USING ELECTRICAL RESISTIVITY FOR PREDICTING SOIL TEXTURE AS AN INPUT TO MODEL WATER AND CHLORIDE TRANSPORT IN A VARIABLY SATURATED HILLSLOPE IN THE PIEDMONT OF GEORGIA, USA

by

MARIA TERESA TANCREDI

(Under the Direction of Nandita Gaur)

ABSTRACT

Currently available soil datasets, presenting low resolution and shallow depth, are inadequate for activities that warrant a detailed characterization of the sub-surface. I investigated an alternative method to determine spatially distributed soil texture and developed a high-resolution soil texture map by importing soil electrical resistivity, relative depth of investigation, and weekly antecedent rainfall in an Artificial Neural Network (ANN) framework. I used the predicted textures to investigate the effects of varying domain complexity in simulations of hydrological and chloride transport at the hillslope scale using a Richard's equation based single porosity model. I compared the outcomes with ones developed using the traditional pedological soil horizon-based division of the sub-surface. In both models, the hillslope received a constant sub-surface flux input of solute and fed into a lake. The ANN accurately predicted soils, but the information did not have a major effect on the single porosity, hydrological simulations.

INDEX WORDS: Soil texture, Electrical resistivity, Piedmont, Water flux, Solute movement, Hillslope, HYDRUS (2D/3D).

USING ELECTRICAL RESISTIVITY FOR PREDICTING SOIL TEXTURE AS AN INPUT TO MODEL WATER AND CHLORIDE TRANSPORT IN A VARIABLY SATURATED HILLSLOPE IN THE PIEDMONT OF GEORGIA, USA

by

MARIA TERESA TANCREDI

BS, Università degli Studi di Padova, Italy, 2017

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

This thesis was part of a dual MS program between the University of Georgia and the Università degli Studi di Padova, and it was submitted to the Università degli Studi di Padova in partial fulfillment of the requirements for the degree of Laurea Magistrale (Master's Degree) in Sustainable Agriculture, Classe LM-69, D.M. 16/03/2007.

MASTER OF SCIENCE

ATHENS, GEORGIA

2020

USING ELECTRICAL RESISTIVITY FOR PREDICTING SOIL TEXTURE AS AN INPUT TO MODEL WATER AND CHLORIDE TRANSPORT IN A VARIABLY SATURATED HILLSLOPE IN THE PIEDMONT OF GEORGIA, USA

by

MARIA TERESA TANCREDI

Major Professor:	Nandita Gaur
Committee:	Daniel Markewitz
	Gary L. Hawkins
	Francesco Morari

Electronic Version Approved:

Ron Walcott
Dean of the Graduate School
The University of Georgia
December 2020

ACKNOWLEDGEMENTS

I would like to thank Dr. Nandita Gaur for the trust she put in me since the beginning of this project and for having motivated me to explore every available opportunity in the last two years. I would also like to thank Dr. David E Radcliffe for his support along the way and the always stimulating discussions. I would like to thank Dr. Francesco Morari for giving me the chance and strongly encouraging me to participate in the Dual Degree between the University of Georgia and the Università degli Studi di Padova. I would like to thank Dr. Matthew Levi, Matthew Thibodeaux, Charlotte Adams, Samuele Ceolin, Jill Mullican, and all the students that worked with me in the Environmental Soil Physics lab. I could have never accomplished this research without your help. I would like to thank the members of my Advisory Committee for listening to me and guiding me along the research. I would like to thank Dr. Brigette Haram for her supervision and her constant availability during the project. I would like to acknowledge the Gwinnett County Department of Water Resources and the Georgia Water Resources Institute at Georgia Tech for sponsoring this research. Last but not least, I would like to thank my family and my friends for believing in me and supporting me constantly.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Overview

Septic systems, also called Onsite Wastewater Treatment Systems (OWTS), are a decentralized method for wastewater treatment. These systems are usually adopted in locations where population expansion has moved past the reach of the sewer lines connected to the local wastewater treatment plant. According to Ade O. Oke (2019) the State of Georgia has 600,000 septic tank systems already in use, and over 30,000 new ones are installed every year. Each year, among the total number of systems installed in Georgia, about 10,000 undergo repairs due to failure (Ade O. Oke, 2019).

A septic system failure is defined as the occurrence of one or more of the following situations: 1) slow drains or sewage reentering the house, 2) ponding or emergence of the effluents on the soil surface, and/or 3) improper treatment of the wastewater at the drainfield level (Ade O. Oke, 2019; Ceolin, 2019). System failures can generate economic losses, and threaten public health and the environment. The concentration of pathogens in drinking and recreational waters can increase as a consequence of contamination from failing OWTS. Diseases can also be spread through food crops if those are irrigated with water infected by septic effluents. High levels of nitrogen (as nitrate) in drinking water, which can also be originated from septic effluents, can be harmful especially to children. Furthermore, these failures can cause soil and water quality degradation (Gunady, Shishkina et al., 2015). Excess nutrient in water can cause harmful algae blooms as part of a process called eutrophication. The most noticeable effects of eutrophication occur at the end of the algal growing season when great amounts of dissolved oxygen are consumed by decomposer microorganisms digesting the algae. The drastic reduction in oxygen content in the water body, also called hypoxia, can harm its wildlife and seriously impact the water quality, particularly in terms of water taste and odor. Cyanobacteria blooms can release into the water several toxins as well as taste and odor compounds, i.e. *2-methylisoborneol* (MIB) and *geosmin*, that make water

unpleasant to drink and smell (Graham, Loftin et al., 2010). To prevent and manage those problems, it is crucial to monitor the state of the OWTS, the diffusion of pathogens, and the nutrient supply to water bodies, in particular in terms of nitrogen, N, and phosphorus, P (Lewis, Wurtsbaugh et al., 2011; Schindler, 1971).

Clean Water Act and Total Maximum Daily Load

In 1948, the Senate and House of Representatives of the United States of America enacted the "Federal Water Pollution Control Act" to "restore and maintain the chemical, physical, and biological integrity of the Nation's waters" ("Federal Water Pollution Control Amendments of 1972," 1972). The Act was edited and expanded in 1972 and its name was changed to the "Clean Water Act" (CWA). The U.S. Environmental Protection Agency (EPA) is appointed as the administrator of the act, in collaboration with federal, state, and tribal regulatory partners. The CWA requires the definition of specific water quality standards for different water bodies according to their "designated uses". Those standards are used to assess if the water bodies are: 1) supporting designated use, 2) not supporting designated use, or 3) in a state of assessment pending. If a water body falls in the second category and does not already have a Total Maximum Daily Load (TMDL), one must be created for it. TMDLs are "the sum of the individual waste load allocations (WLAs) for point sources and load allocations (Las) for nonpoint sources, as well as natural background for a given waterbody. The TMDL must also include a margin of safety (MOS), either implicitly or explicitly, that accounts for the uncertainty in the relationship between pollutant loads and the water quality response of the receiving water body." (The Georgia Department of Natural Resources Environmental Protection Division Atlanta, 2017). In 2017, the State of Georgia, through the Georgia Environmental Protection Division (GA EPD) of the Georgia Department of Natural Resources (GA DNR), promulgated the final TMDL for Lake Sidney Lanier, after one segment of the lake was listed as "not supporting designated use" and a second one as having an "assessment pending" for Chlorophyll *a*. Chlorophyll *a* is an algae pigment and it is usually utilized as a reference for measuring potential excess nutrient loads to water bodies. The TMDL for Lake Sidney Lanier lists several possible sources of

nutrients to the water body. There are both point sources, including municipal and industrial sources, and non-point sources such as urban runoff, agricultural runoff, and septic system failures. The urban development in the area surrounding the lake, for example, can be a good indicator of the importance of these sources.

In the Lake Sidney Lanier watershed, within the Chattahoochee River Basin, the growing population is generating an expansion of the areas adopting OWTS. A summary of the estimated increase in the number of OWTS for the counties in the watershed can be seen in Table 1. The data refer to the period from 2006 to 2011 and also show the approximate number of septic systems that have undergone repair (The Georgia Department of Natural Resources Environmental Protection Division Atlanta, 2017).

Table 1.1. The estimated increase in OWTS in the counties in the Lake Sidney Lanier Watershed from 2006 to 2011 and the approximate number of septic systems that have been repaired. The values in the table are sourced from The Georgia Department of Natural Resources Environmental Protection Division Atlanta (2017).

County	***Existing OWTS (2006)***	***Existing OWTS (2011)***	***Increase in OWTS (2007 – 2011)***	***Repaired OWTS (2007-2011)***
Dawson	*8954*	*9372*	*418*	*172*
Forsyth	*31946*	*32907*	*961*	*1173*
Gwinnett	*64702*	*65192*	*490*	*1550*
Habersham	*14507*	*15259*	*752*	*245*
Hall	*47108*	*48489*	*1381*	*1377*
Lumpkin	*11462*	*12314*	*852*	*71*
Towns	*8538*	*9179*	*641*	*43*
Union	*13390*	*14198*	*808*	*182*
White	*10717*	*11276*	*559*	*217*
TOTAL	*211324*	*218186*	*6862*	*5030*

According to the final TMDL for Lake Lanier, total phosphorus loading should not exceed 0.09 mg/L (0.25 pounds per acre-foot of lake volume) per year, while total nitrogen should not exceed 4 mg/L (10.88 pounds per acre-foot) in the photic zone. The photic zone is the section of a water body where light can pass through and allow for algae and phytoplankton to perform photosynthesis (Sheppard, 1982). These requirements can be met only if there is knowledge regarding the movement of nutrients from point and non-point sources to the lake. In this research, I focused on water and solute movement at the hillslope scale. Hillslopes can be identified as the area comprised between the top of a ridge and the closest stream in a valley (Fan, Clark et al., 2019). This information serves as a base to produce an estimation of the total load of the above nutrients from non-point sources, specifically OWTS. In the case of solutes movement in the soil, it is important to keep in mind the reactions and fate of the molecules with the media surrounding them.

Organic nitrogen and ammonium (NH_4^+) are the most common form of nitrogen found in domestic wastewater. The anaerobic environment of the septic tank and the presence of heterotrophic bacteria convert urea and other types of organic nitrogen into ammonium, through a process called mineralization. When the effluent moves into the drainfield, the changes in the general conditions (presence of oxygen and different kinds of microbes) prompt further transformations of the nitrogen. The bacterial genus *Nitrosomonas* and *Nitrobacter* exploit oxygen and inorganic carbon or CO_2 to oxidize NH_4^+ to nitrite (NO_2^-) and NO_2^- to nitrate (NO_3^-), respectively. Ammonium can also be adsorbed by soil particles or released through ammonia volatilization.

Organic phosphorus and polyphosphate are the main forms of phosphorous found in wastewater. In the anaerobic conditions of the tank biological and biogeochemical processes turn most of the P into orthophosphates. Organic P reaching the drainfield is also converted into orthophosphates. Some of the P remains in the solid phase in the tank. Adsorption and mineral precipitation in the soil are the two main processes for P removal. Additionally, both N and P are in part removed from the subsurface by vegetation, as they are plant macronutrients. A more comprehensive review of contaminants' fate and transport in OWTS is provided by Lusk, Toor et al. (2017).

Ceolin (2019) conducted a *Preliminary assessment of onsite wastewater treatment system effects on Lake Lanier water quality.* The study was performed using monitoring wells to estimate the contribution of nitrogen and phosphorus from domestic OWTS along the shoreline. A 'box model' of one of the study sites was also developed with HYDRUS (2D/3D) to assess the relevance of different factors affecting the load to the lake. The 'box model' consisted of four layers (topsoil, subsoil, saprolite, and bedrock) and replicated all the characteristics of the site, except for the topography. The layers were represented in a slope, but their thickness was constant, and their boundaries were flat. The model was based on Richard's Equation using single porosity. The work of Ceolin (2019) showed that the two investigated structures (both based on a soil horizon division of the sub-surface) underestimated chloride concentrations and simulated inaccurate temporal trends. These results suggested the possible need for a more detailed definition of the soil domain that could also capture preferential flow. Additionally, the models developed by Ceolin (2019) were not validated with an independent dataset, thus requiring further analysis of the accuracy of the calibrated parameters. My research expanded that work with an evaluation of the possible insights deriving from the use of near-surface geophysical methods paired with traditional sampling techniques to model water and chloride movement. I selected chloride because its low reactivity and low plant uptake allowed us to use it as a tracer for water coming from the OWTS (Alhajjar, Chesters et al., 1990).

Background

The Piedmont Physiographic Region stretches for 1600km, from New Jersey and southeastern Pennsylvania to eastern Alabama (Figure 1.1). The bedrock is usually covered by a mantel called regolith composed of a layer of saprolite with residual soil on top (Miller, 1990). The bedrock-mantel structure and the topography have an impact on the hydrologic system of the area and can set the boundaries for water tables. Recharge areas are usually in the upland slopes, while rivers and creeks are mostly situated in flood plains. The water table tends to be deeper where the mantel is thicker and most often lies in the

saprolite, at the top of the bedrock. Its shape can also be pictured as a reversed profile of the area, being deeper under the uplands and shallower in the plains (LeGrand, Back et al., 1988).

Figure 1.1. The highlighted area represents the U.S. Piedmont geologic region. It extends for 1600km, from New Jersey and southeastern Pennsylvania to eastern Alabama.

In Georgia, the eastern Blue Ridge and the Piedmont region present a structure consisting of thrust sheets and melange. The thickness of soil usually decreases with increasing altitude and steepness of the area, presenting deeper argillic Bt horizons moving south toward the Fall Zone. The Fall Zone (or Fall Line) comprises the area where the eastern Piedmont meets the Coastal Plain (Roberts, 2001). In Georgia, the Fall Zone is located approximately between the 33rd and 34th meridians (Markewich, Pavich et al., 1990). The typical soils in the outer Piedmont developed from saprolite and are often classified as Hapludults. The bulk density and clay content of these soils are signals of the intense chemical weathering in the formation process. The clay fractions are mainly kaolinite and hydroxyl-interlayered vermiculite (HIV). This composition is common in soils developing on saprolite that formed from silicate minerals in temperate climates (Zauyah, Schaefer et al., 2018). The presence of the argillic horizon inhibits rapid rainfall infiltration, making the macropores the main path for water influx (Markewich,

Pavich et al., 1990). The bedrock in the area surrounding Lake Sidney Lanier is constituted by Lower Paleozoic and Precambrian felsic gneiss, schist, granite Gneiss, and Granite (Miller, 1990).

Lake Sidney Lanier (Figure 1.2) is a man-made reservoir and one of the major lakes in the Blue Ridge and the Piedmont areas in Georgia. The lake is situated at about 48 km northeast of Atlanta and it was created in 1956 through the construction of the Buford Dam by the US Army Corps of Engineers. The deepest point of the lake is in the original Chattahoochee River channel on the north side of Buford Dam (Figure 1.2). The location is at 277.67 m above sea level and a depth of 48.77 m when the lake is at its winter pool full level (326.44 m MSL). The highest elevation for the lake is 328.33 m MSL and was recorded in April of 1964, while the lowest elevation is 323 m MSL and was registered in December of 2007 (US Army Corps of Engineers, 2019).

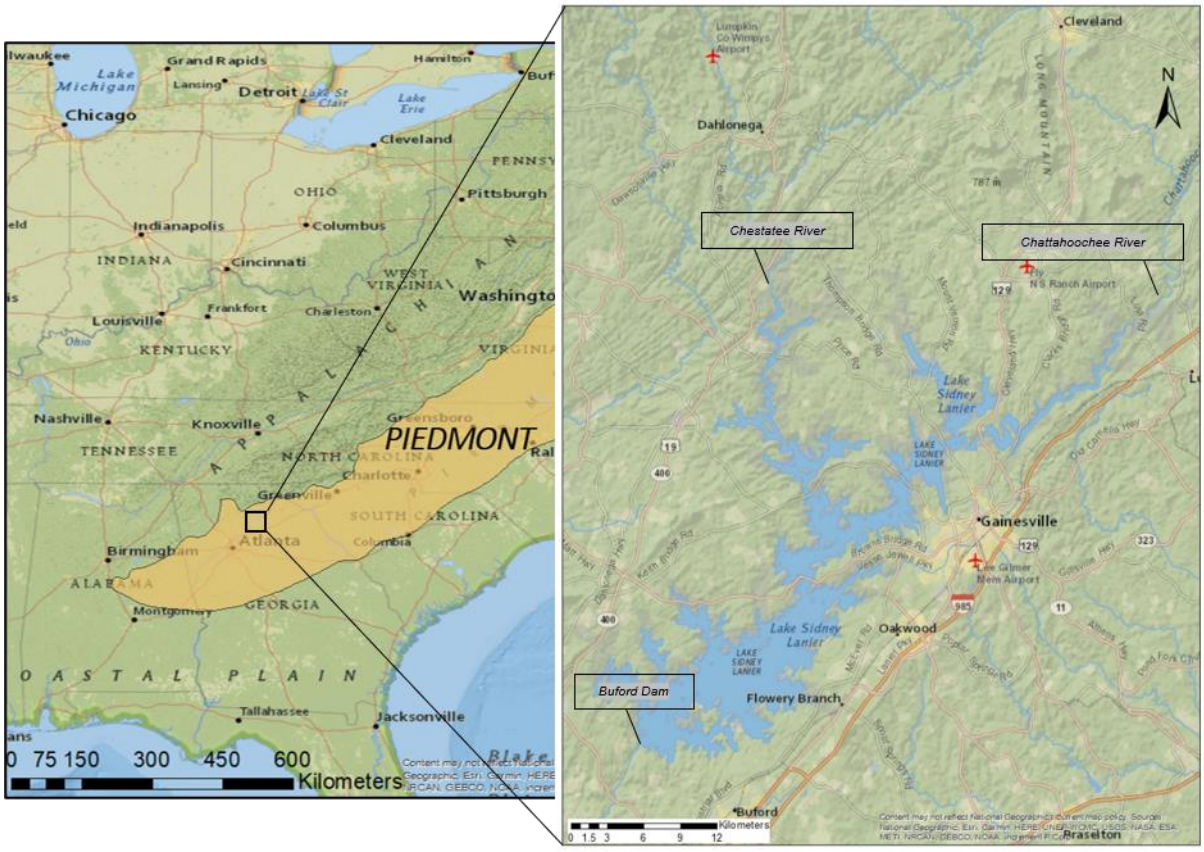


Figure 1.2. The highlighted area on the left image represents the U.S. Piedmont geologic Region. The picture on the right shows Lake Sidney Lanier, a man-made reservoir situated at about 48 km northeast of Atlanta. Most of the inflow to the lake comes from the Chestatee and the Chattahoochee Rivers.

The officially designated uses for the lake are Drinking water supplies and Recreation, but the lake also provides benefits for flood control, hydropower generation, fish and wildlife management, and navigation. The majority of the inflow to the lake comes from the Chestatee and the Chattahoochee Rivers. The reservoir is particularly important for the region since it provides water supply to the cities of Atlanta, Buford, Cumming, Gainesville, and the entire Forsyth and Gwinnett Counties. The area covered by this research falls within the borders of Gwinnett, Forsyth, and Hall Counties, on the southern side of Lake Sidney Lanier in the Piedmont Region of Georgia (Figure 1.3).

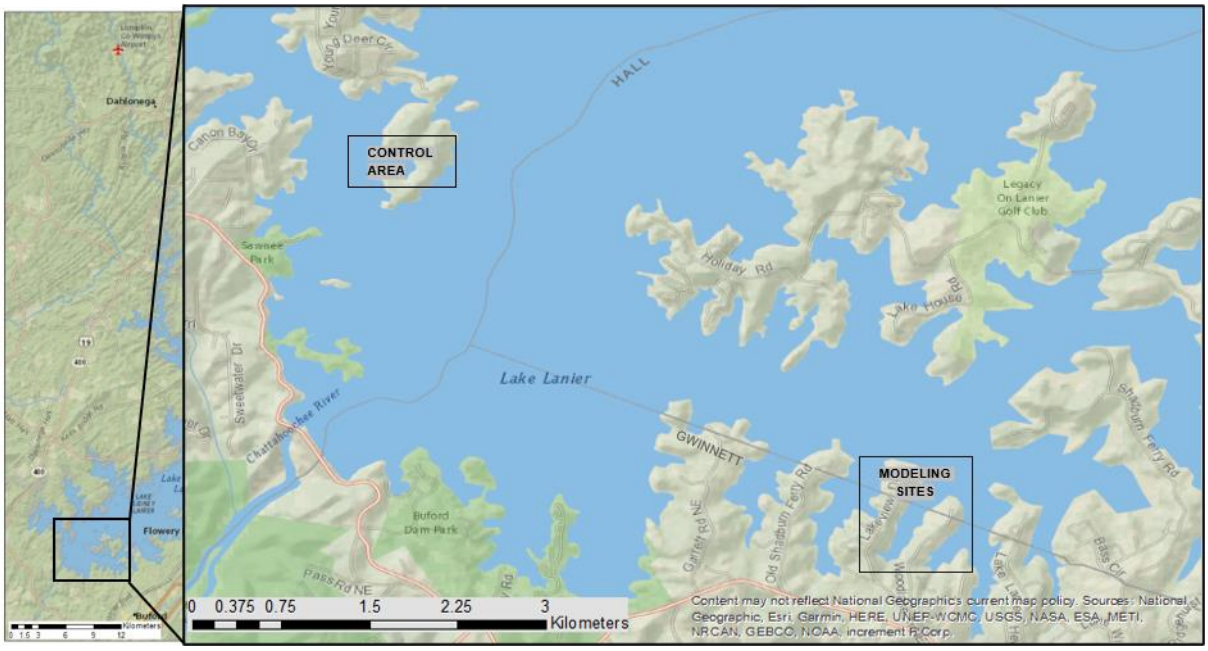


Figure 1.3. The southern side of Lake Sidney Lanier enclosing the study area for this research. The three study sites are situated in an area enclosed by Gwinnett, Forsyth, and Hall Counties.

As previously stated, the water quality of waterbodies where there is a high density of OWTS is a relevant topic for many densely populated areas. Currently, most of the evaluations are carried on through the use of monitoring wells for water sample collection (Ceolin, 2019; Jantrania, 2015; McDowell, Brick et al., 2005; Morrissey, Johnston et al., 2015). These wells usually extend from the surface to the water-table to intercept the nutrient plume developing from the OWTS.

In the peculiar conditions of the Piedmont region, sub-surface flowpaths could be especially varied from idealized systems. The correct flowpaths have to be assessed to monitor and determine OWTS discharge of nutrients, therefore predicting transport processes solely through conventional methods of water sampling could be challenging (Englert, Kemna et al., 2016).

In the Piedmont region, the capacity of a well to yield water is strictly connected to its position relative to the bedrock fractures. Wells that do not penetrate a fracture will carry low to no water and those that connect to small and shallow fractures could soon see the water content decrease because of fracture drainage. Additionally, wells situated on slopes usually contain less water (Miller, 1990).

The difficulties associated with the heterogeneity of the subsurface, with mapping the fractures, and with estimating water flowpaths can make the exploitation of wells unsuitable for monitoring purposes. In these conditions, alternative methods could prove more adequate and efficient. This research studied the potential of Electrical Resistivity Tomography (ERT), in determining parameters helpful to assess hydraulic properties at the hillslope scale in the Piedmont Region of Georgia.

Geophysical methods, in particular electrical resistivity methods, have been widely employed since the 1970s to monitor reservoirs and to detect groundwater pollution (Reynolds, 2011b). Hinnell, Ferré et al. (2010) report that there is a growing consensus on the benefit that subsurface imaging obtained through geophysics provides to hydrologic models. Zeng, Meng et al. (2016) stress the ability of ERT to provide information about water content, near-surface structures, and hydro-geological interfaces and processes.

Furthermore, for several models, including the one developed by Ceolin (2019), the sub-surface is usually simplified and doesn't consider the actual heterogeneities of the system. Heterogeneity in weathering of the different rocks, and spatial variation of soil types, and soil layers can alter water movement. These simplifications compared to the real complexity of the subsurface can degrade model predictions of critical nutrient flux attributes.

Project scopes

The objective of this project was to assess whether, at the hillslope scale, data obtained from non-invasive, non-conventional methods, specifically Electrical Resistivity Tomography (ERT), can generate information on hydraulic properties of the subsurface and ultimately characterize preferential flow paths for water and nutrients contained in it. I investigated two main questions to reach this goal:

1. Can soil resistivity values, processed through an Artificial Neural Network framework, characterize soil texture and subsurface heterogeneity?
2. How does domain complexity for soil layers affect the accuracy of models predicting nutrient movement in a variably saturated hillslope?

References

Ade O. Oke, R. (2019). *Manual for On-Site Sewage Management Systems*. Georgia Department of Public Health

Alhajjar, B. J., Chesters, G., & Harkin, J. M. (1990). Indicators of chemical pollution from septic systems. *Groundwater, 28*(4), 559-568.

Ceolin, S. (2019). *Preliminary Assessment of Onsite Wastewater Treatment System Effects on Lake Lanier Water Quality*. (Master of Science of Crop and Soil Sciences). University of Georgia and University of Padova,

Englert, A., Kemna, A., Zhu, J.-f., Vanderborght, J., Vereecken, H., & Yeh, T.-C. J. (2016). Comparison of smoothness-constrained and geostatistically based cross-borehole electrical resistivity tomography for characterization of solute tracer plumes. *Water Science and Engineering, 9*(4), 274-286. doi:10.1016/j.wse.2017.01.002

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., Dietrich, W. E., Flores, A., Grant, G., Kirchner, J. W., Mackay, D. S., McDonnell, J. J., Milly, P. C. D., Sullivan, P. L., Tague, C., Ajami, H., Chaney, N., Hartmann, A., Hazenberg, P., McNamara, J., Pelletier, J., Perket, J., Rouholahnejad-Freund, E., Wagener, T., Zeng, X., Beighley, E., Buzan, J., Huang, M., Livneh, B., Mohanty, B. P., Nijssen, B., Safeeq, M., Shen, C., Verseveld, W., Volk, J., & Yamazaki, D. (2019). *Hillslope Hydrology in Global Change Research and Earth System Modeling*. Water Resources Research, (55, 2).

Federal Water Pollution Control Amendments of 1972, § 1251 (1972).

Graham, J. L., Loftin, K. A., Meyer, M. T., & Ziegler, A. C. (2010). Cyanotoxin mixtures and taste-and-odor compounds in cyanobacterial blooms from the Midwestern United States. *Environ Sci Technol, 44*(19), 7361-7368. doi:10.1021/es1008938

Gunady, M., Shishkina, N., Tan, H., & Rodriguez, C. (2015). A Review of On-Site Wastewater Treatment Systems in Western Australia from 1997 to 2011. *Journal of Environmental and Public Health, 2015*, 1-12. doi:10.1155/2015/716957

Hinnell, A. C., Ferré, T. P. A., Vrugt, J. A., Huisman, J. A., Moysey, S., Rings, J., & Kowalsky, M. B. (2010). Improved extraction of hydrologic information from geophysical data through coupled hydrogeophysical inversion. *Water Resources Research, 46*(4). doi:10.1029/2008wr007060

Jantrania, A. (2015). Monitoring Protocol for on-Site Systems. Retrieved from https://vernonjames.ces.ncsu.edu/monitoring-protocol-for-on-site-systems/

LeGrand, H. E., Back, W., Rosenshein, J. S., & Seaber, P. R. (1988). Region 21, Piedmont and Blue Ridge. In *Hydrogeology* (Vol. O-2, pp. 201-208): Geological Society of America.

Lewis, W. M., Wurtsbaugh, W. A., & Paerl, H. W. (2011). Rationale for Control of Anthropogenic Nitrogen and Phosphorus to Reduce Eutrophication of Inland Waters. *Environmental Science & Technology, 45*(24), 10300-10305. doi:10.1021/es202401p

Lusk, M. G., Toor, G. S., Yang, Y.-Y., Mechtensimer, S., De, M., & Obreza, T. A. (2017). A review of the fate and transport of nitrogen, phosphorus, pathogens, and trace organic chemicals in septic systems. *Critical Reviews in Environmental Science and Technology, 47*(7), 455-541. doi:10.1080/10643389.2017.1327787

Markewich, H. W., Pavich, M. J., & Buell, G. R. (1990). Contrasting soils and landscapes of the Piedmont and Coastal Plain, eastern United States. *Geomorphology, 3*(3-4), 417-447. doi:10.1016/0169-555x(90)90015-i

McDowell, W., Brick, C., Clifford, M., Frode-Hutchins, M., Harvala, J., & Knudsen, K. (2005). Septic System Impact on Surface waters [Press release]

Miller, J. A. (1990). *Ground Water Atlas of the United States: Segment 6, Alabama, Florida, Georgia, South Carolina* (730G). Retrieved from http://pubs.er.usgs.gov/publication/ha730G

Morrissey, P. J., Johnston, P. M., & Gill, L. W. (2015). The impact of on-site wastewater from high density cluster developments on groundwater quality. *Journal of Contaminant Hydrology, 182*, 36-50. doi:10.1016/j.jconhyd.2015.07.008

Reynolds, J. M. (2011b). *An Introduction to Applied and Environmental Geophysics* (I. John Wiley & Sons Ed.).

Roberts, D. C. (2001). *A field guide to geology: Eastern North America*: Houghton Mifflin Harcourt.

Schindler, D. W. (1971). Carbon, Nitrogen, and Phosphorus and the Eutrophication of Freshwater Lakes. *7*(4), 321-329. doi:10.1111/j.1529-8817.1971.tb01527.x

Sheppard, C. R. C. (1982). Photic zone. In *Beaches and Coastal Geology.* (pp. 636-636). Boston, MA: Springer US.

The Georgia Department of Natural Resources Environmental Protection Division Atlanta, G. (2017). *Total Maximum Daily Load Evaluation for Lake Lanier in the Chattahoochee River Basin for Chlorophyll a*.

US Army Corps of Engineers, M. D. (2019). Lake Sidney Lanier FAQ. Retrieved from https://www.sam.usace.army.mil/Missions/Civil-Works/Recreation/Lake-Sidney-Lanier/Questions/

Zauyah, S., Schaefer, C. E. G. R., & Simas, F. N. B. (2018). Chapter 3 - Saprolites. In G. Stoops, V. Marcelino, & F. Mees (Eds.), *Interpretation of Micromorphological Features of Soils and Regoliths (Second Edition)* (pp. 37-57): Elsevier.

Zeng, R. Q., Meng, X. M., Zhang, F. Y., Wang, S. Y., Cui, Z. J., Zhang, M. S., Zhang, Y., & Chen, G. (2016). Characterizing hydrological processes on loess slopes using electrical resistivity tomography – A case study of the Heifangtai Terrace, Northwest China. *Journal of Hydrology*, *541*, 742-753. doi:10.1016/j.jhydrol.2016.07.033

CHAPTER 2

DETERMINING SPATIALLY DISTRIBUTED SOIL TEXTURE USING ELECTRICAL RESISTIVITY DATA IN AN ARTIFICIAL NEURAL NETWORK FRAMEWORK[1]

[1] Tancredi, M.T. and N. Gaur. To be submitted to *Geochemistry, Geophysics, Geosystems*.

Abstract

Soil is a fundamental part of many ecosystems and soil information can be crucial for a range of disciplines, from waste management and construction to global challenges such as climate change and food security. For this reason, the scientific community has been working to provide easily available and affordable soil data to the public. The aim of this study was to investigate alternative methods to determine soil textures for the Piedmont Physiographic Region in Georgia. I focused on the use of near-surface geophysical methods, specifically electrical resistivity tomography, paired with Artificial Neural Network (ANN) frameworks to estimate high resolution soil textural information. Soil resistivity data were collected for several months across the same hillslope and used as input information for the ANN. Other inputs included in the framework were relative depth of investigation and weekly antecedent rainfall. The models were tested and validated both on- and off-site. Depth proved to be a very strong textural predictor for the area. When used as the only input, I obtained values of $R^2 = 0.82$ and RMSE= 4.84% for the off-site validation. The additional information provided by rainfall and resistivity increased the accuracy of the model, improving the R^2 and RMSE values for the off-site validation to 0.91 and 3.53%, respectively. Estimations of soil textural classes based on the model with the best performance reflected those obtained from the soil cores collected at the sites. There is good potential in the use of electrical resistivity tomography to obtain soil information at the hillslope scale within the Piedmont Physiographic Region of Georgia.

Introduction

In the past decades, there has been a widespread acknowledgment of the role of soil for many ecosystem services and the need to obtain soil information for resolving compelling challenges ranging from wastewater disposal to climate change to food security (Keesstra, Bouma et al., 2016; Omuto, Nachtergaele et al., 2013; Sanchez, Ahamed et al., 2009). Specifically, decision making in agriculture, waste management systems, and construction regularly require detailed characterization of soil physical and hydraulic properties. These properties are also useful in environmental protection as they represent base information for hydrological models applied to predict water movement and transport of nutrients and pollutants. Solute movement in the soil is tightly linked to the characteristics of the subsurface (note that in this study "subsurface" and "subsurface profile" are used interchangeably to describe the section of the Earth surface comprising soil, saprolite, bedrock, water and organic matter). Soil pore distribution, which is a function of soil physical properties such as texture and structure, influences hydrodynamic dispersion, can determine preferential flow paths for water, and solutes that can be adsorbed on soil particles. Thus, a proper determination of soil characteristics is fundamental when studying solute transport. In an attempt to gather and provide stakeholders and policy-makers with the necessary soil data, efforts have been made to extend and improve proximal soil sensing and digital soil mapping (Grunwald, Thompson et al., 2011; Hartemink, Krasilnikov et al., 2013; Hengl, de Jesus et al., 2014).

Currently, it is possible to obtain free soil characterization information through work developed by the U.S. Department of Agriculture's (USDA) Natural Resources Conservation Service (NRCS). USDA-NRCS has developed three online soil geographic databases. Amongst them, Soil Survey Geographic (SSURGO) provides the highest level of detail. SSURGO soil maps have scales ranging from 1:12,000 to 1:63,360 and a maximum depth of investigation usually ranging between 150 and 200 cm. These maps are a good planning tool, but onsite sampling and study are still required for intensive use at finer scales (United States Department of Agriculture, Natural Resources Conservation Service et al., 1995). On the other hand, onsite analysis is mostly contained to point information and is not necessarily representatives of the entire area. The combination of SSURGO data and local analysis could partially

help to solve this issue. However, onsite soil surveys can be expensive, time-consuming (Banton, Cimon et al., 1997; Hazreek, Saad et al., 2015) and they are often limited to the upper layers of soil, not allowing for a complete description of the subsurface profile. The conventional and most common techniques for lab-based soil texture determination are sieving and sedimentation. Sedimentation can be conducted either with the pipette or hydrometer method (Soil Science Division Staff, 2017). These procedures can be precise in texture determination but require soil collection, processing, and an equipped laboratory. As such, they are labor-intensive, invasive, and expensive. In areas where soil types can be varied even in relatively close distances, soil texture determination should be based on a large number of samples. In those locations, alternative methods could speed up the analysis and significantly reduce the associated costs.

The peculiar characteristics of the Piedmont Physiographic Region, which stretches for 1600 km from New Jersey and southeastern Pennsylvania to eastern Alabama, make the use of conventional methods to determine sub-surface features particularly difficult. The bedrock is mainly constituted of igneous and metamorphic rocks that present distinct weathering profiles (Ohta & Arai, 2007), thus creating a heterogeneous subsurface in terms of layering and soil types (Figure 2.1A and 2.1B).

Figure 2.1. The images show exposed soil profiles on the shores of Lake Sidney Lanier, where the study took place. Notice the transition from residual soils to the saprolite layer (white arrows) and the differential weathering conditions in the soils of the Piedmont Physiographic Region. In A, the saprolite is disposed at an angle, while in B it is more horizontally layered. These different orientations show the effect of tectonic transformations.

The bedrock is usually covered by a mantel, called regolith, composed of a layer of saprolite with residual soil on top (Figure 2.2) (Miller, 1990). Saprolite is bedrock material that has weathered in place but still retains the original lithic fabric (Zauyah, Schaefer et al., 2018). The characteristics of saprolite influence the type of developing soil and depend on the parent material. The thickness of the mantel can vary between 0 and 40 meters (LeGrand, Back et al., 1988) and depends on the position (usually thicker at the interfluve and thinner on the slopes) and weathering responses of the parent material (Pavich, Leo et al., 1989).

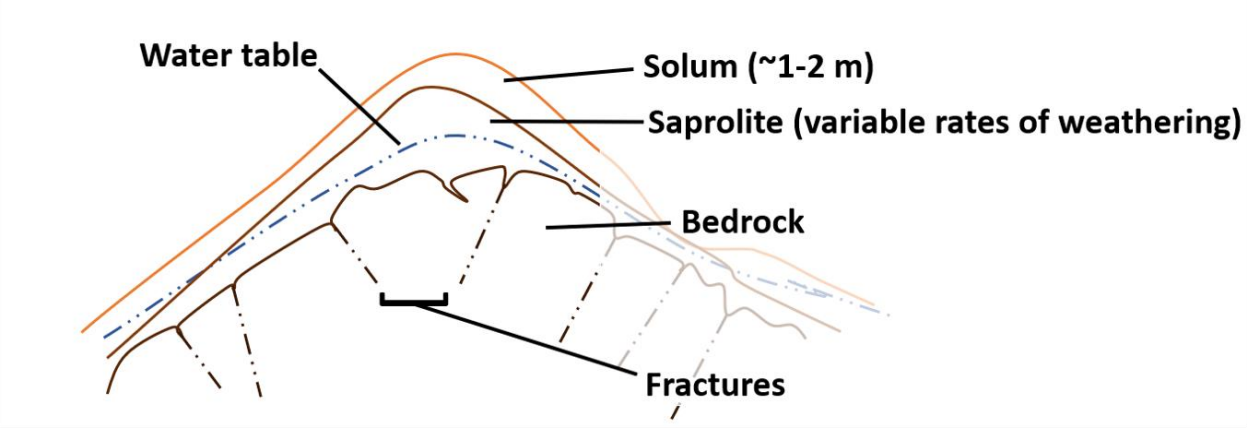


Figure 2.2. The Piedmont is characterized by rolling terrain. Residual thin soils and thick layers of saprolite with a metamorphic origin and a history of differential erosion constitute a mantel called regolith laying above the bedrock. The bedrock presents a system of fractures allowing for the water to flow. The water table lays mostly in the saprolite, just above the fractured bedrock. Adapted from Heath (1980).

Considering the complexity of the Piedmont system and the limitations of the conventional techniques, I explored a non-destructive geophysics based method, Electrical Resistance Tomography (ERT), as an alternative tool for time-efficient and detailed soil texture and sub-surface layering estimation at the hillslope scale in the Piedmont Region of Georgia.

Geophysical methods, in particular electrical resistivity methods, have been widely employed since the 1970s to monitor reservoirs and to detect groundwater pollution (Reynolds, 2011a). There is a growing consensus on the benefits that subsurface imaging obtained through geophysics provides to hydrologic models (Hinnell, Ferré et al., 2010) and the ability of ERT to give information about water

content, near-surface structures, and hydro-geological interfaces and processes (Zeng, Meng et al., 2016). More details on the principles on which this technique is based is provided in the Methods section.

Soils have intermediate electrical properties and those are mostly determined by their physical and chemical characteristics such as texture, salinity of the fluid present in its pores, and water content (Samouëlian, Cousin et al., 2005). Using a simplified approach, one can consider the total resistivity signal measured on the soil by ERT as composed by a dynamic (water and ions content) and a static (soil texture) component. The estimation of soil structures from resistivity values can also be considered as the first step of an uncoupled hydrogeophysical inversion. This process has been widely explored in the past 30 years with different techniques (Cassiani, Bruno et al., 2006; Deiana, Cassiani et al., 2008; French & Binley, 2004; Hubbard, Rubin et al., 1999; Kemna, Vanderborght et al., 2002; Sandberg, Slater et al., 2002; Singha, Day-Lewis et al., 2015; Vereecken, Binley et al., 2006) and is composed of three distinct procedures, the first one being the inversion of the geophysical data collected in the field to assess the geophysical properties of the area (Camporese, Cassiani et al., 2015; Hinnell, Ferré et al., 2010). In the past decades, the relationship between soil resistivity and other soil properties, such as bulk density, soil suction, angle of internal friction, plasticity index, and degree of soil weathering has been shown (Irfan & Syed, 2012; Piegari & Di Maio, 2013; Son, Oh et al., 2010). Studies have also investigated the correlation with texture, effective particle size, and mineralogical composition of particles to identify soil types. The literature shows correlation coefficients (based on a linear regression) between sand, silt and clay, and electrical conductivity (the inverse of electrical resistance) between 0.45 and 0.64. The electrical resistivity method was successfully applied for soil identification and to obtain insights on effective soil particle size (Banton, Cimon et al., 1997; Hazreek, Saad et al., 2015; Kouchaki, 2017; Siddiqui & Osman, 2013). In most studies, the established relationships were defined through different types of regression. Even though regression can capture part of the correlation, the derived information could be less than ideal, due to the complexity of the relationship between soil and its properties. Siddiqui, Pathan et al. (2015) showed that Artificial Neural Networks (ANN) would perform better than simple regression models at predicting soil properties.

ANNs are one of the most popular forms of machine learning. Machine learning is mainly used to process and mine large amounts of data for meaningful information in an automated fashion (Maglogiannis, 2007). Through machine learning, computers can learn from experience, by example, and by analogy to improve their performances over time (Negnevitsky, 2005). With regards to ANNs, the underlying concept is to build a model of reasoning resembling the human brain. A multi-layer neural network is composed of several units, called neurons, which are connected by weighted links. These neurons can be grouped into three types: input, output, and hidden. Input units contain the information to be processed, output neurons contain the results of the processing activity and hidden units are the ones processing the information and moving them from the input to the output (Kotsiantis, Zaharakis et al., 2006). The weighted links connecting the neurons determine the importance of each input. Neural networks can be trained through consecutive adjustments of the weight of these links. To build an ANN, we have to:

1. decide its architecture (determine the total number of neurons and the connections among them),
2. pick the learning algorithm we want the ANN to use, and
3. train the network through a series of examples.

As it is for brain neurons, the neurons in ANNs are activated when the signals are greater or equal to a set threshold value. For the ANNs, the activation function is the weighted sum of the different inputs reaching the neuron. The most common activation functions are step function, sign function, sigmoid function, and linear function (Negnevitsky, 2005). Their graphical representation and correspondent equation are shown in Figure 2.3.

In this work I investigated the utility of soil resistivity, depth, and rainfall to predict percentages of sand, silt, and clay (soil texture). I also explored how different ANN architectures and combinations of the inputs could influence these predictions.

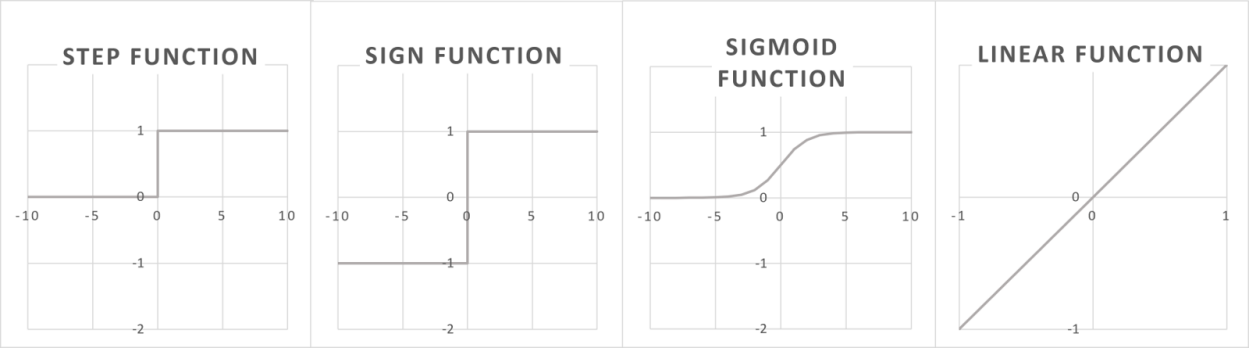


Figure 2.3. Graphical representation and equations of the four most common activation functions used in Artificial Neural Networks.

Methods

Field study site and sampling protocol

This study was conducted from March 2019 to July 2020, in Buford and Hall Counties (Georgia, USA), in two coves of the southern part of Lake Sidney Lanier. Lake Sidney Lanier is a man-made reservoir situated at about 48 km northeast of Atlanta. It was created in 1956 through the construction of the Buford Dam and is currently operated by the US Army Corps of Engineers. The coves enclosing my study sites were mapped at scales ranging from 1:15,800 to 1:20,000 for the SSURGO database, and are the results of multiple soil survey areas (Soil Survey Staff, 2019).

The bedrock in the area surrounding the lake is constituted of Lower Paleozoic and Precambrian felsic gneiss, schist, and granite (Miller, 1990), and the geological properties are characteristics of the Piedmont province. Three locations, referred to as site 1, site 2, and site 3 were selected for this study. The sites are located on either side of a small lake cove. Sites 1 and 2 are in an area that I called “Finger 1” and site 3 is on “Finger 2” (Figure 2.4). The location of the sites will not be specified any further to protect the privacy of the landowners. Site 1 has a “developed, open space” land cover (Homer, Fry et al., 2012) and, according to the SSURGO database, is characterized by two soil types, Pacolet sandy loam, 6 to 10 percent slopes, moderately eroded and Pacolet-Saw complex, 15 to 45 percent slopes, stony. The

latter characterizes most of the area and the entire surveyed section. The average elevation of the site is 334.1 m MSL. Site 2 presents a mix of land covers, specifically “developed, open space”, “developed, medium intensity” and “evergreen forest” (Homer, Fry et al., 2012). In the SSURGO database, two thirds of the study site are characterized by Pacolet soils, 10 to 15 percent slopes, eroded, with the remaining area being Chestatee stony sandy loam, 15 to 25 percent slopes. The average elevation of the site is 332.9 m MSL. Finally, for site 3 the land covers are “developed, medium intensity”, “mixed forest”, and “evergreen forest” (Homer, Fry et al., 2012). From the SSURGO, the soil types are Madison sandy clay loam, 10 to 15 percent slopes, moderately eroded on the top of the ridge and Madison sandy clay loam, 15 to 45 percent slopes, eroded for the rest of the hillslope. The average elevation of the site is 331.4 m MSL.

Figure 2.4. Aerial map of the two areas enclosing the three study sites. Sites 1 and 2 are situated on “Finger 1”, while site 3 is on “Finger 2”.

At all sites, ERT measurements were taken on three transects to visualize the approximate spatial variation in the entire area of each study site. Two transects were parallel to the shoreline, while one was perpendicular to the shoreline. At sites 1 and 2 I collected measurements monthly, thus allowing us to observe how the resistivity of the same portion of subsurface changed during the seasons. At both sites, the transect ends were marked on the ground with plastic pins to keep the position of the electrodes consistent over seven months of data collection (non-continuously from May 2019 to February 2020). At site 3 data were collected just once.

ERT measurements and inversion

Electrical Resistivity Tomography, ERT, (also called Electrical Resistivity Imaging, ERI) is one of the geophysical techniques that exploit electric currents and their properties to detect variations in the ground. Diverse materials allow for electrons to flow through them differently. Metals and electrolytes, for example, show great conductivity, while other substances like air and ice are considered insulating since electron movement in them is very small (Samouëlian, Cousin et al., 2005).

Ohm's Law (Equation 2.1) defines the factors determining current flux in a conductor as:

$V = IR$ [2.1]

where: V is the voltage, or potential drop, across the material (conductor) measured in volts (V), I is the current in amps (A), and R is the resistance measured in ohms (Ω). The resistance is what determines the potential drop between two faces of a uniform cube through which the current is passing. R is directly proportional to the length L between the two faces of the cube and inversely proportional to its cross-sectional area A (Figure 2.5).

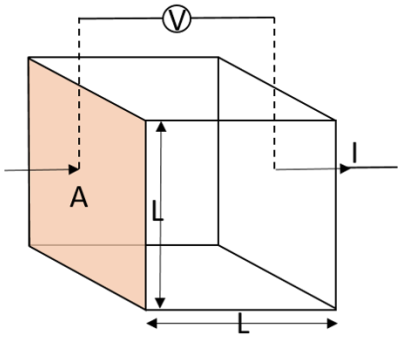


Figure 2.5. Basic concept of resistivity effect on an applied current I and the resultant potential drop V in a block of side length L and area A. $\rho = VA/IL$ (Ωm) from $R = V/I$ and $R = \rho(L/A)$.

These relations can be expressed as

$R = \rho(L/A)$ [2.2]

where: ρ is referred to as the 'true' resistivity. Resistivity is a constant of proportionality, it is measured in ohm-meters (Ωm) and it is a property of the material (Milsom & Ager, 2011; Reynolds, 2011a).

ERT data were collected using a SuperSting™ R8 IP Earth Resistivity IP Meter from Advanced Geosciences Inc. (AGI). The instrument has eight channels, allowing to measure the electric potential between nine electrodes simultaneously for each current injection. A power of 100W is supplied through a 12 V DC battery. Max current is 500mA and max error for the standard deviation is 2.0%. Every electrode consists of a stainless-steel stake, about 30 cm long, and a stainless-steel spring. The spring is used to connect the stake to the metallic contact of the multi-electrode cable. The SuperSting™ can record data through multielectrode cable systems, thus giving the possibility to measure the resistivity for every possible group of four electrodes and generate a 2D pseudo-section in a single measurement event. The pseudo-section is a V-shaped graphical representation of the different resistivities measured. The transect length defines the x-axis, while the y-axis shows the depth of the survey. The maximum depth is determined both by the position of the electrodes, or array, and by the transect length, usually about 20% of the total length (Reynolds, 2011a). In the pseudo-section, each resistivity value is plotted having the center of the quadrupole as the x-value and a distance relative to the positions of the dipoles (varying with the array) as the y-value (Reynolds, 2011a; Santarato, Zeid et al., 2015).

Prior to the beginning of every measurement event, the contact resistance of all electrodes was tested, to ensure good contact of the electrodes to the ground and guarantee good quality of the measured data. The system I used allows the collection of data in automatic mode according to the parameters chosen by the surveyor. In my case, the electrodes were displayed on a linear transect and spaced 1 m apart from one other. The total transect length varied with the characteristics of the sites, but it was always at least 28 m.

Data were inverted singularly using AGI EarthImager™ 2D. Information about the topography of the site was accounted for in the inversion process. The objective of the inversion is to find a model that is as close as possible to the reality of the subsurface. To this attempt, the data are elaborated through a non-linear parameterization. A first guess model is elaborated, and least-squares optimization is applied until the differences between observed and modeled data are minimized. The degree of misfit of those

differences is indicated by the percentage root mean squared (RMS) error (Miller, Routh et al., 2008; Milsom & Ager, 2011; Reynolds, 2011a; Samouëlian, Cousin et al., 2005). The RMS is calculated as:

$$RMS = \sqrt{\frac{\sum_{i=1}^{N}\left(\frac{d_i^{Pred} - d_i^{Meas}}{d_i^{Meas}}\right)^2}{N}} * 100\% \qquad [2.3]$$

where N is the total number of measurements, di^{Pred} is the predicted data, and di^{Meas} is the measured data. A single erroneous data point could lead to a large RMS error; therefore, the RMS error does not indicate the percentage of bad data points (Personal communication with AGIUSA Geophysical Engineer).

Soil samples collection and soil texture classification

Soil cores were hand-augered at 11 locations on site 1 across a 42 m hillslope, and at 4 locations, and at 2 locations across two 28 m hillslopes on sites 2 and 3 respectively (Figure 2.6).

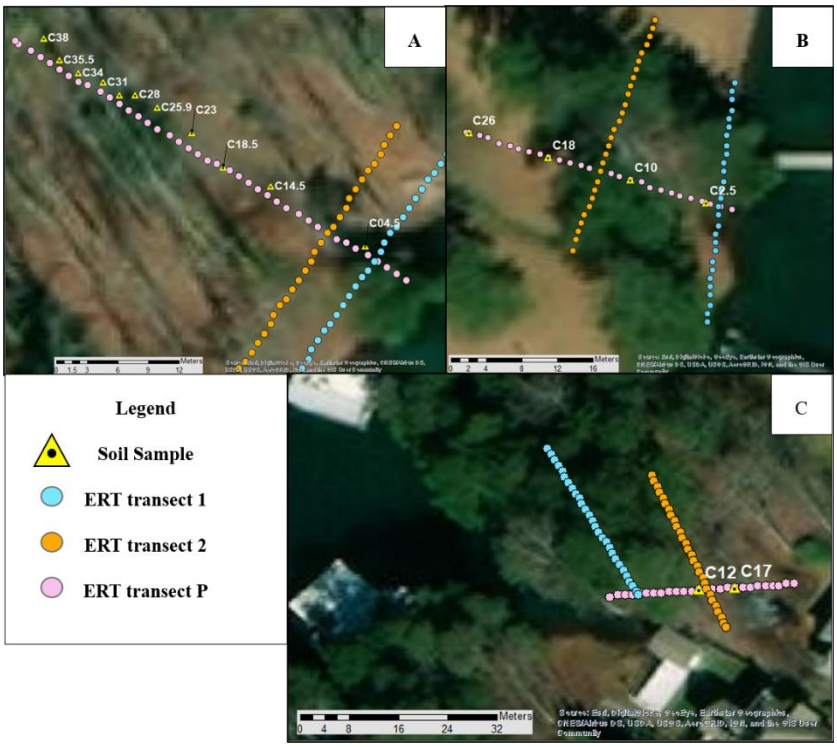


Figure 2.6. Location of the soil cores collected at sites 1(A), 2(B), and 3(C). The cores were collected next to the ERT transect P, perpendicular to the shoreline, and spatially distributed on the entire length of the hillslope to capture as much variability as possible. Transects 1 and 2 are parallel to the shoreline.

Samples were collected as 30-cm increments from the surface to a maximum depth of 615 cm. The total length of the soil cores varied among locations according to characteristics of the subsurface. The hand auger I used had a maximum depth of 7.5 m, but, in some cases, big tree roots or rocks caused my cores to be shallower. I then performed particle size analysis on the samples to determine soil texture at different depths using the LS 13 320 Laser Diffraction Particle Size Analyzer (LDPSA). The instrument estimates the size distribution of soil particles suspended in solution through light scattering intensity as a function of light wavelength and the scattering angle. A beam of light encountering a particle is scattered with an angle that depends on the particle size, creating a unique scattering pattern. Analyzing all the generated patterns, it is possible to determine the particle size distribution of the analyzed sample (Xu, 2002).

Before proceeding with the analysis, each soil sample was ground to avoid the presence of aggregates and sieved to remove rocks or remaining root parts. Samples were then passed through a soil splitter to ensure the representativity of the smaller amount used for the analysis. Hexametaphosphate was added to two replicates for each soil and used as a dispersing agent to separates the smaller particles suspended in solution. After 16 hours of shaking at 120 oscillations per minute, the sample was analyzed using the LDPSA.

For soils collected at site 1, we used 0.25 g for soils that looked more clayey, 0.4 g for those that looked more sandy, and intermediate weights for those that didn't show a clear content. The distinction helped speed up the analysis process, reducing the number of times the LDPSA had to dilute the samples with more clay to be able to detect the diffracted beam of light (Ceolin, 2019). Since sites 2 and 3 had a reduced number of locations, I used 0.5 g of soil for every sample, independently from their apparent texture.

Artificial Neural Network to predict soil texture

The ERT data and the results from the soil texture analysis from site 1 were used to create five models predicting soil texture and its changes with depth. The models were developed using Artificial

Neural Networks. The five models have different structures, thus allowing us to compare the different levels of complexity and their performances. They are presented from the simplest to the most complex. The inputs were resistivity data at the different depths, corresponding depth of investigation, and antecedent weekly rainfall. Depth had the strongest correlation with the percentages of sand, silt, and clay (Figure 2.7), so I developed a model having depth as the sole input.

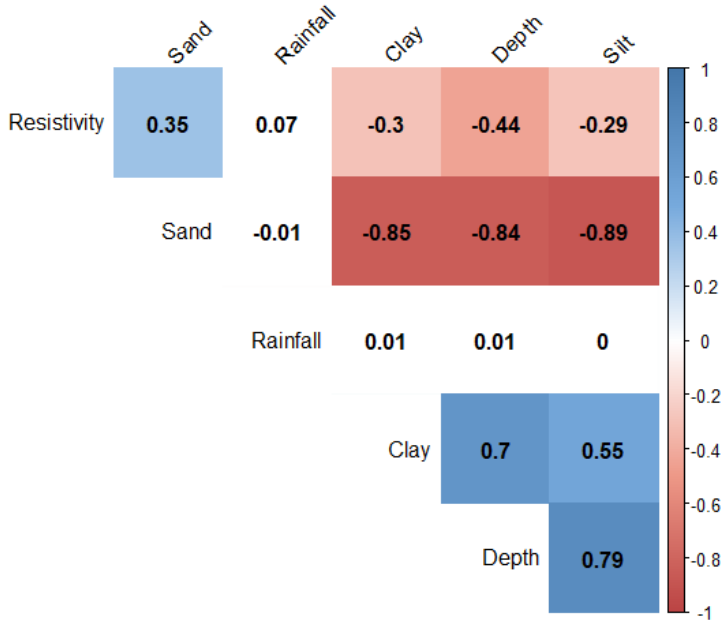


Figure 2.7. Correlation matrix of the inputs (depth, resistivity, and rainfall) and outputs (percentages of sand, silt, and clay) for the ANN models. Depth is expressed as negative values (from 0 at the surface to -4.85 for the deepest included location). The correlation is calculated using the Spearman method, the color represents a correlation value varying from -1 to +1, and the number at the center of the square indicates the associated correlation coefficient.

In accordance with the theoretical conceptualization, resistivity and sand are positively correlated, as sand is usually not able to retain water and solutes and, similarly to rocks, does not show high inherent conducting properties. Silt and clay, instead, because of their smaller size can create the conditions for retention of water, thus favoring the movement of electrons and showing lower resistivity values. Additionally, clay particles are characterized by the presence of negative charges around their surface, that further enhance their conducting properties.

Looking at the correlation between depth and textures, the values reaffirm what is expected for the soil characterization of the area (it is important, when reading the matrix, to remember that depth was expressed in negative terms as moving deeper from the surface). In the Piedmont region clay content is usually limited to the Bt horizon, between 0.07 and 1 m for Cecil, Pacolet, and Saw soil series (Soil Survey Staff, Natural Resources Conservation Service et al., 2020). Since the data used to create the matrix extend to -4.85 m, clay is shown to have a positive correlation with increasing distance from the surface. On the contrary, sand shows a negative correlation at deeper depths because its content increases when moving into the saprolite layers.

My ANN models (Figure 2.8) were developed in MATLAB (Version R2019b), starting from the Neural Net Fitting app (The MathWorks Inc., 2019). Model 1 has one single input, depth (I used the relative depth of investigation provided by the ERT to keep the profile consistent and be able to compare the models), three outputs, and one hidden layer with 20 neurons (Figure 2.8 A). Model 2 has three inputs, three outputs, and one hidden layer with 20 neurons (Figure 2.8 B). In model 3, resistivity and rain are imported together in the first hidden layer (20 neurons). Depth is processed in a second hidden layer, together with the outcomes of the first calculation (Figure 2.8 C). In the fourth model, each input is imported separately following the order depth, resistivity, and rainfall (Figure 2.8 D). Model 5 has the same structure as model four, but the inputs are imported as resistivity, rainfall, and depth (Figure 2.8 E). Additionally, this model was trained on data from both sites 1 and 2. Most of the soil collected at site 2 was constituted by fill material, so I included it to emphasize the effect of resistivity and rainfall over depth. In both models 4 and 5, the hidden layer processing depth has a linear transformation function.

All models have an additional last hidden layer with neurons that normalize the results before calculating the output. The inputs are imported in the models as decimal values, instead of percentages. The ‘softmax’ neural transfer function is used to normalize the values on a probability distribution that sums up to 1. This normalization process ensures that, together, the three predicted percentages give a total of 100. Soil texture information, expressed in percentages of sand, silt, and clay particles of the

locally extracted soil samples, was used as targets to train the models. The models were trained with the Levenberg-Marquardt backpropagation algorithm (function 'trainlm').

Coordinates of the locations of the ERT data and the soil samples were collected in the field and plotted on ArcMap 10.5.1. This step allowed us to pair the soil samples with the closest ERT data points. Once this relation was established, the available data were arranged to link observed percentages of sand, silt, and clay to correspondent resistivity values, for the entire available depth of the soil cores. The procedure was repeated for all the resistivity data collected over the months, thus allowing the model to pick how a single texture could show a range of values according to seasonal changes.

Before being fed to the model, the complete dataset was split into two subsets, one for training and one for testing the model. To avoid bias in the accuracy of the model due to the quality and quantity of information provided by every single soil core, I performed training and testing of the model on all possible location combinations using a holdout method. Since my resistivity values were collected on the same locations for several months, I made sure that if a location were included in the training set, that same location would have not been present in the testing one to keep the training data independent of the testing data. The dataset splitting process was accomplished through resampling without repetition.

I also analyzed how the accuracy of the prediction changed when increasing the number of inputs fed to the model for its training. After training and testing, I calculated values of R^2, RMSE, and a third Objective Function (OF) for models created changing the number of locations selected for the training, from a minimum of three to a maximum of eight. Values of R^2 and RMSE were used to determine the best number of locations to include to have the best predictions. Since the ranges of RMSE could be very different from that of R^2, I developed an OF that would have its range constrained between 0 and 1, give equal weighting to both RMSE and R^2 and consequently be used to find the best absolute model among those predicted with the best number of locations for the training. It was calculated as:

$$0.5 * \left(1 - R_i^2\right) + 0.5 * \left(\frac{RMSE_i - RMSE_{min}}{RMSE_{max} - RMSE_{min}}\right) \qquad [2.4]$$

where R_i^2 and $RMSE_i$ are the values for every single model, while $RMSE_{min}$ and $RMSE_{max}$ are the minimum and maximum RMSE values among all models predicted with the same number of locations.

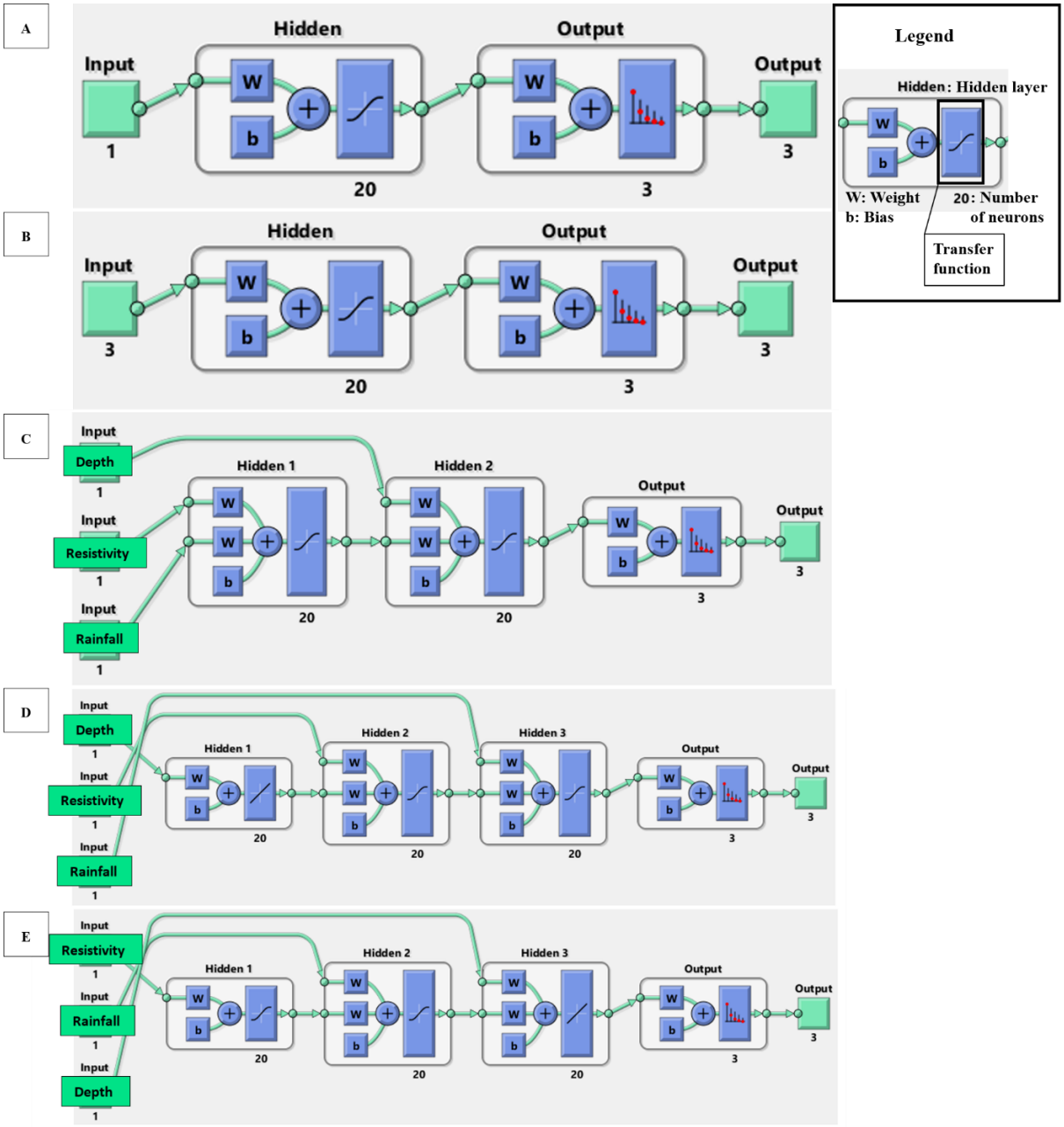


Figure 2.8. Schematic representation of the artificial neural networks developed to predict soil textures (A = model 1, B = model 2, C = model 3, D = model 4, E = model 5). The first model has depth as the only input, while the other models also have resistivity and rainfall. All the models have three outputs (percentages of sand, silt, and clay). The inputs are fed to one or more hidden layers, with 20 neurons. Most hidden layers have a tansigmoid transfer function, but models 4 and 5 have a linear transfer function for the layer processing depth. The outputs are always normalized before being finalized to ensure that the three predictions together sum up to 100% of the soil sample.

Validation of the models

For all my five structures, the model showing the lowest OF value was selected to create predictions of the soil textures for sites 1, 2, and 3. In my validation, I looked at correspondence in percentages of sand, silt, and clay. For the model with the best absolute performances I also compared the observed and predicted textural classes.

For the on-site validation, I looked at the R^2, RMSE of the selected model and at the comparison between observed and predicted textural classes. The outputs were converted into soil textural classes using The Soil Texture Wizard functions in RStudio Version 1.2.5019 (Moeys, 2018).

The second validation was performed off-site to test the model and its ability to give accurate predictions in different locations than the area used for the training process. Percentages of sand, silt, and clay for site 2 were separately calculated using resistivity data collected over six months, while the ones for site 3 were calculated over a single ERT survey. The predictions for site 2 were compared to check the possible effects of seasonality. For both sites, the locations were selected along the perpendicular ERT transect and spaced to cover most of the hillslope. At site 2, three out of the four locations were in a fill area.

Results

Figure 2.9 shows the lake levels measured at Buford Dam and the weekly antecedent rainfall (used as input for the ANN models) for the period between May 2019 and February 2020. The inverted resistivity values measured in proximity to the soil cores, at sites 1, 2 and 3 are shown in Figures 2.10, 2.11, and 2.12, respectively. The images show the results of the inversion process adjusted to the real topography of the sites. The time-series of the inverted images show the changes in subsurface resistivities throughout the seasons. The effect of the changing soil water content is very well captured, exhibiting both areas where the heavy rainfalls do not seem to affect soil resistivity (Figures 2.10, in the boxes on the right, and 2.11) as well as sections where we observe a quite broad variability of values (Figure 2.10, in the boxes on the left).

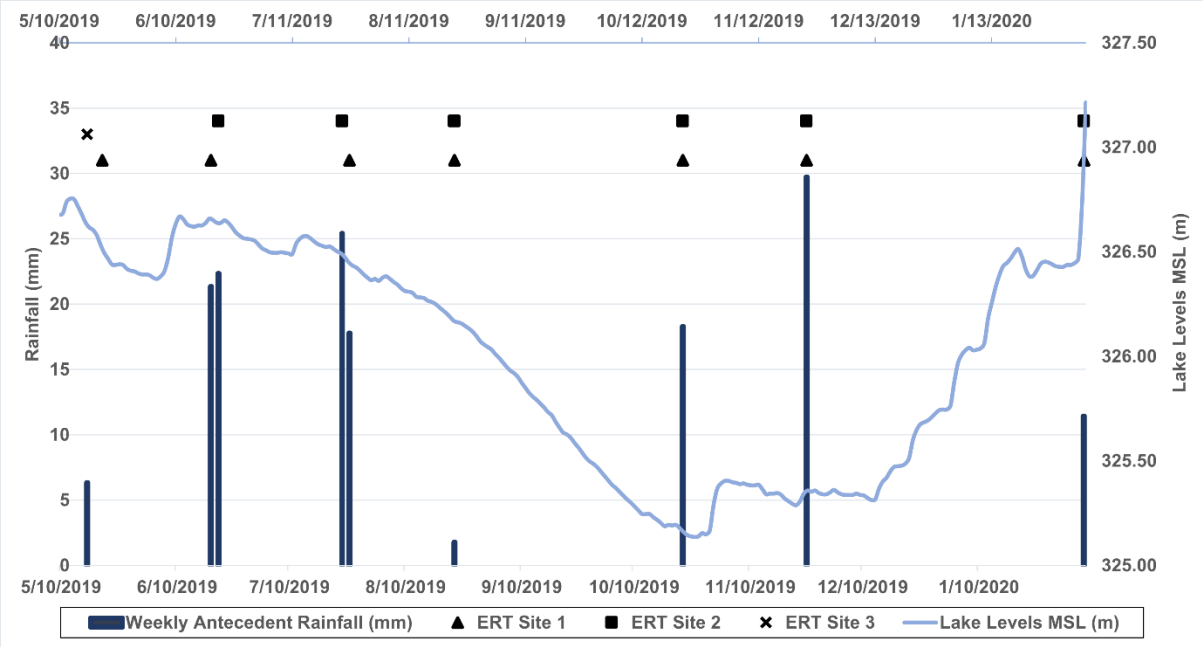


Figure 2.9. The figure shows the lake levels (line) measured at Buford Dam and the weekly antecedent rainfall used as input for the ANN models (bars) for the period between May 2019 and February 2020. Triangle, rectangle, and x symbols indicate the dates of the ERT surveys at the sites.

In the middle part of the slopes, particularly for site 1, higher resistivity is evident in October and November, when the lake level is very low. At site 3, one of the transects parallel to the lake (transect 1) was collected very close to a dry saprolite outcrop. The portion of saprolite was clearly represented in the ERT inverted profile as an area of higher resistivity (Figure 2.13). This transect was particularly important because it allowed us to ground truth my interpretation of the inverted resistivity data and confirm the possibility to identify saprolite features using ERT.

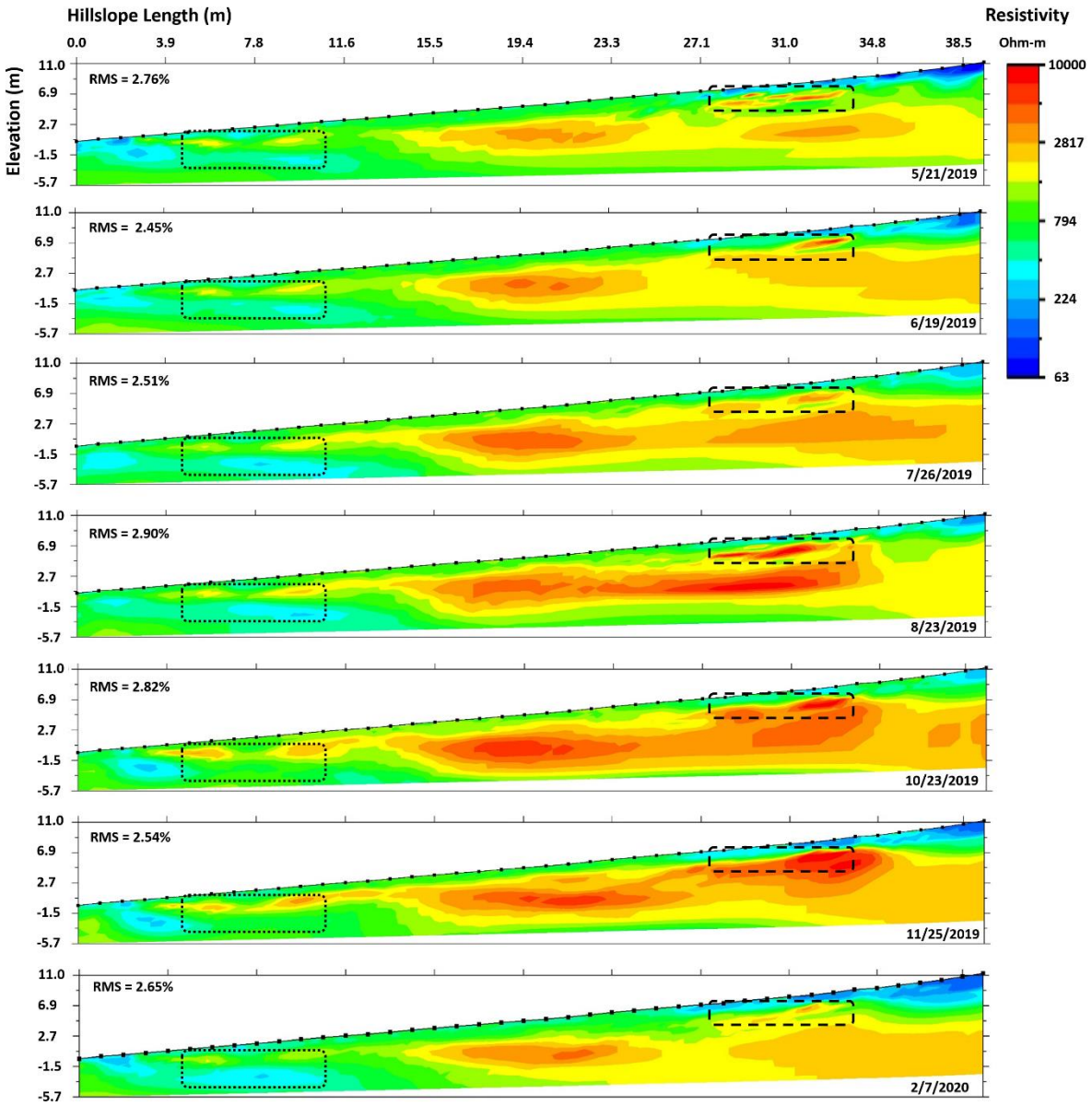


Figure 2.10 Inverted resistivity values measured on the perpendicular transect at site 1, from May 2019 to February 2020. The black boxes highlight how the different wetness conditions can influence the resistivity of the same area. In the dashed area we can observe a thin layer of soil that remains highly resistive throughout the seasons. In the dotted section the lowest part of the box shows an area that is usually saturated, having a lower resistivity, but that changes in resistivity on the surface.

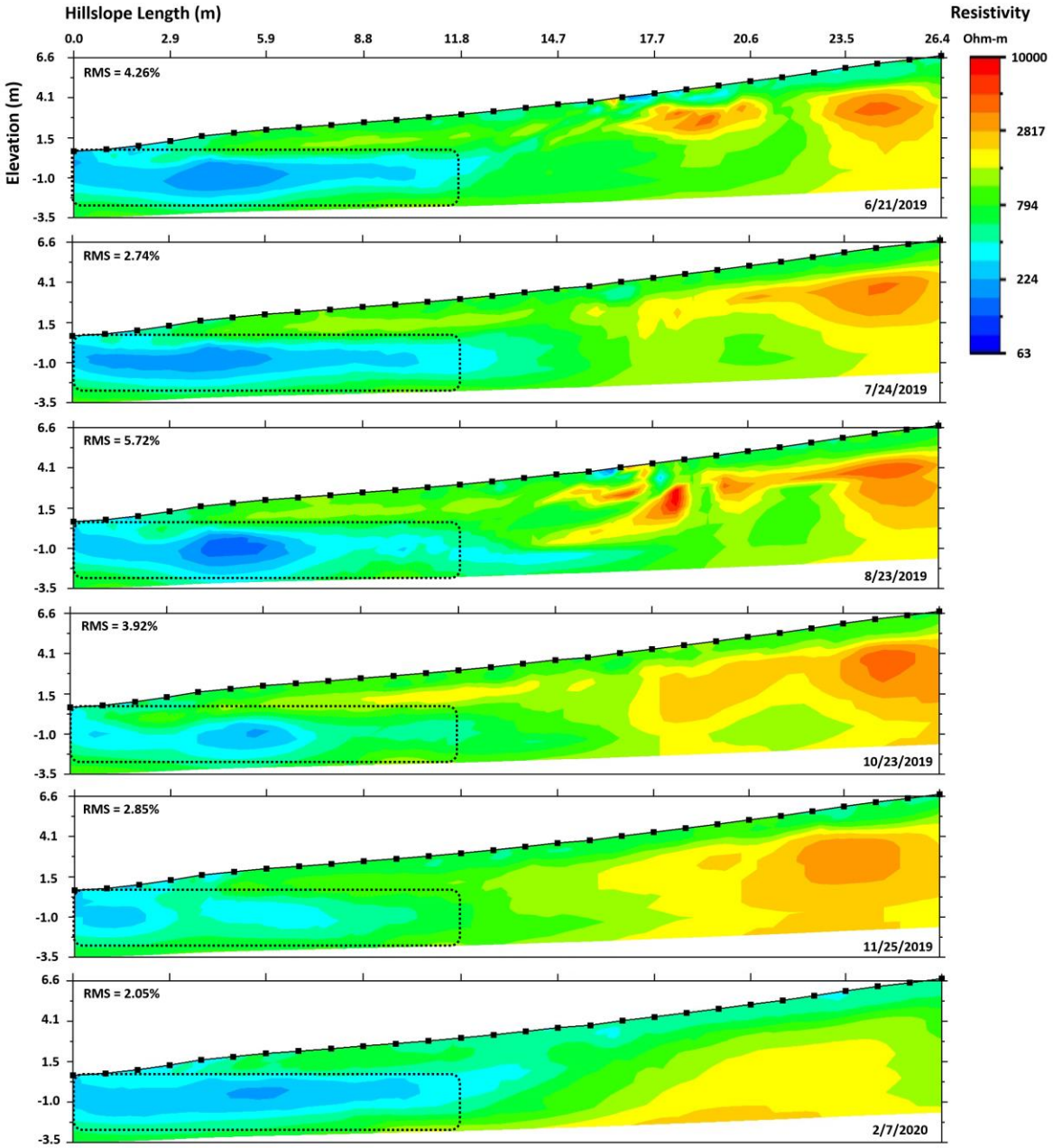


Figure 2.11 Inverted resistivity values measured on the perpendicular transect at site 2, from June 2019 to February 2020. The dotted box highlights an area, close to the shore, that stays saturated in the various months, thus consistently showing low resistivity values.

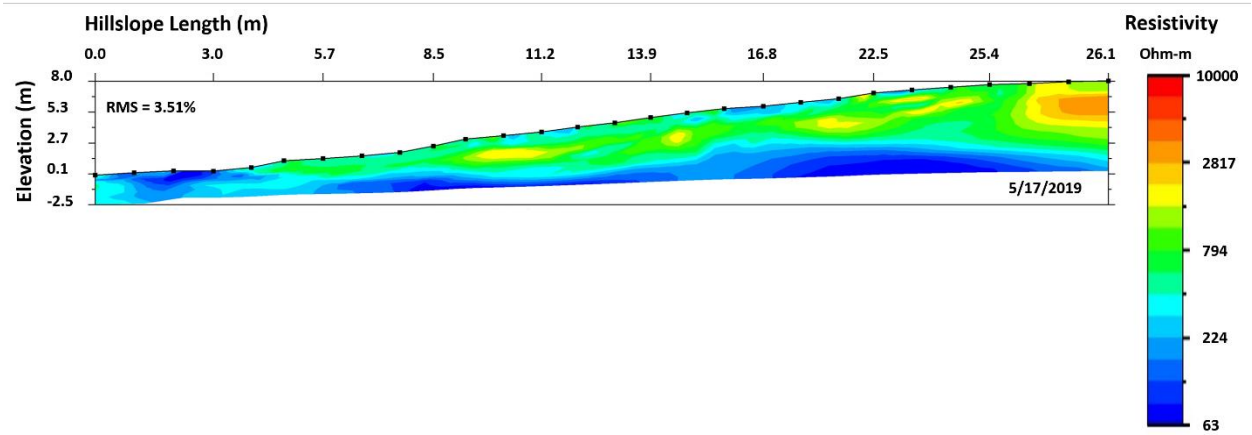


Figure 2.12 Inverted resistivity values measured on the perpendicular transect at site 3. The data were collected in May 2019.

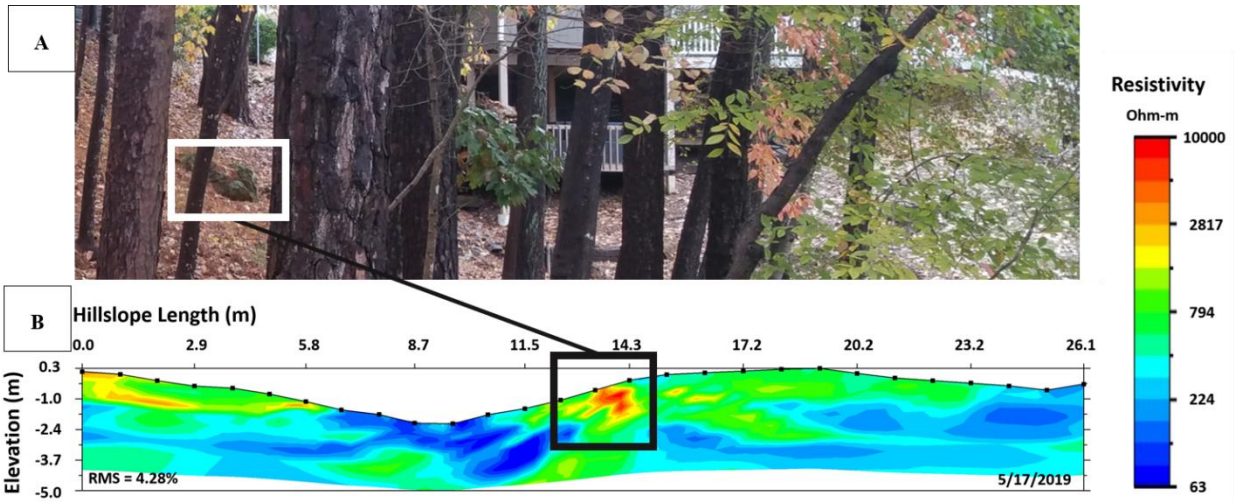


Figure 2.13. The figure shows a saprolite outcrop (in the white box in figure A) and that was clearly represented in the ERT inverted profile of transect 1 at site 3 (in the black box in figure B).

The results from the LPSA, for all the cores collected at the three sites, are presented both in terms of textural classes (Figure 2.14) and percentages of sand, silt, and clay (Figure 2.15).

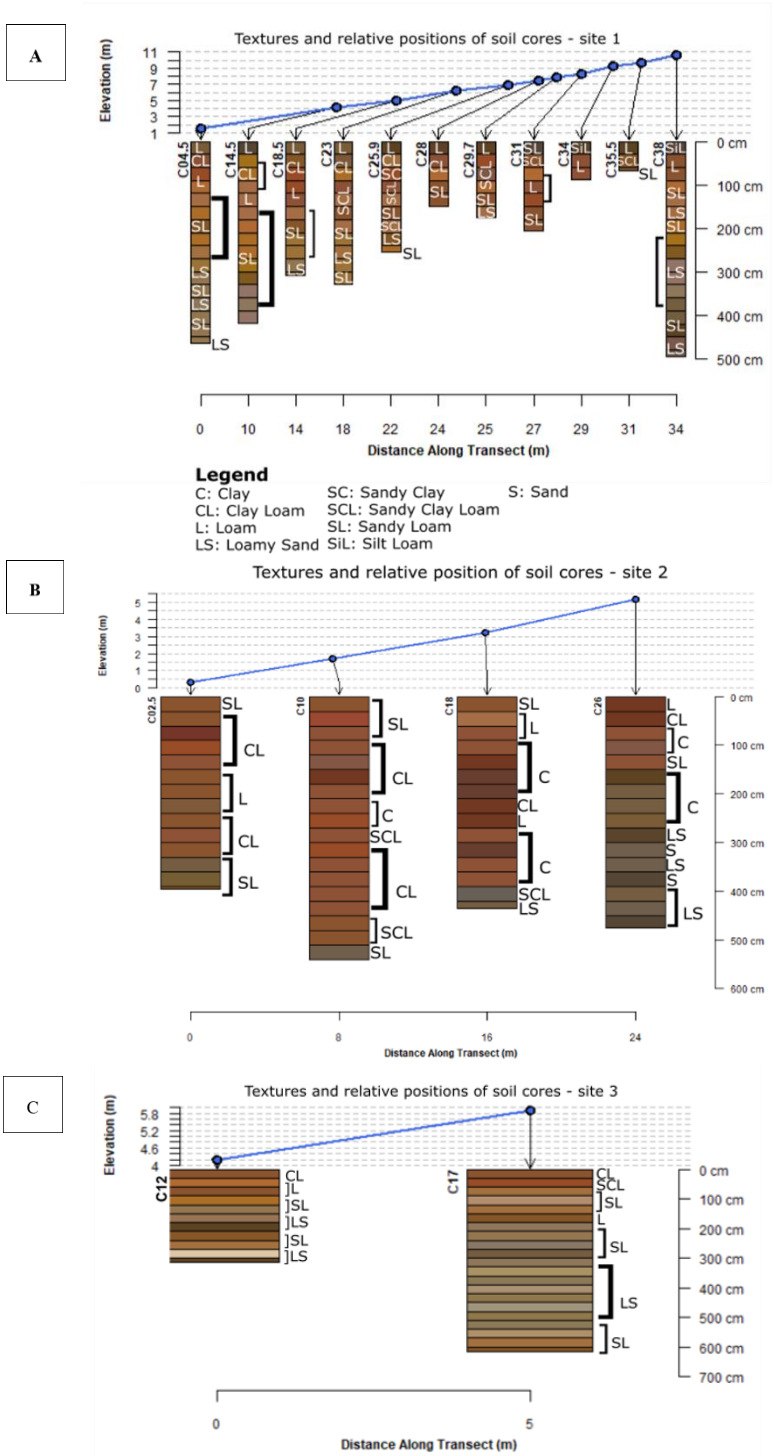


Figure 2.14 The graph shows the relative locations, length and textures of the soil core collected along the hillslopes at sites 1(A), 2(B), and 3(C). The textural classes are presented for the 30 cm increments. The colors were determined moist using the Munsell Soil Color Chart (Munsell, 2010).

Site 1 has the highest number of soil cores and the highest depth variability. Clay content ranges from 1.92 to 38.2 % at site 1, from 2.02 to 56.6 % at site 2, and from 2.59 to 33.3 % at site 3. Silt content ranges from 11.4 to 51.5 % at site 1, from 9.65 to 34.5 % at site 2, and from 13.27 to 42.8 % at site 3. Sand content ranges from 27.7 to 82.6 % at site 1, from 19.8 to 87.6 % at site 2, and from 23.85 to 83.8 % at site 3.

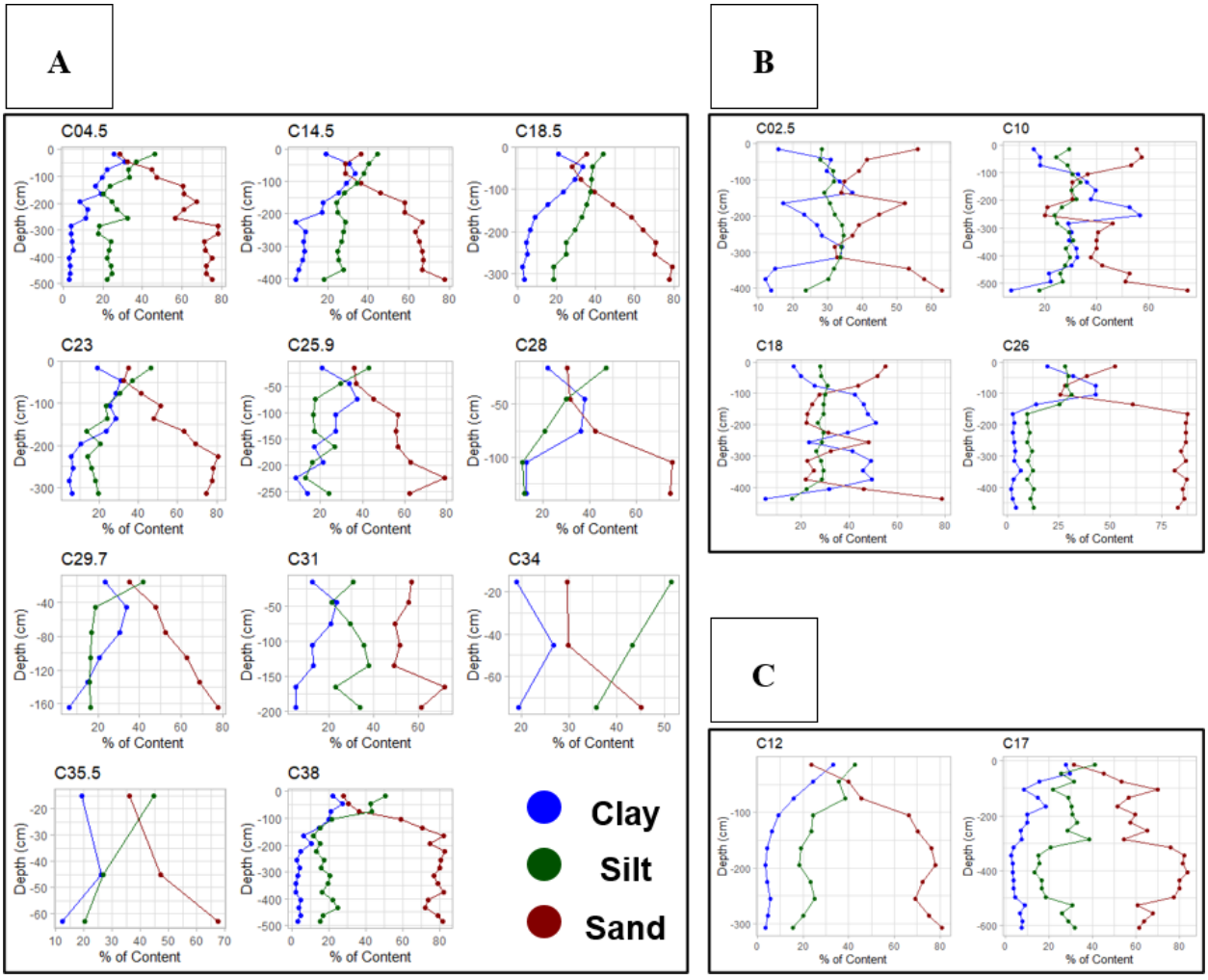


Figure 2.15 The graph shows how the percentages of clay, sand, and silt, change with depth at sites 1(A), 2(B), and 3(C). Since the samples were collected at the 30 cm increments, the depth used for the graph is the middle of the soil sample.

The results from the LPDSA were used to train the five ANN models. Figure 2.16 shows the ranges of R^2, RMSE, and OF for the models, when changing the number of locations included for the training (from a minimum of three to a maximum of eight). A model with a large spread in its R^2, RMSE,

or OF values would imply a large dependence on the data used to train the model while a smaller spread would imply a more robust model. All models, but model 4, had the best performances when trained with data from 8 locations. For an area as big as my study site, a training data with 8 spatial locations would typically be obtained in a Level 3 or 4 detailed survey by a soil scientist. Model 4 predicted better when trained with 7 locations.

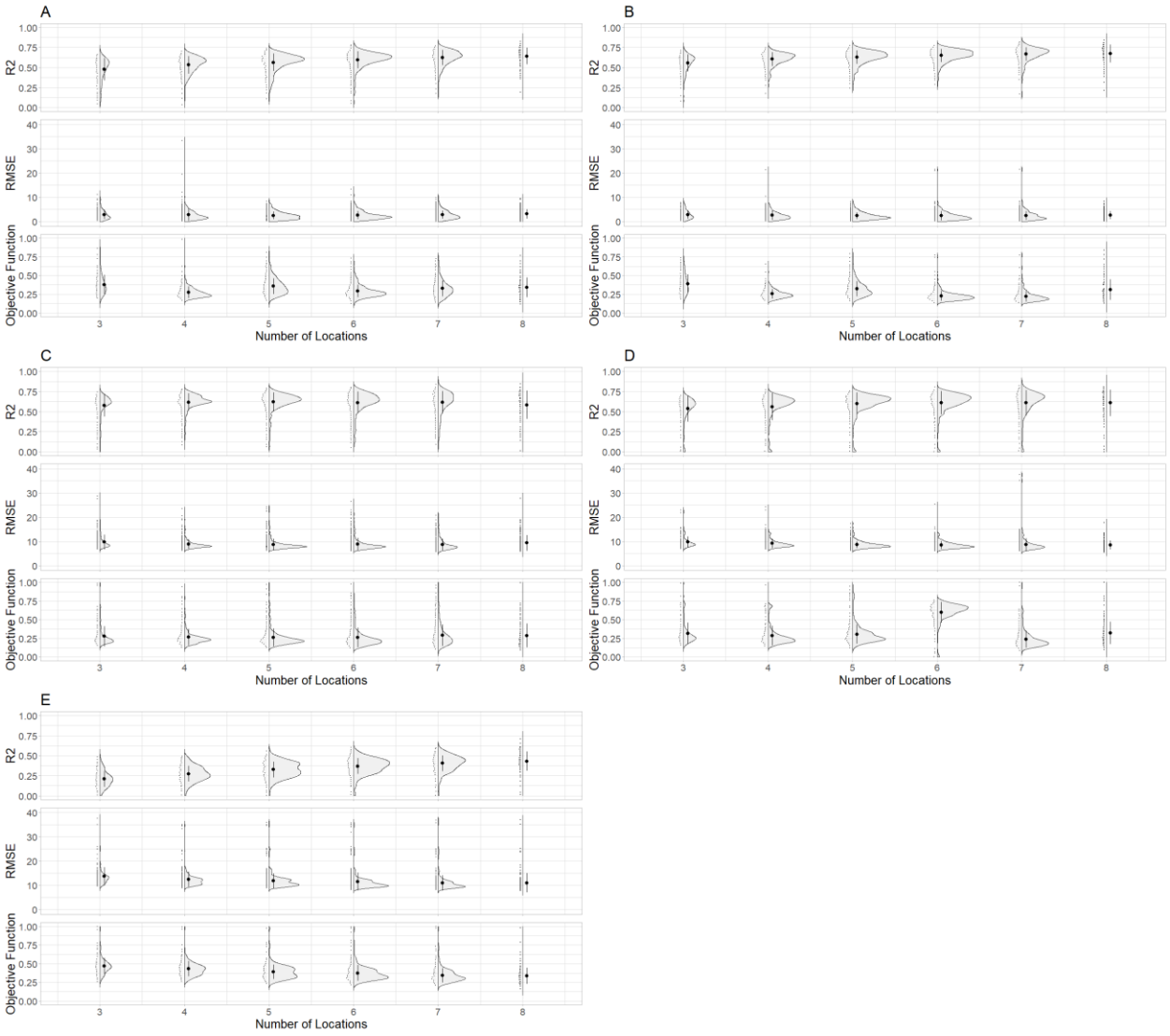


Figure 2.16. The graphs show the R^2, RMSE, or OF values for the 5 structures (A = model 1, B = model 2, C = model 3, D = model 4, E = model 5), when increasing the number of locations included for the training (from 3 to 8). The distribution of the obtained models is not constant. The locations used for the training are selected as possible combinations of numbers from 1 to 11. As we increase the locations used for the training, the number of available locations remaining for testing decreases. Therefore, the number of models increases when going from 3 to 5 training locations and decreases when going from 6 to 8.

The majority of the models had the highest R^2 values around 0.8 and RMSE values around 10%. The R^2 and RMSE values associated with the predictions on the three sites for each of the models are presented in table 2.1.

Table 2.1. R^2, RMSE, and OF values when predicting soil textures for the three sites using the model with the lowest objective function for each of the five structures investigated in the study. In model 5, all the soil cores collected at site 2 where used for the training. For this reason, it was not possible to calculate the performance of the model on that site.

	Predictions for site 1			Predictions for site 2			Predictions for site 3		
MODEL	R^2	RMSE	OF	R^2	RMSE	OF	R^2	RMSE	OF
1	0.78	0.16	0.11	0.01	19.09	0.98	0.82	4.84	0.21
2	0.76	0.18	0.12	0.01	19.69	1.00	0.91	3.53	0.13
3	0.84	6.37	0.24	0.02	17.98	0.95	0.98	3.11	0.09
4	0.82	5.39	0.23	0.02	18.27	0.95	0.91	4.88	0.17
5	0.71	8.28	0.35	-	-	-	0.86	5.69	0.21

As suggested by the strong correlation between depth and distribution of sand, silt, and clay, even model 1 showed good performances, both on and off site. The R^2 and RMSE values obtained from model 1 were compared to predictions based on a simple regression (not presented here). Additionally, soil textures were estimated through the use of different types of regression (linear and multiple multivariate regression) to confirm the need of a more complex tool like ANN, but the results from the regressions will not be presented in this paper. The ANN provided better predictions and had an in-built constraint to keep the sum of the three outputs equal to 100%, a detail that was not possible to obtain through the simple linear regression. Probably also because of this strong correlation, none of the models was able to predict the textures of the fill area at site 2. Nevertheless, the models were able to accurately estimate the

textures at site 3, suggesting that they could be applied off site in a non-fill region within the Piedmont Region of Georgia.

To determine the model with the best performance, I recalculated the OF, taking in consideration only the models presented in table 2.1, and then averaged the OF for each structure (using only the data from sites 1 and 3, since model 5 is missing the predictions on site 2). I found that model 2 had the best overall performance considering the calibration and validation results together, so I used it to calculate the soil textural classes on the entire hillslopes at both sites 1 and 3 (Figures 2.17 and 2.18 respectively).

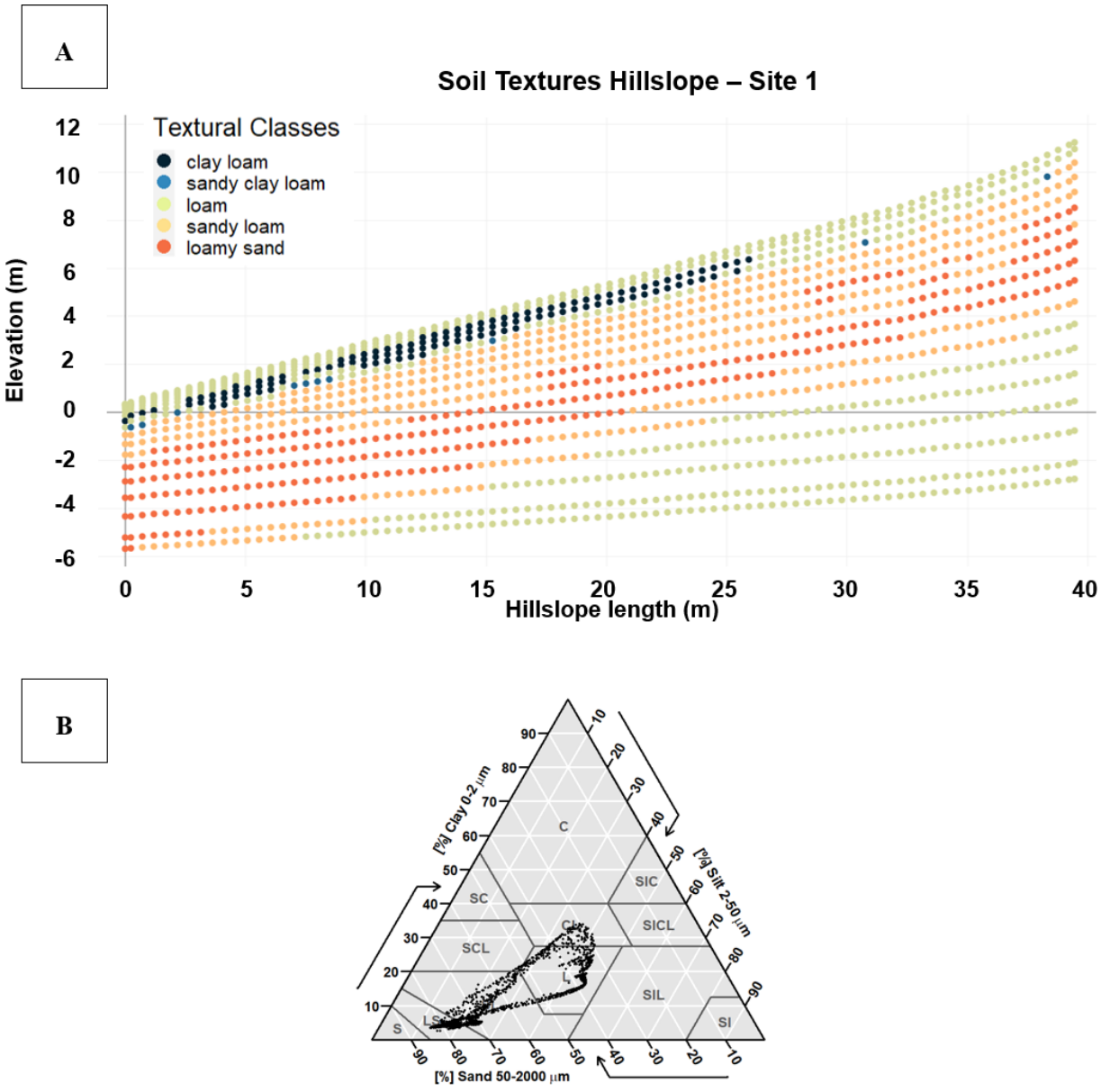


Figure 2. 17. Estimated soil textures for the entire hillslope at site 1 based off model 2. The predictions were calculated for all six months and then averaged, before using them to obtain the textural classes. The textures present are loam, clay loam, sandy clay loam, loamy sand, and sandy loam. Figure A displays the textures with depth, following the actual topography of the site, while figure B is a simple plot of the results on the textural triangle.

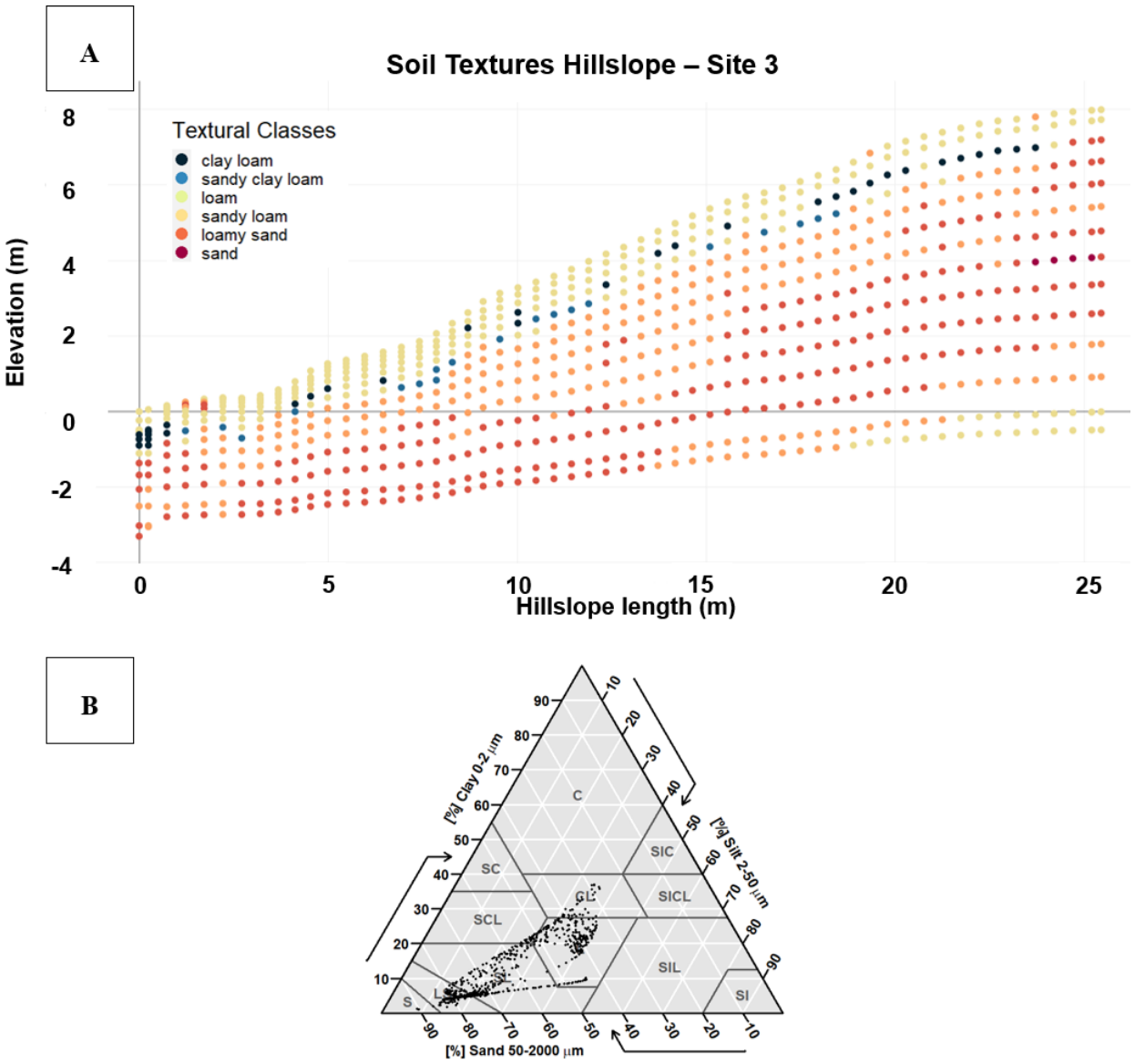


Figure 2. 18. Estimated soil textures for the entire hillslope at site 3 based off model 2. The textures present are loam, clay loam, sandy clay loam, loamy sand, and sandy loam. Figure A displays the textures with depth, following the actual topography of the site, while figure B is a simple plot of the results on the textural triangle.

One of the main limitations of the models, evident in the plots, is the creation of point predictions across the surveyed area as opposed to have continuous estimations. The location of the point predictions is similar to the grid used for the inversion process of the ERT data. The grids usually get coarser with depth. Even though some software allows the customization of the grid, ERT data are innately less

accurate in their measurements at deeper locations, because of the loss of strength in the signal. This principle should be kept in mind if the inversion grid is changed.

Finally, it is important to consider that the soil samples used for this study (both in the models training process and validation) were collected at 30 cm increments, without any consideration of the soil horizons. This procedure might have caused, at the interface of different layers, an artificial mixing of soils and a possible effect in the consequently measured textures.

Conclusion

This study aimed to investigate the ability of near-surface geophysical methods, specifically electrical resistivity tomography, to predict soil textures for the Piedmont Physiographic Region when imported in an Artificial Neural Network framework. I concluded that even though resistivity and rain did not have a very strong correlation with percentages of sand, silt, and clay, when included in an ANN framework with depth, they were able to improve the accuracy of the models. The models, based on a single site, were able to estimate the texture at a different location with natural landscape, but could not be used in areas with fill material. The predictions can be based on a single ERT survey, as shown for site 3, suggesting that seasonality does not have a major effect. The models can provide a finer level of information than what is currently available from SSURGO, also covering deeper portions of the subsurface. It would be valuable to repeat the validation of the model to an even broader area of the Piedmont, to evaluate the extent of its ability. Finally, the model could be refined by including information about features coarser than sand, that for this study were not taken into consideration. These data could improve the models mainly in the predictions within the saprolite layer and could give important insights on the porosity associated with the layers and not just their texture.

References

Banton, O., Cimon, M.-A., & Seguin, M.-K. (1997). Mapping Field-Scale Physical Properties of Soil with Electrical Resistivity. *Soil Science Society of America Journal, 61*(4), 1010-1017. doi:10.2136/sssaj1997.03615995006100040003x

Camporese, M., Cassiani, G., Deiana, R., Salandin, P., & Binley, A. (2015). Coupled and uncoupled hydrogeophysical inversions using ensemble Kalman filter assimilation of ERT-monitored tracer test data. *Water Resources Research, 51*(5), 3277-3291. doi:10.1002/2014wr016017

Cassiani, G., Bruno, V., Villa, A., Fusi, N., & Binley, A. M. (2006). A saline trace test monitored via time-lapse surface electrical resistivity tomography. *Journal of Applied Geophysics, 59*(3), 244-259. doi:10.1016/j.jappgeo.2005.10.007

Ceolin, S. (2019). *Preliminary Assessment of Onsite Wastewater Treatment System Effects on Lake Lanier Water Quality*. (Master of Science of Crop and Soil Sciences). University of Georgia and University of Padova,

Deiana, R., Cassiani, G., Villa, A., Bagliani, A., & Bruno, V. (2008). Calibration of a Vadose Zone Model Using Water Injection Monitored by GPR and Electrical Resistance Tomography. *Vadose Zone Journal, 7*(1), 215. doi:10.2136/vzj2006.0137

French, H., & Binley, A. (2004). Snowmelt infiltration: monitoring temporal and spatial variability using time-lapse electrical resistivity. *Journal of Hydrology, 297*(1-4), 174-186. doi:10.1016/j.jhydrol.2004.04.005

Grunwald, S., Thompson, J. A., & Boettinger, J. L. (2011). Digital Soil Mapping and Modeling at Continental Scales: Finding Solutions for Global Issues. *Soil Science Society of America Journal, 75*(4), 1201-1213. doi:10.2136/sssaj2011.0025

Hartemink, A. E., Krasilnikov, P., & Bockheim, J. G. (2013). Soil maps of the world. *Geoderma, 207-208*, 256-267. doi:https://doi.org/10.1016/j.geoderma.2013.05.003

Hazreek, Z. A. M., Saad, R., Wijeyesekera, D., Ahmad, F., Ahmad Tajudin, S. A., Madun, A., & Hj. Bakar, I. (2015). Soil Identification using Field Electrical Resistivity Method. *Journal of Physics Conference Series, 622*, 1-7. doi:10.1088/1742-6596/622/1/012030

Heath, R. (1980). Basic elements of ground-water hydrology with reference to conditions in North Carolina: US Geological Survey Water-Resources Investigations Open-File Report 80-44, 86 p. 1984. *Ground-water regions of the United States: US Geological Survey Water-Supply Paper, 2242*(78), 1989.

Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B. M., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J. G. B., Walsh, M. G., & Gonzalez, M. R. (2014). SoilGrids1km--global soil information based on automated mapping. *PloS one, 9*(8), e105992-e105992. doi:10.1371/journal.pone.0105992

Hinnell, A. C., Ferré, T. P. A., Vrugt, J. A., Huisman, J. A., Moysey, S., Rings, J., & Kowalsky, M. B. (2010). Improved extraction of hydrologic information from geophysical data through coupled hydrogeophysical inversion. *Water Resources Research, 46*(4). doi:10.1029/2008wr007060

Homer, C. H., Fry, J. A., & C.A., B. (2012). The National Land Cover Database. Retrieved 11/25/20, from U.S. Geological Survey Fact Sheet 2012-3020 https://www.mrlc.gov/

Hubbard, S. S., Rubin, Y., & Majer, E. (1999). Spatial correlation structure estimation using geophysical and hydrogeological data. *Water Resources Research, 35*(6), 1809-1825. doi:10.1029/1999wr900040

Irfan, F., & Syed, B. (2012). Integrating Geo-Electrical and Geotechnical Data for Soil Characterization. *International Journal of Applied Physics and Mathematics, 2*, 104-106. doi:10.7763/IJAPM.2012.V2.63

Keesstra, S. D., Bouma, J., Wallinga, J., Tittonell, P., Smith, P., Cerdà, A., Montanarella, L., Quinton, J. N., Pachepsky, Y., van der Putten, W. H., Bardgett, R. D., Moolenaar, S., Mol, G., Jansen, B., & Fresco, L. O. (2016). The significance of soils and soil science towards realization of the United Nations Sustainable Development Goals. *SOIL, 2*(2), 111-128. doi:10.5194/soil-2-111-2016

Kemna, A., Vanderborght, J., Kulessa, B., & Vereecken, H. (2002). Imaging and characterisation of subsurface solute transport using electrical resistivity tomography (ERT) and equivalent transport models. *Journal of Hydrology, 267*(3-4), 125-146. doi:10.1016/s0022-1694(02)00145-2

Kotsiantis, S. B., Zaharakis, I. D., & Pintelas, P. E. (2006). Machine learning: a review of classification and combining techniques. *Artificial Intelligence Review, 26*(3), 159-190. doi:10.1007/s10462-007-9052-3

Kouchaki, B. M. (2017). *Laboratory Resistivity Measurements for Soil Characterization.* University of Arkansas, Fayetteville, Retrieved from http://scholarworks.uark.edu/etd/2455

LeGrand, H. E., Back, W., Rosenshein, J. S., & Seaber, P. R. (1988). Region 21, Piedmont and Blue Ridge. In *Hydrogeology* (Vol. O-2, pp. 201-208): Geological Society of America.

Maglogiannis, I. G. (2007). *Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in EHealth, HCI, Information Retrieval and Pervasive Technologies*: IOS Press.

Miller, C. R., Routh, P. S., Brosten, T. R., & McNamara, J. P. (2008). *Application of Time-Lapse ERT Imaging to Watershed Characterization*. Department of Geosciences. Boise State University. Geosciences Faculty Publications and Presentations.

Miller, J. A. (1990). *Ground Water Atlas of the United States: Segment 6, Alabama, Florida, Georgia, South Carolina* (730G). Retrieved from http://pubs.er.usgs.gov/publication/ha730G

Milsom, J., & Ager, E. (2011). *Field Geophysics* (L. John Wiley and Sons, Publication Ed. Fourth ed.).

Moeys, J. (2018). The soil texture wizard: R functions for plotting, classifying, transforming and exploring soil texture data. *CRAN. R-Project*.

Munsell, C. (2010). *Munsell soil color charts : with genuine Munsell color chips*: 2009 year revised. Grand Rapids, MI : Munsell Color, 2010.

Negnevitsky, M. (2005). *Artificial Intelligence a Guide to Intelligent Systems* (P. E. Limited Ed. Second ed.).

Ohta, T., & Arai, H. (2007). Statistical empirical index of chemical weathering in igneous rocks: A new tool for evaluating the degree of weathering. *Chemical Geology, 240*(3-4), 280-297. doi:10.1016/j.chemgeo.2007.02.017

Omuto, C., Nachtergaele, F., & Rojas, R. V. (2013). *State of the Art Report on Global and regional Soil Information: Where are we? Where to go?* : Food and Agriculture Organization of the United Nations Rome.

Pavich, M. J., Leo, G., Obermeier, S., & Estabrook, J. (1989). *Investigations of the characteristics, origin, and residence time of the upland residual mantle of the Piedmont of Fairfax County, Virginia* (2330-7102). Retrieved from

Piegari, E., & Di Maio, R. (2013). Estimating soil suction from electrical resistivity. *Natural Hazards and Earth System Sciences, 13*(9), 2369.

Reynolds, J. M. (2011a). An Introduction to Applied and Environmental Geophysics. In J. W. S. Incorporated (Ed.), (2 ed., pp. 289 - 347).

Samouëlian, A., Cousin, I., Tabbagh, A., Bruand, A., & Richard, G. (2005). Electrical resistivity survey in soil science: a review. *Soil and Tillage Research, 83*(2), 173-193. doi:10.1016/j.still.2004.10.004

Sanchez, P. A., Ahamed, S., Carré, F., Hartemink, A. E., Hempel, J., Huising, J., Lagacherie, P., McBratney, A. B., McKenzie, N. J., Mendonça-Santos, M. d. L., Minasny, B., Montanarella, L., Okoth, P., Palm, C. A., Sachs, J. D., Shepherd, K. D., Vågen, T.-G., Vanlauwe, B., Walsh, M. G., Winowiecki, L. A., & Zhang, G.-L. (2009). Digital Soil Map of the World. *Science, 325*(5941), 680. doi:10.1126/science.1175084

Sandberg, S. K., Slater, L. D., & Versteeg, R. (2002). An integrated geophysical investigation of the hydrogeology of an anisotropic unconfined aquifer. *Journal of Hydrology, 267*(3-4), 227-243. doi:10.1016/s0022-1694(02)00153-1

Santarato, G., Zeid, N. A., & Bignardi, S. (2015). *Lezioni di geofisica applicata* (libreriauniversitaria.it Ed.).

Siddiqui, F. I., & Osman, S. B. A. B. S. (2013). Simple and multiple regression models for relationship between electrical resistivity and various soil properties for soil characterization. *Environmental Earth Sciences, 70*(1), 259-267.

Siddiqui, F. I., Pathan, D. M., Osman, S. B. A. B. S., Pinjaro, M. A., & Memon, S. (2015). Comparison between regression and ANN models for relationship of soil properties and electrical resistivity. *Arabian Journal of Geosciences, 8*(8), 6145-6155.

Singha, K., Day-Lewis, F. D., Johnson, T., & Slater, L. D. (2015). Advances in interpretation of subsurface processes with time-lapse electrical imaging. *29*(6), 1549-1576. doi:10.1002/hyp.10280

Soil Science Division Staff. (2017). *Soil survey manual. C. Ditzler, K. Scheffe, and H.C. Monger (eds.). USDA Handbook 18. Government Printing Office, Washington, D.C.*

Soil Survey Staff, Natural Resources Conservation Service, & Agriculture., U. S. D. o. (2020). Official Soil Series Descriptions. Retrieved from https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/geo/?cid=nrcs142p2_053587

Soil Survey Staff, N. R. C. S., United States Department of Agriculture. (2019). Web Soil Survey. Retrieved from http://websoilsurvey.sc.egov.usda.gov/.

Son, Y., Oh, M., & Lee, S. (2010). Estimation of soil weathering degree using electrical resistivity. *Environmental Earth Sciences, 59*(6), 1319-1326.

The MathWorks Inc. (2019). MATLAB.

United States Department of Agriculture, Natural Resources Conservation Service, & Survey, N. C. S. (1995). *Soil Survey Geographic (SSURGO) Data Base - Data Use Information*. Miscellaneous Publication

Vereecken, H., Binley, A., Cassiani, G., Revil, A., & Titov, K. (2006). Applied Hydrogeophysics. In (pp. 1-8): Springer Netherlands.

Xu, R. (2002). *Particle Characterization : Light Scattering Methods*. New York: Springer.

Zauyah, S., Schaefer, C. E. G. R., & Simas, F. N. B. (2018). Chapter 3 - Saprolites. In G. Stoops, V. Marcelino, & F. Mees (Eds.), *Interpretation of Micromorphological Features of Soils and Regoliths (Second Edition)* (pp. 37-57): Elsevier.

Zeng, R. Q., Meng, X. M., Zhang, F. Y., Wang, S. Y., Cui, Z. J., Zhang, M. S., Zhang, Y., & Chen, G. (2016). Characterizing hydrological processes on loess slopes using electrical resistivity tomography – A case study of the Heifangtai Terrace, Northwest China. *Journal of Hydrology, 541*, 742-753. doi:10.1016/j.jhydrol.2016.07.033

CHAPTER 3

DETERMINING THE EFFECT OF DOMAIN COMPLEXITY FOR SOIL LAYERS AND WEATHERING STATE ON WATER AND CHLORIDE TRANSPORT AT THE HILLSLOPE SCALE[2]

[2] Tancredi, M.T. and N. Gaur. To be submitted to *Journal of Geophysical Research: Earth Surface*.

Abstract

Simulation models are always faced with a choice between using parsimonious quantitative structures and those with a high level of spatial and temporal detail. Parsimonious models are preferred since they are easier to calibrate while still accomplishing a desired level of prediction. However, such models may be unable to predict accurately under complex conditions resulting from the interaction between certain boundary inputs and site-specific domain heterogeneity. Soil wetness, topography, and soil properties have been identified as the main factors influencing hillslope heterogeneity. The peculiar geology of the Piedmont Physiographic Region of Georgia presents two main challenges related to topography and soil properties: the presence of a layer of saprolite with variable thickness and weathering rates and the rapidly changing depth to bedrock. In this study, I compared the hydrological and chloride transport predictions for a variably saturated hillslope using two models with varying domain complexity. The first model domain follows a soil horizon-based division of the sub-surface while the second model has a detailed soil texture classification based on Electrical Resistance Tomography of the region. The main differences exist in the number and textural classification of layers in the saprolite and the depth to bedrock. The modeled hillslope receives a constant sub-surface flux input of solute (akin to a septic system) and feeds into a water body. Both models accurately predict water table elevations (R^2 values of 0.97 and 0.96), and the primary differences being attributed to different transmissivity values for the models (related to the depth to bedrock). For chloride transport, the parsimonious model has an overall better performance (R^2=0.63) than the complex model (R^2=0.52). In both cases, anyway, the solute concentrations are underestimated, with an overall RMSE=14.06 mg/L for the parsimonious model and RMSE=23.48 mg/L for the complex model. The study shows that, when using a single porosity transport model, increased complexity for soil layering and weathering state does not show the occurrence of funnel-type preferential flow in the saprolite. Transmissivity is the most important characteristic when simulating water movement. Revised chloride input values, together with a dual porosity conceptual model based on macropore preferential flow might be necessary to better capture solute movement in the area.

Introduction

Quantitative, processed based models for soil hydrologic flow prediction continue to grow in spatial and temporal complexity with computational power. The number of parameters included in a model, however, should reflect the availability of data that can constrain it, rather than the computational power. The choice between parsimonious models and models with a high level of site-specific should be based on the "principle of parsimony". Box and Jenkins (1976) defined it as "the smallest possible number of parameters for adequate representation". This principle is particularly important because past studies showed how the performances of a model can plateau, regardless of the increasing level of complexity, but that increasing the number of parameters can increase the level of uncertainty, as the uncertainty of the parameters spreads to the predictions (Wagener, Lees et al., 2001).

In hydrologic models it is possible to modulate the level of complexity. The number of parameters will change when: 1) changing the number of dimensions (1D, 2D, 3D), 2) going from a micropore to a hillslope or watershed scale, 3) changing the conceptual model upon which calculations are based, and 4) varying the types and number of soil layers specified in the study. All these factors can affect the outcomes of the simulations and, since they are connected, the selection of one aspect will condition the others. In my study, for example, I first determined the scale, and then I selected the level of complexity for the other factors, as my objective was to study water and solute movement at the hillslope scale in the Piedmont Region of Georgia.

Hillslopes are pivotal landscape elements that organize water availability in a watershed (Fan, Clark et al., 2019). There has been substantial research to better characterize the processes determining water transport at the hillslope scale (Graham, Van Verseveld et al., 2010; Haught & Meerveld, 2011; Jia, Shao et al., 2013). Even though soil wetness, topography, and soil properties have been identified as the main factors influencing hillslope connections (McGuire & McDonnell, 2010), there still is no unique answer on how those factors should be conceptualized and included in models. Due to the importance of topography and soil properties and their characteristics in my study area, I evaluate both water and solute movement, and perform an analysis on the effect of their domain complexity.

The peculiar geology of the Piedmont Physiographic Region of Georgia presents two main challenges related to topography and soil properties: the presence of a layer of saprolite with variable thickness and weathering rates, and the rapidly changing depth to bedrock (LeGrand, Back et al., 1988; Miller, 1990; Pavich, 1996). The definition of saprolite, so far, has focused mainly on the material's origin and physical appearance than on its properties (Amoozegar, Schoeneberger et al., 1991). Saprolite is weathered bedrock that retains the structure of the parent material (Zauyah, Schaefer et al., 2018). It usually has a higher clay content than the parent rock and is characterized as very friable and easy to excavate (Pavich, 1996). Thickness and weathering rates of the saprolite are strictly linked to the parent material (Pavich, Leo et al., 1989). In the Piedmont region, the bedrock is mainly constituted by igneous and metamorphic rocks. Rocks with diverse characteristics (i.e. age and type) can be adjoining due to tectonic transformations. In general, two predominant rock types can be identified according to their composition, the first resembling granite (i.e. granite, granite gneiss, mica schist, slate, and rhyolite) and the second diorite (i.e. diorite, gabbro, hornblende gneiss, and andesite). The two rock types present distinct weathering profiles (Ohta & Arai, 2007). The presence of these structures and the tectonic transformations that occurred in the area, cause high variability of the depth to bedrock, ranging from completely exposed to buried for 40 m under a layer of regolith (soil and saprolite together) (LeGrand, Back et al., 1988). On average, the saprolite layer in the eastern Georgia Piedmont extends for 15 m (Schroeder & West, 2005). In the southern coves of Lake Sidney Lanier, within 2.5 km of my study site, depth to bedrock ranged from 7 to 69 m (Figure 3) (Drilling boring logs, personal communication with US Army Corps of Engineers).

Furthermore, the bedrock comprises a network of fractures that, together with the regolith (saprolite and residual soils laying on top of it), create a dual-media system for recharge and discharge (LeGrand, Back et al., 1988). In the Piedmont, it is possible to observe preferential flow due to the different structures within the saprolite and at the interface with the bedrock (Pavich, Leo et al., 1989).

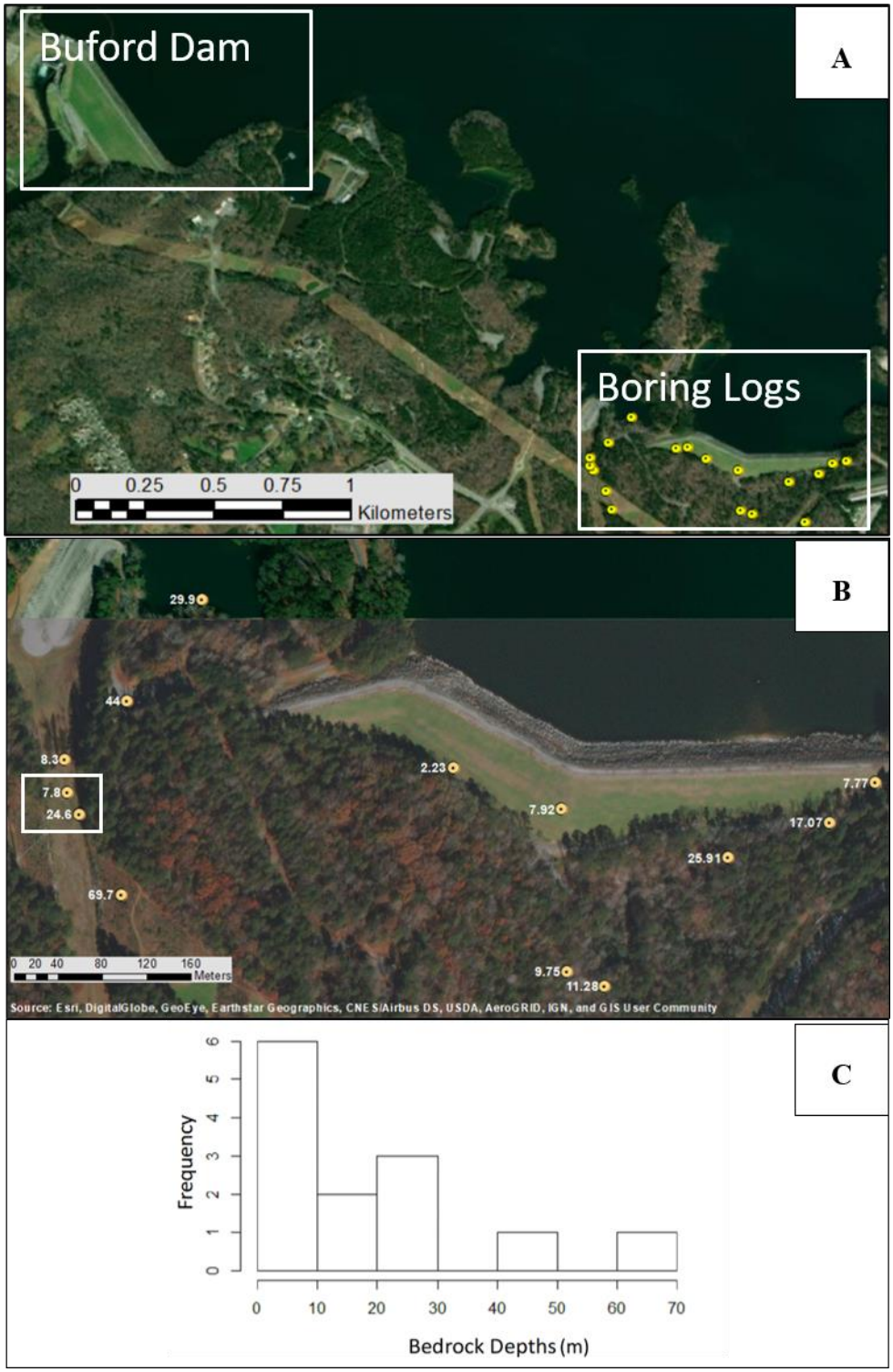


Figure 3.1. Image A shows the locations of the boring logs in respect to Buford Dam, in Buford, GA. B is a blow-up of the area containing the boring logs and shows the depth to bedrock, in meters, as reported by the Us Army Corps of Engineers. In the highlighted box in B, depth to bedrock changes by 17 m with a 10 m distance. Image C presents the same data as a summary of the frequencies of the depths to bedrock.

As a consequence of the variability of the subsurface in my study area, I evaluated simulations of the same hillslope with two structures: a parsimonious model and a more complex one. The parsimonious model followed a horizon-based characterization of the subsurface with the top of the bedrock starting close to the surface. Conversely, the more complex model had a finer description of the soil layers and their weathering state, and the saprolite layer extending deeper into the subsurface. With the complex model, I also investigated the presence of funnel-type preferential flow in the saprolite layer. This type of flow usually results from textural discontinuity in inclined soil layers, specifically when the overlapping layers present very different porosities (Ju & Kung, 1993). For funnel flow to occur, it is necessary to have a fine textured layer on top of a coarser one, and both layers need to be at an angle (Walter, Kim et al., 2000). Because of the difference in porosity, the pressure head of the wetting front in the top layer is less than the water-entry potential, thus not enabling water to flow from the top layer into the bottom one. In this case, since the layers are in a slope, water starts moving laterally at the interface of the two layers, before being able to move again downwards (Kung, 1990a, 1990b). In the hills of the Piedmont region, funnel flow could develop at the interface between the Bt horizon and the saprolite layer, presenting a coarse sandy-like texture, or within the saprolite layer itself, because of the effect of the different weathering state. As it will be presented in the Materials and Methods section, the area of my study site presented some features in the saprolite that could suggest the occurrence of funnel-type preferential flow.

Materials and methods

Study area and data collection

For this study, I selected a house located in one of the southern coves of Lake Sidney Lanier and a small nearby island without residential structures. The island served as the control site to estimate the natural background concentrations of chloride in groundwater. The lake is a man-made reservoir situated about 48 km northeast of Atlanta. The US Army Corps of Engineers, which currently own the water body, created it in 1956 through the construction of the Buford Dam. The house I selected was located at the top of a hill and had an onsite wastewater treatment system (OWTS). The property is defined as

"developed, open space" (Homer, Fry et al., 2012) and, according to the SSURGO database, is characterized by two soil types, Pacolet sandy loam, 6 to 10 percent slopes, moderately eroded and Pacolet-Saw complex, 15 to 45 percent slopes, stony. The first occupying the area between the top of the ridge and the house and the latter characterizing the rest of the hillslope. The drainfield extends in the backyard and ends at about 50 m from the lakeshore (Figure 3.2). The location of the site will not be specified any further to protect the privacy of the homeowners.

Figure 3.2. Aerial photo of the modeled site. The yellow square outlines the approximate borders of the property parcel extended to the middle of the lake cove. The light blue square shows the location of the drainfield, adjacent to the house (orange square).

At both locations, we installed monitoring wells to gather information on water table elevations and collect water samples to measure chloride concentration in groundwater. At the modeled site, we installed 18 wells along the shoreline, alternating shallow and deep wells, two wells (W19 and W20) at 6 m from the shore, five on the rest of the hillslope, and one in the front yard as a local control. At the control site, we had nine wells distributed around the island (Figures 3.3A and 3.3B).

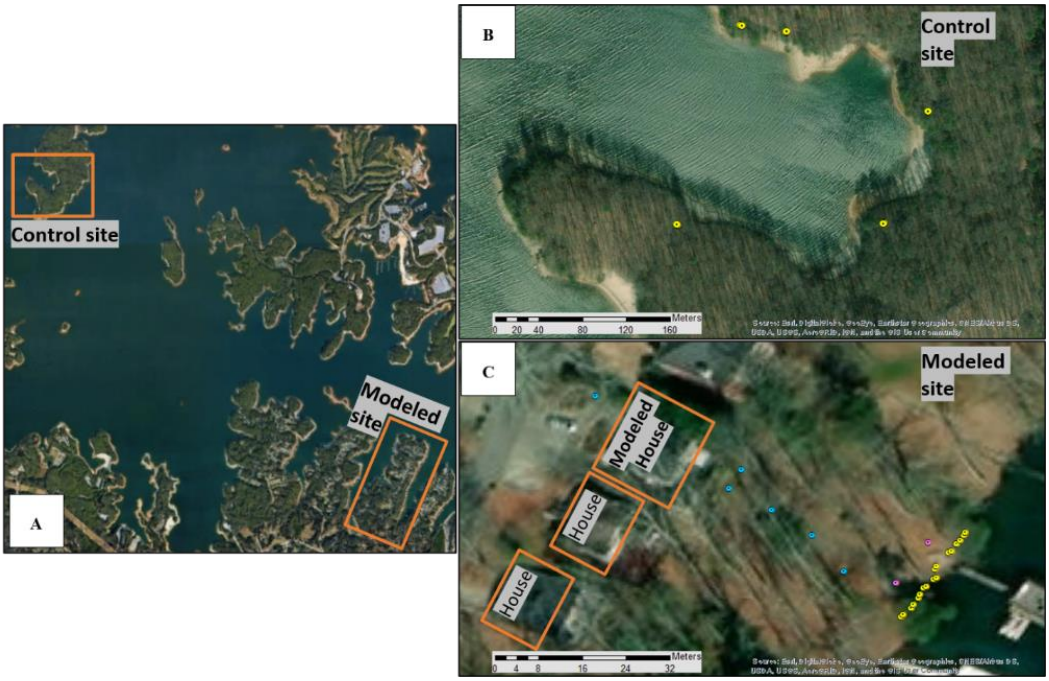


Figure 3.3. A shows the locations of the control site (1) and the modeled site (2). B and C show the locations of the monitoring wells (circles with a central black dot) at the control (B) and the modeled (C) sites. At the modeled site we have a line of shoreline wells, yellow, with alternating shallow and deep ones, two wells (W19 and W20) at 6 meters from the shore, pink, and six upslope wells, blue. The well in the front yard serves as a local control as it is outside the area of influence of the OWTS.

Each monitoring well has a diameter of 5.1 cm and is composed of a circular valve box cover, a lock top, a PVC riser (that allows access to the desired depth), and a 1.5 m long PVC screen (Ceolin, 2019). The installation process took place in September and October 2018 at the modeled site, and between August and September 2019 at the island. During well installation, we augered 11 cm diameter holes to different depths according to the location. At the shoreline, the maximum depth of the installed wells was about 1.5 m below the water table at the time of installation for shallow wells and 3 m below the water table for deep wells. For the upslope wells, the depth ranged between 2.10 m and 2.40 m below the water table. Where we did not reach the water table at the time of the installation, we augered until hitting refusal. For this work, the term refusal indicates a surface that is not penetrated after 5 minutes of continuous manual or mechanical augering. A mechanical auger by Little Beaver, Inc. was used for all well installations while some soil samples were collected at intermediate locations by hand augering (Ceolin, 2019). The maximum depth was 7.30 m, as limited by the mechanical auger. We placed the PVC

case in the center of the boring and then added sand around the PVC case for the entire length of the screen. The sand is used to reduce clogging of the screen and to allow clean water flux into the well. Finally, we added a mixture of bentonite, sand, and excavated soil to fill the hole almost up to the top. A well schematic and a picture of an installed one are shown in Figures 3.4A and 3.4B.

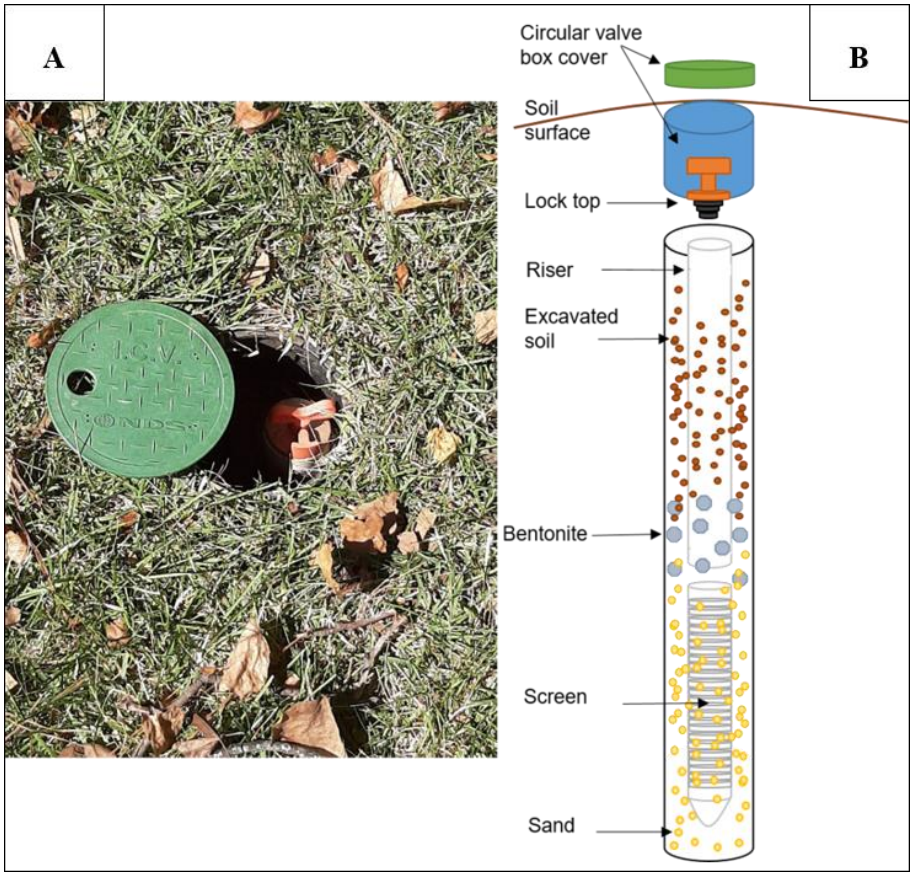


Figure 3.4. Figure A is an example of an installed well, while B is the well schematics.

At both locations, we measured water table elevation and collected water samples monthly, except during March and April 2020 due to the Covid-19 restrictions. More details on sampling procedures and analysis can be found in Ceolin (2019). The monthly water table elevations and chloride concentrations were the observed values used to calibrate the simulations. At the modeled site, we also collected soil moisture values (using the ML3 ThetaProbe Soil Moisture Sensor by Delta-T Devices Ltd), in the first 10 cm of soil. Soil moisture was measured monthly for 11 months, from February 2019 to

February 2020. The soil moisture values were collected at every meter, between the house and the shoreline, and averaged to represent the water content in five sections of the hillslope. I divided the hillslope in five sections so to capture water content near the lake (at 2 m from the shoreline, to ensure that the area would never go under water), in the middle parts of the slope (at 9, 19, and 29 m from the shoreline) and near the drainfield (at 37 m from the shoreline). These measurements were also included in the calibration process.

In addition, I performed a field-based estimation of saturated hydraulic conductivity (*Ks*) to have an initial estimate for my calibration process. Ks values were collected at three different depths using separate methods. The depths were based on a soil survey of the site and were selected to have an estimate of the saturated hydraulic conductivity of the topsoil, the subsoil, and the saprolite. For the surface, we used a single ring infiltrometer (SATURO) by METER Group, Inc. with a 5 cm deep insertion ring in the soil. For the subsoil, we excavated soil in a core up to 60 cm, the top of the clayey layer, to place a borehole permeameter (Aardvark Permeameter by Soilmoisture Equipment Corp). The measurements for the first two layers were collected on two replicates at three locations (top of the slope, middle of the slope, and foot of the slope). For the saprolite layer, we performed an aquifer test using the deep wells installed at the shoreline. We deployed a peristaltic pump to generate a drop in the water table and consequent water movement. The changes were measured through pressure transducers by METER Group, Inc. A detailed description of the locations, procedures, equations, and results for the three types of measurements is available in Ceolin (2019).

Hillslope simulations

I developed the two simulations using HYDRUS (2D/3D) Version 3.02. The software allows, among other features, to model water and solute movement in two dimensions (Šimůnek, Šejna et al., 2018). The program provides a numerical solution to the Richards (1931) equation (Equation 3.1) for transient water flow:

$$\frac{\partial\theta(h)}{\partial t} = -\frac{\partial}{\partial z}\left(K(h)\frac{\partial h}{\partial z}\right) + \frac{\partial K(h)}{\partial z} - S(h) \qquad [3.1]$$

where the coefficient *K(h)* is the unsaturated hydraulic conductivity and is a function of the dependent variable, either pressure head (*h*) or water content (θ). In my simulation, I selected pressure head as the dependent variable for the equation.

For the solute transport, HYDRUS (2D/3D) uses the Advection Dispersion Equation, ADE, (Equation 3.2):

$$\frac{\partial c}{\partial t} = D_e\frac{d^2c}{dz^2} - v\frac{\partial c}{\partial z} \qquad [3.2]$$

where *c* is the solute concentration, D_e is the effective dispersion coefficient, and *v* is the mean pore water velocity (Radcliffe & Simunek, 2018)**.**

In both simulations, the slope is long 152 m, of which 75 m represents a hill, and 72 m constitutes a lake floor. The topography mimics the real elevations from the top of the hill to the deepest point of the lake cove. Elevations were obtained through contour lines and bathymetry measurements provided by Gwinnett County (Ceolin, 2019).

The upper boundary conditions for water flux vary for the two sections of the slope. The area comprised between the top of the hill and the shoreline is a seepage face, and the lakebed is a variable head (used to simulate lake fluctuations). Water movement to the left of the ridge top and the right of the cove falls outside of the modeled area, and there is no water flux through the deepest part of the bedrock (because of the absence of fractures). I replicated these assumptions by assigning "no flux" boundaries to the three remaining sides of the section (Figure 3.5A). For the solute movement, I kept the same "no flux" boundary conditions but I used a "third-type" boundary condition for the entire slope (Figure 3.5B). The "third-type" boundary condition is mass conservative and is the one recommended by Šimůnek, Van Genuchten et al. (2012). Root water uptake is included in both hillslopes using the Feddes, Bresler et al. (1974) model and selecting parameters for "deciduous fruit" to simulate the presence of trees. I also considered plant passive uptake of solute up to a maximum concentration of 5 mg/L, since that is the

average concentration measured at my control site. Roots extend up to 3 m, and their density decreases with depth.

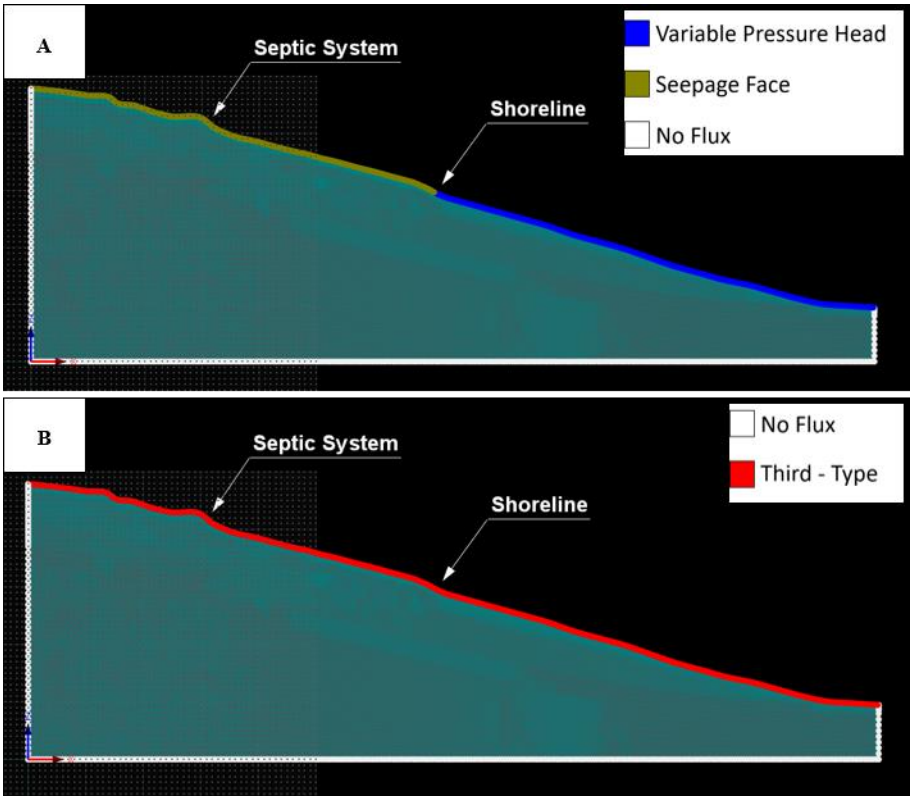


Figure 3.5. Image A shows the flux boundary conditions for both models and image B shows the solute boundary conditions. The arrows pointing at the shoreline and the septic system can be used as a reference to the real site.

The hillslopes receive a constant sub-surface flux input of water and solute from seven points selected for nodal recharge to simulate a septic system (Figure 3.6). The concentration value for the solute input was obtained from a sample taken in the septic tank in May 2019. The water flux entering the system was calculated from water usage data for the household. Other factors influencing water content and flux are daily values of rainfall, transpiration, and lake levels. The lake levels and amounts of rainfall were obtained from the USGS Water Data (waterdata.usgs.gov/nwis), while the evapotranspiration values come from a University of Georgia weather station located in Oakwood (GA) (www.georgiaweather.net). In my model, I express the evapotranspiration values only in terms of transpiration, as I am assuming that the tree canopy shades the surface and that most of the water is leaving the soil through transpiration.

Water movement follows the Van Genuchten (1980) – Mualem (1976) (Equation 3.3) model for single-porosity with no hysteresis.

$$K(S_e) = K_s S_e^l \left[1 - (1 - S_e^{1/m})^m\right]^2 \qquad [3.3]$$

where: S_e is the effective soil water saturation, l is the pore-connectivity parameter, and m is a fitting parameter equal to $1 - 1/n$, $n > 1$, with n being another fitting parameter.

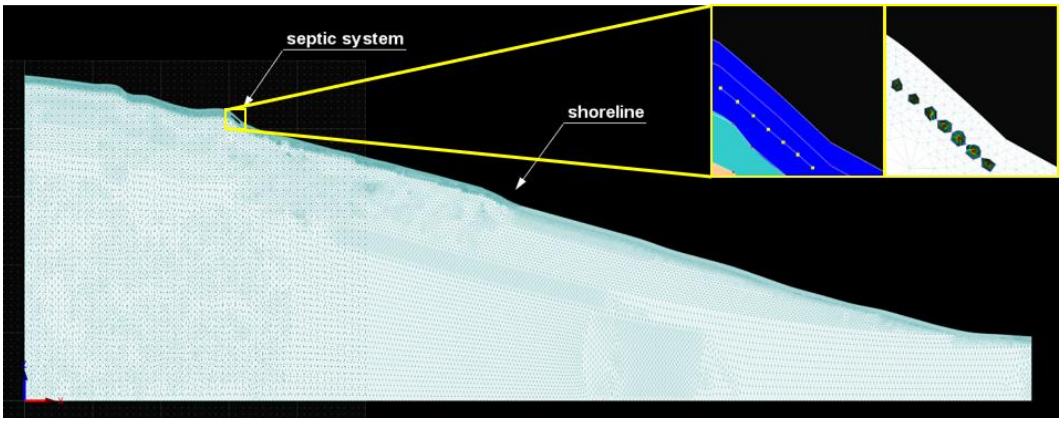


Figure 3.6. The septic drainfield in the model is simulated by the presence of seven recharge nodes. The nodes are located at 0.55 m from the surface. The two zoomed squares show how the nodes appear as "geo-objects" or "on the mesh". Water and chloride are pumped into the system from the nodes at a daily frequency.

I decided to use the single-porosity conceptual model because, even though some big trees are present on my site, I am assuming that there are not enough to generate substantial macropore preferential flow. The first domain represents a traditional soil horizon-based division of the sub-surface. A level 3 soil survey was conducted at my site and found the Pacolet series to be prevalent. The survey identified, in the right portion of the site (facing the lake), an area showing the Saw series. The presence of Saw at the site reaffirmed the variability of depth to bedrock in the area, as this soil series is characterized by shallower bedrock than Pacolet (bedrock starting at -0.74 m from the surface for Saw and deeper than 2 m for Pacolet) (Soil Survey Staff, Natural Resources Conservation Service et al., 2020). The horizon depths were determined based on these results and the domain was divided into five distinct layers, specifically topsoil, subsoil, saprolite, fractured bedrock, and lakebed. The topsoil is 0.3 m deep and has a sandy loam texture, the subsoil extends from -0.3 to -0.9 m and has a clay texture, the saprolite layer stretches from -

0.9 m to an average depth of -5.4 m, and the fractured bedrock covers the remaining part of the section. The top of the bedrock is not uniform. I established its position using the depth at which we reached refusal, either during the installation of piezometers or the collection of soil samples.

The domain of the second model, called “detailed model” or “complex model”, was simulated using information obtained through electrical resistivity tomography (ERT). The main differences from the first simulation are multiple saprolite layers and a deeper bedrock. ERT data were collected between the house and the shoreline for 6 months and included in an Artificial Neural Network framework for the estimation of soil texture. Other inputs to the model were relative depth of investigation and weekly antecedent rainfall. The model used to make predictions had an R^2 value of 0.74 and an RMSE value of 3.04%. For this model, the top part (0 m to about -1.60 m) of the soil profile is characterized by loam, clay loam, and sandy clay loam textures. The clay loam and sandy clay loam constitute the Bt horizon but are different from the parsimonious model as they are not in a homogeneous textural layer. Saprolite starts at -1.60 m and extends for 14 m and is divided into 4 sub-sections. The first section is 2.11 m thick (on average) and has a sandy loam texture, the second section is 2.16 m deep and has a loamy sand texture. Starting at -6.68 m I have again a sandy loam texture and finally, starting at -10.82 m to the top of the bedrock I estimate a silt loam texture. Additionally, the upper two layers of the saprolite show three features which mirror areas of the subsurface that, during the ERT survey, were showing higher resistivity values than the surrounding materials (Figure 3.7).

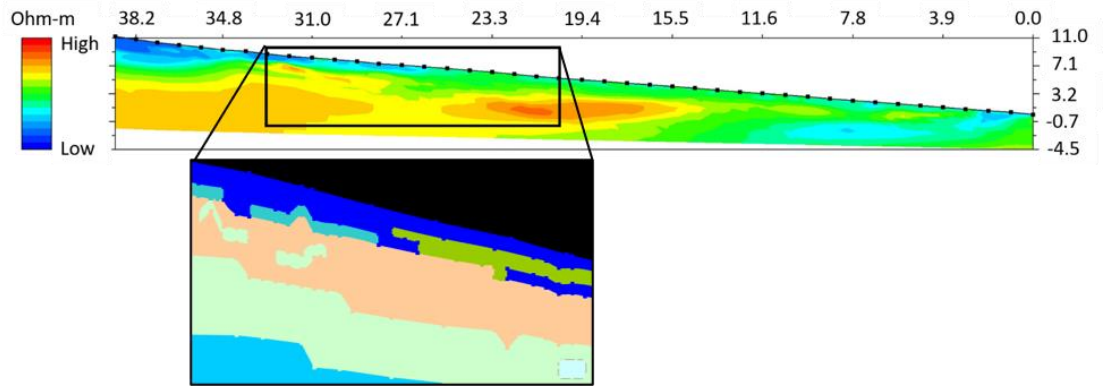


Figure 3.7. The top image is an example of the electrical resistivity data collected at the study site. The second image shows how high resistivity features (inside the black box) are replicated in the simulation as areas with a different texture than the surrounding surface.

I attempted to estimate the depth to bedrock using ERT data. The deepest point in my surveys was -14 m and the bedrock was not detected. Based on the average depth of bedrock typically observed I assumed fractured bedrock to start where the survey depth ended. The two domain structures are presented in Figure 3.8 A and B

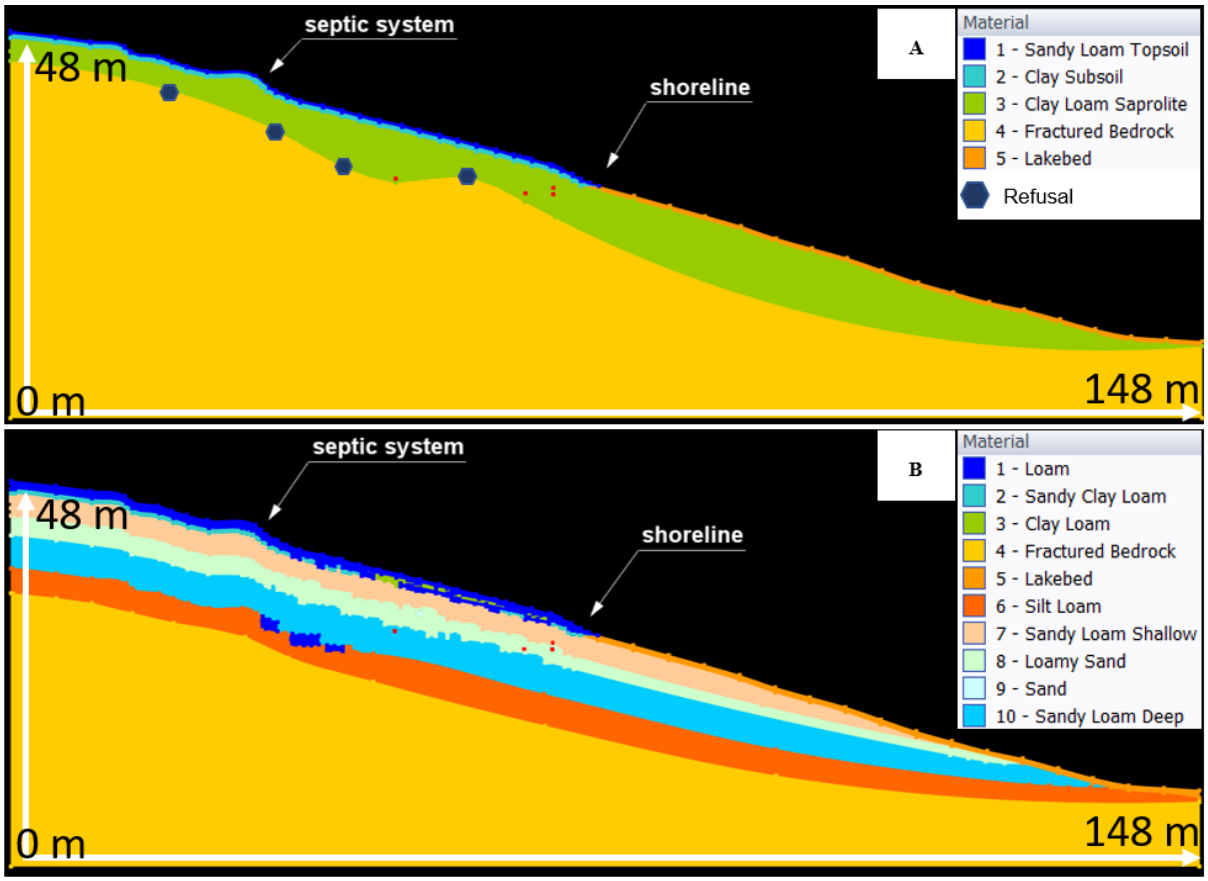


Figure 3.8. A shows the domain properties, in terms of materials, for the parsimonious model. The dark hexagons point out the locations where we hit refusal during the installation of the monitoring wells or the collection of the soil samples. The locations were used to identify the top of the bedrock. B presents the materials of the detailed model. In this case, the saprolite is constituted by three materials divided into four layers (from the top, shallow sandy loam, loamy sand, deep sandy loam, and silt loam). The bedrock starts at 14 m of depth, where my ERT survey stopped. In both figures, the red points represent the monitoring wells.

The calibration process consisted of inverse modeling to estimate soil hydraulic parameters, namely soil water content (θ_s), n and α (the fitting parameters of the water retention equation by Van Genuchten (1980)), and K_s. The Marquardt-Levenberg optimization scheme was used for parameter

estimation and the same basic steps were followed for both models. I calibrated the hydraulic parameters and only once these were considered acceptable I did proceed with the solute parameters. Both calibrations were based on values specified for several observation points (Figure 3.9). I had five observation points on the surface that allowed an inclusion of the soil moisture data collected in the field and four observation points in the saprolite layer to simulate the bottom of my monitoring wells. Since I was modeling a two-dimensional surface, I represented my shoreline wells as two points and attributed to them distinct averages of the shallow and deep wells. I also used a single point to indicate wells 19 and 20 that were situated at 6 m from the shoreline. Of the upslope wells, only well 22 was included in the model, as all the other wells were dry for either all or most of the sampling period.

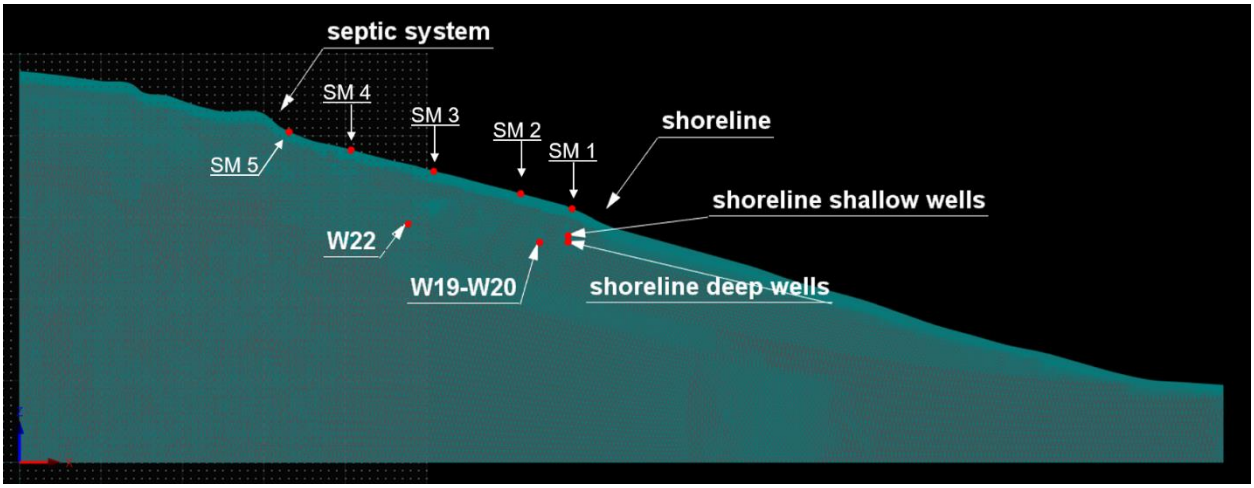


Figure 3.9. The image presents the locations of the observation points (they are the same for both models). The comments indicate the locations, placed where the bottom of the monitoring wells would be, used to target the water table elevations and chloride content. The five dots along the surface (SM) are the observation points used for the superficial soil moisture.

For both models, the calibration values were accepted when: 1) there were no further reductions in the sum of squares error (SSQ), 2) the model stopped after running for the maximum number of selected iterations (10 for hydraulic parameters and four for solute parameters), or 3) the model would not converge, but the values calculated in the last available iteration had acceptable R^2, RMSE, and mass balance values when used to run a "0 iteration" calibration (same as a forward solution, but with performance data on predicted vs observed values).

As the first step for our calibration, I ran a 10-year "warm-up", using the forward solution option in the HYDRUS (2D/3D) with hydraulic parameters for each layer estimated using the Rosetta Lite DLL (Dynamically Linked Library) (Schaap, Leij et al., 2001). For the parsimonious model, I started from the textural classes determined during the soil survey for the Pacolet soil series, while for the complex model I imported the percentages in sand, silt, and clay estimated by the ANN framework. In both cases, instead of the K_s values predicted by the Rosetta Lite DLL model, the values measured in the field were used as initial estimates. The lake levels, rainfall, and transpiration values were those for the decade 2008-2018. The water input coming from the nodal recharge was set as the average water consumption for the household for the same time period and it was equal to 0.0372 m^2/d (spread over the seven recharge nodes). The water usage was transformed from m^3/d into m^2/d (since I was working in a 2D set-up) by dividing the volumetric use by the width of the parcel. The results from the last day of the 10 years model were imported as the initial conditions for the calibration.

For both simulations, I started my calibration from the top layers (topsoil and subsoil). The selected parameters were saturated soil water content (θ_s), n and α (the fitting parameters of the water retention equation by Van Genuchten (1980)), and K_s. Once those values were accepted, I moved to the calibration of K_s for the saprolite. The K_s value for the bedrock was derived from the final results of Ceolin (2019). The initial uncalibrated parameters used for the calibration are presented in Table 3.1 for the parsimonious model and 3.2 for the complex model.

Table 3.1. Summary of the uncalibrated hydraulic parameters for the parsimonious model. The K_s values for topsoil, subsoil and saprolite were measured in the field (Ceolin, 2019). All the other values were estimated using the Rosetta Lite DLL model.

Soil Layer	***θ_r [-]***	***θ_s [-]***	***α [1/m]***	***n [-]***	***K_s [m/day]***	***l [-]***
Sandy loam topsoil	*0.039*	*0.387*	*2.67*	*1.45*	*0.98*	*0.5*
Clay subsoil	*0.098*	*0.459*	*1.50*	*1.25*	*0.062*	*0.5*
Saprolite	*0.079*	*0.442*	*1.58*	*1.41*	*11.6*	*0.5*
Fractured Bedrock	*0.053*	*0.10*	*3.53*	*3.18*	*0.05*	*0.5*
Lakebed	*0.098*	*0.459*	*1.50*	*1.25*	*0.072*	*0.5*

Table 3.2. Summary of the uncalibrated hydraulic parameters for the complex model. The K_s values were measured in the field (Ceolin, 2019). All other values were estimated using the Rosetta Lite DLL included in the software. The K_s value for Sandy Clay Loam was estimated using the Rosetta Lite DLL since it was not possible to link the layer to any of the field measurements.

Soil Layer	***θ_r [-]***	***θ_s [-]***	***α [1/m]***	***n [-]***	***K_s [m/day]***	***l [-]***
Loam	*0.066*	*0.412*	*0.96*	*2.31*	*0.98*	*0.5*
Sandy Clay Loam	*0.065*	*0.397*	*2.08*	*1.37*	*0.314*	*0.5*
Clay Loam	*0.077*	*0.415*	*1.58*	*1.39*	*0.157*	*0.5*
Fractured Bedrock	*0.053*	*0.1*	*3.53*	*3.18*	*0.05*	*0.5*
Lakebed	*0.098*	*0.459*	*1.5*	*1.25*	*0.072*	*0.5*
Silt Loam	*0.053*	*0.424*	*0.459*	*1.68*	*11.6*	*0.5*
Sandy Loam Shallow	*0.037*	*0.388*	*2.99*	*1.41*	*11.6*	*0.5*
Loamy Sand	*0.033*	*0.390*	*4.45*	*1.59*	*11.6*	*0.5*
Sand	*0.044*	*0.383*	*3.89*	*2.36*	*2.57*	*0.5*
Sandy Loam Deep	*0.037*	*0.388*	*2.99*	*1.41*	*11.6*	*0.5*

The minimum and maximum limits set for the calibration, reported in Tables 3.3 and 3.4, were drawn from the literature cited within the software unless the field measured values showed a broader range. The observed values used were the soil moistures and the water table elevations. The water table elevations were included as meters of water above the bottom of the well (or pressure head). The parameters obtained from the calibration of the top layers were accepted and fixed during the calibration of the saprolite. For the saprolite, I only calibrated K_s using water table elevations as target values. During the saprolite calibration, I attributed a weight of 5 to well 22 since the well was the farthest away and possibly had the least influence of the lake levels on its water table elevations. For the second model, all four saprolite layers were calibrated together. I characterized the two layers with sandy loam texture as a separate material (shallow and deep), even though they had the same texture, to allow for distinct calibrations of the two. The data used as observed values in the calibration were collected from October 2018 to the end of July 2019. I included 29 observed values for soil moisture and 55 for pressure head. The water usage for the calibration period, and consequent water pulses amount from the recharge nodes, was equal to 0.0478 m^2/d.

Table 3.3. Minimum and maximum limits set for the calibration for the parsimonious model.

Soil Layer	***θ_s [-]***	***α [1/m]***	***n [-]***	***K_s [m/day]***
Sandy loam topsoil min	*0.387*	*2.67*	*0.322*	*0.28*
Sandy loam topsoil max	*0.454*	*2742*	*1.89*	*2.56*
Clay subsoil min	*0.38*	*0.8*	*0.131*	*0.0144*
Clay subsoil max	*0.459*	*1e+007*	*4.67*	*0.148*
Saprolite min	-	-	-	*7.73*
Saprolite max	-	-	-	*40*

Table 3.4. Minimum and maximum limits set for the calibration for the complex model.

Soil Layer	θ_s [-]	α [1/m]	n [-]	K_s [m/day]
Loam min	*0.399*	*0.96*	*0.22*	*0.097*
Loam max	*0.454*	*10180*	*3*	*2.56*
Sandy Clay Loam min	*0.33*	*2.1*	*0.25*	*0.01*
Sandy Clay Loam max	*0.436*	*8089*	*2.04*	*0.475*
Clay Loam min	*0.39*	*1.191*	*0.194*	*0.055*
Clay Loam max	*0.442*	*66630*	*2.81*	*0.34*
Silt Loam min	-	-	-	*0.108*
Silt Loam max	-	-	-	*40*
Sandy Loam Shallow min	-	-	-	*0.382*
Sandy Loam Shallow max	-	-	-	*40*
Loamy Sand min	-	-	-	*0.861*
Loamy Sand max	-	-	-	*40*
Sandy Loam Deep min	-	-	-	*0.382*
Sandy Loam Deep max	-	-	-	*40*

Once the hydraulic parameters were calibrated and validated, I used them to run a new "warm-up" period. The new "warm-up" lasted for 20 years (I repeated the 10 year conditions twice), to allow for the establishment of the chloride plume. I set a daily pulse of 22.7 mg/L of chloride in the recharge node (as measured from the septic tank at the site) and starting background concentrations of 5 mg/L of chloride. The background concentration, applied to the entire slope, was estimated from the average concentration measured at the control site. The results from the last day of the 20 years were imported as initial conditions for the solute parameters' calibration. For this step, I decided to calibrate only the transverse (λ_T) and lateral (λ_L) dispersivities. The dispersivities were calibrated for the top layers, the saprolite, and the bedrock. Initial estimates for the top layers and the fractured bedrock were from

Bradshaw and Radcliffe (2013). For the saprolite, I estimated the starting values of λ_L and λ_T. Bulk density (ρ_b) was derived from θ_s, using the following equation (Equation 3.4):

$$\rho_b = \rho_s * (1 - \theta_s) \qquad [3.4]$$

where ρ_s is the soil particle density and θ_s is substituted to the soil porosity since, for saturated soils, all pores are filled by water. For the bedrock, I used a density value of 2.75 m/m^3 (Smithson, 1971). For all the other layers, soil particle density is assumed to be 2.65 m/m^3, as that is the approximate density for most soil particles (Radcliffe & Simunek, 2018).

The values for the adsorption isotherm coefficient (K_d) were derived from Qafoku, Sumner et al. (2000), according to the measured pH values for our soil cores at different depths. I used the values from Bradshaw and Radcliffe (2013) for the bedrock and the lakebed (assumed to be a clayey layer), since I didn't have pH information on those two layers. The initial uncalibrated values for the two models are presented in tables 3.5 and 3.6. In this case, also, the data used as target values for the calibration were collected from October 2018 and the end of July 2019, for a total of 35 observed values. I attributed a weight of 2 for the observation points for wells 19 and 20, and well 22. The decision was based on the fact that the 2 points representing the shoreline wells have the same x coordinate, so I wanted to equalize the effect of each well on the location of the plume.

Table 3.5. Uncalibrated values for the chloride transport and reaction simulation used in the parsimonious model.

Soil Layer	***ρ_b [g/m^3]***	***λ_L [m]***	***λ_T [m]***	***K_d [m^3/g]***
Sandy loam topsoil	*1.609*	*0.15*	*0.05*	*0.171*
Clay subsoil	*1.434*	*0.15*	*0.05*	*0.195*
Saprolite	*1.479*	*0.5*	*0.004*	*0.311*
Fractured Bedrock	*2.475*	*0.15*	*0.05*	*0.25*
Lakebed	*1.434*	*0.15*	*0.05*	*0.25*

Table 3.6. Uncalibrated values for the chloride transport and reaction simulation used in the complex model.

Soil Layer	ρ_b [g/m^3]	λ_L [m]	λ_T [m]	K_d [m^3/g]
Loam	*1.518*	*0.15*	*0.05*	*0.195*
Sandy Clay Loam	*1.611*	*0.15*	*0.05*	*0.218*
Clay Loam	*1.586*	*0.15*	*0.05*	*0.218*
Fractured Bedrock	*2.475*	*0.15*	*0.05*	*0.250*
Lakebed	*1.434*	*0.15*	*0.05*	*0.250*
Silt Loam	*1.527*	*0.5*	*0.004*	*0.311*
Sandy Loam Shallow	*1.621*	*0.5*	*0.004*	*0.311*
Loamy Sand	*1.615*	*0.5*	*0.004*	*0.311*
Sand	*1.634*	*0.5*	*0.004*	*0.311*
Sandy Loam Deep	*1.621*	*0.5*	*0.004*	*0.311*

Results

The observed values, used for the calibration and validation of the two models, are presented in Figure 3.10. The values for the shoreline wells and wells 19 and 20 have been averaged, as explained previously. For all the months, the chloride concentrations were higher in the right section of the site (facing the lake), suggesting a development of the plume towards that side of the hillslope. In March 2019, at the shoreline, we measured values equal to 123.4 mg/L, on the right side, versus 7.57 mg/L on the left side.

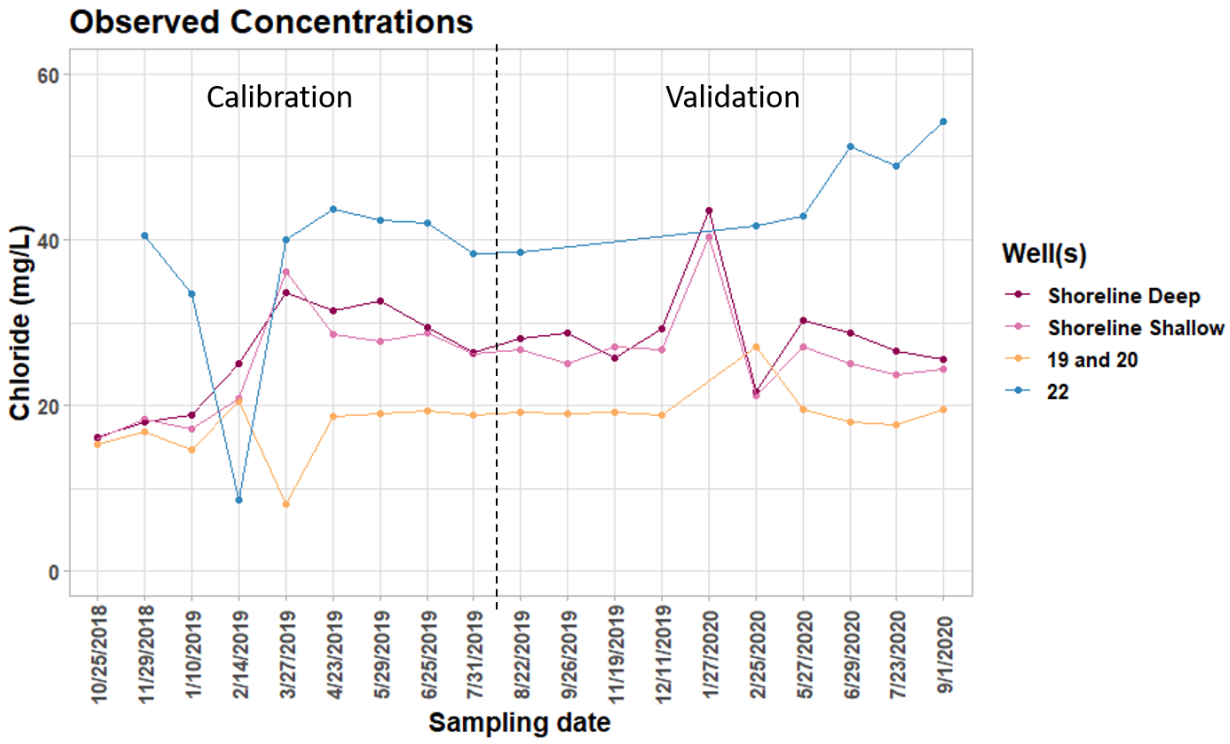


Figure 3.10. The graph shows the observed chloride concentrations used to calibrate and validate the two models.

The calibrated hydraulic parameters for the two models are presented in table 3.7 and 3.8.

Table 3.7. Calibrated hydraulic parameters, parsimonious model. For the saprolite, K_s was the only calibrated value.

Soil Layer	*θ_s [-]*	*α [1/m]*	*n [-]*	*K_s [m/day]*
Sandy loam topsoil	*0.393*	*0.751*	*2.16*	*2.11*
Clay subsoil	*0.459*	*1.1*	*1.86*	*0.014*
Saprolite	-	-	-	*19.5*

Table 3.8. Calibrated hydraulic parameters, complex model. For the saprolite, K_s was the only calibrated value.

Soil Layer	*θ_s [-]*	*α [1/m]*	*n [-]*	*K_s [m/day]*
Loam	*0.427*	*0.96*	*2.43*	*1.06*
Sandy Clay Loam	*0.392*	*2.29*	*1.20*	*0.314*
Clay Loam	*0.401*	*1.57*	*1.26*	*0.168*
Silt Loam	-	-	-	*0.108*
Sandy Loam Shallow	-	-	-	*15.3*
Loamy Sand	-	-	-	*16.5*
Sandy Loam Deep	-	-	-	*0.382*

The above values were used to re-calculate the "warm-up", that was used for the initial conditions to the validations of surface soil moisture (Figure 3.11) and water table elevations (Figure 3.12)

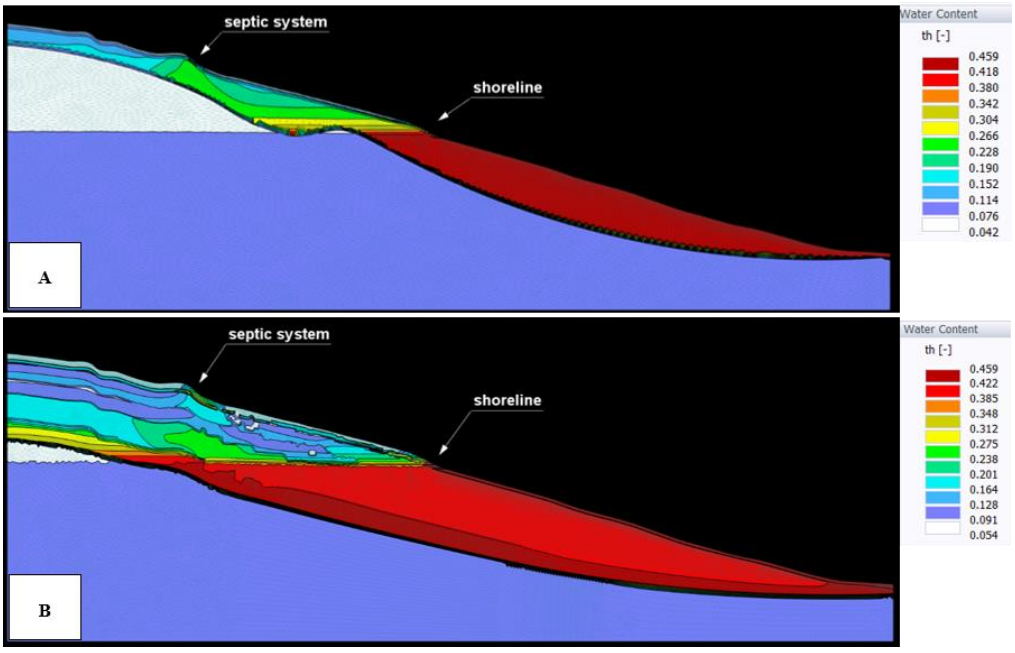


Figure 3.11. The images (A for the parsimonious model and B for the complex model) show the water content for the two models on the last day of the "warm-up" period. The values are used as initial conditions for the validation.

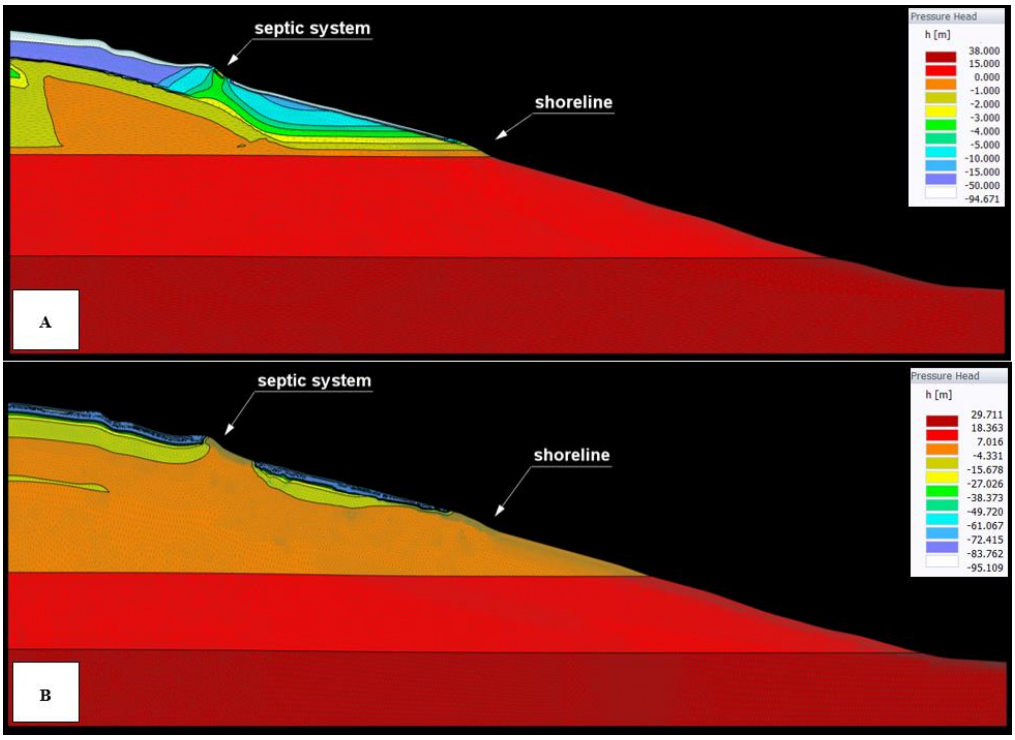
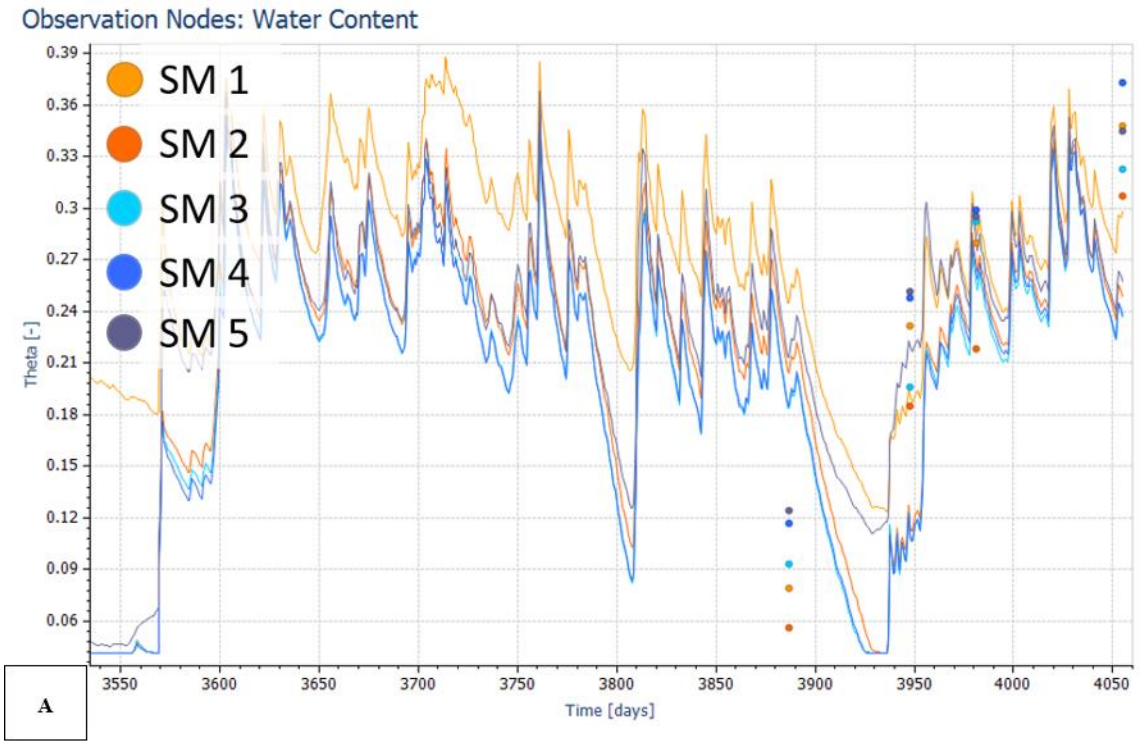


Figure 3.12. The images (A for the parsimonious model and B for the complex model) show the values of pressure head for the two models on the last day of the "warm-up" period. The values are used as initial conditions for the validation.

The validation period goes from August 2019 to September 2020, with 20 observed values for soil moisture and 39 for pressure head. The results of the validation are presented in Figures 3.13 and 3.14.



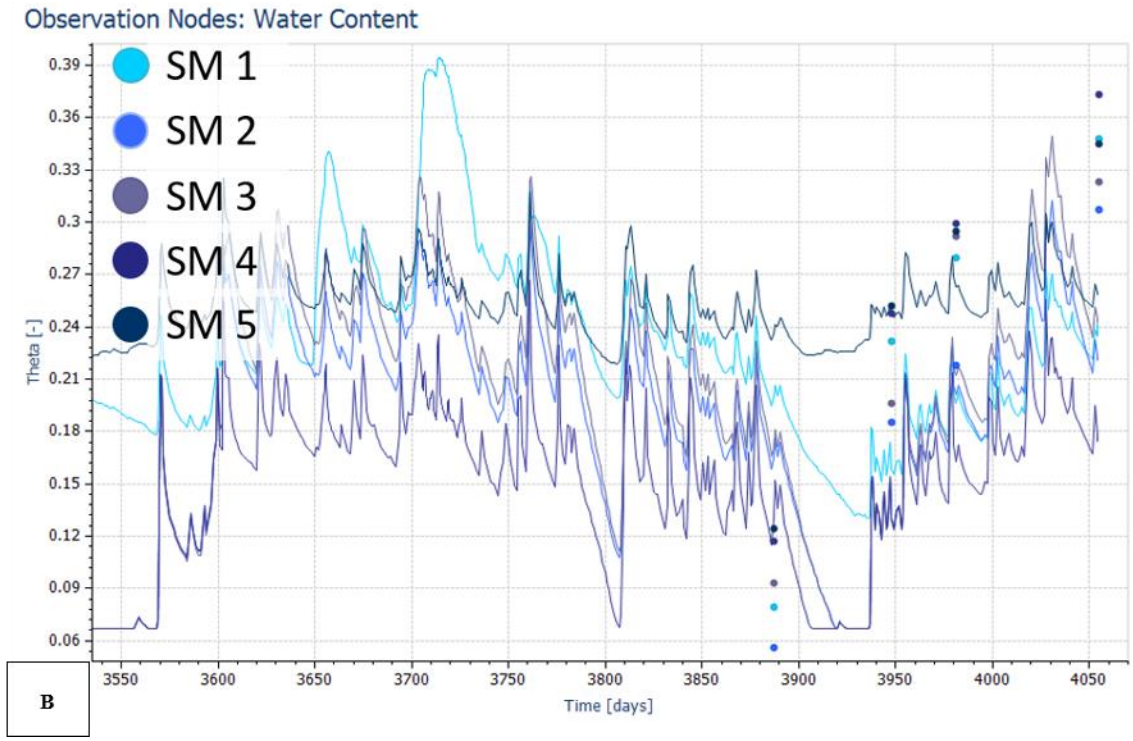


Figure 3.13. The two graphs show the predicted (lines) versus observed (point) water content values for the parsimonious model (A) and the complex model (B). The observation points are identified as "SM", standing for soil moisture, and numbered to distinguish their location.

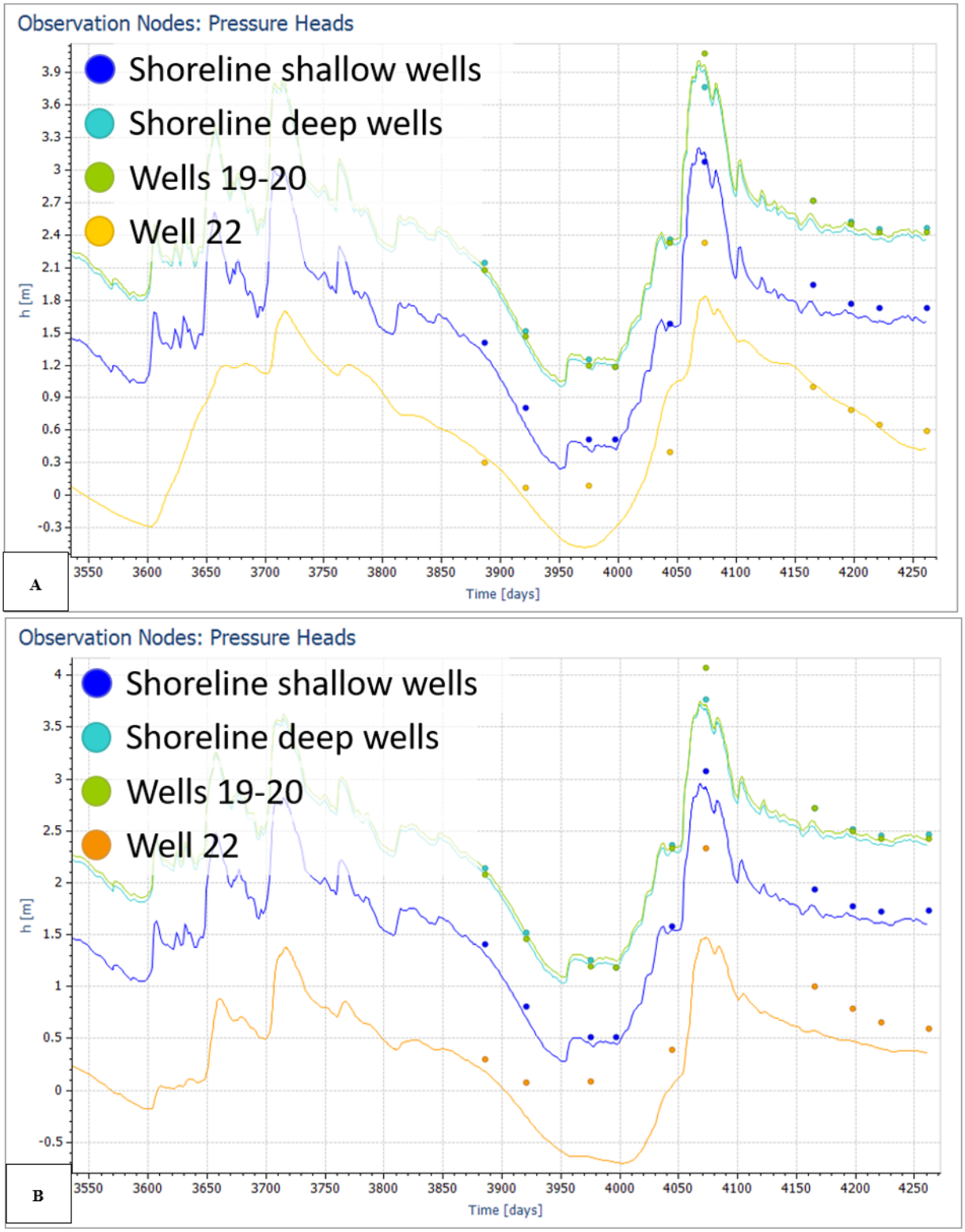


Figure 3.14. The two graphs show the predicted (lines) versus observed (point) pressure head values for the parsimonious model (A) and the complex model (B). Both models replicate the observed fluctuations.

The calibrated dispersivity values for chloride are presented in Tables 3.9 and 3.10. Again, the calibrated values were used to run a "warm-up" period to create the initial conditions for the validation.

Table 3.9. Summary of the calibrated values for lateral and transverse dispersivities for the parsimonious model.

Soil Layer	λ_L ***[m]***	λ_T ***[m]***
Sandy loam topsoil	*0.252*	*0.025*
Clay subsoil	*0.254*	*0.025*
Saprolite	*0.165*	*0.004*
Fractured Bedrock	*0.083*	*0.004*

Table 3.10. Summary of the calibrated values for lateral and transverse dispersivities for the complex model.

Soil Layer	λ_L ***[m]***	λ_T ***[m]***
Loam	*0.196*	*0.015*
Sandy Clay Loam	*0.15*	*0.015*
Clay Loam	*0.15*	*0.015*
Fractured Bedrock	*0.020*	*0.015*
Silt Loam	*0.15*	*0.05*
Sandy Loam Shallow	*0.5*	*0.005*
Loamy Sand	*0.015*	*0.004*
Sandy Loam Deep	*0.015*	*0.004*

The initial conditions are shown in Figure 3.15. The development of the plume is influenced by the position of the bedrock. In the parsimonious model, the plume goes primarily towards the lake. Its direction changes, going vertically, only after reaching a depression in the bedrock, at about 6 m

downslope in the drainfield. The deepest point reached is at –17.15 m, in the upper part of the fractured bedrock. In the complex model, the plume moves horizontally for 5.5 m, but extends up to 18 m deeper (vertical flux), under the area of the drainfield, before starting to move primarily toward the shoreline (horizontal flux). In both models, the initial flux seems to be favored by the high Ks of the saprolite (only for the upper layers in the complex model). For the complex model, the absence of a shallow bedrock causes the bulk of the horizontal flux to mainly happen in the deep sandy loam, silt loam, and top of the fractured bedrock layers. These layers are characterized by very low K_s and dispersivities values. Starting from those initial conditions, I performed a validation using 35 observed values, collected between August 2019 and September 2020.

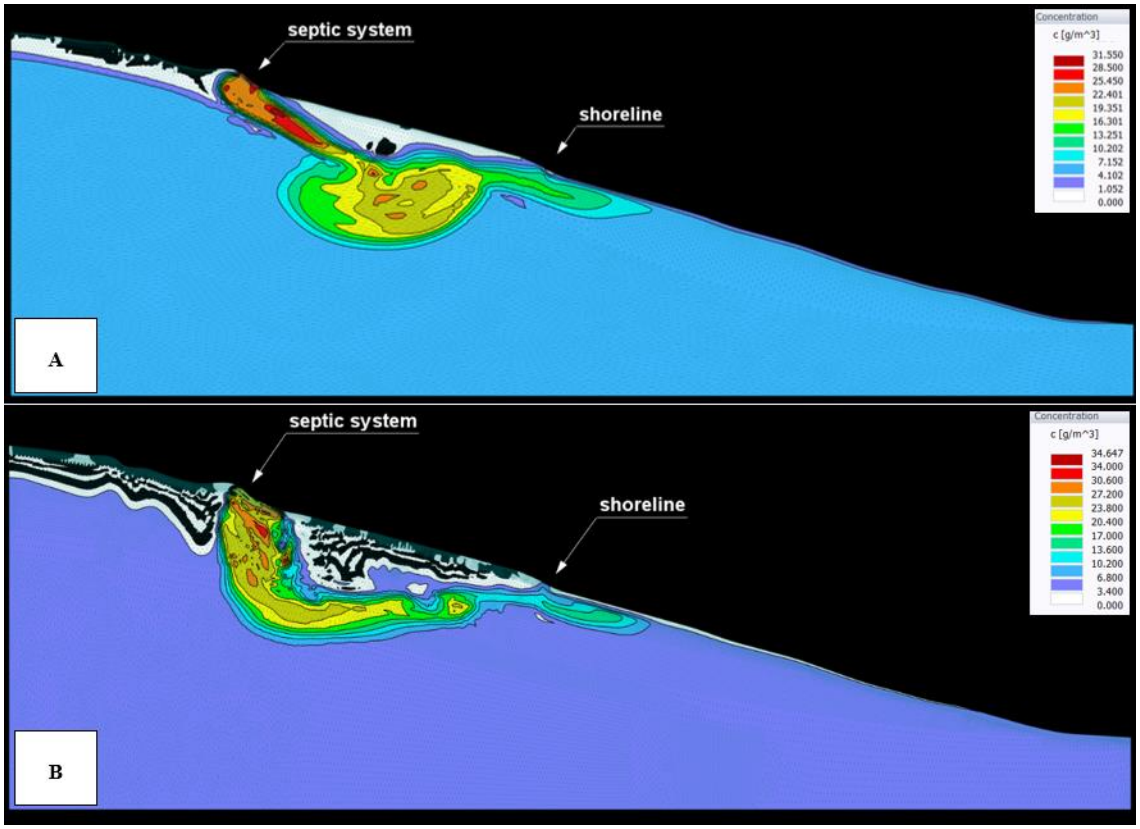


Figure 3.15. The images (A for the parsimonious model and B for the complex model) show the chloride concentrations for the two models on the last day of the "warm-up" period. The values are used as initial conditions for the validation.

In the detailed model, during the first 80 days of the validation, at the interface between the Bt horizon, characterized by sandy clay loam texture, and the first saprolite layer, presenting sandy loam texture, the chloride movement resembles the presence of funnel-type preferential flow (Figure 3.16). A small section of the plume, in red, starts moving laterally in the clayey layer before infiltrating in the saprolite. The results of the validation are reported in Figure 3.17.

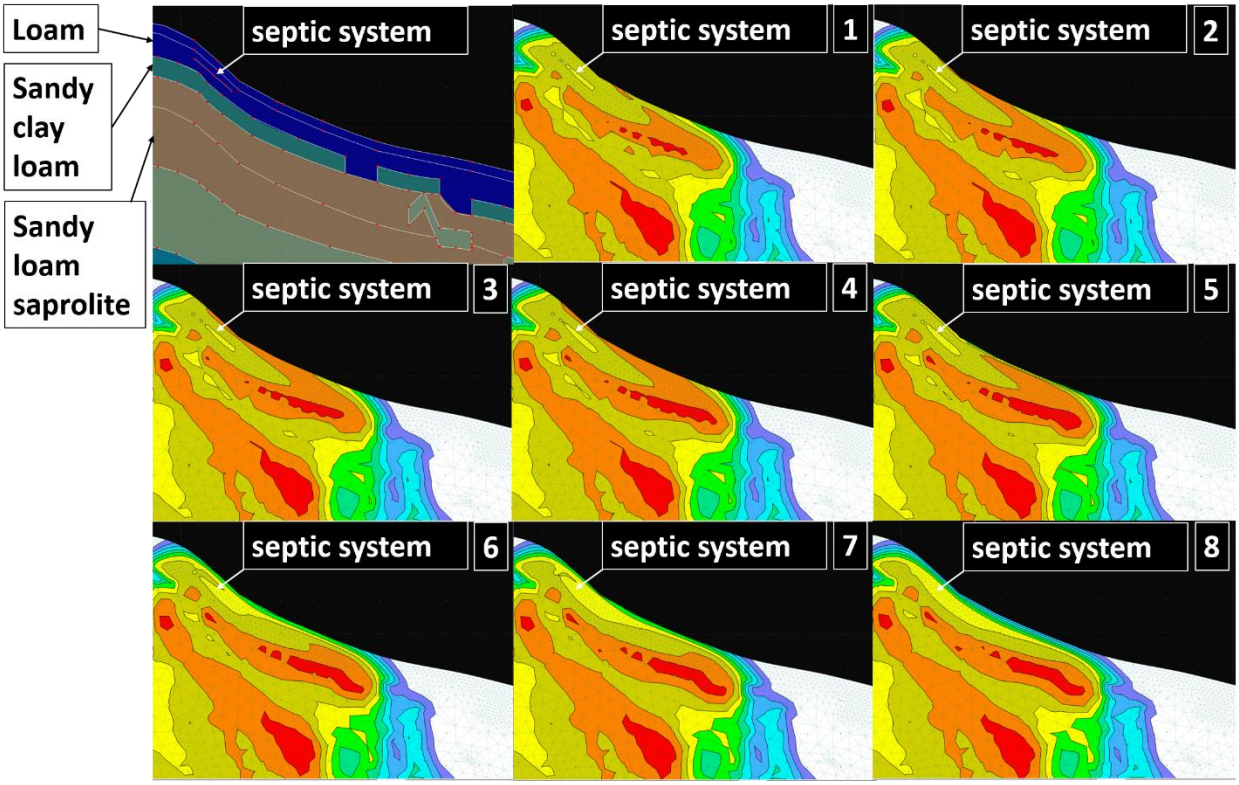


Figure 3.16. The images show the movement of chloride through funnel-type preferential flow. The first box on the upper left corner represents the domain properties of the area and the position of the recharge nodes simulating the septic system. The other frames, numbered from 1 to 8, capture the development of the plume at 10 days intervals. Frame 1 shows the state of the plume at the last day of the "warm-up" period.

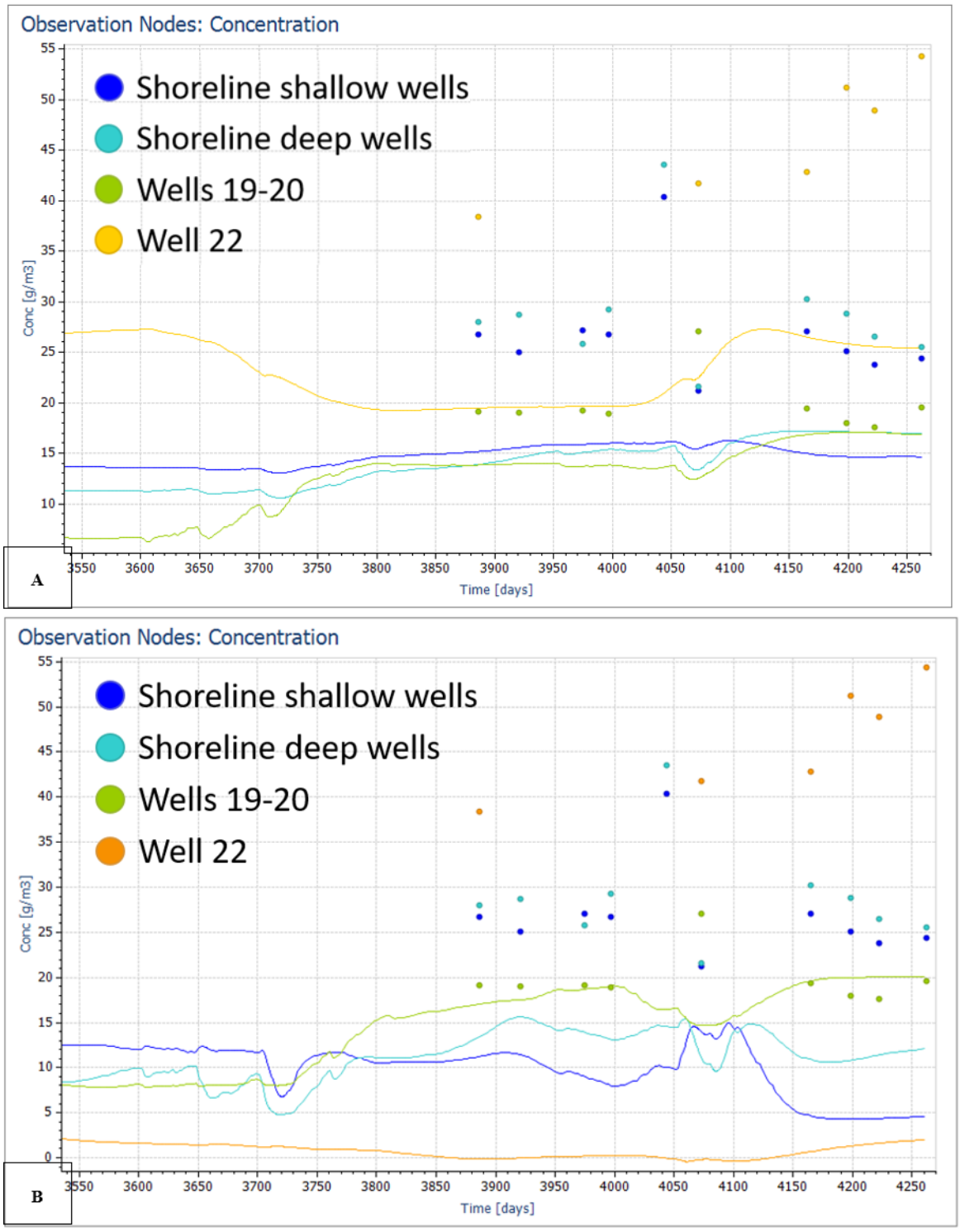


Figure 3.17. Predicted chloride concentrations (lines) versus the observed values (points). The two models show very different patterns, but for the most part, they underestimate the chloride concentrations.

The three validations were performed separately to understand the ability of the two models to accurately predict all the different phenomena. For both models, the resulting values of R^2, RMSE, and mass balance error are summarized in table 3.11.

Table 3.11. Values of R^2, RMSE, and mass balance error for the three validations of both models.

	Parsimonious Model			***Complex Model***		
	R^2	*RMSE*	*mass balance error*	*R^2*	*RMSE*	*mass balance error*
Water content - Model	*0.20*	*0.09 cm^3/cm^3*	*0.03%*	*0.21*	*0.09 cm^3/cm^3*	*0.20%*
Water content - SM 1	*0.25*	*0.09 cm^3/cm^3*	-	*0.26*	*0.09 cm^3/cm^3*	-
Water content - SM 2	*0.13*	*0.09 cm^3/cm^3*	-	*0.43*	*0.07 cm^3/cm^3*	-
Water content - SM 3	*0.36*	*0.08 cm^3/cm^3*	-	*0.55*	*0.08 cm^3/cm^3*	-
Water content - SM 4	*0.21*	*0.10 cm^3/cm^3*	-	*0.83*	*0.13 cm^3/cm^3*	-
Water content - SM 5	*0.55*	*0.06 cm^3/cm^3*	-	*0.83*	*0.07 cm^3/cm^3*	-
Pressure Head - Model	*0.97*	*0.18 m*	*0.02%*	*0.96*	*0.24 m*	*0.19%*
Pressure Head - Shoreline Shallow Wells	*0.99*	*0.11 m*	-	*1.00*	*0.11 m*	-
Pressure Head - Shoreline Deep Wells	*0.99*	*0.10 m*	-	*1.00*	*0.09 m*	-
Pressure Head - Wells 19 and 20	*1.00*	*0.06 m*	-	*0.99*	*0.12 m*	-
Pressure Head - Well 22	*0.77*	*0.33 m*	-	*0.89*	*0.46 m*	-
Chloride concentration - Model	*0.63*	*14.06 mg/L*	*0.03%*	*0.52*	*23.48 mg/L*	*0.19%*
Chloride concentration - Shoreline Shallow Wells	*0.29*	*12.34 mg/L*	-	*0.01*	*19.44 mg/L*	-
Chloride concentration - Shoreline Deep Wells	*0.03*	*14.26 mg/L*	-	*0.18*	*16.87 mg/L*	-
Chloride concentration - Wells 19 and 20	*0.33*	*6.12 mg/L*	-	*0.67*	*4.35 mg/L*	-
Chloride concentration - Well 22	*0.49*	*22.45 mg/L*	-	*0.88*	*45.65 mg/L*	-

When looking at soil moisture at the surface, neither model performs well (Table 3.11). Even though the predictions seem to fall within the observed ranges, both models miss the lowest and highest values. The differences in the trends between observed and predicted water content values highlight a discrepancy in the timing of drying and wetting periods for observations and predictions. This discrepancy suggests that the source of error could reside in the data used for rainfall and transpiration, rather than phenomena linked to runoff, intensity of rain, runoff, and evaporation rates. Even though the two weather stations used to gather rainfall and evapotranspiration are close to the study site, local rainfall events could have been registered at the weather station without also happening at the modeled location, and vice versa. The main differences in predictions are probably linked to the soil texture chosen for the surface layer, that for the parsimonious model is a sandy loam while for the complex model is a loam.

Moving to predictions for the water table elevations, both models show very high accuracy, with R^2 values of 0.97 for the parsimonious model and 0.96 for the complex model. The shoreline wells and wells 19 and 20, show R^2 values even closer to 1. The predicted fluctuations follow the same pattern as the observed values. Water table levels at the shoreline are highly influenced by the lake levels. The main discrepancies are linked to the highest peaks and the predictions for well 22. When looking only at these two aspects, the parsimonious model has higher accuracy than the complex model. This effect can be explained by transmissivity. Transmissivity [$L^2\ T^{-1}$] is a hydraulic property that describes the ability of water to move or transmit through the saturated zone in the direction normal to the base of the aquifer (Equation 3.5) (Stewart & Howell, 2003):

$$T = KB \qquad [3.5]$$

where K is the hydraulic conductivity and B is the thickness of the aquifer, in our case represented by the saprolite layer. In this case, the different values of transmissivity are linked to the different depths to bedrock used for the two models. At the same location, in the parsimonious model we find the bedrock with very low K_s, while at site 2 we are still in the saprolite layers, characterized by higher hydraulic conductivity values. This difference allows the water table in the first model to raise higher than in the second, thus showing better performances for the peaks and well 22. These results reiterate the findings

from Tromp-van Meerveld and Weiler (2008) and the importance of the inclusion of the bedrock in hillslope scale models.

Finally, for the chloride transport predictions, both models underestimate most of the concentrations and the finer description of the soil layers, does not seem to generate funnel flow in the saprolite. In particular, the three features presented in Figure 3.7, showing a different texture than the surrounding area, were never intercepted by the plume, neither during the "warm-up" phase, nor during the validation. At the interface between the Bt horizon and the saprolite, vertically below the recharge nodes, the movement of some sections of the plume, having higher chloride concentrations than the rest of the area, seemed to show funnel flow, by moving laterally in the clayey layer before infiltrating the saprolite. Nevertheless, in this case the movement was happening within the boundaries of the plume, thus not showing an effect on the movement towards the monitoring wells or the shoreline. The initial values used for the warm-up and calibration proved particularly important since the calibration process is based on the optimization of local minima. An initial calibration starting from lower values of lateral dispersivity produced worse results than the one reported here. Wells 19 and 20 are the only ones, for which the predicted values are close to the observed ones (with the complex model showing a higher R^2 and lower RMSE values than the parsimonious model), even though in both cases the predicted trends appear to be opposite to the observed ones (Figure 3.16). In parsimonious model, the predictions for W22 are closer to the observed values than in the complex model but in both cases the RMSE values are high, 22.45 mg/L for the parsimonious model and 45.65 mg/L for the complex model. Among all wells, well 22 is the only one where the models correctly predict the temporal trends (Figure 3.16). For this observation point in particular, depth to bedrock is playing an important role, as pointed out for the initial conditions. In the complex model, the section of the plume showing the higher concentrations appears to be below well 22, thus affecting the predicted levels. Neither model seems to capture the behavior of the shoreline wells (R^2 values of 0.29 and 0.03 for the parsimonious model, and of 0.01 and 0.18 for the complex model). Interestingly, the complex model shows more variability in chloride content than the parsimonious one, but the predicted fluctuations are lagged in respect to the observed ones.

In both my models, the chloride input from the recharge nodes was kept constant, thus probably affecting their ability to predict fluctuations in the chloride concentrations as shown in the literature. In their study, Bradshaw, Radcliffe et al. (2013) reported that observed and predicted chloride concentrations in fact mirrored the variations in input. The use of recharge nodes in HYDRUS (2D/3D) does not allow the modeler to change the input value over time in a single run and, for my study, the septic tank was sampled only once, eliminating the possibility to have variable concentrations for the model (because of time constraint reasons and Covid-19 safety measures, I was not able to replicate the measurement). The detected chloride value is in the low range among concentrations reported in the literature. Other studies (Bradshaw & Radcliffe, 2013; Snider, Roy et al., 2017) showed that chloride concentrations in septic tanks can be variable over time, registering values between 22.7 and 160 mg/L. The concentration measured at our site falls on the lower range of values, suggesting that over the months, the input coming from the tank could have been higher than what we measured in our only sampling event. The low concentration could partially explain why all the predicted values are underestimated. Additionally, at the shoreline, the chloride values higher than 100 mg/L could be also linked to external sources of solute that were not accounted for. The plume coming from the neighboring house could have been intercepted by some of our monitoring wells. The discrepancies in observed and predicted concentrations at the shoreline could be caused by this external source of chloride, thus reaffirming the complexity of the subsurface in the area and the challenge in modeling three dimensional processes in a two dimensional setting. I exclude that the highest concentrations measured at the shoreline were caused by water coming from the lake, as they were all in a restricted area and a separate study taking samples from the lake was measuring much lower concentrations.

Conclusions

In this study, I investigated the relevance of domain complexity for soil layering and weathering state on predictions for water and solute movement at the hillslope scale. The two simulations highlighted the central role of the saprolite layer and the bedrock. The bulk of water and solute movement in the

hillslope was happening in the saprolite. It is fundamental to keep studying the properties and processes happening in this material, to produce accurate predictions in the Piedmont region. It appears that even the characterization of the saprolite layer with different textures did not produce funnel-type preferential flow in the saprolite.

Depth to bedrock appeared to be highly related to both water table elevations (because of transmissivity) and chloride concentrations. Other geophysical methods could prove more helpful than the ERT in the establishment of the top of the bedrock. It could be valuable to keep investigating their use in the Piedmont region for that scope.

Most of the observed chloride concentrations were higher than the input measured in the septic tank. For my study I sampled the septic tank only once (for time constraint reasons and Covid-19 safety measures, we were not able to replicate the measurement). When compared to other values in the literature, the concentration measured at our site falls on the lower range of values, suggesting that over the months, the input coming from the tank could have been higher than what measured in my only sampling event. At the shoreline, the higher chloride values could be also linked to external sources of solute that were not accounted for (possibly the interference of a plume from the neighboring house that was intercepted by our wells).

The temporal variations in the chloride observations were much larger than the chloride simulations. This occurrence could be explained by two phenomena, 1) the variable chloride concentrations in the septic tank, replicated in the wells through fast flow along the hillslope, 2) the possibility of macropore preferential flow. Exploring the effect of these two phenomena could help getting a better understanding of the predominant processes affecting solute movement at the hillslope scale.

This study stressed the importance of saprolite and depth to bedrock in hydraulic and solute transport processes at the hillslope scale in the Piedmont Region. Even though the parsimonious model had the best performances overall, the complex model presented more dynamic predictions in chloride concentrations along the hillslope. This difference suggests that a parsimonious model could provide

accurate average loads of a solute along a hillslope to a water body, while a complex model could be more apt in replicating the temporal variations of solute movement along the hillslope.

References

Amoozegar, A., Schoeneberger, P. J., & Vepraskas, M. J. (1991). *Characterization of soils and saprolites from the Piedmont region for waste disposal purposes*. Retrieved from https://repository.lib.ncsu.edu/bitstream/handle/1840.4/1909/NC-WRRI-255.pdf?sequence=1&isAllowed=y

Box, G. E., & Jenkins, G. M. (1976). Time series analysis: Forecasting and control San Francisco. *Calif: Holden-Day*.

Bradshaw, J. K., & Radcliffe, D. E. (2013). Nitrogen fate and transport in a conventional onsite wastewater treatment system installed in a clay soil: Experimental results. *Vadose Zone Journal, 12*(3), 1-20.

Bradshaw, J. K., Radcliffe, D. E., Šimůnek, J., Wunsch, A., & McCray, J. E. (2013). Nitrogen fate and transport in a conventional onsite wastewater treatment system installed in a clay soil: A nitrogen chain model. *Vadose Zone Journal, 12*(3), 1-20.

Ceolin, S. (2019). *Preliminary Assessment of Onsite Wastewater Treatment System Effects on Lake Lanier Water Quality*. (Master of Science of Crop and Soil Sciences). University of Georgia and University of Padova,

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., Dietrich, W. E., Flores, A., Grant, G., Kirchner, J. W., Mackay, D. S., McDonnell, J. J., Milly, P. C. D., Sullivan, P. L., Tague, C., Ajami, H., Chaney, N., Hartmann, A., Hazenberg, P., McNamara, J., Pelletier, J., Perket, J., Rouholahnejad-Freund, E., Wagener, T., Zeng, X., Beighley, E., Buzan, J., Huang, M., Livneh, B., Mohanty, B. P., Nijssen, B., Safeeq, M., Shen, C., Verseveld, W., Volk, J., & Yamazaki, D. (2019). *Hillslope Hydrology in Global Change Research and Earth System Modeling*. Water Resources Research, (55, 2).

Feddes, R. A., Bresler, E., & Neuman, S. P. (1974). Field test of a modified numerical model for water uptake by root systems. *Water Resources Research, 10*(6), 1199-1206.

Graham, C. B., Van Verseveld, W., Barnard, H. R., & McDonnell, J. J. (2010). Estimating the deep seepage component of the hillslope and catchment water balance within a measurement uncertainty framework. *Hydrological Processes, 24*(25), 3631-3647. doi:10.1002/hyp.7788

Haught, D. R. W., & Meerveld, H. J. (2011). Spatial variation in transient water table responses: differences between an upper and lower hillslope zone. *Hydrological Processes, 25*(25), 3866-3877. doi:10.1002/hyp.8354

Jia, X., Shao, M. A., Wei, X., & Wang, Y. (2013). Hillslope scale temporal stability of soil water storage in diverse soil layers. *Journal of Hydrology, 498*, 254-264. doi:10.1016/j.jhydrol.2013.05.042

Ju, S.-H., & Kung, K.-J. S. (1993). Simulating Funnel-Type Preferential Flow and Overall Flow Property Induced by Multiple Soil Layers. *Journal of Environmental Quality, 22*(3), 432-442. doi:10.2134/jeq1993.00472425002200030006x

Kung, K. J. S. (1990a). Preferential flow in a sandy vadose zone: 1. Field observation. *Geoderma, 46*(1), 51-58. doi:https://doi.org/10.1016/0016-7061(90)90006-U

Kung, K. J. S. (1990b). Preferential flow in a sandy vadose zone: 2. Mechanism and implications. *Geoderma, 46*(1), 59-71. doi:https://doi.org/10.1016/0016-7061(90)90007-V

LeGrand, H. E., Back, W., Rosenshein, J. S., & Seaber, P. R. (1988). Region 21, Piedmont and Blue Ridge. In *Hydrogeology* (Vol. O-2, pp. 201-208): Geological Society of America.

McGuire, K. J., & McDonnell, J. J. (2010). Hydrological connectivity of hillslopes and streams: Characteristic time scales and nonlinearities. *Water Resources Research, 46*(10), n/a-n/a. doi:10.1029/2010wr009341

Miller, J. A. (1990). *Ground Water Atlas of the United States: Segment 6, Alabama, Florida, Georgia, South Carolina* (730G). Retrieved from http://pubs.er.usgs.gov/publication/ha730G

Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research, 12*(3), 513-522.

Ohta, T., & Arai, H. (2007). Statistical empirical index of chemical weathering in igneous rocks: A new tool for evaluating the degree of weathering. *Chemical Geology, 240*(3-4), 280-297. doi:10.1016/j.chemgeo.2007.02.017

Pavich, M. J. (1996). Appalachian Piedmont regolith: relations of saprolite and residual soils to rock-type. *Geotechnical Special Publication*(63).

Pavich, M. J., Leo, G., Obermeier, S., & Estabrook, J. (1989). *Investigations of the characteristics, origin, and residence time of the upland residual mantle of the Piedmont of Fairfax County, Virginia* (2330-7102). Retrieved from

Qafoku, N. P., Sumner, M. E., & Radcliffe, D. E. (2000). Anion transport in columns of variable charge subsoils: Nitrate and chloride. *Journal of Environmental Quality, 29*(2), 484-493.

Radcliffe, D. E., & Simunek, J. (2018). *Soil physics with HYDRUS: Modeling and applications*: CRC press.

Richards, L. A. (1931). Capillary conduction of liquids through porous mediums. *Physics, 1*(5), 318-333. doi:10.1063/1.1745010

Schaap, M. G., Leij, F. J., & Van Genuchten, M. T. (2001). Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of Hydrology, 251*(3-4), 163-176.

Schroeder, P. A., & West, L. T. (2005). Weathering profiles developed on granitic mafic and ultramafic terrains in the area of Elberton, Georgia. *Georgia Geological Society Guidebook, 25*, 55-80.

Šimůnek, J., Šejna, M., & van Genuchten, M. T. (2018). New features of version 3 of the HYDRUS (2D/3D) computer software package. *Journal of hydrology and hydromechanics, 66*(2), 133-142.

Šimůnek, J., Van Genuchten, M. T., & Šejna, M. (2012). The HYDRUS software package for simulating the two-and three-dimensional movement of water, heat, and multiple solutes in variably-saturated porous media. *Technical manual*.

Smithson, S. B. (1971). Densities of metamorphic rocks. *Geophysics, 36*(4), 690-694. doi:10.1190/1.1440205

Snider, D. M., Roy, J. W., Robertson, W. D., Garda, D. I., & Spoelstra, J. (2017). Concentrations of artificial sweeteners and their ratios with nutrients in septic system wastewater. *Groundwater Monitoring & Remediation, 37*(3), 94-102.

Soil Survey Staff, Natural Resources Conservation Service, & Agriculture., U. S. D. o. (2020). Official Soil Series Descriptions. Retrieved from https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/geo/?cid=nrcs142p2_053587

Stewart, B. A., & Howell, T. (2003). *Encyclopedia of Water Science (Print)*: CRC press.

Tromp-van Meerveld, I., & Weiler, M. (2008). Hillslope dynamics modeled with increasing complexity. *Journal of Hydrology, 361*(1-2), 24-40.

Van Genuchten, M. T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal, 44*(5), 892-898.

Wagener, T., Lees, M. J., & Wheater, H. S. (2001). A toolkit for the development and application of parsimonious hydrological models. *Mathematical models of small watershed hydrology, 2*, 1-34.

Walter, M. T., Kim, J.-S., Steenhuis, T. S., Parlange, J.-Y., Heilig, A., Braddock, R. D., Selker, J. S., & Boll, J. (2000). Funneled flow mechanisms in a sloping layered soil: Laboratory investigation. *Water Resources Research, 36*(4), 841-849. doi:https://doi.org/10.1029/1999WR900328

Zauyah, S., Schaefer, C. E. G. R., & Simas, F. N. B. (2018). Chapter 3 - Saprolites. In G. Stoops, V. Marcelino, & F. Mees (Eds.), *Interpretation of Micromorphological Features of Soils and Regoliths (Second Edition)* (pp. 37-57): Elsevier.

CONCLUSIONS AND FUTURE WORK

In this work, I tried to assess the ability of data obtained from Electrical Resistivity Tomography (ERT), to provide necessary information for water and solute movement at the hillslope scale in the Piedmont Region in Georgia. I first explored the possibility of determining soil textures from the electrical resistivity of soil, the relative depth of investigation, and the weekly antecedent rainfall in an Artificial Neural Network framework. Then, I included these textures in a 2-D model for water and chloride movement. The predictions were compared to those obtained from a model presenting a horizon-based characterization of the subsurface, thus allowing us to perform an analysis of the effect of domain complexity for soil layering and weathering state of the saprolite. This analysis aimed at identifying the best structure to predict nutrient movement from OWTS to Lake Lanier at the hillslope scale. For our study, I limited the solute transport predictions to chloride since it is an indicator of the septic plume. Its low reactivity and low plant uptake make it a much easier solute to model than others present in OWTS (as N or P). Please note that since Cl^- is a conservative anion, it will behave differently than N, and P. Anyway, it can help understand the physics of plume development and movement.

For the first part of the study, I concluded that soil textures in the Piedmont region in Georgia can be predicted from soil electrical resistivity, depth, and antecedent rainfall using an ANN framework. The model, based on a single site, accurately estimated the textures in a different location. The validation of the model should be repeated in an even broader area of the Piedmont, to evaluate the extent of its ability. Furthermore, the predictions within the saprolite layer could be improved by including information about features coarser than sand, that for this study were not taken into consideration. Those features could give better insight into the structure and porosity of the saprolite, rather than characterizing it only for its texture.

Moving to the second question, I determined that both models performed well when simulating changes in pressure head and that the main differences were ascribable to different transmissivity values

of the two resulting aquifers (related to the depth to bedrock). For chloride transport, the study showed that, when using a single porosity transport model, increased level complexity for soil layering and weathering state seemed not to generate the funnel-type preferential flow in the saprolite. Rather, even in this case, the depth to bedrock influenced the development and extent of the plume. It is important, though, to keep in mind that our ANN model was limited to a textural description of the saprolite layers. Additional information on the structure and the actual porosity of these layers could have produced different solutions since the presence of features larger than sand can generate soil macroporosity. This observation reiterates the value of including particles larger than sand when describing the subsurface and its properties. For both models, the predicted chloride concentrations were lower than the observed ones. The lower predictions could be caused by the input concentration used in the model. The value was based on a single measurement from the septic tank, that appears to be lower than what found in the literature. Unfortunately, we were not able to repeat the measurement. Nevertheless, the models performed better than the ones explored in Ceolin (2019), thus showing the importance of a calibration starting from measured parameters and a validation based on an independent set of data.

Most of the observed chloride concentrations were higher than the input measured in the septic tank. At the shoreline, this phenomenon could also be linked to external sources of solute that were not accounted for (possibly the interference of a plume from the neighboring house that was intercepted by our wells). This occurrence shows that, when studying the effects and loads of nutrients from OWTS, isolating inputs from a single system can be challenging if, in some cases, not possible. Therefore, it is important to consider the possible occurrence of overlapping plumes and the need to monitor, sample and model several neighboring houses at the same time. I exclude that the highest concentrations measured at the shoreline were caused by water coming from the lake, as they were all in a restricted area and a separate study taking samples from the lake was measuring much lower concentrations.

The temporal variations in the chloride observations were much larger than the chloride simulations indicating the possibility of: 1) the effect of variable chloride inputs from the OWTS, 2) the presence of macropore preferential flow that was not considered in our models. On this point, the

complex model appeared to better predict the temporal variations in chloride content, while the simple model, with lower RMSE values, was more accurate in determining a total load of chloride from the septic system to the shoreline. Additionally, the proven effect of the position of the bedrock in the development of the plume, it seems necessary to explore alternative geophysical methods to have more insights into its depth and shape.

Differently from chloride, phosphorus presents the possibility of colloid facilitated transport which could cause it to move more rapidly and to greater distances than those anticipated through a soil matrix. For this reason, I think that the inclusion of preferential flow would be even more critical when trying to model this nutrient for an accurate estimate of the load to the lake.

APPENDIX A

LISTS OF SYMBOLS AND THEIR PHYSICAL QUANTITIES

Physical quantities

M: mass

L: length

T: time

I: electric current

Symbols

V: voltage [$M\ L^2\ T^{-3}\ I^{-1}$]

R: electrical resistance [$M\ L^2\ T^{-3}\ I^{-2}$]

ρ: electrical resistivity [$M\ L^2\ T^{-3}\ I^{-2}$]

K(h): unsaturated hydraulic conductivity [$L\ T^{-1}$]

h: pressure head [L]

$\theta(h)$: volumetric water content, dimensionless but sometimes expressed as [$L^3\ L^{-3}$]

t: time [T]

z: vertical dimension [L]

c: solute concentration [$M\ L^{-3}$]

D_e: effective dispersion coefficient [$L^2\ T^{-1}$]

v: mean pore water velocity [$L\ T^{-1}$]

S_e: effective soil water saturation, dimensionless

l: pore-connectivity parameter, dimensionless

m: fitting parameter, dimensionless

n: fitting parameter, dimensionless

α: fitting parameter, [L^{-1}]

K_s: saturated hydraulic conductivity [$L\ T^{-1}$]

θ_r: residual volumetric content, dimensionless but sometimes expressed as [$L^3\ L^{-3}$]

θ_s: saturated volumetric water content, dimensionless but sometimes expressed as [$L^3\ L^{-3}$]

λ_L: lateral dispersivity [L]

λ_T: transverse dispersivity [L]

ρ_b: bulk density [M L^{-3}]

ρ_s: soil particle density [M L^{-3}]

K_d: adsorption isotherm coefficient [L^3 M^{-1}]

T: transmissivity [L^2 T^{-1}]

K: hydraulic conductivity [L T^{-1}]

B: thickness of the aquifer [L]

APPENDIX B

ERT RESULTS FROM TRANSECTS 1 AND 2 AT SITES 1, 2, AND FROM TRANSECT 2 AT SITE 3

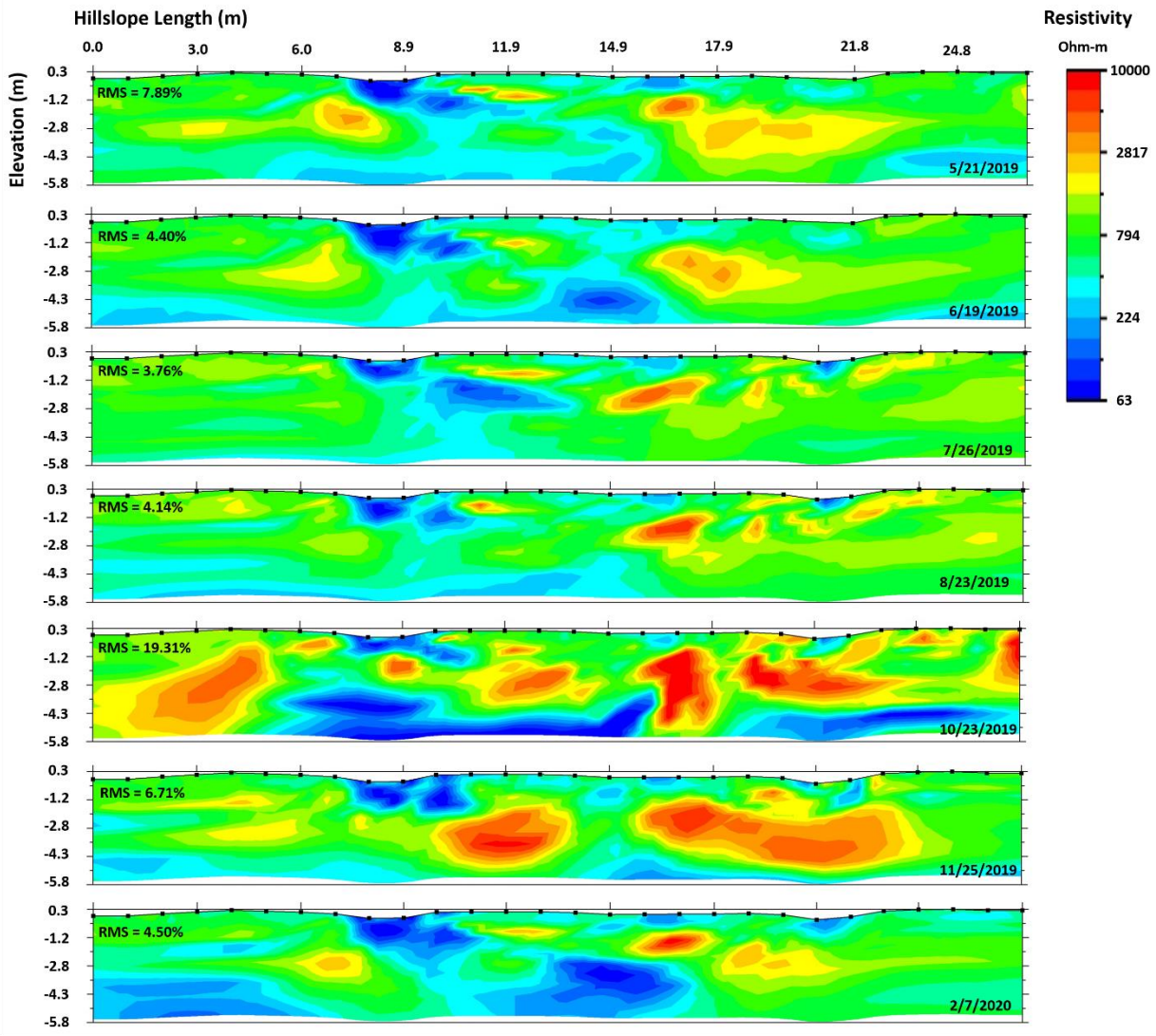


Figure B.1. Inverted resistivity values measured on the transect 1 (parallel to the lake) at site 1, from May 2019 to February 2020.

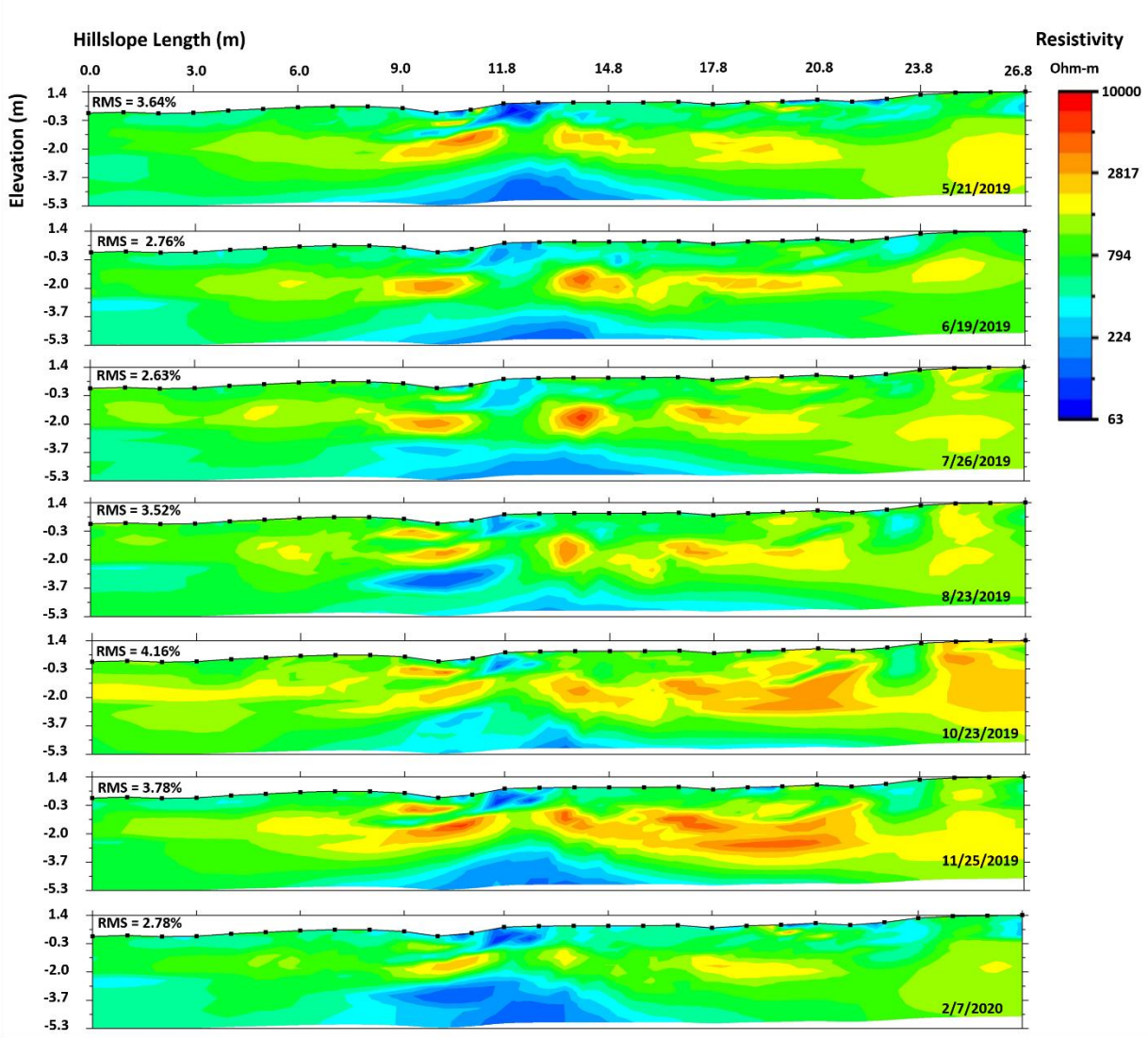


Figure B.2. Inverted resistivity values measured on the transect 2 (parallel to the lake) at site 1, from May 2019 to February 2020.

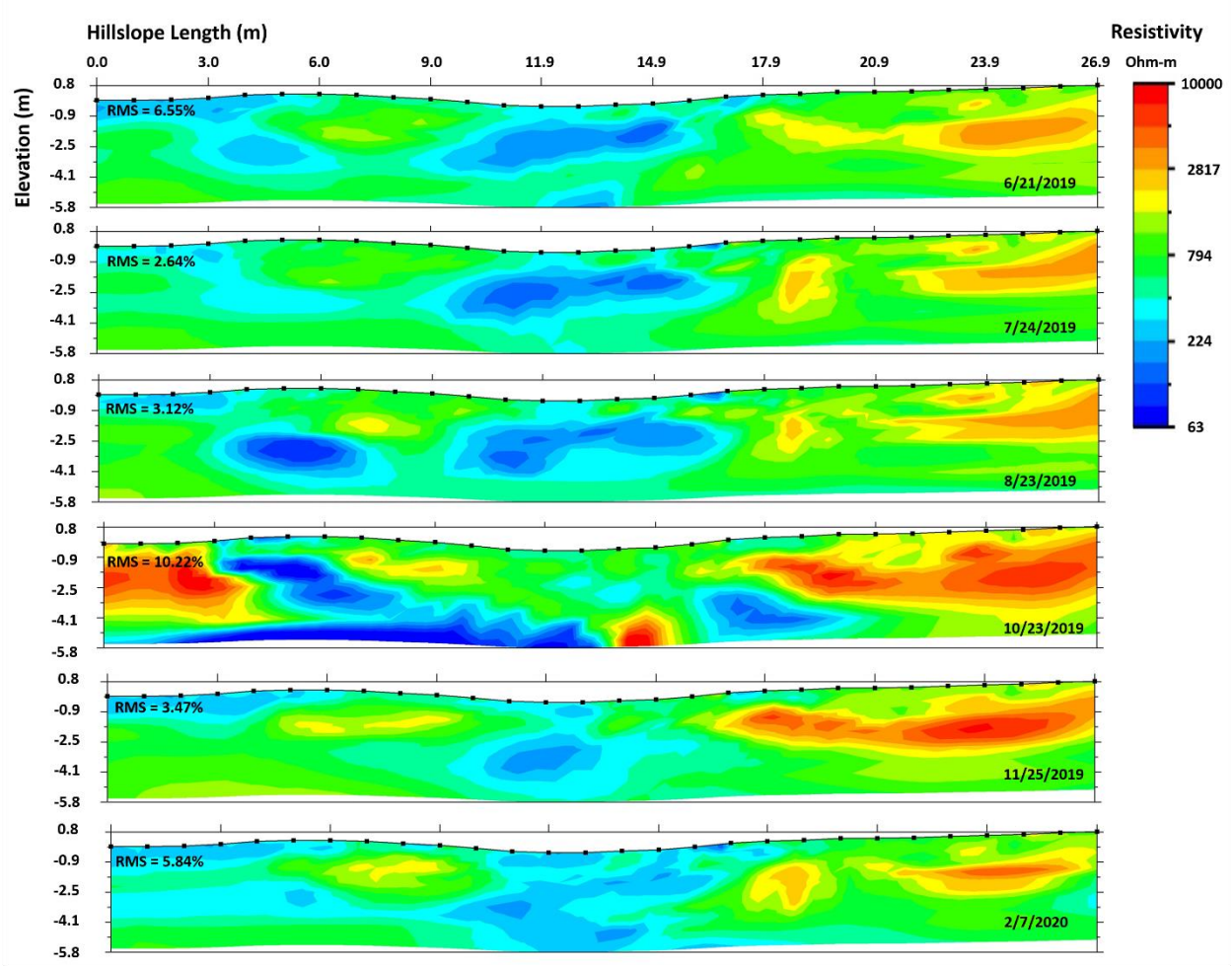


Figure B.3. Inverted resistivity values measured on the transect 1 (parallel to the lake) at site 2, from June 2019 to February 2020.

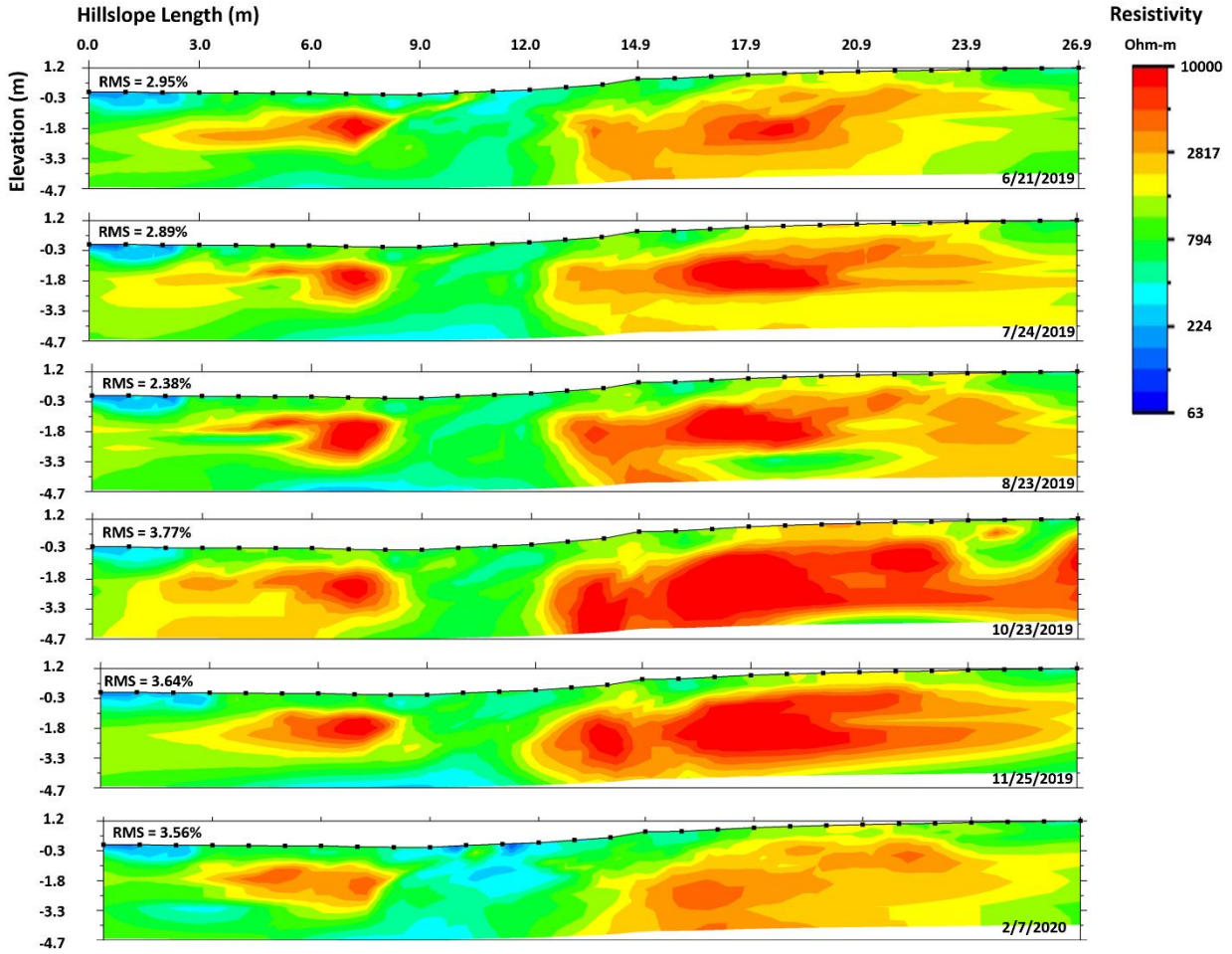


Figure B.4. Inverted resistivity values measured on the transect 2 (parallel to the lake) at site 2, from June 2019 to February 2020. For this transect, 21 to 30% of the data has been removed, for every transect, as they were categorized as "noisy data" from the inversion software.

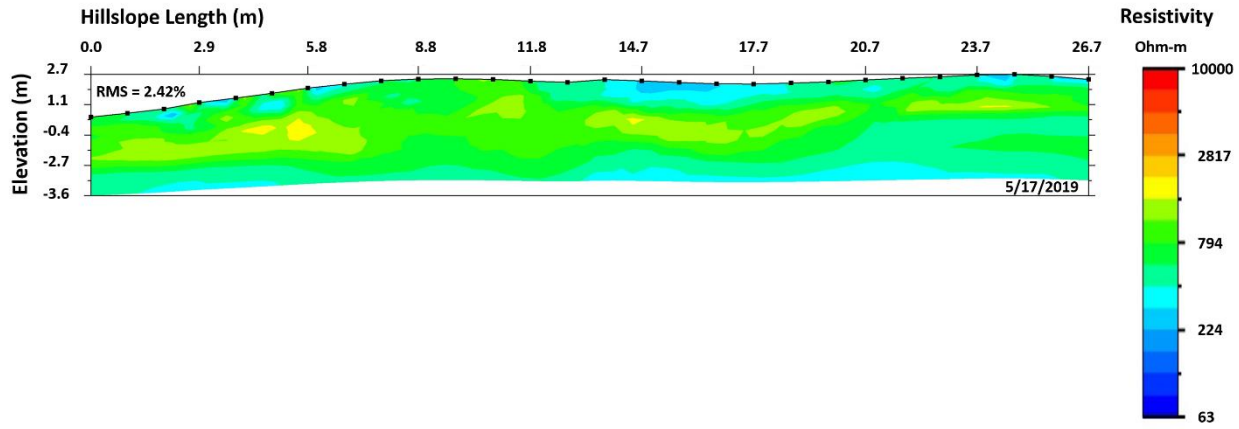


Figure B.5. Inverted resistivity values measured on the transect 2 (parallel to the lake) at site 3, in May 2019.

APPENDIX C

RESULTS OF THE LASER PARTICLE SIZE ANALYSER FOR SITES 1, 2 AND 3

Table C.1. Results of texture determination performed using the Laser Particle Size Analyzer performed on the soil cores collected at site 1.

Sample	*Clay %*	*Sand %*	*Silt %*	*Depth increments (cm)*
C04.5	*25.4*	*28.6*	*46.0*	*0 - 30*
C04.5	*31.0*	*32.2*	*36.8*	*30 - 60*
C04.5	*22.2*	*44.7*	*33.1*	*60 - 90*
C04.5	*19.7*	*47.0*	*33.3*	*90 - 120*
C04.5	*16.1*	*60.4*	*23.4*	*120 - 150*
C04.5	*19.0*	*60.9*	*20.2*	*150 - 180*
C04.5	*8.5*	*66.9*	*24.6*	*180 - 210*
C04.5	*12.2*	*60.7*	*27.1*	*210 - 240*
C04.5	*11.1*	*56.6*	*32.2*	*240 -270*
C04.5	*4.1*	*77.9*	*18.0*	*270 - 300*
C04.5	*4.2*	*77.8*	*17.9*	*300 - 360*
C04.5	*4.6*	*71.2*	*24.3*	*300 - 360*
C04.5	*5.1*	*71.6*	*23.3*	*360 - 390*
C04.5	*3.1*	*74.9*	*22.0*	*390 - 420*
C04.5	*3.7*	*72.2*	*24.1*	*420 - 450*
C04.5	*3.4*	*72.2*	*24.4*	*450 - 480*
C04.5	*2.9*	*74.8*	*22.3*	*480 - 495*
C14.5	*18.90*	*36.32*	*44.78*	*0 - 30*
C14.5	*30.77*	*28.86*	*40.37*	*30 - 60*
C14.5	*33.33*	*28.89*	*37.78*	*60 - 90*
C14.5	*29.10*	*36.33*	*34.57*	*90 - 120*
C14.5	*25.37*	*46.24*	*28.40*	*120 - 150*

C14.5	*17.68*	*57.92*	*24.40*	*150 - 180*
C14.5	*17.16*	*57.89*	*24.95*	*180 - 210*
C14.5	*4.32*	*66.90*	*28.78*	*210 - 240*
C14.5	*8.74*	*63.33*	*27.94*	*240 - 270*
C14.5	*8.03*	*65.26*	*26.70*	*270 - 300*
C14.5	*8.47*	*66.82*	*24.71*	*300 - 330*
C14.5	*7.50*	*67.21*	*25.29*	*330 - 360*
C14.5	*5.69*	*66.56*	*27.75*	*360 - 390*
C14.5	*4.14*	*77.92*	*17.93*	*390 - 420*
C18.5	*20.92*	*35.30*	*43.78*	*0 - 30*
C18.5	*33.40*	*28.07*	*38.53*	*30 - 60*
C18.5	*29.62*	*32.27*	*38.11*	*60 - 90*
C18.5	*23.49*	*39.23*	*37.27*	*90 - 120*
C18.5	*15.54*	*49.06*	*35.40*	*120 - 150*
C18.5	*8.84*	*58.21*	*32.95*	*150 - 180*
C18.5	*6.37*	*64.11*	*29.52*	*180 - 210*
C18.5	*4.39*	*70.73*	*24.88*	*210 - 240*
C18.5	*5.00*	*70.23*	*24.77*	*240 - 270*
C18.5	*2.80*	*78.91*	*18.28*	*270 - 300*
C18.5	*3.81*	*77.53*	*18.65*	*300 - 330*
C23	*19.03*	*34.83*	*46.14*	*0 - 30*
C23	*30.73*	*32.35*	*36.93*	*30 - 60*
C23	*28.29*	*41.27*	*30.44*	*60 - 90*
C23	*25.47*	*51.29*	*23.25*	*90 - 120*
C23	*28.16*	*47.69*	*24.15*	*120 - 150*

C23	*23.44*	*63.26*	*13.29*	*150 - 180*
C23	*10.28*	*69.16*	*20.55*	*180 - 210*
C23	*5.56*	*80.60*	*13.84*	*210 - 240*
C23	*6.34*	*77.95*	*15.70*	*240 - 270*
C23	*4.35*	*77.92*	*17.73*	*270 - 300*
C23	*6.06*	*74.55*	*19.39*	*300 - 330*
C25.9	*20.76*	*36.14*	*43.10*	*0 - 30*
C25.9	*33.59*	*36.99*	*29.41*	*30 - 60*
C25.9	*37.28*	*45.31*	*17.41*	*60 - 90*
C25.9	*27.01*	*56.56*	*16.44*	*90 - 120*
C25.9	*27.00*	*55.96*	*17.04*	*120 - 150*
C25.9	*16.87*	*56.56*	*26.57*	*150 - 180*
C25.9	*21.28*	*62.62*	*16.09*	*180 - 210*
C25.9	*8.29*	*79.02*	*12.69*	*210 - 240*
C25.9	*13.89*	*62.32*	*23.79*	*240 - 270*
C28	*22.35*	*30.46*	*47.19*	*0 - 30*
C28	*38.16*	*31.90*	*29.93*	*30 - 60*
C28	*36.15*	*42.82*	*21.03*	*60 - 90*
C28	*13.12*	*75.43*	*11.44*	*90 - 120*
C28	*12.91*	*74.79*	*12.30*	*120 - 150*
C29.7	*23.32*	*35.20*	*41.48*	*0 - 30*
C29.7	*33.83*	*47.63*	*18.54*	*30 - 60*
C29.7	*30.60*	*52.66*	*16.75*	*60 - 90*
C29.7	*20.52*	*62.93*	*16.55*	*90 - 120*
C29.7	*15.05*	*68.87*	*16.09*	*120 - 150*

C29.7	*6.00*	*77.65*	*16.35*	*150 - 175*
C31	*12.49*	*56.70*	*30.80*	*0 - 30*
C31	*23.25*	*55.63*	*21.12*	*30 - 60*
C31	*21.03*	*49.33*	*29.64*	*60 - 90*
C31	*12.70*	*51.77*	*35.52*	*90 - 120*
C31	*13.00*	*49.15*	*37.85*	*120 - 150*
C31	*5.24*	*71.71*	*23.04*	*150 - 180*
C31	*5.33*	*61.01*	*33.66*	*180-210*
C34	*18.91*	*29.58*	*51.51*	*0 - 30*
C34	*26.82*	*29.85*	*43.32*	*30 - 60*
C34	*19.30*	*45.04*	*35.66*	*60 - 87*
C35.5	*19.15*	*36.15*	*44.70*	*0 - 30*
C35.5	*26.09*	*47.20*	*26.70*	*30 - 60*
C35.5	*12.28*	*67.35*	*20.37*	*60 - 67*
C38	*21.95*	*27.68*	*50.37*	*0 - 30*
C38	*27.02*	*30.43*	*42.56*	*30 - 60*
C38	*21.06*	*35.90*	*43.03*	*60 - 90*
C38	*19.75*	*58.95*	*21.30*	*90 - 120*
C38	*14.44*	*70.35*	*15.21*	*120 - 150*
C38	*6.37*	*82.07*	*11.56*	*150 - 180*
C38	*10.33*	*74.79*	*14.88*	*180 - 210*
C38	*4.65*	*82.56*	*12.79*	*210 - 240*
C38	*2.28*	*80.55*	*17.16*	*240 - 270*
C38	*4.17*	*80.02*	*15.81*	*270 - 300*
C38	*2.94*	*76.72*	*20.34*	*300 - 330*

C38	*2.08*	*78.91*	*19.02*	*330 - 360*
C38	*1.92*	*82.02*	*16.06*	*360 - 390*
C38	*4.58*	*73.41*	*22.01*	*390 - 420*
C38	*3.70*	*71.92*	*24.38*	*420 - 450*
C38	*4.36*	*78.97*	*16.67*	*450 - 480*
C38	*3.15*	*81.65*	*15.19*	*480 - 495*

Table C.2. Results of texture determination performed using the Laser Particle Size Analyzer performed on the soil cores collected at site 2.

Sample	***Clay %***	***Sand %***	***Silt %***	***Depth increments (cm)***
C02.5	*15.8*	*55.85*	*28.35*	*0-30*
C02.5	*30.9*	*41.35*	*27.75*	*30-60*
C02.5	*29.6*	*38.9*	*31.5*	*60-90*
C02.5	*33.45*	*34.8*	*31.75*	*90-120*
C02.5	*37.05*	*33.9*	*29.05*	*120-150*
C02.5	*17.15*	*52.2*	*30.65*	*150-180*
C02.5	*23.2*	*44.7*	*32.1*	*180-210*
C02.5	*26.8*	*39*	*34.2*	*210-240*
C02.5	*28.35*	*37.15*	*34.5*	*240-270*
C02.5	*34.15*	*32.05*	*33.8*	*270-300*
C02.5	*33.6*	*32.8*	*33.6*	*300-330*
C02.5	*14.8*	*53.4*	*31.8*	*330-360*
C02.5	*12*	*57.8*	*30.2*	*360-390*
C02.5	*13.6*	*62.85*	*23.55*	*390-395*

C10	*15.45*	*55.4*	*29.15*	*0-30*
C10	*18.2*	*57.35*	*24.45*	*30-60*
C10	*17.95*	*53.45*	*28.6*	*60-90*
C10	*32.85*	*36.55*	*30.6*	*90-120*
C10	*36*	*30.5*	*33.5*	*120-150*
C10	*39.65*	*30.25*	*30.1*	*150-180*
C10	*37.6*	*30.45*	*31.95*	*180-210*
C10	*52.6*	*21*	*26.4*	*210-240*
C10	*56.6*	*19.8*	*23.6*	*240-270*
C10	*29*	*46.25*	*24.75*	*270-300*
C10	*30.2*	*40.55*	*29.25*	*300-330*
C10	*29.35*	*39.9*	*30.75*	*330-360*
C10	*32.1*	*39.8*	*28.1*	*360-390*
C10	*32.45*	*37.8*	*29.75*	*390-420*
C10	*30.2*	*41.95*	*27.85*	*420-450*
C18	*16.95*	*54.95*	*28.1*	*0-30*
C18	*20.05*	*51.65*	*28.3*	*30-60*
C18	*25.50*	*43.55*	*30.95*	*60-90*
C18	*42.25*	*27.70*	*30.05*	*90-120*
C18	*46.05*	*24.65*	*29.3*	*120-150*
C18	*47.85*	*22.70*	*29.45*	*150-180*
C18	*50.90*	*22.10*	*27*	*180-210*
C18	*39.25*	*31.30*	*29.45*	*210-240*
C18	*23.15*	*48.15*	*28.7*	*240-270*
C18	*41.40*	*32.35*	*26.25*	*270-300*

C18	*49.05*	*22.75*	*28.2*	*300-330*
C18	*45.65*	*25.15*	*29.2*	*330-360*
C18	*49.35*	*21.95*	*28.7*	*360-390*
C18	*31.80*	*46.00*	*22.2*	*390-420*
C18	*5.17*	*78.45*	*16.385*	*420-450*
C26	*19.60*	*52.40*	*28.00*	*0-30*
C26	*31.90*	*38.65*	*29.45*	*30-60*
C26	*43.00*	*28.15*	*28.85*	*60-90*
C26	*42.90*	*25.85*	*31.25*	*90-120*
C26	*13.80*	*60.75*	*25.45*	*120-150*
C26	*2.71*	*87.60*	*9.70*	*150-180*
C26	*3.45*	*86.90*	*9.65*	*180-210*
C26	*2.70*	*86.65*	*10.65*	*210-240*
C26	*3.41*	*86.85*	*9.74*	*240-270*
C26	*3.71*	*84.45*	*11.85*	*270-300*
C26	*3.22*	*86.80*	*9.98*	*300-330*
C26	*6.40*	*81.05*	*12.56*	*330-360*
C26	*3.10*	*87.15*	*9.75*	*360-390*
C26	*2.03*	*85.10*	*12.88*	*390-420*
C26	*2.78*	*86.05*	*11.18*	*420-450*

Table C.3. Results of texture determination performed using the Laser Particle Size Analyzer performed on the soil cores collected at site 3.

Sample	*Clay %*	*Sand %*	*Silt %*	*Depth increments (cm)*
C12	*33.35*	*23.85*	*42.80*	*0-30*
C12	*24.25*	*40.00*	*35.75*	*30-60*
C12	*16.10*	*45.55*	*38.35*	*60-90*
C12	*9.31*	*66.40*	*24.29*	*90-120*
C12	*6.36*	*70.10*	*23.54*	*120-150*
C12	*4.45*	*76.35*	*19.20*	*150-180*
C12	*3.59*	*77.85*	*18.56*	*180-210*
C12	*4.47*	*72.25*	*23.29*	*210-240*
C12	*5.54*	*69.25*	*25.21*	*240-270*
C12	*4.67*	*75.15*	*20.19*	*270-300*
C12	*3.75*	*80.70*	*15.56*	*300-330*
C17	*27.55*	*31.40*	*41.05*	*0-30*
C17	*29.35*	*45.15*	*25.50*	*30-60*
C17	*15.35*	*53.20*	*31.45*	*60-90*
C17	*8.27*	*69.95*	*21.79*	*90-120*
C17	*14.65*	*56.45*	*28.90*	*120-150*
C17	*18.45*	*51.35*	*30.20*	*150-180*
C17	*9.78*	*59.60*	*30.62*	*180-210*
C17	*9.77*	*57.30*	*32.93*	*210-240*
C17	*6.73*	*65.00*	*28.27*	*240-270*
C17	*7.25*	*54.20*	*38.55*	*270-300*
C17	*3.38*	*75.90*	*20.72*	*300-330*

C17	*2.59*	*82.25*	*15.16*	*330-360*
C17	*3.18*	*81.35*	*15.47*	*360-390*
C17	*2.98*	*83.75*	*13.27*	*390-420*
C17	*3.61*	*79.95*	*16.44*	*420-450*
C17	*3.54*	*79.85*	*16.62*	*450-480*
C17	*4.16*	*77.40*	*18.45*	*480-510*
C17	*8.77*	*60.70*	*30.54*	*510-540*
C17	*6.49*	*67.60*	*25.92*	*540-570*
C17	*7.75*	*63.45*	*28.81*	*570-600*
C17	*7.15*	*61.25*	*31.61*	*600-615*

APPENDIX D

PREDICTING WATER TABLE ELEVATIONS FROM SOIL RESISTIVITY USING BINOMIAL REGRESSION

Overview

Onsite Wastewater Treatment Systems (OWTS), or septic systems, are a decentralized method of treating wastewater. In the State of Georgia, 600,000 septic tank systems are already in use, over 30,000 new ones are installed every year, and about 10,000 of the total undergo repairs annually due to failure (Ade O. Oke, 2019). System failures can generate economic losses, threaten public health and the environment, and cause soil and water quality degradation (Gunady, Shishkina et al., 2015). To prevent and manage those problems, it is crucial to monitor the state of the OWTS, the diffusion of pathogens and the nutrient load to water bodies (Lewis, Wurtsbaugh et al., 2011; Schindler, 1971).

In the Lake Sidney Lanier watershed, within the Piedmont Region, the growing urbanization is causing an increase in the number of OWTS. Furthermore, in 2017, the final Total Maximum Daily Load for Lake Sidney Lanier suggested that nutrient inputs, including those from OWTS, may need to be reduced, thus creating the necessity to assess the inputs from them (The Georgia Department of Natural Resources Environmental Protection Division Atlanta, 2017).

Determining the contribution from septic systems is challenging due to their diffuse nature. In addition, in the peculiar geologic conditions of the region, sub-surface flowpaths, that are one of the main mechanisms of nutrient transport to surface water bodies, could be especially varied from idealized systems. Hence, data from conventional methods of water sampling and water table level determination through isolated installed wells and piezometers (USDA, 2019; USGS, 2019) could be less efficient for predicting transport and discharge processes (Englert, Kemna et al., 2016). Proper soil characterization is also crucial for the creation of an accurate hydrological model.

Aware of the complexity of the Piedmont and of the limitations of current techniques, we explored non-destructive methods as a possibly more time-efficient and cost-effective alternative. This research is studying the potential of Electrical Resistivity Tomography (ERT) in determining parameters to model nutrient movement at the hillslope scale in the Piedmont Region of Georgia. ERT is a geophysical technique that provides a spatially distributed high-resolution image of the resistivities of the sub-surface. The success of this technique is attributed to the variable electrical conductivities of the

different sub-surface geologic materials (also a function of soil water content). Using a simplified approach, we can consider the total resistivity signal measured on the soil as composed of a dynamic (water and ions content) and a static (soil texture) component. We focused on the extraction of the contribution of each of these individual components. Soil texture was characterized using Artificial Neural Network at a horizontal spatial resolution of 0.5 m and an average vertical resolution of 0.47 m up to a depth of about 6 m. We then predict information regarding water table elevation.

Materials and Methods

The project takes place in 3 adjacent coves in the southern part of Lake Sidney Lanier. A total of five lake-shore houses have been selected as study sites, based on three selection parameters: age of the OWTS, distance from the shoreline, and water use. Further details about the determination of the study sites can be found in Ceolin (2019). Shoreline and upslope monitoring wells have been installed at each site to intercept the nutrient plume coming from the OWTS and monitor fluctuations in water table elevation over the months.

At all sites, with minor changes, ERT measurements were taken on three transects to visualize the approximate spatial variations of the entire area. Two transects were parallel to the shoreline, while one was perpendicular to it. In order to separate the static and dynamic components of the resistivity data, as well as observing seasonal resistivity variations, at two sites monthly measurements were made over the same transect for six months. Monitoring wells were drilled adjacent to the ERT transects to collect ground truth data to generate the water table elevation (Figure D.1).

We employed binomial logistic regression analysis to detect water table elevation at the different sites. This type of regression allows predicting a dichotomous (1 or 0 type) dependent variable from one or several independent variables. For our model, the dependent variable was soil saturation with respect to the depth of the water table. A value of 1 indicated saturated soil, while 0 was unsaturated soil, respectively below or above the water table. The independent variables were:

1. Resistivity values (Figure D.2);

2. Specific depth at which the resistivity value has been measured;
3. Weekly antecedent rainfall in mm, from the USGS Water Data for USA database.

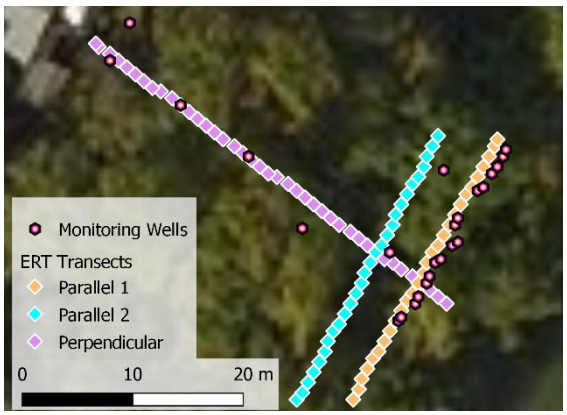


Figure D.1. Monitoring wells and transects where resistivity data were collected with the ERT at one of the study sites.

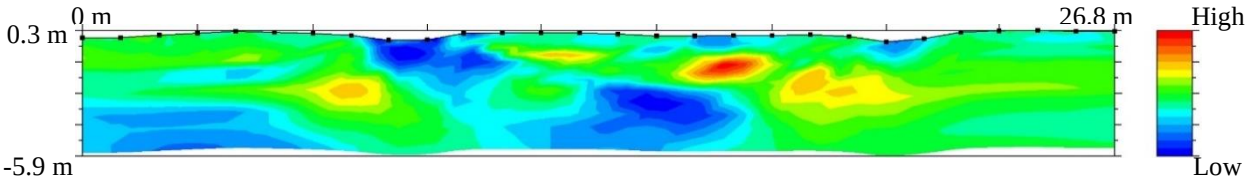


Figure D.2. Example of inverted resistivity data collected with the ERT at the transect "Parallel 1" from Figure D.1. The color scale shows the resistivity values in Ω-m. Blue is low resistivity and red is high resistivity.

Preliminary results of the water table model

After running our regression analysis, we checked the coefficients obtained for every independent variable and their significance for the model, using P-values, to decide whether we had to exclude any of them (Table D.1).

Table D.1. Binomial regression coefficients for every independent variable and relative significance. Significance codes: very highly significant (<0.001) '***'; highly significant (<0.01) '**'; significant (<0.05) '*'.

	Intercept	***Depth***	***Resistivity***	***Antecedent Rainfall***
Coefficient value	*-3.69 ****	*-1.78 ****	*-0.0003 ****	*-0.0205 **

So far, the model has been tested at two independent study sites in the vicinity of Lake Sidney Lanier. Table D.2 shows values of accuracy and error for the model, considering together the predictions for both sites.

Table D.2. Covariance matrix showing true positives and false positive. Total observed saturated is 533 and total observed unsaturated is 455.

	Observed Saturated	*Observed Unsaturated*
Predicted Saturated	*369/533 (69.23%)*	*27/445 (6.07%)*
Predicted Unsaturated	*164/533 (30.77%)*	*418/445 (93.93%)*

We compared our predictions with total number of points where soil was observed as saturated, 533 times, or unsaturated, 455 times, to create a covariance matrix. We also calculated the distance between the observed and predicted depth of the water table. The minimum error, registered at the shoreline, in the detection is 0.46 m, while the maximum error registered above the midslope, is 2.71 m. Figure 3 shows examples of predictions for a perpendicular (a) and a parallel transect (b).

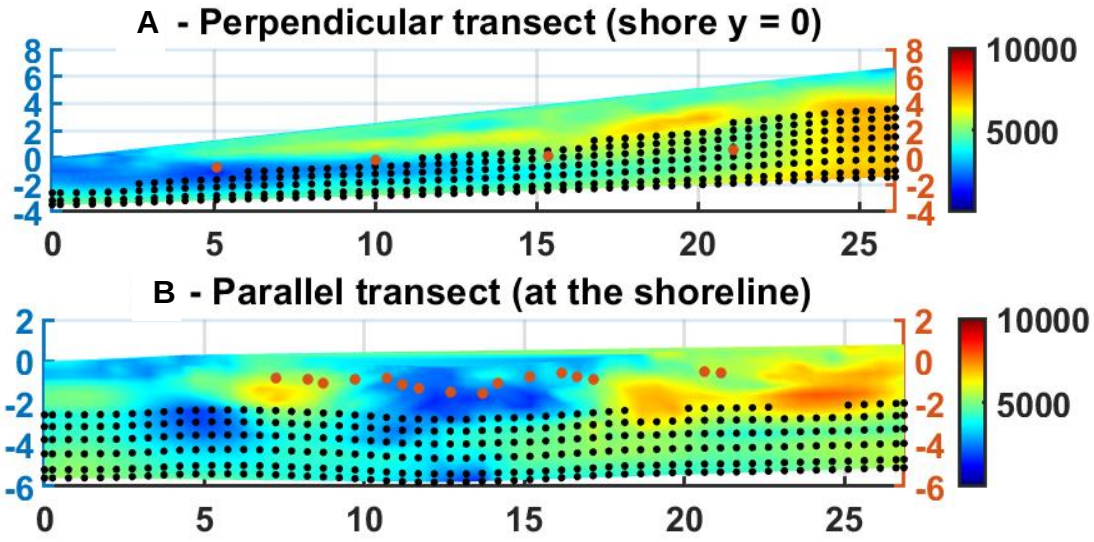


Figure D.3. Water table level predictions on a perpendicular (A) and a parallel transect (B). The black dots represent predicted saturated soil locations, while the red ones show the actual water table positions as measured in the monitoring wells. The background and the color scale show the inverted resistivity values in Ω-m at the time of the prediction. Blue is low resistivity and red is high resistivity (10000 can be assumed to be rock).

Conclusions and Future work

The model for the prediction of the water table level is promising. The errors in the detection might be influenced by the software grid for inversion and calculation of the resistivity data. We will also investigate influence of seasonality on the quality of the prediction.

Finally, after the soil texture (to be submitted for peer-reviewed publication) and water table model are developed, we should be able to extract the remainder of the resistivity signal, corresponding to water ionic strength. This could give us insights about ionic electrical conductivity at a high 2D spatial resolution and thus help determine groundwater contamination. Also, as was done in this study, we can use ERT information to determine the distribution of thick saprolite layers and identify locations for installing monitoring wells.

References

Ade O. Oke, R. (2019). *Manual for On-Site Sewage Management Systems*. Georgia Department of Public Health

Ceolin, S. (2019). *Preliminary Assessment of Onsite Wastewater Treatment System Effects on Lake Lanier Water Quality*. (Master of Science of Crop and Soil Sciences). University of Georgia and University of Padova,

Englert, A., Kemna, A., Zhu, J.-f., Vanderborght, J., Vereecken, H., & Yeh, T.-C. J. (2016). Comparison of smoothness-constrained and geostatistically based cross-borehole electrical resistivity tomography for characterization of solute tracer plumes. *Water Science and Engineering, 9*(4), 274-286. doi:10.1016/j.wse.2017.01.002

Gunady, M., Shishkina, N., Tan, H., & Rodriguez, C. (2015). A Review of On-Site Wastewater Treatment Systems in Western Australia from 1997 to 2011. *Journal of Environmental and Public Health, 2015*, 1-12. doi:10.1155/2015/716957

Lewis, W. M., Wurtsbaugh, W. A., & Paerl, H. W. (2011). Rationale for Control of Anthropogenic Nitrogen and Phosphorus to Reduce Eutrophication of Inland Waters. *Environmental Science & Technology, 45*(24), 10300-10305. doi:10.1021/es202401p

Schindler, D. W. (1971). Carbon, Nitrogen, and Phosphorus and the Eutrophication of Freshwater Lakes. *7*(4), 321-329. doi:10.1111/j.1529-8817.1971.tb01527.x

The Georgia Department of Natural Resources Environmental Protection Division Atlanta, G. (2017). *Total Maximum Daily Load Evaluation for Lake Lanier in the Chattahoochee River Basin for Chlorophyll a*.

USDA. (2019). Water Table Determinations. Retrieved from https://www.nrcs.usda.gov/

USGS. (2019). How can I find the depth to the water table in a specific location? Retrieved from https://www.usgs.gov/