VA: VADEQ.

Virginia DEQ. (2015b). Municipal Separate Storm Sewer System (MS4) Permits. Retrieved October 16, 2015, from http://www.deq.virginia.gov/Programs/Water/StormwaterManagement/VSMPPermits/MS4Permits.aspx.

Wang, L. Z., Lyons, J., and Kanehl, P. (2001). Impacts of urbanization on stream habitat and fish across multiple spatial scales. *Environmental Management*, *28*(2), 255–266. doi:10.1007/s002670010222.

Zielinski, J., and Brown, T. (2003). *Inappropriate discharge detection and elimination: What Phase I communities are doing to address the problem*. Ellicott City, MD.

APPENDIX A – RISK MAPPING PROCEDURES

Individual Risk Factors

1. Land Cover
    a. Spatial Analysis
        i. Begin with land cover data for COR and FFX in “nlcd2011_clip”
        ii. Reclassify (Spatial Analyst → Reclass → Reclassify Tool) this raster into the desired categories of land cover (input=nlcd2011_clip, reclass field=Land_Cover {see list below, NoData → 0}, output=nlcd_reclass2). There are 15 different classifications within the study areas which were grouped into eight (8) new categories. Try a simple reclassification scheme first (below), then see if it would make sense to split any categories to evaluate further.
            1. Blank_
            2. Developed, Open Space
            3. Developed, Low Intensity
            4. Developed, Medium Intensity
            5. Developed, High Intensity
            6. Barren Land
            7. Agricultural: Hay Pasture, Cultivated Crops
            8. Undeveloped: Open Water, Deciduous Forest, Evergreen Forest, Mixed Forest, Shrub/Scrub, Herbaceous, Woody Wetlands, Emergent Herbaceous Wetlands
        iii. Convert (Conversion Tools → From Raster → Raster to Polygon Tool) the reclassified raster into a polygon layer for selection purposes (Input=nlcd_reclass2, field=LAND_COVER, output=Land_Cover_Types_2, check radio box to simplify polygons).
        iv. Overlay COR_IDDE_Reports_2_prj onto “Land_Cover_Types_2”.
        v. Apply a Selection by Location (Target=COR_IDDE_Reports_2_prj, Source=Land_Cover_Types_2, Method=within).
        vi. Report the number of records captured and not captured by “Land_Cover_Types_2”, for both ID_Type and Pollutant_2, in the excel spreadsheet.
        vii. Repeat steps iv-vi with “FFX_IDID_Reports_prj”
    b. Sensitivity Analysis
        i. Create an attribute table using the Build Raster Attribute Table tool (Data Management → Raster → Raster Properties → Build Raster Attribute Table) for the nlcd_reclass2 integer raster layer (input raster = nlcd_reclass2, uncheck the overwrite radio button).
        ii. Create a geodatabase table (name=land_cover_risks, alias=Land Cover Risks, accept default configuration settings) with the following fields:
            1. “Grid_Key” (short integer)
            2. “value” (short integer)
            3. “Land_Cover” (string)
            4. “Risk_Index” (double)

      iii. Fill in the table with the data for each Land Cover value and the corresponding decimal index value to be assigned.
         1. Developed, Medium Intensity (3 → 3.0)
         2. Developed, High Intensity (4 → 2.5)
         3. Developed, Low Intensity (2 → 2.0)
         4. Developed, Open Space (1 → 1.5)
         5. Undeveloped (Forest, wetland, etc.) (7 → 1.0)
         6. Agricultural (6 → 1.0)
         7. Barren (5 → 1.0)
         8. Blank (0 → 0.0)
      iv. Join land_cover_risks table to the nlcd_reclass2 attribute table (Attribute Table Join Field = OBJECTID, table = land_cover_risks, table join field = Grid_Key, check radio button to keep all records).
      v. Use the Lookup (Reclass) tool (Spatial Analyst Tools → Reclass → Lookup) to create a new raster from the integer raster with the decimal risk index values (input raster = nlcd_reclass2, Lookup field = land_cover_risks.Risk_Index, output raster = nlcd_fac).
      vi. Clip Raster (Data Management → Raster → Raster Processing → Clip) to the geometric extents of the Jurisdictional Boundaries (input raster=nlcd_fac, output extent=Study_Area_Jurisdictions, check radio button to use input features, output=nlcd_fac_c, NoData value=null).

2. Land Use
   a. Spatial Analysis
      i. Begin with "COR_zoning_prj" and "Fairfax_Ex_Land_Use_Generalized_prj" as they are the best representations available for true land use in each jurisdiction.
      ii. Add a string field (50 characters) to "Fairfax_Ex_Land_Use_Generalized_prj" called Data_Source to indicate the layer the data originated from.
      iii. Field calculate Data_Source = "Fairfax_Ex_Land_Use_Generalized" for current features
      iv. Append (Data Management → General → Append) "COR_zoning_prj" to "Fairfax_Ex_Land_Use_Generalized_prj" (input= COR_zoning_prj, target= Fairfax_Ex_Land_Use_Generalized_prj, schema type= NO_TEST, map field COR_BASE_ZONING to CATEG)
      v. Rename Layer "COR_FFX_Land_Use_Zoning"
      vi. Field calculate Data Source = "COR_zoning" for appended features
      vii. Add a field "Land_Use_Zoning" (string, 50) for general land use/zoning category
      viii. Field calculate "Land_Use_Zoning" based on original data codes to group polygons into one of the following categories of dominant land uses (dominant is defined as the most common land use, or in the case of mixed development zoning, the land use with the highest perceived risk):
         1. Agricultural (Agricultural, RA)
         2. Commercial (CG, CLS, CN, Commercial, D, MXPUD)
         3. Industrial (AD, I-1, I-2, Industrial light and heavy, IPUD, UF, Utilities)
         4. Institutional (IN, INPUD, Institutional)
         5. Recreational and Open Space (Open land not forested or developed, Public, Recreation, ROS)

        6. Residential (High-density Residential, Low-density Residential, Medium-density Residential, MX, R-12, R-3, R-5, R-7, RM-1, RM-2, RMF)
    ix. Overlay COR_IDDE_Reports_2_prj onto “COR_FFX_Land_Use_Zoning”.
    x. Apply a Selection by Location (Target=COR_IDDE_Reports_2_prj, Source=COR_FFX_Land_Use_Zoning, Method=within).
    xi. Report the number of records captured and not captured by “COR_FFX_Land_Use_Zoning”, for both ID_Type and Pollutant_2, in the excel spreadsheet.
    xii. Repeat steps ix-xi with “FFX_IDID_Reports_prj”
b. Sensitivity Analysis
    i. Begin with the Land_Use_Zoning vector layer (exported all data from COR_FFX_Land_Use_Zoning) and add a field “Land_Use” (short integer) to number each land use.
    ii. Field calculate “Land_Use” based on the following numberings:
        1. Commercial → 1
        2. Residential → 2
        3. Industrial → 3
        4. Recreation/Open Space → 4
        5. Institutional → 5
        6. Agricultural → 6
    iii. Convert to raster (Conversion Tools → To Raster → Polygon to Raster) based on “Land_Use” field (input=Land_Use_Zoning, valuefield=Land_Use, output=landuse, cell-assign-type=cell_center, priorityfield=none, cell size=30).
    iv. Reclassify (Spatial Analyst → Reclass → Reclassify Tool) to ensure NoData values are mapped to zero (input=landuse, reclass field=value {keep original values for 1, 2, 3, NoData → 0}, output=landuse_r).
    v. Create a geodatabase table (name=land_use_risks, alias=Land Use Risks, accept default configuration settings) with the following fields:
        1. “Grid_Key” (short integer)
        2. “value” (short integer)
        3. “Land_Use” (string)
        4. “Risk_Index” (double)
    vi. Fill in the table with the data for each Land Use value and the corresponding decimal index value I wish to assign.
        1. Commercial (1 → 3.0)
        2. Residential (2 → 3.0)
        3. Industrial (3 → 2.5)
        4. Recreation/Open Space (4 → 2.0)
        5. Institutional (5 → 1.5)
        6. Agricultural (6 → 1.0)
        7. NoData (0 → 0)
    vii. Join land_use_risks table to the landuse_r attribute table (Attribute Table Join Field = OBJECTID, table = land_use_risks, table join field = Grid_Key, check radio button to keep all records).

        viii. Use the Lookup (Reclass) tool (Spatial Analyst Tools → Reclass → Lookup) to create a new raster from the integer raster with the decimal risk index values (input raster = landuse_r, Lookup field = land_use_risks.Risk_Index, output raster = landuse_fac).
        ix. Clip Raster (Data Management → Raster → Raster Processing → Clip) to the geometric extents of the Jurisdictional Boundaries (input raster=landuse_fac, output extent=Study_Area_Jurisdictions, check radio button to use input features, output=landuse_fac_c, NoData value=null).

3. Population Density
    a. Spatial Analysis
        i. Begin with the downloaded csv file with population totals by block group, and the imported, projected, and appended census block groups for the study areas (COR_FFX_2010census_BlockGroups_prj)
        ii. Pre-process the csv file to prepare for join. Separate the GEO.id field to obtain the GEOID10 number for joining to the block groups, then rename this field GEO_ID. Rename the population column "Population_Total" and remove the second row of headers. Also remove extraneous data from the cell separation operation. Save the file as a .xlsx file.
        iii. Add a field to the "COR_FFX_2010census_BlockGroups_prj" layer called "Pop_Total" of type double.
        iv. Join the population excel table to the "COR_FFX_2010census_BlockGroups_prj" by linking the GEO_ID and GEOID10 fields, respectively.
        v. Field calculate the "Pop_Total" field = "Population_Total" and remove the join.
        vi. Add a field to "COR_FFX_2010census_BlockGroups_prj" layer called "Pop_Density" of type double.
        vii. Field calculate "Pop_Density" = ("Pop_Total" * 1000000) / "Shape_Area" to obtain the average population density (people/km^2) by block group.
        viii. Export the data by jurisdiction for continued analysis ("FFX_BlockGroups_PopDensity" and "COR_BlockGroups_PopDensity").
        ix. Symbolize block groups by "Pop_Density" (Natural Breaks (Jenks) w/ 3 classes, then round to nearest 10 people/sqmi). The thresholds for moderate and high risk for the City of Roanoke were 850 and 1650, and for Fairfax they were 3290 and 9390.
        x. Overlay COR_IDDE_Reports_2_prj onto "COR_BlockGroups_PopDensity".
        xi. Apply a Selection by Location (Target=COR_IDDE_Reports_2_prj, Source=COR_BlockGroups_PopDensity, Method=within).
        xii. Report the number of records captured and not captured by "COR_BlockGroups_PopDensity", for both ID_Type and Pollutant_2, in the excel spreadsheet.
        xiii. Repeat steps x-xii with "FFX_IDID_Reports_prj" and "FFX_BlockGroups_PopDensity"
    b. Sensitivity Analysis
        i. Begin with FFX_BlockGroups_PopDensity and add a field of type Double called "Risk_Index."
        ii. Field calculate "Risk_Index" based on population bins for FFX data (range 1-3).

        iii. Convert to raster (Conversion Tools → To Raster → Polygon to Raster) based on "Risk_Index" field (input=FFX_BlockGroups_PopDensity, valuefield=Risk_Index, output=ffx_pop, cell-assign-type=cell_center, priorityfield=none, cell size=30).
        iv. Reclassify (Spatial Analyst → Reclass → Reclassify Tool) to ensure NoData values are mapped to zero (input=ffx_pop, reclass field=value {keep original values for 1, 2, 3, NoData → 0}, output=ffx_pop_r).
        v. Copy Raster Tool (Data Management → Raster → Raster Dataset → Copy Raster) – convert risk-index raster to pixel type = 32_Bit_Float. This format enables raster calculator operations (input=ffx_pop_r, output=ffx_pop_fac, configuration keyword=<null>, ignore background value=<null>, NoData value=<null>, Pixel Type=32_BIT_FLOAT).
        vi. Clip Raster (Data Management → Raster → Raster Processing → Clip) to the geometric extents of the Jurisdictional Boundaries (input raster=ffx_pop_fac, output extent=FFX_County_Border_prj, check radio button to use input features, output=ffx_pop_fac_c, NoData value=null).
        vii. Repeat steps i-vi for COR_BlockGroups_PopDensity

4. Percent Impervious
    a. Spatial Analysis
        i. Begin with NLCD impervious cover data for COR and FFX in "nImp2011_clip"
        ii. Perform a raster calculation (Spatial Analyst Tools → Map Algebra → Raster Calculator) to obtain the area in each cell that is impervious from the given percentage (Expression = nImp2011_clip * 9, Output Raster = imperv_area).
        iii. Calculate the zonal statistics (Spatial Analyst Tools → Zonal → Zonal Statistics as Table) for the total impervious area within each watershed (input zone data = Study_Area_Watersheds_prj, zone field = NAME, input value raster = imperv_area, output table = watershed_impervious, check box to Ignore NoData, statistics type = ALL).
        iv. Create two new fields in Study_Area_Watersheds_prj of type double called "Impervious_Area" and "Impervious_Pct"
        v. Join the watershed_impervious table to Study_Area_Watersheds_prj and calculate "Impervious_Area" = SUM. Remove Join.
        vi. Calculate "Impervious_Pct" = "Impervious_Area" / "Shape_Area"
        vii. Symbolize watersheds by "Impervious_Pct" using Jenks Natural Breaks, 3 classes
        viii. Export each jurisdiction's watersheds with impervious data to new feature classes in the "Processed_Data_MapReady" feature dataset (COR_WS_Impervious and FFX_WS_Impervious).
        ix. Symbolize by Jenks Natural Breaks with 3 classes, round to nice numbers. Thresholds for moderate and high risk in Roanoke were 0.04 and 0.45, while in Fairfax they were 0.05 and 0.20.
        x. Overlay "FFX_IDID_Reports_prj" onto "FFX_WS_Impervious"
        xi. Apply a Selection by Location (Target= FFX_IDID_Reports_prj, Source= FFX_WS_Impervious, Method=within).

        xii. Report the number of records captured and not captured by “FFX_WS_Impervious”, for both ID_Type and Pollutant_2, in the excel spreadsheet.
        xiii. Repeat steps viii - xii with COR data
    b. Sensitivity Analysis
        i. Begin with FFX_WS_Impervious and add a field of type Double called “Risk_Index.”
        ii. Field calculate “Risk_Index” based on imperviousness bins for FFX (range 1-3).
        iii. Convert to raster (Conversion Tools → To Raster → Polygon to Raster) based on “Risk_Index” field (input= FFX_WS_Impervious, valuefield=Risk_Index, output=ffx_imp, cell-assign-type=cell_center, priorityfield=none, cell size=30).
        iv. Reclassify (Spatial Analyst → Reclass → Reclassify Tool) this raster to ensure NoData values are mapped to zero (input=ffx_imp, reclass field=value {keep original values for 1, 2, 3, NoData → 0}, output=ffx_imp_r).
        v. Copy Raster Tool (Data Management → Raster → Raster Dataset → Copy Raster) – convert risk-index raster to pixel type = 32_Bit_Float. This format enables raster calculator operations (input=ffx_imp_r, output=ffx_imp_fac, configuration keyword=<null>, ignore background value=<null>, NoData value=<null>, Pixel Type=32_BIT_FLOAT).
        vi. Clip Raster (Data Management → Raster → Raster Processing → Clip) to the geometric extents of the Jurisdictional Boundaries (input raster=ffx_imp_fac, output extent=FFX_County_Border_prj, check radio button to use input features, output=ffx_imp_fac_c, NoData value=null).
        vii. Repeat steps i-vi for COR_WS_Impervious

5. Development Age
    a. Spatial Analysis
        i. Begin with COR data: citywide12012015.shp (imported), RealImprov.xlsx
        ii. Add new fields to COR_citywide_12012015 layer: “BuildOut_Year” (double) and Dev_Age (double)
        iii. Create “RealImprov” pivot table to produce a table by parcel with attributes LRSN, Oldest Structure Date, Newest Structure Date. I copied these values into a new sheet with blank formatting (Year_Built.xlsx): LRSN, Old_YrBuilt, New_YrBuilt, Count_YrBuilt
        iv. Join new build out table to parcels (LRSN) and calculate BuildOut_Year = New_YrBuilt; remove join
        v. Calculate “Dev_Age” = 2016 - “BuildOut_Year”
        vi. Perform a Spatial Join (Analysis Tools → Overlay → Spatial Join) of the parcels to the watersheds to find the average of all developed parcels in each watershed (target features = COR_Watersheds_Clipped, join features = COR_citywide_12012015, output feature class = COR_WS_Dev_Age, join operation = JOIN_ONE_TO_ONE, check radio box to keep all target features, map fields {accept all watershed fields, delete all parcel fields except Dev_Age, map Dev_Age parcel field with rule = MEDIAN}, match option = INTERSECT, search radius = null).
        vii. Symbolize watersheds as desired using “Dev_Age” field

    viii. Return to start with Fairfax data: Fairfax_Parcels layer and Tax_Administrations_Real_Estate__Dwelling_Data.csv
    ix. Create a pivot table for the Tax_Administrations_Real_Estate__Dwelling_Data.csv and determine the following for each parcel: PIN, Old_YRBLT, New_YRBLT, Count_YRBLT, Old_YRREMOD, New_YRREMOD, Count_YRREMOD. Copy values into a new spreadsheet with blank formatting (FFX_BuildOutYears.xlsx).
    x. Add new fields to Fairfax_Parcels of type double called BuildOut_Year and Dev_Age
    xi. Separate Fairfax_Parcels into four layers with about 22500 features a piece to perform joins and calculates.
    xii. Export FFX_BuildOutYears.xlsx to a gdb table (Fairfax_Parcels_BuildOut_Year).
    xiii. Join Fairfax_Parcels_BuildOut_Year to Fairfax_Parcels_Subset# via Parcel_PIN field
    xiv. Calculate BuildOut_Year field = New_YRBLT and remove the join
    xv. Calculate “Dev_Age” = 2016 - “BuildOut_Year”
    xvi. Append all 16 parcel layers back into one feature class (Fairfax_Parcels_DevAge)
    xvii. Perform a Spatial Join (Analysis Tools → Overlay → Spatial Join) of the parcels to the watersheds to find the average of all developed parcels in each watershed (target features = Fairfax_Watersheds_Clipped, join features = Fairfax_Parcels_DevAge, output feature class = FFX_WS_Dev_Age, join operation = JOIN_ONE_TO_ONE, check radio box to keep all target features, map fields {accept all watershed fields, delete all parcel fields except Dev_Age, map Dev_Age parcel field with rule = MEDIAN}, match option = INTERSECT, search radius = null).
    xviii. Project (Data Management → Projections and Transformations → Project) the median data to Processed_Data_MapReady dataset (input=FFX_WS_Dev_Age_Median, output= FFX_WS_Dev_Age_Median_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).
    xix. Symbolize watersheds as desired using “Dev_Age” and compare to COR data. Decided to symbolize by Jenks Natural Breaks 3 classes, no rounding needed. Thresholds for moderate and high risk in Roanoke were 52 and 66, while in Fairfax they were 30 and 41.
    xx. Overlay “FFX_IDID_Reports_prj” onto “FFX_WS_Dev_Age_Median_prj”
    xxi. Apply a Selection by Location (Target= FFX_IDID_Reports_prj, Source= FFX_WS_Dev_Age_Median_prj, Method=within).
    xxii. Report the number of records captured and not captured by “FFX_WS_Dev_Age_Median_prj”, for both ID_Type and Pollutant_2, in the excel spreadsheet.
    xxiii. Repeat steps xx - xxiv with COR data
b. Sensitivity Analysis
    i. Begin with FFX_WS_Dev_Age_Median_prj and add a field of type Double called “Risk_Index.”

        ii. Field calculate “Risk_Index” based on development age bins for FFX (range 1-3).
        iii. Convert to raster (Conversion Tools → To Raster → Polygon to Raster) based on “Risk_Index” field (input= FFX_WS_Dev_Age_Median_prj, valuefield=Risk_Index, output=ffx_dev, cell-assign-type=cell_center, priorityfield=none, cell size=30).
        iv. Reclassify (Spatial Analyst → Reclass → Reclassify Tool) this raster to ensure NoData values are mapped to zero (input=ffx_dev, reclass field=value {keep original values for 1, 2, 3, NoData → 0}, output=ffx_dev_r).
        v. Copy Raster Tool (Data Management → Raster → Raster Dataset → Copy Raster) – convert risk-index raster to pixel type = 32_Bit_Float. This format enables raster calculator operations (input=ffx_dev_r, output=ffx_dev_fac, configuration keyword=<null>, ignore background value=<null>, NoData value=<null>, Pixel Type=32_BIT_FLOAT).
        vi. Clip Raster (Data Management → Raster → Raster Processing → Clip) to the geometric extents of the Jurisdictional Boundaries (input raster=ffx_dev_fac, output extent=FFX_County_Border_prj, check radio button to use input features, output=ffx_dev_fac_c, NoData value=null).
        vii. Repeat steps i-vi for COR_WS_Dev_Age_Median_prj

6. Outfall Density
    a. Spatial Analysis
        i. Export and clip all watersheds to jurisdictions
        ii. Export, clip, and dissolve streams within each watershed in each jurisdiction
        iii. Export outfalls for each jurisdiction
        iv. Create a buffer around the dissolved stream reaches (input=FFX_NHD_Streamlines_clip_dissolve, output=FFX_NHD_Streamlines_clip_dissolve_buffer, linear unit=50 ft, side type=full, end type=round, method=planar, dissolve type=none).
        v. Perform a Spatial Join (Analysis Tools → Overlay → Spatial Join) between the outfalls and the stream buffer (target=FFX_NHD_Streamlines_clip_dissolve_buffer, join=FFX_Outfalls, output=FFX_Outfall_Density, Join Operation=JOIN_ONE_TO_ONE, Keep all targets, field map defaults, Match Option=contains, search radius=<null>).
        vi. Add a field “Outfall_Density” of type “double” to the spatial join layer.
        vii. Field calculate “Outfall_Density” using the following equation, in units of outfalls per stream mile: Outfall_Density = (Join_Count * 5280 ft/mi * 100 ft) / (Shape_Area (ft^2))
        viii. Add a field “Outfall_Density” of type “double” to “Fairfax_Watersheds_Clipped”
        ix. Join “FFX_Outfall_Density” to “Fairfax_Watersheds_Clipped” and copy the density value into the watersheds feature class. Remove join.
        x. Project (Data Management → Projections and Transformations → Project) data to Processed_Data_MapReady dataset (input= Fairfax_Watersheds_Clipped, output= Fairfax_Watersheds_Clipped_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).

        - xi. Add a field of type double called "Outfall_Density_Metric" and recalculate the density from imperial to metric: Metric = Outfall Density / 1.6 kilometers/mile.
        - xii. Symbolize with Jenks Natural Breaks 3 Classes. Thresholds for moderate and high risk in Roanoke were 1.9 and 6.8, while in Fairfax they were 1.7 and 4.5.
        - xiii. Overlay FFX_IDID_Reports_prj onto "Fairfax_Watersheds_Clipped_prj".
        - xiv. Apply a Selection by Location (Target= FFX_IDID_Reports_prj, Source= Fairfax_Watersheds_Clipped_prj, Method=within).
        - xv. Report the number of records captured and not captured by "Fairfax_Watersheds_Clipped_prj", for both ID_Type and Pollutant_2, in the excel spreadsheet.
        - xvi. Repeat steps i-xiii with COR data
    - b. Sensitivity Analysis
        - i. Begin with Fairfax_Watersheds_Clipped_prj and add a field of type Double called "Risk_Index."
        - ii. Field calculate "Risk_Index" based on outfall density bins for FFX (range 1-3).
        - iii. Convert to raster (Conversion Tools → To Raster → Polygon to Raster) based on "Risk_Index" field (input=Fairfax_Watersheds_Clipped_prj, valuefield=Risk_Index, output=ffx_ofd, cell-assign-type=cell_center, priorityfield=none, cell size=30).
        - iv. Reclassify (Spatial Analyst → Reclass → Reclassify Tool) this raster to ensure NoData values are mapped to zero (input=ffx_ofd, reclass field=value {keep original values for 1, 2, 3, NoData → 0}, output=ffx_ofd_r).
        - v. Copy Raster Tool (Data Management → Raster → Raster Dataset → Copy Raster) – convert risk-index raster to pixel type = 32_Bit_Float. This format enables raster calculator operations (input=ffx_ofd_r, output=ffx_ofd_fac, configuration keyword=<null>, ignore background value=<null>, NoData value=<null>, Pixel Type=32_BIT_FLOAT).
        - vi. Clip Raster (Data Management → Raster → Raster Processing → Clip) to the geometric extents of the Jurisdictional Boundaries (input raster=ffx_ofd_fac, output extent=FFX_County_Border_prj, check radio button to use input features, output=ffx_ofd_fac_c, NoData value=null).
        - vii. Repeat steps i-vi for Roanoke_Watersheds_Clipped_prj

7. Aging Sanitary Infrastructure
    - a. Spatial Analysis
        - i. Begin with sewer line work for City of Roanoke (WVWA_sewerline_all)
        - ii. Add fields for Sewer_Date (date), Age_Years (double), Aging_Inf (short integer)
        - iii. Field Calculate Sewer_Date based on the Install Date field (consult maintenance and year renewed fields where available, use most recent date).
        - iv. Use the following statement to calculate the Age_Years field:

          DateDiff("d", [Sewer_Date], Now()) / 365
        - v. Query for all sewer lines satisfying the following criteria and field calculate Aging_Inf field based on relative risk level (1, 2, or 3):
            1. Age_Years <=15, OR (Material = {NULL, Brick, Clay Tile, Fusible PVC, HDPE, Orange burg, PE, PP, PVC, PVC[non-domain]x4, Reinforced Plastic Pipe, Vitrified Clay Pipe, Unknown, Other} AND Age_Years <= 25)

    2. (Material = {NULL, Brick, Clay Tile, Fusible PVC, HDPE, Orange burg, PE, PP, PVC, PVC[non-domain]x4, Reinforced Plastic Pipe, Vitrified Clay Pipe, Unknown, Other} AND 25 < Age_Years <= 50), OR (Material = {Asbestos Cement, 'C', Cast Iron, 'CMP', Concrete Pipe, Concrete Segments Bolted, 'CSU', Ductile Iron Pipe, Fiberglass Reinforced Pipe, 'PCCP', Reinforced Concrete Pipe, 'SB', Steel Pipe, 'TTE'} AND 15 < Age_Years <= 30)
    3. Any pipe with Age_Years > 50 or NULL, or (Material = {Asbestos Cement, 'C', Cast Iron, 'CMP', Concrete Pipe, Concrete Segments Bolted, 'CSU', Ductile Iron Pipe, Fiberglass Reinforced Pipe, 'PCCP', Reinforced Concrete Pipe, 'SB', Steel Pipe, 'TTE'} AND Age_Years > 30)

vi. Intersect (Analysis Tools → Overlay → Intersect) pipes with watersheds to ensure that all pipes are broken along watershed boundaries (input features=WVWA_sewerline_all, COR_Watersheds_Clipped, output=WVWA_sewerline_intersect, JoinAttributes=ALL, XY Tolerance=NULL, Output Type=INPUT).
vii. Add a field to the intersect feature class called "DistXScore" (double)
viii. Field calculate "DistXScore" = "Shape_Length" * "Aging_Inf"
ix. Use the Spatial Join tool (Analysis Tools → Overlay → Spatial Join) to join the sanitary sewer linework to each watershed (target=COR_Watersheds_Clipped, join features=WVWA_sewerline_intersect, output=COR_WS_Aging_Sanitary_Sewer, Join Operation=JOIN_ONE_TO_ONE, check radio button to keep all target features, field map {keep all COR WS fields, sum on Shape_Length_1 and DistXScore fields, delete rest of intersect fields}, Match Option=Contains, Search Radius=null).
x. Add field to COR_WS_Aging_Sanitary_Sewer called "Aging_San_Score" (double).
xi. Field Calculate "Aging_San_Score" = "DistXScore" / "Shape_Length_1"
xii. Project (Data Management → Projections and Transformations → Project) the data to Processed_Data_MapReady dataset (input=COR_WS_Aging_Sanitary_Sewer, output= COR_WS_Aging_Sanitary_Sewer_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).
xiii. Symbolize by Jenks Natural Breaks with 3 classes and compare to Fairfax data.
xiv. Return to start with Fairfax: Fairfax_WW_REHABS, Fairfax_WW_SEWERLINE
xv. Add fields for Sewer_Year (double), Age_Years (double), Aging_Inf (short)
xvi. Join the REHABS layer to the SEWERLINES layer using the "Pipe_ID" field
xvii. Field Calculate Sewer_Year based on the YEARBUILT field (consult YEAR from REHABS where available, use most recent year).
xviii. Field Calculate "Age_Years" = 2016 - "Sewer_Year"
xix. Query for all sewer lines satisfying the following criteria and field calculate Aging_Inf field based on relative risk level (1, 2, or 3):

        1. Age_Years <=15, OR (Material = {CLAY, INSITUFORM, LINER, N/A, SLIPLINE, TRANSITE, PVC, HDPE, NULL, Blanks} AND Age_Years <= 25)
        2. (Material = {CLAY, INSITUFORM, LINER, N/A, SLIPLINE, TRANSITE, PVC, HDPE, NULL, Blanks} AND 25 < Age_Years <= 50), OR (Material = {FIBERGLASS, PRE-STRESS, AC, CIP, RCP, CONC, DIP, STEEL, GALVANIZED} AND 15 < Age_Years <= 30)
        3. Any pipe with Age_Years > 50 or NULL, or (Material = {FIBERGLASS, PRE-STRESS, AC, CIP, RCP, CONC, DIP, STEEL, GALVANIZED} AND Age_Years > 30)
    - xx. Intersect (Analysis Tools → Overlay → Intersect) pipes with watersheds to ensure that all pipes are broken along watershed boundaries (input features=Fairfax_WW_SEWERLINE, Fairfax_Watersheds_Clipped, output=Fairfax_WW_SEWERLINE_intersect, JoinAttributes=ALL, XY Tolerance=NULL, Output Type=INPUT).
    - xxi. Add a field to the intersect feature class called “DistXScore” (double)
    - xxii. Field calculate “DistXScore” = “SHAPE_Length” * “Aging_Inf”
    - xxiii. Use the Spatial Join tool (Analysis Tools → Overlay → Spatial Join) to join the sanitary sewer linework to each watershed (target=Fairfax_Watersheds_Clipped, join features=Fairfax_WW_SEWERLINE_intersect, output=Fairfax_WS_Aging_Sanitary_Sewer, Join Operation=JOIN_ONE_TO_ONE, check radio button to keep all target features, field map {keep all Fairfax WS fields, sum on SHAPE_Length_1 and DistXScore fields, delete rest of intersect fields}, Match Option=Contains, Search Radius=null).
    - xxiv. Add field to Fairfax_WS_Aging_Sanitary_Sewer called “Aging_San_Score” (double).
    - xxv. Field Calculate “Aging_San_Score” = “DistXScore” / “SHAPE_Length_1”
    - xxvi. Project (Data Management → Projections and Transformations → Project) the data to Processed_Data_MapReady dataset (input=Fairfax_WS_Aging_Sanitary_Sewer, output= Fairfax_WS_Aging_Sanitary_Sewer_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).
    - xxvii. Symbolize by Jenks Natural Breaks with 3 classes and compare to Fairfax data. Thresholds used for moderate and high risk in Roanoke were 2.4 and 2.7, while Fairfax they were 1.8 and 2.4.
    - xxviii. Overlay “FFX_IDID_Reports_prj” onto “Fairfax_WS_Aging_Sanitary_Sewer_prj”
    - xxix. Apply a Selection by Location (Target= FFX_IDID_Reports_prj, Source= Fairfax_WS_Aging_Sanitary_Sewer_prj, Method=within).
    - xxx. Report the number of records captured and not captured by “Fairfax_WS_Aging_Sanitary_Sewer_prj”, for both ID_Type and Pollutant_2, in the excel spreadsheet.
    - xxxi. Repeat steps xxviii - xxx with COR data
- b. Sensitivity Analysis

        i. Begin with Fairfax_WS_Aging_Sanitary_Sewer_prj and add a field of type Double called “Risk_Index.”
        ii. Field calculate “Risk_Index” based on Aging_San_Score bins for FFX (range 1-3).
        iii. Convert to raster (Conversion Tools → To Raster → Polygon to Raster) based on “Risk_Index” field (input= Fairfax_WS_Aging_Sanitary_Sewer_prj, valuefield=Risk_Index, output=ffx_ass, cell-assign-type=cell_center, priorityfield=none, cell size=30).
        iv. Reclassify (Spatial Analyst → Reclass → Reclassify Tool) this raster to ensure NoData values are mapped to zero (input=ffx_ass, reclass field=value {keep original values for 1, 2, 3, NoData → 0}, output=ffx_ass_r).
        v. Copy Raster Tool (Data Management → Raster → Raster Dataset → Copy Raster) – convert risk-index raster to pixel type = 32_Bit_Float. This format enables raster calculator operations (input=ffx_ass_r, output=ffx_ass_fac, configuration keyword=<null>, ignore background value=<null>, NoData value=<null>, Pixel Type=32_BIT_FLOAT).
        vi. Clip Raster (Data Management → Raster → Raster Processing → Clip) to the geometric extents of the Jurisdictional Boundaries (input raster=ffx_ass_fac, output extent=FFX_County_Border_prj, check radio button to use input features, output=ffx_ass_fac_c, NoData value=null).
        vii. Repeat steps i-vi for COR_WS_Aging_Sanitary_Sewer_prj

8. Drainage Density
    a. Spatial Analysis
        i. Begin with Watersheds and drainage polylines for Fairfax County (Fairfax_Watersheds_Clipped and Fairfax_STORMNET_ARCS feature classes).
        ii. Use the Intersect Tool (Analysis Tools → Overlay → Intersect) to separate lines which cross watershed boundaries, and to clip features to the county’s jurisdiction of local watersheds (input features=Fairfax_STORMNET_ARCS, Fairfax_Watersheds_Clipped, output=Fairfax_ARCS_Watersheds_Intersect, JoinAttributes=ONLY_FID, XY Tolerance=NULL, Output Type=INPUT).
        iii. Use Spatial Join Tool (Analysis Tools → Overlay → Spatial Join) to find the sum of conveyance kilometers in each watershed (Target=Fairfax_Watersheds_Clipped, Join=Fairfax_ARCS_Watersheds_Intersect, Output=Fairfax_Watershed_Drainage_Density, Join Operation=JOIN_ONE_TO_ONE, Check radio box to keep all target features, Accept field map defaults, but change Merge Rule for SHAPE_Length_1 to sum the Intersect features’ shape lengths, Match Option=CONTAINS, Search Radius=NULL).
        iv. Add a field of type “Double” called “Drainage_Density”
        v. Field Calculate “Drainage_Density” = “SHAPE_Length_1” * 3281 / “Shape_Area” (results in units of 1/km or km/km^2)
        vi. Project (Data Management → Projections and Transformations → Project) the data to Processed_Data_MapReady dataset (input=Fairfax_Watershed_Drainage_Density, output=

Fairfax_Watershed_Drainage_Density_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).

vii. Symbolize using Jenks Natural Breaks with 3 classes and compare to COR data

viii. Return to start with City of Roanoke data (COR_MS4_conveyances, COR_Watersheds_Clipped).

ix. Use the Intersect Tool (Analysis Tools → Overlay → Intersect) to separate lines which cross watershed boundaries, and to clip features to the city's jurisdiction of local watersheds (input features=COR_MS4_conveyances, COR_Watersheds_Clipped, output=COR_conveyances_Watersheds_Intersect, JoinAttributes=ONLY_FID, XY Tolerance=NULL, Output Type=INPUT).

x. Use Spatial Join Tool (Analysis Tools → Overlay → Spatial Join) to find the sum of conveyance kilometers in each watershed (Target=COR_Watersheds_Clipped, Join= COR_conveyances_Watersheds_Intersect, Output=COR_Watershed_Drainage_Density, Join Operation=JOIN_ONE_TO_ONE, Check radio box to keep all target features, Accept field map defaults, but change Merge Rule for SHAPE_Length_1 to sum the Intersect features' shape lengths, Match Option=CONTAINS, Search Radius=NULL).

xi. Add a field of type "Double" called "Drainage_Density"

xii. Field Calculate "Drainage_Density" = "SHAPE_Length_1" * 3281 / "Shape_Area" (results in units of 1/km or km/km^2)

xiii. Project (Data Management → Projections and Transformations → Project) the data to Processed_Data_MapReady dataset (input=COR_Watershed_Drainage_Density, output= COR_Watershed_Drainage_Density_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).

xiv. Symbolize using Jenks Natural Breaks with 3 classes and compare to Fairfax data

xv. Breaks were similar enough to use a standard binning for both MS4s, natural breaks taken to nice numbers (Low = 0 - 2.2, Moderate = 2.2 – 6.5, High = > 6.5).

xvi. Overlay "FFX_IDID_Reports_prj" onto "Fairfax_Watershed_Drainage_Density_prj"

xvii. Apply a Selection by Location (Target= FFX_IDID_Reports_prj, Source= Fairfax_Watershed_Drainage_Density_prj, Method=within).

xviii. Report the number of records captured and not captured by "Fairfax_Watershed_Drainage_Density_prj", for both ID_Type and Pollutant_2, in the excel spreadsheet.

xix. Repeat steps xvi - xviii with COR data

b. Sensitivity Analysis

i. Begin with Fairfax_Watershed_Drainage_Density_prj and add a field of type Double called "Risk_Index."

ii. Field calculate "Risk_Index" based on drainage density bins for FFX (range 1-3).

iii. Convert to raster (Conversion Tools → To Raster → Polygon to Raster) based on "Risk_Index" field (input= Fairfax_Watershed_Drainage_Density_prj,

        valuefield=Risk_Index, output=ffx_wdd, cell-assign-type=cell_center, priorityfield=none, cell size=30).
        iv. Reclassify (Spatial Analyst → Reclass → Reclassify Tool) this raster to ensure NoData values are mapped to zero (input=ffx_wdd, reclass field=value {keep original values for 1, 2, 3, NoData → 0}, output=ffx_wdd_r).
        v. Copy Raster Tool (Data Management → Raster → Raster Dataset → Copy Raster) – convert risk-index raster to pixel type = 32_Bit_Float. This format enables raster calculator operations (input=ffx_wdd_r, output=ffx_wdd_fac, configuration keyword=<null>, ignore background value=<null>, NoData value=<null>, Pixel Type=32_BIT_FLOAT).
        vi. Clip Raster (Data Management → Raster → Raster Processing → Clip) to the geometric extents of the Jurisdictional Boundaries (input raster=ffx_wdd_fac, output extent=FFX_County_Border_prj, check radio button to use input features, output=ffx_wdd_fac_c, NoData value=null).
        vii. Repeat steps i-vi for COR_Watershed_Drainage_Density_prj

9. Generating Site Density
    a. Spatial Analysis
        i. Begin with COR data (GenSites_ECHO_pts layer, already geolocated)
        ii. Project (Data Management → Projections and Transformations → Project) data to Processed_Data_MapReady dataset (input= GenSites_ECHO_pts, output= COR_GenSites_ECHO_pts_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).
        iii. Use Kernel Density Tool (Spatial Analyst → Density → Kernel Density) to get raster layer with density of generating sites per square kilometer (input= COR_GenSites_ECHO_pts_prj, population field=none, output=cor_gensite_d, output cell size=30 m, search radius=<null>, area units=square kilometers, output values are= densities, method=planar).
        iv. Reclassify be Jenks Natural Breaks with three classes and compare to FFX data.
        v. Return to start with FFX data (EPA_ECHO_FairfaxCounty_12-18-15.xlsx)
        vi. Add EPA_ECHO_FairfaxCounty_12-18-15.xlsx to the map, right click on the table and select Display XY Data. Select "fac_long" as the X-Field, "fac_lat" as the Y-Field, and the GCS NAD 1983 coordinate system, and click OK.
        vii. A temporary layer will be created displaying the spatial location of those records with values for latitude and longitude. Verify the locations of these points using Google Maps and Fairfax_Address_Points layer.
        viii. Export the data into the Fairfax_County feature dataset (Fairfax_ECHO_test1).
        ix. Select the records in Fairfax_ECHO_test1 with null latitude and longitude, but with values for the address and city fields. Remove those which are from cities outside of Fairfax County's jurisdictional boundary. Export these features to another layer (Fairfax_ECHO_null_xy).
        x. Process the "fac_street" field to match the format of the "ADDRESS_1" field of Fairfax_Address_Points, then join the two tables based on those two fields.
        xi. Manually go in and add addresses to the joined addresses for points which did not join properly. Export the address points to a new layer (Fairfax_ECHO_Final_pts). Remove join from Fairfax_Address_Points layer.

        - xii. Select by location and by attributes from Fairfax_ECHO_test1 to retrieve points inside Fairfax County jurisdiction which do not have null coordinates.
        - xiii. Append these points to the Fairfax_ECHO_Final_pts layer (input=Fairfax_ECHO_test1, target=Fairfax_ECHO_Final_pts, schema type=NO TEST, map fields as appropriate).
        - xiv. Project (Data Management → Projections and Transformations → Project) data to Processed_Data_MapReady dataset (input=Fairfax_ECHO_Final_pts, output= Fairfax_ECHO_Final_pts_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).
        - xv. Use Kernel Density Tool (Spatial Analyst → Density → Kernel Density) to get raster layer with density of generating sites per square kilometer (input=Fairfax_ECHO_Final_pts_prj, population field=none, output raster=ffx_gensite_d, output cell size=30 m, search radius=<null>, area units=square kilometers, output values are= densities, method=planar).
        - xvi. Reclassify by Jenks Natural Breaks with three classes and compare to COR data.
        - xvii. Ranges are close enough that common binning can be implemented (COR actually has slightly higher densities). Use bins that also reflect CWP high risk density threshold at 10 sites per square mile (4 sites per square kilometer). Low risk = 0 – 2, Moderate risk = 2 – 4, High risk = > 4 sites per square km.
        - xviii. Reclassify raster layer (Spatial Analyst → Reclass → Reclassify Tool) to code it based on desired cutoffs (Input=ffx_gensite_d, reclass field=value {0 - 2 → 1, 2 - 4 → 2, and > 4 → 3, NoData → 0}, output=ffx_gensite_r).
        - xix. Convert (Conversion Tools → From Raster → Raster to Polygon Tool) the reclassified raster into a polygon layer for selection purposes (Input=ffx_gensite_r, field=value, output=Fairfax_Generating_Site_Density, check radio box to simplify polygons).
        - xx. Overlay “FFX_IDID_Reports_prj” onto “Fairfax_Generating_Site_Density”
        - xxi. Apply a Selection by Location (Target= FFX_IDID_Reports_prj, Source= Fairfax_Generating_Site_Density, Method=within).
        - xxii. Report the number of records captured and not captured by “Fairfax_Generating_Site_Density”, for both ID_Type and Pollutant_2, in the excel spreadsheet.
        - xxiii. Repeat steps xviii - xxii with COR data
    - b. Sensitivity Analysis
        - i. Copy Raster Tool (Data Management → Raster → Raster Dataset → Copy Raster) – convert risk-index raster to pixel type = 32_Bit_Float. This format enables raster calculator operations (input=ffx_gensite_r raster, output=ffx_gen_fac, configuration keyword=<null>, ignore background value=<null>, NoData value=<null>, Pixel Type=32_BIT_FLOAT).
        - ii. Clip Raster (Data Management → Raster → Raster Processing → Clip) to the geometric extents of the Fairfax County jurisdictional boundary (input raster=ffx_gen_fac, output extent=FFX_County_Border_Prj, check radio button to use input features, output=ffx_gen_fac_c, NoData value=null).
        - iii. Repeat with cor_gensite_r data

10. Infrastructure Access Density
    - a. Spatial Analysis

i. Begin with Fairfax County Data: Fairfax_STORMNET_NODES
ii. Query access points from Fairfax_STORMNET_NODES (select by attributes, TYPE_NAME = {Bioretention, Control Structure, Curb Inlet, Dry Pond, Grated Inlet, Infall, Parking Lot, Roof Top, Sand Filter, Trench, Wet Pond, Wetland, Yard Inlet, End of Pipe, End Wall, Inlet Pipe, Headwall, End Section}).
iii. Export access points to new layer (Export: Selected Features, output=Fairfax_Access_Pts).
iv. Project (Data Management → Projections and Transformations → Project) data to Processed_Data_MapReady dataset (input=Fairfax_Access_Pts, output=Fairfax_Access_Pts_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).
v. Use Kernel Density Tool (Spatial Analyst → Density → Kernel Density) to get raster layer with density of SW access points per square kilometer (input=Fairfax_Access_Pts_prj, population field=none, output=ffx_inf_acc, output cell size=30 m, search radius=blank, area units=squarekilometers, output values are=densities, method=planar).
vi. Symbolize the infrastructure access density using Jenks Natural Breaks (3 Classes) and round to nearest appropriate units. Compare densities with Roanoke's data.
vii. Return to start with Roanoke data: COR_MS4_nodes and COR_MS4_SCMs
viii. Feature to Point Tool (Data Management → Features → Feature to Point) to convert SCMs to points (input=COR_MS4_SCMs, output= COR_MS4_SCMs_Pts, uncheck "Inside Feature").
ix. Query access points from COR_MS4_nodes (select by attributes, Type = {CB, PE, POS}).
x. Make a copy of COR_MS4_SCMs_Pts in the gdb and rename it COR_Access_Pts
xi. Append (Data Management → General → Append) selected features from COR_MS4_nodes to COR_MS4_SCMs_Pts (input dataset=COR_MS4_nodes, Target dataset=COR_Access_Pts, Schema Type=NO_TEST, map fields as needed)
xii. Project (Data Management → Projections and Transformations → Project) data to Processed_Data_MapReady dataset (input=COR_Access_Pts, output=COR_Access_Pts_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).
xiii. Use Kernel Density Tool (Spatial Analyst → Density → Kernel Density) to get raster layer with density of SW access points per square kilometer (input=COR_Access_Pts_prj, population field=none, output=cor_inf_acc, output cell size=30 m, search radius=blank, area units=squarekilometers, output values are=densities, method=planar).
xiv. Symbolize by Jenks Natural Breaks (3 Classes) and compare with Fairfax data.
xv. Roanoke and Fairfax data on infrastructure access density are similar enough to warrant using the same intervals for classification:
    1. Low Risk = 0 – 75 access points per square kilometer
    2. Moderate Risk = 75 – 200
    3. High Risk = > 200

        xvi. Reclassify raster layer (Spatial Analyst → Reclass → Reclassify Tool) to code it based on desired risk levels (input=cor_inf_acc, reclass field=value {0 – 75 → 1, 75 – 200 → 2, > 700 → 3, NoData → 0}, output=cor_inf_acc_r).
        xvii. Convert (Conversion Tools → From Raster → Raster to Polygon Tool) the reclassified raster into a polygon layer for selection purposes (Input=cor_inf_acc_r, field=value, output=COR_Infrastructure_Access_Density, check radio box to simplify polygons).
        xviii. Overlay "COR_IDDE_Reports_2_prj" onto "COR_Infrastructure_Access_Density"
        xix. Apply a Selection by Location (Target= COR_IDDE_Reports_2_prj, Source= COR_Infrastructure_Access_Density, Method=within).
        xx. Report the number of records captured and not captured by "COR_Infrastructure_Access_Density", for both ID_Type and Pollutant_2, in the excel spreadsheet.
        xxi. Repeat steps xvi - xx with FFX data
    b. Sensitivity Analysis
        i. Copy Raster Tool (Data Management → Raster → Raster Dataset → Copy Raster) – convert risk-index raster to pixel type = 32_Bit_Float. This format enables raster calculator operations (input=ffx_inf_acc_r raster, output=ffx_iad_fac, configuration keyword=<null>, ignore background value=<null>, NoData value=<null>, Pixel Type=32_BIT_FLOAT).
        ii. Clip Raster (Data Management → Raster → Raster Processing → Clip) to the geometric extents of the Fairfax County jurisdictional boundary (input raster=ffx_iad_fac, output extent=FFX_County_Border_Prj, check radio button to use input features, output=ffx_iad_fac_c, NoData value=null).
        iii. Repeat with cor_inf_acc_r data
11. Construction Site Density
    a. Spatial Analysis
        i. Begin with Roanoke City data in Construction_General_Permits_Selection.xlsx (filtered out from the original dataset 2014_CGP_ACTIVE_WEB.xlsx)
        ii. Add the RoanokeCity sheet to the map, right click on the table and select Display XY Data. Select "Site_Longitude" as the X-Field, "Site_Latitude" as the Y-Field, the GCS NAD 1983 coordinate system, click OK.
        iii. A temporary layer will be created displaying the spatial location of those records with values for latitude and longitude. Verify the locations of these points using Google Maps and citystreets layer.
        iv. Export the data into the City of Roanoke feature dataset (COR_Active_Construction_Site_pts).
        v. Project (Data Management → Projections and Transformations → Project) data to Processed_Data_MapReady dataset (input=COR_Active_Construction_Site_pts, output=COR_Active_Construction_Site_pts_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).
        vi. Use Kernel Density Tool (Spatial Analyst → Density → Kernel Density) to get raster layer with density of generating sites per square kilometers (input= COR_Active_Construction_Site_pts_prj, population field=none,

    output=cor_cgp_dens, output cell size=30 m, search radius=<null>, area units=square kilometers, output values are= densities, method=planar).

    vii. Reclassify be Jenks Natural Breaks with three classes and compare to FFX data.
    viii. Return to start with FFX data (Construction_General_Permits_Selection.xlsx) (filtered out from the original dataset 2014_CGP_ACTIVE_WEB.xlsx)
    ix. Add FairfaxCounty sheet to the map, right click on the table and select Display XY Data. Select "Site_Longitude" as the X-Field, "Site_Latitude" as the Y-Field, the GCS NAD 1983 coordinate system, click OK.
    x. A temporary layer will be created displaying the spatial location of those records with values for latitude and longitude. Verify the locations of these points using Google Maps and Fairfax_Address_Points layer.
    xi. Export the data into the City of Roanoke feature dataset (FFX_Active_Construction_Site_pts).
    xii. Project (Data Management → Projections and Transformations → Project) data to Processed_Data_MapReady dataset (input=FFX_Active_Construction_Site_pts, output=FFX_Active_Construction_Site_pts_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).
    xiii. Use Kernel Density Tool (Spatial Analyst → Density → Kernel Density) to get raster layer with density of generating sites per square kilometers (input= FFX_Active_Construction_Site_pts_prj, population field=none, output=ffx_cgp_dens, output cell size=30 m, search radius=<null>, area units=square kilometers, output values are= densities, method=planar).
    xiv. Reclassify be Jenks Natural Breaks with three classes and compare to FFX data.
    xv. Datasets were evaluated and it was decided that individual Jenks Natural Breaks should be used to set the thresholds for each jurisdiction. The thresholds for moderate and high risk in Roanoke were 0.15 and 0.40, while in Fairfax they were 0.25 and 0.75.
    xvi. Reclassify raster layer (Spatial Analyst → Reclass → Reclassify Tool) to code it based on desired cutoffs (Input=ffx_cgp_dens, reclass field=value {natural breaks, then round to nice numbers}, output=ffx_cgp_r).
    xvii. Convert (Conversion Tools → From Raster → Raster to Polygon Tool) the reclassified raster into a polygon layer for selection purposes (Input=ffx_cgp_r, field=value, output=Fairfax_Active_Construction_Site_Density, check radio box to simplify polygons).
    xviii. Overlay "FFX_IDID_Reports_prj" onto "Fairfax_Active_Construction_Site_Density"
    xix. Apply a Selection by Location (Target= FFX_IDID_Reports_prj, Source= Fairfax_Active_Construction_Site_Density, Method=within).
    xx. Report the number of records captured and not captured by "Fairfax_Active_Construction_Site_Density", for both ID_Type and Pollutant_2, in the excel spreadsheet.
    xxi. Repeat steps xvi - xx with COR data

b. Sensitivity Analysis
    i. Copy Raster Tool (Data Management → Raster → Raster Dataset → Copy Raster) – convert risk-index raster to pixel type = 32_Bit_Float. This format

enables raster calculator operations (input=ffx_cgp_r raster, output=ffx_cgp_fac, configuration keyword=<null>, ignore background value=<null>, NoData value=<null>, Pixel Type=32_BIT_FLOAT).

    ii. Clip Raster (Data Management → Raster → Raster Processing → Clip) to the geometric extents of the Fairfax County jurisdictional boundary (input raster=ffx_cgp_fac, output extent=FFX_County_Border_Prj, check radio button to use input features, output=ffx_cgp_fac_c, NoData value=null).
    iii. Repeat with cor_cgp_r data

12. Aging Septic System Density
    a. Spatial Analysis
        i. Begin with data for Fairfax County: Septic_Systems_October_2014.shp, imported to gdb as Fairfax_Septic_Systems_Oct2014; Tax_Administrations_Real_Estate__Dwelling_Data.csv
        ii. Convert parcels to points (Data Management Tools → Features → Feature to Point) (input=Fairfax_Septic_Systems_Oct2014, output= Fairfax_Septic_Systems_Pts, uncheck radio box for Inside)
        iii. Create a pivot table for the Tax_Administrations_Real_Estate__Dwelling_Data.csv and determine the following for each parcel: PIN, Old_YRBLT, New_YRBLT, Count_YRBLT, Old_YRREMOD, New_YRREMOD, Count_YRREMOD. Copy values into a new spreadsheet with blank formatting (FFX_BuildOutYears.xlsx).
        iv. Add a field to Fairfax_Septic_Systems_Pts of type double called BuildOut_Year
        v. Join FFX_BuildOutYears.xlsx to Fairfax_Septic_Systems_Pts via Parcel_PIN field
        vi. Calculate BuildOut_Year field = New_YRBLT and remove the join
        vii. Was the parcel with septic developed more than 30 years ago? Query parcels for BuildOutYear < 1986.
        viii. Export the aging septic systems to a new layer (Fairfax_Aging_Septic_Pts).
        ix. Project (Data Management → Projections and Transformations → Project) data to Processed_Data_MapReady dataset (input= Fairfax_Aging_Septic_Pts, output=Fairfax_Aging_Septic_Pts _prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).
        x. Use Kernel Density Tool (Spatial Analyst → Density → Kernel Density) to get raster layer with density of septic parcels per square kilometer (input=Fairfax_Aging_Septic_Pts_prj, pop field=none, output=ffx_septic_d, cell size=30 m, search radius=null, area units=square_kilometers, output values are=densities, method=planar).
        xi. Symbolize by Jenks Natural Breaks in 3 classes and compare to COR data.
        xii. Return to start with COR data: citywide12012015.shp, WVWA_Services.txt, RealImprov.xlsx, Septic Targets from VDH.docx
        xiii. Add new fields to COR_citywide_12012015 layer: “WVWA_Service” (text) and “BuildOut_Year” (double)
        xiv. Join WVWA_Services table to COR_citywide_12012015 layer (TAXID on parcels to TAX_NUM on WVWA_Services) and calculate WVWA_Service field = ServiceType; remove join

xv. Create “RealImprov” pivot table to produce a table by parcel with attributes LRSN, Oldest Structure Date, Newest Structure Date. I copied these values into a new sheet with blank formatting (Year_Built.xlsx): LRSN, Old_YrBuilt, New_YrBuilt, Count_YrBuilt

xvi. Join new build out table to parcels (LRSN) and calculate BuildOut_Year = New_YrBuilt; remove join

xvii. Is there a structure over $38k present on the lot? Query parcels for DWELLING_1 >= 38000

xviii. If “yes,” does the parcel have sewer service? Select from selected features WVWA_Service IS NULL OR WVWA_Service NOT LIKE ‘%SW%’

xix. If “no,” probable septic system. Check these hot spots against VDH list of known septic areas to ensure accuracy.

xx. Is build out year over 30 years ago? Query selected features for BuildOut_Year < 1986 OR BuildOut_Year IS NULL

xxi. If “yes,” export parcels (COR_Aging_Septic_Parcels)

xxii. Convert parcels to points (Data Management Tools → Features → Feature to Point) (input= COR_Aging_Septic_Parcels, output= COR_Aging_Septic _Pts)

xxiii. Project (Data Management → Projections and Transformations → Project) data to Processed_Data_MapReady dataset (input= COR_Aging_Septic _Pts, output= Roanoke_Aging_Septic _Pts_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).

xxiv. Use Kernel Density Tool (Spatial Analyst → Density → Kernel Density) to get raster layer with density of septic parcels per square kilometer (input= Roanoke_Aging_Septic _Pts_prj, pop field=none, output=cor_septic_d, cell size=30 m, search radius=null, area units=square_kilometers, output values are=densities, method=planar).

xxv. Symbolize by Jenks Natural Breaks and compare with Fairfax data.

xxvi. Ranges by Natural Breaks are similar enough that the same ranges for septic density risk levels are warranted in each locality. Start with natural breaks, and set the high risk category to the metric equivalent of CWP recommendation (> 100 per sqmi = > 40 per sqkm). Low = 0 – 15, Moderate = 15 – 40, High = > 40 old fields per square kilometer

xxvii. Reclassify (Spatial Analyst Tools → Reclass → Reclassify) to assign IDP risk ranking (input=cor_septic_d, reclass field=value {0-50 → 1, 50-100 → 2, > 100 → 3, NoData → 0}, output=cor_septic_r)

xxviii. Convert (Conversion Tools → From Raster → Raster to Polygon Tool) the reclassified raster into a polygon layer for selection purposes (Input=cor_septic_r, field=value, output=Roanoke_Aging_Septic_System_Density, check radio box to simplify polygons).

xxix. Overlay “COR_IDDE_Reports_2_prj” onto “Roanoke_Aging_Septic_System_Density”

xxx. Apply a Selection by Location (Target= COR_IDDE_Reports_2_prj, Source= Roanoke_Aging_Septic_System_Density, Method=within).

        xxxi. Report the number of records captured and not captured by “Roanoke_Aging_Septic_System_Density”, for both ID_Type and Pollutant_2, in the excel spreadsheet.
        xxxii. Repeat steps xxvii - xxxi with FFX data
    b. Sensitivity Analysis
        i. Copy Raster Tool (Data Management → Raster → Raster Dataset → Copy Raster) – convert risk-index raster to pixel type = 32_Bit_Float. This format enables raster calculator operations (input=ffx_septic_r raster, output=ffx_sep_fac, configuration keyword=<null>, ignore background value=<null>, NoData value=<null>, Pixel Type=32_BIT_FLOAT).
        ii. Clip Raster (Data Management → Raster → Raster Processing → Clip) to the geometric extents of the Fairfax County jurisdictional boundary (input raster=ffx_sep_fac, output extent=FFX_County_Border_Prj, check radio button to use input features, output=ffx_sep_fac_c, NoData value=null).
        iii. Repeat with cor_septic_r data

13. Dog License Density
    a. Spatial Analysis
        i. Begin with City of Roanoke data: COR_ActiveDogAddresses.xls, COR_citystreets layer
        ii. Prepare the spreadsheet of dog license addresses for geocoding. Ensure all features have full address, city, and state fields of type “string.” Save in a local folder.
        iii. Pre-process citystreets layer to include columns of type “string” for Lo_Left_Addr, Hi_Left_Addr, Lo_Right_Addr, Hi_Right_Addr, CITY, and STATE.
        iv. Create Address Locator (Geocoding Tools → Create Address Locator) (used style=US Dual Ranges, file=COR_citystreets as Primary Table, fields mapped: Feature ID→OBJECTID, From Left→Lo_Left_Addr, To Left→Hi_Left_Addr, From Right→Lo_Right_Addr, To Right→Hi_Right_Addr, Street Name→STR_NAME, Suffix Type→SFX_TYPE, Suffix Direction→SFX_DIR, Left City or Place→City, Right City or Place→City, Left State→, Right State→State). Make sure to save address locator in a basic folder in a local directory.
        v. Geocode addresses for dog licenses (Geocoding Tools → Geocode Addresses) to a point FC (settings: input table=Dog_Licenses, input address locator= Dog_License_Locator_2, select the multiple fields radio button, fields mapped: Street→Full_Address, City→, State→, output= COR_Dog_Licenses)
        vi. Inspect resulting feature classes and tables for matches, accuracy, and to rematch any unmatched or tied addresses.
        vii. Export matched address points to the City of Roanoke Feature Dataset.
        viii. Project (Data Management → Projections and Transformations → Project) the data to Processed_Data_MapReady dataset (input=COR_Dog_Licenses_Export, output= COR_Dog_Licenses_Export_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).
        ix. Use Kernel Density Tool (Spatial Analyst → Density → Kernel Density) to get raster layer with density of dog licenses per square kilometer

(input=COR_Dog_Licenses_Export_prj, population field=none, output raster=cor_dog_lic_d, output cell size=30 m, search radius=<null>, area units=square kilometers, output values are= densities, method=planar).

x. Classify densities by Jenks Natural Breaks 3 classes and compare to Fairfax data

xi. Return to start with Fairfax data: Fairfax_Address_Points layer, FFX CO Dog Licenses 2015 Sheet 1.xlsx, and FFX CO Dog Licenses 2015 Sheet 2.xlsx

xii. Process the Fairfax dog license tables by removing the first two rows (so that the first row is the column headers) and calculating a new field “Full_Address” based on the six (6) major address fields in the spreadsheet: StrNo, Addn, Pre, Street, StrType, and Post. Ensure all features have full address, city, and state fields of type “string.” Save in a local folder

xiii. Pre-process Fairfax_Roadway_Centerlines layer to include “string” columns for RW_TYPE_US, RW_PREFIX, RW_SUFFIX, Lo_Left_Addr, Hi_Left_Addr, Lo_Right_Addr, Hi_Right_Addr, Jurisdiction, and State, which match the terms used in the dog licenses spreadsheet. Save in a geodatabase outside of all feature datasets (FFX_Roadway_Centerlines_edit).

xiv. Create Address Locator (Geocoding Tools → Create Address Locator) (used style=US Dual Ranges, file=FFX_Roadway_Centerlines_edit as Primary Table, fields mapped: Feature ID→OBJECTID, From Left→Lo_Left_Addr, To Left→Hi_Left_Addr, From Right→Lo_Right_Addr, To Right→Hi_Right_Addr, Prefix Direction→RW_PREFIX, Street Name→RW_NAME, Suffix Type→RW_TYPE_US, Suffix Direction→RW_SUFFIX, Left City or Place→L_JURISDIC, Right City or Place→R_JURISDIC, Left State→, Right State→State). Make sure to save address locator in a basic folder in a local directory.

xv. Geocode addresses for dog licenses (Geocoding Tools → Geocode Addresses) to a point FC (settings: input table=Sheet1 of FFX_CO_Dog_Licenses_2015_ShtX_Processed, input address locator= FFX_Dog_License_Locator, select the multiple fields radio button, fields mapped: Street→Full_Address, City→City, State→State, output= FFX_Dog_Licenses_X).

xvi. Inspect resulting feature classes and tables for matches, accuracy, and to rematch any unmatched or tied addresses. Repeat steps xv and xvi for both Fairfax Dog License spreadsheets.

xvii. Export matched address points to the Fairfax County Feature Dataset.

xviii. Append feature classes for each spreadsheet into one feature class (input=FFX_Dog_Licenses_2_Export, target=FFX_Dog_Licenses_1_Export, schema type=NO_TEST, accept default field maps).

xix. Project (Data Management → Projections and Transformations → Project) the data to Processed_Data_MapReady dataset (input=FFX_Dog_Licenses_1_Export, output= FFX_Dog_Licenses_Export_prj, output coordinate system=NAD_1983_Contiguous_USA_Albers).

xx. Use Kernel Density Tool (Spatial Analyst → Density → Kernel Density) to get raster layer with density of dog licenses per square kilometer (input=FFX_Dog_Licenses_Export_prj, population field=none, output raster=ffx_dog_lic_d, output cell size=30 m, search radius=<null>, area units=square kilometers, output values are= densities, method=planar).

xxi. Classify densities by Jenks Natural Breaks 3 classes and compare to Fairfax data
xxii. Natural breaks are similar enough to use the same breaks for both MS4s: Low = 0 – 30, Moderate = 30 – 100, High = > 100. Continue with both datasets.
xxiii. Reclassify raster layer (Spatial Analyst → Reclass → Reclassify Tool) to code it based on desired cutoffs (Input=ffx_dog_lic_d, reclass field=value {0-30→1, 30-100→2, >100→3, NoData→0}, output=ffx_dog_lic_r).
xxiv. Convert (Conversion Tools → From Raster → Raster to Polygon Tool) the reclassified raster into a polygon layer for selection purposes (Input=ffx_dog_lic_r, field=value, output=Fairfax_Dog_License_Density, check radio box to simplify polygons).
xxv. Overlay "FFX_IDID_Reports_prj" onto "Fairfax_Dog_License_Density"
xxvi. Apply a Selection by Location (Target= FFX_IDID_Reports_prj, Source= Fairfax_Dog_License_Density, Method=within).
xxvii. Report the number of records captured and not captured by "Fairfax_Dog_License_Density", for both ID_Type and Pollutant_2, in the excel spreadsheet.
xxviii. Repeat steps xxiii - xxvii with COR data

b. Sensitivity Analysis
    i. Copy Raster Tool (Data Management → Raster → Raster Dataset → Copy Raster) – convert risk-index raster to pixel type = 32_Bit_Float. This format enables raster calculator operations (input=ffx_dog_lic_r raster, output=ffx_dog_fac, configuration keyword=<null>, ignore background value=<null>, NoData value=<null>, Pixel Type=32_BIT_FLOAT).
    ii. Clip Raster (Data Management → Raster → Raster Processing → Clip) to the geometric extents of the Fairfax County jurisdictional boundary (input raster=ffx_dog_fac, output extent=FFX_County_Border_Prj, check radio button to use input features, output=ffx_dog_fac_c, NoData value=null).
    iii. Repeat with cor_dog_lic_r data

Composite Map Production

1. Fairfax County Unweighted
    a. Use the Raster Calculator (Spatial Analyst Tools → Map Algebra → Raster Calculator) to combine the 13 risk factors for Fairfax County into a composite map with equal weights. Use the equation below:

ffx_idp_eq = 1 * “ffx_dog_fac_c” + 1 * “ffx_ass_fac_c” + 1 * “ffx_wdd_fac_c” + 1 * “ffx_dev_fac_c” + 1 * “ffx_imp_fac_c” + 1 * “ffx_cgp_fac_c” + 1 * “ffx_gen_fac_c” + 1 * “ffx_sep_fac_c” + 1 * “ffx_iad_fac_c” + 1 * “ffx_ofd_fac_c” + 1 * “ffx_pop_fac_c” + 1 * “landuse_fac_c” + 1 * “nlcd_fac_c”

    b. Reclassify Raster (Spatial Analyst Tools → Reclass → Reclassify) into three levels of risk (input=ffx_idp_eq, reclass field=value {0-13 → 1 (Low), 13-26 → 2 (Moderate), 26-39 → 3 (High), NoData → NoData}, output=ffx_idp_eq_r).
    c. Symbolize and produce map.

2. City of Roanoke Unweighted
    a. Use the Raster Calculator (Spatial Analyst Tools → Map Algebra → Raster Calculator) to combine the 13 risk factors for Roanoke into a composite map with equal weights. Use the equation below:

cor_idp_eq = 1 * “cor_dog_fac_c” + 1 * “cor_ass_fac_c” + 1 * “cor_wdd_fac_c” + 1 * “cor_dev_fac_c” + 1 * “cor_imp_fac_c” + 1 * “cor_cgp_fac_c” + 1 * “cor_gen_fac_c” + 1 * “cor_sep_fac_c” + 1 * “cor_iad_fac_c” + 1 * “cor_ofd_fac_c” + 1 * “cor_pop_fac_c” + 1 * “landuse_fac_c” + 1 * “nlcd_fac_c”

    b. Reclassify Raster (Spatial Analyst Tools → Reclass → Reclassify) into three levels of risk (input=cor_idp_eq, reclass field=value {0-13 → 1 (Low), 13-26 → 2 (Moderate), 26-39 → 3 (High), NoData → NoData}, output=cor_idp_eq_r).
    c. Symbolize and produce map.

3. Fairfax County Weighted Scheme 1
    a. Use the Raster Calculator (Spatial Analyst Tools → Map Algebra → Raster Calculator) to combine the 13 risk factors for Fairfax County into a weighted composite map. Use the equation below:

ffx_idp_wt1 = 1 * “ffx_dog_fac_c” + 1 * “ffx_ass_fac_c” + 1 * “ffx_wdd_fac_c” + 1 * “ffx_dev_fac_c” + 2 * “ffx_imp_fac_c” + 1 * “ffx_cgp_fac_c” + 1 * “ffx_gen_fac_c” + 1 * “ffx_sep_fac_c” + 1 * “ffx_iad_fac_c” + 1 * “ffx_ofd_fac_c” + 2 * “ffx_pop_fac_c” + 2 * “landuse_fac_c” + 1 * “nlcd_fac_c”

    b. Reclassify Raster (Spatial Analyst Tools → Reclass → Reclassify) into three levels of risk (input=ffx_idp_wt1, reclass field=value {0-16 → 1 (Low), 16-32 → 2 (Moderate), 32-48 → 3 (High), NoData → NoData}, output=ffx_idp_wt1_r).
    c. Symbolize and produce map.

4. City of Roanoke Weighted Scheme 1
    a. Use the Raster Calculator (Spatial Analyst Tools → Map Algebra → Raster Calculator) to combine the 13 risk factors for Roanoke into a weighted composite map. Use the equation below:

cor_idp_wt1 = 1 * “cor_dog_fac_c” + 1 * “cor_ass_fac_c” + 1 * “cor_wdd_fac_c” + 1 * “cor_dev_fac_c” + 2 * “cor_imp_fac_c” + 1 * “cor_cgp_fac_c” + 1 * “cor_gen_fac_c” + 1 * “cor_sep_fac_c” + 1 * “cor_iad_fac_c” + 1 * “cor_ofd_fac_c” + 2 * “cor_pop_fac_c” + 2 * “landuse_fac_c” + 1 * “nlcd_fac_c”

    b. Reclassify Raster (Spatial Analyst Tools → Reclass → Reclassify) into three levels of risk (input=cor_idp_wt1, reclass field=value {0-16 → 1 (Low), 16-32 → 2 (Moderate), 32-48 → 3 (High), NoData → NoData}, output=cor_idp_wt1_r).
    c. Symbolize and produce map.

5. Fairfax County Weighted Scheme 2
    a. Use the Raster Calculator (Spatial Analyst Tools → Map Algebra → Raster Calculator) to combine the 13 risk factors for Fairfax County into a second weighted composite map. Use the equation below:

ffx_idp_wt2 = 0.5 * “ffx_dog_fac_c” + 1 * “ffx_ass_fac_c” + 0.5 * “ffx_wdd_fac_c” + 1 * “ffx_dev_fac_c” + 2 * “ffx_imp_fac_c” + 0.5 * “ffx_cgp_fac_c” + 0.5 * “ffx_gen_fac_c” + 1 * “ffx_sep_fac_c” + 0.5 * “ffx_iad_fac_c” + 0.5 * “ffx_ofd_fac_c” + 2 * “ffx_pop_fac_c” + 2 * “landuse_fac_c” + 0.5 * “nlcd_fac_c”

    b. Reclassify Raster (Spatial Analyst Tools → Reclass → Reclassify) into three levels of risk (input=ffx_idp_wt2, reclass field=value {0-12.5 → 1 (Low), 12.5-25 → 2 (Moderate), 25-37.5 → 3 (High), NoData → NoData}, output=ffx_idp_wt2_r).
    c. Symbolize and produce map.

6. City of Roanoke Weighted Scheme 2
    a. Use the Raster Calculator (Spatial Analyst Tools → Map Algebra → Raster Calculator) to combine the 13 risk factors for Roanoke into a second weighted composite map. Use the equation below:

cor_idp_wt2 = 0.5 * “cor_dog_fac_c” + 1 * “cor_ass_fac_c” + 0.5 * “cor_wdd_fac_c” + 1 * “cor_dev_fac_c” + 2 * “cor_imp_fac_c” + 0.5 * “cor_cgp_fac_c” + 0.5 * “cor_gen_fac_c” + 1 * “cor_sep_fac_c” + 0.5 * “cor_iad_fac_c” + 0.5 * “cor_ofd_fac_c” + 2 * “cor_pop_fac_c” + 2 * “landuse_fac_c” + 0.5 * “nlcd_fac_c”

   b. Reclassify Raster (Spatial Analyst Tools → Reclass → Reclassify) into three levels of risk (input=cor_idp_wt2, reclass field=value {0-12.5 → 1 (Low), 12.5-25 → 2 (Moderate), 25-37.5 → 3 (High), NoData → NoData}, output=cor_idp_wt2_r).
   c. Symbolize and produce map.

7. Fairfax County Weighted Scheme 3
   a. Use the Raster Calculator (Spatial Analyst Tools → Map Algebra → Raster Calculator) to combine the 13 risk factors for Fairfax County into a third weighted composite map. Use the equation below:

ffx_idp_wt3 = 1 * “ffx_dog_fac_c” + 1.5 * “ffx_ass_fac_c” + 1.5 * “ffx_wdd_fac_c” + 2 * “ffx_dev_fac_c” + 2 * “ffx_imp_fac_c” + 1.5 * “ffx_cgp_fac_c” + 2 * “ffx_gen_fac_c” + 1 * “ffx_sep_fac_c” + 1.5 * “ffx_iad_fac_c” + 1 * “ffx_ofd_fac_c” + 1.5 * “ffx_pop_fac_c” + 2 * “landuse_fac_c” + 2 * “nlcd_fac_c”

   b. Reclassify Raster (Spatial Analyst Tools → Reclass → Reclassify) into three levels of risk (input=ffx_idp_wt3, reclass field=value {0-20.5 → 1 (Low), 20.5-41 → 2 (Moderate), 41-61.5 → 3 (High), NoData → NoData}, output=ffx_idp_wt3_r).
   c. Symbolize and produce map.

8. City of Roanoke Weighted Scheme 3
   a. Use the Raster Calculator (Spatial Analyst Tools → Map Algebra → Raster Calculator) to combine the 13 risk factors for Roanoke into a third weighted composite map. Use the equation below:

cor_idp_wt3 = 1 * “cor_dog_fac_c” + 1.5 * “cor_ass_fac_c” + 1.5 * “cor_wdd_fac_c” + 2 * “cor_dev_fac_c” + 2 * “cor_imp_fac_c” + 1.5 * “cor_cgp_fac_c” + 2 * “cor_gen_fac_c” + 1 * “cor_sep_fac_c” + 1.5 * “cor_iad_fac_c” + 1 * “cor_ofd_fac_c” + 1.5 * “cor_pop_fac_c” + 2 * “landuse_fac_c” + 2 * “nlcd_fac_c”

   b. Reclassify Raster (Spatial Analyst Tools → Reclass → Reclassify) into three levels of risk (input=cor_idp_wt3, reclass field=value {0-20.5 → 1 (Low), 20.5-41 → 2 (Moderate), 41-61.5 → 3 (High), NoData → NoData}, output=cor_idp_wt3_r).
   c. Symbolize and produce map.

APPENDIX B – R CODE FOR FISHER EXACT TESTS

TWO-TAILED FISHER EXACT TEST

```
setwd("[path to working directory]")
.pardefault <- par(no.readonly = T)
x <- read.csv("[name of jurisdiction data file].csv", header = TRUE, skip = 0, na.string=c("","NA"))

factors<-ordered(x$Factor)

output=matrix(NA,25,25,dimnames=list(factors, factors))
output=output[-1,-1] #removes "Number in Group" from output table

row<-2 #start with row 2, since row 1 is the total number
row.compare<-row+1 #the first row compared to row 2 in a contingency table
k<-23 #number of comparisons made with the first row
col<-6 #column indicating discharge or pollutant cluster in data file
#Outer loop cycles through row
#Inner loop cycles through row.compare - compares row 2 with 3,4,5..., then row 3 with 4,5,6...
for(i in seq(1,23,1)){
 for(j in 1:k){ #k decreases by 1 for each iteration of outer loop
  #the structure for calling rows and columns in a data frame is dataframename[row,column]
  #name the two factors
  factor1name<-as.character(x[row,1])
  factor2name<-as.character(x[row.compare,1])
  #assign values to cells 1-4
  n<-x[1,col] #grabs the total number from the first row
  factor1Y<-x[row,col]
  factor1N<-n - factor1Y
  factor2Y<-x[row.compare,col]
  factor2N<-n - factor2Y
  m<-matrix(c(factor1Y,factor1N,factor2Y,factor2N),nrow=2,
dimnames=list(c("Y","N"),c(factor1name,factor2name)))
  [name of cluster].test<-fisher.test(m, alternative="t") #two-tailed fisher test on m, change name for
each cluster

  output[row-1,row.compare-1] <- [name of cluster].test$p.value #change name for each cluster

  row.compare<-row.compare+1
 }
 #set parameters for next iteration of outer loop
 row<-row+1
 row.compare<-row+1
 k<-k-1
}
output

write.csv(output, file="[name of cluster and jurisdiction tested]_2_tailed.csv")
```

ONE-TAILED FISHER EXACT TEST

```
setwd("[path to working directory]")
.pardefault <- par(no.readonly = T)
x <- read.csv("[name of jurisdiction data file].csv", header = TRUE, skip = 0, na.string=c("","NA"))

factors<-ordered(x$Factor)

output=matrix(NA,25,25,dimnames=list(factors, factors))
output=output[-1,-1] #removes "Number in Group" from output table

row<-2 #start with row 2, since row 1 is the total number
row.compare<-row+1 #the first row compared to row 2 in a contingency table
k<-23 #number of comparisons made with the first row
col<-6 #column indicating discharge or pollutant cluster in data file
#Outer loop cycles through row
#Inner loop cycles through row.compare - compares row 2 with 3,4,5..., then row 3 with 4,5,6...
for(i in seq(1,23,1)){
  for(j in 1:k){ #k decreases by 1 for each iteration of outer loop
    #the structure for calling rows and columns in a data frame is dataframename[row,column]
    #name the two factors
    factor1name<-as.character(x[row,1])
    factor2name<-as.character(x[row.compare,1])
    #assign values to cells 1-4
    n<-x[1,col] #grabs the total number from the first row
    factor1Y<-x[row,col]
    factor1N<-n - factor1Y
    factor2Y<-x[row.compare,col]
    factor2N<-n - factor2Y
    m<-matrix(c(factor1Y,factor1N,factor2Y,factor2N),nrow=2,
dimnames=list(c("Y","N"),c(factor1name,factor2name)))
    #tests which factor has stronger effect; null= factor2 performs better than factor1 (OR<=1)
    [name of cluster].test2 <- fisher.test(m, alternative="g") #change name for each cluster
    #tests which factor has stronger effect; null= factor1 performs better than factor2 (OR>=1)
    [name of cluster].test3 <- fisher.test(m, alternative="l") #change name for each cluster
    if([name of cluster].test3$p.value<=0.05) {
      output[row-1,row.compare-1] <- [name of cluster].test3$p.value #change name for each cluster
    }
    if([name of cluster].test2$p.value<=0.05) {
      output[row.compare-1,row-1] <- [name of cluster].test2$p.value #change name for each cluster
    }
        row.compare<-row.compare+1
  }
  #set parameters for next iteration of outer loop
  row<-row+1
  row.compare<-row+1
  k<-k-1
}
output
write.csv(output, file="[name of cluster and jurisdiction tested]_1_Tailed.csv")
```

APPENDIX C – SPATIAL ANALYSIS DATA

Table C-1: Discharges captured by individual risk factors during the spatial overlay process for the City of Roanoke.

	Illicit Connection	Illegal Dumping or Spill	Fecal Bacteria	Chloride	Construction	FOG, Surfactants, Trash	Hazardous Materials	Unknown
Records in Cluster	7	83	9	0	9	39	31	2
1. Land Cover								
Developed, Open Space	0	0	0	0	0	0	0	0
Developed, Low Intensity	0	13	1	0	1	0	11	0
Developed, Medium Intensity	4	31	6	0	3	14	10	2
Developed, High Intensity	3	39	2	0	5	25	10	0
Barren	0	0	0	0	0	0	0	0
Agricultural	0	0	0	0	0	0	0	0
Undeveloped	0	0	0	0	0	0	0	0
2. Land Use								
Agricultural	0	0	0	0	0	0	0	0
Commercial	1	52	5	0	4	33	9	2
Industrial	3	8	0	0	2	0	9	0
Institutional	0	4	1	0	0	1	2	0
Recreation and Open Space	0	6	1	0	0	1	4	0
Residential	3	13	2	0	3	4	7	0
3. Population Density	0	9	0	0	1	1	6	1
4. Percent Impervious	2	35	3	0	5	18	11	0
5. Development Age	4	51	5	0	6	24	19	1
6. Outfall Density	0	3	0	0	0	0	3	0
7. Aging Sanitary Infrastructure	2	36	3	0	4	21	10	0
8. Drainage Density	2	36	3	0	4	20	11	0
9. Generating Site Density	3	50	5	0	5	26	17	0
10. Infrastructure Access Density	2	38	5	0	4	24	7	0
11. Construction Site Density	1	20	3	0	1	9	7	1
12. Aging Septic System Density	0	2	0	0	0	0	2	0
13. Dog License Density	0	10	0	0	1	2	6	1

Table C-2: Discharges captured by individual risk factors during the spatial overlay process for the County of Fairfax.

	Illicit Connection	**Illegal Dumping or Spill**	**Fecal Bacteria**	**Chloride**	**Construction**	**FOG, Surfactants, Trash**	**Hazardous Materials**	**Unknown**
Records in Cluster	23	77	8	8	12	30	36	6
1. Land Cover								
Developed, Open Space	4	13	2	1	5	4	5	0
Developed, Low Intensity	7	12	3	4	4	2	4	2
Developed, Medium Intensity	7	25	1	1	3	10	15	2
Developed, High Intensity	1	23	1	1	0	12	10	0
Barren	0	0	0	0	0	0	0	0
Agricultural	0	1	0	0	0	0	1	0
Undeveloped	4	3	1	1	0	2	1	2
2. Land Use								
Agricultural	0	0	0	0	0	0	0	0
Commercial	7	26	2	1	1	14	13	2
Industrial	1	16	0	0	1	8	8	0
Institutional	1	3	1	1	1	0	1	0
Recreation and Open Space	7	8	1	3	0	3	7	1
Residential	7	24	4	3	9	5	7	3
3. Population Density	1	0	0	0	0	0	1	0
4. Percent Impervious	14	55	7	4	7	21	27	3
5. Development Age	9	43	6	5	5	17	16	3
6. Outfall Density	6	11	1	3	3	4	6	0
7. Aging Sanitary Infrastructure	11	34	1	3	6	14	18	3
8. Drainage Density	11	46	6	5	7	17	19	3
9. Generating Site Density	10	28	2	1	2	11	19	3
10. Infrastructure Access Density	11	49	3	1	7	22	25	2
11. Construction Site Density	10	23	2	0	4	11	16	0
12. Aging Septic System Density	1	4	0	1	0	2	1	1
13. Dog License Density	1	20	2	0	4	9	4	2

Table C-3: Example contingency table created in R for the Fisher exact test. Data taken from City of Roanoke aging septic system density risk factor results.

	Illicit Connection	**Illegal Dumping and Spill**
Captured	0	2
Not Captured	7	81

Table C-4: R output showing p-values from two-tailed Fisher exact tests on the Roanoke illicit connections cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	1E+00	7E-02	2E-01	1E+00	1E+00	1E+00	1E+00	1E+00	2E-01	1E+00	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	5E-01	7E-02	5E-01	5E-01	1E+00
LC_Developed_Low_Intensity	NA	NA	7E-02	2E-01	1E+00	1E+00	1E+00	1E+00	1E+00	2E-01	1E+00	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	5E-01	7E-02	5E-01	5E-01	1E+00
LC_Developed_Medium_Intensity	NA	NA	NA	1E+00	7E-02	7E-02	7E-02	7E-02	3E-01	1E+00	7E-02	7E-02	1E+00	7E-02	7E-02	6E-01	7E-02	1E+00	3E-01	6E-01	1E+00	6E-01	6E-01	7E-02
LC_Developed_High_Intensity	NA	NA	NA	NA	2E-01	2E-01	2E-01	2E-01	6E-01	1E+00	2E-01	2E-01	1E+00	2E-01	2E-01	1E+00	2E-01	1E+00	6E-01	1E+00	1E+00	1E+00	1E+00	2E-01
LC_Barren	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	2E-01	1E+00	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	5E-01	7E-02	5E-01	5E-01	1E+00
LC_Agricultural	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	2E-01	1E+00	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	5E-01	7E-02	5E-01	5E-01	1E+00
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	2E-01	1E+00	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	5E-01	7E-02	5E-01	5E-01	1E+00
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	2E-01	1E+00	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	5E-01	7E-02	5E-01	5E-01	1E+00
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	1E+00	1E+00	6E-01	1E+00	1E+00	1E+00	1E+00	6E-01	1E+00	1E+00	3E-01	1E+00	1E+00	1E+00
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	2E-01	1E+00	2E-01	2E-01	1E+00	2E-01	1E+00	6E-01	1E+00	1E+00	1E+00	1E+00	2E-01
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	5E-01	7E-02	5E-01	5E-01	1E+00
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	5E-01	7E-02	5E-01	5E-01	1E+00
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	2E-01	1E+00	2E-01	1E+00	6E-01	1E+00	1E+00	1E+00	1E+00	2E-01
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	5E-01	1E+00	2E-01	1E+00	5E-01	7E-02	5E-01	5E-01	1E+00
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	1E+00	2E-01	1E+00	5E-01	7E-02	5E-01	5E-01	1E+00
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	1E+00	1E+00	1E+00	6E-01	1E+00	1E+00	5E-01
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	1E+00	5E-01	7E-02	5E-01	5E-01	1E+00
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	1E+00	1E+00	1E+00	1E+00	2E-01
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	3E-01	1E+00	1E+00	1E+00
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	1E+00	1E+00	5E-01
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	6E-01	7E-02
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	5E-01
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-5: R output showing p-values from two-tailed Fisher exact tests on the Roanoke illegal dumping and spills cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	1E-04	3E-11	1E-14	1E+00	1E+00	1E+00	1E+00	3E-21	7E-03	1E-01	3E-02	1E-04	3E-03	2E-01	3E-14	5E-01	3E-20	5E-07	6E-13	1E-20	2E-13	2E-13	1E-03
LC_Developed_Low_Intensity	NA	NA	3E-03	2E-05	1E-04	1E-04	1E-04	1E-04	6E-10	4E-01	4E-02	1E-01	1E+00	5E-01	2E-02	4E-05	5E-03	4E-09	2E-01	3E-04	2E-09	1E-04	1E-04	7E-01
LC_Developed_Medium_Intensity	NA	NA	NA	3E-01	3E-11	3E-11	3E-11	3E-11	2E-03	4E-05	2E-07	4E-06	3E-03	1E-04	4E-08	3E-01	6E-09	5E-03	9E-02	6E-01	3E-03	5E-01	5E-01	3E-04
LC_Developed_High_Intensity	NA	NA	NA	NA	1E-14	1E-14	1E-14	1E-14	6E-02	1E-07	2E-10	6E-09	2E-05	4E-07	3E-11	1E+00	4E-12	1E-01	3E-03	6E-01	9E-02	8E-01	8E-01	1E-06
LC_Barren	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	3E-21	7E-03	1E-01	3E-02	1E-04	3E-03	2E-01	3E-14	5E-01	3E-20	5E-07	6E-13	1E-20	2E-13	2E-13	1E-03
LC_Agricultural	NA	NA	NA	NA	NA	NA	1E+00	1E+00	3E-21	7E-03	1E-01	3E-02	1E-04	3E-03	2E-01	3E-14	5E-01	3E-20	5E-07	6E-13	1E-20	2E-13	2E-13	1E-03
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	1E+00	3E-21	7E-03	1E-01	3E-02	1E-04	3E-03	2E-01	3E-14	5E-01	3E-20	5E-07	6E-13	1E-20	2E-13	2E-13	1E-03
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	3E-21	7E-03	1E-01	3E-02	1E-04	3E-03	2E-01	3E-14	5E-01	3E-20	5E-07	6E-13	1E-20	2E-13	2E-13	1E-03
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-13	3E-16	2E-14	6E-10	3E-12	3E-17	4E-02	2E-18	9E-01	9E-07	1E-02	1E+00	2E-02	2E-02	1E-11
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-01	8E-01	4E-01	1E+00	2E-01	2E-07	1E-01	4E-12	2E-02	2E-06	2E-12	1E-06	1E-06	8E-01
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	7E-01	4E-02	2E-01	1E+00	6E-10	7E-01	3E-15	7E-04	8E-09	9E-16	3E-09	3E-09	2E-01
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-01	6E-01	5E-01	2E-08	3E-01	1E-13	5E-03	2E-07	5E-14	8E-08	8E-08	4E-01
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	2E-02	4E-05	5E-03	4E-09	2E-01	3E-04	2E-09	1E-04	1E-04	7E-01
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-01	8E-07	6E-02	2E-11	4E-02	7E-06	7E-12	3E-06	3E-06	1E+00
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	8E-11	1E+00	3E-16	2E-04	1E-09	9E-17	5E-10	5E-10	8E-02
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	9E-12	9E-02	5E-03	8E-01	6E-02	9E-01	9E-01	2E-06
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-17	4E-05	2E-10	7E-18	6E-11	6E-11	3E-02
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-06	3E-02	1E+00	4E-02	4E-02	8E-11
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-02	2E-06	1E-02	1E-02	7E-02
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-02	1E+00	1E+00	2E-05
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-02	3E-02	3E-11
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E-05
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-05
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-6: R output showing p-values from two-tailed Fisher exact tests on the Roanoke fecal bacteria cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	1E+00	9E-03	5E-01	1E+00	1E+00	1E+00	1E+00	3E-02	1E+00	1E+00	1E+00	5E-01	1E+00	1E+00	3E-02	1E+00	3E-02	2E-01	2E-01	3E-02	2E-01	2E-01	1E+00
LC_Developed_Low_Intensity	NA	NA	5E-02	1E+00	1E+00	1E+00	1E+00	1E+00	1E-01	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E-01	1E+00	1E-01	6E-01	6E-01	1E-01	6E-01	6E-01	1E+00
LC_Developed_Medium_Intensity	NA	NA	NA	2E-01	9E-03	9E-03	9E-03	9E-03	1E+00	9E-03	5E-02	5E-02	2E-01	9E-03	9E-03	1E+00	9E-03	1E+00	3E-01	3E-01	1E+00	3E-01	3E-01	9E-03
LC_Developed_High_Intensity	NA	NA	NA	NA	5E-01	5E-01	5E-01	5E-01	3E-01	5E-01	1E+00	1E+00	1E+00	5E-01	5E-01	3E-01	5E-01	3E-01	1E+00	1E+00	3E-01	1E+00	1E+00	5E-01
LC_Barren	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	3E-02	1E+00	1E+00	1E+00	5E-01	1E+00	1E+00	3E-02	1E+00	3E-02	2E-01	2E-01	3E-02	2E-01	2E-01	1E+00
LC_Agricultural	NA	NA	NA	NA	NA	NA	1E+00	1E+00	3E-02	1E+00	1E+00	1E+00	5E-01	1E+00	1E+00	3E-02	1E+00	3E-02	2E-01	2E-01	3E-02	2E-01	2E-01	1E+00
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	1E+00	3E-02	1E+00	1E+00	1E+00	5E-01	1E+00	1E+00	3E-02	1E+00	3E-02	2E-01	2E-01	3E-02	2E-01	2E-01	1E+00
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	3E-02	1E+00	1E+00	1E+00	5E-01	1E+00	1E+00	3E-02	1E+00	3E-02	2E-01	2E-01	3E-02	2E-01	2E-01	1E+00
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-02	1E-01	1E-01	3E-01	3E-02	3E-02	1E+00	3E-02	1E+00	6E-01	6E-01	1E+00	6E-01	6E-01	3E-02
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	5E-01	1E+00	1E+00	3E-02	1E+00	3E-02	2E-01	2E-01	3E-02	2E-01	2E-01	1E+00
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E-01	1E+00	1E-01	6E-01	6E-01	1E-01	6E-01	6E-01	1E+00
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E-01	1E+00	1E-01	6E-01	6E-01	1E-01	6E-01	6E-01	1E+00
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	5E-01	3E-01	5E-01	3E-01	1E+00	1E+00	3E-01	1E+00	1E+00	5E-01
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	3E-02	1E+00	3E-02	2E-01	2E-01	3E-02	2E-01	2E-01	1E+00
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-02	1E+00	3E-02	2E-01	2E-01	3E-02	2E-01	2E-01	1E+00
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-02	1E+00	6E-01	6E-01	1E+00	6E-01	6E-01	3E-02
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-02	2E-01	2E-01	3E-02	2E-01	2E-01	1E+00
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	6E-01	1E+00	6E-01	6E-01	3E-02
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	6E-01	1E+00	1E+00	2E-01
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	1E+00	1E+00	2E-01
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	6E-01	3E-02
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	2E-01
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-7: R output showing p-values from two-tailed Fisher exact tests on the Roanoke construction cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	1E+00	2E-01	3E-02	1E+00	1E+00	1E+00	1E+00	8E-02	5E-01	1E+00	1E+00	2E-01	1E+00	1E+00	8E-02	1E+00	3E-02	1E+00	3E-02	9E-03	8E-02	8E-02	1E+00
LC_Developed_Low_Intensity	NA	NA	6E-01	1E-01	1E+00	1E+00	1E+00	1E+00	3E-01	1E+00	1E+00	1E+00	6E-01	1E+00	1E+00	3E-01	1E+00	1E-01	1E+00	1E-01	5E-02	3E-01	3E-01	1E+00
LC_Developed_Medium_Intensity	NA	NA	NA	6E-01	2E-01	2E-01	2E-01	2E-01	1E+00	1E+00	2E-01	2E-01	1E+00	6E-01	2E-01	1E+00	2E-01	6E-01	6E-01	6E-01	3E-01	1E+00	1E+00	6E-01
LC_Developed_High_Intensity	NA	NA	NA	NA	3E-02	3E-02	3E-02	3E-02	1E+00	3E-01	3E-02	3E-02	6E-01	1E-01	3E-02	1E+00	3E-02	1E+00	1E-01	1E+00	1E+00	1E+00	1E+00	1E-01
LC_Barren	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	8E-02	5E-01	1E+00	1E+00	2E-01	1E+00	1E+00	8E-02	1E+00	3E-02	1E+00	3E-02	9E-03	8E-02	8E-02	1E+00
LC_Agricultural	NA	NA	NA	NA	NA	NA	1E+00	1E+00	8E-02	5E-01	1E+00	1E+00	2E-01	1E+00	1E+00	8E-02	1E+00	3E-02	1E+00	3E-02	9E-03	8E-02	8E-02	1E+00
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	1E+00	8E-02	5E-01	1E+00	1E+00	2E-01	1E+00	1E+00	8E-02	1E+00	3E-02	1E+00	3E-02	9E-03	8E-02	8E-02	1E+00
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	8E-02	5E-01	1E+00	1E+00	2E-01	1E+00	1E+00	8E-02	1E+00	3E-02	1E+00	3E-02	9E-03	8E-02	8E-02	1E+00
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	8E-02	8E-02	1E+00	3E-01	8E-02	1E+00	8E-02	1E+00	3E-01	1E+00	6E-01	1E+00	1E+00	3E-01
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	5E-01	1E+00	1E+00	5E-01	6E-01	5E-01	3E-01	1E+00	3E-01	2E-01	6E-01	6E-01	1E+00
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	2E-01	1E+00	1E+00	8E-02	1E+00	3E-02	1E+00	3E-02	9E-03	8E-02	8E-02	1E+00
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	1E+00	1E+00	8E-02	1E+00	3E-02	1E+00	3E-02	9E-03	8E-02	8E-02	1E+00
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	2E-01	1E+00	2E-01	6E-01	6E-01	6E-01	3E-01	1E+00	1E+00	6E-01
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	3E-01	1E+00	1E-01	1E+00	1E-01	5E-02	3E-01	3E-01	1E+00
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	8E-02	1E+00	3E-02	1E+00	3E-02	9E-03	8E-02	8E-02	1E+00
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	8E-02	1E+00	3E-01	1E+00	6E-01	1E+00	1E+00	3E-01
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-02	1E+00	3E-02	9E-03	8E-02	8E-02	1E+00
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-01	1E+00	1E+00	1E+00	1E+00	1E-01
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-01	5E-02	3E-01	3E-01	1E+00
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E-01
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	6E-01	5E-02
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	3E-01
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-01
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-8: R output showing p-values from two-tailed Fisher exact tests on the Roanoke FOG, surfactants, and trash cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	1E+00	3E-05	2E-10	1E+00	1E+00	1E+00	1E+00	6E-16	1E+00	1E+00	1E+00	1E-01	1E+00	1E+00	6E-10	1E+00	5E-11	2E-03	6E-07	6E-10	7E-08	2E-08	5E-01
LC_Developed_Low_Intensity	NA	NA	3E-05	2E-10	1E+00	1E+00	1E+00	1E+00	6E-16	1E+00	1E+00	1E+00	1E-01	1E+00	1E+00	6E-10	1E+00	5E-11	2E-03	6E-07	6E-10	7E-08	2E-08	5E-01
LC_Developed_Medium_Intensity	NA	NA	NA	2E-02	3E-05	3E-05	3E-05	3E-05	2E-05	3E-05	3E-04	3E-04	1E-02	3E-04	3E-05	4E-02	3E-05	1E-02	3E-01	5E-01	4E-02	3E-01	2E-01	1E-03
LC_Developed_High_Intensity	NA	NA	NA	NA	2E-10	2E-10	2E-10	2E-10	7E-02	2E-10	3E-09	3E-09	1E-06	3E-09	2E-10	1E+00	2E-10	1E+00	5E-04	2E-01	1E+00	4E-01	5E-01	3E-08
LC_Barren	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	6E-16	1E+00	1E+00	1E+00	1E-01	1E+00	1E+00	6E-10	1E+00	5E-11	2E-03	6E-07	6E-10	7E-08	2E-08	5E-01
LC_Agricultural	NA	NA	NA	NA	NA	NA	1E+00	1E+00	6E-16	1E+00	1E+00	1E+00	1E-01	1E+00	1E+00	6E-10	1E+00	5E-11	2E-03	6E-07	6E-10	7E-08	2E-08	5E-01
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	1E+00	6E-16	1E+00	1E+00	1E+00	1E-01	1E+00	1E+00	6E-10	1E+00	5E-11	2E-03	6E-07	6E-10	7E-08	2E-08	5E-01
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	6E-16	1E+00	1E+00	1E+00	1E-01	1E+00	1E+00	6E-10	1E+00	5E-11	2E-03	6E-07	6E-10	7E-08	2E-08	5E-01
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-16	2E-14	2E-14	2E-11	2E-14	6E-16	4E-02	6E-16	1E-01	7E-08	7E-04	4E-02	3E-03	6E-03	3E-13
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E-01	1E+00	1E+00	6E-10	1E+00	5E-11	2E-03	6E-07	6E-10	7E-08	2E-08	5E-01
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	4E-01	1E+00	1E+00	1E-08	1E+00	1E-09	1E-02	7E-06	1E-08	1E-06	4E-07	1E+00
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-01	1E+00	1E+00	1E-08	1E+00	1E-09	1E-02	7E-06	1E-08	1E-06	4E-07	1E+00
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-01	1E-01	4E-06	1E-01	4E-07	2E-01	8E-04	4E-06	2E-04	7E-05	7E-01
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E-08	1E+00	1E-09	1E-02	7E-06	1E-08	1E-06	4E-07	1E+00
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-10	1E+00	5E-11	2E-03	6E-07	6E-10	7E-08	2E-08	5E-01
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-10	8E-01	1E-03	3E-01	1E+00	5E-01	6E-01	1E-07
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-11	2E-03	6E-07	6E-10	7E-08	2E-08	5E-01
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-04	1E-01	8E-01	2E-01	4E-01	1E-08
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-02	1E-03	2E-02	1E-02	5E-02
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-01	8E-01	7E-01	5E-05
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	6E-01	1E-07
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	8E-06
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-06
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-9: R output showing p-values from two-tailed Fisher exact tests on the Roanoke hazardous materials cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	3E-04	8E-04	8E-04	1E+00	1E+00	1E+00	1E+00	2E-03	2E-03	5E-01	1E-01	1E-02	2E-02	2E-01	1E-02	5E-01	7E-07	1E-02	3E-04	7E-08	3E-04	8E-04	2E-02
LC_Developed_Low_Intensity	NA	NA	1E+00	1E+00	3E-04	3E-04	3E-04	3E-04	8E-01	8E-01	1E-02	7E-02	4E-01	3E-01	3E-02	4E-01	1E-02	2E-01	4E-01	1E+00	7E-02	1E+00	1E+00	3E-01
LC_Developed_Medium_Intensity	NA	NA	NA	1E+00	8E-04	8E-04	8E-04	8E-04	1E+00	1E+00	2E-02	1E-01	6E-01	4E-01	6E-02	6E-01	2E-02	1E-01	6E-01	1E+00	4E-02	1E+00	1E+00	4E-01
LC_Developed_High_Intensity	NA	NA	NA	NA	8E-04	8E-04	8E-04	8E-04	1E+00	1E+00	2E-02	1E-01	6E-01	4E-01	6E-02	6E-01	2E-02	1E-01	6E-01	1E+00	4E-02	1E+00	1E+00	4E-01
LC_Barren	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	2E-03	2E-03	5E-01	1E-01	1E-02	2E-02	2E-01	1E-02	5E-01	7E-07	1E-02	3E-04	7E-08	3E-04	8E-04	2E-02
LC_Agricultural	NA	NA	NA	NA	NA	NA	1E+00	1E+00	2E-03	2E-03	5E-01	1E-01	1E-02	2E-02	2E-01	1E-02	5E-01	7E-07	1E-02	3E-04	7E-08	3E-04	8E-04	2E-02
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	1E+00	2E-03	2E-03	5E-01	1E-01	1E-02	2E-02	2E-01	1E-02	5E-01	7E-07	1E-02	3E-04	7E-08	3E-04	8E-04	2E-02
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	2E-03	2E-03	5E-01	1E-01	1E-02	2E-02	2E-01	1E-02	5E-01	7E-07	1E-02	3E-04	7E-08	3E-04	8E-04	2E-02
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	4E-02	2E-01	8E-01	6E-01	1E-01	8E-01	4E-02	7E-02	8E-01	8E-01	2E-02	8E-01	1E+00	6E-01
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-02	2E-01	8E-01	6E-01	1E-01	8E-01	4E-02	7E-02	8E-01	8E-01	2E-02	8E-01	1E+00	6E-01
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	7E-01	1E-01	3E-01	1E+00	1E-01	1E+00	6E-05	1E-01	1E-02	7E-06	1E-02	2E-02	3E-01
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	7E-01	1E+00	5E-01	7E-01	1E-03	5E-01	7E-02	2E-04	7E-02	1E-01	7E-01
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	3E-01	1E+00	1E-01	2E-02	1E+00	4E-01	4E-03	4E-01	6E-01	1E+00
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	1E+00	3E-01	8E-03	1E+00	3E-01	2E-03	3E-01	4E-01	1E+00
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-01	1E+00	3E-04	3E-01	3E-02	4E-05	3E-02	6E-02	5E-01
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-01	2E-02	1E+00	4E-01	4E-03	4E-01	6E-01	1E+00
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-05	1E-01	1E-02	7E-06	1E-02	2E-02	3E-01
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-02	2E-01	8E-01	2E-01	1E-01	8E-03
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-01	4E-03	4E-01	6E-01	1E+00
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	7E-02	1E+00	1E+00	3E-01
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	7E-02	4E-02	2E-03
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	3E-01
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-01
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-10: R output showing p-values from two-tailed Fisher exact tests on the Roanoke unknown cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	1E+00	3E-01	1E+00	1E+00	1E+00	1E+00	1E+00	3E-01	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
LC_Developed_Low_Intensity	NA	NA	3E-01	1E+00	1E+00	1E+00	1E+00	1E+00	3E-01	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
LC_Developed_Medium_Intensity	NA	NA	NA	3E-01	3E-01	3E-01	3E-01	3E-01	1E+00	3E-01	3E-01	3E-01	3E-01	1E+00	3E-01	3E-01	3E-01	3E-01	1E+00	3E-01	1E+00	3E-01	3E-01	1E+00
LC_Developed_High_Intensity	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	3E-01	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
LC_Barren	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	3E-01	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
LC_Agricultural	NA	NA	NA	NA	NA	NA	1E+00	1E+00	3E-01	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	1E+00	3E-01	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	3E-01	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-01	3E-01	3E-01	3E-01	1E+00	3E-01	3E-01	3E-01	3E-01	1E+00	3E-01	1E+00	3E-01	3E-01	1E+00
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-11: R output showing p-values from two-tailed Fisher exact tests on the Fairfax illicit connections cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	5E-01	5E-01	3E-01	1E-01	1E-01	1E+00	1E-01	5E-01	3E-01	3E-01	5E-01	5E-01	3E-01	7E-01	6E-02	3E-01	1E-01	1E-01	6E-03	2E-01	6E-02	6E-02	3E-01
LC_Developed_Low_Intensity	NA	NA	1E+00	5E-02	9E-03	9E-03	5E-01	9E-03	1E+00	5E-02	5E-02	1E+00	1E+00	5E-02	1E+00	4E-01	5E-02	5E-01	5E-01	7E-02	8E-01	4E-01	4E-01	5E-02
LC_Developed_Medium_Intensity	NA	NA	NA	5E-02	9E-03	9E-03	5E-01	9E-03	1E+00	5E-02	5E-02	1E+00	1E+00	5E-02	1E+00	4E-01	5E-02	5E-01	5E-01	7E-02	8E-01	4E-01	4E-01	5E-02
LC_Developed_High_Intensity	NA	NA	NA	NA	1E+00	1E+00	3E-01	1E+00	5E-02	1E+00	1E+00	5E-02	5E-02	1E+00	1E-01	2E-03	1E+00	4E-03	4E-03	8E-05	1E-02	2E-03	2E-03	1E+00
LC_Barren	NA	NA	NA	NA	NA	1E+00	1E-01	1E+00	9E-03	1E+00	1E+00	9E-03	9E-03	1E+00	2E-02	2E-04	1E+00	6E-04	6E-04	7E-06	1E-03	2E-04	2E-04	1E+00
LC_Agricultural	NA	NA	NA	NA	NA	NA	1E-01	1E+00	9E-03	1E+00	1E+00	9E-03	9E-03	1E+00	2E-02	2E-04	1E+00	6E-04	6E-04	7E-06	1E-03	2E-04	2E-04	1E+00
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	1E-01	5E-01	3E-01	3E-01	5E-01	5E-01	3E-01	7E-01	6E-02	3E-01	1E-01	1E-01	6E-03	2E-01	6E-02	6E-02	3E-01
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	9E-03	1E+00	1E+00	9E-03	9E-03	1E+00	2E-02	2E-04	1E+00	6E-04	6E-04	7E-06	1E-03	2E-04	2E-04	1E+00
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-02	5E-02	1E+00	1E+00	5E-02	1E+00	4E-01	5E-02	5E-01	5E-01	7E-02	8E-01	4E-01	4E-01	5E-02
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	5E-02	5E-02	1E+00	1E-01	2E-03	1E+00	4E-03	4E-03	8E-05	1E-02	2E-03	2E-03	1E+00
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-02	5E-02	1E+00	1E-01	2E-03	1E+00	4E-03	4E-03	8E-05	1E-02	2E-03	2E-03	1E+00
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	5E-02	1E+00	4E-01	5E-02	5E-01	5E-01	7E-02	8E-01	4E-01	4E-01	5E-02
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-02	1E+00	4E-01	5E-02	5E-01	5E-01	7E-02	8E-01	4E-01	4E-01	5E-02
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-01	2E-03	1E+00	4E-03	4E-03	8E-05	1E-02	2E-03	2E-03	1E+00
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	1E-01	4E-01	4E-01	4E-02	5E-01	2E-01	2E-01	1E-01
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-03	1E+00	1E+00	6E-01	8E-01	1E+00	1E+00	2E-03
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-03	4E-03	8E-05	1E-02	2E-03	2E-03	1E+00
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	4E-01	1E+00	1E+00	1E+00	4E-03
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-01	1E+00	1E+00	1E+00	4E-03
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	6E-01	6E-01	8E-05
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	8E-01	8E-01	1E-02
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	2E-03
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-03
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-12: R output showing p-values from two-tailed Fisher exact tests on the Fairfax illegal dumping and spills cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	1E+00	4E-02	9E-02	1E-04	1E-03	2E-02	1E-04	3E-02	7E-01	2E-02	3E-01	6E-02	1E-04	8E-01	4E-09	4E-02	1E-02	9E-02	8E-12	8E-07	6E-08	4E-04	2E-01
LC_Developed_Low_Intensity	NA	NA	2E-02	5E-02	3E-04	2E-03	3E-02	3E-04	1E-02	5E-01	3E-02	5E-01	4E-02	3E-04	1E+00	1E-09	6E-02	5E-03	5E-02	2E-12	3E-07	2E-08	2E-04	2E-01
LC_Developed_Medium_Intensity	NA	NA	NA	9E-01	6E-09	9E-08	4E-06	6E-09	1E+00	1E-01	4E-06	1E-03	1E+00	6E-09	1E-02	2E-04	2E-05	7E-01	9E-01	2E-06	6E-03	1E-03	2E-01	5E-01
LC_Developed_High_Intensity	NA	NA	NA	NA	3E-08	5E-07	2E-05	3E-08	7E-01	3E-01	2E-05	4E-03	1E+00	3E-08	3E-02	5E-05	8E-05	5E-01	1E+00	4E-07	2E-03	3E-04	9E-02	7E-01
LC_Barren	NA	NA	NA	NA	NA	1E+00	2E-01	1E+00	2E-09	1E-05	2E-01	6E-03	1E-08	1E+00	7E-04	3E-20	1E-01	4E-10	3E-08	8E-24	6E-17	2E-18	1E-12	5E-07
LC_Agricultural	NA	NA	NA	NA	NA	NA	6E-01	1E+00	4E-08	1E-04	6E-01	3E-02	2E-07	1E+00	5E-03	1E-18	4E-01	7E-09	5E-07	3E-22	2E-15	5E-17	2E-11	6E-06
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	2E-01	2E-06	2E-03	1E+00	2E-01	9E-06	2E-01	5E-02	3E-16	1E+00	4E-07	2E-05	1E-19	3E-13	1E-14	2E-09	2E-04
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	2E-09	1E-05	2E-01	6E-03	1E-08	1E+00	7E-04	3E-20	1E-01	4E-10	3E-08	8E-24	6E-17	2E-18	1E-12	5E-07
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-01	2E-06	8E-04	9E-01	2E-09	8E-03	4E-04	9E-06	9E-01	7E-01	5E-06	9E-03	2E-03	2E-01	4E-01
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-03	1E-01	2E-01	1E-05	4E-01	1E-07	7E-03	5E-02	3E-01	3E-10	1E-05	1E-06	3E-03	6E-01
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	9E-06	2E-01	5E-02	3E-16	1E+00	4E-07	2E-05	1E-19	3E-13	1E-14	2E-09	2E-04
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-03	6E-03	6E-01	4E-12	4E-01	2E-04	4E-03	4E-15	2E-09	1E-10	4E-06	2E-02
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-08	2E-02	9E-05	4E-05	6E-01	1E+00	1E-06	3E-03	6E-04	1E-01	6E-01
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	7E-04	3E-20	1E-01	4E-10	3E-08	8E-24	6E-17	2E-18	1E-12	5E-07
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-10	1E-01	3E-03	3E-02	5E-13	9E-08	6E-09	8E-05	1E-01
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-15	1E-03	5E-05	4E-01	4E-01	7E-01	2E-02	4E-06
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-06	8E-05	1E-18	2E-12	9E-14	1E-08	6E-04
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	2E-05	2E-02	6E-03	4E-01	2E-01
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-07	2E-03	3E-04	9E-02	7E-01
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-02	2E-01	1E-03	2E-08
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	7E-01	2E-01	3E-04
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	8E-02	4E-05
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-02
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-13: R output showing p-values from two-tailed Fisher exact tests on the Fairfax fecal bacteria cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	1E+00	1E+00	1E+00	5E-01	5E-01	1E+00	5E-01	1E+00	5E-01	1E+00	1E+00	6E-01	5E-01	1E+00	1E+00	5E-01	1E+00	1E+00	4E-02	1E-01	1E-01	1E+00	1E+00
LC_Developed_Low_Intensity	NA	NA	6E-01	6E-01	2E-01	2E-01	6E-01	2E-01	1E+00	2E-01	6E-01	6E-01	1E+00	2E-01	6E-01	1E+00	2E-01	1E+00	1E+00	1E-01	3E-01	3E-01	6E-01	1E+00
LC_Developed_Medium_Intensity	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	3E-01	1E+00	1E+00	6E-01	1E+00	1E+00	1E+00	1E-02	4E-02	4E-02	1E+00	1E+00
LC_Developed_High_Intensity	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	3E-01	1E+00	1E+00	6E-01	1E+00	1E+00	1E+00	1E-02	4E-02	4E-02	1E+00	1E+00
LC_Barren	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	5E-01	1E+00	1E+00	1E+00	8E-02	1E+00	1E+00	2E-01	1E+00	5E-01	5E-01	1E-03	7E-03	7E-03	1E+00	5E-01
LC_Agricultural	NA	NA	NA	NA	NA	NA	1E+00	1E+00	5E-01	1E+00	1E+00	1E+00	8E-02	1E+00	1E+00	2E-01	1E+00	5E-01	5E-01	1E-03	7E-03	7E-03	1E+00	5E-01
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	3E-01	1E+00	1E+00	6E-01	1E+00	1E+00	1E+00	1E-02	4E-02	4E-02	1E+00	1E+00
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	1E+00	1E+00	1E+00	8E-02	1E+00	1E+00	2E-01	1E+00	5E-01	5E-01	1E-03	7E-03	7E-03	1E+00	5E-01
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	1E+00	1E+00	6E-01	5E-01	1E+00	1E+00	5E-01	1E+00	1E+00	4E-02	1E-01	1E-01	1E+00	1E+00
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	8E-02	1E+00	1E+00	2E-01	1E+00	5E-01	5E-01	1E-03	7E-03	7E-03	1E+00	5E-01
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	3E-01	1E+00	1E+00	6E-01	1E+00	1E+00	1E+00	1E-02	4E-02	4E-02	1E+00	1E+00
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-01	1E+00	1E+00	6E-01	1E+00	1E+00	1E+00	1E-02	4E-02	4E-02	1E+00	1E+00
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	8E-02	3E-01	1E+00	8E-02	6E-01	6E-01	3E-01	6E-01	6E-01	3E-01	6E-01
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	2E-01	1E+00	5E-01	5E-01	1E-03	7E-03	7E-03	1E+00	5E-01
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	1E+00	1E+00	1E+00	1E-02	4E-02	4E-02	1E+00	1E+00
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	1E+00	1E+00	1E-01	3E-01	3E-01	6E-01	1E+00
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	5E-01	1E-03	7E-03	7E-03	1E+00	5E-01
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	4E-02	1E-01	1E-01	1E+00	1E+00
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-02	1E-01	1E-01	1E+00	1E+00
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E-02	4E-02
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	4E-02	1E-01
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-02	1E-01
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-14: R output showing p-values from two-tailed Fisher exact tests on the Fairfax chloride cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	3E-01	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	6E-01	6E-01	1E+00	6E-01	1E+00	1E+00	1E+00	1E+00	3E-01	1E-01	1E-01	6E-01	1E+00
LC_Developed_Low_Intensity	NA	NA	3E-01	3E-01	8E-02	8E-02	3E-01	8E-02	3E-01	8E-02	3E-01	1E+00	1E+00	8E-02	1E+00	3E-01	3E-01	3E-01	8E-02	1E+00	1E+00	1E+00	1E+00	8E-02
LC_Developed_Medium_Intensity	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	6E-01	6E-01	1E+00	6E-01	1E+00	1E+00	1E+00	1E+00	3E-01	1E-01	1E-01	6E-01	1E+00
LC_Developed_High_Intensity	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	6E-01	6E-01	1E+00	6E-01	1E+00	1E+00	1E+00	1E+00	3E-01	1E-01	1E-01	6E-01	1E+00
LC_Barren	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	2E-01	2E-01	1E+00	2E-01	1E+00	1E+00	1E+00	1E+00	8E-02	3E-02	3E-02	2E-01	1E+00
LC_Agricultural	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	2E-01	2E-01	1E+00	2E-01	1E+00	1E+00	1E+00	1E+00	8E-02	3E-02	3E-02	2E-01	1E+00
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	6E-01	6E-01	1E+00	6E-01	1E+00	1E+00	1E+00	1E+00	3E-01	1E-01	1E-01	6E-01	1E+00
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	2E-01	2E-01	1E+00	2E-01	1E+00	1E+00	1E+00	1E+00	8E-02	3E-02	3E-02	2E-01	1E+00
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	6E-01	6E-01	1E+00	6E-01	1E+00	1E+00	1E+00	1E+00	3E-01	1E-01	1E-01	6E-01	1E+00
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	2E-01	2E-01	1E+00	2E-01	1E+00	1E+00	1E+00	1E+00	8E-02	3E-02	3E-02	2E-01	1E+00
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	6E-01	1E+00	6E-01	1E+00	1E+00	1E+00	1E+00	3E-01	1E-01	1E-01	6E-01	1E+00
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	2E-01	1E+00	6E-01	6E-01	6E-01	2E-01	1E+00	6E-01	6E-01	1E+00	2E-01
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	1E+00	6E-01	6E-01	6E-01	2E-01	1E+00	6E-01	6E-01	1E+00	2E-01
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	1E+00	1E+00	1E+00	1E+00	8E-02	3E-02	3E-02	2E-01	1E+00
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	6E-01	6E-01	2E-01	1E+00	6E-01	6E-01	1E+00	2E-01
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	3E-01	1E-01	1E-01	6E-01	1E+00
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	3E-01	1E-01	1E-01	6E-01	1E+00
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	3E-01	1E-01	1E-01	6E-01	1E+00
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	8E-02	3E-02	3E-02	2E-01	1E+00
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	8E-02
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	6E-01	3E-02
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	3E-02
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-15: R output showing p-values from two-tailed Fisher exact tests on the Fairfax construction cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	1E+00	7E-01	4E-02	4E-02	4E-02	4E-02	4E-02	2E-01	2E-01	2E-01	4E-02	2E-01	4E-02	7E-01	7E-01	4E-02	4E-01	1E+00	7E-01	1E+00	7E-01	1E+00	1E+00
LC_Developed_Low_Intensity	NA	NA	1E+00	9E-02	9E-02	9E-02	9E-02	9E-02	3E-01	3E-01	3E-01	9E-02	1E-01	9E-02	1E+00	4E-01	9E-02	6E-01	1E+00	4E-01	1E+00	4E-01	7E-01	1E+00
LC_Developed_Medium_Intensity	NA	NA	NA	2E-01	2E-01	2E-01	2E-01	2E-01	6E-01	6E-01	6E-01	2E-01	4E-02	2E-01	1E+00	2E-01	2E-01	1E+00	1E+00	2E-01	7E-01	2E-01	4E-01	1E+00
LC_Developed_High_Intensity	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	3E-04	1E+00	2E-01	5E-03	1E+00	5E-01	9E-02	5E-03	4E-02	5E-03	1E-02	9E-02
LC_Barren	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	3E-04	1E+00	2E-01	5E-03	1E+00	5E-01	9E-02	5E-03	4E-02	5E-03	1E-02	9E-02
LC_Agricultural	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	1E+00	3E-04	1E+00	2E-01	5E-03	1E+00	5E-01	9E-02	5E-03	4E-02	5E-03	1E-02	9E-02
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	1E+00	3E-04	1E+00	2E-01	5E-03	1E+00	5E-01	9E-02	5E-03	4E-02	5E-03	1E-02	9E-02
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00	3E-04	1E+00	2E-01	5E-03	1E+00	5E-01	9E-02	5E-03	4E-02	5E-03	1E-02	9E-02
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	3E-03	1E+00	6E-01	3E-02	1E+00	1E+00	3E-01	3E-02	2E-01	3E-02	7E-02	3E-01
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	3E-03	1E+00	6E-01	3E-02	1E+00	1E+00	3E-01	3E-02	2E-01	3E-02	7E-02	3E-01
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	3E-03	1E+00	6E-01	3E-02	1E+00	1E+00	3E-01	3E-02	2E-01	3E-02	7E-02	3E-01
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-04	1E+00	2E-01	5E-03	1E+00	5E-01	9E-02	5E-03	4E-02	5E-03	1E-02	9E-02
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-04	4E-02	7E-01	3E-04	1E-02	1E-01	7E-01	2E-01	7E-01	4E-01	1E-01
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	5E-03	1E+00	5E-01	9E-02	5E-03	4E-02	5E-03	1E-02	9E-02
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	2E-01	1E+00	1E+00	2E-01	7E-01	2E-01	4E-01	1E+00
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-03	9E-02	4E-01	1E+00	7E-01	1E+00	1E+00	4E-01
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	9E-02	5E-03	4E-02	5E-03	1E-02	9E-02
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	9E-02	4E-01	9E-02	2E-01	6E-01
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-01	1E+00	4E-01	7E-01	1E+00
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	7E-01	1E+00	1E+00	4E-01
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	7E-01	1E+00	1E+00
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	4E-01
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	7E-01
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-16: R output showing p-values from two-tailed Fisher exact tests on the Fairfax FOG, surfactants, and trash cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	7E-01	1E-01	4E-02	1E-01	1E-01	7E-01	1E-01	1E-02	3E-01	1E-01	1E+00	1E+00	1E-01	1E+00	5E-06	7E-01	7E-02	7E-02	2E-05	9E-04	9E-04	1E-02	2E-01
LC_Developed_Low_Intensity	NA	NA	2E-02	5E-03	5E-01	5E-01	1E+00	5E-01	9E-04	8E-02	5E-01	1E+00	4E-01	5E-01	7E-01	1E-07	1E+00	1E-02	1E-02	5E-07	5E-05	5E-05	9E-04	4E-02
LC_Developed_Medium_Intensity	NA	NA	NA	8E-01	8E-04	8E-04	2E-02	8E-04	4E-01	8E-01	8E-04	6E-02	2E-01	8E-04	1E-01	4E-03	2E-02	1E+00	1E+00	9E-03	1E-01	1E-01	4E-01	1E+00
LC_Developed_High_Intensity	NA	NA	NA	NA	1E-04	1E-04	5E-03	1E-04	8E-01	4E-01	1E-04	2E-02	8E-02	1E-04	4E-02	2E-02	5E-03	1E+00	1E+00	4E-02	3E-01	3E-01	8E-01	6E-01
LC_Barren	NA	NA	NA	NA	NA	1E+00	5E-01	1E+00	2E-05	5E-03	1E+00	2E-01	5E-02	1E+00	1E-01	8E-10	5E-01	3E-04	3E-04	4E-09	6E-07	6E-07	2E-05	2E-03
LC_Agricultural	NA	NA	NA	NA	NA	NA	5E-01	1E+00	2E-05	5E-03	1E+00	2E-01	5E-02	1E+00	1E-01	8E-10	5E-01	3E-04	3E-04	4E-09	6E-07	6E-07	2E-05	2E-03
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	5E-01	9E-04	8E-02	5E-01	1E+00	4E-01	5E-01	7E-01	1E-07	1E+00	1E-02	1E-02	5E-07	5E-05	5E-05	9E-04	4E-02
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	2E-05	5E-03	1E+00	2E-01	5E-02	1E+00	1E-01	8E-10	5E-01	3E-04	3E-04	4E-09	6E-07	6E-07	2E-05	2E-03
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	2E-05	3E-03	3E-02	2E-05	1E-02	6E-02	9E-04	6E-01	6E-01	1E-01	6E-01	6E-01	1E+00	3E-01
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-03	2E-01	5E-01	5E-03	3E-01	7E-04	8E-02	6E-01	6E-01	2E-03	4E-02	4E-02	2E-01	1E+00
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	5E-02	1E+00	1E-01	8E-10	5E-01	3E-04	3E-04	4E-09	6E-07	6E-07	2E-05	2E-03
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	7E-01	2E-01	1E+00	1E-06	1E+00	3E-02	3E-02	3E-06	3E-04	3E-04	3E-03	1E-01
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-02	1E+00	2E-05	4E-01	1E-01	1E-01	6E-05	3E-03	3E-03	3E-02	4E-01
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-01	8E-10	5E-01	3E-04	3E-04	4E-09	6E-07	6E-07	2E-05	2E-03
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-06	7E-01	7E-02	7E-02	2E-05	9E-04	9E-04	1E-02	2E-01
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-07	9E-03	9E-03	1E+00	3E-01	3E-01	6E-02	2E-03
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-02	1E-02	5E-07	5E-05	5E-05	9E-04	4E-02
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	2E-02	2E-01	2E-01	6E-01	8E-01
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-02	2E-01	2E-01	6E-01	8E-01
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4E-01	4E-01	1E-01	4E-03
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	6E-01	7E-02
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	7E-02
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-01
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-17: R output showing p-values from two-tailed Fisher exact tests on the Fairfax hazardous materials cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	1E+00	2E-02	2E-01	5E-02	2E-01	2E-01	5E-02	6E-02	5E-01	2E-01	8E-01	8E-01	2E-01	1E+00	3E-06	2E-01	9E-04	9E-03	3E-07	9E-03	9E-04	2E-03	1E+00
LC_Developed_Low_Intensity	NA	NA	7E-03	1E-01	1E-01	4E-01	4E-01	1E-01	2E-02	3E-01	4E-01	5E-01	5E-01	4E-01	7E-01	7E-07	4E-01	3E-04	3E-03	5E-08	3E-03	3E-04	7E-04	1E+00
LC_Developed_Medium_Intensity	NA	NA	NA	3E-01	1E-05	1E-04	1E-04	1E-05	8E-01	1E-01	1E-04	7E-02	7E-02	1E-04	4E-02	3E-02	1E-04	5E-01	1E+00	8E-03	1E+00	5E-01	6E-01	7E-03
LC_Developed_High_Intensity	NA	NA	NA	NA	9E-04	6E-03	6E-03	9E-04	6E-01	8E-01	6E-03	6E-01	6E-01	6E-03	4E-01	8E-04	6E-03	5E-02	2E-01	1E-04	2E-01	5E-02	9E-02	1E-01
LC_Barren	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	7E-05	5E-03	1E+00	1E-02	1E-02	1E+00	2E-02	8E-11	1E+00	1E-07	4E-06	4E-12	4E-06	1E-07	4E-07	1E-01
LC_Agricultural	NA	NA	NA	NA	NA	NA	1E+00	1E+00	6E-04	3E-02	1E+00	6E-02	6E-02	1E+00	1E-01	2E-09	1E+00	2E-06	4E-05	9E-11	4E-05	2E-06	6E-06	4E-01
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	1E+00	6E-04	3E-02	1E+00	6E-02	6E-02	1E+00	1E-01	2E-09	1E+00	2E-06	4E-05	9E-11	4E-05	2E-06	6E-06	4E-01
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	7E-05	5E-03	1E+00	1E-02	1E-02	1E+00	2E-02	8E-11	1E+00	1E-07	4E-06	4E-12	4E-06	1E-07	4E-07	1E-01
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-01	6E-04	2E-01	2E-01	6E-04	1E-01	9E-03	6E-04	2E-01	6E-01	2E-03	6E-01	2E-01	3E-01	2E-02
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3E-02	1E+00	1E+00	3E-02	8E-01	1E-04	3E-02	1E-02	8E-02	1E-05	8E-02	1E-02	3E-02	3E-01
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-02	6E-02	1E+00	1E-01	2E-09	1E+00	2E-06	4E-05	9E-11	4E-05	2E-06	6E-06	4E-01
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	6E-02	1E+00	4E-05	6E-02	6E-03	4E-02	4E-06	4E-02	6E-03	1E-02	5E-01
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-02	1E+00	4E-05	6E-02	6E-03	4E-02	4E-06	4E-02	6E-03	1E-02	5E-01
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-01	2E-09	1E+00	2E-06	4E-05	9E-11	4E-05	2E-06	6E-06	4E-01
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E-05	1E-01	3E-03	2E-02	1E-06	2E-02	3E-03	5E-03	7E-01
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-09	2E-01	6E-02	8E-01	6E-02	2E-01	1E-01	7E-07
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-06	4E-05	9E-11	4E-05	2E-06	6E-06	4E-01
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	8E-02	6E-01	1E+00	1E+00	3E-04
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-02	1E+00	6E-01	8E-01	3E-03
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-02	8E-02	5E-02	5E-08
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6E-01	8E-01	3E-03
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	3E-04
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	7E-04
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-18: R output showing p-values from two-tailed Fisher exact tests on the Fairfax unknown cluster.

	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Commercial	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Population_Density	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Construction_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	Dog_License_Density
LC_Developed_Open_Space	NA	5E-01	5E-01	1E+00	1E+00	1E+00	5E-01	1E+00	5E-01	1E+00	1E+00	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	2E-01	2E-01	2E-01	2E-01	5E-01
LC_Developed_Low_Intensity	NA	NA	1E+00	5E-01	5E-01	5E-01	1E+00	5E-01	1E+00	5E-01	5E-01	1E+00	1E+00	5E-01	5E-01	1E+00	1E+00	1E+00	5E-01	1E+00	1E+00	1E+00	1E+00	1E+00
LC_Developed_Medium_Intensity	NA	NA	NA	5E-01	5E-01	5E-01	1E+00	5E-01	1E+00	5E-01	5E-01	1E+00	1E+00	5E-01	5E-01	1E+00	1E+00	1E+00	5E-01	1E+00	1E+00	1E+00	1E+00	1E+00
LC_Developed_High_Intensity	NA	NA	NA	NA	1E+00	1E+00	5E-01	1E+00	5E-01	1E+00	1E+00	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	2E-01	2E-01	2E-01	2E-01	5E-01
LC_Barren	NA	NA	NA	NA	NA	1E+00	5E-01	1E+00	5E-01	1E+00	1E+00	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	2E-01	2E-01	2E-01	2E-01	5E-01
LC_Agricultural	NA	NA	NA	NA	NA	NA	5E-01	1E+00	5E-01	1E+00	1E+00	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	2E-01	2E-01	2E-01	2E-01	5E-01
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	5E-01	1E+00	5E-01	5E-01	1E+00	1E+00	5E-01	5E-01	1E+00	1E+00	1E+00	5E-01	1E+00	1E+00	1E+00	1E+00	1E+00
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	1E+00	1E+00	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	2E-01	2E-01	2E-01	2E-01	5E-01
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	5E-01	1E+00	1E+00	5E-01	5E-01	1E+00	1E+00	1E+00	5E-01	1E+00	1E+00	1E+00	1E+00	1E+00
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	2E-01	2E-01	2E-01	2E-01	5E-01
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	2E-01	1E+00	1E+00	5E-01	1E+00	2E-01	1E+00	2E-01	2E-01	2E-01	2E-01	5E-01
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	1E+00	1E+00	1E+00	1E+00	5E-01	1E+00	5E-01	5E-01	5E-01	5E-01	1E+00
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	2E-01	1E+00	5E-01	1E+00	2E-01	1E+00	1E+00	1E+00	1E+00	1E+00
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	5E-01	1E+00	2E-01	1E+00	2E-01	2E-01	2E-01	2E-01	5E-01
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	1E+00	2E-01	1E+00	2E-01	2E-01	2E-01	2E-01	5E-01
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	5E-01	1E+00	1E+00	1E+00	1E+00	1E+00
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5E-01	1E+00	5E-01	5E-01	5E-01	5E-01	1E+00
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	1E+00	1E+00	1E+00	1E+00	1E+00
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2E-01	2E-01	2E-01	2E-01	5E-01
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00	1E+00
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00	1E+00
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00	1E+00
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1E+00
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-19: R output showing p-values from one-tailed Fisher exact tests on the Roanoke illicit connections cluster.

	LC_Developed_Medium_Intensity	Development_Age	LC_Developed_High_Intensity	LU_Industrial	LU_Residential	Generating_Site_Density	Infrastructure_Access_Density	Percent_Impervious	Drainage_Density	Aging_Sanitary_Infrastructure	LU_Commercial	Construction_Site_Density	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Institutional	LU_Recreation_and_Open_Space	Population_Density	Outfall_Density	Aging_Septic_System_Density	Dog_License_Density
LC_Developed_Open_Space	3.5E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	3.5E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Barren	3.5E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	3.5E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	3.5E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	3.5E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	3.5E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	3.5E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	3.5E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	3.5E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	3.5E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	3.5E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-20: R output showing p-values from one-tailed Fisher exact tests on the Roanoke illegal dumping and spills cluster.

	LU_Commercial	Development_Age	Generating_Site_Density	LC_Developed_High_Intensity	Infrastructure_Access_Density	Drainage_Density	Aging_Sanitary_Infrastructure	Percent_Impervious	LC_Developed_Medium_Intensity	Construction_Site_Density	LC_Developed_Low_Intensity	LU_Residential	Dog_License_Density	Population_Density	LU_Industrial	LU_Recreation_and_Open_Space	LU_Institutional	Outfall_Density	Aging_Septic_System_Density	LC_Developed_Open_Space	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural
LC_Developed_Open_Space	1.4E-21	4.9E-21	1.7E-20	5.0E-15	1.4E-14	1.1E-13	1.1E-13	3.0E-13	1.5E-11	2.6E-07	7.3E-05	7.3E-05	7.3E-04	1.6E-03	3.3E-03	1.4E-02	NA	NA	NA	NA	NA	NA	NA	NA
LC_Barren	1.4E-21	4.9E-21	1.7E-20	5.0E-15	1.4E-14	1.1E-13	1.1E-13	3.0E-13	1.5E-11	2.6E-07	7.3E-05	7.3E-05	7.3E-04	1.6E-03	3.3E-03	1.4E-02	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	1.4E-21	4.9E-21	1.7E-20	5.0E-15	1.4E-14	1.1E-13	1.1E-13	3.0E-13	1.5E-11	2.6E-07	7.3E-05	7.3E-05	7.3E-04	1.6E-03	3.3E-03	1.4E-02	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	1.4E-21	4.9E-21	1.7E-20	5.0E-15	1.4E-14	1.1E-13	1.1E-13	3.0E-13	1.5E-11	2.6E-07	7.3E-05	7.3E-05	7.3E-04	1.6E-03	3.3E-03	1.4E-02	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	1.4E-21	4.9E-21	1.7E-20	5.0E-15	1.4E-14	1.1E-13	1.1E-13	3.0E-13	1.5E-11	2.6E-07	7.3E-05	7.3E-05	7.3E-04	1.6E-03	3.3E-03	1.4E-02	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	1.0E-18	3.5E-18	1.2E-17	1.8E-12	4.7E-12	3.2E-11	3.2E-11	8.2E-11	3.0E-09	2.1E-05	2.5E-03	2.5E-03	1.6E-02	2.8E-02	4.9E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	1.4E-17	4.6E-17	1.5E-16	1.6E-11	4.2E-11	2.7E-10	2.7E-10	6.7E-10	2.1E-08	9.3E-05	7.8E-03	7.8E-03	4.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	1.4E-16	4.6E-16	1.4E-15	1.2E-10	2.9E-10	1.7E-09	1.7E-09	4.1E-09	1.2E-07	3.3E-04	1.9E-02	1.9E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	8.1E-15	2.5E-14	7.4E-14	3.2E-09	7.6E-09	4.0E-08	4.0E-08	9.0E-08	1.9E-06	2.4E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	2.6E-13	7.7E-13	2.2E-12	5.2E-08	1.2E-07	5.4E-07	5.4E-07	1.1E-06	1.9E-05	1.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	1.3E-12	3.6E-12	1.0E-11	1.8E-07	3.8E-07	1.7E-06	1.7E-06	3.5E-06	5.2E-05	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	5.6E-12	1.6E-11	4.2E-11	5.6E-07	1.2E-06	4.9E-06	4.9E-06	9.7E-06	1.3E-04	3.4E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	3.0E-10	7.8E-10	2.0E-09	1.1E-05	2.1E-05	7.5E-05	7.5E-05	1.4E-04	1.3E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	3.0E-10	7.8E-10	2.0E-09	1.1E-05	2.1E-05	7.5E-05	7.5E-05	1.4E-04	1.3E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	4.4E-07	9.5E-07	2.0E-06	1.7E-03	2.7E-03	6.7E-03	6.7E-03	1.0E-02	4.6E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	9.1E-04	1.5E-03	2.5E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	6.3E-03	9.8E-03	1.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	9.7E-03	1.5E-02	2.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	9.7E-03	1.5E-02	2.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	2.1E-02	3.1E-02	4.3E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	3.0E-02	4.3E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-21: R output showing p-values from one-tailed Fisher exact tests on the Roanoke fecal bacteria cluster.

	LC_Developed_Medium_Intensity	LU_Commercial	Infrastructure_Access_Density	Generating_Site_Density	Development_Age	Construction_Site_Density	Percent_Impervious	Drainage_Density	Aging_Sanitary_Infrastructure	LC_Developed_High_Intensity	LU_Residential	LC_Developed_Low_Intensity	LU_Institutional	LU_Recreation_and_Open_Space	LC_Developed_Open_Space	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Industrial	Population_Density	Outfall_Density	Aging_Septic_System_Density	Dog_License_Density
LC_Developed_Open_Space	4.5E-03	1.5E-02	1.5E-02	1.5E-02	1.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Barren	4.5E-03	1.5E-02	1.5E-02	1.5E-02	1.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	4.5E-03	1.5E-02	1.5E-02	1.5E-02	1.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	4.5E-03	1.5E-02	1.5E-02	1.5E-02	1.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	4.5E-03	1.5E-02	1.5E-02	1.5E-02	1.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	4.5E-03	1.5E-02	1.5E-02	1.5E-02	1.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	4.5E-03	1.5E-02	1.5E-02	1.5E-02	1.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	4.5E-03	1.5E-02	1.5E-02	1.5E-02	1.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	4.5E-03	1.5E-02	1.5E-02	1.5E-02	1.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	4.5E-03	1.5E-02	1.5E-02	1.5E-02	1.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	2.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	2.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	2.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-22: R output showing p-values from one-tailed Fisher exact tests on the Roanoke construction cluster.

	Development_Age	LC_Developed_High_Intensity	Generating_Site_Density	Percent_Impervious	LU_Commercial	Infrastructure_Access_Density	Drainage_Density	Aging_Sanitary_Infrastructure	LC_Developed_Medium_Intensity	LU_Residential	LU_Industrial	LC_Developed_Low_Intensity	Population_Density	Construction_Site_Density	Dog_License_Density	LC_Developed_Open_Space	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Institutional	LU_Recreation_and_Open_Space	Outfall_Density	Aging_Septic_System_Density
LC_Developed_Open_Space	4.5E-03	1.5E-02	1.5E-02	1.5E-02	4.1E-02	4.1E-02	4.1E-02	4.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Barren	4.5E-03	1.5E-02	1.5E-02	1.5E-02	4.1E-02	4.1E-02	4.1E-02	4.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	4.5E-03	1.5E-02	1.5E-02	1.5E-02	4.1E-02	4.1E-02	4.1E-02	4.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	4.5E-03	1.5E-02	1.5E-02	1.5E-02	4.1E-02	4.1E-02	4.1E-02	4.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	4.5E-03	1.5E-02	1.5E-02	1.5E-02	4.1E-02	4.1E-02	4.1E-02	4.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	4.5E-03	1.5E-02	1.5E-02	1.5E-02	4.1E-02	4.1E-02	4.1E-02	4.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	4.5E-03	1.5E-02	1.5E-02	1.5E-02	4.1E-02	4.1E-02	4.1E-02	4.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	4.5E-03	1.5E-02	1.5E-02	1.5E-02	4.1E-02	4.1E-02	4.1E-02	4.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	4.5E-03	1.5E-02	1.5E-02	1.5E-02	4.1E-02	4.1E-02	4.1E-02	4.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	2.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	2.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	2.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	2.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-23: R output showing p-values from one-tailed Fisher exact tests on the Roanoke FOG, surfactants, and trash cluster.

	LU_Commercial	Generating_Site_Density	LC_Developed_High_Intensity	Infrastructure_Access_Density	Development_Age	Aging_Sanitary_Infrastructure	Drainage_Density	Percent_Impervious	LC_Developed_Medium_Intensity	Construction_Site_Density	LU_Residential	Dog_License_Density	LU_Institutional	LU_Recreation_and_Open_Space	Population_Density	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Industrial	Outfall_Density	Aging_Septic_System_Density
LC_Developed_Open_Space	3.0E-16	2.3E-11	8.8E-11	3.2E-10	3.2E-10	1.1E-08	3.5E-08	2.9E-07	1.5E-05	1.2E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	3.0E-16	2.3E-11	8.8E-11	3.2E-10	3.2E-10	1.1E-08	3.5E-08	2.9E-07	1.5E-05	1.2E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Barren	3.0E-16	2.3E-11	8.8E-11	3.2E-10	3.2E-10	1.1E-08	3.5E-08	2.9E-07	1.5E-05	1.2E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	3.0E-16	2.3E-11	8.8E-11	3.2E-10	3.2E-10	1.1E-08	3.5E-08	2.9E-07	1.5E-05	1.2E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	3.0E-16	2.3E-11	8.8E-11	3.2E-10	3.2E-10	1.1E-08	3.5E-08	2.9E-07	1.5E-05	1.2E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	3.0E-16	2.3E-11	8.8E-11	3.2E-10	3.2E-10	1.1E-08	3.5E-08	2.9E-07	1.5E-05	1.2E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	3.0E-16	2.3E-11	8.8E-11	3.2E-10	3.2E-10	1.1E-08	3.5E-08	2.9E-07	1.5E-05	1.2E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	3.0E-16	2.3E-11	8.8E-11	3.2E-10	3.2E-10	1.1E-08	3.5E-08	2.9E-07	1.5E-05	1.2E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	3.0E-16	2.3E-11	8.8E-11	3.2E-10	3.2E-10	1.1E-08	3.5E-08	2.9E-07	1.5E-05	1.2E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	8.9E-15	4.8E-10	1.7E-09	5.8E-09	5.8E-09	1.8E-07	5.0E-07	3.7E-06	1.4E-04	7.1E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	8.9E-15	4.8E-10	1.7E-09	5.8E-09	5.8E-09	1.8E-07	5.0E-07	3.7E-06	1.4E-04	7.1E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	8.9E-15	4.8E-10	1.7E-09	5.8E-09	5.8E-09	1.8E-07	5.0E-07	3.7E-06	1.4E-04	7.1E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	1.3E-13	5.1E-09	1.7E-08	5.5E-08	5.5E-08	1.4E-06	3.8E-06	2.5E-05	7.1E-04	2.4E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	1.1E-11	2.1E-07	6.3E-07	1.8E-06	1.8E-06	3.3E-05	7.9E-05	4.1E-04	7.1E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	3.4E-08	1.1E-04	2.6E-04	5.9E-04	5.9E-04	5.0E-03	9.2E-03	2.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	1.0E-05	6.1E-03	1.1E-02	2.0E-02	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	3.5E-04	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	1.6E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	3.2E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	3.4E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-24: R output showing p-values from one-tailed Fisher exact tests on the Roanoke hazardous materials cluster.

	Development_Age	Generating_Site_Density	LC_Developed_Low_Intensity	Percent_Impervious	Drainage_Density	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	Aging_Sanitary_Infrastructure	LU_Commercial	LU_Industrial	LU_Residential	Infrastructure_Access_Density	Construction_Site_Density	Population_Density	Dog_License_Density	LU_Recreation_and_Open_Space	Outfall_Density	LU_Institutional	Aging_Septic_System_Density	LC_Developed_Open_Space	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural
LC_Developed_Open_Space	3.3E-08	3.6E-07	1.7E-04	1.7E-04	1.7E-04	4.1E-04	4.1E-04	4.1E-04	9.9E-04	9.9E-04	5.3E-03	5.3E-03	5.3E-03	1.2E-02	1.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Barren	3.3E-08	3.6E-07	1.7E-04	1.7E-04	1.7E-04	4.1E-04	4.1E-04	4.1E-04	9.9E-04	9.9E-04	5.3E-03	5.3E-03	5.3E-03	1.2E-02	1.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	3.3E-08	3.6E-07	1.7E-04	1.7E-04	1.7E-04	4.1E-04	4.1E-04	4.1E-04	9.9E-04	9.9E-04	5.3E-03	5.3E-03	5.3E-03	1.2E-02	1.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	3.3E-08	3.6E-07	1.7E-04	1.7E-04	1.7E-04	4.1E-04	4.1E-04	4.1E-04	9.9E-04	9.9E-04	5.3E-03	5.3E-03	5.3E-03	1.2E-02	1.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	3.3E-08	3.6E-07	1.7E-04	1.7E-04	1.7E-04	4.1E-04	4.1E-04	4.1E-04	9.9E-04	9.9E-04	5.3E-03	5.3E-03	5.3E-03	1.2E-02	1.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	3.7E-06	3.0E-05	5.3E-03	5.3E-03	5.3E-03	1.1E-02	1.1E-02	1.1E-02	2.1E-02	2.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	3.7E-06	3.0E-05	5.3E-03	5.3E-03	5.3E-03	1.1E-02	1.1E-02	1.1E-02	2.1E-02	2.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	2.0E-05	1.4E-04	1.6E-02	1.6E-02	1.6E-02	2.9E-02	2.9E-02	2.9E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	8.1E-05	5.1E-04	3.7E-02	3.7E-02	3.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	8.1E-04	3.9E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	8.1E-04	3.9E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	2.1E-03	9.0E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	2.1E-03	9.0E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	2.1E-03	9.0E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	1.0E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	1.0E-02	3.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	3.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	3.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	3.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-25: R output showing p-values from one-tailed Fisher exact tests on the Roanoke unknown cluster.

	LC_Developed_Medium_Intensity	LU_Commercial	Population_Density	Construction_Site_Density	Development_Age	Dog_License_Density	LC_Developed_Open_Space	LC_Developed_Low_Intensity	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Industrial	LU_Institutional	LU_Recreation_and_Open_Space	LU_Residential	Outfall_Density	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	Percent_Impervious	Drainage_Density	Aging_Sanitary_Infrastructure
LC_Developed_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Barren	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-26: R output showing p-values from one-tailed Fisher exact tests on the Fairfax illicit connections cluster.

	Percent_Impervious	Infrastructure_Access_Density	Drainage_Density	Aging_Sanitary_Infrastructure	Generating_Site_Density	Construction_Site_Density	Development_Age	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LU_Commercial	LU_Recreation_and_Open_Space	LU_Residential	Outfall_Density	LC_Developed_Open_Space	LC_Undeveloped	LC_Developed_High_Intensity	LU_Industrial	LU_Institutional	Population_Density	Aging_Septic_System_Density	Dog_License_Density	LC_Barren	LC_Agricultural	LU_Agricultural
LC_Barren	3.4E-06	1.0E-04	1.0E-04	1.0E-04	2.8E-04	2.8E-04	7.4E-04	4.6E-03	4.6E-03	4.6E-03	4.6E-03	4.6E-03	1.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	3.4E-06	1.0E-04	1.0E-04	1.0E-04	2.8E-04	2.8E-04	7.4E-04	4.6E-03	4.6E-03	4.6E-03	4.6E-03	4.6E-03	1.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	3.4E-06	1.0E-04	1.0E-04	1.0E-04	2.8E-04	2.8E-04	7.4E-04	4.6E-03	4.6E-03	4.6E-03	4.6E-03	4.6E-03	1.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	3.8E-05	8.3E-04	8.3E-04	8.3E-04	2.1E-03	2.1E-03	4.9E-03	2.3E-02	2.3E-02	2.3E-02	2.3E-02	2.3E-02	4.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	3.8E-05	8.3E-04	8.3E-04	8.3E-04	2.1E-03	2.1E-03	4.9E-03	2.3E-02	2.3E-02	2.3E-02	2.3E-02	2.3E-02	4.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	3.8E-05	8.3E-04	8.3E-04	8.3E-04	2.1E-03	2.1E-03	4.9E-03	2.3E-02	2.3E-02	2.3E-02	2.3E-02	2.3E-02	4.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	3.8E-05	8.3E-04	8.3E-04	8.3E-04	2.1E-03	2.1E-03	4.9E-03	2.3E-02	2.3E-02	2.3E-02	2.3E-02	2.3E-02	4.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	3.8E-05	8.3E-04	8.3E-04	8.3E-04	2.1E-03	2.1E-03	4.9E-03	2.3E-02	2.3E-02	2.3E-02	2.3E-02	2.3E-02	4.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	3.8E-05	8.3E-04	8.3E-04	8.3E-04	2.1E-03	2.1E-03	4.9E-03	2.3E-02	2.3E-02	2.3E-02	2.3E-02	2.3E-02	4.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Open_Space	2.9E-03	2.9E-02	2.9E-02	2.9E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	2.9E-03	2.9E-02	2.9E-02	2.9E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	1.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	3.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	3.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	3.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	3.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	3.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-27: R output showing p-values from one-tailed Fisher exact tests on the Fairfax illegal dumping and spills cluster.

	Percent_Impervious	Infrastructure_Access_Density	Drainage_Density	Development_Age	Aging_Sanitary_Infrastructure	Generating_Site_Density	LU_Commercial	LC_Developed_Medium_Intensity	LU_Residential	LC_Developed_High_Intensity	Construction_Site_Density	Dog_License_Density	LU_Industrial	LC_Developed_Open_Space	LC_Developed_Low_Intensity	Outfall_Density	LU_Recreation_and_Open_Space	Aging_Septic_System_Density	LC_Undeveloped	LU_Institutional	LC_Agricultural	LC_Barren	LU_Agricultural	Population_Density
LC_Barren	3.9E-24	1.7E-20	7.6E-19	2.8E-17	5.2E-13	1.8E-10	1.2E-09	2.9E-09	7.1E-09	1.7E-08	1.7E-08	2.3E-07	6.4E-06	7.0E-05	1.5E-04	3.3E-04	3.2E-03	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	3.9E-24	1.7E-20	7.6E-19	2.8E-17	5.2E-13	1.8E-10	1.2E-09	2.9E-09	7.1E-09	1.7E-08	1.7E-08	2.3E-07	6.4E-06	7.0E-05	1.5E-04	3.3E-04	3.2E-03	NA	NA	NA	NA	NA	NA	NA
Population_Density	3.9E-24	1.7E-20	7.6E-19	2.8E-17	5.2E-13	1.8E-10	1.2E-09	2.9E-09	7.1E-09	1.7E-08	1.7E-08	2.3E-07	6.4E-06	7.0E-05	1.5E-04	3.3E-04	3.2E-03	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	1.7E-22	6.2E-19	2.6E-17	8.7E-16	1.2E-11	3.3E-09	1.9E-08	4.6E-08	1.1E-07	2.5E-07	2.5E-07	2.9E-06	6.4E-05	5.7E-04	1.2E-03	2.3E-03	1.7E-02	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	5.7E-20	1.5E-16	5.1E-15	1.4E-13	1.1E-09	2.0E-07	9.9E-07	2.2E-06	4.6E-06	9.8E-06	9.8E-06	8.5E-05	1.2E-03	7.6E-03	1.3E-02	2.3E-02	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	5.7E-20	1.5E-16	5.1E-15	1.4E-13	1.1E-09	2.0E-07	9.9E-07	2.2E-06	4.6E-06	9.8E-06	9.8E-06	8.5E-05	1.2E-03	7.6E-03	1.3E-02	2.3E-02	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	6.5E-19	1.4E-15	4.6E-14	1.2E-12	6.8E-09	9.8E-07	4.5E-06	9.4E-06	1.9E-05	4.0E-05	4.0E-05	3.0E-04	3.5E-03	1.9E-02	3.1E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	1.9E-15	2.2E-12	5.1E-11	9.3E-10	1.8E-06	1.1E-04	3.9E-04	7.1E-04	1.3E-03	2.2E-03	2.2E-03	1.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	2.4E-13	1.7E-10	3.0E-09	4.3E-08	3.8E-05	1.4E-03	3.9E-03	6.3E-03	1.0E-02	1.6E-02	1.6E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	1.0E-12	6.1E-10	1.0E-08	1.3E-07	9.0E-05	2.7E-03	7.2E-03	1.1E-02	1.8E-02	2.7E-02	2.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Open_Space	3.9E-12	2.0E-09	3.1E-08	3.8E-07	2.0E-04	5.1E-03	1.3E-02	1.9E-02	2.9E-02	4.3E-02	4.3E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	1.7E-10	5.5E-08	6.7E-07	6.4E-06	1.6E-03	2.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	1.2E-08	2.2E-06	2.0E-05	1.4E-04	1.4E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	2.0E-07	2.3E-05	1.7E-04	9.4E-04	4.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	2.0E-07	2.3E-05	1.7E-04	9.4E-04	4.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	4.8E-07	4.7E-05	3.1E-04	1.6E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	1.1E-06	9.3E-05	5.8E-04	2.8E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	2.4E-06	1.8E-04	1.0E-03	4.6E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	1.1E-05	6.0E-04	3.0E-03	1.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	5.1E-04	1.2E-02	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	3.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-28: R output showing p-values from one-tailed Fisher exact tests on the Fairfax fecal bacteria cluster.

	Percent_Impervious	Development_Age	Drainage_Density	LU_Residential	LC_Developed_Low_Intensity	Infrastructure_Access_Density	LC_Developed_Open_Space	LU_Commercial	Generating_Site_Density	Construction_Site_Density	Dog_License_Density	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Undeveloped	LU_Institutional	LU_Recreation_and_Open_Space	Outfall_Density	Aging_Sanitary_Infrastructure	LC_Barren	LC_Agricultural	LU_Agricultural	LU_Industrial	Population_Density	Aging_Septic_System_Density
LC_Barren	7.0E-04	3.5E-03	3.5E-03	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	7.0E-04	3.5E-03	3.5E-03	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	7.0E-04	3.5E-03	3.5E-03	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	7.0E-04	3.5E-03	3.5E-03	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	7.0E-04	3.5E-03	3.5E-03	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	7.0E-04	3.5E-03	3.5E-03	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	5.1E-03	2.0E-02	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	5.1E-03	2.0E-02	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	5.1E-03	2.0E-02	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	5.1E-03	2.0E-02	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	5.1E-03	2.0E-02	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	5.1E-03	2.0E-02	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	5.1E-03	2.0E-02	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Open_Space	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-29: R output showing p-values from one-tailed Fisher exact tests on the Fairfax chloride cluster.

	Development_Age	Drainage_Density	LC_Developed_Low_Intensity	Percent_Impervious	LU_Recreation_and_Open_Space	LU_Residential	Outfall_Density	Aging_Sanitary_Infrastructure	LC_Developed_Open_Space	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LC_Undeveloped	LU_Commercial	LU_Institutional	Infrastructure_Access_Density	Aging_Septic_System_Density	Generating_Site_Density	LC_Barren	LC_Agricultural	LU_Agricultural	LU_Industrial	Population_Density	Construction_Site_Density	Dog_License_Density
LC_Barren	1.3E-02	1.3E-02	3.8E-02	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	1.3E-02	1.3E-02	3.8E-02	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	1.3E-02	1.3E-02	3.8E-02	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	1.3E-02	1.3E-02	3.8E-02	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	1.3E-02	1.3E-02	3.8E-02	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	1.3E-02	1.3E-02	3.8E-02	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	1.3E-02	1.3E-02	3.8E-02	3.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-30: R output showing p-values from one-tailed Fisher exact tests on the Fairfax construction cluster.

	LU_Residential	Infrastructure_Access_Density	Percent_Impervious	Drainage_Density	Aging_Sanitary_Infrastructure	LC_Developed_Open_Space	Development_Age	LC_Developed_Low_Intensity	Construction_Site_Density	Dog_License_Density	LC_Developed_Medium_Intensity	Outfall_Density	Generating_Site_Density	LU_Commercial	LU_Industrial	LU_Institutional	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LC_Undeveloped	LU_Agricultural	LU_Recreation_and_Open_Space	Population_Density	Aging_Septic_System_Density
LC_Developed_High_Intensity	1.7E-04	2.3E-03	2.3E-03	2.3E-03	6.9E-03	1.9E-02	1.9E-02	4.7E-02	4.7E-02	4.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Barren	1.7E-04	2.3E-03	2.3E-03	2.3E-03	6.9E-03	1.9E-02	1.9E-02	4.7E-02	4.7E-02	4.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	1.7E-04	2.3E-03	2.3E-03	2.3E-03	6.9E-03	1.9E-02	1.9E-02	4.7E-02	4.7E-02	4.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	1.7E-04	2.3E-03	2.3E-03	2.3E-03	6.9E-03	1.9E-02	1.9E-02	4.7E-02	4.7E-02	4.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	1.7E-04	2.3E-03	2.3E-03	2.3E-03	6.9E-03	1.9E-02	1.9E-02	4.7E-02	4.7E-02	4.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	1.7E-04	2.3E-03	2.3E-03	2.3E-03	6.9E-03	1.9E-02	1.9E-02	4.7E-02	4.7E-02	4.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	1.7E-04	2.3E-03	2.3E-03	2.3E-03	6.9E-03	1.9E-02	1.9E-02	4.7E-02	4.7E-02	4.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	1.7E-04	2.3E-03	2.3E-03	2.3E-03	6.9E-03	1.9E-02	1.9E-02	4.7E-02	4.7E-02	4.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	1.4E-03	1.4E-02	1.4E-02	1.4E-02	3.4E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	1.4E-03	1.4E-02	1.4E-02	1.4E-02	3.4E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	1.4E-03	1.4E-02	1.4E-02	1.4E-02	3.4E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	6.1E-03	4.5E-02	4.5E-02	4.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	2.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	5.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	5.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	5.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-31: R output showing p-values from one-tailed Fisher exact tests on the Fairfax FOG, surfactants, and trash cluster.

	Infrastructure_Access_Density	Percent_Impervious	Development_Age	Drainage_Density	LU_Commercial	Aging_Sanitary_Infrastructure	LC_Developed_High_Intensity	Generating_Site_Density	Construction_Site_Density	LC_Developed_Medium_Intensity	Dog_License_Density	LU_Industrial	LU_Residential	LC_Developed_Open_Space	Outfall_Density	LU_Recreation_and_Open_Space	LC_Developed_Low_Intensity	LC_Undeveloped	Aging_Septic_System_Density	LC_Barren	LC_Agricultural	LU_Agricultural	LU_Institutional	Population_Density
LC_Barren	4.1E-10	1.8E-09	3.1E-07	3.1E-07	8.4E-06	8.4E-06	6.2E-05	1.6E-04	1.6E-04	4.0E-04	9.7E-04	2.3E-03	2.6E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	4.1E-10	1.8E-09	3.1E-07	3.1E-07	8.4E-06	8.4E-06	6.2E-05	1.6E-04	1.6E-04	4.0E-04	9.7E-04	2.3E-03	2.6E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	4.1E-10	1.8E-09	3.1E-07	3.1E-07	8.4E-06	8.4E-06	6.2E-05	1.6E-04	1.6E-04	4.0E-04	9.7E-04	2.3E-03	2.6E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	4.1E-10	1.8E-09	3.1E-07	3.1E-07	8.4E-06	8.4E-06	6.2E-05	1.6E-04	1.6E-04	4.0E-04	9.7E-04	2.3E-03	2.6E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	4.1E-10	1.8E-09	3.1E-07	3.1E-07	8.4E-06	8.4E-06	6.2E-05	1.6E-04	1.6E-04	4.0E-04	9.7E-04	2.3E-03	2.6E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	7.2E-08	2.7E-07	2.7E-05	2.7E-05	4.5E-04	4.5E-04	2.4E-03	5.1E-03	5.1E-03	1.1E-02	2.1E-02	4.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	7.2E-08	2.7E-07	2.7E-05	2.7E-05	4.5E-04	4.5E-04	2.4E-03	5.1E-03	5.1E-03	1.1E-02	2.1E-02	4.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	7.2E-08	2.7E-07	2.7E-05	2.7E-05	4.5E-04	4.5E-04	2.4E-03	5.1E-03	5.1E-03	1.1E-02	2.1E-02	4.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	4.8E-07	1.7E-06	1.3E-04	1.3E-04	1.7E-03	1.7E-03	7.7E-03	1.5E-02	1.5E-02	2.9E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Open_Space	2.4E-06	8.0E-06	4.6E-04	4.6E-04	5.1E-03	5.1E-03	2.0E-02	3.6E-02	3.6E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	2.4E-06	8.0E-06	4.6E-04	4.6E-04	5.1E-03	5.1E-03	2.0E-02	3.6E-02	3.6E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	1.0E-05	3.2E-05	1.4E-03	1.4E-03	1.3E-02	1.3E-02	4.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	3.3E-04	8.5E-04	1.8E-02	1.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	8.5E-04	2.1E-03	3.4E-02	3.4E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	2.0E-03	4.6E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	4.4E-03	9.5E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	4.4E-03	9.5E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	9.1E-03	1.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	3.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	3.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-32: R output showing p-values from one-tailed Fisher exact tests on the Fairfax hazardous materials cluster.

	Percent_Impervious	Infrastructure_Access_Density	Generating_Site_Density	Drainage_Density	Aging_Sanitary_Infrastructure	Construction_Site_Density	Development_Age	LC_Developed_Medium_Intensity	LU_Commercial	LC_Developed_High_Intensity	LU_Industrial	LU_Recreation_and_Open_Space	LU_Residential	Outfall_Density	LC_Developed_Open_Space	LC_Developed_Low_Intensity	Dog_License_Density	LC_Agricultural	LC_Undeveloped	LU_Institutional	Population_Density	Aging_Septic_System_Density	LC_Barren	LU_Agricultural
LC_Barren	2.0E-12	3.9E-11	7.3E-08	7.3E-08	2.2E-07	1.8E-06	1.8E-06	4.8E-06	3.3E-05	4.7E-04	2.5E-03	5.7E-03	5.7E-03	1.2E-02	2.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	2.0E-12	3.9E-11	7.3E-08	7.3E-08	2.2E-07	1.8E-06	1.8E-06	4.8E-06	3.3E-05	4.7E-04	2.5E-03	5.7E-03	5.7E-03	1.2E-02	2.7E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	4.5E-11	7.9E-10	1.0E-06	1.0E-06	2.8E-06	2.0E-05	2.0E-05	5.0E-05	2.9E-04	3.2E-03	1.4E-02	2.8E-02	2.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	4.5E-11	7.9E-10	1.0E-06	1.0E-06	2.8E-06	2.0E-05	2.0E-05	5.0E-05	2.9E-04	3.2E-03	1.4E-02	2.8E-02	2.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	4.5E-11	7.9E-10	1.0E-06	1.0E-06	2.8E-06	2.0E-05	2.0E-05	5.0E-05	2.9E-04	3.2E-03	1.4E-02	2.8E-02	2.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	4.5E-11	7.9E-10	1.0E-06	1.0E-06	2.8E-06	2.0E-05	2.0E-05	5.0E-05	2.9E-04	3.2E-03	1.4E-02	2.8E-02	2.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	4.5E-11	7.9E-10	1.0E-06	1.0E-06	2.8E-06	2.0E-05	2.0E-05	5.0E-05	2.9E-04	3.2E-03	1.4E-02	2.8E-02	2.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	2.6E-08	3.3E-07	1.4E-04	1.4E-04	3.4E-04	1.6E-03	1.6E-03	3.3E-03	1.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	2.6E-08	3.3E-07	1.4E-04	1.4E-04	3.4E-04	1.6E-03	1.6E-03	3.3E-03	1.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Open_Space	1.3E-07	1.5E-06	4.7E-04	4.7E-04	1.0E-03	4.3E-03	4.3E-03	8.3E-03	2.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	5.6E-07	5.7E-06	1.3E-03	1.3E-03	2.7E-03	1.0E-02	1.0E-02	1.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	2.1E-06	1.9E-05	3.2E-03	3.2E-03	6.3E-03	2.1E-02	2.1E-02	3.6E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	2.1E-06	1.9E-05	3.2E-03	3.2E-03	6.3E-03	2.1E-02	2.1E-02	3.6E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	7.2E-06	5.9E-05	7.1E-03	7.1E-03	1.3E-02	4.0E-02	4.0E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	6.3E-05	4.2E-04	2.7E-02	2.7E-02	4.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	9.1E-04	4.5E-03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	4.0E-03	1.6E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	7.8E-03	2.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	7.8E-03	2.8E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	2.5E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	4.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	4.2E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-33: R output showing p-values from one-tailed Fisher exact tests on the Fairfax unknown cluster.

	LU_Residential	Generating_Site_Density	Percent_Impervious	Development_Age	Drainage_Density	Aging_Sanitary_Infrastructure	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LC_Undeveloped	LU_Commercial	Infrastructure_Access_Density	Dog_License_Density	LU_Recreation_and_Open_Space	Aging_Septic_System_Density	LC_Developed_Open_Space	LC_Developed_High_Intensity	LC_Barren	LC_Agricultural	LU_Agricultural	LU_Industrial	LU_Institutional	Population_Density	Outfall_Density	Construction_Site_Density
LC_Developed_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Low_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_Medium_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Developed_High_Intensity	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Barren	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LC_Undeveloped	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Agricultural	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Commercial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Industrial	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Institutional	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Recreation_and_Open_Space	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LU_Residential	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Population_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outfall_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Infrastructure_Access_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Septic_System_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Generating_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Construction_Site_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Percent_Impervious	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Development_Age	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Drainage_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aging_Sanitary_Infrastructure	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dog_License_Density	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table C-34: Relative ranking of risk factors for each discharge and pollutant cluster in the City of Roanoke. The dark shade indicates risk factors which had at least one significant one-tailed Fisher test with p <= 0.01, while the light shade indicates risk factors which had at least one significant one-tailed Fisher test with p <= 0.05. Unshaded cells indicate risk factors which generated no significant results and were ranked based on their performance across all clusters.

Cluster	Illicit Connections	Illegal Dumping or Spills	Fecal Bacteria	Chloride	Construction	FOG, Surfactants, Trash	Hazardous Materials	Unknown
n	7	83	9	0	9	39	31	2
1	LC_Developed_Medium_Intensity	LU_Commercial	LC_Developed_Medium_Intensity	NA	Development_Age	LU_Commercial	Development_Age	LC_Developed_Medium_Intensity
2	Development_Age	Development_Age	LU_Commercial	NA	LC_Developed_High_Intensity	Generating_Site_Density	Generating_Site_Density	LU_Commercial
3	LC_Developed_High_Intensity	Generating_Site_Density	Infrastructure_Access_Density	NA	Generating_Site_Density	LC_Developed_High_Intensity	LC_Developed_Low_Intensity	Population_Density
4	LU_Industrial	LC_Developed_High_Intensity	Generating_Site_Density	NA	Percent_Impervious	Infrastructure_Access_Density	Percent_Impervious	Construction_Site_Density
5	LU_Residential	Infrastructure_Access_Density	Development_Age	NA	LU_Commercial	Development_Age	Drainage_Density	Development_Age
6	Generating_Site_Density	Drainage_Density	Construction_Site_Density	NA	Infrastructure_Access_Density	Aging_Sanitary_Infrastructure	LC_Developed_Medium_Intensity	Dog_License_Density
7	Infrastructure_Access_Density	Aging_Sanitary_Infrastructure	Percent_Impervious	NA	Drainage_Density	Drainage_Density	LC_Developed_High_Intensity	Generating_Site_Density
8	Percent_Impervious	Percent_Impervious	Drainage_Density	NA	Aging_Sanitary_Infrastructure	Percent_Impervious	Aging_Sanitary_Infrastructure	LC_Developed_High_Intensity
9	Drainage_Density	LC_Developed_Medium_Intensity	Aging_Sanitary_Infrastructure	NA	LC_Developed_Medium_Intensity	LC_Developed_Medium_Intensity	LU_Commercial	Infrastructure_Access_Density
10	Aging_Sanitary_Infrastructure	Construction_Site_Density	LC_Developed_High_Intensity	NA	LU_Residential	Construction_Site_Density	LU_Industrial	Percent_Impervious
11	LU_Commercial	LC_Developed_Low_Intensity	LU_Residential	NA	LU_Industrial	LU_Residential	LU_Residential	Drainage_Density
12	Construction_Site_Density	LU_Residential	LC_Developed_Low_Intensity	NA	LC_Developed_Low_Intensity	Dog_License_Density	Infrastructure_Access_Density	Aging_Sanitary_Infrastructure
13	Dog_License_Density	Dog_License_Density	LU_Institutional	NA	Population_Density	LU_Institutional	Construction_Site_Density	LU_Residential
14	Population_Density	Population_Density	LU_Recreation_and_Open_Space	NA	Construction_Site_Density	LU_Recreation_and_Open_Space	Population_Density	LU_Industrial
15	LC_Developed_Low_Intensity	LU_Industrial	LU_Industrial	NA	Dog_License_Density	Population_Density	Dog_License_Density	LC_Developed_Low_Intensity
16	LU_Recreation_and_Open_Space	LU_Recreation_and_Open_Space	Dog_License_Density	NA	LU_Recreation_and_Open_Space	LU_Industrial	LU_Recreation_and_Open_Space	LU_Recreation_and_Open_Space
17	LU_Institutional	LU_Institutional	Population_Density	NA	LU_Institutional	LC_Developed_Low_Intensity	Outfall_Density	LU_Institutional
18	Aging_Septic_System_Density	Outfall_Density	Aging_Septic_System_Density	NA	Aging_Septic_System_Density	Aging_Septic_System_Density	LU_Institutional	Aging_Septic_System_Density
19	Outfall_Density	Aging_Septic_System_Density	Outfall_Density	NA	Outfall_Density	Outfall_Density	Aging_Septic_System_Density	Outfall_Density
20	LC_Developed_Open_Space	LC_Developed_Open_Space	LC_Developed_Open_Space	NA	LC_Developed_Open_Space	LC_Developed_Open_Space	LC_Developed_Open_Space	LC_Developed_Open_Space
21	LC_Undeveloped	LC_Undeveloped	LC_Undeveloped	NA	LC_Undeveloped	LC_Undeveloped	LC_Undeveloped	LC_Undeveloped
22	LC_Agricultural	LC_Agricultural	LC_Agricultural	NA	LC_Agricultural	LC_Agricultural	LC_Agricultural	LC_Agricultural
23	LC_Barren	LC_Barren	LC_Barren	NA	LC_Barren	LC_Barren	LC_Barren	LC_Barren
24	LU_Agricultural	LU_Agricultural	LU_Agricultural	NA	LU_Agricultural	LU_Agricultural	LU_Agricultural	LU_Agricultural

Table C-35: Relative ranking of risk factors for each discharge and pollutant cluster in the County of Fairfax. The dark shade indicates risk factors which had at least one significant one-tailed Fisher test with p <= 0.01, while the light shade indicates risk factors which had at least one significant one-tailed Fisher test with p <= 0.05. Unshaded cells indicate risk factors which generated no significant results and were ranked based on their performance across all clusters.

Cluster	Illicit Connections	Illegal Dumping or Spills	Fecal Bacteria	Chloride	Construction	FOG, Surfactants, Trash	Hazardous Materials	Unknown
n	23	77	8	8	12	30	36	6
1	Percent_Impervious	Percent_Impervious	Percent_Impervious	Development_Age	LU_Residential	Infrastructure_Access_Density	Percent_Impervious	LU_Residential
2	Infrastructure_Access_Density	Infrastructure_Access_Density	Development_Age	Drainage_Density	Infrastructure_Access_Density	Percent_Impervious	Infrastructure_Access_Density	Generating_Site_Density
3	Drainage_Density	Drainage_Density	Drainage_Density	LC_Developed_Low_Intensity	Percent_Impervious	Development_Age	Generating_Site_Density	Percent_Impervious
4	Aging_Sanitary_Infrastructure	Development_Age	LU_Residential	Percent_Impervious	Drainage_Density	Drainage_Density	Drainage_Density	Development_Age
5	Generating_Site_Density	Aging_Sanitary_Infrastructure	LC_Developed_Low_Intensity	LU_Recreation_and_Open_Space	Aging_Sanitary_Infrastructure	LU_Commercial	Aging_Sanitary_Infrastructure	Drainage_Density
6	Construction_Site_Density	Generating_Site_Density	Infrastructure_Access_Density	LU_Residential	LC_Developed_Open_Space	Aging_Sanitary_Infrastructure	Construction_Site_Density	Aging_Sanitary_Infrastructure
7	Development_Age	LU_Commercial	LC_Developed_Open_Space	Outfall_Density	Development_Age	LC_Developed_High_Intensity	Development_Age	LC_Developed_Low_Intensity
8	LC_Developed_Low_Intensity	LC_Developed_Medium_Intensity	LU_Commercial	Aging_Sanitary_Infrastructure	LC_Developed_Low_Intensity	Generating_Site_Density	LC_Developed_Medium_Intensity	LC_Developed_Medium_Intensity
9	LC_Developed_Medium_Intensity	LU_Residential	Generating_Site_Density	LC_Developed_Open_Space	Construction_Site_Density	Construction_Site_Density	LU_Commercial	LC_Undeveloped
10	LU_Commercial	LC_Developed_High_Intensity	Construction_Site_Density	LC_Developed_Medium_Intensity	Dog_License_Density	LC_Developed_Medium_Intensity	LC_Developed_High_Intensity	LU_Commercial
11	LU_Recreation_and_Open_Space	Construction_Site_Density	Dog_License_Density	LC_Developed_High_Intensity	LC_Developed_Medium_Intensity	Dog_License_Density	LU_Industrial	Infrastructure_Access_Density
12	LU_Residential	Dog_License_Density	LC_Developed_Medium_Intensity	LC_Undeveloped	Outfall_Density	LU_Industrial	LU_Recreation_and_Open_Space	Dog_License_Density
13	Outfall_Density	LU_Industrial	LC_Developed_High_Intensity	LU_Commercial	Generating_Site_Density	LU_Residential	LU_Residential	LU_Recreation_and_Open_Space
14	LC_Developed_Open_Space	LC_Developed_Open_Space	LC_Undeveloped	LU_Institutional	LU_Commercial	LC_Developed_Open_Space	Outfall_Density	Aging_Septic_System_Density
15	LC_Undeveloped	LC_Developed_Low_Intensity	LU_Institutional	Infrastructure_Access_Density	LU_Industrial	Outfall_Density	LC_Developed_Open_Space	Construction_Site_Density
16	LC_Developed_High_Intensity	Outfall_Density	LU_Recreation_and_Open_Space	Aging_Septic_System_Density	LU_Institutional	LU_Recreation_and_Open_Space	LC_Developed_Low_Intensity	LC_Developed_Open_Space
17	LU_Industrial	LU_Recreation_and_Open_Space	Outfall_Density	Generating_Site_Density	LC_Developed_High_Intensity	LC_Developed_Low_Intensity	Dog_License_Density	Outfall_Density
18	LU_Institutional	Aging_Septic_System_Density	Aging_Sanitary_Infrastructure	Construction_Site_Density	LU_Recreation_and_Open_Space	LC_Undeveloped	LC_Agricultural	LC_Developed_High_Intensity
19	Population_Density	LC_Undeveloped	LU_Industrial	Dog_License_Density	LC_Undeveloped	Aging_Septic_System_Density	LC_Undeveloped	LU_Industrial
20	Aging_Septic_System_Density	LU_Institutional	Aging_Septic_System_Density	LU_Industrial	Aging_Septic_System_Density	LU_Institutional	LU_Institutional	LU_Institutional
21	Dog_License_Density	LC_Agricultural	Population_Density	Population_Density	Population_Density	Population_Density	Population_Density	Population_Density
22	LC_Agricultural	Population_Density	LC_Agricultural	LC_Agricultural	LC_Agricultural	LC_Agricultural	Aging_Septic_System_Density	LC_Agricultural
23	LC_Barren	LC_Barren	LC_Barren	LC_Barren	LC_Barren	LC_Barren	LC_Barren	LC_Barren
24	LU_Agricultural	LU_Agricultural	LU_Agricultural	LU_Agricultural	LU_Agricultural	LU_Agricultural	LU_Agricultural	LU_Agricultural

APPENDIX D – EXTENDED DISCUSSION OF RISK FACTOR UNCERTAINTY

CATEGORICAL VARIABLES

Land Cover

The land cover risk factor is a description of the risk of discharge due to the existing land cover, indicated by the 2011 National Land Cover Dataset (NLCD). This risk factor indicates a general level of development in a region, which may have importance in determining the type or number of discharges present in certain areas. The NLCD for each jurisdiction was reclassified into seven (7) categories describing general land cover, as well as intensity of urban land cover. The uncertainty in this dataset lies in its resolution and its age. The NLCD classifies land cover in 30x30 meter cells, which is generally considered to be a fine (sharp) resolution for regional or national analyses; however, it has been proposed that the NLCD is too coarse to be used for local applications at the scale of an individual jurisdiction such as Fairfax County or the City of Roanoke. Given the coarse scale of the data, there is a chance that any given cell may actually contain multiple land cover types. As different land cover types appeared to capture different types and quantities of illicit discharges, this issue of resolution most likely impacted the performance of this risk factor. In addition, as the data was most likely 5 years old at the time of analysis, it may not accurately represent the true land cover conditions at present, further adding to the uncertainty produced by the resolution of the NLCD.

Land Use

The land use risk factor describes the land use on any given parcel, which gives a better indication of the types of activities and potential pollution that may be present, when compared to land cover. Land use data was available from Fairfax County. For Roanoke, land use data was unavailable, and zoning data was used as a surrogate as it is related to the types of land uses permitted on a parcel. There is uncertainty in the zoning data for Roanoke as it reflects a 5-year old ordinance, and several parcels are suspected to have changed zoning since that time. For this reason, data for Fairfax County may be more accurate than data for Roanoke. It should be noted that neither jurisdiction reported the methods used to collect these datasets, and both datasets lost resolution when converted to a 30x30 meter raster grid. From a sensitivity standpoint, it was hypothesized early on that this risk factor may interact with other risk factors, land uses that are perceived as high-risk are often more developed and more heavily drained by infrastructure.

CONTINUOUS VARIABLES EVALUATED OVER SUBUNITS

Population Density

Population density was defined as the number of people per square kilometer, measured within block groups for each jurisdiction, which was the smallest census unit for which population data was available. The risk factor functions as an indicator of development intensity as more people living in the same area could indicate that the region is more urbanized. The input data for this factor is from the US Census Bureau and the USDA geospatial Data Gateway, and is of high quality. It was hypothesized that this risk factor would be more affected by raster conversion given the relatively small size of the analysis units (block groups) when compared to 30x30 meter raster cells. Resolution was lost around the perimeters of these block groups since a 30-meter grid could easily span multiple parcels, throwing them into a cell from an adjoining block group. From a sensitivity standpoint, it is important to note that this risk factor is represented by a continuous variable which requires binning to come up with risk categories. While this binning can be as accomplished through GIS techniques such as Jenks Natural Breaks, it is still somewhat arbitrary in its manner of defining what constitutes high risk. The chosen thresholds significantly impacted the risk factor's performance by limiting the area defined as high risk to a handful of block groups.

Percent Impervious

The percent impervious risk factor was defined as the proportion of area covered by impervious surfaces such as concrete or rooftops, evaluated over individual watersheds. Percent impervious is a well-known indicator of development, and has a direct effect on watershed hydrology. It also creates more potential surfaces for the accumulation and wash-off of pollutants, and was therefore tested as a risk factor for illicit discharges. The input datasets for this factor included watershed polygons for each jurisdiction, and the 2011 NLCD impervious cover dataset. Two major problems were noted that play into the uncertainty in the percent impervious risk factor – the resolution of the NLCD impervious raster and the watershed layers. Given the large cell size of the raster (30x30m), some cells overlapped watersheds, which had a slight effect on the watershed impervious calculations. A slight loss in resolution was also expected when watershed polygons were converted to raster cells; however, given the relatively large size of the watersheds used in this analysis, this loss was presumed to be small. The watersheds increased uncertainty in the percent impervious calculations as the watersheds were clipped to the jurisdictional boundaries. It was noted that watersheds with a small fraction of area inside a jurisdiction had disproportionately high impervious percentages. Also, the size of watersheds differs significantly between the two jurisdictions, as Fairfax County delineated larger watersheds over a larger jurisdictional area. This discrepancy resulted in overall lower impervious percentages for Fairfax County as each watershed contained more cells, smoothing out the effect of any highly impervious areas. Sensitivity of the risk level to the binning of percentage (a continuous variable) was also suspected to be high.

It should be noted that many jurisdictions, including Fairfax County and Roanoke, maintain local datasets describing impervious cover at a much higher resolution than the NLCD. These datasets are often used to manage stormwater utilities or other policies. In this study, the NLCD was used instead of these layers to demonstrate the capability of a free, publicly-available dataset. However, the accuracy of this risk factor could be significantly improved by using local high-resolution datasets over the NLCD 30m raster data.

Development Age

Development age was defined as the number of years since a structure was built on a parcel, and these parcel ages were then aggregated over watersheds to determine the median age of each watershed. Development age, by this definition, provides an idea of the age of infrastructure such as stormwater pipes or water/sanitary hookups for residences and businesses. However, in this analysis the risk factor was considered a surrogate for age of infrastructure as data for age of infrastructure was not available for both case study jurisdictions. Nonetheless, both jurisdictions had high quality data going into production. Uncertainty was centered on the number of null development age attributes which existed in Fairfax (8%) and Roanoke (21%). Based on the data sources used for this information, nulls can be interpreted as either completely undeveloped, or developed in some way which does not involve a structure or dwelling (e.g. parking lot, scrap yard, etc.). Some of these “semi-developed” parcels may still have MS4 infrastructure or other utilities or land uses present which over time could contribute to cross-connections and failing infrastructure problems. The median development age was used for watershed calculations to mitigate for the extent of nulls in the data, and the “intersect” method was also used for the spatial join operation in GIS to account for parcels which may overlap multiple watersheds. The final risk map incurred some uncertainty due to the conversion of watershed polygons to 30x30 m raster grids, but this was much less uncertainty than if parcels had been used as the base unit for raster conversion. The binning threshold for development age (a continuous variable) also directly affected the extent of high-risk area.

Outfall Density

Outfall density measures the number of outfalls from the MS4 network per kilometer of jurisdictional waters or streams. In theory, the higher the outfall density, the more connections there are between the MS4 and the stream, which may indicate a dense stormwater drainage network with many catchments of different sizes. This risk factor used three different layers for each jurisdiction. Watersheds and outfall nodes were collected from jurisdictional sources. However, jurisdictions did not have comparable useful streamlines, so the National Hydrography Dataset (NHD) was used. Uncertainty is present in all layers, although the data from Fairfax is generally in better shape than the data from Roanoke. Watersheds are generally high quality, which is important as they are the subunit of analysis used for overlays and raster conversion. However, watersheds were typically only delineated using topography with no consideration for subsurface drainage, some loss of resolution along watershed boundaries occurred during raster conversion. NHD data are of much lower quality as they are digitized at a much smaller scale compared to the scale of this analysis. Stream data was missing from NHD layers if streams were too small, and some of the streamlines no longer existed due to development or stream burial and rerouting. Also, it's important to note that little to no stream data was available in several watersheds in each jurisdiction due to (1) the finer scale of watershed delineation, and (2) the geographic location of the jurisdiction relative to the streams. In Roanoke, the Mason Creek, Back Creek, and Trout Run watersheds all contained no length of stream inside the city as they were either fringe watersheds, or the stream was too small or buried such that the NHD dataset did not include it. Fairfax outfall data was nearly complete, although Fairfax defined two different types of outfalls. Roanoke outfall data is not as complete, but is more accurate around streams. For this reason a buffer selection method was chosen; however, buffer distance significantly impacted density values since outfalls were located at variable distances from the stream in both jurisdictions. As with other continuous variables the binning thresholds for outfall density values impacted the risk designations.

Aging Sanitary Infrastructure

The process used to evaluate the sanitary network involved rating material and age characteristics of pipes to calculate a weighted sum of risk in each watershed. Individual pipes were assigned a risk level (1, 2, 3) based on pipe age and material, and this risk level was multiplied by the pipe length. Pipe scores were then summed over the entire watershed and divided by the total length of sanitary pipe in the watershed to produce a watershed risk score between 1 and 3. These scores were binned into three risk levels to produce the final risk factor layer. This risk factor uses watersheds as the base layer to evaluate the condition of the sanitary sewer network and the potential for the sanitary sewer system to contribute illicit discharges to the drainage area through leaks, overflows, or illicit connections. The watershed is used as the unit of analysis since it is the smallest hydrologic unit available to both jurisdictions which would catch a sanitary illicit discharge. Using watersheds resulted in some inaccuracy given that these watersheds are not completely delineated to subsurface infrastructure and therefore do not completely represent the drainage areas which would catch leaking sewage or overflows. In addition, some resolution was lost during the raster map conversion, but this uncertainty is small compared to the uncertainty in the delineations.

Watershed scores were quite variable and were inaccurate in some cases. Some watersheds had scores of zero because they were completely served by septic systems. Other fringe watersheds had unusually low or high scores due to the presence of only a handful of new or old pipes, respectively. Therefore, fringe watersheds exhibited more uncertainty than interior watersheds. In addition, although age and material information was relatively widely available, about 7% of each jurisdictional dataset lacked one or both of these attributes. Since these pipes were spread across the jurisdiction, the effect was not localized to one watershed, but still created a source of uncertainty. Lastly, both jurisdictions are still in the process of verifying the sanitary network and are continually adding to the databases, meaning that the data used for this analysis do not reflect the true sanitary network conditions at present.

Drainage Density

Drainage density is typically defined as the length of stream channels divided by the watershed drainage area; however, this definition was modified for application to urban drainage networks in this study. The definition used in this study was the length of drainage conveyances (open channels, pipes, and streams) divided by the watershed drainage area. This risk factor used the MS4 drainage infrastructure to provide a qualitative indication of drainage network development in the watershed. The more built out the drainage network, the higher the drainage density (kilometers of conveyances per square kilometer of area) and the quicker the watershed will drain to its outlet. Several problems arose from using local datasets to calculate this risk factor. The watersheds that were used were generally not delineated to include the effects of subsurface infrastructure on the watershed boundary. In addition, watersheds which are small in size (due to clipping or its natural state) often showed exaggerated densities in either the high or low direction. Some watersheds had no infrastructure included due to incomplete surveying or clipping of the watershed along the jurisdictional fringes. These sources of uncertainty were further compounded when the map was converted to raster format, resulting in a loss of resolution around the watershed boundaries.

The MS4 conveyance datasets used to calculate drainage lengths were also highly uncertain. Some localities have include pieces of infrastructure along streams (i.e. where streams have been piped or channelized) but generally do not contain the full stream network. Since including the stream network from NHD data would have significantly overestimated the density by including overlapping features, the NHD features were not included. The consequence of not including the NHD data was that most watershed drainage densities were slightly underestimated. The lack of connectivity between conveyances in both locality’s datasets suggests that neither dataset is truly complete, and that more surveys would reveal more conveyances and increase density values across the board.

CONTINUOUS VARIABLES EVALUATED OVER JURISDICTIONAL BOUNDARIES

Generating Site Density

The generating site risk factor is defined as the number of potential generating sites per square kilometer, where potential generating sites are facilities ranging from gas stations to restaurants to industrial plants which inherently carry the risk of contributing illicit discharges of various pollutants to the watershed. Data was retrieved from the EPA’s ECHO database to illustrate the use of free, publicly-available data which may be used to create this risk factor. Uncertainty in this risk factor is due both to the non-spatial data format of input data and the extent of data included in these records. Data must be geospatially located, which is an arduous process that sometimes results in improper matches, or the inability to locate records due to a lack of data describing the business location. However, diligent geolocation and verification of resulting data has led to datasets for Roanoke and Fairfax which are 71% and 93% spatially accurate, respectively. The majority of unmatched records for these municipalities are actually outside of the jurisdictional boundaries, although a few within the boundary could not be located. Therefore, the data actually has a higher certainty than indicated by these statistics. It is important to note, however, that these are not complete lists of sites which could potentially produce an illicit discharge. Localities often have more detailed data which they may wish to combine with this information to consider other types of facilities. Raster conversion and binning of the continuous density values also contributed to uncertainty in this risk factor.

Infrastructure Access Density

Infrastructure access density is defined as the number of access points to the MS4 network per square kilometer, where access points are defined as catch basins, pond outlet structures, pipe ends, and other features open to the atmosphere, allowing for materials to enter the MS4 network. Uncertainty regarding infrastructure access density is limited to the node and Stormwater Control Measure (SCM) feature classes. While the data collection procedures used in both municipalities are similar and comprehensive, the Fairfax dataset is more complete, given the time and additional resources that Fairfax has put into verifying their stormwater network. Roanoke has begun a comprehensive infrastructure verification program in the last few years, but has not progressed as far as Fairfax. As such, there is more uncertainty surrounding the completeness and accuracy of the data in the City of Roanoke. However, Fairfax County has a larger range of node types which had to be filtered for this process. As detailed descriptions of these types were not given in metadata, it is possible that some nodes may have been included or omitted when in reality they are not considered access points.

Both datasets went through the same set of processes in the GIS environment, and the most impactful processes were the kernel density function and the raster to polygon tool. The kernel density function was used to determine the density of stormwater infrastructure access points, and generated a smoothly tapered surface representing density values. This continuous surface was broken when the reclassify tool was used to bin the continuous density variable into a set of risk levels, which were used for the final map. The threshold used to bin these variables also directly affected the amount of area designated as high risk. The overlay was made by converting the raster to a vector format using the raster to polygon tool. These processes may have had some effects on the final data layers used to carry out the risk analysis.

Construction Site Density

Construction site density measured the number of active construction sites per square kilometer. Data for construction sites was limited to publicly available DEQ data on Construction General Permits (CGP), which generally do not cover individual sites < 1 acre. Municipalities may want to augment this dataset with some of their own permit data based on lower areas of disturbed land or other parameters. As far as the DEQ's data is concerned, the level of accuracy is relatively high due to the ability to geospatially locate all records in the GIS environment, although some points may still be displaced from the actual construction site by 500 feet based on verification findings. The choice of the binning threshold for continuous density values was challenging for this risk factor because the Roanoke CGP dataset only contained 33 records across the whole jurisdiction. With very few data points available and a small range of calculated densities, this risk factor is most likely not representative of the true conditions in the City of Roanoke, and would benefit from the inclusion of other types of building permits. It was observed as with other continuous variables that the choice for this threshold significantly impacted the number of discharges captured in high-risk areas. There was also some loss of resolution due to the raster to polygon conversion.

Aging Septic System Density

Aging septic system density measures the number of aging septic systems per square kilometer, which provides an indication of the potential for aging or failing septic systems to contribute a significant pollutant load to streams or stormwater networks. Data for septic system locations in Fairfax was immediately available from the County, and the data are of high quality. However, the City of Roanoke did not have data available to show exact locations of septic parcels, although the Virginia Department of Health (VDH) maintained a list of neighborhoods or local areas which were known septic hot spots. This list and data on parcel structure values, utility service, and buildout year, were used to develop and verify a selection algorithm to identify the most probable locations of septic systems in the city, and determine their age. It is also important to note that development age was used as a surrogate for actual septic tank age in both jurisdictions, as age information on septic tanks was not readily available from local or state health departments. This is an added source of uncertainty on top of location accuracy.

The risk factor layers for both jurisdictions were affected by the use of a density function instead of direct parcel overlay. Large parcels translated to one point over a large area and produced localized lower density values, causing some areas to be assigned lower risk. Uncertainties associated with raster to polygon conversion and the binning of the continuous density variable also resulted in a loss of data resolution. For the City of Roanoke, uncertainty is also present due to the selection algorithm used to identify potential septic parcels. While the algorithm correctly identified major clusters of septic parcels, it also identified scattered individual parcels around the city which created higher density and risk values in areas which may or may not be on septic.

Dog License Density

Dog license density measured the number of dogs per square kilometer, which is considered an IDP risk factor because of the high potential for improper disposal of dog waste, as well as polluted runoff from dog parks or high density dog neighborhoods. This risk factor has very different levels of uncertainty between jurisdictions. The dataset for the City of Roanoke contains 5394 records of dog licenses, and the address and road centerline data available were high quality, allowing for more than 99% of the records to be geospatially located; therefore, dog license density for Roanoke was calculated using a nearly complete spatial dataset, which significantly reduces uncertainty. Fairfax County contains a much larger dataset of 82510 records; however, only 69% of these records could be geospatially located, and due to the quantity of remaining records and project time constraints it was infeasible to manually locate additional records. Furthermore, the address attributes in the Fairfax County road centerlines layer were not in the proper format for geocoding and required heavy manipulation beyond what was required for Roanoke. In addition, the license records had multiple misspelled entries for local towns within the County jurisdiction, which complicated the geocoding process. Due to the combined data quality issues in the road centerlines and license records, multiple geographic regions within Fairfax County contained no spatially located license records, which significantly impacted the density calculations and distorted the risk factor. Both jurisdictions experienced uncertainty due to distance offsets from actual parcels arising from the geolocation process, potential inaccuracies in addresses given on license records, and the inability to map PO Boxes or addresses outside of the jurisdictional area. In addition, binning the continuous density variable and converting from raster to vector formats created uncertainty in the final risk factor layers.