Evaluating the Impact of Geophysical Variables on the Bacterial Contamination of Private Wells in Southern Ontario

By Erik Wright
Supervised by Dr. Geoff Hall and Dr. Anna Majury
Examined by Dr. Stephen Brown

Abstract
Bacterial contamination in groundwater is not often analyzed beyond local and regional scales. The presence of bacterial contaminants in the groundwater of southern Ontario has been observed, and several areas have been identified as being at higher relative risk of E. coli contamination (Krolik et al, 2013). This discovery by Public Health Ontario has prompted further research to evaluate relationships between observed levels of bacterial contamination and the geological setting of private wells. Using approximately 38,000 geocoded wells, statistical analysis was performed to compare overburden depth, soil texture, bedrock classification, and surficial deposit classification to observed levels of total coliform bacteria and E. coli. A very weak negative correlation was observed between overburden depth and E. coli, but not total coliform bacteria. One-Way ANOVA identified statistically significant differences among the means of total coliform bacteria counts in bedrock and surficial units, but not for E. coli. One-Way ANOVA identified statistically significant differences among the means of both total coliform bacteria and E. coli in wells situated in soils of different texture.
## Table of Contents

1 - Introduction ............................................................................................................................ 4
  1.1 - Foreword ............................................................................................................................. 4
  1.2 - Purpose .................................................................................................................................. 4
  1.3 - Hypothesis ........................................................................................................................... 5

2 - Literature Review ...................................................................................................................... 6
  2.1 - Well Water in Rural Ontario ............................................................................................... 6
  2.2 - Indicator Bacteria ............................................................................................................... 6
  2.3 - Aquifers ............................................................................................................................... 7
  2.4 - Overburden .......................................................................................................................... 8
    Texture ......................................................................................................................................... 8
    Horizons ..................................................................................................................................... 8
    Bacteria Survival Rates in Overburden ....................................................................................... 9
    Using Overburden to Evaluate Groundwater Vulnerability ...................................................... 10
    Conclusions about Overburden ................................................................................................. 11
  2.5 Surficial Geology ..................................................................................................................... 11
    Glacial Till .................................................................................................................................. 12
    Lacustrine Sediments ................................................................................................................ 12
    Glacio-marine sediments .......................................................................................................... 13
    Glacio-fluvial sediments ............................................................................................................ 13
    Conclusions about Surficial Geology ....................................................................................... 14
  2.6 - Bedrock Geology ................................................................................................................... 14
    Limestone and Dolostone Aquifers .......................................................................................... 14
    Shale .......................................................................................................................................... 16
    Sandstone and Arkose Aquifers .............................................................................................. 16
    Fracturing .................................................................................................................................. 17
    Porosity ...................................................................................................................................... 18
  2.7 - Conclusions .......................................................................................................................... 19

3 - Methodology ............................................................................................................................. 20
  3.1 - Rural Drinking Water Wells .............................................................................................. 20
  3.2 - Bedrock Geology ................................................................................................................ 21
  3.3 - Surficial Geology ................................................................................................................ 24
  3.4 - Overburden Depth .............................................................................................................. 25
  3.5 - Soil Survey Complex .......................................................................................................... 26
  3.6 - Spatial Join ........................................................................................................................... 28
  3.7 - Spearman Correlation ........................................................................................................ 30
  3.8 - One-Way Analysis of Variance .......................................................................................... 31

4 - Results ....................................................................................................................................... 32
  4.1 - Overburden Depth .............................................................................................................. 32
  4.2 - Bedrock Geology ................................................................................................................ 33
  4.3 - Surficial Geology ................................................................................................................ 34
  4.4 - Soil Texture .......................................................................................................................... 35
5 - Discussion ................................................................................................................................................. 37
5.1 - Survival Rates of Bacteria ..................................................................................................................... 37
5.2 - Quantitative Distribution of Total Coliform Bacteria and E. coli .............................................................. 38
5.3 - Overburden Depth ..................................................................................................................................... 39
5.4 - Soil Type .................................................................................................................................................. 41
5.5 - Bedrock Geology ..................................................................................................................................... 42
5.6 - Surficial Geology ..................................................................................................................................... 44
6 - Conclusions .................................................................................................................................................. 45
7 - References .................................................................................................................................................... 46
1 - Introduction

1.1 - Foreword
Groundwater is found in the void spaces of soils and bedrock, and is a water source for human and natural systems. In Canada, it is estimated that groundwater is a source of drinking water for approximately 12% of Canadians (Heather et al, 2015). Even in Ontario where fresh surface water is abundant, 23% of all water use is sourced from groundwater (Simpson, 2010). Therefore, it is imperative that groundwater sources are secured from contamination. The quality of groundwater can be impacted by land use in rural areas increasing the risk of contamination by bacteria, notably due to agricultural activities and the use of septic systems (Levison and Novakowski, 2009; Peed et al. 2011). The associated overburden and geology of these systems may further influence groundwater quality consequent to these land use activities (Kaufman et al, 2011).

Bacterial contaminants of water may include *Escherichia coli* (*E. coli*), a group of bacteria normally associated with the digestive system of mammals, and that comprises 75% of all waterborne fecal coliform bacteria (Altherholt et al, 2013). Specific strains of *E. coli* such as *E. coli* O157:H7 are pathogenic and, when consumed, may have human health consequences, including vomiting, diarrhea and stomach cramping (PHAC, 2015). Groundwater with elevated levels of *E. coli* bacteria is considered to be fecally contaminated and unsafe to drink. Drivers of bacterial contamination in rural regions include poorly managed and leaky septic systems, proximal grazing of livestock or wildlife, and shallow or improperly sealed private wells (Alevi et al, 2013).

Bacterial contamination of groundwater, as a consequence of the aforementioned land use activities, may be impacted by the amount of overburden present, the soil texture, and the characteristics of bedrock and surficial geology in the regions where wells are located (Allen and Morrison, 1973). These geophysical factors regulate the pathways water travels over land and through the subsurface to the water table. Water serves as the vector for transport of bacteria over land and in the subsurface (Wilkinson et al, 1994). Investigating these variables in relation to observed well water quality will add much needed information in support of risk assessment and mitigation resulting in improved public health outcomes.

1.2 - Purpose
The purpose of this study was to evaluate the role of several geophysical variables on the rate of total coliform and *E. coli* contamination in rural drinking water wells in southern Ontario.
based on historical well water quality data (Krolik et al., 2013). Geological variables investigated include the depth of overburden to bedrock, soil texture, the classification of the bedrock, and the classification of surficial geology. Suggestions based on the research findings that may inform the modeling of bacterial contamination and the evaluation of rural drinking water sources will be provided.

1.3 - Hypothesis
It was predicted that rural drinking water wells would have lower observed concentrations of total coliform bacteria and/or E. coli when overburden depth was greatest, meaning that bacterial concentrations are negatively correlated with overburden depth. It was also predicted that different types of bedrock geology, surficial geology and soil texture would impact bacterial contamination differently based on mean bacterial concentrations within each geophysical group. These predictions were made with a literature review supporting the assumption that the variables used represented the primary components of the hydrological cycle relevant to the transportation of coliform bacteria.
2 – Literature Review

2.1 - Well Water in Rural Ontario
In rural areas, most households do not receive their drinking water from a central treatment and distribution facility. In Canada, it is estimated that between 14-15% of Canadians living primarily in rural areas use a private well pumped from groundwater (CWWA, 2012). These households rely on in situ water treatment systems, if any. Available methods of treatment include chlorination, ozonation, and UV disinfection. Environmental setting remains one of the principal variables determining vulnerability for contaminated groundwater. Understanding the pathways by which contamination reaches a well via groundwater is important from an environmental and human health standpoint.

In Ontario, Regulation 903, amended under the Ontario Water Resources Act, governs the construction of a well. These regulations require a well have at least 6m of watertight casing, disinfected with 50-200 milligrams of chlorine, and sealed (Ontario, 2014). However, there are no regulations in Ontario regarding drinking water from wells for private use. Well maintenance, water treatment and well water monitoring are entirely the responsibility of the well water owner, making this particular water source, and the population that relies upon it, particularly vulnerable.

2.2 - Indicator Bacteria
Bacterial indicators are studied to evaluate the quality of a sample taken from a body of surface water or groundwater. Bacterial indicators are measured as a concentration expressed as the number of colony forming units per volume of water (USEPA, 2015). Microbiological tests are used to determine the concentration of a bacterial indicator to evaluate the possible presence of pathogens in a water body, with the assumption that samples will be representative of the water body. Therefore, it is imperative that bacterial indicators are used that captures the most accurate level of contamination and indicates presence of pathogens.

Because microbiological tests are evaluating a given genus of bacteria known to be associated with a pathogen, these bacteria are only indicators of contamination that may be used to infer water quality, and are not an absolute measure (Ashbolt et al., 2001). Broader indicators of contamination, like total coliform, test for bacteria from a range of sources not limited to humans and agricultural land use (Health Canada, 2013). Although important to consider from a human health standpoint, these tests may not be diagnostic of a given pathogen and a source useful for risk assessment and mitigation. Using *E. coli* as an assay for bacterial contamination provides indication
with greater certainty of fecal contamination, and has been found to consistently predict gastrointestinal illness derived from drinking water (Doyle and Erickson, 2006). However, use of *E. coli* as an indicator of fecal contamination in groundwater is limited by populations of *E. coli* naturalized in the environment (Ishii *et al.*, 2006). While total coliform and *E. coli* remain principal indicators of the presence of harmful bacteria in groundwater, it is important to recognize the challenges in inferring how representative indicators may be. Evolving methods of enhancing the indication of fecal contamination by bacteria include the use of microbial source tracking (Krolik *et al*, 2014), which allows diagnosis of fecal contamination and source to be made with much greater certainty.

2.3 - Aquifers

An aquifer contains and transmits groundwater in the fractures or pores of a geological unit, or in the void space of unconsolidated geological material composed of soils and rocks (Arthur and Saffer, 2016). Aquifers may be defined through an anthropogenic lens as units from which water can be pumped for human use. Unconfined aquifers are characterized by phreatic conditions, and do not have a confining unit of impermeable clay or bedrock creating pressure, and contain the water table line. A confined aquifer is situated between two aquitards at depth, and exists under pressure. Confining layers may either be non-lithified clay or silt deposits, or impermeable rock such as shale, siltstone, quartz, or carbonates. As result, unconfined aquifers commonly receive other natural or anthropogenic inputs from the land surface (Reynolds, 2017). As described below, water will enter these units over time during the hydrological cycle by traveling through overburden. Rock units that are not concealed by overburden will receive water without prior residence in soil.

Water is the primary mode of transport for bacteria in porous media like soil and rock during overland and subsurface flow. Common point sources of *E. coli* contamination included agricultural inputs, livestock manure, leaking septic systems, feces of wild animals (BC Ministry of Environment, 2007) or naturalized *E. coli* (Ishii *et al*, 2006). Therefore, the discussion of geophysical influences on rates of bacterial contamination will primarily concern how each regulates the flow of water from a point source of bacterial contamination to an aquifer drawn from by wells for drinking in rural areas, and how this may be reflected in statistical analysis.
2.4 - Overburden
Overburden influences groundwater contamination by moderating the paths by which water can travel over land and through the subsurface, and will generally provide more security from groundwater contamination when a unit is thicker (Levison and Novakowski, 2009). Overburden is a general term to describe the unconsolidated rock and soil overlying bedrock. The type of overburden regulates the rate at which water can enter the subsurface and drain through the pathways of soil pores to enter the zone of saturation (Bache, 2006; Marechal et al, 2007). The depth of overburden will generally influence the fraction of bacterial contamination that enters the groundwater, with contaminants decreasing in a gradient towards bedrock (Gerba, 1975), assuming there are no preferential pathways for water to flow. Shallow soil profiles generally provide less opportunity for bacterial contaminants to interact with the mitigating filtering effects of soil (Levison and Novakowski, 2009; Borchardt et al. 2007). The residence time of water and the mitigating effects of soil will vary greatly with the properties of the soil (Simin et al., 2013; Bachmann et al., 2006), meaning that it is logical to expect this would be reflected in the observed levels of bacteria in water with transportation from a source given bacterial die off rates.

Texture
The ability of soil to hold water when drained by gravity is known as its field capacity. Field capacity is greatly affected by a soil’s bulk density; i.e., the mass of soil of a given volume as a function of soil texture (Gardiner and Miller, 2008). The size of soil particles will determine its bulk density and the size of pores which water can then occupy. Finer soil texture from smaller soil particles will result in smaller pores, giving better filtration for fecal contaminants by reducing connectivity, and increasing surface area and surface tension (Gardiner and Miller, 2008; Conboy, 2000). Coarse grained overburden will generally have much larger pore spaces that drain quicker (provided the soil is not poorly sorted), allowing for the quick mobilization of fecal contaminants. Organic matter content will also have a strong modifying influence by forming aggregates that amend soil structure to allow more water to be held (Scott, 2016). However, Mubiru et al. (2000) state that finer soil texture facilitate the accumulation of organic matter, thereby increasing survival rates of fecal bacteria.

Horizons
Overburden profiles of soil and unconsolidated sediments may be characterized by one or more horizons in their overburden profiles. These horizons may have distinct properties that could influence the survivability and transportation of bacterial contaminants in groundwater. ‘A’
horizons are surface soils rich in organic matter vulnerable to erosion, and are eluvial due to constant flushing by overland flow. ‘B’ horizons are characterized by lower organic matter, and higher clay content from the flushing of the ‘A’ horizon. ‘C’ horizons are deep in ground and are relatively unaffected by the migration of materials above (Gardiner and Miller, 2008). Therefore, overburden cannot be treated as a homogenous medium, and horizons will have different biological, chemical and physical properties that could potentially produce variations in the survivability and transportation of bacteria.

While the literature explicitly exploring this proposition is limited, a useful example is a study by Dejonghe from 2000. Dejonghe evaluated the applicability of bioaugmentation of plasmids by bacteria, and observed different assemblages in indigenous bacteria of two soil horizons from the same profile. This was attributed to differences in organic matter and microbial communities between horizons (Dejonghe, 2000). While this study does not evaluate the survivability of coliform bacteria in horizons, it demonstrates that it would be logical to expect differences in how well they survive in different soil horizons.

Soil texture variations among horizons of a vertical profile will create differences in hydraulic conductivity between units, meaning that residence time in different parts of a profile could be experienced. Empirical relationships demonstrate that reduced hydraulic conductivity of a profile is driven by thicker units and those with lower hydraulic conductivity, as expressed by Equation 1. Equation 1 uses the sum of each unit’s thickness divided by the sum of products of each unit’s thickness and hydraulic conductivity (Reynolds, 2017).

\[
K = \frac{\sum b_i}{\sum \frac{b_i}{K_i}}
\]

Equation 1 – Calculation of Vertical Hydraulic Conductivity for Multiple Units using the thickness of each unit in stratigraphic succession (b) and the hydraulic conductivity of each unit (K).

**Bacteria Survival Rates in Overburden**

When water is contaminated with bacteria and suspended in porous media, bacteria will survive for days (Franz et al, 2014). Soils can limit transportation of bacteria by removing organisms through adsorption to grains. The strength of this impact will depend on water
chemistry and the adsorptive forces of the grain (Borchardt et al., 2007). As an example it is estimated that enteric (originating in the intestine) bacteria can survive for 2-4 months in soil and groundwater. Survival times could be extended locally in cases where there are continual bacterial inputs (Jamieson et al., 2002). An estimate by Rogers et al. (2011) using quantitative PCR in silt loam and silty-clay loam inoculated with swine or cow manure slurry serves as a second example. E. coli was recovered anywhere from 41 to 120 days during trials with cow manure when kept at 10°C. This has been consistent with other cited studies like that by Reddy et al. (1981), who evaluated soils treated with organic wastes. Therefore, there it appears that there is a large window of time in which bacterial contaminants can survive in overburden and potentially be transported through overburden to a private well.

**Moisture Content**

Moisture content is believed to be a strong factor in the survival of fecal bacteria in overburden, but is often evaluated by proxy of soil texture and organic matter as factors influencing how moisture is retained in soil pores (Reddy et al., 1981). Mubiru observed in an experiment that E. coli had lower mortality rates in soils of a finer texture with richer organic matter, noting that this and other studies attributed results to the availability of water (Mubiru et al., 2000). A study by Reddy et al. (1981) observed that E. coli in a Weld fine sandy loam had a half-life range from 37.8 hours to 57.4 hours when moisture content ranged from 10% to 50%, respectively, demonstrating that when all other variables are controlled, moisture content does improve the survival of bacteria. Experiments did not use moisture content of greater than 50%, though is it predicted that survival rates may decrease in water logged soil due to a decreased availability of oxygen.

**Using Overburden to Evaluate Groundwater Vulnerability**

The vulnerability of an aquifer to contamination may be expressed using the DRASTIC method, a common method of classifying aquifers for vertical vulnerability (Ersoy and Gultekin, 2013). Weighting factors for the method are the depth to water, the net recharge, the aquifer media, soil media, topography slope, impact of vadose zone, and hydraulic conductivity. The DRASTIC method is a technique of classifying aquifers for vertical vulnerability; horizontal flow is not directly captured, although topography is a factor (Ersoy and Gultekin, 2013). Methods have also been developed to classify distinct types of aquifers, such as those with karst topography (Hammouri and El-Naqa, 2008). The DRASTIC method is well correlated with regional distributions of nitrate, total coliform and fecal coliform (Jamrah et al., 2007; Hammouri and El-Naqa, 2008).
overlap in some variables between DRASTIC and those affecting survival of bacteria, creates an opportunity for further development. This method demonstrates the importance of an overburden profile in securing an aquifer from contamination via vertical migration.

A second example of a vulnerability assessment for groundwater relevant to the study area of Southern Ontario is the EPIK method. EPIK is a method for quantifying the vulnerability of karstic environments. Karst topography occurs during the chemical weathering of carbonate units common in southern Ontario, creating channels for preferential flow (White, 2002). These large openings in rock could serve as macropores allowing for transportation of bacteria. These environments provide a unique challenge for quantifying vulnerability, as such a preferential flow network would not be found in other bedrock systems (White, 2002). The EPIK system uses epikarst, protective cover, infiltration conditions and karst network development as quantitative variables to assess vulnerability (Gogu and Dassargues, 2000). This method still uses overburden as an important consideration, but recognizes that, once water infiltrates, it will be subject to a wide range of other conditions.

**Conclusions about Overburden**

Soil physical properties may increase the residence of contaminated water (rather than draining to the zone of saturation), but in certain cases this may also facilitate the survival of bacteria. Further research may be required to address this complex issue, and it is difficult to distinguish whether soil properties had a mitigating affect on enteric bacteria, or if the soil simply retained water long enough for bacteria to die off. Other factors, outside of the scope of this particular review, but that are important to consider include: interactions with microorganisms (Semenov et al., 2010), a high pH, the presence of cations and their exchange capacity, and low flow rate, all of which may work to reduce bacterial contamination in soil (Gerba, 1975; Jamieson et al., 2002). In analyzing rates of bacterial contamination in the well water of southern Ontario, overburden depth is considered an important factor that can reduce the effects of contaminated water by increasing residence time of water, reducing the survival rates of bacteria, and by acting as a barrier to private wells.

**2.5 Surficial Geology**

Surficial geology refers to geological deposits less than 1.81 million years old, although it is noted that deposits in Ontario are often less than 45,000 years old due to the timing of the last ice
Surficial deposits will be present below topsoil, and are a key component to understanding how groundwater flows and becomes contaminated. Surficial deposits are unconsolidated geological material that has been physically and chemically weathered during transit to a site of deposition, or formed in place. Properties of surficial deposits are controlled by their depositional environments. Understanding these environments will provide insight into how surficial deposits may influence the vulnerability of groundwater to fecal contamination. As acknowledged by Goss et al, surficial soils in southern Ontario reflect parent deposits, and are highly variable across the region. They note that many rural wells are drawn directly from glacial overburden sediments rather than bedrock units (1998).

**Glacial Till**

Glacial till is a surficial deposit associated with the mass movement of glaciers that carried sediment and scouring the landscape. These sediments are generally silt sized in Ontario, meaning that they are capable of transmitting groundwater flow assuming compacted clay is not present. Within this grain size, mineralogy can vary across the region and have been known to influence groundwater chemistry. Because these deposits are highly heterogeneous, they vary spatially, making them difficult to evaluate as a factor in groundwater transportation (Hinton et al., 1994). Hinton et al. also note that tills exhibit a wide range of horizontal hydraulic conductivities ($2.6 \times 10^{-5}$ m/s to $1.8 \times 10^{-9}$ m/s), meaning that any attempt to characterize glacial till for the purpose of evaluating regional dynamics of bacterial contamination in private wells will have broad assumptions. While glacial till is heterogeneous in nature, in can still have characteristic parent material with unique properties such as clast fabric (Menzies, 2006) that separates it from other forms of till in southern Ontario.

**Lacustrine Sediments**

Lacustrine sediments are those that accumulate in lake bottoms. Glacio-lacustrine sediments were deposited by the melting of a glacier into a lake through direct deposition or meltwater inflows (VIU, 2012). When a lake drains lacustrine sediments are left behind. The nature of the sediment physical and chemical properties will depend on the origin of sediment, water chemistry, lake trophic state, catchment size, interactions with groundwater and regional climate (Schnurrenber et al, 2003). These deposits are likely to be characterized by finer texture and improved sorting within an individual horizon (Narbonne, 2017). Lacustrine sediments will be laterally deposited in layers associated with settling in a lake, meaning that in areas of lacustrine surficial geology, multiple homogenous layers with different properties may be encountered rather
than a heterogeneous mixture. As an example, Rodvang and Simpkins (2001) found in review that lacustrine sediments had lower hydraulic conductivity than till. Many of these deposits are young and have not been weathered yet, making them effective aquitards that lowered rates of nitrate contamination in groundwater.

**Glacio-marine sediments**

Glacio-marine sediments are those deposited by inland seas present during glaciation due to isostatic depression. When these seas retreated, sediments were deposited. These sediments will exhibit characteristics associated with marine environments and salt water/fresh water interactions that do not reflect the current physical geography (Harrap, 2015; VIU, 2012). Glaciomarine sediments are often characterized as being dominated by marine clays and silts that deposit during marine recession (Albuquerque, 2007). In addition, these environments are often more turbid than lacustrine environments with sediment redeposition occurring through mass movements, resulting in debris and turbidite deposits mixing with primary glaciomarine sediments. Within these heterogeneous sediments, gradients of particle size, deposit rates and organic content can be observed with increasing distance from the glacier, with sediment accumulating most rapidly near the glacier (Gornitz, 2008). Despite these complex environments, Albuquerque states that glaciomarine sediments dominated by marine silt and clay often act as confining layers (Albuquerque, 2007), potentially preventing groundwater from being contaminated by an overlying source of bacterial contamination.

**Glacio-fluvial sediments**

Glacio-fluvial sediments are those that originate proximal to a melting glacier. This form of sedimentation is dominated by the distance material is deposited from the glacier. Sediments nearest the glacier are the coarsest grains associated with turbulent deposition, often leading to braided channeling. Grain size will decrease further downstream from the glacier of origin to eventually behave like a conventional river system (VIU, 2012). These surficial deposits can facilitate or mitigate the transport of fecal contamination (in particular through the formation of paleochannels) (Reynolds, 2017). Paleochannels of coarse-grained material may create preferential flow, rapidly transporting bacterial contaminated water between less conductive units.

As an example, Levy et al. (2007) found that by obtaining unconsolidated sediment cores from a glacio-fluvial outwash plan and running flow interruption experiments, the degree of heterogeneity was a direct control of how bacteria were transported. Bacteria transport varied greatly among the cores due to differences in heterogeneity. More bacteria were recovered in cores
with larger median grain size, and peak bacteria concentrations arrived more quickly in cores that were more poorly sorted. Therefore, understanding the glacio-fluvial setting of a well can provide insight into how vulnerable it may be to contamination by bacteria.

**Conclusions about Surficial Geology**

In post-glacial environments like southern Ontario, understanding the wide range of depositional environments is important in order to evaluate the unique environments for groundwater. In sediments directly overlying bedrock, the character of sediments is reflective of their depositional environment. Because of their location, it is logical to expect that they would often contain the water table, making them a key component of hydrological systems that could influence how bacteria are transported from a source to a private well in unconsolidated systems. Literature directly exploring the importance of surficial deposits was limited, and was often contained to brief descriptions in site descriptions. While overburden can be characterized without interpreting historical depositional environments, surficial deposits are important to understand given the differences from topsoil. Because of the size of these units, surficial deposits are an important component in understanding the bacterial contamination of groundwater beyond local scales.

**2.6 - Bedrock Geology**

Bedrock is formed by processes that produce a wide range of rock types, each of which creates a unique environment for groundwater. The composition and structure of bedrock will influence how contaminated water is transported once it travels through the overburden to the bedrock. The properties of bedrock that will most strongly influence the extent of bacterial contamination are pore space and fracturing, the two forms of void space in bedrock that can provide a vector for water. Several sedimentary and metamorphic rock types are found throughout the study area and their physical and chemical properties will be discussed in detail below in relation to their ability to prevent the spread of fecal contamination in groundwater.

**Limestone and Dolostone Aquifers**

Limestone is composed of calcium carbonate and the product of chemical weathering prior to formation. Dolomite may also be present depending on the presence of warm salty water during the deposition environment, resulting in the addition of magnesium carbonate (Tindall, 1975).
Carbonate rock units have chemical properties that create vulnerable hydrological environments known as karst topography. Freeze and Cherry list 1.00e-6 to 1.00e-9 m/s as the hydraulic conductivity of limestone, rising to 1.00e-2 to 1.00 1.00e-6 m/s in karst limestone (1979). The relatively low hydraulic conductivity of limestone can be attributed to the crystalline formation of calcium carbonate, creating tightly held grains with very little porosity. Carbonates may prevent flow of fecal contaminated groundwater, until it is chemically weathered. Limestone karst topography is variable and presents challenges in evaluating how water may be transported in limestone bedrock.

Fractures in soluble carbonate rock units can be chemically weathered during the movement of ground water, resulting in the creation of large void spaces in bedrock units that characterize karst. Erosion of pore spaces and fractures can create larger pathways of preferential flow that will increase the transport of groundwater, potentially transporting coliform bacteria. Karst environments are heterogenous in nature and dependent on the organization of underground flow routes (Lubianetzky, 2014). When interpreting the effects of bedrock on rates of bacterial contamination across southern Ontario, karst topography must be recognized as a factor that may introduce variability.

An investigation by the Ontario Geological Survey and Golder Associates established that karst topography is found throughout southern Ontario, but that this topography is generally limited to areas of limestone and dolostone that have less than a meter of overburden. This was attributed to the parent material of overburden containing carbonates, meaning that water typically exceeds its capacity to dissolve ions before it reaches bedrock (Gagnon-Nandram, 2016). These results suggest vulnerability of groundwater to bacterial contamination in areas of shallow overburden and limestone bedrock, the most common form of bedrock in the study area. This has also been observed in other regions, where the soil profile overlying vulnerable limestone aquifers plays a large role in the transport of bacteria from the surface to groundwater (Celico et al., 2004, Conboy and Goss, 2000). Therefore, in evaluating the effects of bedrock on rates of bacterial contamination in Southern Ontario, limestone bedrock should be considered a key contributor when accompanied by shallow overburden. Multivariate statistics would be required to evaluate the co-occurrence of limestone or dolostone, thin overburden and bacterial contaminants using the data from this study.
**Shale**

Shale is formed from silt and clay deposits, often after marine deposition or during the overtopping of banks in fluvial environments. This process of compressing very fine-grained sediments yields layers of sedimentary rock that are virtually impermeable to flow, with porosity rarely exceeding 2%. Shale is effective as an aquitard, and will securing confined groundwater from bacterial contamination (Freeze and Cherry, 1979). However as discussed below, studies have found that shale aquitards assumed to be impermeable may still be fractured.

During a survey of rural groundwater quality using 300 wells from southern Ontario, Conboy and Goss compared regional soil maps, physiography, and geology with well records to evaluate the influences on groundwater quality. Using binomial logistical regressions, it was found that wells drilled into shale units fell within the largest group of low risk well, and not a single high risk well was reported as having shale present, likely due to its ability to act as an aquitard. Conboy and Goss also note that shale in southern Ontario is younger than those in other regions, typically originating in the mid-Devonian. These shale units are typically thicker and more continuous than those elsewhere in Canada (Conboy and Goss, 2000). This study proves that shale units in southern Ontario will generally prevent the spread of bacterial contamination to private wells. When evaluating the influence of bedrock on fecal contamination, this study indicates that private wells with low or no contamination may correlate with protection by a shale unit.

**Sandstone and Arkose Aquifers**

Sandstone is a sedimentary rock of composed of consolidated coarse siliciclastics. The size of sand grains (> 0.0625mm) result in greater pore space between the grains relative to other rock types as a result of how grains become oriented. Cementation and matrix that forms between these grains, as well as the level of sorting in grains prior to consolidation will influence the amount of porosity beyond the texture of the clasts. Porosity can vary greatly from 2% in extreme cases of cementation and can reach as high as 22%. Based on figures provided by Freeze and Cherry, porosity is most commonly between 8 – 12 % (1979). This property is of hydrological significance, and allows sandstone to form regional aquifers that hold large amounts of water. When evaluating the effects of sandstone on the transport of bacterial contaminated water, porosity should be understood as a primary control. Sandstone units were the most common type of unit in southern Ontario, excluding carbonates and shale, and were often accompanied by arkose, a variation of sandstone rich in feldspar. Sandstone plays an important role in the occurrence of bacterial
contamination in rural water. Given the porosity of these units, sandstone presents a potential vector for contamination once bacteria travel through overburden to reach bedrock, but spread will be limited by the hydraulic conductivity of the unit.

During a large-scale estimate of hydraulic conductivity in sandstone aquifers by the USGS for the Dakota Aquifer using 494 wells, it was estimated that mean hydraulic conductivity averaged 1.951e-5 meters per second (approximately 1.65m/day). Here, sandstone originated as Paleozoic sedimentary deposits (Bredehoeft et al., 1983). Similarly, sandstone in Southern Ontario originates in the Lower to Middle Paleozoic era (University of Waterloo, n.d.), meaning that this may prove as a useful estimate given similarities in age (and thus time allotted for weathering). Freeze and Cherry list 1.00e-6 – 1.00e-10 m/s as a range of possible hydraulic conductivity values (1979).

During a study by Krapac et al., the impacts of deep pit manure storage systems through monthly groundwater sampling were studied in areas of different geological vulnerability after cracks in the concrete walls were observed. The first site had 6m of silt and clay overlying 7-36m of shale and more than 30m of limestone. The second site had 6m of silt diamict overlying sandstone used as an active water resource. Both sites were proximal to a manure storage system as well as fields where manure was spread. When compared to background levels, groundwater in sandstone near the manure storage and as well as at a house 800m away had comparable elevated levels of nitrate, whereas groundwater at the shale site was not impacted (Krapac et al., 2002). This demonstrates that sandstone is a potential vector for bacteria.

**Fracturing**

Geological units of all types may experience structural effects such as jointing and fracturing that will create preferred paths of travel for groundwater and a vector for bacterial contamination (Conboy and Goss, 2000). The direction and flow rate that is permitted by joints and fractures will control how contamination is transported (Allen and Morrison, 1973). It is estimated that fractured rock typically accounts for a total porosity of 2 to 5%. If it is a crystalline rock, this may comprise effective porosity (Reynolds, 2017). Fracturing will typically occur when a rock unit is under compression or tension, and could be associated with weathering of a cleavage surface, tectonic activity, freezing and thawing of pore water, or water pressure in rock slopes (Hungr and Evans, 1989). The consequence of fracturing is that pathways for the preferential flow of water are created, meaning that once contaminated water enters fractured bedrock, it can travel in a route
that is hard to predict. The extent of flow in bedrock fractures is a function of fracture network density, geometry and connectivity (Gabrielli et al., 2012). Characterizing the vulnerability of these systems can be complicated due to complex structures and heterogeneity, requiring knowledge of the extent of fractures and their connectivity, and the hydraulic conductivity of units connected by fractures (Lubianetzky, 2014). As noted by Personne et al., this level of detail typically isn’t available due to the natural complexity of fissured environments (1998).

The nature of bedrock fractures will impact the extent which contaminated groundwater can travel. It is often assumed that confined aquifers are impermeable given their confining units, but fractures can threaten the integrity of these aquitards. During a study by Borchardt et al. in which 30 wells were tested that drew from an aquifer confined by a shale aquitard, 7 wells tested positive for enteric viruses, They attribute this to fractures by postulating that viruses would have to be rapidly transported to survive in the aquifer, and are small enough to travel through these pathways. This study replicated results observed in a larger study by Abbaszadega (2003), demonstrating that confined systems long assumed to be impermeable may still experience contaminant inputs. (2007).

Unlike properties like porosity, grain size, and sorting, degree of fracturing is difficult to evaluate with regional data. Fracturing will be a large source of uncertainty when assessing the rates of bacterial contamination in private wells. Fracturing cannot be accounted for in the statistical analysis using bedrock type, but remains an important local consideration.

**Porosity**

In sedimentary rocks, porosity is an important vector for groundwater that will influence how bacteria are transported. The effective porosity of bedrock is the porosity available for fluid flow (Reynolds, 2017). Primary porosity describes the natural characteristic of a geological unit to allow water to move prior to chemical weathering. Secondary porosity refers to the pores that result from chemical leaching of minerals or the creation of a fracture via fluid movement, and accounts for the most fluid movement in rock. This is because despite porosity of sedimentary rocks like sandstones, compaction and cementation greatly reduce how pores are interconnected. (Gilbert, n.d.). Water will only move through bedrock when pore spaces are large enough to support connected movement. The speed of groundwater movement is a function of the size of the pores and pressure gradient (Swanson, n.d.). Without the influence of fractures, the ability of a rock unit
to conduct water will control how far contaminated water may travel from the source during the survival period of bacteria to become drawn into a rural drinking water well. Therefore, rock type and subsequent porosity and hydraulic conductivity may be understood as a primary control on the transport of bacteria in groundwater, with fractures introducing variability and faster modes of transport.

2.7 - Conclusions

Bacterial contamination of groundwater is a complicated issue that transcends many anthropogenic and natural processes. Bacterial contamination of groundwater is an important environmental issue and an excellent example of how natural and human systems are interconnected. The lack of centrally distributed water in rural areas creates an inherent dependence on ecosystem services of filtration that provide clean groundwater. By evaluating how anthropogenic and natural sources of bacteria interact with a geophysical setting to contaminate groundwater, a greater understanding of the hydrology of these areas is gained. This is important information for officials concerned with the health of rural communities and users of private wells. This information will allow stakeholders to better understand the dynamics of bacterial contamination in private wells, and provides indications of which environments are most vulnerable to contamination by coliform bacteria.
3 - Methodology

To investigate how observed bacterial contaminants (total coliform and \textit{E. coli}) at a concentration (colony forming units per volume of water) in private wells were related to the geophysical setting at the location of the well, spatial data describing overburden depth, soil texture, bedrock classification and surficial deposit classification were analyzed using statistics and GIS. The purpose of the statistical analysis was to quantitatively investigate the strength of the relationship between the concentration of bacteria observed in drinking water wells and the geophysical setting of the well. This statistical analysis used total coliform and \textit{E. coli} contamination data from private drinking water wells at sampling locations as a dependent variable, and one of the four geophysical datasets regarding environmental setting as the independent variable. Data was analyzed with either a Spearman’s Correlation or One-Way Analysis of Variance (ANOVA).

3.1 - Rural Drinking Water Wells

A point class shapefile was obtained from Public Health Ontario (Krolik \textit{et al.}, 2013). The shapefile contained approximately 47,000 data elements, each of which described the location where a well water sample was taken, and the results of laboratory analysis for \textit{E. coli} and total coliform bacteria present in the given sample. Samples were given to Public Health Ontario Laboratories during 2012 by households. The purpose of these sample submissions was for households to receive indication of the quality of their privately-sourced drinking water from Public Health Ontario, as evaluated through microbiological lab analysis. This dataset covered a large area of Southern Ontario, as illustrated in Figure 1. Wells were geocoded using a civic address. Point elements do not represent the exact location from which the water was drawn. This is a limitation of the dataset when using it to evaluate a relationship between the contamination of a well and the environmental setting. This dataset originally contained 47048 records. The final dataset of rural drinking water wells suitable for statistical analysis contained 37692 records after a data cleaning process (3.7). The target attribute fields for statistical analysis from this dataset were total coliform and \textit{E. coli} count fields. These fields contain integers from 0-80 that signified the number of colony forming units per volume of water. This count was obtained by culturing a submitted drinking water sample. These fields were continuous variables, and would serve as the dependent variable during two rounds of statistical analyses.
3.2 - Bedrock Geology

A polygon shapefile describing bedrock geology was obtained from the Ontario Geological Survey (2011). The attribute data associated with the polygons of the geological units provided information about the classification and age of the units. The field of attribute data selected for analysis was a rock description field that named rock units in a standardized manner by geological surveyors. Unique classifiers of bedrock were translated to unique integers mapped in a legend using find and replace functions of spreadsheet software (3.7 – Data Cleaning). This set of integers was classified as a categorical variable because it did not represent a continuous phenomenon or categories in a given order. This dataset was then used as an independent variable during statistical analysis. A map of the data may be viewed in Figure 2. Using the spatial join described in section 3.6, the point shapefile of wells was found to overlap different types of bedrock units at frequencies supplied in Table 1.
Figure 2 – Map of the bedrock geology dataset used for spatial join and statistical analysis
Table 1 – Frequency of bedrock units underlying a private well in the target dataset

<table>
<thead>
<tr>
<th>Bedrock Unit Underlying a Private Well</th>
<th>Frequency</th>
<th>Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amphibolite, gabbro, diorite, mafic gneisses</td>
<td>11</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Commonly layered biotite gneisses and migmatises</td>
<td>314</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>Conglomerate, sandstone, shale, dolostone</td>
<td>518</td>
<td>1.4</td>
<td>2.2</td>
</tr>
<tr>
<td>Conglomerate, wacke, quartz arenite, arkose, limestone, siltstone, chert, minor</td>
<td>770</td>
<td>2.0</td>
<td>4.3</td>
</tr>
<tr>
<td>Diorite, gabbro, peridotite, pyroxenite, anorthosite, derived metamorphic rocks</td>
<td>168</td>
<td>0.4</td>
<td>4.7</td>
</tr>
<tr>
<td>Dolostone, sandstone</td>
<td>3056</td>
<td>8.1</td>
<td>12.8</td>
</tr>
<tr>
<td>Flows, tuffs, breccias, minor iron formation, minor metasedimentary rocks</td>
<td>354</td>
<td>0.9</td>
<td>13.8</td>
</tr>
<tr>
<td>Granitic and syenitic gneisses</td>
<td>67</td>
<td>0.2</td>
<td>13.9</td>
</tr>
<tr>
<td>Granitic gneisses with metasedimentary xenoliths, migmatises, injection gneisses,</td>
<td>345</td>
<td>0.9</td>
<td>14.9</td>
</tr>
<tr>
<td>Granodiorite, granite, syenite, pegmatite, alkalic granite, migmatitic gneisses</td>
<td>72</td>
<td>0.2</td>
<td>15.1</td>
</tr>
<tr>
<td>Granodiorite, tonalite, monzogranite, syenogranite; derived gneisses and</td>
<td>495</td>
<td>1.3</td>
<td>16.4</td>
</tr>
<tr>
<td>Limestone, dolostone, shale</td>
<td>3207</td>
<td>8.5</td>
<td>24.9</td>
</tr>
<tr>
<td>Limestone, dolostone, shale, arkose, sandstone</td>
<td>15513</td>
<td>41.2</td>
<td>66.0</td>
</tr>
<tr>
<td>Limestone, dolostone, shale, sandstone, gypsum, salt</td>
<td>1380</td>
<td>3.7</td>
<td>69.7</td>
</tr>
<tr>
<td>Marble, calc-silicate rocks, skarn, tectonic breccias</td>
<td>1390</td>
<td>3.7</td>
<td>73.4</td>
</tr>
<tr>
<td>Nepheline syenite, alkalic syenite, fenite; associated mafic, ultramafic and</td>
<td>90</td>
<td>0.2</td>
<td>73.6</td>
</tr>
<tr>
<td>Quartzofeldspathic gneisses, pelitic to semi-pelitic gneisses, calc-silicate gneisses</td>
<td>18</td>
<td>0.0</td>
<td>73.7</td>
</tr>
<tr>
<td>Sandstone, dolostone, limestone</td>
<td>259</td>
<td>0.7</td>
<td>74.4</td>
</tr>
<tr>
<td>Sandstone, shale, dolostone, siltstone</td>
<td>5167</td>
<td>13.7</td>
<td>88.1</td>
</tr>
<tr>
<td>Shale</td>
<td>344</td>
<td>0.9</td>
<td>89.0</td>
</tr>
<tr>
<td>Shale, limestone, dolostone, siltstone</td>
<td>3368</td>
<td>8.9</td>
<td>97.9</td>
</tr>
<tr>
<td>Tectonites, straight gneisses, porphyroclastic gneisses, unsubdivided gneisses in</td>
<td>183</td>
<td>0.5</td>
<td>98.4</td>
</tr>
<tr>
<td>Tonalite, granodiorite, monzonite, granite, syenite; derived gneisses</td>
<td>604</td>
<td>1.6</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>37693</strong></td>
<td><strong>100.0</strong></td>
<td></td>
</tr>
</tbody>
</table>
3.3 - Surficial Geology

A polygon shapefile describing surficial geology was obtained from the Ontario Geological Survey (2007). The attribute field that described depositional environment and sediment type in a standardized process by geological survey was selected for analysis. This was translated to integers mapped in a legend using find and replace functions. This set of integers was classified as a categorical variable because it did not represent a continuous phenomenon or categories in a given order. This dataset was then used as an independent variable during statistical analysis. A map of the dataset may be viewed in Figure 3. The frequency distribution of units underlying the private well dataset may be viewed in Table 2.

Figure 3 – Surficial Geology dataset used for spatial join and statistical analysis
Table 2 – Frequency of surficial units underlying a private well in the target dataset

<table>
<thead>
<tr>
<th>Surficial Unit Underlying a Private Well</th>
<th>Frequency</th>
<th>Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bedrock</td>
<td>6363</td>
<td>16.9</td>
<td>16.9</td>
</tr>
<tr>
<td>Fluvial deposits</td>
<td>514</td>
<td>1.4</td>
<td>18.2</td>
</tr>
<tr>
<td>Glaciofluvial ice-contact deposits</td>
<td>2354</td>
<td>6.2</td>
<td>24.5</td>
</tr>
<tr>
<td>Glaciofluvial outwash deposits</td>
<td>2182</td>
<td>5.8</td>
<td>30.3</td>
</tr>
<tr>
<td>Glaciolacustrine deposits</td>
<td>7932</td>
<td>21.0</td>
<td>51.3</td>
</tr>
<tr>
<td>Glaciomarine and marine deposits</td>
<td>3374</td>
<td>9.0</td>
<td>60.3</td>
</tr>
<tr>
<td>Lacustrine deposits</td>
<td>1173</td>
<td>3.1</td>
<td>63.4</td>
</tr>
<tr>
<td>Organic deposits</td>
<td>574</td>
<td>1.5</td>
<td>64.9</td>
</tr>
<tr>
<td>Till</td>
<td>13227</td>
<td>35.1</td>
<td>100.0</td>
</tr>
</tbody>
</table>

3.4 - Overburden Depth

A point shapefile that describes overburden depth in southern Ontario was obtained from the Ontario Ministry of Mining and Northern Development. The shapefile was created by compiling well records, geotechnical borehole records, oil and gas well records, and existing publications (Gao et al. 2006). Each point represented an observation location, and gives a measurement of the depth to bedrock in meters and the method used in measurement. The depth to bedrock measurement field was selected from this dataset for statistical analysis as a continuous independent variable. A map of the dataset may be viewed in Figure 4. Unfortunately, data about the soil horizons was not provided. For this reason, overburden depth will be evaluated as a variable following the observations by Gerba et al. (1975) that fecal contamination will decrease in a soil profile with depth, but with the recognition that soil texture in horizons would significantly change the effects of contamination mitigation (Gardiner and Miller, 2008).
3.5 - Soil Survey Complex

A polygon shapefile mapping soils was obtained from the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA, 2011). This shapefile covered southern Ontario, with the exception of a portion of the Algonquin Park Region. This dataset was produced from the conversion of physical maps drafted after the exhaustive survey of the top soil southern Ontario with the purpose of evaluating agricultural suitability. This shapefile provided detailed information about the classification, composition, stoniness and texture of soils. A data field describing the soil texture as a function of a soil grain size distribution was used as the target for statistical analysis. This independent variable is categorical. Converting these categories to a continuous variable using grain size was initially proposed. However, poorly sorted soils (in particular loams) would not be
well represented using a measure such as median grain size, as well as gleysolic soils. It was decided that grain size as a continuous variable was not representative of the qualitative differences among soils of similar median grain sizes (i.e: well sorted silt and very poorly sorted silt), and the decision was made to use categorical data. A map of the dataset may be viewed in Figure 5. The frequency distribution of units underlying the private well dataset may be viewed in Table 3.

Figure 5 – Map of the soil classification dataset used for spatial join and statistical analysis.

Soil Texture of Southern Ontario

Legend

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Clay</td>
<td>Fine Sand</td>
<td>Fine Sandy Loam</td>
<td>Loam</td>
<td>Loamy Sand</td>
<td>Loamy Very Fine Sand</td>
<td>Silty Clay</td>
<td>Silty Clay Loam</td>
<td>Bedrock</td>
<td>Organic Deposits</td>
</tr>
<tr>
<td>Clay Loam</td>
<td>Clay Loam</td>
<td>Clay Loam</td>
<td>Clay Loam</td>
<td>Clay Loam</td>
<td>Clay Loam</td>
<td>Clay Loam</td>
<td>Clay Loam</td>
<td>Organic Deposits</td>
<td>Organic Deposits</td>
</tr>
<tr>
<td>Clayey Sandy Loam</td>
<td>Clayey Sandy Loam</td>
<td>Clayey Sandy Loam</td>
<td>Clayey Sandy Loam</td>
<td>Clayey Sandy Loam</td>
<td>Clayey Sandy Loam</td>
<td>Clayey Sandy Loam</td>
<td>Clayey Sandy Loam</td>
<td>Organic Deposits</td>
<td>Organic Deposits</td>
</tr>
</tbody>
</table>

Bedrock

Organic Deposits

Gleysolic Loam

Gleysolic Sand

Gleysolic Sandy Loam
Table 3 - Frequency of soil textures underlying a private well in the target dataset

<table>
<thead>
<tr>
<th>Soil Texture at a Private Well</th>
<th>Frequency</th>
<th>Percent</th>
<th>Cumulative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Fine Sand Loam</td>
<td>169</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Clay</td>
<td>1541</td>
<td>4.1</td>
<td>4.5</td>
</tr>
<tr>
<td>Clay Loam</td>
<td>4169</td>
<td>11.1</td>
<td>15.6</td>
</tr>
<tr>
<td>Clayey Sand Loam</td>
<td>3</td>
<td>0.0</td>
<td>15.6</td>
</tr>
<tr>
<td>Fine Sand</td>
<td>537</td>
<td>1.4</td>
<td>17.0</td>
</tr>
<tr>
<td>Gleysolic Loam</td>
<td>220</td>
<td>0.6</td>
<td>17.6</td>
</tr>
<tr>
<td>Gleysolic Sandy Loam</td>
<td>992</td>
<td>2.6</td>
<td>21.2</td>
</tr>
<tr>
<td>Loam</td>
<td>11407</td>
<td>30.3</td>
<td>51.5</td>
</tr>
<tr>
<td>Loamy Fine Sand</td>
<td>669</td>
<td>1.8</td>
<td>52.3</td>
</tr>
<tr>
<td>Loamy Sand</td>
<td>1876</td>
<td>5.0</td>
<td>57.3</td>
</tr>
<tr>
<td>Loamy Very Fine Sandy</td>
<td>64</td>
<td>0.2</td>
<td>57.5</td>
</tr>
<tr>
<td>Organic Matter</td>
<td>1302</td>
<td>3.5</td>
<td>61.0</td>
</tr>
<tr>
<td>Bedrock</td>
<td>946</td>
<td>2.5</td>
<td>63.5</td>
</tr>
<tr>
<td>Sand</td>
<td>1540</td>
<td>4.1</td>
<td>67.6</td>
</tr>
<tr>
<td>Silty Clay</td>
<td>228</td>
<td>0.6</td>
<td>68.2</td>
</tr>
<tr>
<td>Silty Clay Loam</td>
<td>1392</td>
<td>3.7</td>
<td>71.9</td>
</tr>
<tr>
<td>Silt Loam</td>
<td>2343</td>
<td>6.2</td>
<td>78.1</td>
</tr>
<tr>
<td>Sand Loam</td>
<td>8295</td>
<td>22.0</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>37693</strong></td>
<td><strong>100.0</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

3.6 - Spatial Join

The shapefile of well records was used to perform spatial joins with the four additional shapefiles described above. To accomplish this ArcMap (ArcGIS 10.3.1, Esri) was used to overlay the well records shapefile with the bedrock geology, surficial geology, soil texture and overburden depth shapefiles. Before performing a spatial join, each shapefile was projected into the World Geodetic System 1984 geographic coordinate system using the Project Tool in the Data Management toolset. The attribute information associated with a polygon that contained a given well record point was then appended to each well record point in the shapefile using a spatial join. A spatial join is a technique used to append records of attributes tables to one another, where the record is defined a spatial location. When performing the spatial join between the well shapefile and the three polygon shapefiles, the specification was made that well points must intersect the polygon. Instances where this did not occur yielded null values in the output shapefile subsequently removed during data cleaning. The spatial join was performed using one to one cardinality. The overburden depth point shapefile could not be joined to the well records using the above method because two point datasets cannot overlap or contain one another. The overburden depth dataset came with a 500m resolution raster. It was proposed that the raster be joined to well points using
the cell each point intersected as an alternative to performing a spatial join between two point datasets. However, it was found that the distances preserved in the spatial join were much smaller than 500m. This lead to the conclusions that depth measurements are likely more representative when the interpolated raster was not used. Instead, a spatial join was performed between nearest points of overburden depth and well points without intersection. This resulted in the closest overburden measurement being appended to a given well record. This distance measurement was recorded in the attribute data to identify any extreme distances between wells and overburden measurements during data cleaning.

Once the well records had data about their geophysical setting appended in the attribute table through spatial joins, the attribute table was exported as an Excel spreadsheet. Within Excel, data cleaning was performed.

3.7 – Data Cleaning
Records were removed if a well point did not overlap with the three polygon GIS datasets (bedrock, surficial deposits and soil texture). If a mapped unit provided a classification not suitable for analysis (ex. “Lake” or “Urban Area”), this record was also removed from the resultant well and geophysical variable dataset. If the classification of bacterial contamination at a well was “overgrown” (a source of uncertainty that prevented evaluating the true amount of contamination) for either total coliform bacteria or *E. coli*, the record was removed. If an overburden point measurement was more than 250m away from a well point it was deemed to not be representative and removed.

The decision was made to remove each record from a single common dataset if one of these criteria occurred, rather than opting for several larger datasets that preserved records where a fraction of the variables could be evaluated. This allowed for stronger comparisons between tests using a smaller but complete dataset.

Categorical data was converted from text to integers mapped in a legend. This was accomplished by identifying the unique values found within each variable, and assigning these values a unique integer. Once converted, these integers were mapped in a legend for reference purposes. IBM SPSS Statistics 24 required that categorical data be in an integer format for it to be processed.
The surficial geological dataset classified several different types of Till using a place name (ex. Kingston Till). These classifiers were aggregated to a single value (“Till”) to avoid introducing spatial autocorrelation and unnecessary variation.

Once data was properly cleaned and converted accordingly, it was loaded into SPSS Statistics as a new dataset. Each variable (independent or dependent) was defined as being categorical or continuous using the variable view of SPSS Statistics.

3.8 - Spearman Correlation
A Spearman Correlation is a form of bivariate statistical analysis that assesses the statistical significance of a pair of variables. This test assumes that variables exist in a monotonic relationship, meaning that they are either inversely related or proportional to one another. This test requires that data is ordinal, interval or continuous (Laerd, 2013). A Spearman Correlation is typically applied to bivariate data if a Pearson Correlation’s assumptions are rejected due to a variable or the relationship between two not being linear or normally distributed. This test was only applied to the overburden depth variable. Based on a frequency distribution of overburden measurements proximal to geocoded well locations, data appears to be very skewed towards shallow overburden (Figure 6).

![Histogram](image)

Figure 6 – Histogram of the distribution of overburden depth values
3.9 - One-Way Analysis of Variance

A One Way Analysis of Variance (ANOVA) is a form of statistical analysis useful for measuring the statistical differences of the mean between two variables that do not meet the assumptions of a bivariate correlation. An ANOVA test requires the dependent variable to be continuous, the independent variable to be categorical, and there be no significant outliers. The dependent variable should be approximately normally distributed for each category of the independent variable. ANOVA tested categorical data using total coliform bacteria or E. coli concentrations as the dependent variable in separate trials.

A limitation of this test is that the influence of individual categories within a variable cannot be identified with an output. Categories that show greater relative variance can be identified using a scatter plot, but this test can not identify which categories were at greater risk of contamination. This can be accomplished with more complicated forms of statistical analysis such as Principal Components Analysis that create eigenvalues to evaluate factor loads of different categories (van Ewijk, 2017). This was found to be outside the scope of the undergraduate project. The weight of individual categories could be investigated through smaller scale forms of spatial analysis or a principal component analysis as discussed in 5 - Discussion.
4 - Results

4.1 - Overburden Depth
The correlation coefficient for the Spearman's Correlation between total coliform bacteria and overburden depth was .001 and not significant (p< .887), meaning that the test did not pass or meet the expected result of a negative correlation. The result of the Spearman's Correlation may be viewed at Table 4. Conversely, the Spearman's Correlation between overburden depth and E. coli yielded a correlation coefficient of -0.32 (p<0.01). This correlation coefficient demonstrates that there was a moderate inversely proportional relationship between the depth of the overburden and the level of E. coli contamination in groundwater. The result of the Spearman's Correlation may be viewed at Table 5.

<table>
<thead>
<tr>
<th>Spearman’s Correlation</th>
<th>Total Coliform</th>
<th>Overburden Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Coliform Count</td>
<td>Correlation Coefficient</td>
<td>.001</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>1.00</td>
<td>.887</td>
</tr>
<tr>
<td>N</td>
<td>37693</td>
<td>37693</td>
</tr>
<tr>
<td>Overburden Depth</td>
<td>Correlation Coefficient</td>
<td>.001</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>.887</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>37693</td>
<td>37693</td>
</tr>
</tbody>
</table>

Table 4 – Spearman’s Correlation between total coliform bacteria and overburden depth.
<table>
<thead>
<tr>
<th>Spearman's Correlation</th>
<th>$E. coli$</th>
<th>Overburden Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli Count</strong></td>
<td>Correlation Coefficient</td>
<td>-.032**</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>1.00</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>37693</td>
<td>37693</td>
</tr>
<tr>
<td><strong>Overburden Depth</strong></td>
<td>Correlation Coefficient</td>
<td>-.032**</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>37693</td>
<td>37693</td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (two-tailed)**

Table 5 – Spearman’s Correlation between $E. coli$ and overburden depth.

4.2 - Bedrock Geology

A One-Way ANOVA test was performed using total coliform bacteria as the dependent variable and bedrock geology as the independent variable. The result was significant ($p < 0.0001$) with an F-value of 4.921, meaning that the test passed with an expected result. This test demonstrated that total coliform bacteria varied between different types of geological units with statistical significance. The result of this test suggests that the type of bedrock geological unit could be a useful predictor of total coliform bacteria concentrations in the drinking water sampled from private wells. The results from this test may be viewed in Table 6.

<table>
<thead>
<tr>
<th>ANOVA</th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>6125.140</td>
<td>22</td>
<td>278.415</td>
<td>4.921</td>
<td>.000</td>
</tr>
<tr>
<td>Within Groups</td>
<td>2131404.590</td>
<td>37670</td>
<td>56.581</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2137529.730</td>
<td>37692</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6 – One-Way ANOVA between total coliform bacteria and bedrock geology.
The One-Way ANOVA test was repeated using *E. coli* concentrations as the dependent variable and bedrock geology as the independent variable. This test did not pass, and results were not significant (*p* = .690) (Table 7). The differences in *E. coli* concentrations were not found to vary between groups. The F-statistic as an expression of the average variability within groups compared to the average variability between groups was .830, demonstrating that the overall variability was very low. The result of this test demonstrates that bedrock geology type and *E. coli* concentrations do not share a statistical relationship, and that bedrock geology would not likely be a useful predictor of *E. coli* concentrations in the drinking water sampled from private wells.

<table>
<thead>
<tr>
<th>ANOVA</th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>19.320</td>
<td>22</td>
<td>.878</td>
<td>.830</td>
<td>.690</td>
</tr>
<tr>
<td>Within Groups</td>
<td>39842.406</td>
<td>37670</td>
<td>1.058</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>39861.727</td>
<td>37692</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7 – One-Way ANOVA between *E. coli* and bedrock geology.

### 4.3 - Surficial Geology

The One-Way ANOVA tests were repeated for an independent variable of surficial geology unit type, using both total coliform bacteria and *E. coli* concentrations again as the dependent variables in two separate tests. The test using total coliform bacteria found significant differences in the means between groups of surficial geological units (*p* < 0.0001) (Table 8) and an F-Statistic of 7.874. Conversely, the test using the means of *E. coli* concentrations did not observe statistically significant differences between groups of surficial geological units (*p* = .162) (Table 9). These tests demonstrated that surficial geology deposit type could possibly be used as a predictor of total coliform concentrations in private wells, but not of *E. coli* concentrations.
<table>
<thead>
<tr>
<th>ANOVA</th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>3566.952</td>
<td>8</td>
<td>445.869</td>
<td>7.874</td>
<td>.000</td>
</tr>
<tr>
<td>Within Groups</td>
<td>2133962.777</td>
<td>37684</td>
<td>56.628</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2137529.730</td>
<td>37692</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8 – One-Way ANOVA between fecal coliform bacteria and surficial geology

<table>
<thead>
<tr>
<th>ANOVA</th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>12.443</td>
<td>8</td>
<td>1.555</td>
<td>1.471</td>
<td>.162</td>
</tr>
<tr>
<td>Within Groups</td>
<td>39849.284</td>
<td>37684</td>
<td>1.057</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>39861.727</td>
<td>37692</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9 – One-Way ANOVA between E. coli and surficial geology.

4.4 – Soil Texture
Finally, the One-Way ANOVA test was repeated for classification of soil texture as an independent variable. Total coliform was found to significantly differ between different types of soil texture (F=2.063, P=0.005). When repeated for E. coli, it was observed that concentrations varied with a statistical significance of 0.027, and an F-Statistic of 1.732. This test demonstrated that soil texture could be a possible indicator of both total coliform bacteria and E. coli concentrations given the level of observed variance with significance.
<table>
<thead>
<tr>
<th>ANOVA</th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>2104.559</td>
<td>18</td>
<td>116.920</td>
<td>2.063</td>
<td>.005</td>
</tr>
<tr>
<td>Within Groups</td>
<td>2135425.171</td>
<td>37674</td>
<td>56.682</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2137529.730</td>
<td>37692</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 10 – One-Way ANOVA between total coliform bacteria and soil texture.

<table>
<thead>
<tr>
<th>ANOVA</th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>32.964</td>
<td>18</td>
<td>1.831</td>
<td>1.732</td>
<td>.027</td>
</tr>
<tr>
<td>Within Groups</td>
<td>39828.763</td>
<td>37674</td>
<td>1.057</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>39861.727</td>
<td>37692</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 11 – One Way ANOVA between *E. coli* and soil texture.
5 - Discussion

Total coliform and *E. coli* counts often exhibited different statistical relationships with each dependent variable. During the Spearman’s correlation, total coliform bacteria did not demonstrate a statistically significant negative correlation with depth (the expected result), but demonstrated statistically significant levels of variances when compared to all other variables (bedrock geology, surficial geology and soil texture).

Conversely, *E. coli* exhibited a statistical relationship with the overburden depth. *E. coli* varied with soil texture, but did not vary with statistical significance in bedrock and surficial geological units. This difference in results between bacterial indicators was not expected given that *E. coli* are one type of bacteria within the total coliform bacteria parameter. Therefore, it should be expected that they would occur concurrently in soil and groundwater if it were assumed all observed bacteria were fecal sourced. Following this assumption, it is predicted that these differences in statistical significance could be attributed to differences in the distribution of the total coliform and *E. coli* values in the data, or differences in survival rates of *E. coli* in soil and groundwater relative to other types of bacteria.

5.1 - Survival Rates of Bacteria

One possible explanation for the differences in statistical analysis of total coliform and *E. coli* is differences in survival rates. For example, in a study conducted by Sinton et al., (2007), prepared cow manure pats were placed on pasture lands, taking samples over three months for analysis of the decay rates of several different types of bacteria including *E. coli* and enterococci. This process was repeated over four, three month periods (accounting for each season). In most cases *E.coli* had higher decrease coefficients than enterococci. While this experiment was conducted at the surface rather than in the subsurface, it suggests *E. coli* die faster than other types of enteric bacteria when outside of the gastrointestinal tract. This could explain the negative correlation observed between the *E. coli* concentration and the depth of overburden, with *E. coli* dying at rates in deeper soil that would produce a characteristic correlation. Survivability of other coliform bacteria at depth could prevent a characteristic correlation. This proposition is complicated by the fact that total coliform and *E. coli* likely do not always originate from the same source (as discussed in 5.2), making evaluating the observed differences difficult.
Total coliform bacteria with greater environmental resilience and multiple sources could appear in a wider range of environments than *E. coli*. Presence in a wider range of environments could lead to significant differences in the means of total coliform counts. This line of reasoning could explain why total coliform bacteria varied in groups of bedrock and surficial units, but *E. coli* did not. The idea of a faster die off of *E. coli* in the environment is also supported by the variance of these bacteria in units of different soil texture, but not in geological units. This demonstrates that *E. coli* were present in a wider range of soil environments, indicating that survival may be more common in soil but that *E. coli* likely die off before reaching many bedrock units.

Crane and Moore (1985) is a useful example of literature citing differences in survival rates of *E. coli* and other forms of bacteria when outside the digestive tract. This study highlighted other examples where *E. coli* was observed to die off faster than many other forms of coliform bacteria in lab experiments. Examples include a 1979 study which found that *E. coli* had significantly higher die off rates in sterilized and inoculated manure slurry (6.22 k/day > 1.042 k/day) than general coliform bacteria. However, two other studies from 1969 and 1970 used similar methods and found that *E. coli* had lower die off rates than other forms of bacteria analyzed (Crane and Moore, 1985). Evident from these differences and the challenges associated with comparing different die-off coefficients obtained using different experiments is that more study is required to evaluate how survival rates of different types of bacteria are reflected in contaminants of drinking water. Selecting the most relevant indicator bacteria could lead to more effective forms of statistical analysis for Southern Ontario during attempts to identify regional dynamics of bacterial contamination. Further study should include a more recent literature review, or further review of lab experiments.

**5.2 - Quantitative Distribution of Total Coliform Bacteria and E. Coli**

During an exploration of the well data provided by Public Health Ontario, it was also observed that counts of total coliform bacteria and *E. coli* had different distributions. A “zero” *E. coli* count was observed at 97.9% of all wells in the dataset, whereas a “zero” total coliform count was only observed at 79.0% of all wells in the dataset. This observation is summarized in Table 12. This provides a possible explanation for the very weak F-values observed in the ANOVA outputs that used *E. coli* as the dependent variable. Given the drastic difference in frequency of occurrence of each indicator, this confirms that *E. coli* is a better indicator of interactions between a fecal source of bacteria and the geophysical environment. The magnitude of difference suggests that total
coliiform comes from sources other than fecal matter, as previously stated (Health Canada, 2013). Since total coliiform bacteria contamination is more frequent in private drinking water well samples, it should be expected that stronger statistical results would also result (however the results may not be indicative of relationship between geophysical setting and fecal contamination of a well). This means that it is difficult to quantitatively compare the results of total coliiform bacteria and E. coli.

It is possible that the E. coli only occurred in the most vulnerable conditions where survival rate was not a factor. As noted earlier, Jaimeson et al. (2002) state that survival times of bacteria can be elevated in cases where there are continual bacterial inputs. Given the skewness of the data towards total coliiform bacteria, one possible explanation is that E. coli is only observed when there are continual bacterial inputs in vulnerable environments. More evaluation would be required to determine the vulnerability to contamination of groundwater environments where E. coli was present relative to the rest of the region.

<table>
<thead>
<tr>
<th>Bacterial Indicator</th>
<th>Occurrence in the Well Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td>2.1%</td>
</tr>
<tr>
<td>Total Coliform Bacteria</td>
<td>21.0%</td>
</tr>
</tbody>
</table>

Table 12 – Rate of Occurrence of Bacterial Indicators in Well Data

Therefore, it is difficult to evaluate why total coliiform bacteria displayed different statistical results than the E. coli, and whether this can be attributed to differences in quantitative distribution, biological properties of bacteria, or geophysical setting. Subsequent discussion of the environmental variables studied will not attempt to address these differences, instead considering why hypotheses were proven false for the effects of the environmental variables on rates of general bacterial contamination in rural drinking water.

5.3 - Overburden Depth
Despite the numerous sources indicating that the presence of total coliiform should be inversely proportional to the depth of overburden (Gerba, 1975; Conboy, 2000; Borchardt et al, 2007) this was not observed. This provides an indication that the presence of total coliiform
bacteria in groundwater may be governed by processes other than the depth of an overburden profile. In conceptualizing the issue of contamination in private wells and the role of overburden, there are a few possible contributing factors.

It is possible that the control on contamination is soil texture and not just depth. Considering that unconsolidated sand or clays have differences in hydraulic conductivity that are several orders of magnitude (Freeze and Cherry, 1979), it is possible that units with less overburden and finer texture (ex. silt and clay) could actually mitigate flow of contaminated water more than a thicker unit of coarser material (ex. coarse grained sand or till). This was reflected in the One-Way ANOVA test between soil texture and total coliform bacteria, during which total coliform bacteria and E. coli both varied with soil texture. This supports the given proposition that soil texture may play a role in the flow of groundwater and the transportation of bacterial contaminants when overburden is present.

As noted earlier, Jaimeson et al state that survival times of bacteria can be elevated in cases where there are continual bacterial inputs (Jaimeson et al, 2002). This could help to explain why overburden depth was not negatively correlated with rates of total coliform contamination. If the source of bacteria was leakage from a septic system, this could provide a localized but continual source of bacteria, perhaps also creating preferential flow paths for water from the subsurface into groundwater. These sources are also already submerged, thereby circumventing the mitigating effects of any of the soil above the septic system. The overburden value used in the statistical analysis would be an over estimate, thereby reducing a negative correlation due to high levels of bacteria. Because as much as 49% of bacterial contamination in wells could be attributed to human sources in southern Ontario (Krolik et al, 2014), this is a possible explanation. If contamination is human-sourced and originating from a leaking system, a preferential flow path with continual bacterial inputs could exist, consequently circumventing the expected flow paths through the overburden column. Overburden depth may, however, be a useful predictor of total coliform bacteria contamination when above ground and diffuse sources exist (Eg. agricultural inputs).

Another possible explanation is that some residents may be drawing their water from an unconfined aquifer. If the ground water table occurred above bedrock in a zone of saturated porous media, water could theoretically be drawn from a depth that is shallower than the observed depth to the bedrock used for statistical analysis. This would mean that the overburden depth in these
cases is an overestimation, and that coliform bacteria may circumvent the mitigating processes of parts of the column of porous material measured in the data. In these cases, higher levels of total coliform could be observed despite having deep overburden, thus creating a weaker negative correlation between contamination and depth than what was expected.

*E. coli* did not show the same results as total coliform bacteria, and negatively correlated with overburden depth at a rate of -0.032 per meter of overburden. This was closer to the idealized result that was expected, but not a strong enough correlation to conclude the overburden depth at a well is the dominant control of observed levels of bacterial contamination. The difference between the tests for *E. coli* and total coliform bacteria indicates that total coliform generally survives at slightly better rates in deeper overburden profiles. Given the results showing variance with soil texture, these complementing results give evidence that soil texture is potentially an important component of the geophysical environment that may protect groundwater from contamination.

5.4 - Soil Type

As briefly stated, when rates of bacterial contamination were evaluated for variance using soil texture as an independent variable, both total coliform bacteria and *E. coli* varied with statistical significance. This was expected as a logical conclusion given the primary control texture will play on the conductivity of water in unconsolidated sediments, and the support for this concept in literature (Conboy, 2000; Jamieson *et al.*, 2002; Gerba, 1975). Because overburden displayed such weak statistical significance, these results suggest that texture could be more important than depth. Given that these two independent variable categorized the unconsolidated sediments above wells from the same sample set in two different fashions, this is a logical conclusion. It may also be argued that because soil texture has been categorized as being representative of an area rather than overburden as a point measurement, the results for soil texture are more valid. This is because the overburden depth measurement closest to a private well makes the assumption that all bacteria in a well is a product of vertical migration through the overburden of a given depth at the well, when this is likely not the case. Because soil texture may be representative of the larger area that bacteria at a well could be originating from, the results for soil texture may be a more appropriate predictor of contamination.

The F-Statistics for soil texture were lower then those of the bedrock geology and surficial geology tests using total coliform bacteria. This demonstrates that when statistical significance was
identified in the variance of bacteria when compared to a dependent variable, that bacteria varied less with soil texture than other variables. However, soil texture was the only independent variable to identify variance in both total coliform bacteria and *E. coli*. This test suggests that soil texture could be used in other forms of statistical analysis to provide conservative but consistent predictions of different types of bacterial indicators. The lower F-statistics suggests that bacteria may be present in many soil textures, with some variation. The higher F-statistics for total coliform bacteria in bedrock and surficial units suggests that these are more variable environments for total coliform bacteria, but unfortunately these independent variables will not be able to predict *E. coli* as effectively. It is logical to expect less variance in bacterial contamination of wells of different soil textures despite the statistical significance given that soil is typically the most proximal component of the environment to a point source of bacteria if located at the surface. When soil is present, it will always be encountered before deeper surficial deposits (if present) and bedrock. This means that soil textures should show less variance than deeper units of sediment and bedrock, given many bacteria die when travelling to these units.

5.5 - Bedrock Geology

It was expected that bedrock geology would show a difference in the means of bacterial concentrations of both total coliform and *E. coli*, especially given the wide range of bedrock types within the study area. However, this was only observed in total coliform bacteria counts. It was predicted that bedrock would exhibit a wide range of characteristics known to influence the rates and pathways of flow as expressed during literature review. These differences in the ability to regulate flow were expected to have a strong influence on the travel of fecal contaminants from a source to a private well. There a several factors that provide critique of the result that total coliform varied greatly in different units.

As noted by Goss, many people draw from wells drilled into unconsolidated sediments rather than bedrock (Goss *et al.*, 1998), although an exact figure was not provided. In a survey of 400,000 well records in southern Ontario, Carter *et al.* state that very few wells were drilled more than 10 meters into bedrock (2014). If the water quality of many wells is dominated by saturated overburden as a primary control due to only shallow penetration of bedrock, the type of bedrock could have a lower influence on the fecal contamination of the well. If there is no confining unit above the aquifer, bedrock would only protect from lateral flow deep in the aquifer. The aquifer would be very vulnerable to flow through overburden, vulnerability that could be exaggerated by
well pumping for household use to create a drawdown of hydraulic head gradient, potentially
drawing in fecal contaminants without a confining unit to protect from contamination. The lack of
knowledge about the depth, horizons and stratigraphy of each well is the greatest source of error in
using this variable. In reality, confined systems and unconfined wells should be evaluated in
separate forms of analysis because of how different each aquifer is. Because this study uses spatial
joins at a well location to evaluate geophysical variables, the assumption implicit in this dataset is
that groundwater in the area is sourced from primarily vertical migration. This form of spatial
analysis does not provide a mechanism to evaluate horizontal migration. The size of the geophysical
polygon units (bedrock, surficial deposits, soil texture) improve this limitation by having larger
coverages representative of the area around the well point, in addition at the well itself. Many of
these assumptions will not apply to confined aquifers, where wells draw from much older water
less susceptible to the processes occurring at the water table due to a protective confining unit, and
could be sourced from further away through horizontal migration.

It is also important to note many wells are drilled into bedrock with very shallow
overburden. Given the dramatic increase of karst topography of carbonate units in southern
Ontario that was observed when overburden was less than 1 meter (Gagnon-Nandram, 2016), it
has long been assumed that these wells are vulnerable due to the limited mitigating capacity of
overburden and the propensity for karst topography. Given the dominance of carbonate units in
southern Ontario, it is important to note that while bedrock type showed variances in total coliform
bacteria it may not be a useful indicator given the dominance of this type of unit. In these cases,
overburden should be a determining factor. However, the overburden correlations did not show
useful results, meaning that this expectation requires further evaluation, perhaps using co-variance
or codependence of bedrock type, overburden depth and bacterial contamination.

Evident from the above discussion of using bedrock type as an independent variable in
statistical analysis is that a number of assumptions are made about the character of bedrock based
on its composition. When a bedrock unit is identified as a certain type a number of assumptions are
made about the degree of fracturing, the amount and connectivity of pores, the amount of
weathering and the expected level of fecal contamination. For example, units of the same type can
have extremely different hydraulic conductivity values depending on the relative age and degree of
weathering in each unit. As an extreme example, Freeze and Cherry estimate the intrinsic
permeability of immature limestone to be 0.0001 to 1 Darcy, whereas karst limestone has an
intrinsic permeability of 0.1 to 1000 (1979). This holds true as well for sandstones, but at a lower level of difference. Therefore, simply characterizing units as “limestone and dolomite” does not give enough insight into the weathering of the unit. Weathering could be estimated by age of the units, or by using bedrock topography to estimate regional groundwater flow. Fracturing would also require estimation, and presents a regional in evaluating the regional scale dynamics of bacterial contamination in souther Ontario. As discovered during literature review, the flow of groundwater in bedrock is an incredibly complicated process. To properly estimate the likelihood of a well having bacterial contamination as a function of the characteristic of the bedrock, a much wider range of factors will need to be incorporated into statistical analysis. It is predicted that the One-Way ANOVA test did not incorporate enough of these complexities to garner a statistical relationship between bedrock geology and total coliform bacteria or E. coli concentrations.

5.6 - Surficial Geology

The result of surficial geology and historical depositional environment influencing the rates of total coliform bacteria in rural drinking water was expected, however the lack of significance in the results for E. coli was not. In addition to the distribution of grain sizes associated with the energy of the depositional environment, surficial geology could also influence the arrangement of grains and the parent material. While literature has not answered this idea, it is predicted that the depositional environment of surficial geology could influence the arrangement of soil grains, thereby creating a modifying influence on soil structure dominated by grain size distribution. Higher energy environments would likely lead to less hexagonal (tight) stacking, thereby increasing the size of pores and providing larger pathways for flow relative to other depositional environments with similar grain sizes.

One possible explanation for the statistically significant results for surficial geology is that grain size distribution had a strong influence on the rate of fecal contamination, as in the previous discussion of overburden and soil type. Southern Ontario’s surficial geology is driven by the last glaciation. The energy of the diverse environments will determine the types of grains that are suspended, from the braided channels of coarse sand and till originating from meltwater to silt and clay sized particles settling on the bottom of glaciolacustrine environments. While these environments will also determine the thickness of drift deposits, the Froude energy level will also greatly influence the size of grains present (Narbonne, 2017). It is predicted that surficial geology yielded a statistically significant result from One-Way ANOVA because of how the variable
encapsulates thickness, grain stacking and grain size distribution in a single parameter, thereby providing a more detailed summary of the factors regulating the flow of water and bacterial contaminants in the subsurface.

With exception (Rodvang and Simpkins, 2001; Levy et al., 2007), it was found during literature review that few researchers have explicitly incorporated surficial geology into an analysis of observed rates of fecal contamination in groundwater. Literature tended to focus on the effects of topsoil and bedrock geology, without explicit consideration of the possible effects of surficial geological depositional environment. This is to be expected given the evolving understanding of hydrology and the biogeochemistry of rural drinking water sources and bacteria, but based on the results of this analysis it appears this relationship is underrepresented in literature.

6 – Conclusions

The results of the statistical analysis demonstrated that soil texture is the only explanatory variable that could be useful for evaluating concentrations of both total coliform and E. coli at wells on a regional level on a statistical basis, based on the results of a One-Way ANOVA test and Spearman’s Correlation. One-Way ANOVA showed total coliform bacteria concentrations varied among different types of bedrock geology and surficial geology, but E. coli concentrations did not, reducing the viability of these variables for further analysis. The negative correlation observed between overburden depth and E. coli concentration mean that overburden depth may be used as a variable for risk assessment. Because soil texture showed positive results for both indicator bacteria, texture appears to be the more useful way to classify the unconsolidated sediments of private wells contaminated with bacteria on a regional basis in southern Ontario (when compared to surficial deposit classification). Critique of the analysis demonstrated that more variables require analysis such as fracturing, bedrock age, well depth, well abandonment, presence of confining units, and presence of distinct soil horizons, only some of which are feasible for regional analysis. Analysis could be conducted using a more robust test such as a principal components analysis or geographically weighted regression. The results of this experiment should be considered in any future attempts to build on the groundbreaking work by Public Health Ontario and Krolik et al (2013) to understand the regional and environmental dynamics of the bacterial contamination of private wells in southern Ontario. This project demonstrates that the use of GIS and statistics remain key tools to evaluate well contamination as an environmental and human health issue requiring policy, risk assessment, and mitigation in southern Ontario.
References


Dejonghe, W., Goris, J., Fantroussi, S. E., Hofte, M., Vos, P. D., Verstraete, W., & Top, E. M. (2000). Effect of Dissemination of 2,4-Dichlorophenoxyacetic Acid (2,4-D) Degradation Plasmids on 2,4-D Degradation and on Bacterial Community Structure in Two Different Soil Horizons. Applied and Environmental Microbiology, 66(8), 3297-3304.
Doyle, M., & Erickson, M. (2006). The fecal coliform assay, the results of which have led to numerous misinterpretations over the years, may have outlived its usefulness. *Woods End Laboratory*, 1-7.


