2023

Mapping infiltration in an urbanizing mixed-land-use watershed with multi-temporal satellite imagery

Sarah J. Higgins
West Virginia University, sjh00024@mix.wvu.edu

MAPPING INFILTRATION IN AN URBANIZING MIXED-LAND-USE WATERSHED WITH MULTI-TEMPORAL SATELLITE IMAGERY

Sarah J. Higgins

Thesis submitted
to the Davis College of Agriculture, Natural Resources, and Design
at West Virginia University

in partial fulfillment of the requirements for the degree of

Master of Science
in Environmental Soil and Water Sciences

James A. Thompson, Ph.D., Chair
Jason A. Hubbart, Ph.D.
Aaron E. Maxwell, Ph.D.

Division of Plant and Soil Sciences

Morgantown, WV
2023

Keywords: digital soil mapping, infiltration, land use change, urbanization, dynamic soil properties, dynamic soil survey, multi-temporal dynamic covariates, automated dual-head infiltrometers, random forest, multiple linear regression

ABSTRACT

MAPPING INFILTRATION IN AN URBANIZING MIXED-LAND-USE WATERSHED WITH MULTI-TEMPORAL SATELLITE IMAGERY

Sarah J. Higgins, M.S.

Advisor: Dr. James A. Thompson

Digital soil mapping (DSM) is a field of soil science that aims to improve traditional soil maps by producing higher resolution predictive maps of soil properties using spatial environmental data. DSM has historically relied primarily on static environmental covariates—such as slope gradient, slope aspect, and other topographic variables derived from digital terrain models—for predicting static soil properties, like soil texture. Advancements in satellite imagery and statistical modeling improve the accuracy of digital soil maps by incorporating multi-temporal data that can capture landscape-scale change over relatively short periods of time. Adding these dynamic environmental covariates may be especially useful for spatial prediction of dynamic soil properties, like infiltration rate, that are strongly affected by phenomenon that satellite imagery can detect, like land use that changes rapidly due to human activity. Infiltration strongly impacts soil health and hydrologic characteristics in a watershed. Understanding infiltration for sustainable land management is vital for making best management decisions in urbanizing environments like the West Run Watershed in Morgantown, West Virginia. We hypothesized that infiltration could be predicted at a higher accuracy and a finer spatiotemporal scale using digital soil mapping techniques than is currently provided by the current official soil data and maps produced by the National Cooperative Soil Survey. Spatial predictions of infiltration rate were produced for the West Run watershed using both static and dynamic environmental covariates as inputs into multiple linear regression (MLR) and random forest (RF) models, each of which were made using 10-fold cross validation. Training and independent validation sampling locations were selected using a conditioned Latin hypercube sampling scheme and observed saturated hydraulic conductivity of the soil surface was collected using automated dual-head infiltrometers. The MLR and RF models had R^2 of 0.302 and 0.201, respectively. Validation sampling was stratified by the predicted infiltration values of the MLR model. Validation R^2 values for the MLR and RF models were 0.080 and 0.103. The results from this study will benefit the development of a dynamic soil survey and will improve hydrologic models in this and potentially other mixed-land-use watersheds.

ACKNOWLEDGEMENTS

Thank you to the Natural Resources Conservation Service for funding this work and providing me with support throughout the process. I would also like to thank and acknowledge the original caretakers of the West Run Watershed: the tribal nations of the Osage, the Shawandasse Tula, the Massawomeck, and the Monongahela Culture.

I would like to thank my advisor, Dr. Jim Thompson, for his support and guidance throughout my time at West Virginia University. Digital soil mapping was a mystery to me two years ago, and I have thoroughly enjoyed every step of the process of building the skills that have made me the (novice) digital soil mapper I am now! Thank you for your patience and encouragement on that journey. Thanks to you, I feel that I am a much stronger soil scientist today and I am eager to take these new skills out into the world. Thank you also the Stephen Roecker, Suzann Kienast-Brown, Dave White, and Tiffany Allen for your valuable advice and assistance throughout the modeling process. Thank you to my committee members, Dr. Jason Hubbart and Dr. Aaron Maxwell, for sharing your knowledge with me and strengthening my skills as a scientist. Special thanks to Nathan Edwards and Mylee Franklin—your help throughout the sampling was invaluable, and I enjoyed your fantastic company as well. Thank you to all others who helped me accomplish the sampling: Dan Benyei, Lilai Jin, Kinsey Reed, Abby Clark, Ann Tan, James Leonard, Megan Thomas, Debby Cunningham, Wendy Noll, and Joel Gebhard. Thank you also to all the landowners in the West Run watershed who allowed me to sample their properties for your kindness and cooperation.

Thank you also to all members of the WVU Soils Team for allowing me to help coach you and learn more about the soils in this world with you. Watching you all grow into impressive soil scientists has been one of the highlights of my graduate studies and keeps my passion for

pedology alive. Thank you to everyone in the Davis College for providing a welcoming and encouraging environment throughout my studies. Thank you to all my friends I've made here at WVU. Thanks especially to *mi amor*, Francisco Gil, for your encouragement and all our study dates at Blue Moose. You're easily the best thing to come from stats class. Thank you as well to everyone back in Missouri who has been supporting me throughout this journey even while being 700 miles away. You all helped shape me on my journey and inspired me to keep working to make you proud every day. Thanks to my parents for encouraging me to play in the dirt and to my friends for sharing that passion with me. I could never have gotten this far without your support and your faith in me. And of course, thank you to my cat, Ash Andi(sol) Higgins, for her unwavering emotional support these two years. I feel so lucky to study the soils and the environment that I love every day and experience the beauty of Appalachia. This world is so full of wonder, and I hope that I can give back to it even a fraction of what it has given to me.

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION

Dynamic soil properties (DSP), or soil properties that are influenced by anthropogenic disturbance over short periods of time, are gaining more attention from scientists for use in soil health and sustainable watershed management because they are strongly connected to land use practices (Karlen et al., 2019). Patterns in DSPs are inherently linked to land use due to the different vegetation and management characteristics associated with different land uses (Wills et al., 2017). For example, forests and pastures exhibit different vegetation structures and management regimes that may lead to differences in various soil physical and chemical characteristics, even in similar soil types and landscape positions. It is important characterize DSPs across a landscape because they influence a soil's capacity to perform necessary functions like holding water, storing nutrients, and maintaining structural support (Natural Resources Conservation Service, 2022). One DSP critical to soil scientists, hydrologists, and engineers is infiltration, which is the process by which water derived from precipitation or overland flow enters the soil. Infiltration influences erosion, runoff, streamflow, and available water to plants, making it a significant factor in hydrologic models (Horton 1933, Blasch et al., 2006). As a DSP, infiltration is linked to land use. Vegetation characteristics and management techniques change soils' organic matter content and bulk density, which are both strongly related to infiltration rate (Sun et al., 2018). Detailed infiltration maps are necessary to improve the accuracy of hydrologic models and strategically target management on a watershed scale (Vieux, 2001). This is especially true for urbanizing watersheds that encompass several different land uses, such as the West Run watershed in Morgantown, West Virginia. Flooding is an increasingly common problem in Appalachia due to the region's rugged topography and higher occurrences of severe precipitation events (Harvey, 2022). Extreme flooding events are economically devasting in

West Virginia—in fact, in their 2022 state climate summary for West Virginia, the National Oceanic and Atmospheric Administration found flooding as "the costliest and most severe natural hazard for the state" (Runkle et al., 2022). Infrastructure within the West Run watershed is at risk of flood damage, and severe, damaging floods are becoming more common. Each year urbanization in the watershed increases. Understanding the spatiotemporal variation in infiltration offers insight into flooding dynamics in a watershed and provides useful information for flood mitigation strategies.

We propose that infiltration, measured as the saturated hydraulic conductivity of the soil surface, can be modeled using multi-temporal satellite imagery when combined with terrain derivatives traditionally used in digital soil mapping (DSM). Infiltration varies over space and time according to vegetation and land use, both of which are spatially and temporally variable in the West Run watershed. We hypothesize that the temporal resolution offered by multi-temporal satellite imagery can capture this change and thus improve infiltration map accuracy. As such, the goals of this study are to:

1. Represent DSP on a watershed-scale using DSM techniques.
2. Use both static and dynamic environmental covariates to predict DSP.
3. Determine the utility of multi-temporal environmental covariates for creating a DSS.

The objectives of this study are to analyze the links between the spatial variability of infiltration and use these measurements to develop an infiltration raster map that will contribute to a dynamic soil survey and serve as input for distributed hydrological modeling. Study objectives will be accomplished by developing and validating a predictive soil map for a dynamic soil property (infiltration of surface soil) using static and dynamic environmental covariates.

We hypothesize that hydrologic properties of the surface soil will vary among land use types in a mixed-land-use watershed. These differences will be evident in measurements of infiltration rate and saturated hydraulic conductivity. Our specific hypothesis is that spatial predictions of both static and dynamic soil properties will be improved by incorporating dynamic environmental covariates into the modeling workflow.

CHAPTER 2. LITERATURE REVIEW

Soil and water share a close relationship that has been studied since the beginning of soil science as a discipline (Brevik & Hartemink, 2010). Notably, Vasily Dokuchaev (Johnson & Schaetzl, 2015) and Hans Jenny (1941) outlined climate, a function of mean annual temperature and precipitation (in addition to other climatic factors), as one of the key factors controlling pedogenesis and regional soil heterogeneity. At a local scale, topography also influences landscape hydrology and soil formation by controlling the direction and velocity of water flow in a watershed. It also controls subsurface flow and the water table level and is thus an important factor when considering soil water infiltration in a comprehensive water budget (Crave & Gascuel-Odoux, 1997). Water is also a driver of pedogenesis by transporting ions and clay minerals down a profile and creating opportunities for anaerobic soil processes to occur by saturating parts of a soil profile (Johnson et al., 1990). The interactions between soil and water are fundamental to understanding ecosystem function and the watershed response to precipitation events (Hewlett & Hibbert, 1967). Precipitation that falls in an unurbanized watershed is fated to be either intercepted by vegetation, fall directly into surface water bodies, or seep into the soil. In an urban or urbanizing watershed, rainfall may also be intercepted by concrete or other built structures. It then may find its way to nearby soil, surface waters, a rainwater basin, or a stormwater management system. Common measures to quantify this relationship between soil and water include soil moisture content, hydraulic conductivity, and infiltration.

Infiltration

Infiltration is the process of water movement into the soil surface. Infiltration rate is the rate at which water enters the soil, which is commonly measured in units of thickness of water infiltrated over time such as centimeters per second or millimeters per hour (Natural Resources

Conservation Service, 2022). Saturated hydraulic conductivity of the soil surface refers to the movement of water through the soil when the soil is fully saturated and all pores are filled with water. It is a common metric used to estimate the steady-state infiltration rate, which is the infiltration rate that is reached after initial infiltration once all pores are filled with water under a constant head (Lal & Shukla, 2004). Infiltration rate of a soil is considered a DSP, or a soil property that changes over a relatively short spatiotemporal scale with natural and anthropogenic disturbances (Natural Resources Conservation Service, 2022). Soil organic matter content and total porosity share a strong positive relationship to infiltration rate, while bulk density and initial moisture content show a negative relationship to infiltration rate (Shukla et al., 2003; Sun et al., 2018; Anderson et al., 2020). Additions of organic matter to the soil surface promotes aggregation of soil particles, thereby increasing average pore size, improving pore space connectivity, and providing preferential flow paths for water (Shukla et al., 2003). Increased organic matter content and minimal compaction will lower bulk density, increasing infiltration rates (Shukla et al., 2003). Both soil organic matter content and bulk density are soil properties that are strongly influenced by land use and management (Shukla et al., 2003; Sun et al., 2018), making infiltration spatially and temporally dynamic in nature. Soil surface texture, while not a DSP, is a spatially variable soil physical property that also has a strong influence on infiltration due to the inherent porosity characteristics associated with each soil textural class (Saxton et al., 1986; Jabro et al., 1992; Arya et al., 1999).

Several direct methods are used to measure infiltration. Among the most common approaches are using a single-ring infiltrometer, double-ring infiltrometer, permeameters, and tension infiltrometers (Erikson et al., 2013). As the name suggests, a standard single-ring infiltrometer is one ring that is driven into the ground. With some of the ring remaining above

the surface, water is poured into the ring to pool over the surface and saturate the soil. An issue with single-ring infiltrometers is the error caused by sub-surface lateral flow into the surrounding soil outside of the ring. To mitigate this effect, a double-ring infiltrometer uses two separate rings: a 15-cm diameter ring inside a 30-cm diameter ring driven into the ground (Gregory et al., 2005). Once both rings are at the same height, water is pooled in both. There are two ways to measure infiltration rate using the ring methods—both of which require having a known volume of water when beginning the measurement. The falling head test measures the water level (of the inner ring in the case of the double-ring infiltrometer) over time as it infiltrates into the soil. Alternatively, the constant head test maintains a constant water level throughout the entire measurement by adding more water as infiltration occurs. The constant head test is more accurate than the falling head test (Wu et al., 1997). Permeameters can measure soil hydraulic properties in a borehole and can measure soil hydraulic conductivity at various depths in the subsoil, although modified methods exist for using permeameters methods for surface soil infiltration (Ahmed et al., 2014). Two commonly used permeameters are the Guelph permeameter and the Philip-Dunne permeameter. The Guelph permeameter maintains a constant head of water over the soil using a Mariotte reservoir and takes measurements at two different pressure heads. The Guelph method can be modified to measure the soil surface hydraulic conductivity by using a tension infiltrometer, which also connects to a Mariotte reservoir (Reynolds, 1993). Tension infiltrometers use a porous disc that is placed on the surface, which allows measurement of the unsaturated soil hydraulic conductivity using two different pressure heads (Reynolds & Elrick, 1991). For the Philip-Dunne permeameter, a tube is inserted into the borehole and filled with water. This permeameter uses the falling head test. Hydraulic conductivity of the specified depth is measured by recording the time for half of the water in the

tube and infiltrate and when the tube is completely empty (Philip, 1993). Compact constant-head permeameters (also known as Amoozemeters) are also commonly used to measure saturated hydraulic conductivity of an unsaturated soil. This measurement also takes place in an auger bore hole in the field and maintains a constant hydraulic head. This method uses less water than the others, but still takes several hours and requires constant monitoring (Amoozegar, 1989). All methods for infiltration outlined previously are viable and accurate, but either require tens of gallons of water, specialized equipment, constant monitoring, and/or several hours to obtain a single measurement. Automated dual-head infiltrometers are a more recent method that aims to produce accurate results using less water and effort in less time. These infiltrometers use one ring that is driven into the ground and an infiltrometer head that attaches to it and pools water over the ring. The head is connected to a computer system that automatically calculates saturated hydraulic conductivity in the field at two pressure heads, specified in the system's settings. Once these infiltrometers are running, they are relatively "hands-off" and complete the measurement of saturated hydraulic conductivity in about one and a half to two hours, depending on the system settings. (Rivera et al., 2016).

Timely and accurate field measurements of infiltration are important because infiltration parameters are a significant factor in many hydrologic models, including those used to predict erosion, flooding, streamflow, and groundwater levels (Gray et al., 1985; Roo & Riezebos, 1992; Vieux, 2001; Shanafield & Cook, 2014). Lower infiltration rates are often associated with higher rates of runoff and erosion. Erosion occurs when a soil particle becomes detached from its surrounding particles and it transported elsewhere by an erosive agent, such as wind, water, or gravity. Water that does not infiltrate into the soil may become runoff, which is capable of carrying soil particles from the nutrient-rich topsoil with the overland flow (Ellison, 1948).

Raindrop impact from precipitation events and compaction cause soil surface crusting as fine particles become detached and clog surface pores (McIntyre, 1958). This soil surface crusting process, which is especially risky for disturbed bare soils, causes reduced infiltration rates and increased runoff compared to soils where this phenomenon is not occurring (Moore, 1981).

Watershed Management and Land Use

Infiltration is a key environmental process considered in watershed management to achieve soil and water conservation goals (Zimale et al., 2017). Soil infiltration capacity is critical to groundwater recharge, runoff and erosion control, and supplying water to the soil matrix for use by plants (Sun et al., 2018). Soils with high infiltration capacities can also mitigate flood damage during extreme rainfall events (Itsukushima et al., 2021). In addition to being a temporally dynamic soil property, infiltration also is spatially variable across a watershed depending on land use, vegetation, surface texture, bulk density, organic matter content, slope gradient, and hillslope position (Shukla et al., 2003; Schaik, 2009; Sun et al., 2018). As such, site-specific infiltration modeling is necessary for an adequate understanding of soil and water dynamics on a watershed scale, especially in rapidly urbanizing watersheds with significant anthropogenic influence on the soil. Land use is strongly tied to infiltration due to differences in vegetation characteristics and management practices. Management practices that are more intensive and frequent tend to decrease the soil's infiltration capacity through compaction and organic matter disturbance (Anderson et al., 2020). Understanding the land use impact on soil and water interactions is important when considering infiltration dynamics in a mixed-use watershed.

Forested Land

Forests are typically associated with relatively high infiltration due to abundant native forest vegetation and minimal anthropogenic disturbance. Vegetation plays a critical role in watershed soil and water conservation and watershed response to precipitation events. In most forested ecosystems, including the eastern hardwoods of the Appalachian region, the infiltration rate across the various soils exceeds the amount of rainfall in normal precipitation events, preventing significant surface runoff and erosion (Harden et al., 2003). Trees have extensive root systems comprising both large and fine root networks, providing numerous preferential flow paths for water to infiltrate the soil (Sun et al., 2018; Anderson et al., 2020). Additionally, in penetrating into the soil, tree root systems break up any existing compaction and increase the soil porosity, thereby lowering the bulk density and improving the infiltration rate (Harden et al., 2003). Forest soils also receive consistent annual additions of organic matter at the soil surface from leaf litter, raising the soil's organic matter content and maintaining their relatively high infiltration rates (Stuart & Edwards, 2006). Additionally, forest soils have high amounts of soil fauna, like earthworms, which aid in organic matter decomposition and positively affect infiltration through their burrows (Lee & Foster, 1991). These trends are evident in the meta-analysis conducted by Sun et al. (2018), which examined land use conversion experiments across China with infiltration data on bare land, cropland, shrubland, grassland, and forest. Average infiltration rate changes from all land use conversions showed that conversion of any of the other land uses to forest significantly increased infiltration rate (Sun et al., 2018). Forests were also found to have the highest infiltration rate across various land uses in a study by Anderson et al. (2020) in the Lower Mississippi River Valley in Arkansas, USA. More specifically, deciduous forests had a significantly higher infiltration rate than coniferous forests, native prairie,

agricultural lands, and Conservation Reserve Program sites. Conversely, if a forest is severely disturbed through conversion to another land use, the infiltration rate decreases over time while erosion and runoff increase due to the different vegetation characteristics of the other land use (Harden et al., 2003; Sun et al., 2018). For example, in the meta-analysis conducted by Sun et al. (2018), the average infiltration rate of forest to pasture conversions across various studies in China decreased as the complex forest root system was replaced by the smaller and finer root systems of grasses and forbs. And like with other land uses, forest management, such as prescribed fire, timber harvesting, and other silvicultural techniques, will affect infiltration. Prescribed burning is a frequently used silvicultural technique used for various management objectives, including favoring fire-adapted native vegetation, reducing competition, stimulating nutrients, reducing severe fire risk, and combatting invasive species (Fernandes & Botelho, 2003; DiTomaso et al., 2006). Ash is hydrophobic by nature, so ash leftover on the soil after a prescribed burn may temporarily reduce the infiltration rate (Bodí et al., 2014). However, the extent to which this phenomenon occurs depends on the spatially heterogeneous burn severity (Robichaud, 2000). Timber harvest reduces infiltration through removal of trees and soil compaction, specifically on skid trails. The severity of harvest effects on infiltration depends on the type of harvest and management used but given time forest soils can recover from this disturbance (Croke et al., 2001). The infiltration of forest soils is also variable depending on the canopy tree species and the characteristics of the forest floor under them. Leaf litter from different species have different chemical properties and promote different soil microorganisms when undergoing decomposition, which can result in different rates of infiltration and runoff (Neris et al., 2013).

Agricultural Land

Infiltration on agricultural land is highly dependent on management. Conventional tillage is a frequently used in agriculture to break up the soil and prepare a seedbed for planting. Conventional tillage may offer short-term benefits like temporarily increased aeration and infiltration. However, it often causes long-term detrimental effects on the soil after repeated annual use by degrading the soil structure, thereby increasing soil compaction and reducing soil infiltration. No-till practices preserve soil structure and better maintain the soil's infiltration capacity (Hill, 1990). A study led by Shukla et al. (2003) in the agricultural soils in the Appalachian region of Ohio reaffirms this idea. Their study examined the effects of manure application as an organic matter amendment and conventional versus no-till treatments on infiltration rate. The results show that the treatment incorporating manure and no-till practices resulted in the highest cumulative infiltration over three hours. Because no-till agriculture is a less intensive management practice, it prevents soil compaction and is better able to preserve the natural aggregates in pores in the soil when compared to conventional tillage. Incorporating organic matter into the soil is also beneficial as it further encourages the aggregation of soil particles, reducing the bulk density and creating more flow paths for water (Six et al., 2004). Agricultural soils also tend to have reduced infiltration capacity compared to less disturbed soils due to the characteristics of their root systems. Especially in monoculture row crop systems, like corn or soybean, the root systems are not very dense, do not penetrate very deep into the soil, and are mainly compromised of fine roots (Mengel & Barber, 1974). In a meta-analysis led by Sun et al. (2018) in China, results showed that conversion of any other land use except bare land (e.g., shrubland, grassland, and forest) to cropland produced a significantly lower infiltration rate due

to the loss of the diverse and complex root systems of native vegetation, as well as the negative effects of agricultural management practices.

Urbanized Land

Urban environments have a wide variety of surface cover, ranging from impermeable concrete to grassy lawns and parks, making soil and water interactions in developed areas extremely variable, complex, and difficult to characterize. However, soil infiltration and stormwater infrastructure are critical variables in urbanized watersheds when simulating stormflow and runoff (Hossain Anni et al., 2020). Several factors contribute to the complexity and diversity of urban soils. Anthropogenic disturbance by compaction and vegetation removal is extremely common and makes urban soils inherently different from other soils (Ali et al., 2021). Degree of compaction is considered the most important factor limiting the infiltration of urban soils (Pitt et al., 2000) as it contributes to frequent flooding in urban environments through increased levels of runoff.

Along with the physical and economic effects of flooding, runoff from urban environments also has significant environmental implications. Urban runoff carries higher concentrations of pollutants than runoff from forested or agricultural areas, worsening water quality downstream (Yang & Zhang, 2011). As the percentage of urban land in a watershed increases, erosion and runoff are also expected to increase due to the low infiltration rates from compaction associated with urban development (He, 2003; Ali et al., 2021). This effect is more pronounced the closer the urban development is to the stream network (He, 2003; Coulter et al., 2004). Additionally, flood risk is often underestimated in urban environments because the pre-development infiltration rate of the soil is commonly used in models, disregarding the effects of compaction and other disturbance during development (Gregory et al., 2016). This further

emphasizes the importance of developing a better understanding of soil infiltration in the urban environment. Increases in runoff have major implications on the hydrology of a watershed, including possible increased occurrence of flooding events. Modeling potential flood risk from urbanization is vital to sustainable urban planning and flood mitigation (Suriya & Mudgal, 2012). As mentioned previously, vegetation strongly influences infiltration, is often altered from its natural state, and is highly variable in urban environments. For example, sparsely vegetated areas, like medians and landscaping that frequently endure foot and vehicular traffic have a lower infiltration rate and higher runoff than larger and more densely vegetated areas, like urban parks (Ali et al., 2021). Urban green areas can be harnessed to reduce flood risk by increasing the infiltration rate if managed sustainably (Ren et al., 2020). Infiltration of urban soils can be enhanced through practices that reduce compaction or add organic matter to the soil, like restoring natural areas, composting, planting native species, and using porous construction materials (Ali et al., 2021). For example, planting native tree species that have wide root distributions improves the infiltration of urban soils because the roots can penetrate the subsoil and break up compacted layers by lowering the bulk density and increasing the porosity, thereby providing flow paths for water (Xie et al., 2020). At the same time, trees add organic matter to the soil annually through leaf litter. (Xie et al., 2020). Residential urban yards are also a valuable and significant aspect of the urban environment because they have enough infiltration capacity to capture urban runoff and mitigate some flood risk, especially if they are being managed for soil health (Shuster et al., 2014). The infiltration of yards can be enhanced further through minimizing impermeable surface coverage (like concrete patios), planting more trees, and adding organic amendments (Shuster et al., 2014). With all these potential sources of variability, finer scale, site-specific soil infiltration maps are needed to understand the urban soil response to

precipitation events fully (Shuster et al., 2014). Further study of the impact of urban environments on watershed hydrology is essential to advance urban watershed management decision-making and practices.

Digital Soil Mapping

Digital soil mapping (DSM) examines soil-landscape spatial relationships using geographic information systems and other spatial analysis tools (McBratney et al., 2003). It is an extremely useful approach for producing more accurate depictions of soil characteristics at the landscape scale and expanding the utility and information contained in conventional soil survey. DSM is necessary to predict soil properties over a spatial scale of interest. Pedotransfer functions have proven useful for predicting infiltration based on data on other soil properties in the same soil profile (Ghorbani-Dashtaki et al., 2016). However, these methods are not applicable for predicting soil properties across a watershed because they possess no spatial component and cannot be extrapolated over an area. In DSM, all predictor covariates must have a spatial component, meaning there is data available for all locations within an area to be able to make a prediction of a soil property at each location in an area. In DSM, soil classes or properties can be spatially predicted based on geographic position, environmental covariates, and/or other soil attributes from input data layers. The concept for predicting soils is mathematically represented through the *scorpan* model (McBratney et al., 2003), which expands on the five soil-forming factors of climate, organisms, relief, parent material, and time (clorpt), described by Hans Jenny (1941), to include soil properties and spatial position as well. It is formulaically represented as:

$$S = f(s,c,o,r,p,a,n)$$

where S represents either soil classes or attributes at a specific point, s are soil properties, c are climate variables, o represents organisms (flora/fauna/human activity), r includes

relief/topographic variables, *p* are parent materials, *a* represents age, and *n* is the spatial position (McBratney et al., 2003).

The *scorpan* model is more applicable to DSM than the *clorpt* model because it is inherently tied to geospatial position. With the layers of geospatial information (both continuous and categorical) associated with each of the factors, it can be used to predict and map soil information at specific locations, whereas *clorpt* is purely a conceptual model (Thompson et al., 2012). Data for these environmental covariates is obtained through use of remote and proximal sensing methods, like satellite and radar (McBratney et al., 2003; Boettinger et al., 2008).

Table 1. Environmental covariates and their relationship to the scorpan factors.

***scorpan* factor**	**Static covariates**	**Dynamic covariates**
Soil	SSURGO map unit	Brightness
Climate	Solar radiation, TWI	Wetness
Organisms	N/A	NDVI, Greenness, MID-Infrared Index, NDBI
Relief	Slope gradient, slope aspect, slope curvature, landform, topographic ruggedness, relative elevation, hillshade	N/A
Parent materials	Landform, SSURGO map unit	N/A
Age	N/A	N/A
Spatial position	N/A	N/A

Each environmental covariate targets some aspect of the *scorpan* equation (Table 1). Terrain data is useful for modeling soil classes and soil properties, including DSP like infiltration, because topography controls landscape-scale water movement. Slope gradient, slope aspect, slope curvature, topographic wetness index (TWI), solar radiation, topographic ruggedness, landform, and relative elevation will likely be useful predictors as they have been significant in similar DSM projects and have important intrinsic relationships to hydrology

(McBratney, 2003; Zhang et al., 2017; Pahlavan-Rad et al., 2020). Measures of slope steepness affect whether water infiltrates into the soil or runs off with gravity during precipitation events (Moore et al., 1991). Steeper slopes, in some cases, have been shown to have higher infiltration rates than gentler slopes. This phenomenon is attributed to the formation of high permeability micro-terraces and seal layers on steep slopes and more compacted soil crusts on gentle slopes under rainfall (Janeau et al., 2003; Assouline & Ben-Hur, 2006; Ribolzi et al., 2011). Thus, steeper slopes generate less runoff during rainfall events, however they do have greater amounts of soil erosion due to higher sediment concentration in the runoff that does occur (Assouline & Ben-Hur, 2006). Different slope aspects harbor different vegetative communities because northern and eastern slopes are more protected than southern and western slopes, meaning they get less solar radiation and thus retain more moisture. As such, north and east facing slopes are likely to have more species that prefer mesic conditions while south and west facing slopes are likely to have more species that prefer xeric sites (Desta et al., 2004). As discussed previously, vegetation will affect infiltration differently depending on their characteristics, like root distribution (Schaik, 2009). Also, as mesic sites are wetter, they tend to have higher levels of organic matter accumulation, which impacts the infiltration (Sun et al., 2018). Landform is related to measures of steepness and vegetative community, so may also show effects on infiltration in similar ways. Slope curvature is often quantified through the profile curvature, which is the curvature in the direction of the maximum slope, and the plan curvature, which is the form of the slope perpendicular to the direction of the maximum slope (Evans & Cox, 1999). Different profile and plan curvature combinations will produce variable effects on infiltration. Convexity and concavity of hillslopes will affect the infiltration purely through the steepness of their features. As such, exactly opposite hillslopes (similar slope gradients in both the vertical

and horizontal directions, regardless of convexity and concavity) will produce similar infiltration rates because they have roughly similar distributions of slope gradient (Wang & Chen, 2021). The TWI is an index that quantifies soil water content based on slope and the size of the catchment area. As mentioned, these factors play a role in the infiltration rate and are worth considering for the model (Beven & Kirkby, 1979).

Collecting environmental covariate data for DSM purposes has been greatly enhanced with remote sensing such as radar, LiDAR, and multispectral satellite imagery (McBratney et al., 2003). Satellite technology expands the capabilities of DSM use on a landscape scale. Soil reflectance data from satellite imagery have been used to derive soil properties like surface soil texture, surface roughness, surface temperature, and surface soil moisture, and have the potential to be applied to many more (McBratney et al., 2003; Gallo et al., 2018; Fathololoumi et al., 2021). Programs like Landsat, a project funded by the United States to collect remotely sensed imagery from space, offer access to 30-meter spatial and a 16-day temporal resolution global satellite imagery for no cost (NASA, 2022). Spectral data acquired from Landsat 8 Operational Land Imager (OLI) has been used to improve the efficiency and accuracy of DSM efforts (Zeng et al., 2020). One limit to this type of data is that it restricts the information collected to the properties of the surface soil. However, subsoil properties, including mineralogical and parent rock information, can be observed or inferred using radar and gamma radiometry (McBratney et al., 2003). These multiple types of data collected from databases, existing maps, remote sensing, and field measurements can be input as covariates for modeling the soil property of interest (McBratney et al., 2003). Inclusion of both static environmental covariates, like digital terrain model (DTM) derivatives, and dynamic environmental covariates, like multitemporal remotely sensed data, has been shown to improve prediction of both static and dynamic soil properties in

tree-based machine learning regression models, such as Cubist and RF (Fathololoumi et al., 2020). Fathololoumi et al. (2020) found that when including both types of environmental covariates into these two models, Cubist had better model efficiency than RF, but also had the most uncertainty and bias. Machine learning models are a powerful tool in DSM because of their ability to run the models as many times as desired with randomness in the observations and covariates in every run to find the most likely prediction for every individual pixel (McBratney, 2003). A brief overview of DSM projects and methods is provided in the Appendix.

One of the major barriers to DSM using environmental covariates other than DTM derivatives is that soil information cannot be readily retrieved from satellite or radar when the land surface is covered by vegetation for most of the year. This is especially true in the eastern United States, where much of the land is largely covered by deciduous forests. Using multi-temporal satellite images is one technique to resolve this issue (Diek et al., 2016; Gallo et al., 2018; Fathololoumi et al. 2020). Use of multitemporal satellite imagery offers the potential to enhance soil characteristic predictions in DSM applications. Gallo et al. (2018) found that multi-temporal satellite images effectively detect the variation in topsoil's texture. Multi-temporal satellite imagery minimizes the amount of land area covered by vegetation by selectively choosing images with the highest number of bare soil pixels and combining them into a "Bare Soil Composite Image." Multi-temporal composite images were also successfully used in a DSM effort by Diek et al. (2016) to predict the percentage of sand of topsoil using a partial least squares regression. An important caveat to be considered when using multi-temporal satellite imagery in DSM is correcting for the different values for the same pixel of bare soil in different years. Diek et al. (2016) used the empirical line method to correct for these differences. Use of multi-temporal images as a dynamic environmental covariate enhances the prediction of DSP,

like soil moisture, in DSM. Static environmental covariates alone are insufficient to effectively predict DSP because DSP are highly temporally variable and static environmental covariates have no temporal factor (Fathololoumi et al., 2021). There are many ways "multi-temporal" covariates can be applied in DSM models, varying in their use, purpose, and method of creation. For instance, multi-temporal satellite imagery can also be used itself as a covariate for modelling soil properties. Zhu et al. (2010) characterized a "land surface dynamic feedback pattern," which compiled images of six to seven consecutive dry days after a major rain event, measuring the change in soil reflectance as the land dried out. They found that under the same rainfall conditions, different soil types exhibited different temporal drying patterns that were detectable from satellite imagery. Liu et al. (2012) expanded on the findings from Zhu et al. (2010) by using these dynamic feedback patterns as environmental covariates for mapping soil surface texture. They found that using land surface dynamic feedback successfully predicted soil texture and using temporal landscape change as a dynamic environmental covariate offers potential in other DSM projects for predicting other soil properties.

DSM can be used to create detailed maps of DSP to inform management decisions, which is becoming increasingly necessary as society shifts its focus to sustainable development and management. This is especially needed on land uses with high levels of anthropogenic disturbance, like urban areas and intensive agricultural land, where the nature of the disturbance may be a challenge to traditional DSM and the soil's characteristics are highly variable due to the disturbance (Zhang et al., 2017).

Dynamic Soil Survey

The concept of a dynamic soil survey (DSS) is to address the need for soil information that can account for current land conditions, and to have it at a finer resolution for more effective

land management (Natural Resources Conservation Service, 2022). The current state of the soil survey provides valuable information regarding soil type, static soil properties, and recommended management in a polygon format. It also provides DSP estimates; however, it is not detailed enough to accurately capture their transient nature, so it is insufficient to be used in models and management applications. The goal for DSS is that it will exist in raster format to display static and dynamic soil properties and soil type at a much finer resolution (Natural Resources Conservation Service, 2022). Compared to the current soil survey, a DSS will provide land managers with a more accurate depiction of soil spatial and temporal variability (Natural Resources Conservation Service, 2022).

The ideas for needed improvement on the current soil survey that is available on Web Soil Survey (Soil Survey Staff, 2022) are not new, but until relatively recently had no concrete name or methodology. Although not specifically referred to as a "dynamic soil survey," Tugel et al. (2005) expressed the need for soil survey efforts in the future to incorporate elements of "soil change," or DSP. The DSS also evolves from the concepts of hydropedology, a sub-discipline in soil science focused on spatiotemporal interactions between pedological and hydrological processes on a landscape scale (Lin, 2003; Ma et al., 2017). As in the goals of a DSS, the field of hydropedology aims to translate hydropedological interactions to be useful on a management level via mapping, modeling, and monitoring (Ma et al., 2017).

Exploring DSM in the context of infiltration prediction is a relatively recent notion. Pahlavan-Rad et al. (2020) successfully predicted soil water infiltration in an arid Iranian floodplain using multiple linear regression and random forest models with an array of environmental covariates. Their RF model used 10 covariates, with the most important being stream order of channel networks, sand percentage, normalized salinity difference index,

elevation, and normalized difference vegetation index. The multiple linear regression model used only four covariates, with distance from river and sand concentration being the most important. While both models performed well, the authors decided the random forest was better due to its consistency with existing observations. Currently, several projects are active in mapping DSP, which all work towards creating a DSS; however, they are all recent enough that there are no publications on results yet (Natural Resources Conservation Service, 2022). This project will contribute to the creation of a DSS as well. This project's focus is unique compared to the other ongoing ones in that it addresses the application of these mapping and modeling approaches in an urbanizing environment. Also, infiltration is the primary output that will be modeled across the watershed, which has not yet been the focus of a DSS project in an eastern United States temperate deciduous landscape. This project will be the first component of several used to build a DSS in the West Run watershed.

CHAPTER 3. MATERIALS AND METHODS

Study Area

This project was conducted in the West Run watershed (WRW), which covers 23.3 km^2 in Monongalia County, West Virginia. West Run creek flows into the Monongahela River within the larger Ohio River watershed. Monongalia County experiences a humid continental climate (Beck et al. 2018) with four distinct seasons. The mean annual air temperature in Monongalia County is 12.4 degrees Celsius, and the mean annual precipitation is 106.7 cm (Northeast Regional Climate Center, 2022). Elevation in the West Run watershed ranges from 244 m to 427 m above sea level. The lowest areas are where West Run creek enters the Monongahela River and the highest at the headwaters (West Virginia Water Research Institute and West Run Watershed Association, 2008). Geologies in the WRW are mostly from the Paleozoic era, primarily of the Carboniferous age. As such, parent rocks of soils in this area include sandstone, limestone, coal, slate, and shale, but are most commonly sandstone and shale (West Virginia Geologic and Economic Survey, 1913).

The WRW is mixed-use watershed (Table 1; Fig. 1), predominantly supporting forested (42.7%), urbanized (37.7%), and agricultural land (19.4%), among other uses (0.2%) (Hubbart et al., 2022). Agricultural land in this watershed consists mainly of pasture and row crops. Natural vegetation is mixed northern hardwood forests, with the overstory largely dominated by oak and hickory (Griffith & Widmann, 2000). The WRW includes the northern sections of the city of Morgantown, West Virginia, which is rapidly expanding, thus continually increasing the percentage of urban area in this watershed and increasing number of flood incidents (West Virginia Water Research Institute and West Run Watershed Association, 2008).

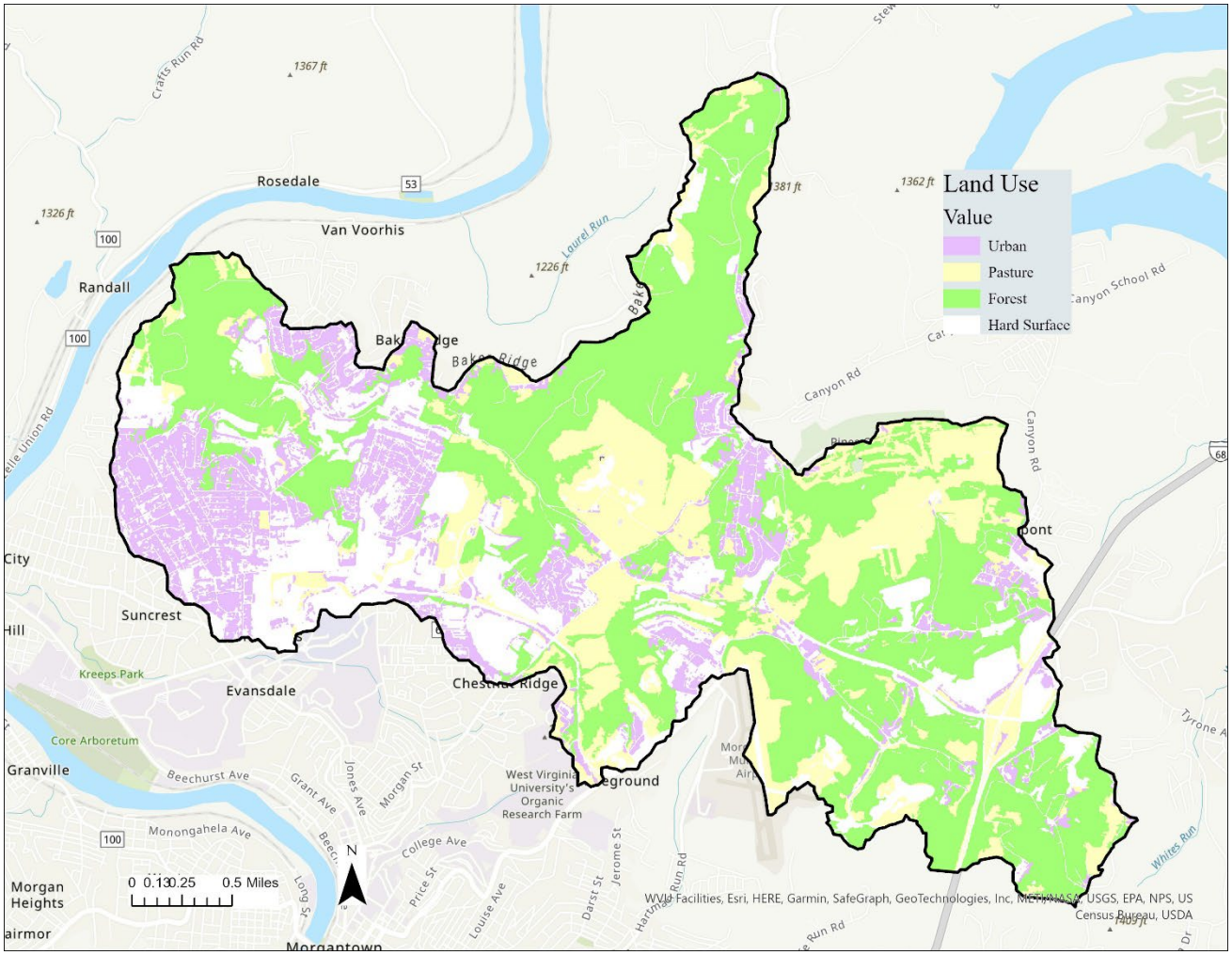


Figure 1. Land uses within the boundary of the WRW. Map created using ArcGIS software by Esri. Source data: World Topographic Map (World Topographic Map—ArcGIS Data Appliance | Documentation) and adapted 2016 National Agriculture Imagery Program (http://wvgis.wvu.edu/data/dataset.php?ID=489)

Twenty-one soil series are mapped in the WRW. The watershed's two most extensive map units are the Culleoka-Westmoreland and Dormont-Guernsey complexes (Table 2). Culleoka, Westmoreland, Dormont, and Guernsey soils are commonly used as pastureland, but their native vegetation is predominantly deciduous hardwoods. Monongahela soils typically support row crops like corn, soy, or pasture. Native vegetation is also primarily deciduous

hardwood species (Wright et al., 1982). The infiltration model created accommodates all land uses and most common soil types in the WRW.

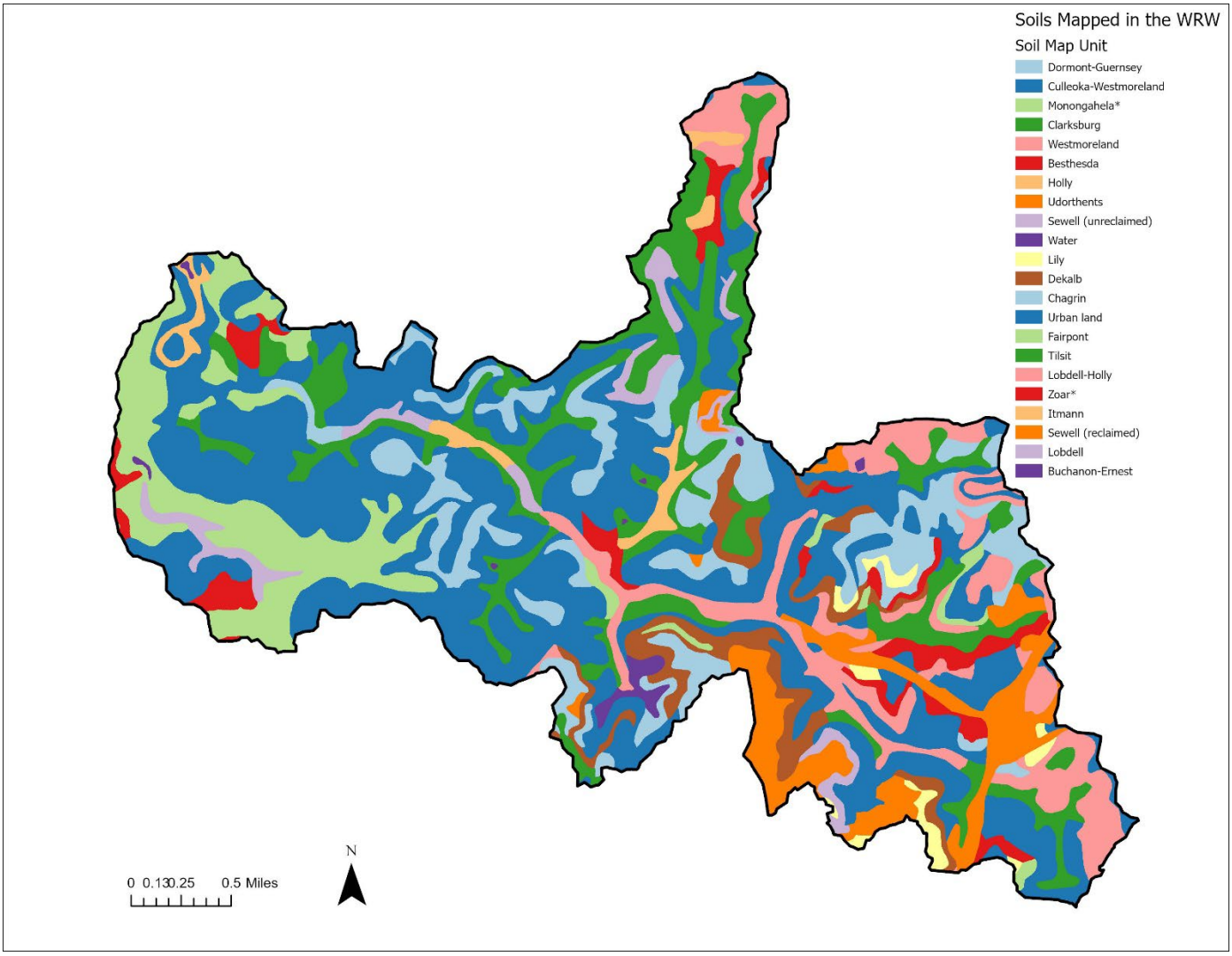


Figure 2. Soil map units located in the WRW. Map created using ArcGIS software by Esri. Source data: Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online at the following link: http://websoilsurvey.sc.egov.usda.gov/. Accessed [02/21/2022].

Table 2. Description of all map units within the WRW and extent mapped (Wright et al., 1982).

Map unit name	Symbol	Coverage (%)
Culleoka-Westmoreland silt loam, 25-35% slopes	CwE	15.212
Culleoka-Westmoreland silt loam, 15-25% slopes	CwD	8.508

Map unit name	Symbol	Coverage (%)
Clarksburg silt loam, 8-15% slopes	CkC	7.599
Urban land-Monongahela complex, 8-15% slopes	UmC	5.983
Udorthents	U1	5.780
Culleoka-Westmoreland silt loam, 8-15% slopes	CwC	5.680
Dormont-Guernsey silt loam, 8-15% slopes	DgC	5.263
Culleoka-Westmoreland silt loam, 35-65% slopes	CwF	3.940
Dormont-Guernsey silt loam, 15-25% slopes	DgD	3.541
Westmoreland silt loam, 8-15% slopes	WeC	2.957
Tilsit silt loam, 8-15% slopes	TlC	2.879
Dekalb very stony loam, 25-35% slopes	DdE	2.868
Lobdell-Holly silt loam	Lh	2.428
Monongahela silt loam, 8-15% slopes	MgC	2.362
Urban land-Culleoka complex, 15-25% slopes	UeD	2.355
Clarksburg silt loam, 15-25% slopes	CkD	2.289
Westmoreland silt loam, 15-25% slopes	WeD	2.059
Dormont-Guernsey silt loam, 3-8% slopes	DgB	1.333
Holly silt loam	Ho	1.311
Westmoreland silt loam, 35-65% slopes	WeE	1.285
Lobdell silt loam	Lb	1.233
Culleoka-Westmoreland silt loam, 3-8% slopes	CwB	1.069
Sewell very channery sandy loam, unreclaimed highwall, 15-25% slopes	ShD	0.880
Tilsit silt loam, 3-8% slopes	TlB	0.846
Westmoreland silt loam, 3-8% slopes	WeB	0.799
Urban land	Uc	0.765
Urban land-Zoar complex, 8-15% slopes	UzC	0.742
Bethesda channery loam, unreclaimed highwall, 15-25% slopes	BhD	0.693
Sewell very channery sandy loam, unreclaimed highwall, 35-65% slopes	ShF	0.650
Lily loam, 3-8% slopes	LaB	0.648
Bethesda loam, reclaimed, 15-25% slopes	BsD	0.570
Buchannon & Ernest very stony soils, 15-25% slopes	BeD	0.490
Zoar silt loam, 8-15% slopes	ZoC	0.483
Bethesda channery loam, unreclaimed highwall, 35-65% slopes	BhF	0.417
Dekalb very stony loam, 35-65% slopes	DdF	0.346
Lily loam, 8-15% slopes	LaC	0.331
Bethesda loam, reclaimed, 35-65% slopes	BsF	0.330
Chagrin silt loam	Cg	0.312
Dekalb channery loam, 3-8% slopes	DaB	0.311
Itmann very channery loam, 15-25% slopes	ImD	0.263
Zoar silt loam, 3-8% slopes	ZoB	0.250
Dekalb channery loam, 8-15% slopes	DaC	0.249
Fairpont channery silt loam, unreclaimed highwall, 35-65% slopes	FhF	0.239

Map unit name	Symbol	Coverage (%)
Urban land-Fairpont complex, reclaimed	FuD	0.225
Sewell channery sandy loam, reclaimed, 15-25% slopes	SlD	0.150
Clarksburg silt loam, 3-8% slopes	CkB	0.149
Fairpont channery silt loam, unreclaimed highwall, 15-25% slopes	FhD	0.146
Bethesda channery loam, unreclaimed, 15-25% slopes	BdD	0.139
Water	W	0.138
Dekalb channery loam, 35-65% slopes	DaE	0.121
Bethesda loam, reclaimed highwall, 15-25% slopes	BaD	0.118
Bethesda channery loam, unreclaimed, 35-65% slopes	BdF	0.091
Itmann very channery loam, 35-65% slopes	ImF	0.089
Sewell channery sandy loam, reclaimed, 0-8% slopes	SlB	0.088

Table 3. Brief description of all soil series present in the WRW (Wright et al., 1982).

Soil Series	Classification	Landform	Parent Material
Bethesda	Loamy-skeletal, mixed, active, acid, mesic Typic Udorthents	Upland	Mine spoil
Buchanan	Fine-loamy, mixed, superactive, mesic Aquic Fragiudults	Upland	Colluvium
Chagrin	Fine-loamy, mixed, active, mesic Dystric Fluventic Eutrudepts	Floodplain	Alluvium
Clarksburg	Fine-loamy, mixed, superactive, mesic Oxyaquic Fragiudalfs	Upland	Colluvium
Culleoka	Fine-loamy, mixed, active mesic Ultic Hapuldalfs	Upland	Colluvium or residuum
Dekalb	Loamy-skeletal, siliceous, active, mesic Typic Dystrudepts	Upland	Residuum
Dormont	Fine-loamy, mixed, superactive, mesic Oxyaquic Hapludalfs	Upland	Residuum

Soil Series	Classification	Landform	Parent Material
Ernest	Fine-loamy, mixed, superactive, mesic Aquic Fragiudults	Upland	Colluvium
Fairpont	Loamy-skeletal, mixed, active, nonacid, mesic Typic Udorthents	Upland	Mine spoil
Guernsey	Fine, mixed, superactive, mesic Aquic Hapludalfs	Upland	Colluvium and residuum
Holly	Fine-loamy, mixed, active, nonacid, mesic Fluvaquentic Endoaquepts	Floodplain	Alluvium
Holly	Fine-loamy, mixed, active, nonacid, mesic Fluvaquentic Endoaquepts	Floodplain	Alluvium
Itmann	Loamy-skeletal, mixed, semiactive, acid, mesic Typic Udorthents	Upland	Mine spoil
Lily	Fine-loamy, siliceous, semiactive, mesic Typic Hapludults	Upland	Residuum
Lobdell	Fine-loamy, mixed, active mesic Fluvaquentic Eutrudepts	Floodplain	Alluvium
Lobdell	Fine-loamy, mixed, active mesic Fluvaquentic Eutrudepts	Floodplain	Alluvium
Monongahela*	Fine-loamy, mixed, semiactive, mesic Typic cFragiudults	Terrace	Alluvium
Sewell	Loamy-skeletal, mixed, semiactive, acid, medic Typic Udorthents	Upland	Mine spoil
Tilsit	Fine-silty, mixed, semiactive, mesic Typic Fragiudults	Upland	Residuum

Soil Series	Classification	Landform	Parent Material
Udorthents	Udorthents	Upland	Cut & fill
Urban land	N/A	N/A	N/A
Water	N/A	N/A	N/A
Westmoreland	Fine-loamy, mixed, active, mesic Ultic Hapuldalfs	Upland	Residuum
Zoar*	Fine, mixed, semiactive, mesic Aquic Hapludults	Terrace	Lacustrine

*also mapped in an urban-land complex

Field & Laboratory Methods

The conditioned Latin hypercube sampling (cLHS) method, a form of stratified random sampling, was employed to select 50 points for data collection. An additional satellite point was selected nearby at each point generated from the cLHS sample. The satellite point was chosen to capture a different land use, soil type, or slope position. This sampling design was selected to align with the approach that has been adopted by the NRCS for other DSS projects. With this site selection process, the goal was to sample approximately 100 points. After removing some inaccessible points outside the watershed or having issues sampling, 86 points were sampled for use in the models. The WRW was stratified based on land use, soil survey map unit, slope gradient, topographic wetness index, normalized difference vegetation index, and slope aspect. The cLHS method was chosen to follow precedent set by similar DSM projects and for its cost function. In a cLHS, a cost raster can be imported to maximize the number of random plots within the West Run watershed that can fall on property owned by West Virginia University (WVU) while still maintaining a random sample to cover all variability in the watershed. Sample points on WVU property were prioritized because they were simpler to access and coordinate than points that fell on privately owned property. Two separate sampling campaigns took place, both using a cLHS with a cost raster to select sample points. The first took place throughout the

growing season of 2022 to collect data to be used as the training dataset to build the model. The second sampling was independent of the first sampling to collect a validation dataset and took place after the model had been built in the spring of 2023. This sampling for the validation dataset was stratified to cover the full range of infiltration measurements from the predictive map produced for infiltration.

At each plot, data were collected on infiltration, saturated hydraulic conductivity (Ksat), soil temperature, antecedent soil water content, slope gradient (%) of the landform of the site, slope gradient (°) of each infiltrometer head at the site, slope aspect (°), depth to bedrock (cm), land use, and dominant vegetation. Ksat was measured using fully automated dual-head infiltrometers (DHI) (Meter Group, WA, USA). The DHI were selected for this project due to the ease, speed, accuracy, and efficiency of their infiltration and Ksat measurements. The DHI were run using the optimized method outlined by Zhang et al. (2019). This method uses one long pressure cycle instead of two short pressure cycles, which Zhang et al. (2019) found to yield steadier infiltration rates. A pressure cycle consists of two distinct pressure head settings, standardly at 5 and 15 cm. Pressure head settings may be changed based on site variability. Sites with exceptionally high infiltration rates may need lower pressure heads, whereas sites with exceptionally low infiltration rates require a higher pressure head. Regardless of site requirements, a difference of 5 cm is always maintained between the high and low pressure heads. Three DHI were run at the same time to get three measurements on infiltration and Ksat at each plot. They were placed lateral to the plot center so that they all fall along the same slope gradient. The three measurements will be averaged to account for any spatial variability of soil properties. Adjacent to each DHI, three measurements of both soil temperature and antecedent soil water content were collected and averaged as well. Soil temperature at a depth of 5 cm was

measured in degrees Celsius with a digital thermometer from VEE GEE Scientific. Antecedent soil water content, as a percentage, was measured using the ML3 ThetaProbe Soil Moisture Sensor from Delta-T Devices, a moisture probe based on time domain reflectometry (Topp et al., 1980; Noborio, 2001).

Basic soil, landform, and vegetation descriptions was collected from each plot. Each of these measurements was taken at the plot center, except for depth to bedrock, which was measured once near each infiltrometer. Depth to bedrock was found using a tile probe as they are a minimally invasive tool. The tile probe used for this project is 124 cm long, which defined the maximum depth to bedrock that could be measured. Bedrock deeper than 124 cm was classified as "greater than 124 cm." Slope gradient was measured using a clinometer and slope aspect was measured using a compass. Land use and vegetation descriptions were made using the judgement and experience of those collecting the data. The soil series was determined after observing the soil profile through an auger boring and using those observations and the descriptions of soil series common in the WRW (Wright et al., 1982), along with the judgement and experience of the data collectors. The soil series determination will assist the subsequent creation of a digital soil map of the watershed. Additionally, three soil cores, with a 5.08-cm diameter and 5.08-cm length each, were collected at each plot where permission to sample was granted and taken back to the laboratory for further analysis of bulk density and particle size distribution. Bulk density was measured using the core method (Blake & Hartge, 1986). Auger descriptions and bulk

density samples were not taken at every site because they are a more invasive process and require landowner permission. Samples taken were archived for future analysis.

Figure 3. Forested sample site.

Figure 4. SATURO device running on forested site.

Modeling Methods

Predictive soil maps of infiltration were created using the random forest (RF) method, a complex of regression trees, in R (Breiman, 2001). RF was chosen for its common use in DSM, specifically in capturing the spatial relationships of DSP (Sihag et al., 2018; Pahlavan-Rad et al., 2020). Additional predictive maps were made using multiple linear regression (MLR) (McBratney et al., 2003). Both static and dynamic covariates were used as inputs in the model with infiltration as the dependent variable. The same covariates were offered to the selection method for each model. To test if the addition of dynamic covariates significantly improves the accuracy of the model for infiltration, another model was produced that only uses static environmental covariates as the explanatory variables. Models were evaluated based on goodness-of-fit parameters like R^2 and the root mean square error.

Environmental Covariates

Environmental covariates were derived from several different sources. Remotely sensed satellite imagery of the WRW is available through the Landsat programs funded by the National Aeronautics and Space Administration (NASA) and the U.S. Geological Survey (NASA, 2022; U.S. Geological Survey, 2022). Spectral data from recent imagery from the Landsat 8 satellite as digital numbers was used, specifically from the visible, near infrared, shortwave infrared, and thermal bands. Sentinel-2 was considered as well, but for the purposes of this project it did not provide better data than what was available through Landsat 8. Images selected were from over the growing season. Spectral indices were used as dynamic covariates in the model. The normalized difference vegetation index (NDVI) and MID-infrared index were used to separate vegetation and bare soil (Carlson & Ripley, 1997; Gallo et al., 2018):

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)}$$

$$MID - Infrared = \frac{(SWIR1 - SWIR2)}{(SWIR1 + SWIR2)}$$

The NDVI, which uses the near-infrared and red bands, is useful in distinguishing green vegetation from bare soil in satellite imagery (Carlson & Ripley, 1997). In addition, it can indicate moisture stress and overall vegetation health (Lotsch et al., 2003). Vegetation will appear greener following precipitation events and will appear less green under drought conditions (Anyamba & Tucker, 2005). Surface soil texture is related to water holding capacity and thus water availability to plants. As such, NDVI has been used to successfully predict surface texture in the southwestern U.S. (Maynard & Levi, 2017). In essence, greener vegetation has access to adequate amounts of soil water, and less green or wilting vegetation does not have enough soil water available. Information from the NDVI could be related to soil water dynamics, and thus infiltration rate. The MID-Infrared Index is typically used for bare straw and crop residue covering agricultural fields (Gallo et al., 2018). It could be useful in this watershed on cool season grasses in either pastures or lawns that tend to appear browner in the summer months. The normalized difference built-up index (NDBI) was used to separate impervious concrete and asphalt surfaces in the watershed from urban development. NDBI has been used as an environmental covariate to model other dynamic soil properties, like soil carbon (John et al., 2020). Since this project takes place in an urbanizing watershed, we expect the NDBI could be useful in the model to detect differences by land use. The NDBI uses the shortwave infrared and near infrared bands (Zha et al., 2003):

$$NDBI = \frac{(SWIR - NIR)}{(SWIR + NIR)}$$

A tasseled cap transformation was performed on the satellite images for an alternative view of the spatial distribution of soil reflectance (brightness), vegetation (greenness), and moisture content (wetness) in the watershed. This enhancement condenses the information in each satellite image to make it easier to interpret and compare (Crist, 1985).

Terrain derivatives were calculated from a DTM of the WRW. A DTM of the watershed was taken from the West Virginia Elevation and LIDAR Download Tool, a feature of the West Virginia Geographic Information System Technical Center, which provides LIDAR data for West Virginia at a one-meter resolution. All terrain covariates were calculated on a five-meter spatial resolution DTM. The static terrain covariates included slope gradient (as a percentage), topographic wetness index, slope aspect (as a factor), land use, planform curvature, profile curvature, total curvature, topographical position index (calculated with scales of 15, 30, and 45), transformed aspect, landform, and terrain roughness index (calculated with windows of three, five, and nine). The Gridded Soil Survey Geographic database contained two static soil covariates- available water storage and soil organic carbon content. Available water storage values were for the top 20 cm of the soil and represent the maximum amount of plant available water a soil can store. Soil organic carbon values represent the organic carbon content in the top 20 cm of the soil (Soil Survey Staff, 2023).

Six years of satellite imagery from Landsat 8 were used to create the mono-temporal dynamic environmental covariates, which represent values for the dynamic covariates on individual dates. All images were from June, during the growing season and had minimal cloud

cover. Dates used were June 22, 2013; June 12, 2015; June 14, 2016; June 1, 2017; June 23, 2019, and June 9, 2020. Landsat 8 satellite imagery for the necessary bands (red, green, blue, NIR, SWIR1 and SWIR2) for these indices is available for download from EarthExplorer at a 30 m spatial resolution. Downloaded images were resampled to a 5-m spatial resolution to match other data layers. NDVI, NDBI, Mid-infrared index, brightness, greenness, and wetness were calculated for each year at a 5-m spatial resolution.

Twelve multi-temporal covariates were calculated from the mono-temporal covariates. Multi-temporal covariates represent the quantifiable trend in the mono-temporal covariates through the years. Six were created simply by subtracting the 2013 layers from the 2022 layers. For example, the NDVI of 2020 minus the NDVI of 2013. This was repeated for each type of covariate. The other six multi-temporal covariates were created by calculating the pixel-by-pixel trend in the covariates across all years. Ultimately, the value for each pixel in the final trend layer represents the slope of the regression line from a linear model of that values of that pixel over the six years. For example, in the case of NDVI, a positive value in the trend layer would mean that pixel has become greener from 2013 to 2020. A value around zero would mean that there was little or no change, and a negative value would indicate that the pixel has become less green since 2013. This was done for each type of dynamic covariate.

Each environmental covariate created was resampled to 5-m by 5-m spatial resolution, clipped to the outline of the WRW, and projected to Transverse Mercator with a coordinate system of NAD 1983 UTM Zone 17N. All environmental covariates raster layers were stacked and the value of each layer at each sample site was extracted to build the predictive model.

Model Development and Validation

All models were developed using R statistical software (R Core Team, 2022). Packages used to create to models include the spatial, raster, glmnet, and caret packages. All random forest models were created using the RF regression tree method (Brungard et al., 2015; Fathololoumi et al., 2020). For the RF models, environmental covariates were selected using recursive feature elimination (RFE) (Brungard et al., 2015). RFE requires inputs for the number of features it should select, and then will fit a model and remove the least important variables in the model until it reaches the best of the specified amount (Gregorutti et al., 2017). Environmental covariates for the MLR models were selected using Least Absolute Shrinkage and Selection Operator (LASSO), a penalized regression approach using a shrinkage parameter (Tibshirani, 1996). LASSO can identify important variables for modeling by shrinking the coefficients of the less significant variables to zero (Hastie & Friedman, 2001). Models were trained using 10-fold cross validation. The performance of all predictive maps was assessed using goodness-of-fit parameters like R^2, lowest mean absolute prediction error (MAE), and the root mean square error (RMSE). The best model for continuous data will have a relatively high R^2, low MAE, and low RMSE.

$$R^2 = 1 - \frac{\sum_{i=1}\left(y_{i_{observed}} - y_{i_{predicted}}\right)^2}{\sum_{i=1}\left(y_{i_{observed}} - \bar{y}_{predicted}\right)^2}$$

$$MAE = \frac{1}{n}\sum_{i=1}\left[\frac{\left|y_{i_{observed}} - y_{i_{predicted}}\right|}{y_{i_{observed}}}\right]$$

$$RMSE = \sqrt{\frac{\Sigma_{i=1}\left(y_{i_{observed}} - y_{i_{predicted}}\right)^2}{n}}$$

Where n is the number of samples, $y_{i_{predicted}}$ are the predicted infiltration values, $y_{i_{observed}}$ are the observed infiltration values, and $\bar{y}_{predicted}$ are the mean of the predicted infiltration values.

The final model was validated with a separate independent sampling. This sampling proceeded similar to the first sampling for training data, using a cLHS design. This sampling was stratified using the infiltration values from the final model and by land use to ensure that all infiltration values and land uses would be represented in the validation data. Twenty points were selected for the validation sample, five on each land use: urban, pasture, forests, and hard surfaces. Validation R^2 concordance, and RMSE of the final models will be found.

CHAPTER 4. RESULTS

In total 86 saturated hydraulic conductivity (Kfs) estimates were collected during the growing season of 2022 using the cLHS sampling design in WRW. Thirty-three sites were urbanized, 27 were in pasture, 23 were forested, and 10 were hard surfaces, which included roads and roofs. Hard surfaces make up a large portion of the WRW since it is an urbanizing area, so they were important to represent in the models. Instead of taking a measurement with an infiltrometer on the pavement and roofs, infiltration was assumed to be zero. Saturated hydraulic conductivity (Kfs) measurements varied across the sampled sites (Table 4). Excluding the hard surfaces (for which infiltration was not measured and were assigned a Kfs value of 0 mm hr^{-1}), the next lowest Kfs value was 12.68 mm hr^{-1}, which occurred on an urbanized site. The highest measured infiltration rate of 1606.56 mm hr^{-1} was also on an urbanized site. However, this observation was removed from the modeling process as it was found to be an outlier based on a chi-squared test. Excluding that observation, the highest measured Kfs was 823.12 mm hr^{-1} on a forested site. The forested land use had the highest median infiltration rates among the sites (Fig. 6). Medians of the urban and pasture sites were similar. The urban land use had the largest variation in the Kfs values (Table 4).

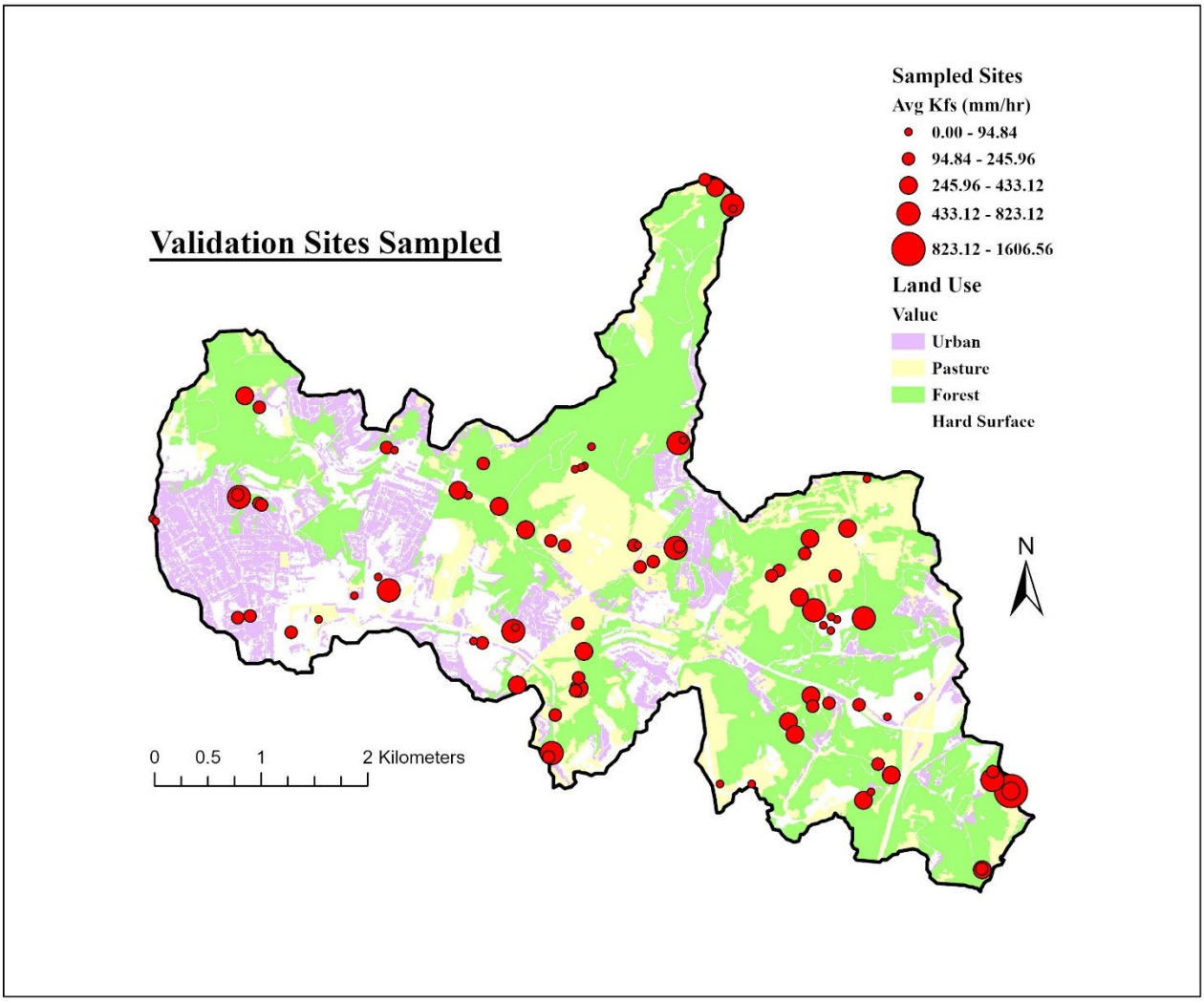


Figure 5. Sites sampled and associated Kfs (mm/hr) for the calibration dataset. Map created using ArcGIS software by Esri. Source data: and adapted 2016 National Agriculture Imagery Program (http://wvgis.wvu.edu/data/dataset.php?ID=489).

Table 4. Summary statistics of calibration sample of infiltration.

Land Use	n	Minimum (mm/hr)	Maximum (mm/hr)	Mean (mm/hr)	Median (mm/hr)	Standard Deviation (mm/hr)
Urban	30	12.69	1606.56	234.6	151.72	299.39
Pasture	25	15.73	433.12	190.14	160.87	128.78
Forest	21	157.87	823.12	406.18	385.68	208.43
Hard Surface	10	0	0	0	0	0
Total	86	0	1606.56	237.03	182.9	245.08

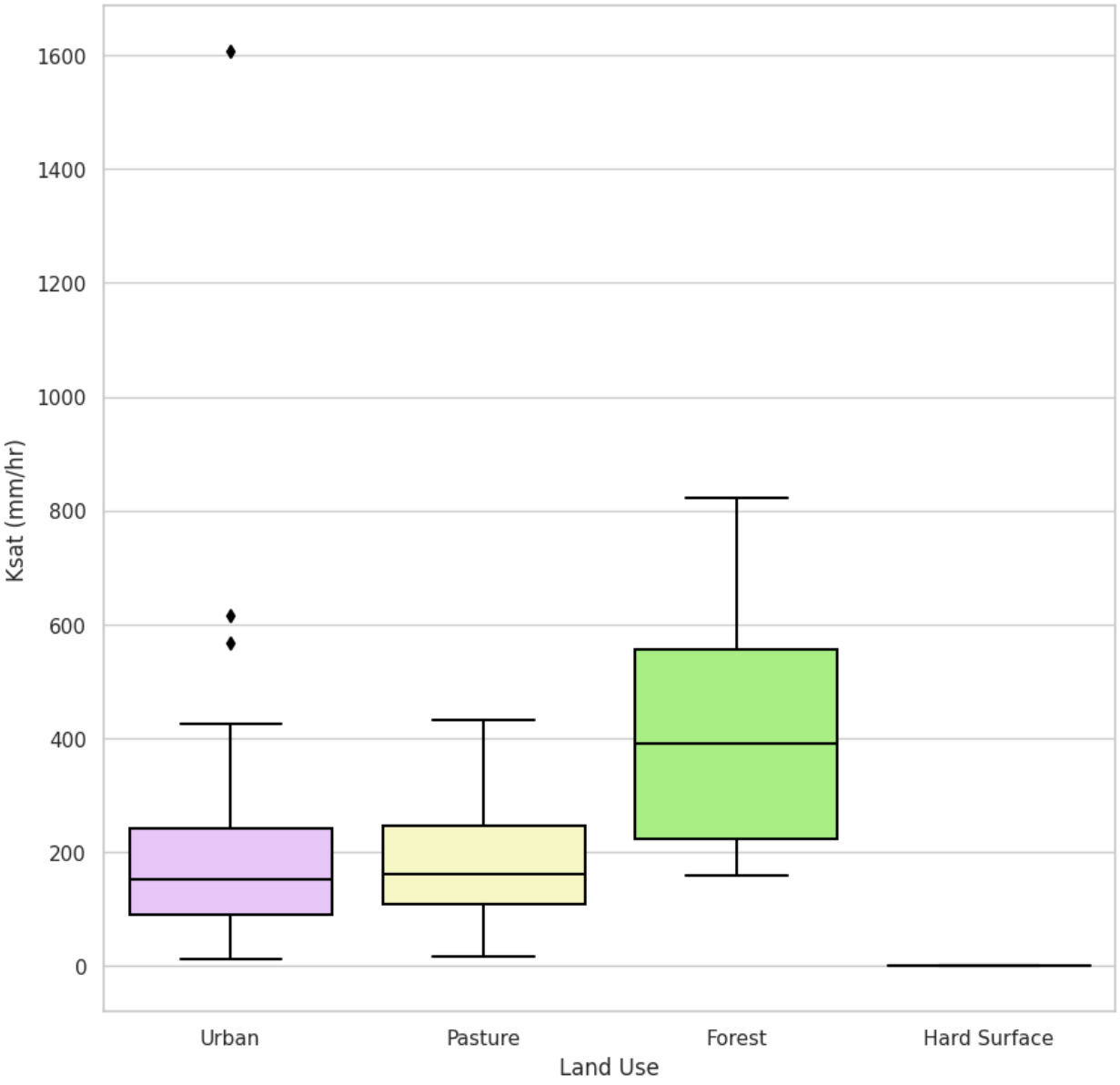


Figure 6. Boxplots of Kfs (mm/hr) by land uses in the WRW.

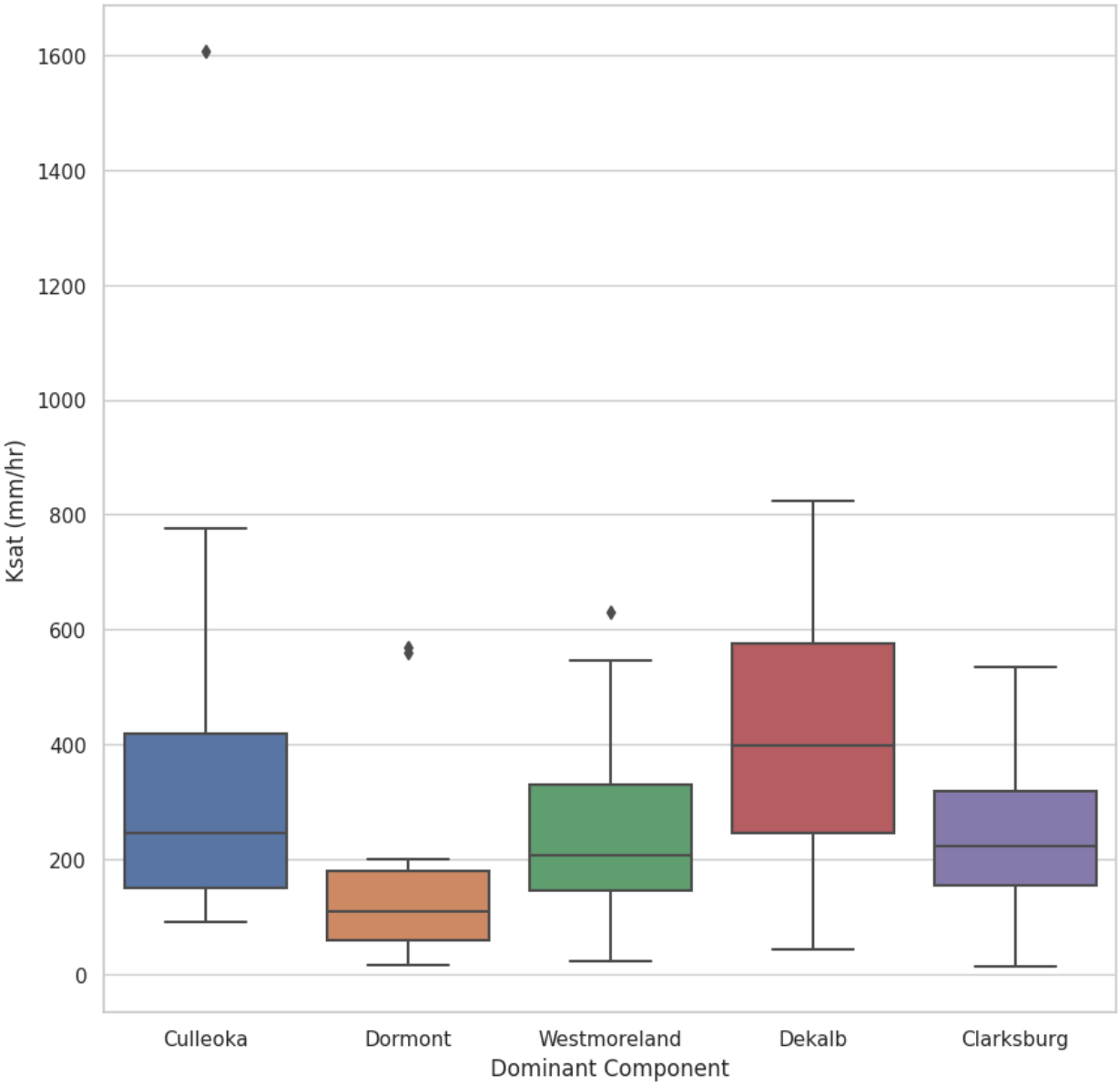


Figure 7. Boxplot of Kfs (mm/hr) by most common (spatial extent) soil series mapped in the WRW.

The square root of the Kfs values was used for model building to meet the normality assumption for linear models. In total, 66 environmental covariates (Table 6) were offered for selection in the models for prediction of square root of Kfs. Of the 66 covariates, 18 were static, 36 were mono-temporal dynamic, and 12 were multi-temporal dynamic (Table 6).

Table 5. All environmental covariates offered for selection for the models.

Mono-temporal dynamic[1]	Multi-temporal dynamic	Static[4]
NDVI	NDVI trend[2]	Slope gradient
NDBI	NDBI trend	Topographic wetness index
MID-IR	MID-IR trend	Slope aspect
Brightness	Brightness trend	Land use
Greenness	Greenness trend	Planform curvature
Wetness	Wetness trend	Profile curvature
	NDVI difference[3]	Total curvature
	NDBI difference	Transformed aspect[5]
	MID-IR difference	Terrain ruggedness index[6]
	Brightness difference	Topographic position index[6]
	Greenness difference	Landform
	Wetness difference	Soil organic carbon (0-20 cm)
		Available water storage (0-20 cm)
		Parent Material

[1] Each mono-temporal dynamic environmental covariates were derived from Landsat 8 on six different dates: 22 June 2013, 12 June 2015, 14 June 2016, 1 June 2017, 23 June 2019, and 9 June 2020.
[2] Each multi-temporal dynamic trend covariate was derived as the slope of the best-fit linear regression model through the six corresponding Landsat 8 scenes from 2013 to 2020.
[3] Each multi-temporal dynamic difference covariate was derived as the difference between the 2020 and 2013 Landsat scenes of the corresponding mono-temporal dynamic covariates.
[4] Each static environmental covariate was calculated from the LiDAR-derived DTM at 5 m, except for soil organic carbon (0-20 cm), available water storage (0-20 cm), parent material, which were derived from the gSSURGO database.
[5] Transformed aspect assigns slopes a value based on north-northeastness (zero), indicating cooler and wetter slopes, and south-southwestness (one), indicating hotter and drier slopes.
[6] Calculated at three spatial extents: 15x15 m, 30x30 m, and 45x45m moving window.

LASSO was used for covariate selection for the MLR model. The best lambda for the LASSO model was 0.988 and the R^2 was 0.313. The LASSO selected nine covariates for use in the MLR model: NDBI of 2017, NDBI of 2019, wetness of 2019, greenness of 2016, wetness of 2015, brightness of 2013, terrain ruggedness index using a window with a radius of nine pixels (45 m), topographical position index using a window with a radius of nine pixels (45 m),, and

soil organic carbon in the top 20 centimeters. To assess multicollinearity, these nine variables forced into a regular MLR and their variance inflation factors were tested. Variables with variance inflation factors greater than five were removed one at a time, as five has been found as a threshold between acceptable and concerning variance inflation factors based on the degree of multicollinearity (Gareth et al., 2013). Based on this, wetness of 2019 was removed. Rerunning the MLR without wetness produced variance inflation factors that were all less than five. Model building proceeded with the remaining eight covariates for the final MLR. The final MLR was developed using a 10-fold cross-validation. The R^2 of the cross-validated MLR model is 0.332, and the adjusted R^2 is 0.302. The RMSE and MAE were 45.70 mm hr^{-1} and 29.42 mm hr^{-1}, respectively.

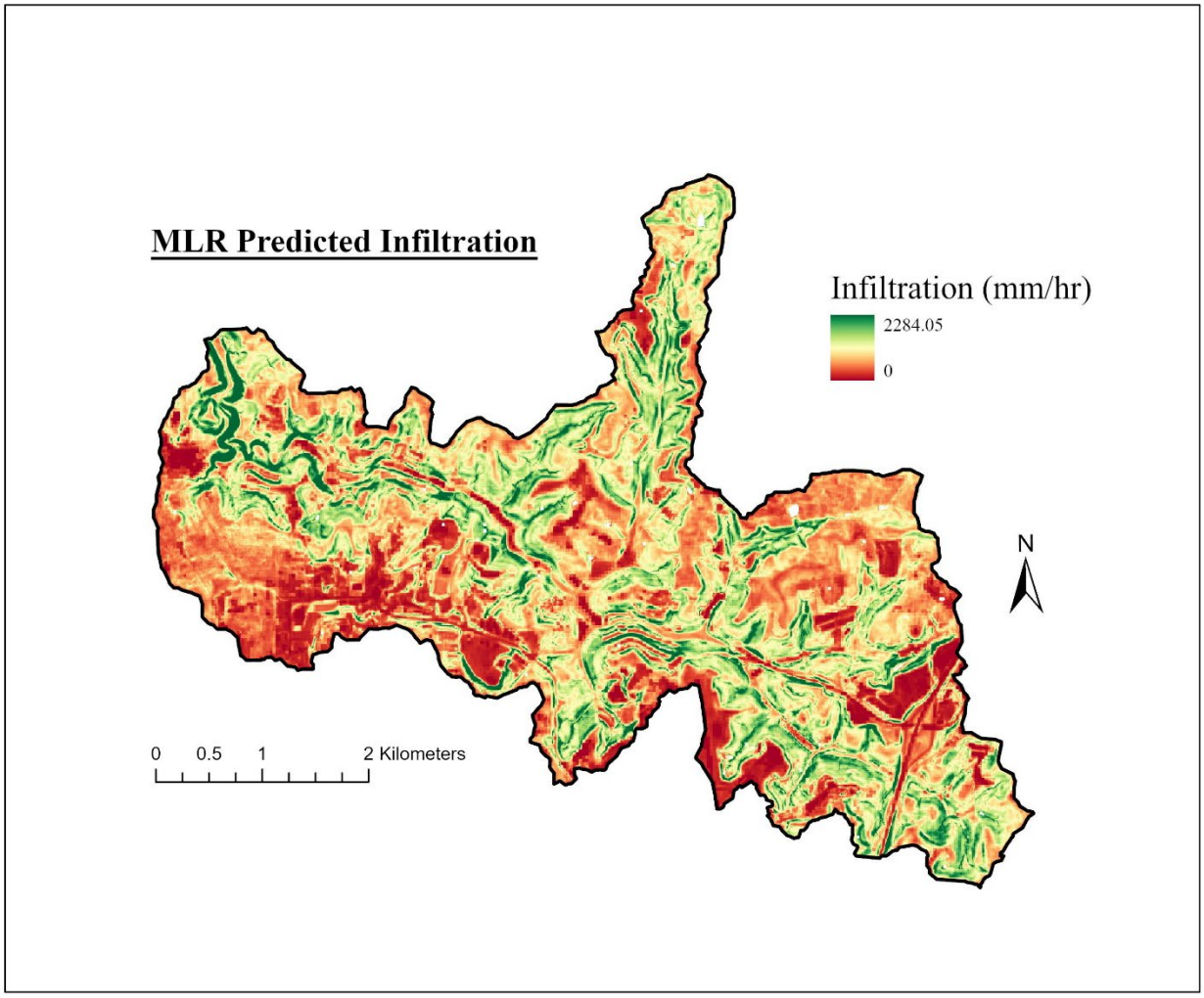


Figure 8. Map of the predicted square root of infiltration rate from the MLR model. Map made by ArcGIS software by Esri.

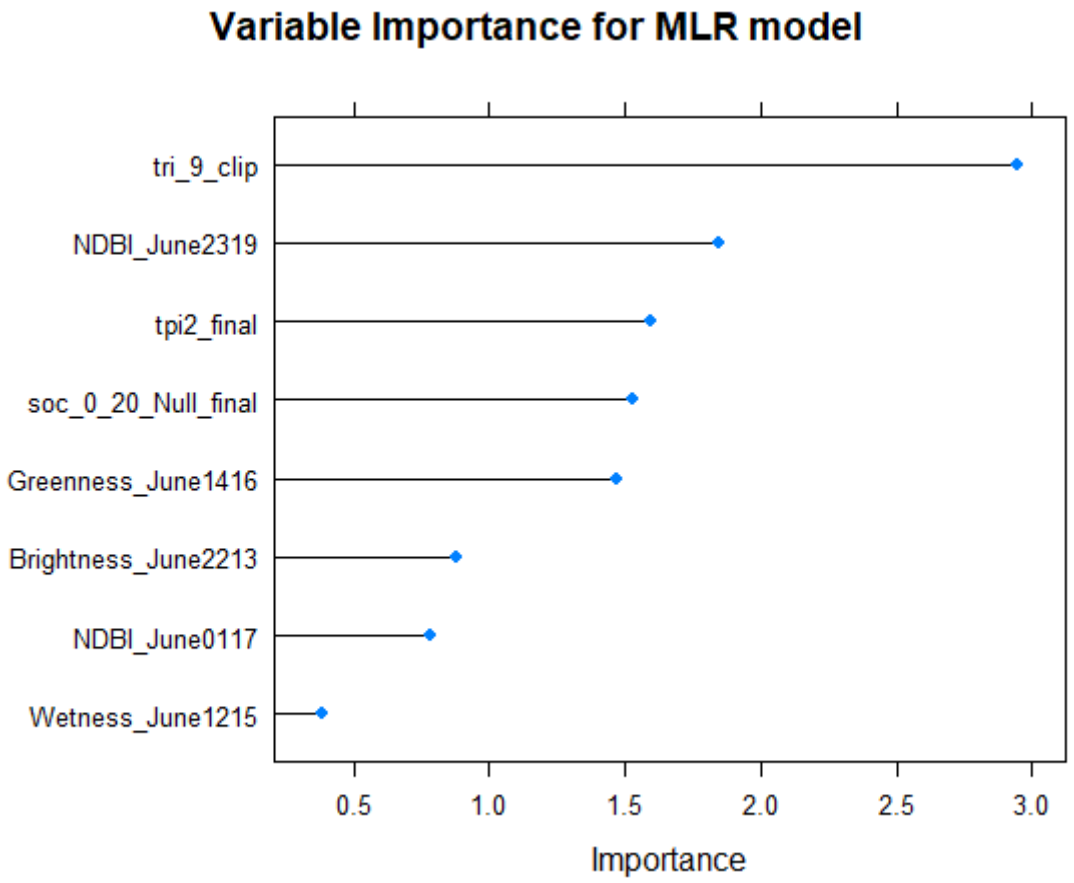


Figure 9. Variable importance in the MLR model.

Table 6. Coefficients of covariates used in the MLR model.

EC	Coefficient (mm/hr)
NDBI 2019	-8.12E+05
NDBI 2017	3.73E+05
Greenness 2016	-2.51E+00
Wetness 2015	8.68E+00
Brightness 2013	-1.36E+01
TRI at 45 m	9.22E+03
TPI at 45 m	3.60E+04
SOC (0-20 cm)	3.12E+01

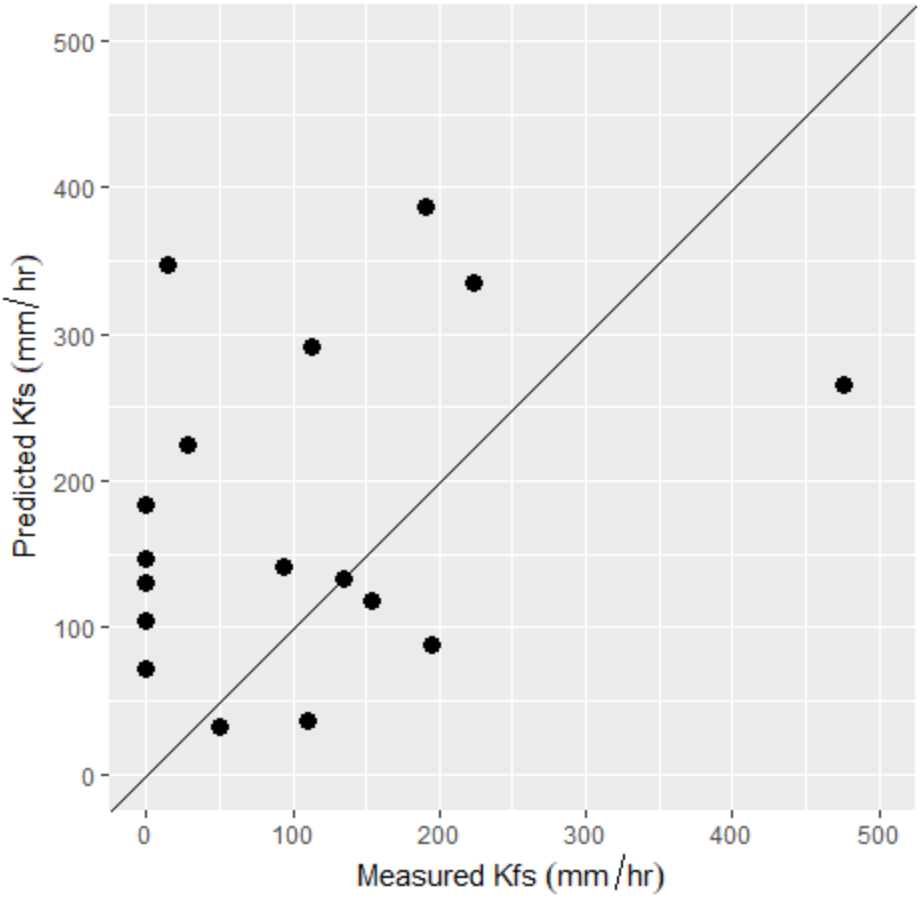


Figure 10. Predicted vs. Observed (measured) Kfs (mm/hr) for the MLR model.

Recursive feature elimination was the covariate selection method for the random forest model. Five covariates selected by recursive feature elimination produced the best model for random forest: slope gradient, NDBI of 2019, wetness of 2015, NDVI of 2019, and NDBI of 2017. The RF model was developed using a 10-fold cross validation. The parameters of the random forest model were the default with 500 as the number of trees and two variables tried at each split. The final RF model explained 20.1% of the variance of infiltration rate in the WRW, which is equivalent to an R^2 of 0.201.

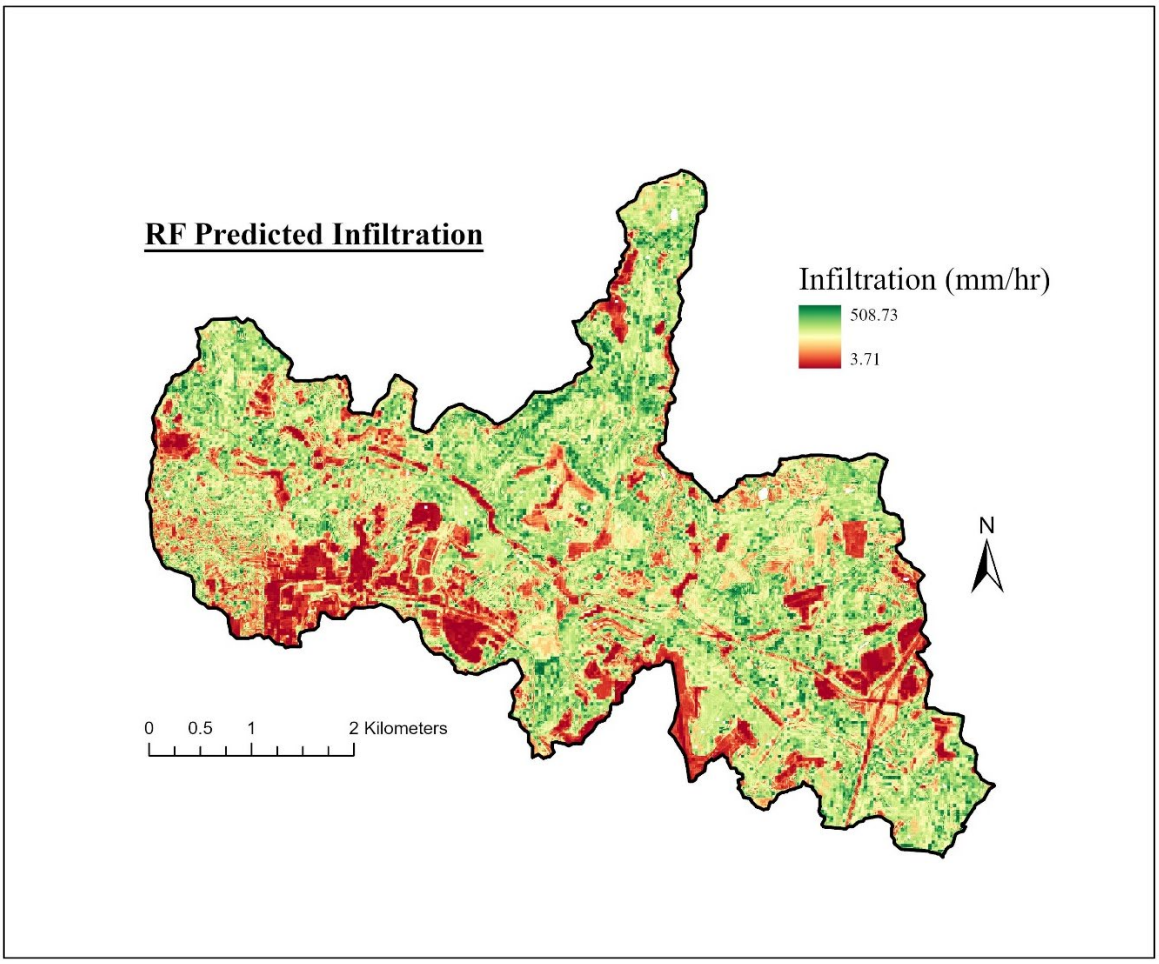


Figure 11. Map of predicted square root of infiltration rate in the RF model. Map made using ArcGIS software from Esri.

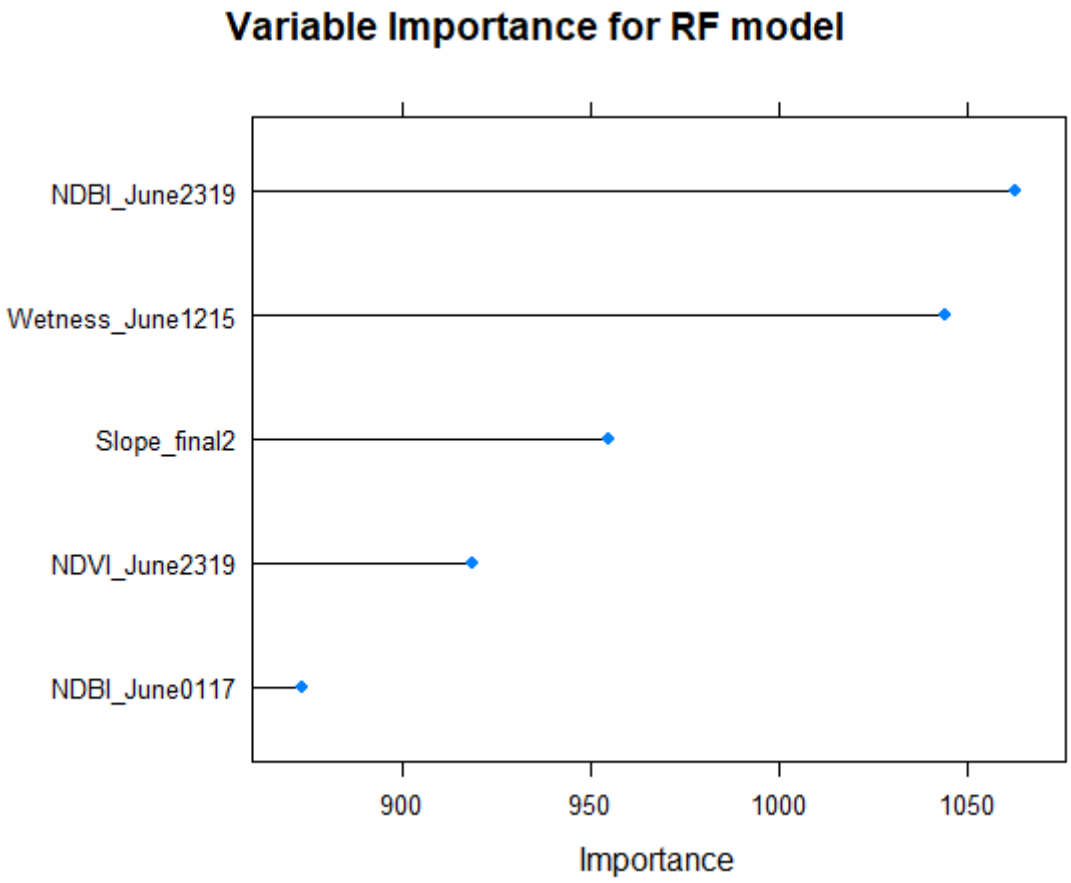


Figure 12. Variable importance in the RF model.

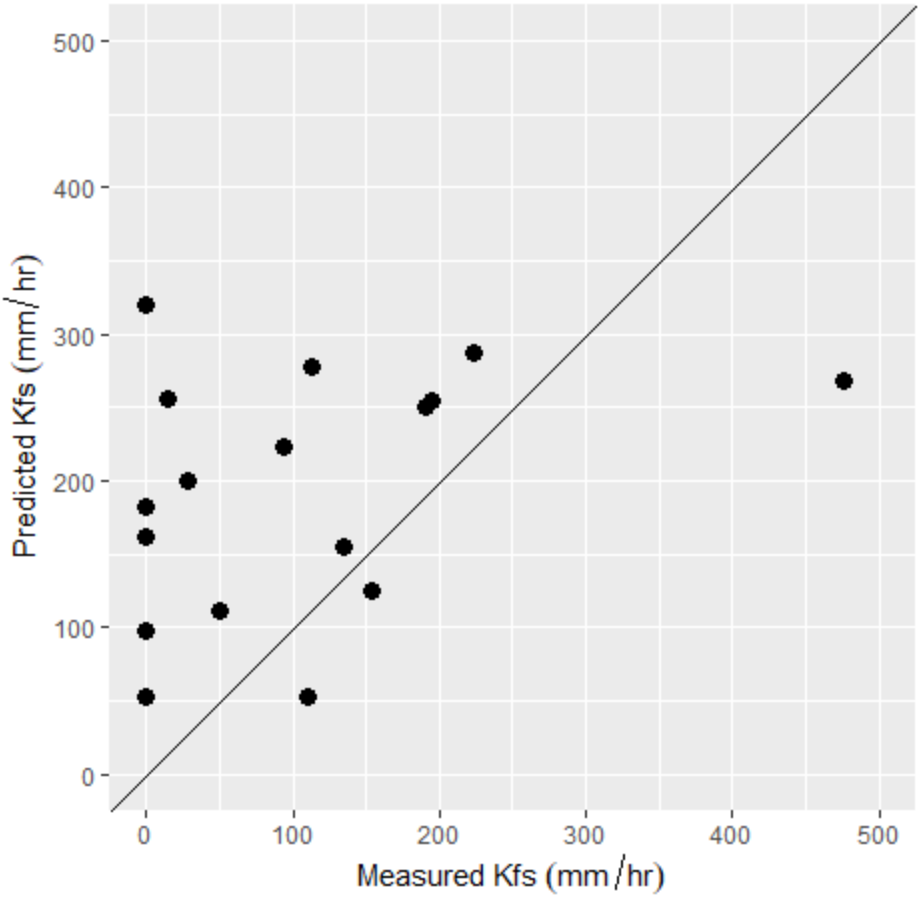


Figure 13. Predicted vs. Observed (measured) Kfs (mm/hr) for the RF model.

The MLR was chosen as the best model because it had the largest R^2 and visually agreed with what was expected in certain watershed areas. As such, this model was used as input into the cLHS for a random sample of validation points. In total, 17 validation points were sampled. This sampling consisted of three urban, five pasture, four forest, and five hard surface sites. The lowest Kfs value of the validation sample (other than the hard surfaces, which were assigned values of 0 mm hr^{-1}) was 42.35 mm hr^{-1}, whichwas on a forested site. The highest value (1426.14 mm hr^{-1}) was also on a forested site. The validation sample yielded an R^2 of 0.080 and 0.103 for the MLR and RF, respectively. The MLR had a concordance value of 0.178 and the RF's concordance was 0.156.

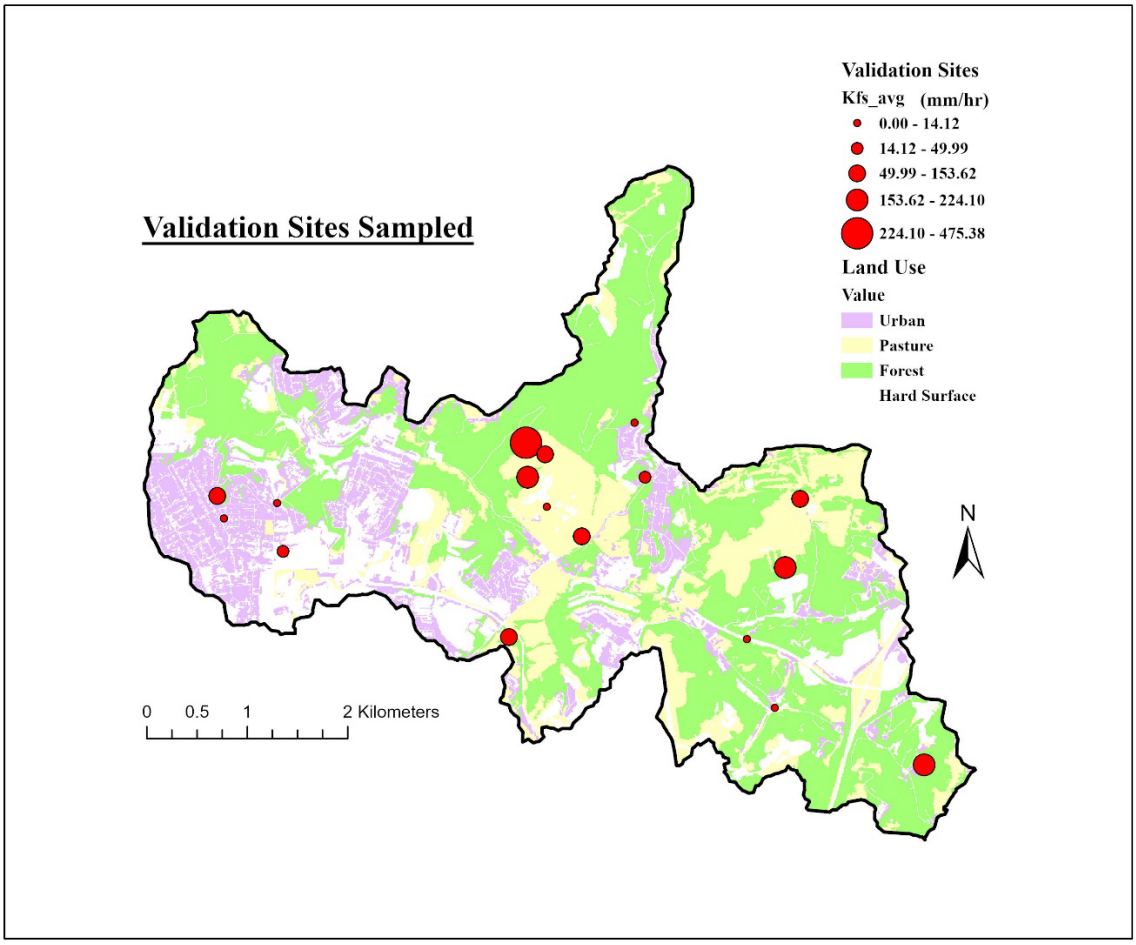


Figure 14. Sites sampled and associated Kfs (mm/hr) for the validation set. Map made using ArcGIS software by Esri.

Table 7. Summary statistics of validation sample of infiltration.

Land Use	n	Minimum (mm/hr)	Maximum (mm/hr)	Mean (mm/hr)	Median (mm/hr)	Standard Deviation (mm/hr)
Urban	3	99.97	153.62	103.7	99.97	48.17
Pasture	5	113.5	585.4	302.53	220.93	191.25
Forest	4	42.35	1426.14	565.66	397.08	613.99
Hard Surface	5	0	0	0	0	0
Total	17	0	1426.14	240.38	113.5	359.64

Table 8. Validation statistics of MLR and RF models.

Model	**R^2**	**Concordance**	**MSE**	**RMSE**
MLR	0.080	0.178	65.76	8.11
RF	0.103	0.156	68.87	8.30

CHAPTER 5. DISCUSSION

The infiltration characteristics by land use generally follow expected trends, with the highest median infiltration on forested land (Sun et al., 2018; Anderson et al., 2020) and urban sites having a high variability (Shuster et al., 2014; Ali et al., 2021). The goals of this study were to represent infiltration, a DSP, over a watershed using DSM methods and to determine the utility of static, mono-temporal dynamic, and multi-temporal dynamic covariates to make that prediction. The resulting predictive models indicate that using these ECs explains some of the variability of infiltration in the WRW, as displayed by their R^2 values of 0.302 and 0.201 for the MLR and RF, respectively. The validation R^2 values were lower, at 0.080 and 0.103 for the MLR and RF, respectively. There was one forested observation in the validation dataset that was extremely overpredicted by both models. This point occurred on the edge of a forest and was completely covered in honeysuckle with no real overstory canopy. This area was bordered by an urban area and had refuse from the nearby residential area mixed in with the soil. We kept this point in the dataset because it is a valid point in the watershed. Still, it threw off the model because it predicted that area as having the infiltration of a normal, less disturbed forest in the watershed. When that point is removed from the validation set, the R^2 values for the MLR and RF increase to 0.138 and 0.134, respectively. Without that point, the concordance increased to 0.243 and 0.188 for the MLR and RF, respectively. The models struggled to predict any of the hard surfaces correctly, as they all should be zero. For the MLR model, the predicted values of the hard surfaces ranged from 71.97 to 182.99 mm hr^{-1}, and for the RF they ranged from 52.52 to 320.55 mm hr^{-1}. To improve the representation of the infiltration rate of hard surfaces, the land use layer was used to force all hard surfaces to have a value of zero while leaving predictions on all other land uses unchanged. The result of these modified prediction maps improved the

goodness-of-fit metrics for the validation samples, producing an R^2 of 0.52 and a concordance of 0.66 for the MLR and an R^2 of 0.65 and a concordance of 0.73 for the RF map outputs.

Evaluating the quality of a DSM model can be somewhat subjective due to the wide variety of applications in which DSM can be applied. Generally, for modeling efforts of variables in nature, like soil, some researchers contend that R^2 values of less than 0.5 should be considered unacceptable predictions. In contrast, those greater than 0.5 but less than 0.75 should be considered acceptable predictions, and those above 0.75 should be considered good predictions (Chang et al., 2001; Shepherd & Walsh, 2002; Wang et al., 2013). This scale of acceptability is not broadly applicable to most DSM projects. Acceptable R^2 and other goodness-of-fit metrics vary by project according to the spatial scale, the quality of the covariates, and the nature of the targeted soil property for prediction. Generally, lower R^2 may be considered more acceptable and useful for DSM projects over a large spatial extent and/or more difficult to predict soil properties, which include many DSPs. Conversely, DSM projects predicting more well-understood soil properties or on smaller spatial extents, higher R^2 values are expected. This is demonstrated by the results of Reddy et al. (2021), who predicted several soil properties across India. In their study, the R^2 varied by soil property, with R^2 ranging from 0.66 to 0.73 for pH and from 0.16 to 0.24 for sand content. Similar R^2 values are observed in Mulder et al. (2016), who mapped various soil properties across France. For predictions of the soil surface, they had a R^2 of 0.55 (0.48 validation) for pH, 0.36 (0.29) for soil organic carbon, 0.33 (0.24) for cation exchange capacity, and 0.24 (0.17) for coarse fragment percentage. These R^2 values decreased when predicting these same soil properties for deeper depths in the soil profile (Mulder et al., 2016). For predicting organic carbon over a watershed in Iran, Fathololoumi et al. (2020), the best models for organic carbon had a R^2 of 0.69 and for sand content 0.59. Over a watershed in

Brazil, Gallo et al. (2016) had R^2 values of 0.56 for sand content, 0.40 for cation exchange capacity, and 0.34 for soil organic matter content. As demonstrated, many different factors influence the evaluation of DSM models. Given that this project occurred in a small watershed and that infiltration is difficult to characterize and not yet well-understood spatially, an adjusted R^2 of 0.302 is adequate for this modeling effort. There are few DSM projects predicting infiltration thus far, so it is difficult to make comparisons. Still, observations from this project will be beneficial in informing future DSM projects for infiltration and other DSPs, especially in highly vegetated environments.

The multi-temporal ECs created were not selected for either the MLR or the RF models, suggesting they were not useful for predicting DSPs. Other types of multi-temporal covariates have been useful in mapping other DSPs, as in Diek et al. (2016), Gallo et al. (2018), and Fathololoumi et al. (2020, 2021). All these researchers explored ways to incorporate multi-temporal dynamic ECs into DSM, resulting in unique methods to create these multi-temporal covariates, such that their end products differ from the type of multi-temporal covariate used in this project. Whereas most of the multi-temporal covariates in the current literature are "multi-temporal" in the sense that they aggregate multiple satellite images taken at different times into a single product to get a complete view of the targeted dynamic covariate, our "multi-temporal" covariate aimed to quantify a change in a dynamic covariate over time by analyzing the pixel-by-pixel linear trend over several satellite images from various years. This presents a challenge for comparing the utility of multi-temporal covariates for DSM, specifically of DSPs, but nonetheless offers valuable information for the vast number of methods and applications that could be used. For instance, Fathololoumi et al.'s (2020, 2021) studies had multi-temporal models that used various dynamic covariates over one growing season. Unlike our study, their

studies took place in a semi-arid environment where vegetation had less of an impact on the study. As such, their most important factors for predicting soil properties were soil surface satellite properties including emissivity, land surface temperature, albedo, and incidence. Their best models were the "multi-temporal" ones comprised of dynamic ECs from any month over the growing season. Being in a vegetated and urbanized environment, our study primarily relied on covariates that reflected the vegetative and urban spatial patterns, like NDVI, NDBI, Greenness, Wetness, and Brightness. For our study, these patterns were best represented through the mono-temporal dynamic covariates instead of our multi-temporal options. Diek et al. (2016) and Gallo et al. (2018) were both presented with the challenge of modeling with satellite imagery in a temperate environment to create a predictive model of sand. They overcame this through their version of multi-temporal dynamic covariates—using multiple satellite images to patch them together into one bare soil composite image. Since these studies were both in highly agricultural regions, their multi-temporal covariates were able to take advantage of the crop management and harvest patterns each year that would expose different areas of the soil. This method is inapplicable for creating a bare soil composite image in heavily forested environments like Appalachia, where most vegetation covers the ground for decades to hundreds of years, as opposed to the cultivation observed each year in most crop systems. Maynard & Levi (2017) created multi-/hyper-temporal covariates in their study that were similar to ours, in which they captured seasonal and yearly trends in a 28-year time series, improving their soil texture models and course fragment content. Our trend covariates in contrast captured the trend over several years to quantify land use change rather than seasonal and yearly effects on dynamic covariates from seasonal/yearly differences in rainfall and moisture. All these DSM studies discussed the benefit from the use of their version of multi-temporal covariates. However, unlike our project,

the targeted variables for prediction were static soil properties. Our project also uniquely applies multi-temporal covariates in an urbanizing, temperate forest environment. The use of multi-temporal covariates for DSP prediction, like for this project, is a more recent goal in the field of DSM and requires more research on how they can be applied.

Despite the multi-temporal dynamic covariates not being selected, the information captured by dynamic ECs is important for the models, as evidenced by the mono-temporal dynamic covariates being most of the ECs that were selected in the best models. In the RF model, four out of five covariates selected were mono-temporal dynamic ECs. The remaining covariate in the MLR model was a static EC. For the MLR, five out of the eight covariates selected were mono-temporal dynamic ECs. Of the remaining three, two were static and one was a soil covariate. This suggests that there is some relationship that exists between dynamic soil properties and dynamic environmental covariates. Static covariates and soil covariates were less important. Still, they helped explain some of the variability in infiltration, indicating that infiltration has some relationship to terrain and soil factors but is more strongly impacted by the characteristics of dynamic covariates. At least one static EC was selected in both the MLR and RF models. Terrain steepness factors, like slope gradient and terrain ruggedness, appear to influence the variability of infiltration as they were used in the RF and MLR models, respectively. Slope gradient as a measure of terrain steepness was the only static covariate used for the RF. The influence of slope gradient on infiltration has been observed by Assouline & Ben-Hur (2006) and Ribolzi et al. (2011), who found that steeper slopes have higher infiltration rates than gentler slopes. Overall, the covariates that best explained variability in infiltration for our study contrast with the findings from Pahlavan-Rad et al. (2020), in which static covariates were more important to their infiltration predictions in Iran. Like our project, they mapped

infiltration using RF and MLR models, and in each model, their most important covariates were related to sand content and factors related to stream networks. Their models only used two dynamic covariates: the normalized difference salinity index (NDSI) and NDVI. The types of important covariates varied between their project and ours, likely due to the different study regions. Pahlavan-Rad et al. (2020) modeled infiltration in an arid, low-relief, sparsely vegetated floodplain where the dominant land cover is agricultural land, abandoned land, saline land, and sand dunes. In our local environment, variability in infiltration is expected to be controlled by different factors. For instance, saline soils are not experienced in West Virginia because our high precipitation levels would wash out salt in WV soils and not allow it to accumulate. As such, NDSI would not be a covariate considered here for influencing infiltration. And since the WRW has more rugged terrain, different terrain covariates will have more influence, like the effect of slope gradient. Disregarding the environmental differences, our results for prediction of infiltration are less adequate than those of Pahlavan-Rad et al. (2020) in comparison. They reported an RMSE of 13.9 mm hr^{-1} and an MAE of 10.9 mm hr^{-1} for their MLR model, which is much lower than those in the current study (Table 8).

For both the MLR and RF model, their reliance on primarily dynamic covariates reinforces that infiltration has an important relationship to land use in the region of our study. Each dynamic EC captures information related to land use patterns in the watershed. Both the MLR and the RF model used NDBI from 2019 to 2017, distinguishing the more highly urbanized watershed areas from the less urbanized ones. Wetness from 2015 was also selected for both models, which captures moisture levels in the soil surface and vegetation canopy. The remaining dynamic covariate for the MLR was NDVI from 2019, which shows the more vegetated areas from the less vegetated areas. The RF used Greenness from 2016 to measure vegetated versus

less vegetation areas and did not use any NDVIs. The RF model also used Brightness from 2013, which captures ground reflectance and tends to mirror the urbanized areas with higher reflectance values due to pavement. As hypothesized, dynamic covariates are important for predicting dynamic soil properties, like infiltration, because they can change quickly over time in accordance with the different vegetation and management characteristics associated with different land uses.

Multi-temporal covariates were not selected for either of the models, suggesting that they were not important in explaining the variability in DSPs like infiltration. The multi-temporal covariates that were offered for selection were calculated based on six years of data spanning from 2013 to 2020 in the watershed. Only six satellite images were used due to the limited availability of images over the watershed that were (1) from Landsat 8, (2) during the growing season, and (3) cloud-free. Appalachia frequently experiences clouds and precipitation, so there are rarely entirely cloud-free days. Because of the small size of the WRW, even a few small clouds and their shadows could cover a relatively large portion of the watershed and render the image unusable. The possibility of using imagery from Sentinel-2 was considered, but after exploring the data it was not found to be any more obstruction-free than Landsat 8. As such, roughly each growing season had one useable image, and years without useable images were skipped. It is likely that having only six years of observations to calculate the multi-temporal layer is insufficient to produce values that would help explain variation in infiltration across the watershed. Land use proportions in the WRW changed from 2013 to 2020, that time frame is not long enough to capture the longer-term impact of land use conversion from decades ago that still has lingering effects. This aligns with results from Zhao et al. (2013), who studied a 79-year chronosequence of a cropland restored to grassland in China and found that infiltration rate

rapidly increased until 16 years after the restoration began and then only slightly increased after. Similarly, Zema et al. (2021) found that infiltration increased along with forest stand age, with the lowest infiltrations occurring in the most recent forests. Sampling throughout the WRW, it becomes evident that not all forests are equal, as Zema et al. (2021) suggest. These subtleties are difficult to pick up from a mono-temporal satellite image, which is best at identifying if a forest is present or not in a certain area based on the characteristics canopy coverage. As such, much of the heterogeneity between forests is lost. For instance, one forest site we sampled at first glance may look like any other forest in the watershed, but upon further investigation, we observed that most of trees appeared younger (around 30 years old) and even-aged. The soil there was extremely compacted for a forest soil and had a relatively slow infiltration rate. According to the landowner, the area we sampled had previously been developed, so the soil properties in that area were still in the process of changing from its land use history. In another forested area in the validation sample, cows that grazed in the nearby field were also allowed to go into the forest, leaving those soil trampled and compacted. Yesilonis et al. (2016) observed the lingering effects of land use history on forests that transitioned from agriculture 50-70 years ago, 12-150 years ago, and historically undisturbed forests. They found that forests that grew on abandoned agricultural land exhibited different soil surface properties than the undisturbed forests: these younger forests had higher bulk densities, lower carbon content, evidence of erosion of the surface horizon, and no organic horizon. The impact of historic land use on soil properties has a lasting effect and in many cases is irreversible, as it sets the soil on a different trajectory of formation than its undisturbed counterparts (Yesilonis et al., 2016). Mono-temporal satellite images are not able to read the story that we are able to observe on the ground, which is why multi-temporal covariates are important due to their potential to quantify the effects of historic

management and use of an area. Additionally, even though by definition DSPs change over relatively short periods of time (in contrast to static soil properties that change over hundreds or thousands of years), seven years is still a short time to observe change in soil properties. Change in DSPs may be more wholly represented through multi-temporal dynamic ECs that span over tens of years, at a minimum. In most cases, it takes several years to observe significant changes in soil properties after changes in vegetation cover and land management.

The goal is for the models produced for DSPs, and the overall DSS, to be generalizable (or transferable) both spatially and temporally. Spatially generalizable would mean that the ECs found to be useful for prediction of infiltration in the WRW would similar be effective for predicting infiltration in other watershed under similar environmental conditions like those throughout other West Virginian watersheds and across Appalachia. Temporally generalizable models are one of the benefits of DSS over the conventional soil survey. This would mean the models could be applied to the variation of conditions and changes the watershed will experience decades in the future. The same types of ECs would be recalculated and updated, and still produce a sufficient prediction of infiltration in the future. It is not known whether the models produced from this project are spatially or temporally transferrable as they currently exist.

CHAPTER 6. CONCLUSIONS & SUGGESTIONS FOR FUTURE RESEARCH

The goals of this study were to represent DSP using DSM methods using static, mono-temporal dynamic, and multi-temporal dynamic covariates. Across the three land uses, forests had the highest infiltration rate. The median infiltration rates of the pasture and urban land uses were similarly low, but the urban land had the largest amount of variability. MLR and RF models were produced, with R^2 values of 0.302 and 0.201, respectively. Results demonstrate that some relationship can be harnessed from satellite imagery and terrain data as representatives are environmental conditions in a watershed and the spatial variability of infiltration that can be used to develop predictive maps. However, with the data collected and ECs used in the models, this relationship is not well represented within the scope of this study. This project provided insight and methods that will be useful in the subsequent larger goal of developing a dynamic soil survey. Namely, creating and employing multi-temporal dynamic covariates that capture trends in ECs across several years offers potential to be useful factors to include for mapping DSPs in future DSM projects. Additionally, although the models were not strong, the covariates selected for the best models confirm that dynamic ECs are important for predicting DSPs.

The results of this study lead to several questions that future DSM/DSS projects should investigate. The expansion of multi-temporal dynamic covariates over several decades may prove useful information for building DSP predictive maps, including infiltration in future applications. Quality, longevity, and continuity of satellite imagery pose a technical challenge to creating these multi-temporal ECs, but the collection of satellite imagery available to the public increases and improves over time, alleviating that restriction. The possibility of using images outside of the growing season could also be explored to quantify seasonal trends. Seasonal differences among the dynamic covariates could also be explored, like using the average,

maximum, and minimum values across each year. For example, using the difference between maximum NDVI, which is likely during the growing season over the summer, and the minimum NDVI, which is likely over the winter, could potentially identify differences in vegetated areas that may look the same when only analyzing across growing seasons. Additionally, the mono-temporal covariates could be further evaluated to find the best scale and method to create them that can pick up variability in infiltration. Along with that, other predictors should be explored for the predicting infiltration and other DSPs and highly vegetated and urbanizing areas, like those found in much of the eastern United States. Additionally, kriging of the residuals of the models could be explored and put back into the models to try to represent some of the unexplained variability in infiltration. More projects addressing the prediction of other DSPs in regions like this will work together to complete a full DSS for the WRW and other watersheds. This work towards a finer resolution, raster soil survey that includes DSPs is extremely important and necessary for optimizing land management for sustainability of natural resources, soil health, and overall watershed health, which would mean encouraging management practices that minimize soil compaction and erosion while encouraging less disturbing methods that would restore natural vegetative cover, lower bulk density, and increase organic matter content. Through identifying current watershed-scale hydrologic patterns, areas that require attention can be addressed with targeted management plans.

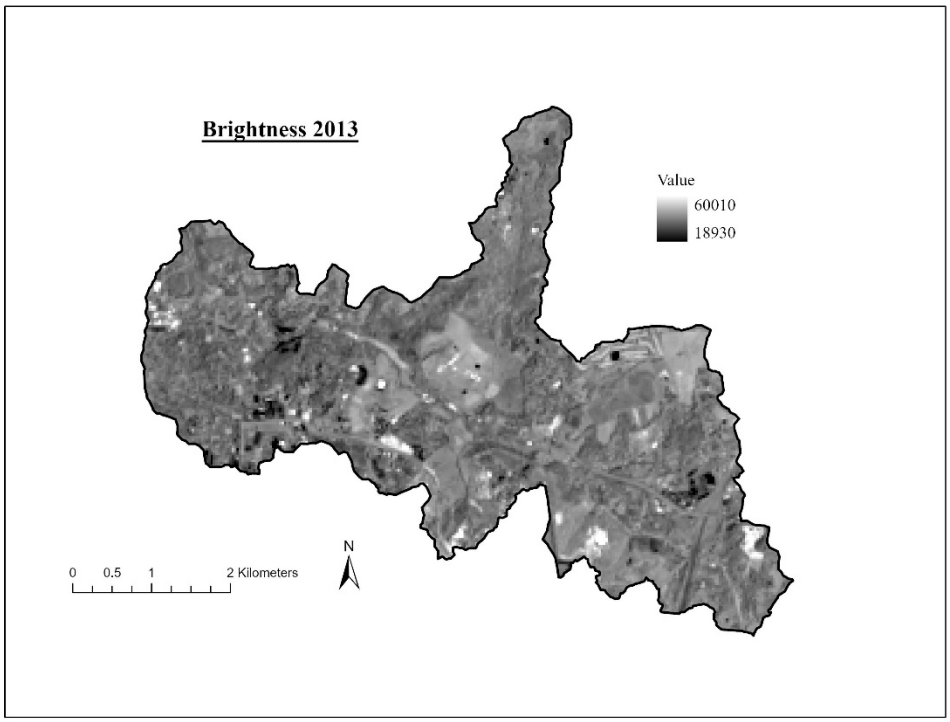
APPENDIX

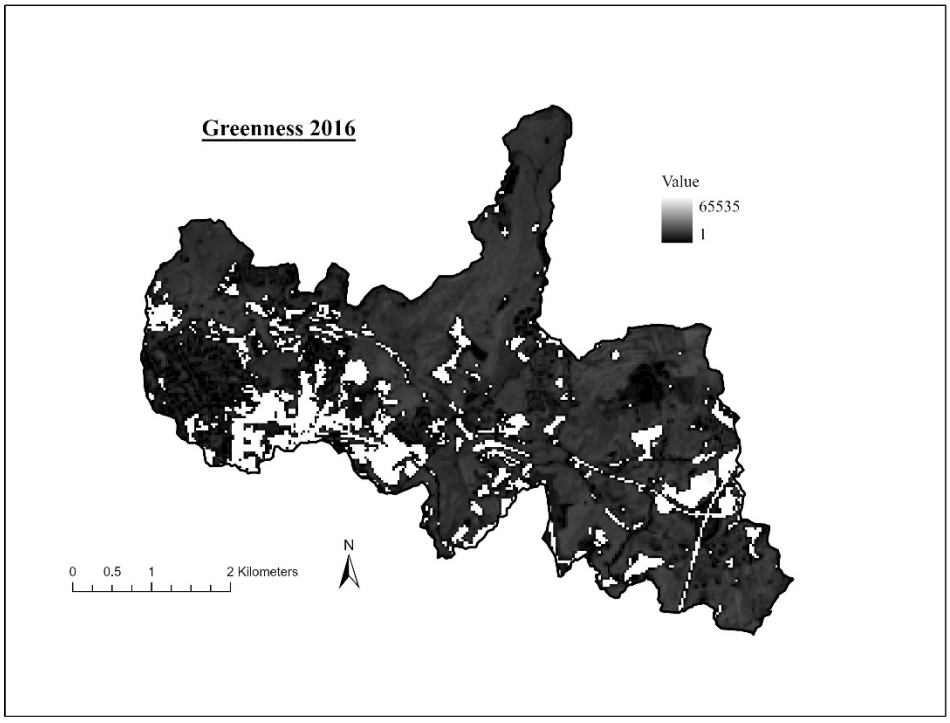
Summary of soil properties predicted in previous DSM projects.

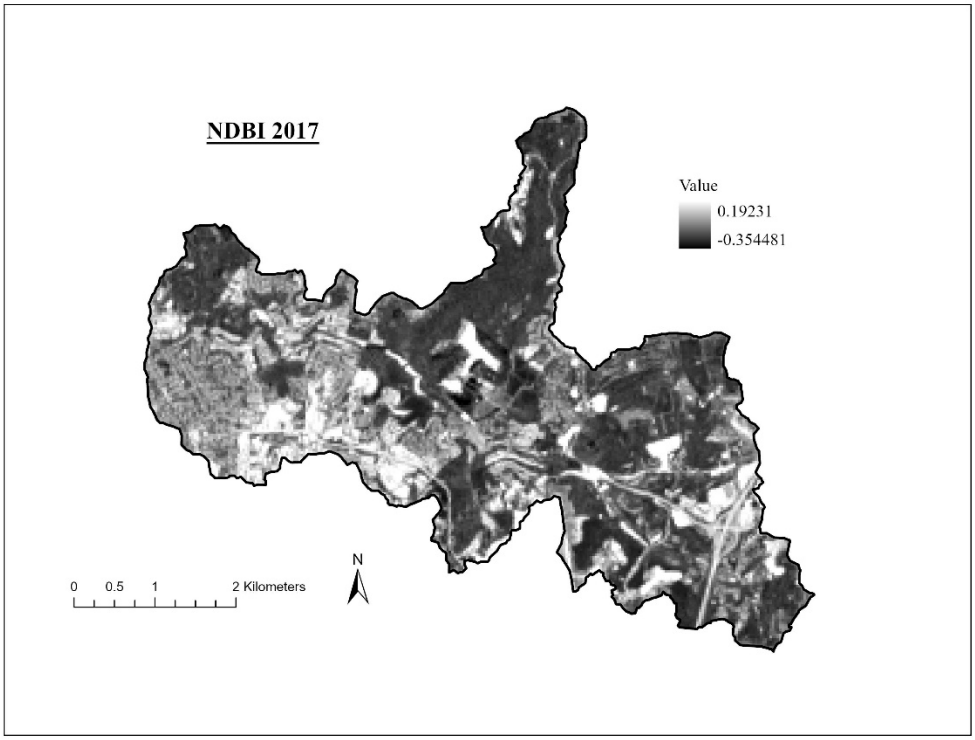
Soil property	Type of property	Modeling Method	Reference
Organic carbon content	Dynamic	Cubist, RF	Fathololoumi et al., 2020
		RF	Grimm et al., 2008
		RF, cubist, artificial neural network. Multiple linear regression, support vector machines	John et al., 2020
Taxonomic classes	Static	RF, bagged classfication tree, classification tree, K nearest neighbor, linear discriminant analysis, linear support vector machines, multinomial logostic regression, multilayer-perceptron neural network, nearest shrunken centroids, radial-basis support vector machines, single-hidden-layer neural networks	Brungard et al., 2015
Infiltration	Dynamic	Multiple linear regression, RF	Pahlavan-Rad et al., 2020
Particle size contents	Static	Partial least squares regression	Diek et al., 2016
		Cubist, RF	Fathololoumi et al., 2020
		Partial least squares regression	Gallo et al., 2018
		Support vector machines	Maynard & Levi, 2017
		Linear regression	Zeng et al., 2020
Moisture content	Dynamic	Partial least squares	Diek et al., 2016
		Cubist, RF	Fathololoumi et al., 2021

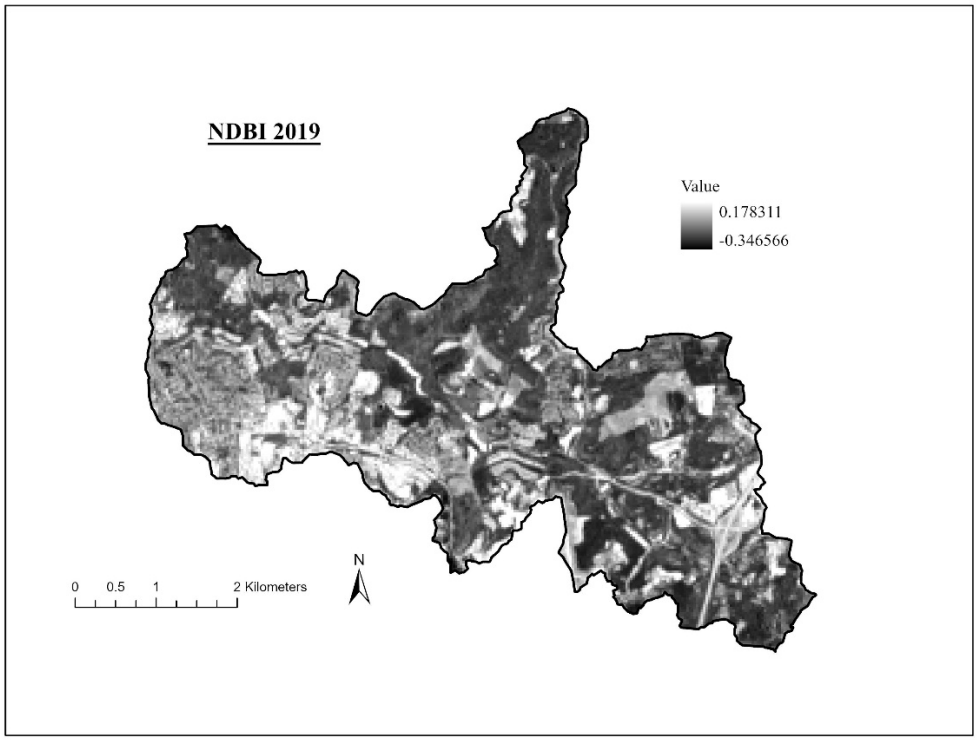
Soil property	Type of property	Modeling Method	Reference
Organic matter content	Dynamic	Partial least squares regression	Diek et al., 2016
		Partial least squares regression	Gallo et al., 2018
Calcium carbonate equivalent	Static	Cubist, RF	Fathololoumi et al., 2020
Cation exchange capacity	Dynamic	Partial least squares regression	Gallo et al., 2018

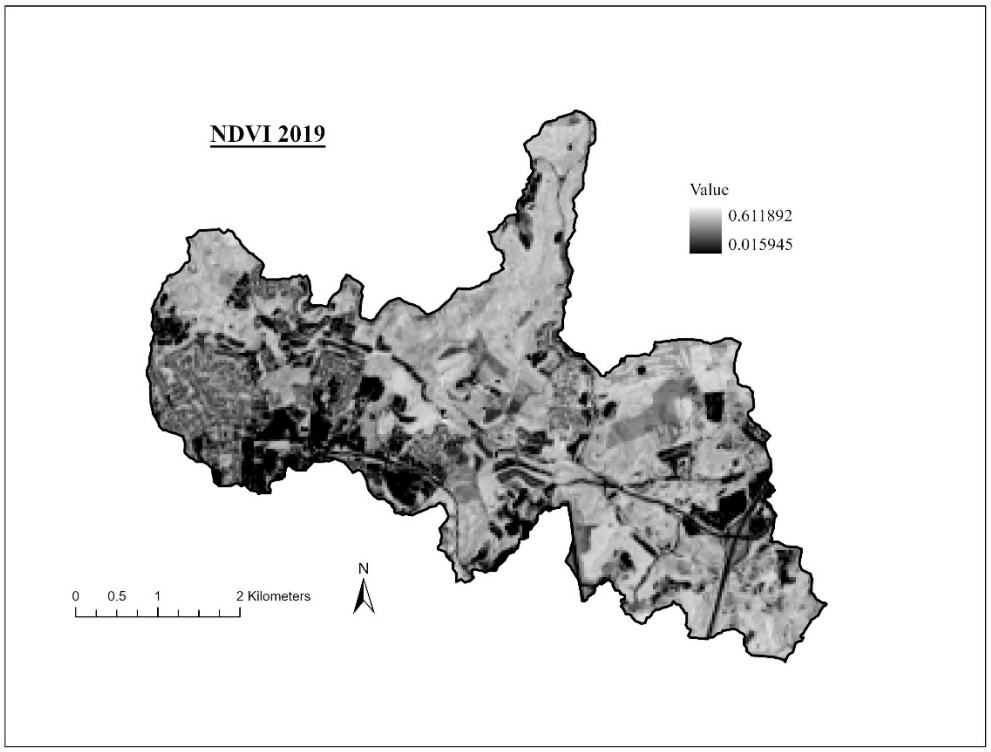
ECs used in MLR and RF models.

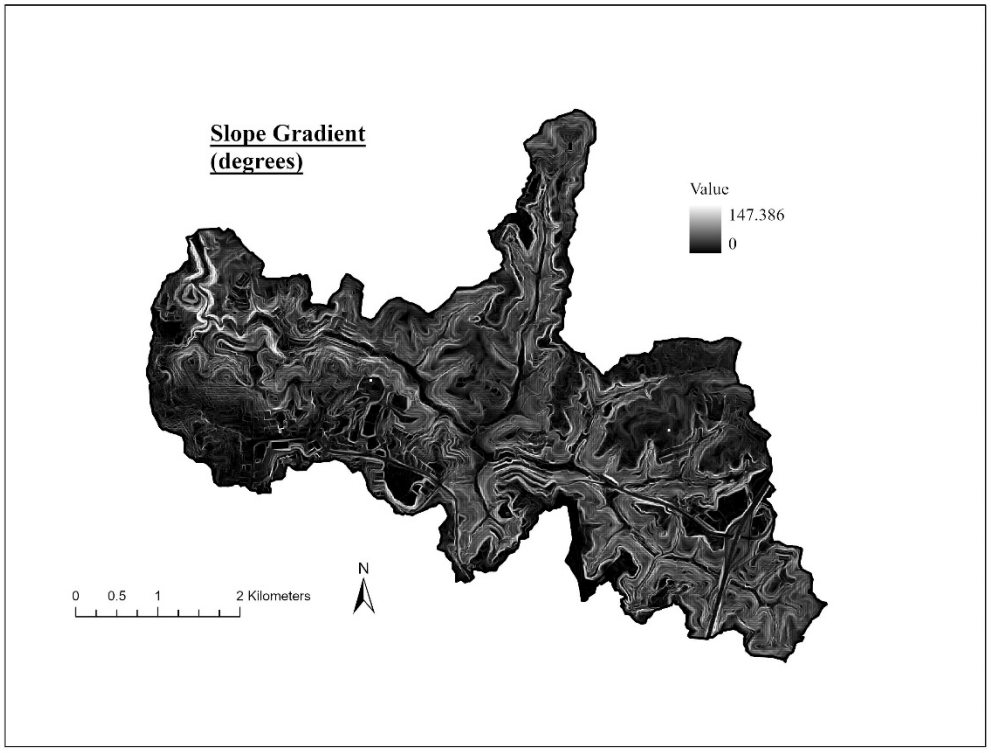


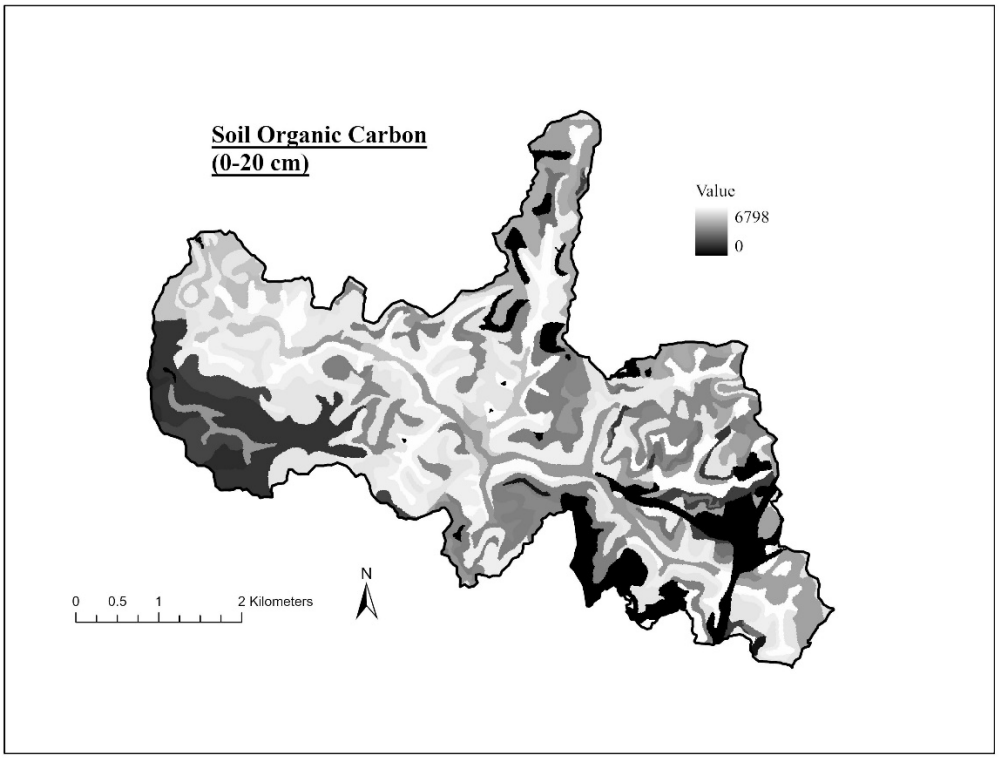


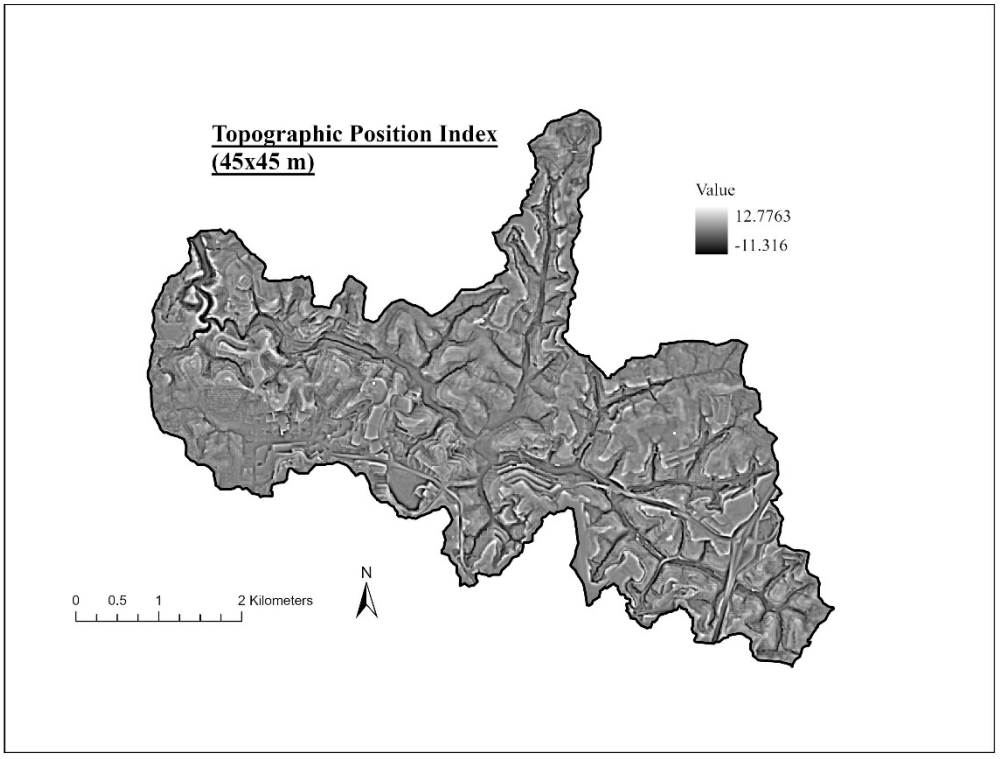

NDBI 2017
Value
0.19231
-0.354481
0 0.5 1 2 Kilometers
N

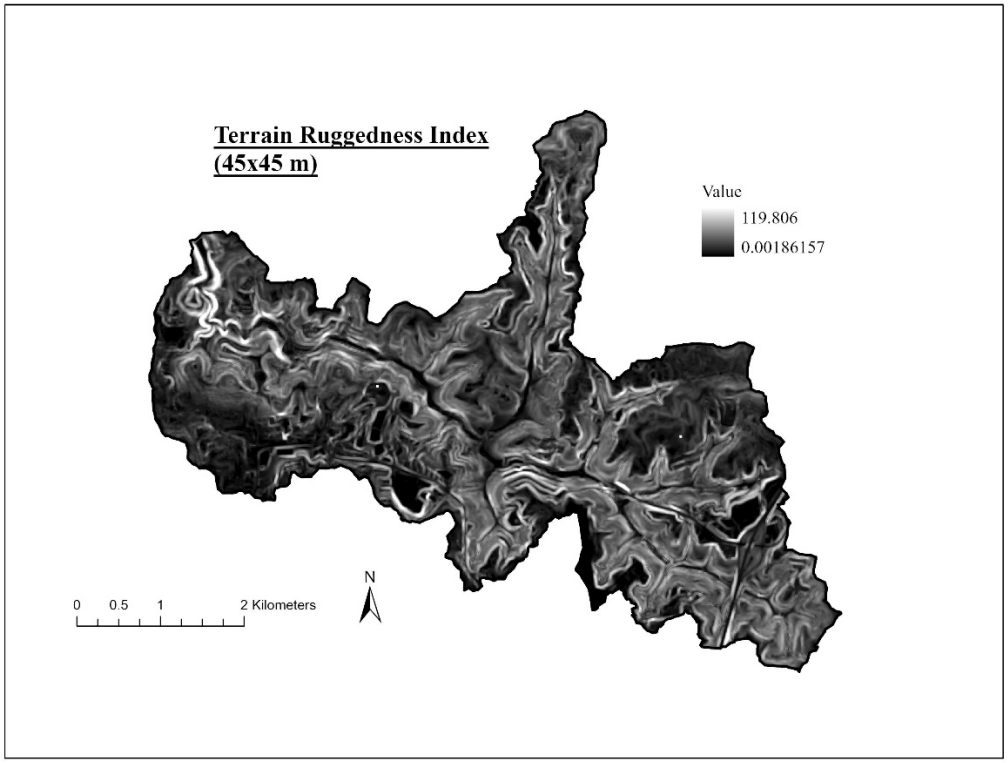

NDBI 2019
Value
0.178311
-0.346566
0 0.5 1 2 Kilometers
N

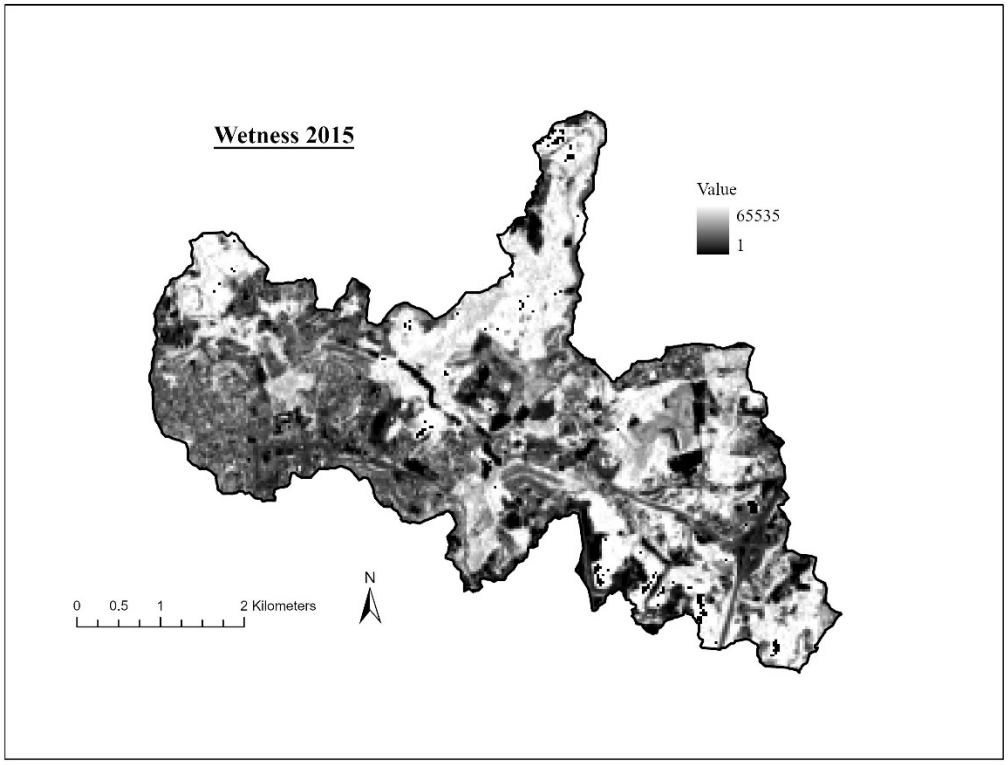

NDVI 2019
Value
0.611892
0.015945
0 0.5 1 2 Kilometers
N


Slope Gradient
(degrees)
Value
147.386
0
0 0.5 1 2 Kilometers
N


Soil Organic Carbon
(0-20 cm)
Value
6798
0
N
0 0.5 1 2 Kilometers


Topographic Position Index
(45x45 m)
Value
12.7763
-11.316
N
0 0.5 1 2 Kilometers


Terrain Ruggedness Index
(45x45 m)
Value
119.806
0.00186157
0 0.5 1 2 Kilometers
N


Wetness 2015
Value
65535
1
0 0.5 1 2 Kilometers
N

LITERATURE CITED

Ahmed, F., Nestingen, R., Nieber, J. L., Gulliver, J. S., & Hozalski, R. M. (2014). A modified Philip–Dunne infiltrometer for measuring the field-saturated hydraulic conductivity of surface soil. *Vadose Zone Journal*, *13*(10).

Ali, Q.S., Pandey, S.K., Chaudhuri, R., Behera, S., & Jeyakumar, L.G. (2021). Development of rainfall-infiltration measurement system and recharge strategies for urban flooding areas: a case study of Delhi, India. Modeling Earth Systems and Environment, 7, 2719-2731.

Amoozegar, A. (1989), A Compact Constant-Head Permeameter for Measuring Saturated Hydraulic Conductivity of the Vadose Zone. Soil Science Society of America Journal, 53: 1356-1361. https://doi.org/10.2136/sssaj1989.03615995005300050009x

Anderson, R. L., Brye, K. R., & Wood, L. S. (2020). Landuse and soil property effects on infiltration into Alfisols in the Lower Mississippi River Valley, USA. Geoderma Regional, 22, e00297. doi:10.1016/j.geodrs.2020.e00297

Anyamba, A., & Tucker, C. J. (2005). Analysis of Sahelian vegetation dynamics using NOAA-AVHRR NDVI data from 1981–2003. Journal of arid environments, 63(3), 596-614.

Arya, L. M., Leij, F. J., Shouse, P. J., & van Genuchten, M. T. (1999). Relationship between the hydraulic conductivity function and the particle-size distribution. Soil Science Society of America Journal, 63(5), 1063-1070.

Assouline, S., & Ben-Hur, M. (2006). Effects of rainfall intensity and slope gradient on the dynamics of interrill erosion during soil surface sealing. Catena, 66(3), 211-220.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific data, 5, 180214. https://doi.org/10.1038/sdata.2018.214

Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrological sciences journal, 24(1), 43-69.

Blake, G. R., & Hartge, K. H. (1986). Bulk density. Methods of soil analysis: Part 1 Physical and mineralogical methods, 5, 363-375.

Blasch, K. W., Ferre, T. P., Hoffmann, J. P., & Fleming, J. B. (2006). Relative contributions of transient and steady state infiltration during ephemeral streamflow. Water Resources Research, 42(8).

Bodí, M. B., Martin, D. A., Balfour, V. N., Santín, C., Doerr, S. H., Pereira, P., ... & Mataix-Solera, J. (2014). Wildland fire ash: production, composition and eco-hydro-geomorphic effects. Earth-Science Reviews, 130, 103-127.

Boettinger, J. L., Ramsey, R. D., Bodily, J. M., Cole, N. J., Kienast-Brown, S., Nield, S. J., ... & Stum, A. K. (2008). Landsat spectral data for digital soil mapping. In Digital soil mapping with limited data (pp. 193-202). Springer, Dordrecht.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Brevik, E. C., & Hartemink, A. E. (2010). Early soil knowledge and the birth and development of soil science. Catena, 83(1), 23-33.

Brungard, C. W., Boettinger, J. L., Duniway, M. C., Wills, S. A., & Edwards Jr, T. C. (2015). Machine learning for predicting soil classes in three semi-arid landscapes. Geoderma, 239, 68-83.

Carlson, T. N., & Ripley, D. A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote sensing of Environment, 62(3), 241-252.

Chang, C. W., Laird, D. A., Mausbach, M. J., & Hurburgh, C. R. (2001). Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. *Soil Science Society of America Journal*, *65*(2), 480-490.

Coulter, C.B., R.K. Kolka, and J.A. Thompson. 2004. Water quality in rural, urban and mixed use watersheds. Journal of the American Water Resources Association. 40:1593-1601.

Crave, A., & Gascuel-Odoux, C. (1997). The influence of topography on time and space distribution of soil surface water content. Hydrological processes, 11(2), 203-210.

Crist, E. P. (1985). A TM tasseled cap equivalent transformation for reflectance factor data. Remote sensing of Environment, 17(3), 301-306.

Croke, J., Hairsine, P., & Fogarty, P. (2001). Soil recovery from track construction and harvesting changes in surface infiltration, erosion and delivery rates with time. Forest Ecology and Management, 143(1-3), 3-12.

Desta, F., Colbert, J. J., Rentch, J. S., & Gottschalk, K. W. (2004). Aspect induced differences in vegetation, soil, and microclimatic characteristics of an Appalachian watershed. Castanea, 69(2), 92-108.

Diek, S., Schaepman, M. E., & De Jong, R. (2016). Creating multi-temporal composites of airborne imaging spectroscopy data in support of digital soil mapping. Remote Sensing, 8(11), 906.

DiTomaso, J. M., Brooks, M. L., Allen, E. B., Minnich, R., Rice, P. M., & Kyser, G. B. (2006). Control of invasive weeds with prescribed burning. Weed technology, 20(2), 535-548.

Durner, W., & Iden, S. C. (2021). The improved integral suspension pressure method (ISP+) for precise particle size analysis of soil and sedimentary materials. Soil and Tillage Research, 213, 105086.

Ellison, W. D. (1948), Soil detachment by water in erosion processes, Eos Trans. AGU, 29(4), 499– 502, doi:10.1029/TR029i004p00499.

Erickson, A. J., Weiss, P. T., & Gulliver, J. S. (2013). Optimizing stormwater treatment practices. *A Handbook of Assessment and Maintenance*, *1*(1), 1-337.

Evans, I. S., & Cox, N. J. (1999). Relations between land surface properties: altitude, slope and curvature. Process modelling and landform evolution, 13-45.

Fathololoumi, S., Vaezi, A. R., Alavipanah, S. K., Ghorbani, A., Saurette, D., & Biswas, A. (2020). Improved digital soil mapping with multitemporal remotely sensed satellite data fusion: A case study in Iran. Science of The Total Environment, 721, 137703.

Fathololoumi, S., Vaezi, A. R., Alavipanah, S. K., Ghorbani, A., Saurette, D., & Biswas, A. (2021). Effect of multi-temporal satellite images on soil moisture prediction using a digital soil mapping approach. Geoderma, 385, 114901.

Fernandes, P. M., & Botelho, H. S. (2003). A review of prescribed burning effectiveness in fire hazard reduction. International Journal of wildland fire, 12(2), 117-128.

Finke, P. A. (2006). Quality assessment of digital soil maps: producers and users perspectives. Developments in Soil Science, 31, 523-631.

Gallo, B. C., Demattê, J. A., Rizzo, R., Safanelli, J. L., Mendes, W. D. S., Lepsch, I. F., ... & Lacerda, M. P. (2018). Multi-temporal satellite images on topsoil attribute quantification and the relationship with soil classes and geology. Remote Sensing, 10(10), 1571.

Gareth, J., Daniela, W., Trevor, H., & Robert, T. (2013). *An introduction to statistical learning: with applications in R*. Spinger.

Ghorbani-Dashtaki, S., Homaee, M., & Loiskandl, W. (2016). Towards using pedotransfer functions for estimating infiltration parameters. *Hydrological Sciences Journal*, *61*(8), 1477-1488.

Gray, D. M., Landine, P. G., & Granger, R. J. (1985). Simulating infiltration into frozen prairie soils in streamflow models. Canadian Journal of Earth Sciences, 22(3), 464-472.

Gregorutti, B., Michel, B., & Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and Computing*, *27*, 659-678.

Gregory, J. H., Dukes, M. D., Miller, G. L., & Jones, P. H. (2005). Analysis of double-ring infiltration techniques and development of a simple automatic water delivery system. *Applied turfgrass science*, *2*(1), 1-7.

Gregory, J.H. & Dukes, Michael & Jones, Pierce & Miller, Grady. (2006). Effect of Urban Soil Compaction on Infiltration Rate. Journal of Soil and Water Conservation. 61. 117-124.

Griffith, D. M., & Widmann, R. H. (2000). Forest statistics for West Virginia: 1989 and 2000 - fs.fed.us. Retrieved February 20, 2022, from https://www.fs.fed.us/ne/newtown_square/publications/resource_bulletins/pdfs/2003/ne_rb157.pdf

Grimm, R., Behrens, T., Märker, M., & Elsenbeer, H. (2008). Soil organic carbon concentrations and stocks on Barro Colorado Island—Digital soil mapping using Random Forests analysis. Geoderma, 146(1-2), 102-113.

Harden, C. P., & Scruggs, P. D. (2003). Infiltration on mountain slopes: a comparison of three environments. Geomorphology, 55(1-4), 5-24.

Harvey, C. (2022, August 22). Three Reasons Appalachia's Risk of Deadly Floods Keeps Rising. *Scientific American*.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). The elements of statistical learning. Springer series in statistics. *New York, NY, USA*.

He, C. (2003). Integration of geographic information systems and simulation model for watershed management. Environ. Model. Softw., 18, 809-813.

Hewlett, J. D., & Hibbert, A. R. (1967). Factors affecting the response of small watersheds to precipitation in humid areas. Forest hydrology, 1, 275-290.

Hill, R. L. (1990). Long-term conventional and no-tillage effects on selected soil physical properties. Soil Science Society of America Journal, 54(1), 161-166.

Horton, R. E. (1933). The role of infiltration in the hydrologic cycle. Eos, Transactions American Geophysical Union, 14(1), 446-460.

Horton, R. E. (1941). An approach toward a physical interpretation of infiltration-capacity 1. Soil science society of America journal, 5(C), 399-417.

Hossain Anni, A., Cohen, S., & Praskievicz, S. (2020). Sensitivity of urban flood simulations to stormwater infrastructure and soil infiltration. Journal of Hydrology, 588, 125028. https://doi.org/10.1016/j.jhydrol.2020.125028

Hubbart, J. A., Kellner, E., & Petersen, F. (2022). A 22-Site Comparison of Land-Use Practices, E-coli and Enterococci Concentrations. *International Journal of Environmental Research and Public Health*, *19*(21), 13907.

Itsukushima, R., Ideta, K., & Takata, H. (2021). Relationship between compaction and infiltration capacity of amended soil for urban flood damage mitigation. Soil Use and Management. doi:10.1111/sum.12705

Jabro, J. D. (1992). Estimation of saturated hydraulic conductivity of soils from particle size distribution and bulk density data. Transactions of the ASAE, 35(2), 557-560.

Janeau, J. L., Bricquet, J. P., Planchon, O., & Valentin, C. (2003). Soil crusting and infiltration on steep slopes in northern Thailand. European Journal of Soil Science, 54(3), 543-554.

Jenny, H., 1941. Factors of Soil Formation, A System of Quantitative Pedology. McGraw-Hill, New York.

John, K., Abraham Isong, I., Michael Kebonye, N., Okon Ayito, E., Chapman Agyeman, P., & Marcus Afu, S. (2020). Using machine learning algorithms to estimate soil organic carbon variability with environmental variables and soil nutrient indicators in an alluvial soil. Land, 9(12), 487.

Johnson, D. L., & Schaetzl, R. J. (2015). Differing views of soil and pedogenesis by two masters: Darwin and Dokuchaev. Geoderma, 237, 176-189.

Johnson, D. L., Keller, E. A., & Rockwell, T. K. (1990). Dynamic pedogenesis: new views on some key soil concepts, and a model for interpreting Quaternary soils. Quaternary Research, 33(3), 306-319.

Karlen, D. L., Veum, K. S., Sudduth, K. A., Obrycki, J. F., & Nunes, M. R. (2019). Soil health assessment: Past accomplishments, current activities, and future opportunities. Soil and Tillage Research, 195, 104365.

Lal, R., & Shukla, M. K. (2004). Principles of soil physics. CRC Press.

Landsat Missions | U.S. Geological Survey. (n.d.). Retrieved March 18, 2022, from https://www.usgs.gov/landsat-missions

Lee, K. E., & Foster, R. C. (1991). Soil fauna and soil structure. *Soil Research*, *29*(6), 745-775.

Lillesand, T., Kiefer, R. W., & Chipman, J. (2015). Remote sensing and image interpretation. John Wiley & Sons.

Lin, H. (2003). Hydropedology: Bridging disciplines, scales, and data. Vadose Zone Journal, 2(1), 1-11.

Lotsch, A., Friedl, M. A., Anderson, B. T., & Tucker, C. J. (2003). Coupled vegetation-precipitation variability observed from satellite and climate records. Geophysical Research Letters, 30(14).

Liu, F., Geng, X., Zhu, A. X., Fraser, W., & Waddell, A. (2012). Soil texture mapping over low relief areas using land surface feedback dynamic patterns extracted from MODIS. Geoderma, 171, 44-52.

Ma, Y. J., Li, X. Y., Guo, L., & Lin, H. (2017). Hydropedology: Interactions between pedologic and hydrologic processes across spatiotemporal scales. Earth-Science Reviews, 171, 181-195.

Malone, B. P., Minasny, B., & McBratney, A. B. (2017). *Using R for digital soil mapping* (Vol. 35). Cham, Switzerland: Springer International Publishing.

Maynard, J. J., & Levi, M. R. (2017). Hyper-temporal remote sensing for digital soil mapping: Characterizing soil-vegetation response to climatic variability. Geoderma, 285, 94-109.

McBratney, A. B., Santos, M. M., & Minasny, B. (2003). On digital soil mapping. Geoderma, 117(1-2), 3-52.

McIntyre, D. S. (1958). Permeability measurements of soil crusts formed by raindrop impact. Soil Science, 85(4): 185-189.

Mengel, D. B., & Barber, S. A. (1974). Development and Distribution of the Corn Root System Under Field Conditions 1. Agronomy Journal, 66(3), 341-344.

Moore, I.D. (1980). Effect of Surface Sealing on Infiltration. Transactions of the ASABE, 24, 1546-1552.

Moore, I. D., Grayson, R. B., & Ladson, A. R. (1991). Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. Hydrological processes, 5(1), 3-30.

Mulder, V. L., Lacoste, M., Richer-De-Forges, A. C., & Arrouays, D. (2016). GlobalSoilMap France: High-resolution spatial modelling the soils of France up to two meter depth. *Science of the Total Environment*, *573*, 1352-1369.

NASA. (n.d.). Landsat science. NASA. Retrieved March 18, 2022, from https://landsat.gsfc.nasa.gov/

Natural Resources Conservation Service. (n.d.). Components of Dynamic Soil Survey. Components of Dynamic Soil Survey | NRCS Soils. Retrieved January 2, 2022, from https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/focusteams/?cid=nrcseprd1710628

Natural Resources Conservation Service. (n.d.). Natural Resources Conservation Service. Dynamic Soil Properties (DSPs) | NRCS Soils. Retrieved February 20, 2022, from https://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/soils/use/?cid=STELPRDB1256067

Neris, J., Tejedor, M., Rodríguez, M., Fuentes, J., & Jiménez, C. (2013). Effect of forest floor characteristics on water repellency, infiltration, runoff and soil loss in Andisols of Tenerife (Canary Islands, Spain). Catena, 108, 50-57.

Noborio, K. (2001). Measurement of soil water content and electrical conductivity by time domain reflectometry: a review. *Computers and electronics in agriculture*, *31*(3), 213-237.

Northeast Regional Climate Center. (n.d.). NRCC summary tables. Retrieved February 17, 2022, from http://www.nrcc.cornell.edu/regional/tables/tables.html

Nussbaum, M., Spiess, K., Baltensweiler, A., Grob, U., Keller, A., Greiner, L., ... & Papritz, A. (2018). Evaluation of digital soil mapping approaches with large sets of environmental covariates. Soil, 4(1), 1-22.

Pahlavan-Rad, M. R., Dahmardeh, K., Hadizadeh, M., Keykha, G., Mohammadnia, N., Gangali, M., ... & Brungard, C. (2020). Prediction of soil water infiltration using multiple linear regression and random forest in a dry flood plain, eastern Iran. Catena, 194, 104715.

Petersen, F., & Hubbart, J. A. (2020). Advancing Understanding of Land Use and Physicochemical Impacts on Fecal Contamination in Mixed-Land-Use Watersheds. Water, 12(4), 1094. doi:10.3390/w12041094

Philip, J. R. (1993). Approximate analysis of falling-head lined borehole permeameter. *Water Resources Research*, *29*(11), 3763-3768.

Pitt, R., Lantrip, J., & O'Connor, T. P. (2000). Infiltration through disturbed urban soils. In Building Partnerships (pp. 1-10).

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reddy, N. N., Chakraborty, P., Roy, S., Singh, K., Minasny, B., McBratney, A. B., ... & Das, B. S. (2021). Legacy data-based national-scale digital mapping of key soil properties in India. *Geoderma*, *381*, 114684.

Ren, X., Hong, N., Li, L., Kang, J., & Li, J. (2020). Effect of infiltration rate changes in urban soils on stormwater runoff process. Geoderma, 363, 114158. https://doi.org/10.1016/j.geoderma.2019.114158

Reynolds, W. D., & Elrick, D. E. (1991). Determination of hydraulic conductivity using a tension infiltrometer. *Soil Science Society of America Journal*, *55*(3), 633-639.

Reynolds, W. D. (1993). Saturated hydraulic conductivity: field measurement. *Soil sampling and methods of analysis*.

Ribolzi, O., Patin, J., Bresson, L. M., Latsachack, K. O., Mouche, E., Sengtaheuanghoung, O., ... & Valentin, C. (2011). Impact of slope gradient on soil surface features and infiltration on steep slopes in northern Laos. Geomorphology, 127(1-2), 53-63.

Rivera, L., Cobos, D., & Campbell, G. (2016, April). Automated Dual-Head Infiltrometer for Measuring Field Saturated Hydraulic Conductivity (Kfs). In *EGU General Assembly Conference Abstracts* (pp. EPSC2016-1057).

Robichaud, P. R. (2000). Fire effects on infiltration rates after prescribed fire in Northern Rocky Mountain forests, USA. Journal of Hydrology, 231, 220-229.

Roo, A.P., & Riezebos, H.T. (1992). Infiltration experiments on loess soils and their implications for modelling surface runoff and soil erosion. Catena, 19, 221-239.

Runkle, J., K.E. Kunkel, R. Frankson, B.C. Stewart, and J. Spaccio (2022). West Virginia State Climate Summary 2022. NOAA Technical Report NESDIS 150-WV. NOAA/NESDIS, Silver Spring, MD, 4 pp.

Saxton, K. E., Rawls, W., Romberger, J. S., & Papendick, R. I. (1986). Estimating generalized soil-water characteristics from texture. Soil science society of America Journal, 50(4), 1031-1036.

Schaik, N.L. (2009). Spatial variability of infiltration patterns related to site characteristics in a semi-arid watershed. Catena, 78, 36-47.

Shanafield, M., & Cook, P. G. (2014). Transmission losses, infiltration and groundwater recharge through ephemeral and intermittent streambeds: A review of applied methods. Journal of Hydrology, 511, 518-529.

Shepherd, K. D., & Walsh, M. G. (2002). Development of reflectance spectral libraries for characterization of soil properties. *Soil science society of America journal*, *66*(3), 988-998.

Shukla, M.K., Lal, R., Owens, L.B., & Unkefer, P.J. (2003). Land use and management impacts on structure and infiltration characteristics of soils in the north Appalachian region of Ohio. Soil Science, 168, 167-177.

Shuster, W. D., Dadio, S. D., Burkman, C. E., Earl, S. R., & Hall, S. J. (2015). Hydropedological assessments of parcel-level infiltration in an arid urban ecosystem. Soil Science Society of America Journal, 79(2), 398-406.

Sihag, P., Tiwari, N. K., & Ranjan, S. (2018). Prediction of cumulative infiltration of sandy soil using random forest approach. Journal of Applied Water Engineering and Research.

Six, J., Bossuyt, H., Degryze, S., & Denef, K. (2004). A history of research on the link between (micro) aggregates, soil biota, and soil organic matter dynamics. Soil and tillage research, 79(1), 7-31.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online at the following link: http://websoilsurvey.sc.egov.usda.gov/. Accessed [02/21/2022].

Soil Survey Staff, The Gridded Soil Survey Geographic (SSURGO) Database for West Virginia. United States Department of Agriculture, Natural Resources Conservation Service. Available online at https://gdg.sc.egov.usda.gov/. January 2023 (202007 official release).

Stuart, G. W., & Edwards, P. J. (2006). Concepts about forests and water. Northern Journal of Applied Forestry, 23(1), 11-19.

Sun, D., Yang, H., Guan, D., Yang, M., Wu, J., Yuan, F., Jin, C., Wang, A., & Zhang, Y. (2018). The effects of land use change on soil infiltration capacity in China: A meta-analysis. The Science of the total environment, 626, 1394–1401. https://doi.org/10.1016/j.scitotenv.2018.01.104

Suriya, S.A., & Mudgal, B. (2012). Impact of urbanization on flooding: The Thirusoolam sub watershed – A case study. Journal of Hydrology, 412, 210-219.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267-288.

Thompson, J.A., S.M. Roecker, S. Grunwald, and P.R. Owens. 2012. Digital Soil Mapping: Interactions with and Applications for Hydropedology. In Lin, H. (Ed.). Hydropedology: Synergistic Integration of Soil Science and Hydrology. Academic Press. p. 665-709.Tugel, A. J., Herrick, J. E., Brown, J. R., Mausbach, M. J., Puckett, W., & Hipple, K. (2005). Soil change, soil survey, and natural resources decision making: a blueprint for action.

Vieux, B. E. (2001). Distributed hydrologic modeling using GIS. In Distributed hydrologic modeling using GIS (pp. 1-17). Springer, Dordrecht.

Wang, S. Q., Li, W. D., Li, J., & Liu, X. S. (2013). Prediction of soil texture using FT-NIR spectroscopy and PXRF spectrometry with data fusion. Soil Science, 178(11), 626-638.

Wang, J., & Chen, L. (2021). The effect of hillslope geometry on Hortonian rainfall-infiltration-runoff processes. Journal of Hydrology, 594, 125962.

West Virginia Water Research Institute and West Run Watershed Association. (2008, July 31). West Run - WV Department of Environmental Protection. Retrieved February 19, 2022, from https://dep.wv.gov/WWE/Programs/nonptsource/WBP/Documents/WP/WestRun_WBP.pdf

West Virginia Geological and Economic Survey. (1913). [County report and maps.]: Marion, Monongalia and Taylor counties.

Wills, S., Williams, C., & Seybold, C. (2017). Assessing dynamic soil properties and soil change. *Soil Science Division Staff, editor, Soil Survey Manual. United States Department of Agriculture*, 481-503.

Wright, E. L. (1982). *Soil survey of Marion and Monongalia Counties, West Virginia*. U.S. Dept. of Agriculture, Natural Resources Conservation Service.

Wu, L., Pan, L., Roberson, M. J., & Shouse, P. J. (1997). Numerical evaluation of ring-infiltrometers under various soil conditions. *Soil Science*, *162*(11), 771-777.

Xie C, Cai S, Yu B, Yan L, Liang A, Che S, The effects of tree root density on water infiltration in urban soil based on a Ground Penetrating Radar in Shanghai, China, Urban Forestry and amp; Urban Greening (2020), doi: https://doi.org/10.1016/j.ufug.2020.126648

Yang, JL., Zhang, GL. Water infiltration in urban soils and its effects on the quantity and quality of runoff. J Soils Sediments **11,** 751–761 (2011). https://doi.org/10.1007/s11368-011-0356-1

Yesilonis, I., Szlavecz, K., Pouyat, R., Whigham, D., & Xia, L. (2016). Historical land use and stand age effects on forest soil properties in the Mid-Atlantic US. *Forest Ecology and Management*, *370*, 83-92.

Zema, D. A., Plaza-Alvarez, P. A., Xu, X., Carra, B. G., & Lucas-Borja, M. E. (2021). Influence of forest stand age on soil water repellency and hydraulic conductivity in the Mediterranean environment. *Science of the Total Environment*, *753*, 142006.

Zeng, C., Yang, L., & Zhu, A. (2020). The generation of soil spectral dynamic feedback using landsat 8 data for digital soil mapping. Remote Sensing, 12(10), 1691.

Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. International journal of remote sensing, 24(3), 583-594.

Zhao, Y., Wu, P., Zhao, S., & Feng, H. (2013). Variation of soil infiltrability across a 79-year chronosequence of naturally restored grassland on the Loess Plateau, China. *Journal of Hydrology*, *504*, 94-103.

Zhang, G. L., Feng, L. I. U., & Sonbg, X. D. (2017). Recent progress and future prospect of digital soil mapping: A review. Journal of Integrative Agriculture, 16(12), 2871-2885.

Zhang, S. Y., Hopkins, I., Guo, L., & Lin, H. (2019). Dynamics of infiltration rate and field-saturated soil hydraulic conductivity in a wastewater-irrigated cropland. Water, 11(8), 1632.

Zhu, A. X., Liu, F., Li, B., Pei, T., Qin, C., Liu, G., ... & Zhou, C. (2010). Differentiation of soil conditions over low relief areas using feedback dynamic patterns. Soil Science Society of America Journal, 74(3), 861-869.

Zimale, FA, Tilahun, SA, Tebebu, TY, et al. Improving watershed management practices in humid regions. Hydrological Processes. 2017; 31: 3294–3301. https://doi.org/10.1002/hyp.11241