USING PORTABLE X-RAY FLUORESCENCE TO PREDICT PHYSICAL AND CHEMICAL PROPERTIES OF CALIFORNIA SOILS

A Thesis

presented to

the Faculty of California Polytechnic State University,

San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Natural Resources and Environmental Sciences

by

Micaela Dyani Frye

August 2022

COMMITTEE MEMBERSHIP

TITLE:	Using Portable X-ray Fluorescence to Predict Physical and Chemical Properties of California Soils
AUTHOR:	Micaela Dyani Frye
DATE SUBMITTED:	August 2022
COMMITTEE CHAIR:	Gordon Rees, Ph.D. Associate Professor of Forest and Range Soils
COMMITTEE MEMBER:	Yamina Pressler, Ph.D. Assistant Professor of Soil Science and Restoration Ecology
COMMITTEE MEMBER:	Stewart Wilson, Ph.D. Assistant Professor of Digital Soil Mapping

ABSTRACT

Using Portable X-ray Fluorescence to Predict Physical and Chemical Properties of California Soils

Micaela Dyani Frye

Soil characterization provides the basic information necessary for understanding the physical, chemical, and biological properties of soils. Knowledge about soils can in turn be used to inform management practices, optimize agricultural operations, and ensure the continuation of ecosystem services provided by soils. However, current analytical standards for identifying each distinct property are costly and time-consuming. The optimization of laboratory grade technology for wide scale use is demonstrated by advances in a proximal soil sensing technique known as portable X-ray fluorescence spectrometry (pXRF). pXRF analyzers use high energy X-rays that interact with a sample to cause characteristic reflorescence that can be distinguished by the analyzer for its energy and intensity to determine the chemical composition of the sample.

While pXRF only measures total elemental abundance, the concentrations of certain elements have been used as a proxy to develop models capable of predicting soil characteristics. This study aimed to evaluate existing models and model building techniques for predicting soil pH, texture, cation exchange capacity (CEC), soil organic carbon (SOC), total nitrogen (TN), and C:N ratio from pXRF spectra and assess their fittingness for California soils by comparing predictions to results from laboratory methods. Multiple linear regression (MLR) and random forest (RF) models were created for each property using a training subset of data and evaluated by R^2, RMSE, RPD and RPIQ on an unseen test set. The California soils sample set was comprised of 480 soil samples from across the state that were subject to laboratory and pXRF analysis in GeoChem mode.

Results showed that existing data models applied to the CA soils dataset lacked predictive ability. In comparison, data models generated using MLR with 10-fold cross validation for variable selection improved predictions, while algorithmic modeling produced the best estimates for all properties besides pH. The best models produced for each property gave RMSE values of 0.489 for pH, 10.8 for sand %, 6.06 for clay % (together predicting the correct texture class 74% of the time), 6.79 for CEC (cmolc/kg soil), 1.01 for SOC %, 0.062 for TN %, and 7.02 for C:N ratio. Where R^2 and RMSE were observed to fluctuate inconsistently with a change in the random train/test splits, RPD and RPIQ were more stable, which may indicate a more useful representation of out of sample applicability. RF modeling for TN content provided the best predictive model overall ($R^2 = 0.782$, RMSE = 0.062, RPD = 2.041, and RPIQ = 2.96). RF models for CEC and TN % achieved RPD values >2, indicating stable predictive models (Cheng et al., 2021). Lower RPD values between 1.75 and 2 and RPIQ >2 were also found for MLR models of CEC, and TN %, as well as RF models for SOC. Better estimates for chemical properties (CEC, N, SOC) when compared to physical properties (texture), may be attributable to a correlation between elemental signatures and organic matter. All models were improved with the addition of categorical variables (land-use and sample set) but came at a great statistical cost (9 extra predictors). Separating models by land type and lab characterization method revealed some improvements within land types, but these effects could not be fully untangled from sample set. Thus, the consortia of characterizing bodies for 'true' lab data may have been a drawback in model performance, by confounding inter-lab errors with predictive errors. Future studies using pXRF analysis for soil property estimation should investigate how predictive models are affected by characterizing method and lab body. While statewide models for California soils provided what may be an acceptable level of error for some applications, models calibrated for a specific site using consistent lab characterization methods likely provide a higher degree of accuracy for indirect measurements of some key soil properties.

Keywords: Portable X-ray fluorescence spectrometry, soil reaction, soil texture, cation exchange capacity, soil organic carbon, prediction models, random forest, California soils

ACKNOWLEDGMENTS

A deep and sincere thanks to my advisor Dr. Gordon Rees, who gifted me with this once in a lifetime opportunity, asked the unique and important question from which this research was based, and was there every step of the way with advice and reassurance. I so appreciate your belief in me to carry out this project. From the bottom of my heart, thank you. It has been a true privilege to be mentored by you.

I'd also like to thank Dr. Yamina Pressler for impressing upon me what it means to be a scientist, in the pursuit of truth, and despite our own self-doubts. Thank you for understanding me and uplifting me when it was the most needed. Thanks to Dr. Stewart Wilson for teaching me about the new age wave of soil science where we can use satellites to predict environmental features on the ground with startling accuracy. Your advice about model construction was valuable in influencing the direction of this research.

A huge thank you to the Agricultural Research Institute and McIntire-Stennis Program for providing the financial resources to make this work possible.

Thank you to my first soil science professors and soil judging coaches. Dr. Brian Needelman, for showing me that soil science is an art, that can be expressed through pottery, music, and celebration. Dr. Martin Rabenhorst, for pushing me and our team to be the best we can and reminding us to always remember the universe.

Thank you to my friends and colleagues from USGS who introduced me to the world of environmental research and made me realize that sometimes teamwork means pulling your friends out of sinkholes. It was a joy navigating uncharted marshes and forests with you all.

Thank you, Peter Weiler, for looking out for me even before we had met, trusting me to help you, and being my friend.

I've relied heavily upon the support and encouragement of my friends and family to complete this project, especially my mom, who was always there to revive my spirits, cheer me on, and give me advice. To my sister Rhea, I would be lost without you, you're my best friend, always. To my niece, Kaia, I can't wait to share with you what I've learned about the secrets of nature and support you through all the passions you'll come to have.

Sean, thank you for being an unwavering source of support to me. It has truly made all the difference to have you by my side. I cherish you deeply, and I can't wait for what comes next.

Lastly, thank you to those explorers of the natural world— past, present, and future, from which knowledge continues to be built upon and with whom, together, we will preserve our most precious shared gift.

"The ancient teachers of this science," said he, "promised impossibilities, and performed nothing. The modern masters promise very little; they know that metals cannot be transmuted, and that the elixir of life is a chimera. But these philosophers, whose hands seem only made to dabble in dirt, and their eyes to pore over the microscope or crucible, have indeed performed miracles. They penetrate into the recesses of nature and show how she works in her hiding places. They ascend into the heavens; they have discovered how the blood circulates, and the nature of the air we breathe. They have acquired new and almost unlimited powers; they can command the thunders of heaven, mimic the earthquake, and even mock the invisible world with its own shadows."

Mary Shelly, *Frankenstein*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

BS	Base saturation
DL	Detection limit
EDXRF	Energy dispersive XRF
FFP	Fit-for-purpose
LOD	Limit of detection
MLR	Multiple linear regression
NRCS	Natural Resources Conservation Service
PLSR	Partial least squares regression
PFT	Pedotransfer function
pXRF	Portable X-ray fluorescence spectrometry
RF	Random forest
RMSE	Root mean square error
RPD	Residual prediction deviation
RPIQ	Ratio of performance to interquartile distance
SDD	silicon drift detector
SI	Sustainable intensification
SOC	Soil organic carbon
SON	Soil organic nitrogen
SVM	Support vector machines
SVR	Support vector regression
WDXRF	Wavelength dispersive XRF

Chapter 1

INTRODUCTION

Soils underpin life on Earth while also playing a major role in today's most pressing environmental challenges, making their effective management more important now than ever before. The sustainable preservation of land ensures that humans and the constructed environment can exist in harmony with the natural world. In essence, an understanding and full consideration of soil bodies is crucial for any successful endeavor where humans impart on the land. Soil characterization provides the basic information necessary for understanding the physical, chemical, and biological properties of soils, which we rely on for food security, maintaining forests and grasslands, and supporting the structures that mark modern human civilization. Analytical data about these soil properties is necessary for taxonomic classification schemes that help organize existing knowledge, foster clear communication, and provide a basis for interpreting the behavior of soils. In California, soils show great variation due to the state's diverse topography, underlying geology, and climatic conditions. Mapping soils across the landscape provides valuable information pertaining to fertility and suitability for a vast range of applications. Overcoming the broad generalizations that weaken predictive mapping requires a high density of ground truth data. This quantitative data allows for appropriate use of soil resources so that human needs can be met while ecological sustainability and environmental safeguards are preserved.

Regionally, changes in land cover and land-use can subdue or enhance effects of climate change (Jia et al., 2019). Therefore, the management decisions made by agriculturists, land managers, and governments of all scales within the coming decades will steer the trajectory of the global carbon balance. The concentration, quality, and dynamics of soil organic carbon (SOC) has a major impact on soil quality, functionality, and health (Lal, 2016). Since the largest share of terrestrial carbon is found in soils (Trumbore, 2009), continued degradation of soil health and loss of SOC could cause massive disruptions to the ecosystem services provided by soils— including water filtration, climate regulation, nutrient cycling, and food production (Garrett et al., 2018; Francaviglia et al., 2018). At smaller scales, implementing best management practices requires knowledge of various soil properties depending on the desired use. At grander scales, monitoring SOC to meet ambitious goals like increasing soil carbon stocks by 4% annually as called for by the "4 per 1000" initiative and improving social, economic, and ecological wellbeing outlined by the United Nations Sustainable Development goals will require streamlined sampling protocols which are reliable and accessible (Trivedi et al., 2018; Lal, 2016). There is evidence that effective and targeted soil management may lead to an increase in soil carbon content, water-holding capacity, and infiltration (Bos et al., 2017). Improving soil health can also decrease erodibility by providing protection against more frequent and severe weather events that contribute to soil erosion (Deryng, 2020). Appropriate soil management is therefore of vital importance for regions relying on maintained soil health for crop production, and to those areas dependent on the exports from agricultural hubs.

California's agricultural sector is large, dynamic, and provides huge economic value on a world-wide scale as a producer and exporter (CDFA, 2019). Despite this vigorous productivity, California's crucial role as the food production epicenter of the United States is threatened by climate change induced pressures, which have continually increased heat extremes and droughts in the state since the 1950s (IPCC, 2021). Changes in temperature, precipitation, snowpack, extreme heat events, and flooding resulting from climate change are expected to prompt extensive spatial shifts in cropland acreage in California (Pathak et al., 2018). In addition to agricultural land cover shifts, grasslands and forests also face significant risk. Continued population growth and urbanization in the state is expected to decrease forest and rangeland areas (Wilson et al., 2016). As a consequence, fragmentation of these wild ecosystems threatens native biodiversity. These anthropogenic land changes cause fundamental shifts in the energy balance of the land and surface-heat budget (Mölders, 2012) and play a major role in terrestrial carbon losses (House et al., 2005).

Deliberate attempts by humans to remove carbon dioxide from the atmosphere in conjunction with controlling emissions to reach "net zero" emissions could limit warming in the long term to 1.4 °C and advert the most catastrophic effects of climate change (IPCC, 2021). Soil carbon sequestration is one such carbon dioxide removal technique with great potential as a carbon sink for a low cost (< $0 - $100/ton) with an estimated reduction in CO2 concentrations between 2 and 5 gigatons/year by 2050 (IPCC, 2018). Soils act as both a buffer to increasing levels of atmospheric CO2 as well as a sink for carbon, but deterioration of land and loss of SOC stores poses ominous threats to ecosystem functioning and human livelihood (Trivedi et al., 2018). Therefore, optimal use of soil resources will focus on preservation, restoration, and informed stewardship by land-users. A growing interest in tracking SOC coupled with a sustained need for soil characterization to understand the current state of soil health is made possible with regular and accurate soil testing regimes. Initial simplistic methods to characterize soil features of interest have been refined over the previous century, giving way to more advanced and accurate methods (Weindorf and Chakraborty, 2020). In fact, current analytical standards often require specialized laboratory procedures for identifying each distinct property (Soil Survey Staff, 2014b). However, certain shortcomings of these methods are apparent. For instance, an accepted laboratory technique for sample elemental detection and quantification uses inductively coupled plasma (ICP) spectrometry. However, sample preparation for this method usually requires acid digestion with strong caustic chemicals including hydrochloric, nitric, and sometimes hydrofluoric acid (US EPA, 1996a; US EPA, 1996b), and can still result in incomplete digestion leading to measurement inaccuracies. Where these specialized methods are possible, traditional laboratory analysis for a range of soil properties can be a costly and time-consuming process, requiring sophisticated equipment and specially trained operators. The barriers to entry for analytical laboratory equipment compel a dire need for reproducible methods to quantify soil properties on a large scale in countries all over the world where soil information is often sparse or inadequate and access to reliable soil testing facilities is limited (Towett et al., 2015).

A critical aspect to soil-testing is ensuring samples are representative of the soil attributes within a field. The high variability of soil across a landscape necessitates more efficient ways to determine soil attributes than traditional approaches. Sensor based technologies offer the opportunity to increase soil knowledge. The optimization of laboratory grade technology for wide scale and field-based use is demonstrated by the advances in a proximal soil sensing technique known as portable X-ray fluorescence (pXRF). pXRF analyzers harness the power and reliability of benchtop XRF analysis in a sophisticated yet compact instrument that can be carried in a single hand and brought into the field. Onsite analysis of elements ranging in concentration from just a few ppm to 100% can be performed in about a minute with these devices. pXRF technology works by emitting high energy X-rays that excite electrons of different elements, causing them to be ejected from their inner shell positions. As outer shells electrons move to fill the inner shell void, a characteristic reflorescence is emitted from the sample (Sharma et al., 2014). The energy and intensity of the egressing fluorescence are measured as electric signals by the pXRF and translated into analytical data representing the specific elements and their concentrations present in the sample. High analytical precision of pXRF instruments (Hall et al., 2011; Goodale et al., 2012) and their close correlation to benchtop XRF (Shefsky, 1997; Guerra et al., 2014; Sarala, 2016) has proven the utility of this technology. The benefits of portable XRF transcend monetary and labor savings by offering real time decision support and increasing sample sizes to achieve a nuanced understanding of the soil environment. As the technology has evolved to become more powerful, the capabilities and utility of the instrument have also expanded—making pXRF analyzers a reliable tool used across a range of disciplines.

While pXRF only measures total elemental abundance, the concentrations of certain elements have been used to develop regression-based and algorithmic models that indirectly predict several different soil characteristics (Radu and Diamond, 2009). pXRF has shown incredible capacity to quickly and accurately predict key soil features such as pH (Sharma et al., 2014), cation exchange capacity (CEC) (Sharma et al., 2015), texture

(Benedet et al., 2020), gypsum quantification (Weindorf et al., 2013), horizonation (Weindorf et al., 2012), salinity (Swanhart et al., 2014), lithologic discontinuities (Weindorf et al., 2015), and C:N ratio (Towett et al., 2015). pXRF methods for soil characterization have the potential to produce rapid, reproducible, and cost-effective estimates with only minimal sample preparation. The draw of this research is achieving good estimates of important soil features without the inherent expense or traditional lag time between sampling and results. Additionally, high spatial resolution of in-situ measurements via pXRF means that a more reliable site assessment can be achieved than would be possible with fewer ex-situ measurements. Utilizing pXRF technology may help optimize agricultural operations in the immediate future by measuring the concentrations of elements important for soil fertility and deriving physical and chemical properties of interest via digital soil morphometrics (Stockmann et al., 2016). Tracking changes across time is necessary for meeting goals that require reliable carbon pools and carbon sequestration to be monitored (D'Amore and Kane, 2016) and also aids in adjusting land use management strategies proactively to discover which practices pay off in the long term.

Methods using pXRF to infer soil properties via regression analysis have produced successful predictions when applied to test samples within that sample set. However, the utility of these models when applied to soils from a different geographic range is unknown. Previous studies necessarily confined by sample availability have produced regionally calibrated and validated models, warranting ongoing investigation into formulating larger scale models. This study aims to evaluate and refine the predictive power of existing published models that use pXRF to characterize soil properties for use on Californian soils collected from throughout the state. If this technology is to be used for regulatory or monitoring purposes in the future, there must be a certain guarantee ofaccuracy and precision of the parameter estimates given. Even if pXRF estimates of soil properties prove to have a lower degree of accuracy than laboratory determinations, this could be compensated for by the higher density sampling permitted with pXRF analysis. Ideally, a mixture of lab characterized and pXRF predicted properties could give a holistic picture of the land with more reasonable time and economic inputs than the current standards. In addition, the overall economic impact of soil testing in California could be greatly reduced with a shift to indirect measurements of soil properties via pXRF analysis when compared to traditional laboratory testing.

pXRF analysis appears to be a timely and elegant solution for predicting a suite of soil properties while significantly cutting down on the volume of samples required for traditional lab analysis, but it's accuracy and limitations must be evaluated on larger spatial scale. The objectives of this study were as follows: (1) evaluate existing models and model building methods that predict soil pH, texture, cation exchange capacity (CEC), soil organic carbon, total nitrogen, and C:N ratio from pXRF spectra and assess their fittingness to California soils by comparing predicted values to results from traditional laboratory methods, (2) use pXRF analysis to create multiple linear regression and random forest models to predict these properties and, (3) assess how the categorical variables of land type and characterization methods affect estimates. To accomplish these objectives, a set of soil samples from across the state of California was characterized using conventional laboratory procedures and using a pXRF analyzer. Several models were created to link elemental concentrations to physical and chemical soil properties.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

California's agricultural sector has consistently achieved tremendous gains in export values (CDFA, 2020)— providing a diversity of crops, including several specialty crops, throughout the country and sustaining a considerable workforce. However, the effects of climate change are already accentuating variable weather events, inducing large scale land use shifts, and challenging existing crop cultivation systems. To cope with rising food demands coupled with fewer acres of arable land, more efficient and sustainable agricultural production systems will be necessary. Targeted soil management achieved through the use of advanced technology is one solution that can help meet these needs by basing management decisions off quantitative data. Achieving high-density, accurate, and timely characterization of soils by proximal sensing techniques is one area which has shown great promise for this cause. Over the last two decades, improvements in portable X-ray florescence spectrometry (pXRF) detectors and miniaturization of internal parts have transformed the niche analytical technique to a widely accepted tool for obtaining total elemental analyses. pXRF works by subjecting a sample to high energy X-rays which cause electrons to be expelled from their inner shell positions and replaced by outer shell electrons (Sharma et al., 2014). This process results in a re-emittance that detected by the pXRF as a characteristic fluorescence and translated into an analyte quantity based upon the unique energy and intensity of the spectral peaks (Bosco, 2013). Using the total elemental profile determined by pXRF (both alone and in tandem with other sensor data) as a proxy, various other physiochemical properties of interest can be estimated (Gozukara et al., 2022). Non-destructive sampling can be achieved both in and ex-situ— but abiding by certain best operating practices is important for ensuring data accurately reflect the chemical composition of the sample in question. This literature review will explore the challenges and opportunities faced by California, technological advancements for resource use management, the working principles behind portable X-ray fluorescence spectroscopy, its use for modeling soil properties, and pertinent operating principles and considerations for use.

2.2 Climate change impacts for California

2.2.1 Overview

California's fertile soils, Mediterranean climate, and extensive groundwater storage and delivery basins have allowed for the establishment and growth of a highly productive agricultural industry that serves as the backbone of the United States food supply. With the fifth largest economy in the world (IMF, 2021), this multi-billion-dollar sector is the nation's sole exporter of several agricultural commodities and specialty crops (CDFA, 2019). However, California faces significant vulnerability from impending climactic shifts lying outside the bounds of natural seasonal variability. The agricultural sector proliferates the amount of greenhouse gas emissions (GHGs) that contribute to climate change while also being disproportionately impacted by the effects of climate change (Deryng, 2020; Jackson et al., 2012), representing a critical challenge to humankind.

2.2.2 Land-use changes

In the state of California, developed urban and suburban areas coalesce with intensively managed farmlands, grazed rangelands, and protected ecosystems to create a diverse mosaic of land uses. Regardless of the preparedness of the state or current efforts to curtail emissions, climate change effects will trigger extensive shifts in land-use throughout California. Projected land use changes in the Central Valley indicate an increase in developed land cover (21,141 km^2/62.9%) and decreases in annual cropland (-30.3%) and rangeland (-7.3%) (Fig 2.1). These land use conversions coincide with population growth estimates in the state— from 39.5 million in 2021 to 44.2 million by 2060 (California Department of Finance, 2021a; 2021b).

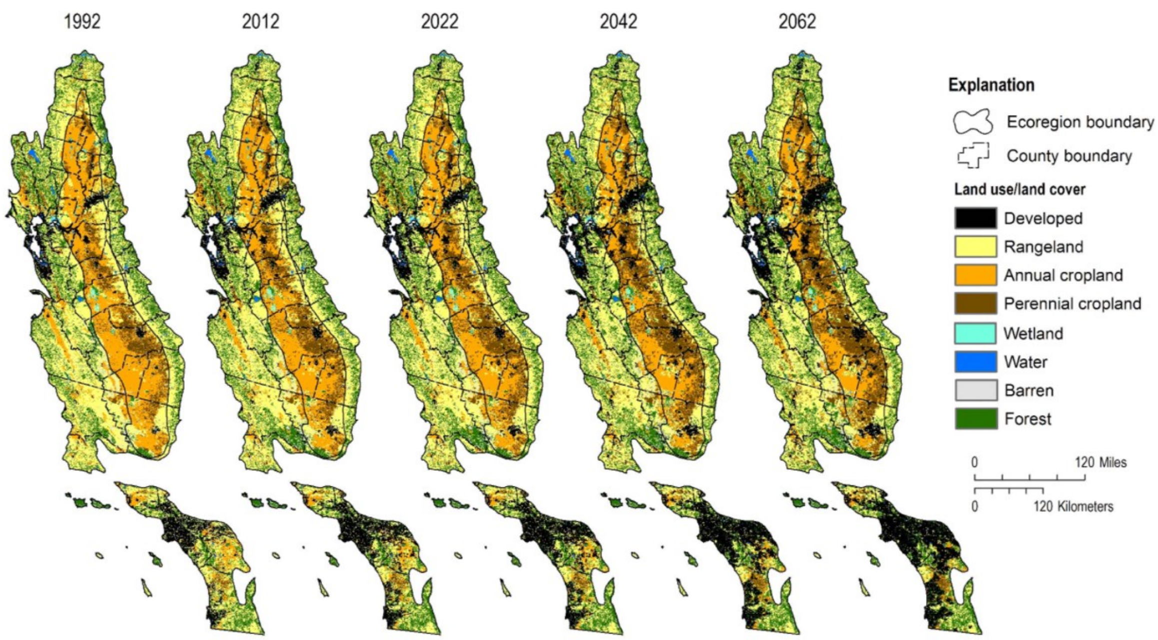

Figure 2.1: Historical and projected land use change in the Central California Foothills, Coastal Mountains, and Central Valley between 1992-2062 under a business-as-usual scenario (Excerpted from Wilson et al., 2016).

In a study by Jackson et al. (2012), a spatially explicit agricultural vulnerability index for the state of California was derived from a framework of 22 land-use, climate, crop, and socioeconomic variables. The Sacramento-San Joaquin Delta, Salinas Valley, the corridor between Merced and Fresno, and the Imperial Valley showed high agricultural vulnerability (Fig. 2.2). Authors suggested that adapting localized approaches could be a benefit in addressing the resiliency or vulnerability of different regions who could use this information in developing climate action plans.

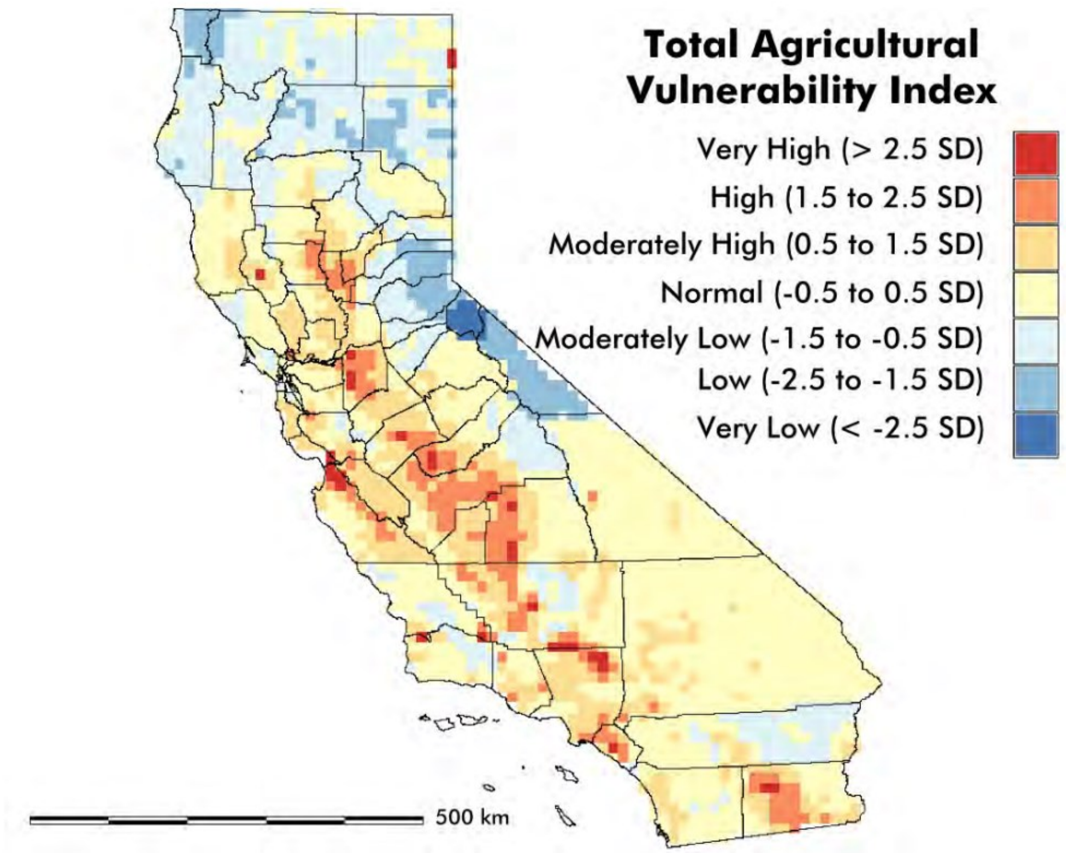

Figure 2.2: The Total Agricultural Vulnerability Index integrates Climate Vulnerability, Crop Vulnerability, Land Use Vulnerability and Socioeconomic Vulnerability to display regions of concern in California (Excerpted from Jackson et al., 2012).

The progressive development and inhabitation of California's urban areas is an extreme case of landcover change. An expansion of urban areas into wild areas results in the replacement of natural surfaces with artificial surfaces. These urban surfaces conduct and store heat, causing higher air temperatures than surrounding rural areas in a phenomenon called urban heat island (UHI) (Vahmani et al., 2016). Higher temperatures can in turn provoke higher energy consumption by the population, which further drives greenhouse gas emissions and warming. Furthermore, urban areas are major sources of carbon dioxide emissions, with up to 97% of human generated carbon emissions coming from cities (Svirejeva-Hopkins et al., 2004). Another consequence of urbanization is the reduction in natural land area serving as terrestrial carbon sinks (Grimmond, 2007). Urbanized areas also change the nature of carbon cycling, with around 50% of carbon from net primary productivity horizontally redistributed to other ecosystems, where decomposition conditions are variable (Svirejeva-Hopkins et al., 2004).

Population increases and urbanization in California have caused considerable impacts on natural ecosystems. The areas in which man-made structures border or coalesce with natural vegetation are called Wild-land Urban Interface (WUI). California has the largest number of houses and highest population of people living in WUI areas which are at high-risk for wildfire (Li et al., 2022). Wildfires have become more severe in the state, with the area of land burned increasing each year (OEHHA, 2018). In 2020, over 10 million acres of land were burned by wildfires, and almost 40% of these acres were in California (CRS, 2021). Deadly wildfire events intensified by climate change release stored carbon in soils and emit carbon dioxide into the atmosphere— 111.7 metric tons of CO_2 from California wildfires in 2020 (Huntsinger and Barry, 2021). Mediterranean vegetation and alpine forest ecosystems, both characteristic of California woodlands, are also particularly vulnerable to wildfire (Fischlin et al., 2007).

Forests and climate change are inextricably tied, with forested land mitigating climate change effects by serving as a carbon sink, but also having the potential to contribute to climate change when forests are burned or destroyed, which releases CO2 emissions. Between the years 1990 to 2020, there has been a global net loss of 178 million hectares of forest, primarily due to agricultural land conversion (FAO, 2020). The need for resources and land for food cultivation must be balanced with conservation goals, which requires more efficient and suitable food systems. Forests provide habitats to most of earth's terrestrial species and preserve genetic diversity (Van Bodegom et al., 2009). A direct benefit of forest conservation is habitat preservation for endangered species (Anderson et al., 2017). To meet the UN's Sustainable Development goals for biodiversity, large scale reforestation efforts will be needed (FAO, 2020).

Habitat loss and fragmentation is also a threat to grassland ecosystems, which support native biodiversity, ranching activities, and recreation (Root et al., 2015). The invasion of non-native grass species from the Mediterranean region have altered the soil carbon balance; in comparison to invasive grasses, native grassland perennials have been shown to increase carbon storage and root biomass, while decreasing soil evaporation to create a cooler and moister microbiome (Root et al., 2015). To ensure that plant and animal species dependent on grassland ecosystems have enough interconnected swaths of habitat to maintain their populations in the face of fire and range shifts, large enough areas of viable grassland habitats need to be established and protected (Klausmeyer et al., 2011). For instance, Gea-Izquierdo et al. (2007) found that in Californian grasslands where non-native species dominate over native grasses, 'islands' of low soil fertility (high C:N ratios) provide a refuge for the native species. This research shows that landscape scale planning of conserved areas can help diverse native grassland species remain in the future.

Mediterranean ecosystems characteristic of California are strongly influenced by a changing climate. As these climate sensitive systems oscillate more widely outside of historic ranges of variability, resource managers are confronted with many unknowns and increasingly complicated risk-benefit analyses. In the design and maintenance of any ecological management plan, an understanding of many mechanisms and existing infrastructure is integral to a holistic approach. For instance, impervious surfaces in urban areas can cause flooding and consequent soil erosion in adjacent natural areas. Thus, the unique conditions of different land types and their interconnected dynamics play a role in soil health and management. Land use change has historically played a huge role in terrestrial carbon losses (40% over the last two centuries) by altering natural carbon fluxes between the soil and the atmosphere (House et al., 2005). Land use is both a driver of and resultant impact from climate change. Understanding how and why land cover changes as well as future threats and opportunities for improved land management is key to addressing the complex interactions between drivers and impacts on ecosystem health.

2.2.3 Impending agricultural challenges

In California, the minimum rate of temperature increase is higher than average global increases, resulting in more frequent and severe droughts and heat waves (Pathak et al., 2018). Agriculture is impacted directly and indirectly by changing climate conditions including temperature shifts, precipitation and snowpack, and extreme weather events.

Understanding the climate sensitivity of crops to future climate conditions is difficult because weather and climate effects conditions must be uncoupled from other yield effecting factors including fertilization, pesticides, and soil health. Modeling the effects of specialty crop production is further complicated by the vast diversity of plants and their physiologies, various cultivation practices, and specific geographic considerations (Auffhammer, 2014; Kerr, 2018; OEHHA, 2014). For major commodities including soy, wheat, cotton, and corn, evidence points to extremely detrimental effects of ‘extreme heat days’ (30ºC) on yields (Auffhammer, 2014). Specialty crops contribute to most of the agricultural value of California, and therefore are of particular importance. The Central Valley is at notable risk due to decline in winter chill hours, which can lead to yield losses in specialty fruit and nut trees if they do not meet their vernalization requirement (OEHA, 2018). Warmer winter temperatures can also effect overwintering in insects and cause them to appear earlier in the season (Shazad et al., 2021).

Perennial crops are slower to adapt to environmental changes and thus more vulnerable than annual crops to substantial alteration (Lobell et al., 2006). For California’s top 14 specialty crops, Kerr et al. (2018) found the highest absolute impacts of temperature increases in the San Joaquin Valley and Central Coast (Fig. 2.3), which contain the top three ranked specialty crop producing counties. The absolute impact metric created by authors considered each county’s overall sensitivity, temperature exposure, and total specialty crop acreage for each crop considered.

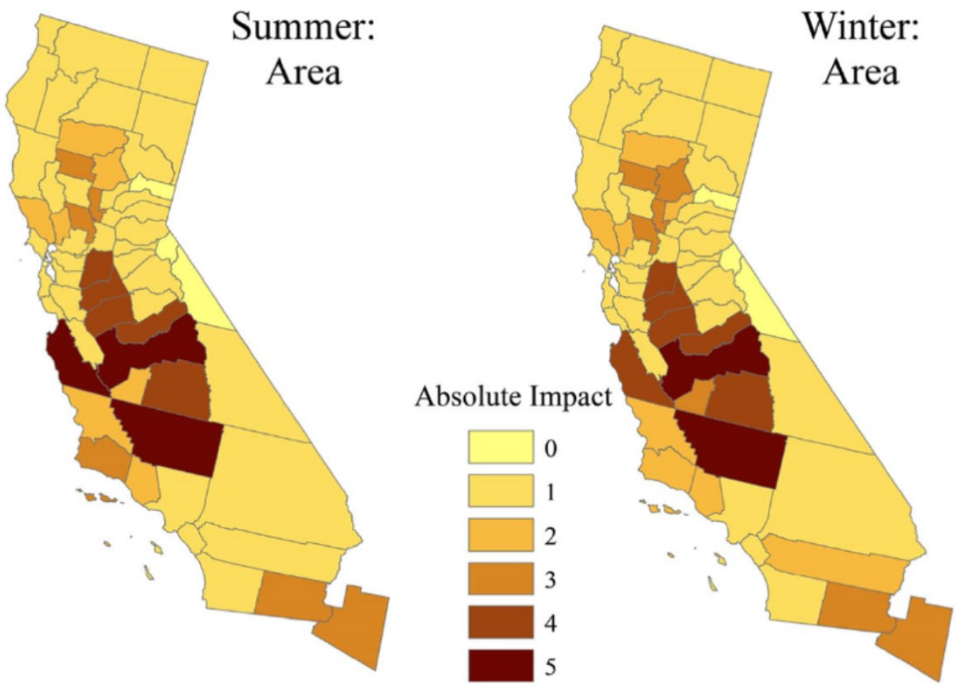

Figure 2.3: Absolute impact data for specialty crops shown on a county basis.1 indicates low sensitivity, 5 indicates high sensitivity, and 0 indicates no specialty crop production (Excerpted from Kerr et al., 2018).

Precipitation by snow and rainfall are California's primary source of water. Snowpack on the Sierras regulates California's water supply and storage; contributing to about 80% of the average annual precipitation (Pathak et al., 2018). However up to 80% of these reserves will be diminished by 2100 (Zilberman and Kaplan, 2014), and heat spells pose the risk of instigating earlier and quicker snowmelt. If soils are high in moisture from snowmelt, spring flooding may not infiltrate into the soil, and would be inaccessible for use throughout the summer (Shahzad, 2021). While the total amount of annual statewide precipitation exhibits no discernable trends, annual variability (dry and wet precipitation extremes) has increased since the 1980s (OEHHA, 2018). For the Southwestern United States, increasing temperatures are expected to coincide with an increase in the frequency and intensity of droughts (IPCC, 2021). General circulation models, a type of climate change simulation, have projected that Southern California and the state as a whole to be 15 to 35% drier by 2100 (Lobell et al., 2006). Drought events are exacerbated by increased atmospheric evaporative demand, which are expected to decrease soil moisture in the Southwestern United States (IPCC, 2021).

Despite the stresses California is expected to endure in the coming decades, according to an economic optimization modeling study by California's Fourth Climate Change Assessment, adaptive decision making in conjunction with technological advances can preserve the economic viability of agriculture in spite of the effects of climate change (Medellín-Azuara, et al., 2018; Bedsworth et al., 2018). Achieving the goal of getting more from less, requires a focus on high-yielding technologies and global technological improvements (Tilman et al., 2011). The successful utilization of technology for use in California agriculture offers the prospect of simultaneously optimizing cultivation outputs and sustainably managing resources.

2.2.4 Adaptive technology

To meet rising demands for food supply with increasing pressure on agricultural systems, effective adaptation to improve productivity will be necessary. Offsetting potential losses in airable farmland, crop yields, and water supply, may require sustainable intensification (SI) practices. The main tenant of SI is producing more output from the same or less land area and lessening negative ecological externalities (Deryng, 2020). According to the Food and Agriculture Organization of the United Nations (2009), it is anticipated that 90% of the growth in crop production globally will come from greater yields and increased cropping intensity.

Beyond California, smallholder farms in tropic and subtropic areas are at a disproportionate risk of being affected by food insecurity, increasing the risk of undernourished and hungry people in these areas (Thornton and Herrero, 2014). In the past, small adaptations by farmers such as shifting planting dates or using different cultivars has allowed them to keep up with gradual environmental changes. However, more significant climate shifts in future will require more drastic adaption in place of incremental changes (Panda, 2018).

Climate smart agriculture (CSA) is a concept based on productivity, adaptation, and mitigation strategies. Climate-smart technologies such as conservation agriculture, precision agriculture, agroforestry, integrated nutrient management, and soil and water conservation attempt to adapt to a changing climate and reduce emissions, while increasing crop productivity (Deryng, 2020).

Technological advances and new inventions for agricultural use have emerged for site-specific management which uses quantitative data to inform best practices. The adoption of these technologies on a farm level has led to automation on multiple scales of farm management, increased resource use efficiency, and reduced labor (Khan et al., 2018). Precision agriculture is a system of gathering, processing, and analyzing spatial, temporal, or individual data to inform managements decisions and improve resource use efficiency, environmental sustainability, and operation profitability (The International Society of Precision Agriculture, 2019).

The adoption of sensor-based technologies for farm management has led to automation on multiple scales, which can in turn reduce operational and labor costs (Khan et al., 2018). Environmental sensors operate via sensor nodes which interact directly with the environment to collect, store, and communicate data to a central database. In-situ sensors and wireless sensor networks (WSN) are increasingly being employed to collect continuous environmental information and assist farmers in activities including irrigation and nutrient management. Advances in wireless communication networks with large deployable ranges and the development of small, low-cost multifunctional sensors offer incredible opportunities to monitor physical, chemical, and microbiological properties across time and space (Chai et al., 2020).

For instance, soil moisture sensing via long-term sensor networks collects continuous in-situ data about soil moisture and temperature which can be used to inform efficient irrigation practices that strike the intermediary balance between crop water stress and excess water application. Figueroa and Pope (2017) used soil moisture probes that collected continuous moisture data (every 15 minutes) at five different depths in fields of avocados, kiwis, and nectarines. Time series analysis for the data involved detection of outliers and recognition of consumption patterns to identify the Root System Water Consumption (RSWC) pattern for each crop, which could be used to recommend an efficient irrigation schedule. Disease monitoring can also be performed by spectral sensors which capture changes in the physical appearance of plants and reveal the spatial distribution of infection to aid in timely and targeted pesticide applications. For example, Castalidi et al. (2017) deployed UAV multi-spectral imaging for weed identification in a cornfield, which led to a decrease in the amount of herbicide applied without an effect on yield.

Barriers to entry for these technologies are apparent, however, with the major drawbacks to some popular spatial and temporal sensors involving ease of implementation, cost, and accessibility. In-situ sensor nodes that operate around the clock can pick up on noise or; malfunction, resulting in errors in the data. Therefore, it is necessary to sift through the data to correct outliers. To overcome these challenges, various clustering techniques including partitioning, hierarchal, density-based, grid-based, and model-based algorithms can be used to analyze high-dimensional time-series data (Singh et al., 2015). Additionally, various desktop software programs which transform sensor data into usable agronomic information can be used as decision support tools. Regardless of how well different sensors might work for precision agriculture, it can be difficult implementing these technologies on the ground. A lack of access to financial capital on a farm scale can severely limit the opportunity to make investments in these types of technology. Another major challenge facilitating direct lines of communication with primary users and the company's technical assistance. However, communication protocols have been shown to increase battery life of sensors (Srbinovska et al., 2015). Aerial imaging or in-situ sensors alone only offer part of the story. An integration of imaging technology and continuous soil data would offer the clearest picture of field conditions and advise a holistic management strategy which bridges the spatial-temporal gap.

For those areas where point measurements are utilized to inform management practices, soil testing is the standard. The Natural Resources Conservation Service has recommended at least 1 composite sample per 20 acres every 3-5 years should be submitted for routine soil testing (NRCS, 2009). Characterizing these samples for properties of importance through a reputable lab can incur significant costs from shipping and lab fees. To save on these expenses, fewer composite samples or less frequent soil testing may occur, but at the cost of less detailed characterization. Losing the spatial variation within a field also compromises the ability to make targeted adjustments to that area. As a consequence, 'one-size-fits-all' management approaches may be taken, which can neglect the nutrient deficits in some areas or overfertilize in others, where excess nutrients no longer contribute to yield gains, and can even cause toxicity.

Inexpensive technology with low barriers of entry for implementation and rapid functionality are a desirable option for farmers interested in precision agriculture practices that can save money by reducing inputs, increasing yields, and improving farm efficiency. Since soils react slowly to change, assessing soil quality over time can be a challenge (Bünemann et al., 2018). Scaling up the spatial and temporal density of soil sampling to achieve a detailed resolution of soil properties can help detect these trends.

The major challenges facing California center around land cover changes where natural ecosystems are converted to urban or agricultural areas. These alterations affect carbon cycling, specifically by increasing greenhouse gas emissions and inhibiting the ability of soils to store carbon (Grimmond, 2007). Overcoming these challenges will require local approaches which leverage modern technology to inform decision making. Preserving, restoring, and monitoring soil health is an obligatory requisite for sustainable intensification of agricultural operations as well as the continued stability of grassland and forest ecosystems. Thus, accurate and timely data about soils, achievable through continued technological advancements, underpins sustainability centered goals.

2.3 Portable XRF for environmental applications

2.3.1 Monitoring for heavy metals

Initial studies that investigated pXRF analyzers for their use in soil focused on heavy metal contamination in urban or industrial soils (Argyraki et al., 1997; Clark et al., 1999; Carr et al., 2008; Chou et al., 2010; Radu et al., 2013). pXRF analysis has since evolved to be an accepted field tool for environmental screening (Ravansari et al., 2020). While investigating heavily polluted soils at the site of a historic silver mine in Ireland, Radu and Diamond (2009) found pXRF analysis gave excellent correlation with laboratory digests for heavy metal concentrations (R^2: 0.99 for Pb, 0.99 for As, 0.96 for Cu, and 0.84 for Zn), and recognized pXRF as a rapid and reliable analytical method for assessing soil pollution. A study measuring peatland lead contamination in the UK used in-situ pXRF measurements to map the spatial distribution of Pb in a 15-hectare peatland (Shuttleworth et al., 2014). Using dried, ground, and homogenized samples, an excellent relationship ($R^2 = 0.99$/RSD = 1.75%) was found between lead levels determined by ex-situ pXRF measurements and ICP-OES data.

Monitoring lead in the urban soil environment has been studied extensively. Lead is a persistent and toxic soil contaminant, which can have devastating effects on human health (WHO, 2021). Through government regulations and widespread education efforts, a continuous and substantial decline in lead exposure to the population has been achieved (Dignam et al., 2019). However, its long-term use and persistence in the environment

Continues to cause public health concerns today, with 500,000 children between 1-5 years old with blood lead levels (BLL) at or above the CDC reference value of 5 µg/dL (Dignam et al., 2019). Because lead is highly insoluble and persists in the soil for centuries, it is important to determine the spatial distribution of lead in urban environments to protect residents from exposure and inform safe land planning efforts. pXRF has shown value in identifying, quantifying, and mapping the presence of lead across the US. At peri-urban agricultural sites in Louisiana, Weindorf et al. (2012) performed on-site interpolation of heavy metal levels and created interpolation maps of enrichment factors from geo-referenced pXRF measurements, revealing the spatial distribution of contamination in the study area. The effect of urban-soil pedogenesis on the legacies of lead from paint and gasoline in Durham, North Carolina was also investigated by Wade et al. (2021). This study used gridded sampling and geospatial analyses in ArcGIS to map the distribution of lead throughout the city, visualize the movement of contaminated soil in the environment, and identify 'hotspots.'

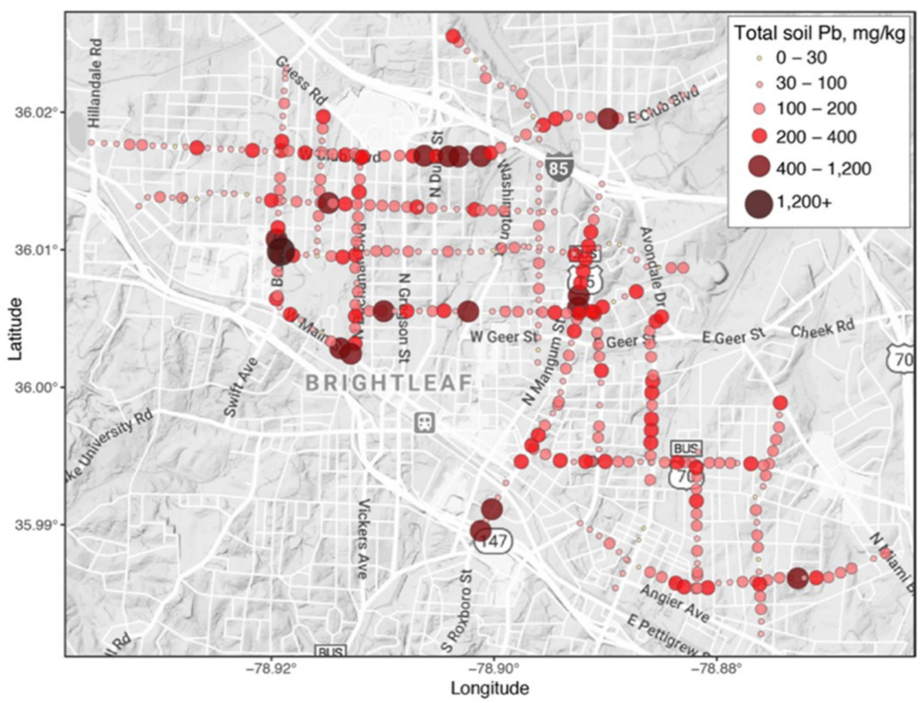

Figure 2.4: Total Pb as determined via pXRF analysis mapped along side streets in Durham, NC (Excerpted from Wade et al., 2021).

For environmental investigations, a single trip to the field to collect all samples can contribute to a cost-efficient sampling plan. The prospect of high sampling density and duplicated measurements is appealing for many uses in the environmental realm including delineating contaminated sites, pedologic descriptions, and civil engineering applications. Integration of pXRF data with GPS data into a geographic information system allows for mapping areas of interest with ease. Official analytical techniques methods compatible with pXRF measurements include NIOSH Method 7702 for airborne lead concentrations (NIOSH, 1998) and USEPA Method 6200 (USEPA, 2007), which outlines procedural considerations for field use of pXRF to determine elemental concentrations in soils and sediments. The establishment of official methods using pXRF is a good indicator of its reliability as a tool for total elemental analysis; but its utility as a predictive model has yet to be incorporated into any official methodologies. Thus, before pXRF calibrated models can be deployed in a meaningful widespread sense, a certain level of accuracy needs to be determined and communicated and certain best practices should be established.

2.4 Modeling soil properties with pXRF

2.4.1 Overview

The pXRF instrument provides multi-elemental data, which has been used successfully as a proxy for predicting other important physical and chemical soil properties and pedogenic processes. Some of these characteristics include pH (Sharma et al., 2014; Wan et al., 2019; Weindorf et al., 2019), cation exchange capacity (Sharma et al., 2015; Li et al., 2018; Wan et al., 2020), soil texture (Weindorf and Zhang, 2011; Zhang and Hartemink, 2020; Benedet et al., 2020; Silva et al., 2020), total carbon and/or nitrogen (Wang et al., 2015; Andrade et al., 2020), parent materials and pedogenesis (Stockmann et al., 2016; Silva et al., 2019; Gozukara et al., 2021), and horizonation (Weindorf et al., 2012; Weindorf et al., 2015). Models typically assess the best predictor elements for the property of interest and use statistical processes to find coefficient values. For some properties, certain elements tend to be key predictors due to the nature of the property in question. For instance, to estimate the quantity of gypsum ($CaSO_4 \cdot 2H_2O$) in soils, Weindorf et al. (2013) created simple and multiple linear regression models using Ca and S concentrations as determined by pXRF and achieved an $R^2 = 0.9127$. Additionally, because weathering indices depend on the fact that the concentration and movement of elements in a soil profile is determined by weathering and leaching processes, pXRF analysis can be used to identify elements in these indices (Stockmann et al., 2016; Zhang and Hartemink, 2019). For example, the Ruxton index (SiO_2/Al_2O_3) and Sesquioxide ratio (Si/Al + Fe) are weathering indices that calculate the ratio of mobile to immobile soil elements in a given horizon (Ruxton, 1968). Where certain oxides are relatively stable and immobile including TiO_2 and Al_2O_3, others including SiO_2, K_2O, and CaO are readily leached down the profile during the weathering process (Sauer et al., 2007; Gozukara et al., 2021). Thus, the index values determined by the concentration of these elements throughout the profile can signify the degree of chemical weathering that has occurred and indicate horizonation boundaries. For instance, to overcome incorrect subjective field classifications and support improved identification of soil parent materials, Gozukara et al. (2021) used pXRF analysis to identify geochemical properties of loess (A and Bt) and terra rosa horizons (2Bt) for soils in the Driftless Area of Wisconsin. Authors found less weathered loess horizons to have higher Si, Ti, and Zr concentrations in comparison to more weathered dolostone bedrock horizons which had higher concentrations of Al and Fe and lower Ruxton index/Sesquioxide ratio values.

The relative novelty of this research and necessity for large modeling datasets compels a wide array of soil collections to build the basis of the models. Published models are calibrated for the areas from which soils came, which may be a single or multiple regions. Research for pXRF modeling of soil properties has been performed in arid regions (Naimi et al., 2022; Towett et al., 2015), tropical areas (Silva et al., 2019; Silva et al., 2020; Silva et al., 2022; Benedet et al., 2022), various states in the United states including Wisconsin (Zhang and Harteman, 2020; Gozukara, et al., 2021), Louisiana (Zhu et al., 2011; Sharma et al., 2015) and Texas (Aldabaa et al., 2015), and throughout the globe (Stockman et al., 2016; Weindorf et al., 2015; Schneider et al., 2016; Wan et al., 2019; Mukhopadhyay et al., 2020; Weindorf et al., 2013). Modeling soil properties from pXRF analysis has been limited in California, despite the state's massive geographic extent within the US and importance for agriculture, forests, and rangelands. Notable studies using CA soils include Sharma et al. (2015) who used a sample set of 450 soils from California and Nebraska farmlands to assess CEC and Rawal et al. (2019) who used 300 samples from California and five other agricultural states to predict base saturation. However, these two studies only examined agricultural soils in California.

No research has been performed for calibrating state-wide models for properties of interest from California soils, taking account of the diverse California land types that exist outside of agricultural production. Existing models are typically calibrated based off of soils which come from a particular area or region, but there is a growing interest in models which facilitate reasonably accurate predictions on a larger geographic scale, such as state-wide scale. According to Weindorf and Chakraborty (2020), customized models calibrated with pXRF typically show considerable accuracy across a given region with relatively similar soil properties. Authors advise that significantly differing soils should have their own calibrated model. However, it is still important to assess the level of accuracy of widespread models constructed with strongly differing soils, because the level of accuracy given by these models may be adequate for some applications and would be more accessible than customized models.

2.4.2 Statistical modeling approaches

Several modeling strategies with varying levels of complexity have been investigated to uncover relationships between elemental concentrations and lab verified soils data. Data modeling and algorithmic modeling are two approaches used to associate predictor variables (x) with response variables (y) as to uncover how they are associated and/or make predictions about response variables from future input variables. Brieman (2001) explains the difference in how x and y are related between these approaches by imaging an intermediary 'box' between independent and dependent variables (Fig. 2.5).

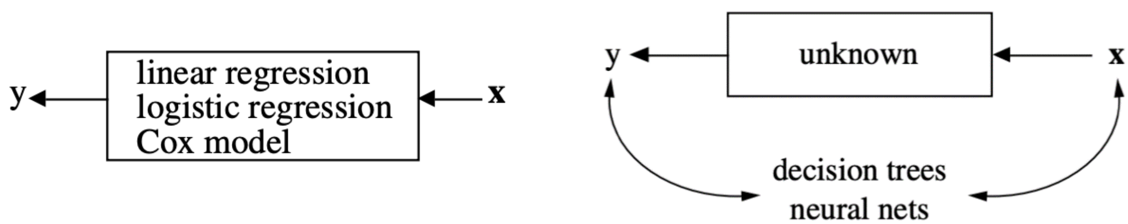

Figure 2.5: In the data modeling approach (left), a stochastic data model relates x and y using variables and coefficients while in the algorithmic modeling approach (right), the relationship between x and y is complex and unknown but can be related through algorithms (Excerpted from Brieman, 2001).

Data modeling is one such technique based on the idea that response variables are a function of the predictor variables, and some 'noise.' Simple linear regression (SLR) models assess the association between a single measured element and a soil property. These models have been unable to produce sufficiently robust predictions for soil pH

(Sharma et al., 2014), but have proven useful for relating Ca to gypsum quantity in soils (Weindorf et al., 2013; Acree et al., 2020)

Multiple linear regression models attempt to explain the response variable as a linear combination of multiple independent x-variables. MLR models can be simply constructed using linear modeling functions of statistical software, wherein the response variable is related to independent variables as shown in Eq. 2.1, with β_0 representing the intercept $\beta_1 \ldots \beta_i$ as the estimated regression coefficients, $x_1 \ldots x_i$ as the predictor variables, and ε as the random error variable, which is assumed to have a mean of 0.

Equation 2.1

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_i x_i + \varepsilon$$

The goal of MLR models in the context of relating spectral data to measured properties is to choose the collection of elements that best describe a certain soil characteristic. MLR models can be expressed as equations or in a table of coefficients. Use of MLR has been successful for predicting soil pH (Sharma et al., 2014), soil texture (Zhu et al., 2011), and CEC (Sharma et al., 2015), for the soils used in these studies.

When model components interact nonlinearly or X values exhibit multicollinearity, MLR models may not be possible, and alterative algorithmic models may be necessary. In addition, while regression models are useful for inference, machine learning is often a better approach for predictive modeling (Brieman, 2001). For instance, decision tree models which can integrate categorical data such as soil color as well quantitative data like total C and N were used to identify parent materials and terra rossa horizons in Wisconsin soils (Gozukara et al., 2021). Decision trees work by iteratively splitting the dataset on different attributes, depending on the reduction in MSE that is produced. During the training phase, decision trees use decision rules from the data attributes to learn relationships and map the data to its output. Overfitting and instability in predictions with small changes in the data are disadvantages of using decision trees. Cubist/M5 models also use a tree structure, wherein each path from the top to bottom of the tree is a rule that divides the dataset into smaller subsets. From these subsets, linear regression models are created, with the models produced at the terminal leaves having been 'smoothed' by the models formed at the above nodes. O'Rouke et al., (2016) used Cubist predictive models to predict agronomic properties from pXRF and Vis-NIR sensor data individually and by averaging the two models together.

Support vector machine (SVM) learning works by identifying a hyperplane that distinctly classifies the datapoints. The data exist as points in n-dimensional space, where n is the number of features. The support vectors which ultimately build the shape of the hyperplane separate different classes of data in a way so that there is the maximum margin between the classes. Unseen data can be classified by the SVM by plotting it against the established hyperplane and seeing where it falls with regards to the vectors. SVM learning can also be used as a regression method called support vector regression or support vector machine regression (SVR/SVMR) which is based upon the same principles as SVM: identify the plane which minimizes error and maximizes the margin between two or more groups. For SVR/SVMR, a continuous variable can be predicted by specifying the margin of error and allowing the algorithm to find the regression model that gives the best approximation. Both SVM and SVR are helpful for visualizing multi-dimensional non-linear patterns and classification of datapoints. SVR has been used to link proximal data to soil characteristics across catenas (Duda et al., 2017) and SVM learning has been used to predict soil texture from proximal data (Benedet et al., 2020).

Partial least squares regression (PLSR) is a linear modeling approach that can be applied for a large number of correlated predictor variables. By preserving those predictors which explain as much covariance between the observations and predictions as possible, the number of predictors is reduced, and a linear regression model is created. Random forest (RF) regression fits several decision trees to train predictions. A number of decision tree regressors are indicated to the algorithm and resultant model outputs from many different subsets of the data are averaged across trees to find the final output (Fig. 2.6). This approach is strong because the averaged predictions from an ensemble of trees reduces the error and variability compared to a single prediction. RF models can be by specifying certain hyperparameters such as the number of trees in the forest and the number of features considered at each node.

Several studies have compared multiple machine learning models for their suitability in predicting certain soil properties. For example, Rawal et al., (2019) applied generalized additive model (GAM), MLR, RF, and regression tree (RT) models for predicting soil base saturation percentage (BSP) and CEC. All four models produced fair residual predication deviations (RPD) with the RT model for BSP and GAM approach for CEC performing the best. Authors advised that GLMs to be preferred over RF models for simplicity and interpretation's sake. Aldabaa et al. (2015) used SVR and PLSR methods to predict soil salinity from pXRF, Vis-NIR, and remotely sensed data. SVMR and PLSR have also been utilized to predict environmental risk based on heavy metals and soil pH from Vis-NIR and pXRF data in the Yunnan Province, China (Wan et al., 2019). Authors found that SVMR from pXRF elemental data gave reasonable predictions of pH. Silva et al., (2020) compared GLM, SVM, and RF algorithms for predicting soil texture of

Brazilian soils from pXRF data and found that SVM provided the best estimates for clay and sand contents, while RF gave the best estimates for silt contents.

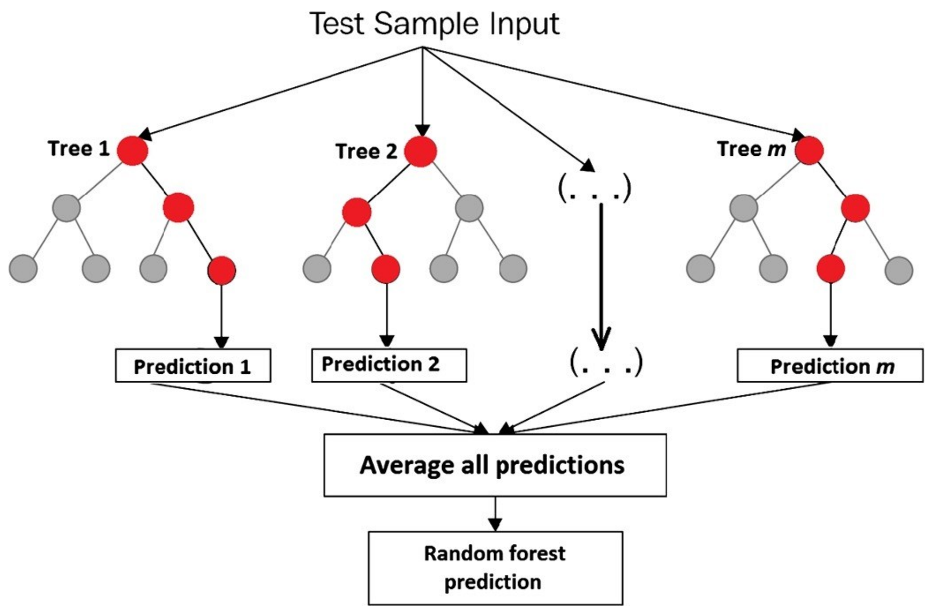

Figure 2.6: RF ensembles use the predictions of many decision trees to produce a final output prediction (Excerpted from Afzal et al., 2020).

2.4.3 Metrics for model performance evaluation

To assess how well a model can predict output values, it is standard to compare the emulated response values (y') to the actual response values (y). For regression models, this typically involves using a subset of the data to establish the model parameters and then assessing how well that model predicts your data using a holdout sub-set for testing. Performance metrics are important for interpreting how well the model fits unseen data and for comparison of various models.

The coefficient of determination (R^2) represents the squared correlation between the predicted values and actual values. It signifies the proportion of the variability of the response that is explained by the predictors through a linear relationship. Past studies have contended that models with an R^2 value of 0.6 - 0.8 provide medium level predictive power, whereas models with $R^2 > 0.8$ are acceptable or highly accurate for indirect prediction of soil properties (Malley et al., 2004; Nduwamungu et al., 2009). But it has also been suggested that the usefulness of the model should instead be based on the reduction of uncertainty it provides for a given purpose (Towett et al., 2015). As more predictor variables are added to a MLR model, R^2 will rise regardless of whether additional variables are related to the response or contribute significantly to the model. A large amount of predictor variables will therefore inflate R^2 and increase the chance of making a type I error (rejecting the null hypothesis when it is actually true). Adjusted R^2 is a metric which helps account for this phenomenon by applying a penalty for unimportant variables in the model. The use of adjusted R^2 can help to explain variance more economically and identify the most parsimonious model. The correlation coefficient (R) is also used to determine the strength and direction between two variables. To validate MLR models for pH and CEC (Sharma et al., 2014; 2015) performed correlation analysis on their validation sub-datasets between the predicted values from their model and the actual values.

Root mean square error (RMSE) is the square root of the average squared error. It is one of the most common ways to quantify how well a model predicts response values. RMSE is a helpful metric for gauging the average distance between the model's predicted values and the actual values from the dataset (the average error magnitude). In other words, how closely concentrated the data are around the line of best fit influences the magnitude of the RMSE. Poor models where predicted values are far from the model line of best fit, will have a relatively high RMSE, whereas a model that predicts y values accurately will have a relatively lower RMSE and be more tightly clustered around the line of best fit. Interpretation of RMSE values is straightforward because it is measured in the same units as the dependent variable.

Another prevalent model assessment metric is residual prediction deviation (RPD), which is the ratio of standard deviation of observed values and RMSE. RPD is a unitless statistic that allows for error to be easily compared (Malley et al., 2004). However, RPD as well as R^2 can be overly influenced if data shows a skewed distribution (Malley et al., 2004). In formulating models to predict various soil properties from near infrared reflectance spectroscopy, Chang et al., (2001) judged their model's performance from the RPD values of the validation set, where an RPD > 2 indicated a stable an accurate model, RPD between 1.4 and 2 were fair models with the potential to be improved with different calibration approaches, and RPD < 1.4 were considered poor models that could not predict the property of interest. RPD and the classification system for interpretation by Chang et al., (2001) have been used to evaluate models characterizing soil properties across catenas (Duda et al., 2017), modeling total carbon and nitrogen (Wang et al., 2015), predicting CEC from pXRF spectra (Sharma et al., 2015), using proximal and remote sensing methods to predict soil salinity (Alabaa et al., 2015), and estimating base saturation of agricultural soils with pXRF data (Rawal et al., 2019). However, the RPD categories created by Chang et al. (2001) are relatively arbitrary, and suitable models have been developed which give considerably lower RPD values (Bellon-Maurel et al., 2010).

The ratio of performance to interquartile distance (RPIQ) is another metric to assess prediction status calculated as the interquartile range divided by the RMSE of the prediction, where a higher RPIQ indicates better model performance. RPIQ is appropriate for non-normally distributed data and better accounts for the spread of the population than RPD (Bellon-Maurel et al., 2010). As an assessment of model performance, RPIQ has been used extensively, including for interpreting results of model averaging of pXRF and Vis-NIR spectra (O'Rouke et al., 2016), classifying soils in Romania with pXRF and Vis-NIR models (Acree et al., 2020), and predicting soil fertility attributes with pXRF and Vis-NIR data (Liu et al., 2021). O'Rouke et al., 2016 used interquartile ranges of RPIQ values from their validation set to categorize predictions as good (>1.03), reasonable (0.77 -1.03), and unreliable (<0.77). In a soil quality study using VNIR reflectance spectrometry, Veum et al. (2015), extended upon the prediction categorization by Chang et al., (2001) to include RPIQ ranges. Their classification defined RPIQ ≥ 3.0 as the most reliable 'Category A' models, RPIQ ≥ 1.9 as 'Category B' models with the potential for improvement, and RPIQ ≥ 1.5 as unsuitable 'Category C' models.

2.4.4 Sensor data fusion for modeling

Proximal sensing simply refers to the use of a sensor which collects signals via a detector when in close proximity to soil (*Soil Science Division Staff, 2017).* A major advantage of this technology is the ability to gather high density measurements which 'fill the gap' between high resolution point data validated in a laboratory and more coarse resolution remote sensing data. Information obtained from proximal sensors can be used to observe spatial variability of a soil property and refine soil survey data by indicating soil map unit boundaries. Some examples of proximal sensors include ground-penetrating radar, time domain reflectometry, electrical resistivity, visible–near–infrared (Vis–NIR) spectroscopy, and portable X-ray fluorescence. Using the spectra obtained by Vis-NIR and pXRF analysis in tandem has proven a popular method for building predictive models relating to soil properties. It should be noted that for both technologies, spectroscopy refers to the science of the interaction between radiated energy and matter, while spectrometry makes sense of the spectra produced by spectroscopy and translates it into quantifiable results. Vis-NIR measures the transmittance, reflectance, and absorbance of light at specific wavelengths across the visible range, allowing for quantitative analysis and material characterization. A sample is illuminated with light across the visible electromagnetic wavelength range (400-800 nm) and the absorbency at discrete wavelengths is graphed to produce a spectrum. This method exhibits similar advantages and limitations to pXRF spectrometry including the ability to collect high sensitivity measurements rapidly and nondestructively with a compact instrument but with the need to distinguish between sample peaks and background noise. Used together, pXRF and Vis-NIR have complementary capacities for assessing soil— with Vis-NIR able to quantify the organic components and minerology, while pXRF can accurately estimate inorganic elements (O'Rourke et al., 2016).

Research into which of these techniques is superior, or if they are best used in tandem for predicting soil properties has shown varied results. Wang et al. (2013) obtained robust models to predict sand and silt contents using the combination of pXRF data and Vis-NIR DRS spectra; however, the use of the combined data did not satisfactorily increase accuracy of clay content prediction in comparison to models trained with pXRF data only. Zhang and Hartemink (2019), also found that pXRF was better at predicting texture than Vis-NIR when used solo— but observed that sensor fusion further improved the prediction. On the other hand, Benedet et al. (2020) found that pXRF and Vis-NIR data produced accurate predictions of soil texture both individually and in tandem. The results of these models were heavily dependent on preprocessing, the sensor dataset, and the algorithms used, but authors found a pXRF dataset and RF algorithm providing the best results. Conversely, Naimi et al. (2022) found Vis-NIR-produced better estimates for soil texture when compared to pXRF. Swetha and Chakraborty (2021) found that a Nix color sensor data improved pXRF based predictions of clay content which were then used as a proxy to predict soil organic carbon content.

For predictions of SOC, Liu et al. (2021) found Vis-NIR alone was a good predictor ($R^2 = 0.77$) while pXRF used alone was an inadequate predictor ($R^2 < 0.32$). Naimi et al., 2022 also found that pXRF was unable to predict SOC in arid soils of the Afar region. However, Wang et al., (2015) found synthesized penalized spline regression (PSR) and RF models using both pXRF and Vis-NIR data were more effective than either proximal sensing technique on its own for predicting total carbon and nitrogen. For distinguishing between parent materials and identifying lithologic discontinuities, Gozukara et al. (2021) found pXRF spectra to have better prediction accuracy than Vis-NIR spectra when using decision trees. In assessing sensor data fusion for predicting soil pH, Wan et al. (2019) observed that pXRF elemental data alone could predict soil pH with reasonable accuracy but predictions were improved with fused pXRF and Vis-NIR data. For CEC predictions, fused Vis-NIR and pXRF data also provided the most accurate and comprehensive predictions when compared to those produced by either single sensor dataset— but pXRF elemental data contributed more to the PLSR fused sensor data (Wan et al., 2020). Model averaging procedures that combined model outcomes from pXRF and Vis-NIR were used by O'Rourke et al. (2017) to greatly improve predictive capacity for soil pH, SOC, total nitrogen, texture, and CEC.

Thus, it is evident that sensor data fusion has yielded mixed results even when tasked with predicting the same properties. When using two sensor datasets to build predictive models, the data collection, preprocessing, and analyzing is made more labor-intensive when compared to using only one of these datasets. In addition, the acquisition cost of these sensors is considerable. An ultimate goal of using proximal sensors to estimate properties of interest soil is to make soil characterization more straightforward. When it comes to using several datasets to build predictive models, it is important to be cautious of incorporating too many predictor variables to 'force' a correlation. Overfitting a model can also occur with enough free variables in overly complex models.

2.5 Existing models of interest

2.5.1 pH

Soil pH is a critical measure for soils and has implications for soil fertility related to buffering capacity and nutrient availability. The need for a standardized and efficient method of measuring pH is imperative for appropriate soil management as soil pH can serve as a measure for terrestrial biogeochemical processes (Kome et al., 2018). In building models to predict soil reaction (pH) from pXRF elemental data, Sharma et al., (2014) used two datasets with two modes of pXRF operation. For this study, soil pH was measured via saturated paste with deionized water. Datasets A and B were divided into 80% modeling and 20% validation subsets. Dataset A was comprised of 100 soil samples across the United States, with 50 coming from supposed alkaline and 50 coming from supposed acidic soils. Scanning via pXRF was conducted using a DP-6000 Delta Premium pXRF (Olympus, Waltham, MA, USA) operated in GeoChem mode for dataset

A samples. Authors used Pearson's correlation on log values of elemental concentrations to identify those elements with significant relationships to pH and then eliminated those samples which were missing any concentrations of the significant elements, leaving n = 57. The resultant model equation shown below (Eq. 2.2) achieved an $R^2 = 0.570$/RMSE = 0.822 on the modeling dataset. The model was validated by running correlation analysis on 15 randomly selected samples, to find an R = 0.433. The addition of clay, sand, and organic matter contents as predictors further improved the model to an $R^2 = 0.825$/RMSE = 0.541.

Equation 2.2

$$pH = 9.7164 – 5.9247 * \log(Al) + 1.8491 * \log(Si) - 2.0419 * \log(Mn) + 1.9212 * \log(Fe) + 2.3906 * \log(K) + 0.4396 * \log(Ca) + 0.6680 * \log(Zn)$$

Dataset B was comprised of 639 samples from across Louisiana scanned by the pXRF using the Soil Mode of operation. Pearson's correlation was used on 15 elements (K, Ca, Cu, Zn, Ti, Cr, Mn, Fe, Co, As, Rb, Sr, Zr, Ba, Pb) to select predictor variables. Authors achieved an $R^2 = 0.772$/RMSE = 0.685 from the modeling dataset using Eq. 2.3 shown below. To validate the MLR model, 20% of samples were randomly selected and used for correlation analyses, returning R = 0.573.

Equation 2.3.

$$pH = 1.4246 – 0.5989 * \log(K) + 1.3739 * \log(Ca) – 0.4426 * \log(Cu) – 0.4296 * \log(Zn) – 0.4220 * \log(Ti) – 1.3528 * \log(Cr) – 6.8667E–02 * \log(Mn) – 0.6366 * \log(Fe) + 0.9780 * \log(Co) + 9.7264E–02 * \log(As) + 1.1561 * \log(Rb) – 5.2320E–02 * \log(Sr) + 1.1699 * \log(Zr) + 1.3802 * \log(Ba) – 0.4718 * \log(Pb)$$

2.5.2 Texture

Soil texture, defined as the relative proportions of sand, silt, and clay, is likely the most important physical characteristic of soils crucial for understanding their behavior, suitability for various applications, and the effect of management practices. Some soil characteristics directly influenced by soil texture include water holding capacity, nutrient retention capacity, rate of chemical weathering and microbial reactions (Weil and Brady, 2017). To assess the viability of using pXRF data to predict soil texture by estimating clay and sand contents, Zhu et al., (2011) analyzed 584 samples from Louisiana and Capulin, New Mexico. Authors scanned the samples using Soil Mode, operating with a sequential 3-beam scan for a total scan time of 90 seconds per sample. Soil texture was also determined via traditional laboratory analysis using the pipette method (Soil Survey Staff, 2004). Authors used a 2/3 modeling and 1/3 validation split of their dataset. Backward stepwise multiple regression analysis with entry significance of 0.5, removal significance of 0.1 and 15 maximum steps was conducted on the modeling sub dataset between lab values of clay and sand precents and 15 predictor elements (K, Ca, Ti, Cr, Mn, Fe, Co, Cu, Zn, As, Rb, Sr, Zr, Ba, and Pb). Predicted clay and sand percentages were subtracted from 100% to find silt contents. The results of the backward stepwise MLR models can be found in Fig. 2.7. Higher Rb concentrations were found to correlate with higher clay and lower sand percentages, while higher Fe concentrations were found to correlate with higher clay and sand percentages. Applying the model to the validation sub-dataset revealed better performance in predicting clay contents (R^2 = 0.975/RMSE = 2.68% for Louisiana soils and R^2 = 0.876/RMSE = 2.66% for Capulin soils) than in predicting sand contents (R^2 = 0.854/RMSE = 5.53% for Louisiana soils and R^2 = 0.891/RMSE = 6.62% for Capulin soils) (Fig. 2.8).

Variable	Louisiana sand	Louisiana clay	Capulin sand	Capulin clay
	Coefficient	Coefficient	Coefficient	Coefficient
Constant	98.5	−0.1	45.7	9.1
K	0.000557	−0.000869		
Ca	−0.000853		0.000326	−0.00031
Ti	−0.00802			−0.00174
Cr				
Mn	−0.00441	−0.00357		
Fe	0.00072	0.00118	0.000293	0.000282
Co	−0.0224			
Cu				
Zn	0.193			
As	−0.372	−0.329		
Rb	−0.412	0.319	−0.424	0.231
Sr	−0.135	−0.0603	0.0223	
Zr			0.0313	−0.0196
Ba	−0.0456	0.0127	−0.0308	0.0112
Pb				
Sample number	284	284	105	105
Outliers	2	1	0	1
R^2	0.86	0.96	0.89	0.78
SE of Estimate	6.05	3.56	6.31	3.33

Figure 2.7: Backward stepwise MLR models produced from the modeling sub-datasets for sand and clay contents of Louisiana and Capulin soils (Excerpted and adapted from Zhu et al., 2011).

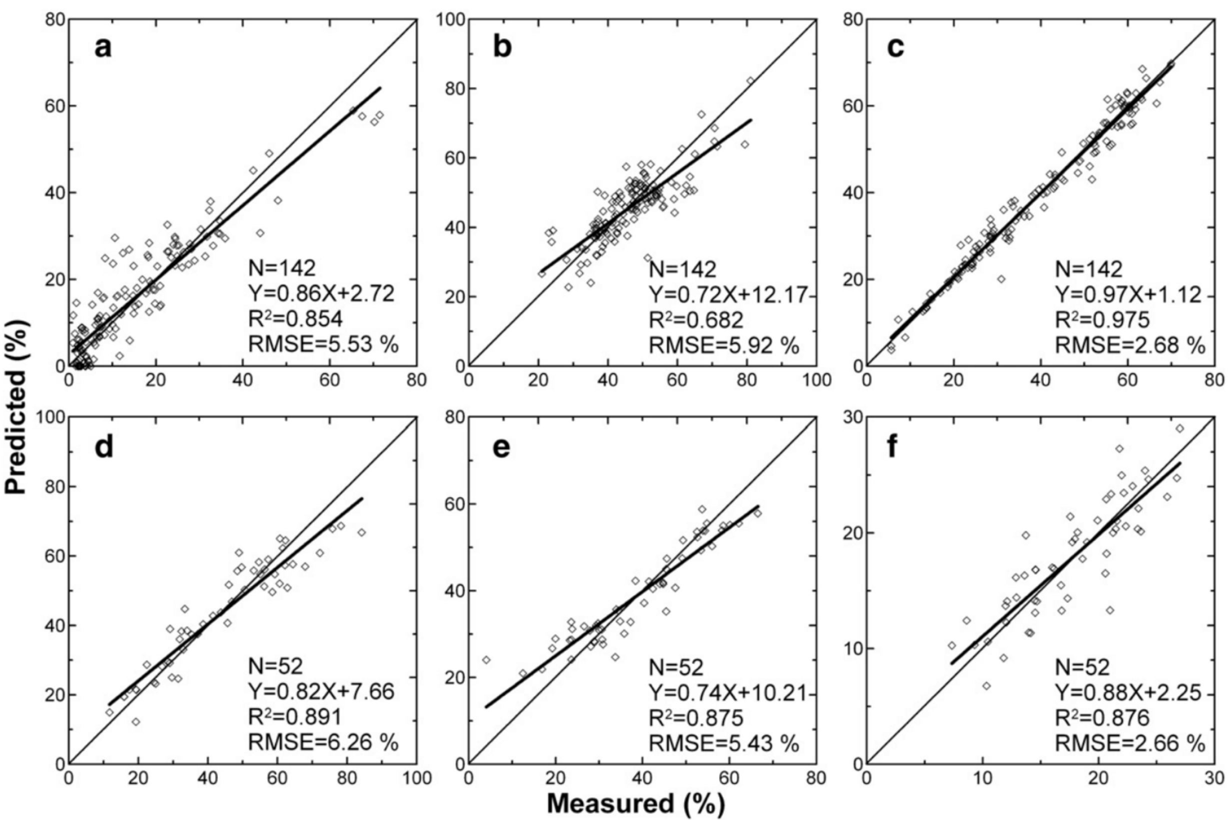

Figure 2.8: Sand, silt, and clay predictions (left to right) for Louisiana (top row) and New Mexico samples (bottom row) (Excerpted from Zhu et al., 2011).

2.5.3 CEC

Cation exchange capacity (CEC) is a useful measure for soil fertility, indicating the portion of exchangeable cations that are electrostatically bound to negatively charged soil surfaces. The nutrient retention of a soil is indicated by CEC because those cations in the exchangeable pool are readily taken up as plant nutrients. A regression model was built by Sharma et al. (2015) to predict CEC based off of 450 agricultural soils from Nebraska and California. For sample collection, three sampling depths were collected from 75 sampling pits in both states. Scanning via pXRF was conducted in Soil Mode, and models were constructed from 360 samples and 15 elements. An 80/20 model training and validation split were conducted on the full dataset. Authors performed stepwise regression analysis which selected eight elements to be included in the model equation (Eq. 2.4) which produced an $R^2 = 0.908$ and RMSE = 2.498 for the modeling dataset (Fig. 2.9). The addition of clay content and SOM as auxiliary predictors improved modeling dataset predictions to $R^2 = 0.926$/ RMSE=2.236. The developed models were validated via correlation analysis, with Eq. 2.4 producing a significant correlation (R = 0.904).

Equation 2.4

$$CEC = 17.2507 - 3.6514E{-}04 * Ca - 3.4957E{-}03 * Ti + 7.0977E{-}02 * V + 7.0991E{-}02 * Cr + 5.9759E{-}04 * Fe + 0.1479 * Cu - 6.2096E{-}02 * Sr + 5.6551E{-}03 * Zr$$

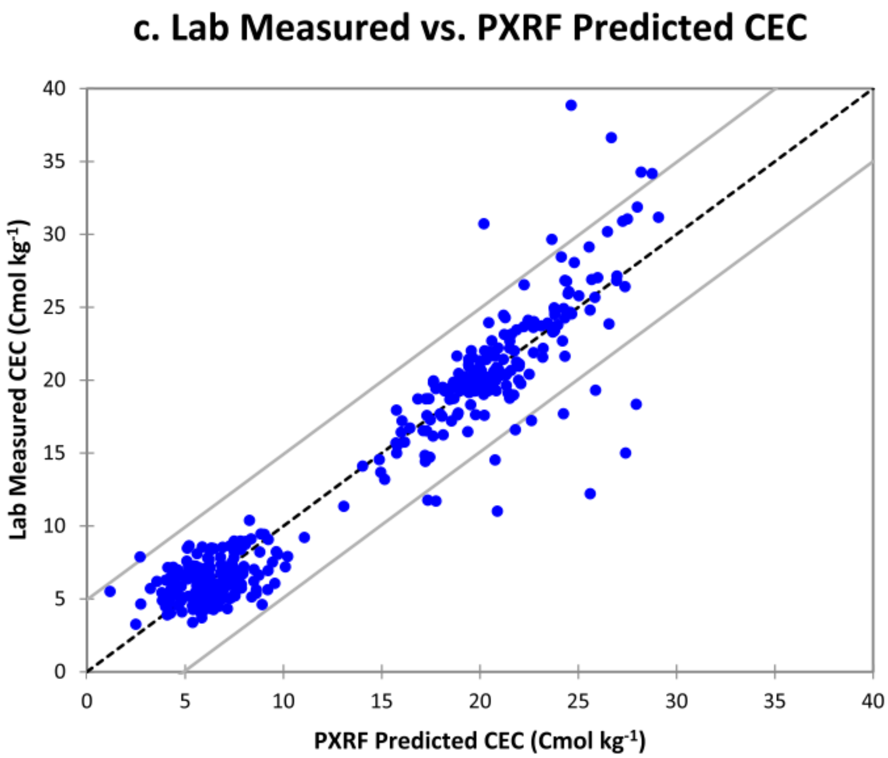

Figure 2.9: The lab measured CEC plotted against pXRF predicted CEC using Eq. 2.4. The dashed line is a 1:1 line and gray lines represent the 95% confidence interval (Excerpted from Sharma et al., 2015).

2.5.4 Soil organic carbon, total nitrogen, and C:N ratio

A limitation of pXRF analysis is the inability to detect light elements (Z scores ≤ 11) and therefore only present an abridged geochemical profile of samples (Duda et al. 2017). Carbon and nitrogen contents are light elements that are important aspects of soil characterization. Nitrogen is a plant essential nutrient imperative in the right quantities for plant health and vigor, while excesses of N can lead to nitrate pollution in waterways. Soil organic matter is comprised of about half organic carbon (Weil and Brady, 2017) and therefore plays an important role in soil quality and the world's carbon balance. Despite not being capable of detecting C or N contents directly, a relatively high portion of the variance in these concentrations explained by XRF may be indicative of elemental signatures which correlate with soil organic matter fractions (Towett et al., 2014).

pXRF has been used with limited success on its own for predicating TC and TN when compared to fused sensor datasets (Wang et al., 2015; Duda et al., 2017). However, Towett et al., (2015) achieved moderate predictive accuracy ($R^2 > 0.60$) when estimating OC and TN from total XRF (TXRF) spectra alone. To develop predicative models for use in Sub-Saharan Africa, Towett et al., (2015) used mid-infrared (MIR) and TXRF spectroscopy individually and in tandem to predict various properties for 700 soil samples. Organic C and total N were determined using flash dynamic combustion with a Flash EA 1112 Elemental Analyzer and TXRF methodology was used to analyze total elemental concentrations in each soil sample using a S2 PICOFOX TXRF spectrometer. Random forest models were built to predict the properties of interest from the raw TXRF elemental concentrations using the ‘randomForest’ library in R (Liaw and Wiener, 2022). To optimize prediction accuracy, RF models grew a prespecified number of classification and regression trees (CART) (ntree= 200) (via bootstrap sampling) with randomly selected variables from the calibration dataset. At each node in the tree a CART algorithm tested the performance of randomly selected variables to determine how the node was to be split. An internal cross-validation was performed by splitting the calibration set into 2/3 in-bag and 1/3 out-of-bag (OOB) subsets. The OOB error predictions provided were justified by comparing these errors to the errors from a 50% holdout set. The RF OOB validation for the TXRF based dataset produced an $R^2 = 0.68$/RMSE= 0.7 for organic carbon and $R^2 = 0.63$/RMSE=0.003 for total nitrogen. The TXRF data used alone performed more poorly than both the MIR and combined MIR + TXRF datasets.

2.6 pXRF instrumentation technology and theory

2.6.1 Excitation sources

XRF works by bombarding a sample with high energy X-ray beams to irradiate a sample via internal sealed radioisotope sources or X-ray tube. Earlier generations of pXRF instruments used a sealed radioisotope source to meet the requirements of minimal mass and no power consumption, but X-ray tubes are the prominent sources used in pXRF analysis today.

The commonly used radioisotope excitation sources include ^{55}Fe, ^{57}Co, ^{109}Cd, and ^{241}Am, which each give off radiation at particular energy levels (Kalnicky and Singhvi, 2001). As a result, each of these sources causes different elements to fluoresce based on their atomic number, making multi-elemental analysis possible only with a combination of isotopes. Use of pXRF with a ^{57}Co radioisotope has been used for decades to detect lead-based paint for public health applications (Guimarães et al., 2015). However, relatively short half-lives for some sources (~272 days for ^{57}Co) means that detection sensitivity degrades over time and isotope replacement is necessary every few years.

An X-ray tube consists of a cathode, anode, and tube envelope, tube housing, and a window. These components are housed within a vacuum sealed envelope necessary to dissipate the heat energy from the X-ray generation and contain radiation. Heating a wire filament made of tungsten causes a beam of electrons to be expelled from the cathode component and accelerated towards and absorbed by the anode component. This collision results in X-rays known as Bremsstrahlung (also called white radiation/breaking radiation) which produce continuous emissions characteristic of the anode material (Kramar, 2017). When the pXRF is aimed at a target, these emitted X-rays interact with the atoms in the substance. If the energy from the emitted X-rays exceeds the shell binding energies of electrons in the K or L orbitals of an atom, an inner shell electron is dislodged. In turn, a characteristic reflorescence indicative of the element is emitted and measured as electric signals by the XRF (Fig. 2.10).

Unlike active sources, which are always ‘on’ and emitting some levels of radiation which can be potentially hazardous, the X-ray tube mechanism only emits X-rays when energized. X-ray tube mechanisms can be modified for specific applications and have a less demanding licensing process when compared to radioisotope sources which present decay characteristics (Nummi, 2015). Other drawbacks to active sources are apparent, including the increased stringency for their use and handling. Additionally, the need for radiation shielding limits the number of sources that can be used in tandem within handheld devices, which makes these sources less bright those using X-ray tubes (Potts and West, 2008). Where radioisotope sources gradually and predictably lose efficacy, X-ray tube mechanisms burn out abruptly, and require replacement by the manufacturing company (Glanzman and Closs, 2007). Radioisotope sources may be a preferred source over X-ray tubes when simplicity, compact construction, low power requirement, and high energy X-rays are needed are needed for the application.

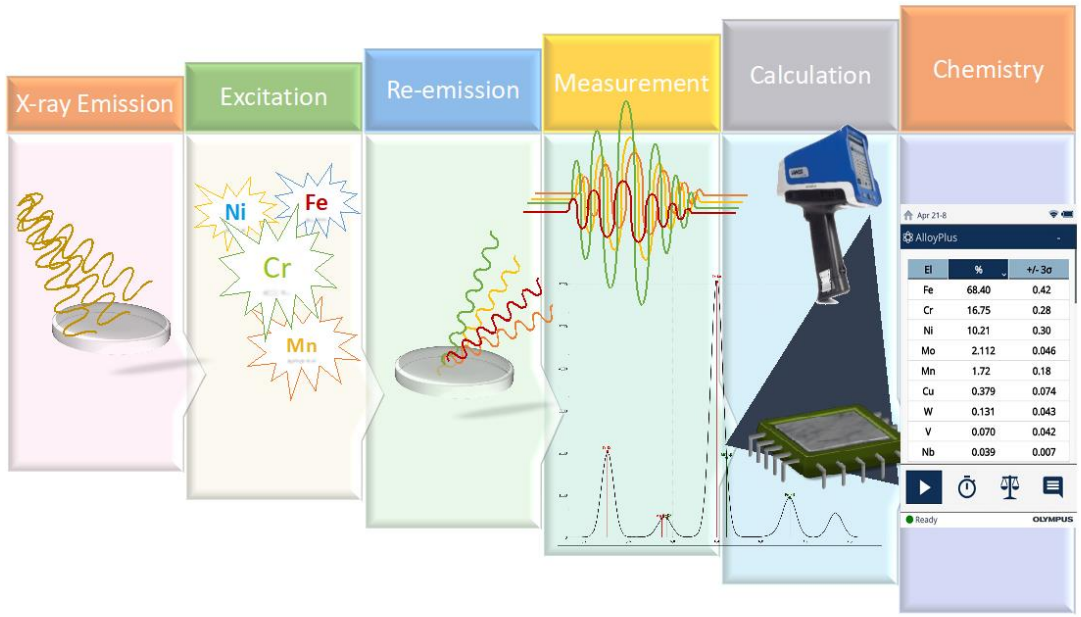

Figure 2.10: pXRF testing process (Image from Olympus Scientific Solutions).

2.6.2 Wavelength vs energy dispersion

There are two main XRF methods used for characterizing elemental composition: energy dispersive XRF (EDXRF) and wavelength dispersive XRF (WDXRF). Prior to the 1970s, wavelength dispersive technology underpinned most X-ray spectrometers. The development of energy dispersive spectrometers in the late 1960s made microanalysis in portable XRF devices possible (Weindorf et al., 2014a). These techniques differ primarily by the way characteristic X-rays are detected and analyzed, with each offering some advantages. Where wavelength dispersion separates X-ray lines based on their wavelengths, energy dispersion separates X-rays based on photon energies.

WDXRF technology is based on Bragg’s law, which states that X-rays of specific wavelengths and diffraction angles will be reflected by crystals when the wavelengths of the scattered x-rays experience constructive interference (Keng, 2015b). Crystals are used in WDXRF to separate and distinguish the wavelengths of each element in the fluorescence spectrum. The crystals physically separate X-rays and diffract them in different directions based on their wavelengths. By fixing the crystal and detector positions, the characteristic wavelengths produced by each element can be quantified (Henry et al., 2016).

In an energy dispersive detection method, the energies of the fluoresced X-rays are directly measured by an internal detector made of a semiconductor material (typically silicon) and transformed into an electric signal. These signals are then processed with a pulse height analyzer (Kalnicky and Singhvi, 2001). The height of the peaks represents the number of return X-rays registered by the instrument and corresponds to the concentration of a particular element (Crumbling et al., 2008).

Since EDXRF provides lower spectral resolution (150-300eV) when compared to WDXRF (5-20eV) the peaks of different elements may overlap, making it difficult to distinguish which elements are present (Wolfgong, 2016). For instance, in Fig. 2.11, at 6.0 keV, the Mn and Cr peak experience spectral interference/peak overlap, which can distort results for these elements. Elements with longer wavelengths are difficult for EDXRF to detect, so the technology is generally only practical for detecting 'heavier' elements (atomic numbers > 11 ((Na)). However calibrations for WDXF are more involved and a high power unit for the X-ray source necessitates a larger and typically more expensive instrument compared to EDXRF instruments (Wolfgong, 2016; Kawahara and Shoji, 2007). pXRF relies on EDXRF, because its components can be miniaturized into a compact device, and the detector can be close to the sample, which allows for highly sensitive measurements from a small amount of sample (Kawahara and Shoji, 2007).

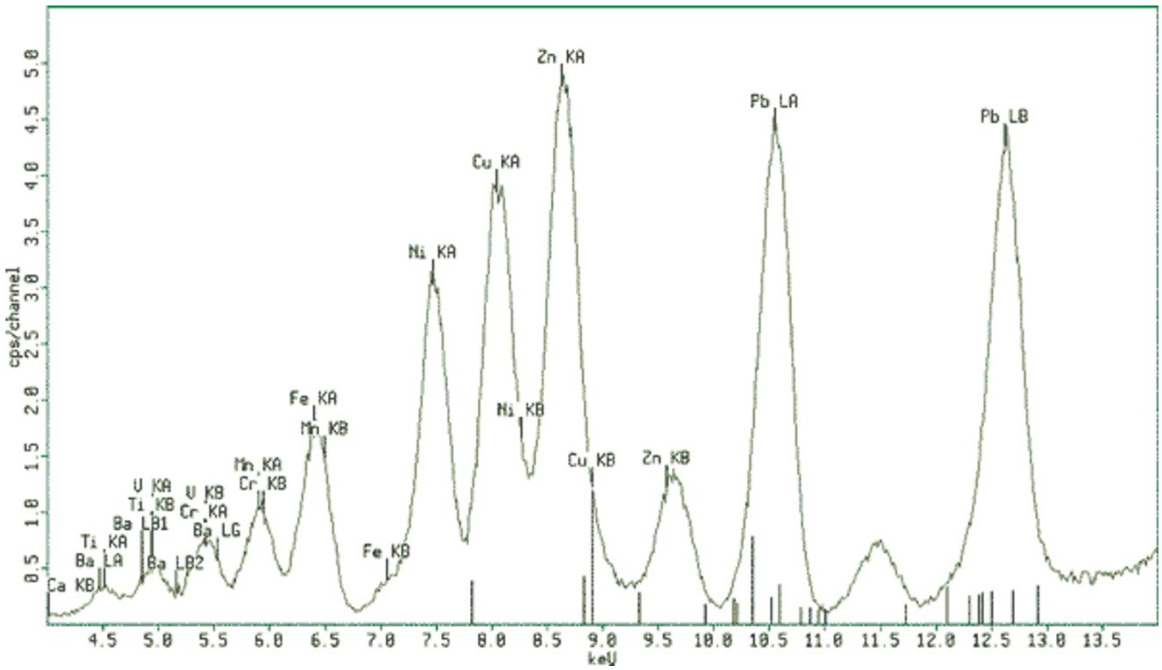

Figure 2.11: An example XRF spectrum with X-ray energy in keV on the x-axis and the number of X-rays observed for that energy level on the y-axis (Excerpted from Crumbling et al., 2008).

2.6.3 Detectors

In addition to the excitation source, a detector and empirical calibration software are major pXRF components. Together, these mechanisms make real-time elemental quantification possible using only the device with no external electronics unit required. Improvements in detectors with lower detection limits and advanced algorithms for spectral corrections have evolved the performance of pXRF technology, facilitating its applications in more arenas than ever before.

The two main categories of detectors are proportional counter detectors including scintillation and gas flow detectors, and solid-state semiconductors including Si-PIN diode and silicon drift detectors (SDD). Detector types vary in their resolution, which reflects their capacity to correctly distinguish the energy level of incident photons that are energetically similar to each other. Generally, a higher resolution output spectrum will have tighter peaks than a lower resolution spectrum with wider peaks. The peaking time refers to the amount of time between a voltage pulse and its peak (Fig. 2.12). A shorter peaking time allows more photons to be collected and distinguished from each other, resulting in more precise readings with increased count rates.

Proportional detectors convert characteristic fluorescence X-ray photons into voltage pulses, with the energy from incoming X-ray photons proportional to the output voltage. In scintillation detectors (also called indirect detectors) radiation interactions occur in a scintillation crystal, where incoming energy is converted into optical photons. These photons are collected in a photodetector and converted to electrical charges. Among the detector types, scintillation detectors offer a wide detection range for incident X-rays but have lowest resolution (Longoni and Fiorini, 2006). In contrast to indirect scintillation detectors, gas flow detectors are a proportional detector which directly converts photoelectric radiation into electric charges received by an output electrode (Longoni and Fiorini, 2006). This is achieved with the use of a cylindrical gas chamber (cathode component) which houses an anode component. When electrons in the cylinder are irradiated, they accelerate towards the anode component and become ionized via collision with the gas atoms. The signal measured by the collision is proportional to the incoming photon's energy. Gas flow detectors have an intermediate resolution between low resolution scintillation detectors and high-resolution solid-state semiconductors

Solid-state semiconductors offer multi-elemental analysis with high sensitivity. The possibility of a high-density ionization chamber to improve resolution was realized when high-purity silicon was used to create silicon lithium, Si(Li), detectors. Si(Li) detectors showed an improved resolution from the proportional detectors, but the need to house these detectors in cryostats for cooling made them large and challenging to handle

(Scholze, 2006). Thermoelectrically cooled Si-PIN detectors have since removed the limitation of these cryogenic cooling mechanisms. Within Si-PIN detectors, silicon crystals interact with incoming X-ray photons to create electron-hole pairs. Depending on the detector, these pairs are created for every ~3.6-3.8 eV of energy lost in the Si. The energy loss can be correlated with the energy of the incoming X-rays to create a spectrum of counts versus energy (Ametek, 2019).

SDDs far outperform Si-PIN detectors in regard to energy resolution and allow for detection of lower-Z elements. Within the detector, electrons ionized by X-rays are caused to drift towards a central anode component by means of an electric field parallel to the surface. The electric field is created with a series of concentric electrodes engraved in the surface (Potts and West, 2008). SDDs have a lower resolution (~40 eV), lower LOD (~3x), and lower peaking time than PIN detectors (Shields, 2020; Hullinger et al., 2009). SDDs can also count approximately ten times more X-rays per second than PIN detectors, making their analysis more sensitive and precise. The complicated equipment used in SDDs substantially increases costs, but the technology has consistently gotten less expensive as demand rises and production is streamlined. Solid-state conductors are preferred for pXRF instruments due to their small size, high resolution and count rates, and fast results (Kalnicky and Singhvi, 2001; Keng, 2015a).

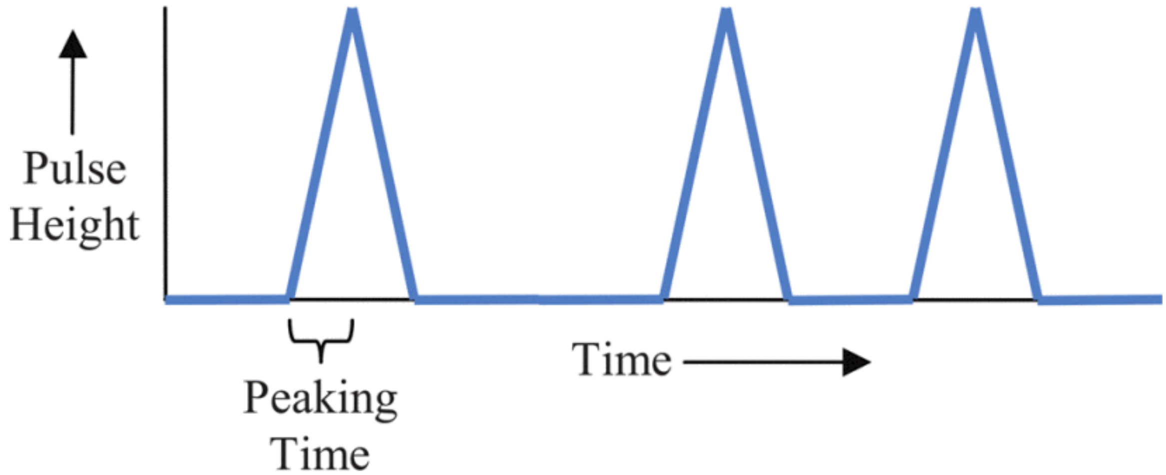

Figure 2.12: When an X-ray photon is absorbed in the detector, the voltage signal passes through a shaping amplifier in order to create peaks which can be distinguished for their energy and intensity (Excerpted from Hullinger et al., 2009) © 2009 IEEE.

2.6.4 Calibration software

Independent from the X-ray source and detector type, a method to accurately discern and quantify fluorescence is crucial for pXRF capabilities. In the past, on-site calibrations for field portable XRF instruments were necessary for specific sites and materials. pXRF instruments can now be calibrated internally using fundamental parameters (FP) established by the manufacturer or using Compton normalization, based off Compton peak ratios (USEPA, 2007).

The FP approach leverages X-ray theory to mathematically account for interelement effects and create quantitative algorithms for a certain sample type (Kalnicky and Singhvi, 2001). Sherman (1955) contrived a computational method from which FP is founded that related measured intensities to sample composition. Certain physical constants are fundamental for each element, such as mass absorption coefficients and excitation efficiencies, so elemental concentrations can be derived as a function of the measured X-ray intensities (Potts and West, 2008). Essentially, the FP approach iterates through and solves a system of equations for many unknowns. To simplify the calculation step, X-ray spectroscopists have used established theory and experiments to formulate approximations and linearize calibration equations. A number of robust algorithms now exist, which account for absorption and enhancement effects to accurately express sample concentrations.

The Compton normalization method relies on the analysis of one certified standard to normalize Compton peaks. Every sample spectrum has backscattered X-ray radiation (Compton scatter) present, but intensity of the Compton peaks varies with the matrix (USEPA, 2007). The ratio between analyte fluorescence intensity to the intensity of Compton scattered radiation for a particular reference material is the normalization factor used to calibrate the instrument. Because Compton scattering is highly dependent on the matrix, for efficacious measurements the SRM used for calibration should have a similar matrix and elemental concentrations to those in the samples being analyzed.

Modern instruments offer different modes of operation, which take advantage of the different calibration techniques. For instance, Soil Mode and GeoChem Mode are popular scanning modes for many pXRF instruments that differ in their calibrations. Soil Mode uses Compton normalization which works well in dilute samples where >85% of the sample is composed of light elements (LE) (elements lighter than magnesium) and no single element exceeds 2-3% concentration. Because soil materials are generally high in quartz, oxides, and organic materials (all LE), and low mineralization (heavy metals) Compton normalization calibrations have traditionally been applied for soil analysis (J. Litofsky, personal communication, January 23, 2022). However, the assumptions of Compton normalization calibrations (dilute samples and no inter-element interferences) are considerable drawbacks to operating in this mode, especially for ores and heavily mineralized samples. Additionally, Compton normalization/Soil mode cannot measure the concentrations of some analytes including Mg, Al, Si, or LE in a sample. Since this calibration approach is computationally simple, it was satisfactory for older instruments with limited processor power.

By comparison, GeoChem modes uses Fundamental Parameters calculations, which are more computationally intensive, but easily managed by modern processors. FP calibration/GeoChem mode is ideal for measuring across the range of concentrations of elements in sample, with the capability to discern ppm level detection in the presence of other elements in the percent range. Modern calibration software based on the fundamental parameters is 'standardless' because the versatile internal calibration detects concentrations from 0.1 ppm to 100% without requiring user input or numerous calibrants (Potts and West, 2008). Fundamental Parameters calibrations determine the total chemistry of the sample, including Mg, Al, Si, and LE.

2.6.5 Fluorescence mechanism

Exposing a material to short wavelength high energy X-rays can cause atoms to become ionized and fluoresce at specific energies. The reflorescence energies can then be categorized and quantified to construct an elemental profile. Incident X-ray photons produced within the device bombard the atoms in the sample and excite inner shell electrons which causes them to be ejected from their position in the K or L orbitals. The electrons will only be expelled from their orbital positions if the X-ray energy exceeds the binding energy for that electron. When outer shell electrons cascade down to regain atom stability by filling the inner shell void, energy is given off by the atom in the form of photons. The energy and intensity of the egressing fluorescence are measured as electric signals by the XRF (Sharma et al., 2014). The energy difference between the two shells is represented by Eq. 2.5, where ΔE is the characteristic X-ray energy, E_1 is the empty shell electron binding energy and E_2 is the donor shell electron binding energy (Kabir, 2013). ΔE and the corresponding energy peak produced by the transition are unique for each element, making qualitative identification of the elements in the sample possible.

Equation 2.5

$$\Delta E = E_1 - E_2$$

A multichannel analyzer produces a digital spectrum of XRF peaks for each element present so that these 'fingerprints' can be transformed to analytical data. The intensity of the reflorescence represents the number of photons being dislodged, to allow for quantitative determination of elemental concentrations.

Depending on which orbital shell an electron is vacated from, the X-ray emission can be classified as a K X-rays (n=1/K-orbital) or L X-rays (n=2/L-orbital). The X-ray emission can be further differentiated with α and β subscripts which indicate the orbital from which an electron cascades down from to fill the hole (Fig. 2.13). For instance, a transition from n=2 to 1 is a K_α X-ray and a transition from n=3 to 1 is a K_β X-ray. Similarly, a transition from n=3 to 2 is an L_α x-ray and from n=4 to 2 is an L_β X-ray (Bosco, 2013). A typical emission spectrum for each element has several peaks indicative of the energy difference between the electron transitions. Thus, measuring appropriate standards to ascertain the resultant peaks alongside unknown samples allows for the relative abundance of elements in the unknown sample to be determined.

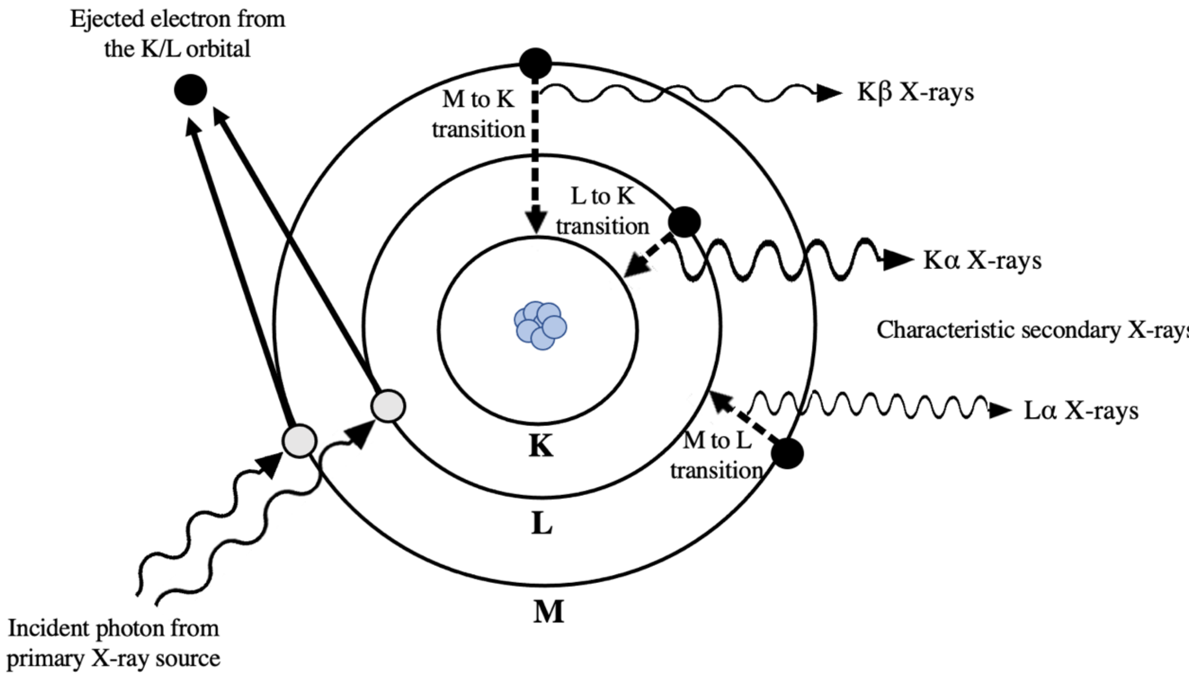

Figure 2.13: Electron transitions within an atom cause characteristic secondary X-rays to be emitted, the energy and intensity of which is measured by the pXRF. The dashed arrows represent ΔE, which is the difference in energy between the 2 quantum states of the electron (Adapted from Kalnicky and Singhvi, 2001).

2.6.6 Interaction of X-rays with matter

When X-rays come in contact with matter, three main interactions take place—fluorescence (photo-electric effect), Compton scatter, and Rayleigh scatter. The proportion of fluoresce to scatter depends upon the thickness (d), density (ρ), and chemical composition of the material (Fig. 2.14).

As discussed in the preceding section, fluorescence occurs when X-ray photons are absorbed in the material, causing an electron in the outer shell of an atom to cascade down to fill the spot of an electron which was ejected from its orbital position by incident X-rays. The fluorescence yield measured by the instrument is the ratio between emitted fluorescent photons and initial vacancies and is dependent upon the atomic number of the element.

Rayleigh and Compton scattering are the portion of X-ray photons which are not absorbed by the material. Rayleigh scattering, also referred to as coherent or elastic scattering, occurs when incoming photons hit electrons which are strongly bound in their orbitals, causing them to oscillate in place and emit radiation at the same frequency as incoming radiation. In this case, the photon's trajectory is deviated but there is no energy transfer (Beckhoff et al., 2006). Compton, or incoherent scattering, occurs when a photon hits an electron and transfers a fraction of its energy to the electron, causing the photon to move off with reduced energy and momentum. The amount of energy transferred is dependent on the angle at which the photon strikes the electron (Beckhoff et al., 2006). The sample composition affects the type and proportion of scatter that occurs (Fig. 2.15) (Potts and West, 2008). Light elements in sample materials cause a high proportion of Compton scatter and low Raleigh scatter because the electrons of these elements are loosely bound in their orbitals. Conversely, the interaction of photons with heavy elements where electrons are tightly bound, eliminates Compton scatter and leaves only Raleigh scatter (Brouwer, 2010). The scattered radiation can be absorbed by the detector, complicating spectrum interpretation.

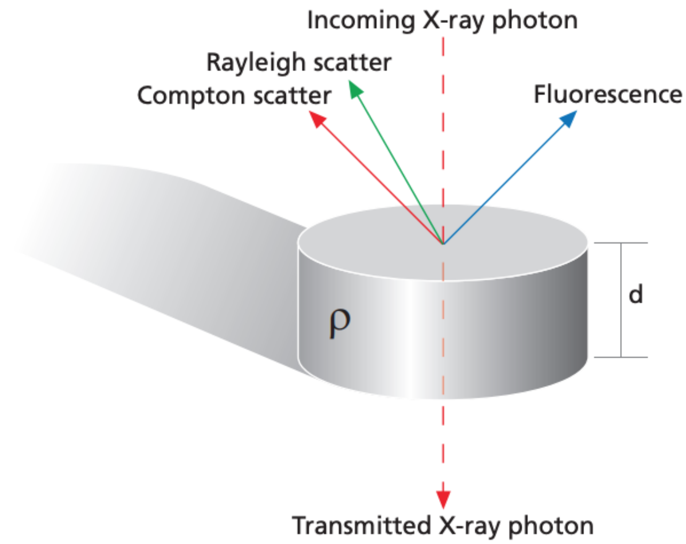

Figure 2.14: X-ray photons coming in contact with matter. While fluorescence returns characteristic measurable X-rays, some of the X-rays are scattered. It is also possible for transmitted photons to travel through the material without interacting with atoms in the material (Excerpted from Brouwer, 2010).

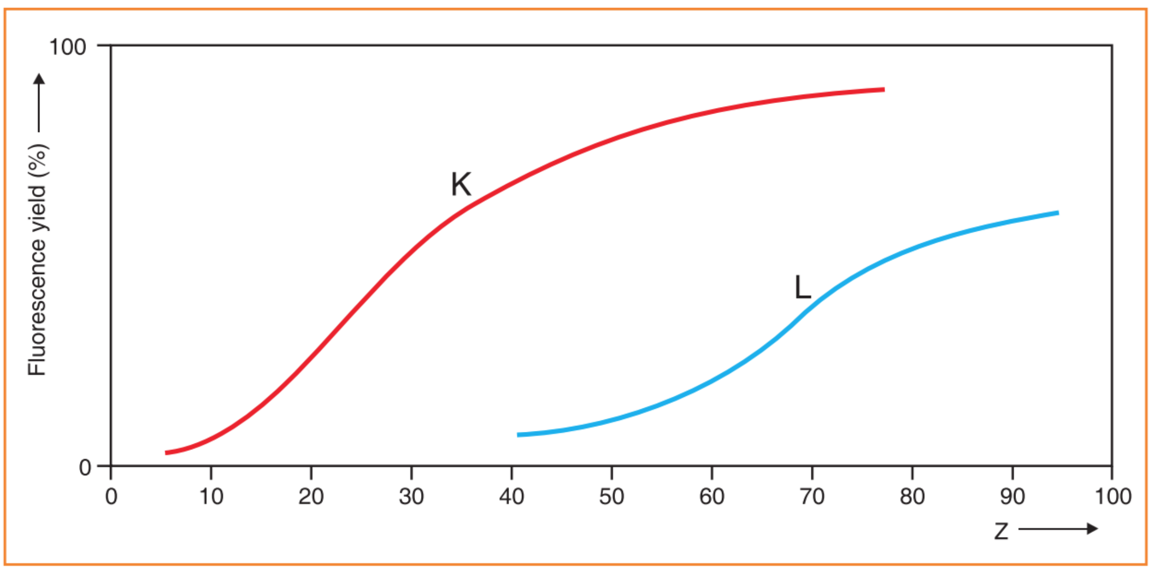

Figure 2.15: The fluorescence yield for K and L electrons. A low yield can be observed for light elements, which makes them difficult to detect and measure (Excerpted from Brouwer, 2010).

2.7 Factors influencing accuracy

2.7.1 Overview

An effective conversion of X-ray intensities into analyte concentrations requires consideration of interfering factors. While careful use of pXRF can give laboratory grade output, lack of consideration for best practices can lead to a misrepresentation of the true nature of the sample. Earlier studies comparing pXRF measurements to wet chemistry techniques correlated the two with variable success— casting doubt on the capabilities of pXRF. However, the EPA's Environmental Technology Verification Program (*ETV*), standard method 6200 (USEPA, 1998; USEPA 1995; USEPA, 2007) and subsequent studies have shown pXRF is capable of accurate repeatable analyses if proper sample preparation procedures mirroring those used for lab analyses are carried out (Radu and Diamond, 2009; Hall et al., 2011; Parsons et al., 2013).

pXRF measurements can be taken in-situ analysis where a profile face or cored samples are scanned directly without any sample preparation. For ex-situ measurements, sample preparation such as air-drying, grinding, and sieving typically precedes analysis. Using pXRF in-situ with soil under field conditions introduces performance variation that has been shown to result in elemental disparities (Ravansari et al., 2020). These sources of variation can include physical matrix effects (soil moisture, heterogeneity, sample thickness, and surface irregularity), chemical matrix effects (absorption and enhancement) and user error (instrument stability). The discrepancy between in and ex-situ measurements leading to error arises from the fact that pXRF instruments are calibrated using dried fine powder reference materials, with little pore space and flat surfaces while soil in situ has field moisture (typically ~5–50%), pore spaces, and an irregular surface (Potts and West, 2008).

2.7.2 Soil moisture

High-water content of a sample can cause X-ray scattering and a dilution effect. Soil water absorbs X-ray radiation from the pXRF and decreases the intensity of egressing fluorescence, which can result in artificially low elemental concentrations (Fischer et al., 2019, Schneider et al., 2016, Ge et al., 2005). Using a dataset of 215 samples, Schneider et al. (2016) found that the elemental concentrations measured by pXRF decreased exponentially as water content increased. Similar findings have been published for permafrost-affected soils. For in-situ analysis of Gelisols using pXRF analysis, Weindorf et al. (2014) found that for in-situ frozen soil, ex-situ re-frozen soil, and a melted soil/water mixture, elemental concentrations were significantly underestimated. They corrected for the denudation of secondary radiation by applying a moisture correction factor based on the total moisture content. Correction equations have also been used by others to effectively mitigate the dilution effect of soils that have a considerable moisture content (Ge et al., 2005; Shuttleworth et al., 2014). Though a high moisture content of field samples has been shown to reduce elemental intensity spectra, a moisture content between 5 to 20% results in minimal error overall (USEPA, 2007).

2.7.3 Soil organic matter content

Organic matter in the soil is crucial for providing nutrients to promote plant growth, increasing water holding capacity, decreasing bulk density, and increasing CEC. While most soils have some amount of organic matter, soils from natural ecosystems tend to have higher SOM levels than agricultural soils (Magdoff and van Es, 2021). However, SOM fractions have been shown to influence pXRF measurements by attenuating the fluorescence signal which in turn lowers the detection accuracy (Chen et al., 2021).

The calibration modes for pXRF instruments developed using organic-free matrices may not account for measurements deviations as a result of soil organic matter. For instance, scanning via a fundamental parameters calibration approach would be affected by OM changes due to the correction for light elements (Shad and Wendler, 2014). Shad and Wendler (2014) suggest that empirical calibrations employ certified reference materials that include organic soils and peaty soils.

Ransavari and Lemke (2018) tested the effect of adding 4 different organic matter surrogates to mineral soils and found that with increasing fraction of organic matter, pXRF concentrations for detected elements decreased due to a dilution effect. These measurement inaccuracies linked to the presence of organic matter would likely be accentuated with in-situ measurements of topsoils, which tend to be enriched with OM from plant residue.

2.7.4 Heterogeneity and sampling uncertainty

Efforts to characterize and quantify environmental properties contain uncertainty, which arises from both sampling and chemical analysis. Spatial heterogeneity, or random distribution of minerals in the environment, is the main source of uncertainty in pXRF data (Crumbling et al., 2010). Boon et al. (2007) assessed uncertainty of measurements in environmental applications and found that sampling contributes over 80% of measurement uncertainty, whereas the analytical component is usually less than 20% uncertainty of the total variance. In a study comparing in-situ and ex-situ pXRF with ICP analyses, Rouillon et al. (2017), found that sampling contributed over 95% of overall measurement errors. According to Ramsey and Boon (2012) since analytical uncertainty is much less important than sampling uncertainty, in-situ and ex-situ measurements can be practically equal in their reliability.

In the field, sample heterogeneity has been shown to have the largest impact on measurement accuracy when compared to laboratory analysis (USEPA, 2007). Since uncertainty from spatial heterogeneity exceeds uncertainty from analytical errors, the most effective way to reduce data uncertainty is to constrain spatial heterogeneity. The "nugget effect" is a phenomenon that can influence measurements when a chunk of soil or crystal of accessory phase minerals causes a particular analyte to be artificially concentrated and thus result in a deceptively high measured concentration (Steiner et al., 2017; Ravansari et al., 2020). Additionally, calculation of in-situ analytical bias is not advised due to the discrepancy between heterogenous surface samples in-situ and dried powder reference materials (Rouillon et al., 2017).

2.7.5 Particle sizes

For multiple samples with the same elemental matrices but different particle sizes, characteristic X-rays will vary in their intensities. Importantly, a sample with very fine particles will give a higher concentration of analyte than for a sample with coarse grains, and these effects are pronounced for low atomic number elements (Potts and West, 2008). For these reasons, in-situ measurements may be less accurate than ex-situ measurements which undergo sample preparation to negate these effects. In order to be consistent with the particle size across a sample set, samples should be ground and sieved to a uniform particle size. Ensuring that particles are a uniform size prevents falsely low and high analyte concentrations and improves accuracy of the readings. In a sampling cup, the physical soil matrix can cause elements to be under or overrepresented if particle sizes are not uniform. With unhomogenized samples, finer particles can settle to the bottom of the sampling cup causing their compositions to go unregistered. Although tedious, manual griding via a mortar and pestle can reduce the particle size of inorganic aluminosilicates to about 40 µm and may be the best option for small amounts of sample (Injuk et al., 2006). If using this method, care must be taken to ensure uniform particle sizes to prevent preferential absorption of secondary X-rays by contrasting particle sizes. Mills and mechanical grinders can also be used, but these require larger quantities of sample and can introduce contamination (Injuk et al., 2006). Homogenizing samples ex-situ via sieving and grinding to achieve a roughly uniform particle size helps to eliminate the fluctuation of field measurements where contrasting particle sizes can cause erroneous measurements.

2.7.6 Sample thickness

Sample thickness may influence elemental concentrations if the sample analyzed is not "infinitely thick." Infinite thickness refers to the minimum thickness that a sample must be to absorb penetrating X-ray beams and remit characteristic fluorescence. At an infinite thickness, 99% of the analyte's return X-rays are generated (Kalnicky and Singhvi, 2001). Critical penetration depth refers to the layer from which the intensity of the secondary X-rays from is measured by the instrument (Markowicz, 2011) and is calculated from Eq. 2.6, where ρ is the sample material's density, and μ_{tot} is the absorption properties of the sample.

Equation 2.6

$$t_{crit} = 4.61/(\rho\mu_{tot})$$

Below this depth, fluorescence photons have a high likelihood of being absorbed by the sample. The critical penetration depth will vary for different photons, since those with high energy penetrate deeper than those with lower energies (Potts, 1999). For example, the energy of the K-line for potassium is 3.31 keV and has a critical penetration depth of 0.03mm within an andesitic silicate rock. Cerium by contrast, has a K-line energy of 34.72 keV and a critical penetration depth of 9.6mm, within the same silicate rock (Potts, 1999). Practically speaking, this means that when the pXRF is used on a sample, the signal for potassium is coming from a layer in the sample much shallower than the layer from which the signal for cerium is coming. This is to say that the pXRF signal is derived from a specific and concentrated area, and disproportionate grain sizes in the bulk sample will affect the output analytics. Samples against a profile in-situ are always infinitely thick, however, a surficial layer of uncontaminated soil only 5mm thick could mask contaminated soil (or vice-versa) and lead to a misrepresentation of the true nature of the sample. Analyzing a sample in a cup that is not infinitely thick can result in artificially low readings, so cups should be filled at least ¾, allowing for sample thickness to remain consistent ("How to Test Soil for Lead," 2020). Typical sized pXRF sampling cups have an outside diameter of 30.7 mm, aperature size of 24.6 mm, and height of 22.9 mm, but larger cup sizes also exist (Chemplex Industries Inc., Palm City, FL).

2.7.7 Surface irregularity

Ideally, the surface of the sample will be entirely flat and aligned perfectly perpendicular to the analytical plane of the device. pXRF devices are calibrated with flat samples, so scanning an irregular surface will result in a reduction of the excitation and detection power owing to the inverse square law effect (Potts et al., 1997). An irregular surface can also introduce air attenuation which is especially consequential for elements with atomic numbers <20 (calcium and lower) (Potts and West, 2008). Without an appropriate correction factor to peak intensities, air gaps as small as 1-2mm can result in inaccuracies (Scholze et al., 2006). Correcting for the effects of surface irregularity can be achieved by determining a normalization factor using the Compton and Rayleigh scattered peak intensities for a limited range (a few mm) of surface irregularity (Marcowicz, 2011).

2.7.8 Chemical matrix effects

The measured concentration of an analyte depends not only on the abundance of that element, but also on the composition of the whole sample. As an X-ray beam travels through a sample, its intensity is affected by other elements in the matrix. Chemical matrix effects manifest as absorption of emitted X-rays (artificially dulling the intensity), and enhancement (artificially enhancing the intensity) (Beckhoff et al., 2006).

The characteristic radiation emitted from atoms when they are excited by incoming radiation is capable of expelling electrons from other atoms in the sample, causing them to fluoresce. While characteristic radiation that is directly produced by the X-ray source is primary fluorescence, secondary fluorescence refers to the characteristic X-rays emitted by atoms that were excited by primary fluorescence (Fig. 2.16). This indirect excitation can enhance the fluorescence intensities and exaggerate the measured count rate registered by the XRF device (Brouwer, 2010). Additional excitation from matrix elements occurs when an atom refluoresces at an energy higher than the critical absorption energy of other elements in the sample (Potts and West, 2008). For instance, the interaction between Fe and Zn can cause both absorption and enhancement effects. Since characteristic X-rays produced by Zn are absorbed strongly by Fe, the reported concentration of Zn may be artificially low. On the other hand, Fe can be enhanced by the characteristic X-rays of Zn, which have an energy close to the K absorption edge of Fe (Potts and West, 2008).

After the discovery of these matrix effects, numerous correction methods were developed, including the Lucas-Tooth and Pyne method, Lachance-Traill method, and Japanese Industrial Standards (JIS) correction (Lucas-Tooth and Pyne, 1963; Lachance and Traill, 1966; The Committee of Iron and Steel Standard Samples, 1982). Modern instruments however, use internal calibration software to correct for these intra- and inter-element interactions to accurately report the elemental concentrations (Glanzman and Closs, 2007).

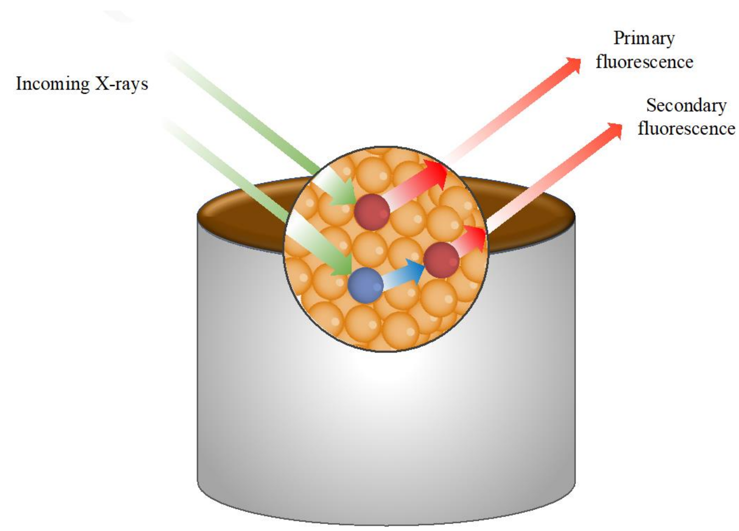

Figure 2.16: Inter-element secondary fluorescence occurs when characteristic X-rays produced by an atom are energetically efficient enough to excite electrons in the inner shells of other atoms in the sample (Adapted from Brouwer, 2010).

2.7.9 Scan time and detection limits

The detection limit (DL) or limit of detection (LOD) refers to the smallest amount of analyte that can be detected in a sample. On a spectrum, an element's peak element must be distinguished from and corrected for the background measurements (noise) beneath the peak intensities (Rousseau, 2001). Detection limits are a function of the specific method, sample preparation, and instrument, and therefore will depend upon the experimental setup and particular matrix (Mantler, 2006). The ability of the instrument to detect if an element is present in the sample or not above some given limit defines the LOD for that element.

Typically, the DL values published by instrument manufacturers represent the concentration equal to three standard deviations of the background intensity for a set of measurements (Rousseau, 2001). Practically, this means that to be considered 'detected,' the area under the peak for an element's signal needs to be at least 3x the background height (Fig. 2.17). The standard deviation (±3 sigma) which drives the LOD calculation is a function of the total number of counts. Therefore, there is a direct relationship between scan time (which determines the total number of counts) and the limit of detection calculation. As shown in Eq. 2.7, the relative standard deviation (σ_M/M) decreases with the number of counts (M) (Friedlander et al., 1981).

Equation 2.7

$$\frac{\sigma_M}{M} = \frac{\sqrt{M}}{M} = \frac{1}{\sqrt{M}}$$

Thus, a shorter scan time results in a higher standard deviation for the concentration of any element when compared to a longer scan time (Fig. 2.18). While a scanning time of 60 or 90s is common for soil analysis, longer scanning times have been shown to increase the accuracy of elemental concentration readings (Weindorf and Chakraborty, 2020). For instance, if aluminum has an LOD of 125 ppm with a 120 second beam condition, halving the beam time to 60 seconds while keeping all else equal would result in a LOD of 250 ppm. However, the increase in accuracy from longer scans must be weighed against the quantity of samples which can be analyzed in the same time frame. While it is true that a longer detection time improves detectability and can decrease measurement variability across replicate measurements (Ransavari et al., 2020), the associated gains in detectability will diminish at some point with increasing scan times reducing detection limits only by the square root of that factor (Potts and West, 2008). The detection limits are improved up until a point when the signal to background noise ratio becomes less optimal (Tighe et al., 2018; Killbride and Hutchings, 2006),

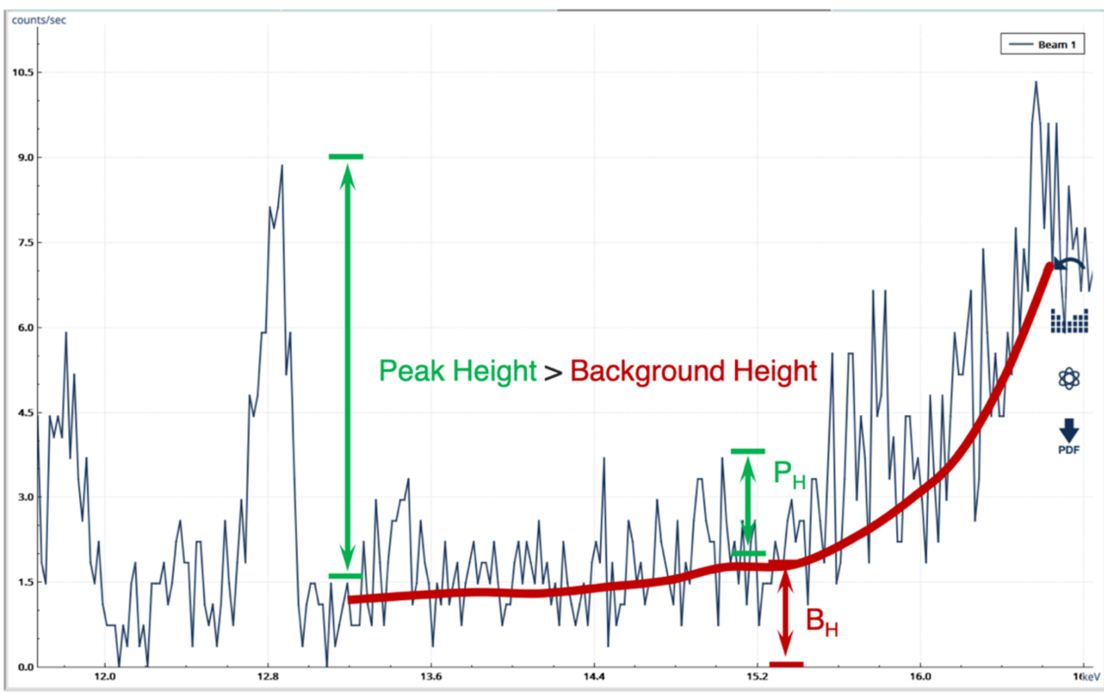

Figure 2.17: An established way to pick up on the detected elements is to only report those where the peak height is at least 3x the background height (Image from Olympus Scientific Solutions, How to Use and Understand LODs).

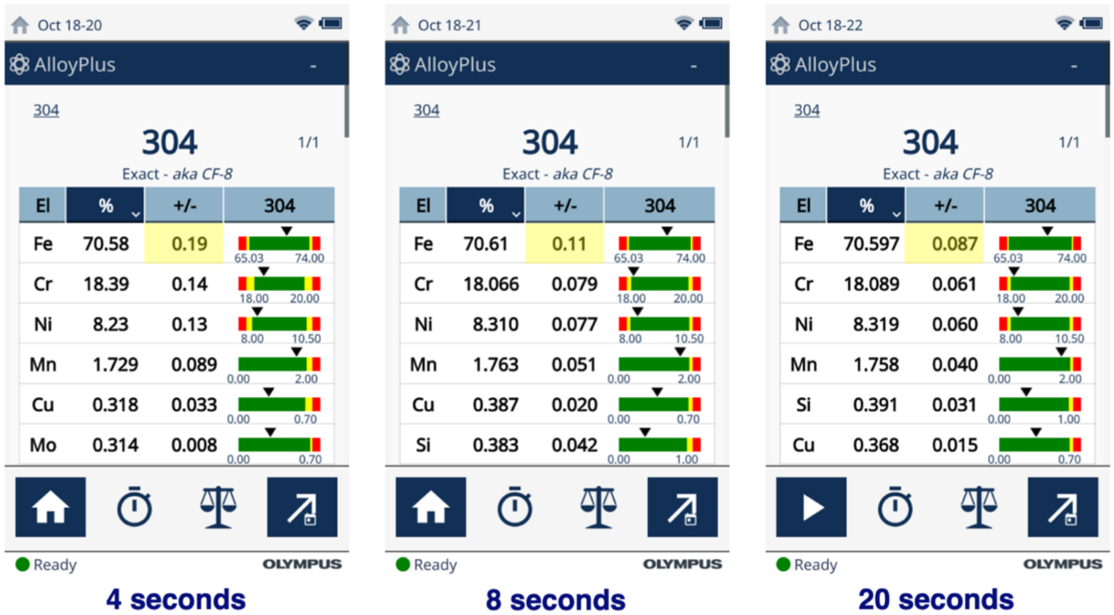

Figure 2.18: Increasing the scan time decreases the standard deviation of the elemental concentrations and captures their presence more consistently. For Fe, the margin of uncertainty decreases from ±0.19 with a 4 second scan to ±0.087 with a 20 second scan (Image from Olympus Scientific Solutions, PMI Workshop- Part 4 -XRF Statistics).

2.7.10 Fit for purpose

The use of quick and efficient in-situ measurements should be weighed against the higher accuracy of ex-situ pXRF analysis, taking specific goals of the study or investigation into account. In other words, the manner of pXRF operation should be tailored to and suited for the particular purpose for which it is being utilized.

In-situ analysis via pXRF greatly increases the uncertainty of the measurements due to the inherent interfering effects (heterogeneity, granularity, moisture, and surface irregularity), making these results only semi-quantitative (Markowicz, 2011). The sensitivity of analysis using a simple ‘point-and-shoot’ methodology is impeded by air space between the target and window while representativity is hindered by heterogeneous mineral soils (Leimere, 2018). However, these limitations are mostly removed in the case of fine-grained and naturally homogenous matrices, such as till. Sarala (2016) and Sarala et al., (2015) observed that pXRF used in Finland for geochemical mineral exploration of till correlated well to ex-situ pXRF and ICP-AES results. Further, Yuan et al. (2021) found in-situ pXRF measurements to produce comparable results to laboratory methods for most elements and were able to achieve high spatial resolution data to observe geochemical patterns resulting from weathering.

Ramsey and Boon (2012) challenged the traditional mindset that in-situ measurements are inherently less reliable than traditional laboratory ex-situ measurements by comparing pXRF measurements of As at a contaminated golf course to hydride generation-AAS method. Authors found that sampling uncertainty consisted over 93% of the total uncertainty for both in-situ and ex-situ measurements. They concluded that despite higher levels of uncertainty for in-situ measurements, as long as the uncertainty is quantified this sampling technique can be more fit for purpose than ex-situ measurements. For some instances, in-situ measurements with water, organisms, coarse fragments, and roots present may more accurately capture the target value than dried, sieved, homogenized, and chemically digested soil used for laboratory analysis, which can cause some bioavailable or volatile analytes to be lost (Ramsey and Boon, 2012). Additionally, high spatial resolution of in-situ measurements means that a more reliable site assessment can be achieved than would be possible with fewer ex-situ measurements. If accuracy of pXRF measurements is monitored, the confidence level of a high volume of pXRF samples will be higher than a handful of lab-analyzed samples, despite a higher analytical uncertainty of pXRF (Lemière, 2018). For the same cost as ICP analyses, much higher resolution sampling can be conducted with pXRF in-situ, which drives down sampling uncertainty considerably (Rouillon et al., 2017). Another advantage of in-situ measurements are the potential cost savings from storage, transportation, and disposal associated with ex-situ samples. For temporal studies and geographically extensive study areas, in-situ pXRF measurements offer tangible cost and labor savings over pXRF samples analyzed in a lab and may more accurately represent the entire sampling target. For instance, Rouillon et al. (2017) found that in-situ pXRF measurements for assessing metal contamination provided over twice as many samples for around half the cost of ex-situ and ICP analysis. Thus, consideration of the study objectives, level of accuracy needed, and economic constraints should be integrated into the sampling plan.

Initial assessments of pXRF instrument reliability were concerned with absolute accuracy of pXRF measurements, but as technology has improved, it has become clear that despite less accurate data, pXRF measurements still provide consistent data sets for geochemical analyses and spatial distribution of elements, with most of the inaccuracy emerging as a result of bias (Lemière, 2018). The use of pXRF for pedological classification or agricultural use to determine soil nutrients might be best served by in-situ measurements in the field to inform immediate decisions or home in on the areas of interest. Intergrade field preparation techniques between point and shoot and laboratory preparation can also be used to improve the precision of in-situ measurements. The “mole heap” technique consists of roughly homogenized loose media which is flattened (Lemière, 2018). A mortar and pestle may also be transported to the field to achieve more uniform particle sizes. In-situ pXRF analysis can be affected by hand movement instability which can change the analyte quantification, but this can be remedied by mounting the pXRF in a small transportable stand or ‘soil foot’ (Fig. 2.19) to allow for stable measurements and consistent sample positioning.

The question of if pXRF analysis is right for the task at hand should be decision comparability, rather than comparability to lab results, which are also imperfect and ‘wrong’ to some extent. This is especially true for methods which require acid digestions to determine the total elemental abundance, because an incomplete digestion can result in artificially low concentrations being reported. Also, despite an imperfect regression between field and lab data, pXRF analysis has the capability to estimate upper concentration limits (UCL) and exposure point concentrations (EPC) for contaminated sites (Crumbling et al., 2010). On the other hand, regulatory decisions that consider public health or environmental impact with a high level of accuracy and precision may be most appropriately determined via ex-situ laboratory confirmation.

Simple sample preparation including drying, sieving, and homogenizing before pXRF scanning still allows for quicker sample analysis than conventional techniques such as ICP-OES. A good way to balance accuracy with practicality is combining low-cost pXRF field measurements as the bulk of the data, with some systematic control analyses in the lab to improve the data quality (Lemière, 2018).

Figure 2.19: Use of a field portable pXRF mount can help address some of the error typical of in-situ analysis, like surface irregularity and sensor instability.

Chapter 3

MATERIALS AND METHODS

3.1 Sample collection

A set of soil samples (n=480) from throughout the state of California were assembled prior to pXRF analysis. The total sample set represents 809 km (~500 miles) across the state (Fig. 3.1). The availability of extra soil material from existing collections at Cal Poly and from other soils departments in the state influenced the resulting dataset. A wide array of samples from California were collected to determine the level of accuracy that could be achieved from predictive models built with soils from a vast geographic range with contrasting properties. For LA Urban, marine terrace, SPR/LBHC Mollisol, and some of the NRCS Chico samples, the land type was provided by the entity or individual who performed the soil sampling. For the NRCS Chico samples which were collected from Lassen National Park and for the UC Merced samples, the sample coordinates were input into Google Earth to identify the land type. The respective land types from which the samples were collected can be viewed in Table 3.1.

Due to the opportunistic sampling design of this study, the different sample sets were characterized by a handful of different labs. As a result, sample characterization was approached using different methods between the sample sets. While it is expected that this approach introduced variability into the lab ‘truth’ measurements, an array of acceptable standard methods was expected to contribute to a robust modeling dataset capable of linking elemental spectra to the soil property of interest in spite of typical analytical or inter-lab variability errors. Summarily, this study relies on the assumption that soils data collected across several labs can be directly compared.

A collection of 159 samples were amassed from a chronosequence of marine terraces at Swanton Pacific Ranch (SPR) in Davenport, CA. For 32 plots, samples were collected at four depths: 0-5 cm, 5-15 cm, 15-50 cm, and 50-100 cm. These samples are referred to as marine terrace samples for the remainder of this report. These samples were characterized by Cal Poly's soils labs for texture, SOC, N, and C:N and by A&L Laboratories for pH and CEC. A total of 39 surface samples were collected from urban forestry sites in Los Angeles, CA and are referred to as LA urban samples in this report. These samples were characterized for all properties by Cal Poly's soils labs.

A total of 218 samples were collected from Landel's-Hill Big Creek Reserve (LHBC) in Monterey County, CA and SPR in Davenport, CA as part of a study investigating the properties and management implications of Mollisols in forest and grassland environments (Clark, 2021). Both locations contained redwood forest and grassland ecosystems with mollic epipedons. At LHBC, 15 pits were established at the southern extent of the LHBC property and at SPR, 28 pits were excavated along five transects. The transects ran through several ecosystem types including redwood forest, mixed evergreen forests, coastal scrub, and coastal grasslands. At each pit, soil material was collected from three different depth classes: 0-10 cm, 10-25 cm, and 25-50 cm. These samples are referred to as SPR/LHBC Mollisol samples in this report. All soil properties assessed in this study were characterized by Cal Poly's soil labs.

An additional four agricultural soils were collected within the vicinity of the University of California Merced and analyzed by the University's soils department. The four samples come from the Atwater, Bear Creek, Alamo, and San Joaquin series. These samples are referred to as UC Merced samples in this report.

A set of 60 samples that were pre-characterized by the Kellogg Soil Survey Laboratory were obtained from storage at the NRCS Chico location. Of this collection, 16 were from Lassen National Park (Project C2008USCA016), 11 were from Lassen Volcanic National Park (Project C2007USCA026), four were from Shasta Co. (Project R2008USCA103), 14 came from Bay Delta MLRA 17 (Project C2016USCA033), three were from Bay Delta Soil Systems Study (Project C2017USCA083), 10 were collected from Bay Delta (Llano Seco) (Project C2014USCA050) and two were from DSP Sutter Co. Prune Orchard (Project C2015USCA019). These samples are referred to as NRCS Chico samples in this report. Lab data for each sample set can be found in Appendix A, and elemental data can be found in Appendix B.

Table 3.1: The land cover categories from which samples in this study were collected.

	Forest	Grassland	Urban	Agriculture	Bay delta	Marine terrace
Number of samples	168	81	39	6	27	159

While marine terrace samples were collected from a grassland environment, they were separated from the grassland category to maintain the distinction of close proximity to a sea cliff and generally higher sand contents.

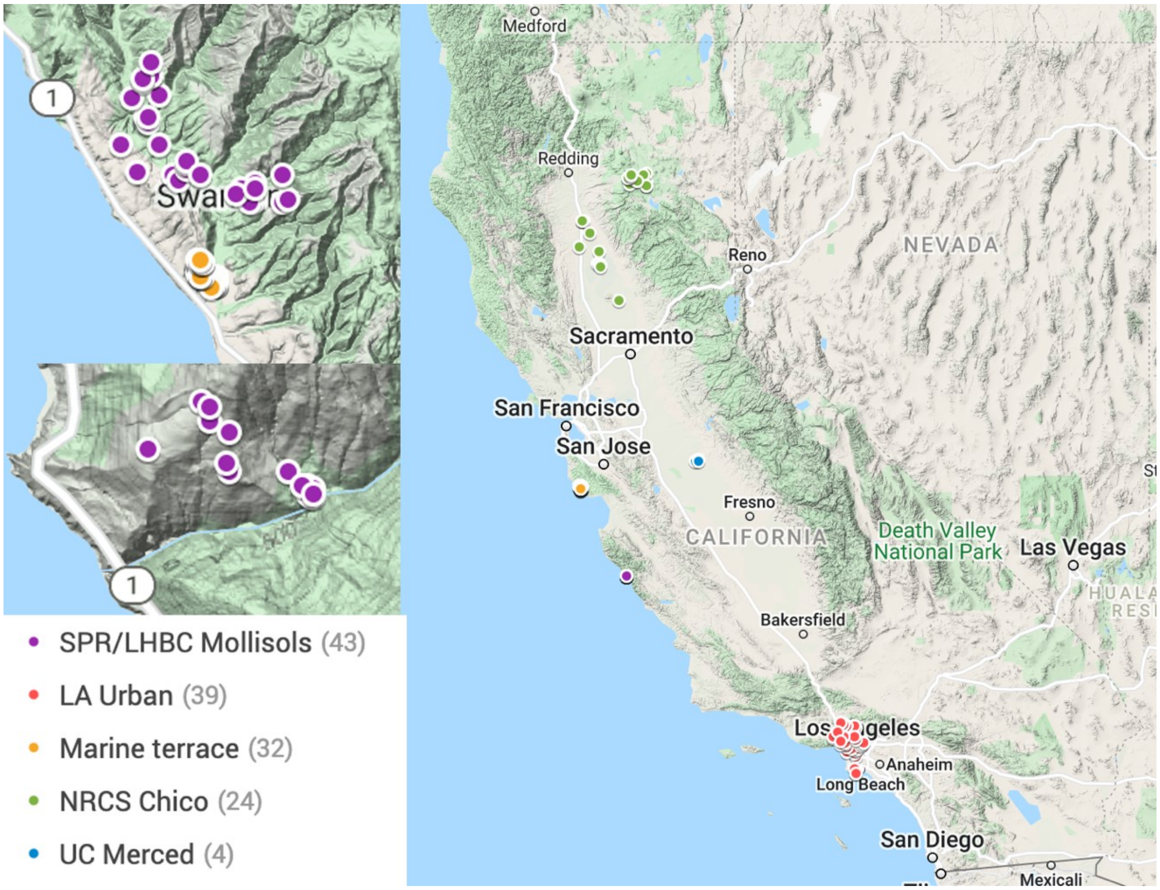

Figure 3.1: Map of sample locations color coded by sample set. Following each sample set is the number of sampling locations for that sample set. Where the number of sites is less than the total number of samples (SPR/LHBC Mollisols, Marine terrace, and NRCS Chico) multiple samples were collected at different depths in the soil profile.

3.2 Laboratory analysis

3.2.1 Sample preparation

Soil samples were left to air-dry prior to analysis and then sieved to ≤2mm to separate the fine earth fraction from any coarse fragments. A subsample of about 40g of the sieved soil was ground with a mortar and pestle into a fine powder to be used for CN and pXRF analysis. The sieved soils and finely ground subsamples were stored in labeled plastic bags on laboratory shelves.

3.2.2 pH

Measurements of pH for the LA urban, SPR/LHBC Mollisols, NRCS Chico, and UC Merced samples were carried out following the standard 1:1 water pH method (4C1a2a1) outlined in the Kellogg Soil Survey Laboratory Methods manual (Soil Survey Staff, 2014b). Prior to taking measurements, a three-point calibration with an acid, base, and neutral buffer solution was carried out to ensure proper pH meter functioning. To prepare the soil water solution, 20g of soil was mixed with 20mL of reverse osmosis (RO) water. The mixture sat for an hour to equilibrate and was stirred occasionally. After an hour, the pH electrode was inserted into the mixture, just above the soil sediment layer and the pH was recorded when the measurement stabilized. Three replicate measurements were taken with the pH probe and averaged to determine the solution pH in water. When the pH electrode was not being actively used, it was kept in a storage solution.

For the marine terrace samples, the saturated paste method (S - 1.10) was carried out as described in the *Soil, Plant, and Water Reference Methods for the Western Region manual* (Gavlak et al., 2005). This approach used a 200g of air-dry soil and deionized water to create a saturated paste which meets the following criteria:

- Does not have free standing water on the surface of the paste.
- Soil paste slides freely and cleanly off a spatula (excluding soils with >40% clay).
- Paste will flow slightly when the container is tipped.
- Soil surface glistens as it reflects light.
- Consolidates easily by tapping after a depression is formed in the paste with a spatula (excluding soils with >70% sand).

After allowing the saturated paste to equilibrate for 4 hours, the saturation characteristics were checked again to ensure they still meet the requirements of a saturated paste. A pH meter equipped with electrodes was standardized to 3 different buffers (pH 4, 7, and 10). The meter was inserted into the soil paste and allowed to stabilize, after which the pH was recorded.

3.2.3 Particle size analysis

In order to determine the textural classes and relative proportions of sand, silt, and clay for LA Urban, SPR/ LHBC Mollisols, marine terrace, and UC Merced samples, particle size analysis (PSA) was carried out using the hydrometer method (S - 14.10) from *Soil, Plant and Water Reference Methods for the Western Region* (Gavlak et al., 2005). Pretreatment of soils for removal of soluble salts, organic matter, carbonates, and iron oxides was not carried out due to logistical constraints. Forgoing H_2O_2 pretreatment for the samples was not anticipated to cause major impacts on the PSA results. While salt and carbonate levels were mostly negligible, some Mollisol soils did have high levels of organic matter, which may have impacted the hydrometer readings of these samples. However, conflicting evidence about the overall impact of forgoing pretreatment exists (Callesen et al., 2018; Ferro and Mirabile, 2009). The degree of error arising from this decision was expected to be minor, compared to inherent theoretical and sampling errors of hydrometer analysis (Black, 1951).

To perform the PSA analysis, a volume of 40 ± 0.05g of sieved and air-dried soil was weighed on an electric balance and placed into metal dispersing cups. The weight of soil was recorded to later correct for the air-dry/oven-dry ratio (AD/OD). Then, 100ml of sodium hexametaphosphate (HMP) solution was added to each cup and samples were left to equilibrate overnight. Extra deionized (DI) water was added to the cups so there was enough solution for the mixing attachment to reach. The dispensing cups were attached to an electric mixer (Hamilton Beach Scovill Model 936-Drink mixer) and set to mix on high speed for five minutes. Additional DI water was used to rinse the soil from the cup into a clean sedimentation cylinder and brought to 1L volume using DI water. A reference cylinder with 100ml HMP brought to 1L volume with DI water was prepared as a reference blank. In some samples, a thick layer of foam formed at the top of the sedimentation cylinder, requiring a couple drops of amyl alcohol to be added as a foam reducer. Using a plunger (rubber disk attached to a rod) the samples were thoroughly mixed using an up and down motion for one minute. Immediately after mixing, a standard hydrometer with the Bouyoucos scale in g L^{-1} was placed into the cylinder and at read at the upper edge of the meniscus surrounding the stem to the closest ±0.05 g/L after 40 seconds. This measurement was recorded and represented the clay and silt fraction (R_{sand}) that was suspended in the cylinder. A measurement of the blank was taken in this same manner (R_{C1}). Samples were allowed to thermally equilibrate based on the temperature of the blank solution. After six hours, the temperature of the blank was taken and used to determine the settling time for clay (ranging from 6 hours and 27 minutes for 28°C to 8 hours and 9 minutes for 18°C). When the settling time was reached, the second density measurement was taken for each sample and the blank (R_{C2}). The second measurement was carried out by placing the clean hydrometer in the cylinder gently as to not disturb the settled particles and again reading to the nearest ±0.05 g/L. The second measurement represented the clay fraction (R_{clay}) of the sample suspended in the cylinder. The air-dry/oven-dry (AD/OD) ratio was determined using method 3D1 from the Kellogg

Soil Survey Laboratory Methods Manual (Soil Survey Staff, 2014b) and used to adjust the air-dry sample weights. Hydrometer measurements of samples and the blank were used to determine the percent sand (Eq. 3.1), clay (Eq. 3.2), and silt (Eq. 3.3).

Equation 3.1

$$Sand\ \% = \frac{(OD\ soil\ mass) - (R_{sand} - R_{C1})}{OD\ soil\ mass} \times 100$$

Equation 3.2

$$Clay\ \% = \frac{(R_{clay} - R_{C2})}{OD\ soil\ mass} \times 100$$

Equation 3.3

$$Silt\ \% = 100 - (Sand\ \% + Clay\ \%)$$

For NRCS Chico samples, standard KSSL PSDA method 3A1a1a (Soil Survey Staff, 2014b) was used to determine soil texture. For this method, 10g of soil was pretreated to remove soluble salts and organic matter. The sample was oven-dried overnight at 110°C to obtain an initial sample weight unaffected by moisture. Then, 10mL of HMP solution and 175mL of RO water was added to the sample and placed on a horizontal shaker set to 120 oscillations per minute to be shaken overnight. Next, the sample was wet sieved using a 300 mesh (0.047mm) sieve to separate the silt and clay fraction (collected underneath the sieve) from the sand fraction which remained on top of the sieve. The silt and clay fraction were transferred to a 1L cylinder and brought to 800mL volume with RO water. A watch glass was placed on top of the cylinder and left to equilibrate overnight while the sand fraction was transferred to an evaporation dish to dry in the oven overnight. The dry sand was sorted with a stack of sieves (in descending order: 1, 0.4, 0.25, 0.1, and 0.047mm) and the sand fraction remaining on top of each sieve was weighed. Clay and silt contents in the 1L cylinder were determined gravimetrically using a 25mL Lowry pipette mounted to an adjustable pipette rack. The temperature of a prepared blank solution was recorded. A hand-stirrer was then used to mechanically stir the silt and clay solution for five minutes and then for an additional 30 seconds using an up and down motion. For the <20μm fraction, an aliquot was retrieved at a depth of 10cm into the suspension and placed into a weighing bottle. For the <2μm fraction, aliquots were retrieved at 4.5, 5, 5.5, or 6.5 hours. A second temperature of the blank was recorded and used to adjust for the pipette depth into the solution. The collected aliquots were oven dried overnight and weighed as residue weights. Particle size fractions could then be determined using Eq. 3.4–3.7, where RW_2 is the <2 μm fraction residue weight, *DW* is the weight of the HMP dispersing agent, *CF* is 1000mL/DV, *DV* is the dispensed pipette volume, *TW* is the total weight of the oven dry sample, RW_{20} is the <20μm residue weight, and SW_i are the weights of the sieved sand fractions.

Equation 3.4

$$Clay\ \% = 100 \text{ x } [(RW_2 - DW) \text{ x } (CF / TW)]$$

Equation 3.5

$$Fine\ Silt\ \% = 100 \text{ x } [(RW_{20} - DW) \text{ x } (CF / TW)] - Clay\ \%$$

Equation 3.6

$$Sand\ \% = \sum (SW_i / TW) \text{ x } 100$$

Equation 3.7

$$Coarse\ Silt\ \% = 100 - (Clay\ \% + Fine\ Silt\ \% + Sand\ \%)$$

3.2.4 CEC

CEC for SPR/LHBC Mollisols samples was determined via a two-step extraction process as described by Clark (2021) which drew on methods taken from the KSSL manual chapter 4B1a (Soil Survey Staff, 2014b), method S - 14.10 and S - 10.10 in *Soil, Plant, and Water Reference Methods for the Western Region* (Miller et al., 2013), and the *Fall 2018 Soil and Water Chemistry Laboratory Manua*l for the California State Polytechnic University of San Luis Obispo (Appel and Stubler, 2018). For this method, 2.5g of soil was combined with 35mL of pH 7 1 *M* ammonium acetate (NH_4OAc) in a centrifuge tube. The samples were placed on an oscillating shaker at 180 cycles/minute for 30 minutes then centrifuged at 2000 rotations/minute. The basic cation extract was filtered from the soil and another 25mL of ammonium acetate was added to ensure full saturation of soil cation exchange sites. Excess ammonium that was unbound to exchange sites was washed from the solution with isopropyl alcohol three times. Then, 35mL of 2 *M* KCL was added to the tube, which was shaken and centrifuged at the previously mentioned settings. The supernatant was decanted into a scintillation vial and frozen until colorimetric analysis. Prior to analysis, extracts were diluted 45x with 2 *M* KCl to ensure absorbance readings would be within the limit of detection for the instrument. Several aliquots were also added to the extract, including two reagents necessary for the colorimetric reaction. Absorbance readings were obtained using Ocean Optics UV-VIS at 650 nm. The results of this method were validated by measuring exchanged ammonium with an ammonia gas electrode.

For LA Urban and UC Merced samples, CEC was determined using an adapted version of UN-FAO methods as described in the *Fall 2019 Soil and Water Chemistry*

*Laboratory Manua*l for the California State Polytechnic University of San Luis Obispo (Appel and Stubler, 2019). In this method, 2.5g of soil was weighed into a falcon tube and combined with 25mL of pH 7 1 *M* NH_4OAc. The samples were shaken for 30 minutes on an oscillating shaker (New Brunswick Scientific Innova 2100 Open-Air Platform Shaker) at 180 cycles per minute for 30 minutes. After being shaken the samples were centrifuged (Eppendorf 5810R Centrifuge, serial no. 0034398) at 2000 rpm for five minutes. The supernatant was poured off of the samples to discard excess unbound ammonium. To rinse off any remining ammonium in the pore water of the sample, 25mL of 91% isopropanol was added to the tubes as a cleansing solution. It was necessary to resuspend the soil pellet that formed at the bottom of the tubes after centrifugation using a vortex mixer (Thermo Scientific, Vortex Maxi Mix II). The soil and isopropanol mixture was placed on the oscillating shaker for five minutes and then centrifuged for five minutes at the previously mentioned settings. Once again, the supernatant was poured off, and another round of cleansing with the isopropanol solution was performed (vortex mixer, oscillating shaker, centrifuge). After the supernatant was poured off for the last time, an additional 10mL aliquot of isopropanol was added to the centrifuge tube to re-suspend the soil pellet so the contents could be transferred to a crucible. Additional isopropanol was used to rinse any remaining soil from within the centrifuge tube. The crucibles were placed in a drying oven set to 30°C for 24 hours. For those samples with coarse particles, the dried soil was ground via mortar and pestle. Then, 1000 ± 250mg of the dried soil was weighed into crucibles for analysis of total N in an Elemantar Vario MAX Cube CN analyzer (Elemetar, Langenselbold, Germany; serial no. 29191038). For every 10 samples, a standard reference material B2178, “Medium Organic Content Sediment” (Elemental Microanalysis Limited) was run for continuing calibration verification. To determine CEC from %N, the total %N was converted to a weight (mg/kg) and then to cmolc/kg. Two example calculations are shown below.

LA Plot 4 (Loamy sand, 6.29% clay)

$$0.087\ \%N \times 10{,}000 = 870\ mg/kg$$

$$\frac{870\ mg}{kg\ soil} \times \frac{cmolc\ NH_3^+}{0.17g\ NH_3^+} \times \frac{1\ g}{1{,}000\ mg} = 5.12\ cmolc/kg\ soil$$

LA Plot 169 (Loam, 25.84% clay)

$$0.728\ \%N \times 10{,}000 = 7{,}280\ mg/kg$$

$$\frac{7{,}280\ mg}{kg\ soil} \times \frac{cmolc\ NH_3^+}{0.17g\ NH_3^+} \times \frac{1\ g}{1{,}000\ mg} = 42.82\ cmolc/kg\ soil$$

Determination of CEC for NRCS Chico samples was carried out following method 4B1a1a as described in the KSSL manual for CEC7 (Soil Survey Staff, 2014b). For this procedure, exchange sites were saturated with an index cation (NH_4^+) using 1 *M* pH 7 NH4OAc solution applied with a mechanical vacuum extractor. The soil was washed with ethanol to remove unabsorbed NH_4^+. Then, the sample was rinsed with 2 *M* KCl and the leachate was analyzed using steam distillation and titration to determine CEC to the nearest 0.1 cmolc/kg soil.

To determine CEC for marine terrace soils, the ammonium replacement method (S - 10.10) as described in *Soil, Plant, and Water Reference Methods for the Western Region* (Gavlak et al., 2005) was performed by A&L laboratory in Modesto, CA. This approach involved weighing 10 ± 0.1g air-dry soil into a 125mL Erlenmeyer flask. Then, 50mL of pH 7 1 *M* NH_4OAc was added to the flask and placed in a reciprocating shaker for 30 minutes. The solution was transferred to a Bucher funnel fitted with Whatman No. 5 filter paper. A 1L vacuum extractor was connected to the Buchner funnel and the solution was leached using 175mL NH_4OAc. The excess NH_4OAc was rinsed from the soil solution in the Buchner funnel with 200mL of ethanol. After rinsing soil in this manner, exchangeable ammonium was replaced by attaching the funnel to a 500mL suction flask and leaching the solution with 225mL of 0.1 *M* HCl. The leachate was brought to a volume of 250mL using DI water and then analyzed for ammonium concentration with an ALPKEM rapid flow analyzer. The analyzer measures indophenol blue at 660 nm produced by the complexation of ammonium and salicylate intensified with sodium nitroprusside. The basic cation concentrations (K, Mg, Ca, Na) of marine terrace samples were also determined by A&L laboratory and reported in ppm. To determine the base saturation of these samples (Eq. 3.8), the ppm of each cation was converted to cmolc/kg soil (Eq. 3.9) and then divided by the CEC, as shown in the example calculation below.

Equation 3.8

$$cation\ concentration\left(\frac{cmolc}{kg\ soil}\right) = (ppm\ of\ cation) \div \left(\frac{atomic\ mass\ of\ cation\ \times\ 10}{charge\ of\ cation}\right)$$

Equation 3.9

$$Base\ saturation\ = \frac{Sum\ of\ basic\ cations\left(\frac{cmolc}{kg}\right)}{CEC\left(\frac{cmolc}{kg}\right)}$$

Marine terrace sample #1

$$121\ ppm\ K \div \left(\frac{39.098\times\ 10}{1}\right) =\ 0.31\ \frac{cmolc}{kg\ soil}\ K$$

$$224\ ppm\ Mg \div \left(\frac{24.305\times\ 10}{2}\right) =\ 1.84\ \frac{cmolc}{kg\ soil}\ Mg$$

$$950\ ppm\ Ca \div \left(\frac{40.078\times\ 10}{2}\right) =\ 4.74\ \frac{cmolc}{kg\ soil}\ Ca$$

$$49\ ppm\ Na \div \left(\frac{22.990\times\ 10}{1}\right) =\ 0.21\ \frac{cmolc}{kg\ soil}\ Na$$

$$Base\ saturation\ = \frac{121\ +\ 224\ +\ 950\ +\ 49\left(\frac{cmolc}{kg}\right)}{15.9\left(\frac{cmolc}{kg}\right)} = 44.70\%$$

3.2.5 SOC, TN, and C:N ratio

Total carbon (TC) and nitrogen contents for the LA urban, SPR/LHBC Mollisol, and marine terrace samples were measured via combustion using an Elemantar Vario MAX Cube CN analyzer. Using an analytical balance (Mettler Toledo, Columbus, OH) CN tube crucibles were filled with 1000 ± 100mg of finely ground soil. For method level quality control, all LA Urban samples were run in duplicate, and 10% of SPR/LHBC samples were duplicated. Empty crucibles were used as blanks and organic analytical standard B2178 was run every 10 samples for continuing calibration verification.

NRCS Chico samples were analyzed by the KSSL for total carbon and nitrogen via combustion techniques 4H2a1 and 4H2a2 (Soil Survey Staff, 2014b). In this method, samples were subjected to high temperatures in an oxygenated CO_2 environment within an elemental analyzer using catalytic tube combustion. The N_2 and CO_2 gases released from the sample were distinguished from each other by adsorption columns and measured using a thermal conductivity detector.

To determine soil organic C content (SOC) for SPR/LHBC Mollisols, a correction to the % TC determined via combustion was applied. Assuming any organic C was the result of calcium carbonate ($CaCO_3$) dissolution, the difference between TC and inorganic C contributed by carbonates was calculated to find the organic C fraction (Soil Survey Staff, 2014a) (Eq. 3.10). A correction was applied to soils that exceeded 120% base saturation (BS) (typically, 100%, but 120% was used for these soils due to error associated with basic cation extractions). To find SOC, the difference between the exchangeable charge and the extracted basic charge was assumed to be associated with calcium carbonate (Eq. 3.11) and was subtracted from TC (Clark, 2021).

Equation 3.10

$$Soil\ organic\ carbon\ \% = Total\ C\ \% - Inorganic\ carbon\ \%\ associated\ with\ CaCO_3$$

Equation 3.11

$$Inorganic\ carbon\ \%\ associated\ with\ CaCO_3 = Ca\ associated\ with\ CaCO_3 \left(\frac{cmolc}{kg\ soil}\right) \times \frac{1\ cmol\ Ca}{2\ cmolc\ charge} \times \frac{1\ mol}{100\ cmol} \times \frac{1\ mol\ CaCO_3}{1\ mol\ Ca} \times \frac{1\ mol\ C}{1\ mol\ CaCO_3} \times \frac{12\ g\ C}{1\ mol\ C} \times \frac{1\ kg}{1000\ g} \times 100$$

For marine terrace samples, the calculated base saturation (as described in the CEC determination section) was used to infer SOC content. If the base saturation exceeded 100%, as was the case for three samples, the sample was discarded due to the presence of carbonates and no way to correct for them. If base saturation was <100%, SOC was assumed to equal TC.

NRCS Chico samples that were classified by the Kellogg National Laboratory, had total carbon and nitrogen, estimated SOC, and CN ratio values reported (methods 4H2a1 and 4H2a2). The estimated organic carbon was calculated using Eq. 3.12. In the case that carbonates were determined (method 4E1a1a1) and reported, TC could be corrected for carbonates. No data in the carbonates section was assumed to mean no carbonates were present in the sample since most samples that go through KSSL are requested for the calcium carbonate equivalent analysis. Thus, no data in that section generally indicates that the sample didn't meet the pretest criteria to be analyzed for carbonates (S. Murphy, personal communication, 30 March 2022). While TC values are reported to the hundredths place, OC values were reported to the tenths place, causing OC values to be higher than TC values in some cases. In the case that there was no data for calcium carbonate equivalent on the sample report, no carbonates were detected, or OC was higher than TC, SOC was reported to be equal to TC. If carbonates were detected in trace amounts or more, the OC value was used for SOC.

Equation 3.12

$$Soil\ organic\ carbon\ \% = Total\ carbon - (CaCO_3 \times 0.12)$$

For LA Urban Samples, soil organic carbon was determined by using the CN analyzer set to a reduced temperature (Pitt et al., 2003). Subjecting the samples to a temperature between 600-650°C results in combustion of the organic fraction, while preventing the loss of inorganic carbon. The same sample preparation process as total C and N was followed for loading the samples (1000 ± 100mg of finely ground soil placed into CN tube crucibles). Empty crucibles were used as blanks and natural reference material B2188 (Elemental Microanalysis) was run every 10 samples for continuing calibration verification. SOC, TN, and C:N ratio were not determined for UC Merced samples. The C:N ratio of all samples for this report is represented as the ratio between SOC and total nitrogen.

3.3 pXRF sample preparation and analysis

Samples were analyzed according to the manufacturer's recommended instructions by placing 3-5g of finely ground soil into XRF sample cups followed by pXRF analysis in a stand mount. A mortar and pestle were used to manually grind the samples into a fine powder. The powder was transferred into double open ended XRF sample cups with caps and a serrated snap-on ring (Cat. No. 1330-SE, Chemplex Industries Inc., Palm City, FL). The soil powder was packed into cups tightly to avoid air pockets which can lead to X-ray attenuation. Cups were sealed using 4 µm Prolene™ films (Cat. No. 426, Chemplex Industries Inc., Palm City, FL), which contain fewer impurities than Mylar™ and Kapton™ films (Laperche and Lemière, 2021).

Figure 3.2: A batch of pXRF cups packed with finely ground soil ready to be scanned (Photograph by the author).

For all samples, scanning was conducted with a handheld VMR model Vanta M Series XRF analyzer (S/N 801741, Olympus, Waltham, MA). The instrument uses a 4-watt X-ray tube with a rhodium (Rh) anode material as an excitation source and was operated at 40 and 10 kV with a large area Silicon Drift Detector (165eV). The GeoChem calibration was used, with two sequential beams set to scan the soil samples for 30 seconds each, so that every complete scan took one minute. The 10kV beam analyzed magnesium up to titanium while the 40kV beam analyzed titanium and heavier elements. Given the greater robustness of a fundamental parameters calibration, as discussed in Chapter 2, it can be reasoned that most applications are better served with GeoChem than with Soil mode, which is why GeoChem mode was used in this study.

Before scanning the samples in each batch, a clean wipe was used to gently wipe off the lens of the pXRF, the quartz blank, and the top films of the standards and samples. An internal calibration or ‘Cal Check’ was performed before the start of each scanning session. To perform this check, the pXRF was placed in the instrument docking station and the Cal Check was initiated. This step checks the detector, X-ray tube, filter wheel, and safety features of the analyzer. After passing the Cal Check, the SiO_2 (quartz) blank was scanned to check for contamination on the analyzer window (EPA, 2007).

To account for within sample variability, differences in particle size and packing density at the small area analyzed by each scan, each sample was reoriented between the two scans (Towett et al., 2015). Accuracy at the instrument level was evaluated with initial calibration verifications and continuing calibration verifications every 10 scans using standard reference materials 2711a “Montana II Soil Moderately Elevated Trace Element Concentrations” and 2706 “New Jersey Soil, Organics and Trace Elements” (NIST, Gaithersburg, MD). The chemistry results for each scan session were exported from the Vanta XRF Analyzer PC Software to .csv files. The results of standard scans were used to monitor precision over time, as discussed further in the Instrument Quality Control section.

3.4 Data processing

After each round of pXRF scans, the chemistry results were exported as .csv files. The .csv exports were manually combined in Excel (Version 16.54, Microsoft, 2021) for each sample set and an extra column labeled "Unique Sample ID" was added to easily average replicate scans and link lab data to pXRF data. Scans of the standard soils, 2711a and 2706, were exported as a .csv file to monitor instrument precision over time.

The combined .csv files were loaded into RStudio for the remaining data processing and analysis (Version 1.3.1093) (RStudio Team, 2020). To average replicate scans into one value for each elemental concentration, a series of iterations through the combined dataframe selected the subset of scans that had matching Unique Sample ID values. For rows with matching Unique Sample IDs, the non '<LOD' values of each column were averaged. If one concentration was '<LOD' and the other concentration was detected, only the detected value was used. In the case that both readings were '<LOD', that elemental concentration was not captured and was henceforth recorded as 'NA'. An example of the concentration averaging is shown in Tables 3.2 and 3.3. The averaged elemental concentrations were used for subsequent modeling and can be found in Appendix B.

Table 3.2: A few rows of the raw replicate scan data loaded into RStudio for averaging.

Unique Sample ID	Mg Concentration	Al Concentration	Si Concentration	P Concentration
NRCS1	5561	72417	228626	1370
NRCS1	<LOD	73645	227576	1311
NRCS2	4116	93591	223757	1052
NRCS2	4874	94745	238797	988
NRCS3	8628	72748	234682	1450
NRCS3	8713	75172	237753	1461
NRCS4	<LOD	102692	216816	365
NRCS4	<LOD	103703	218894	269
NRCS5	19759	67697	205774	852
NRCS5	21975	68747	206118	867

Concentrations are in ppm.

Table 3.3: The averaged concentration values.

Unique Sample ID	Mg Concentration	Al Concentration	Si Concentration	P Concentration
NRCS1	5561	73031	228101	1340.5
NRCS2	4495	94168	226277	1020
NRCS3	8670.5	73960	236217.5	1455.5
NRCS4	<LOD	103197.5	217855	317
NRCS5	20867	68222	205946	859.5

Only the 1st NRCS1 scan was used as the Mg concentration because the 2nd replicate scan of Mg was <LOD. Since both NRCS4 scans were <LOD for Mg, no averaged concentration could be computed and <LOD was returned.

All averaged scans and corresponding lab data were merged based on Unique Sample ID into one at frame which could be used to evaluate and build multiple linear regression and RF models. Importantly, a high proportion of non-detectability for some elements necessitated that they were eliminated prior to modeling (Sharma et al., 2014; Sharma et al., 2015). Thus, Co, Se, Mo, Ag, Cd, Sn, Sb, W, Hg, Bi, Th, and U were excluded from the modeling datasets.

3.5 Instrument quality control

To ensure the pXRF was functioning properly across the duration of the experiment, QC measures were established. The Cal Check was performed and passed at the beginning of each scanning session to make sure the internal components were functioning properly.

A quartz blank was scanned at the beginning of each batch of samples to ensure no contamination was present on the instrument window or sample. To avoid dust or soil from impacting the measurements, a water wipe was used to clean off any dust on the blank and standards. The instrument blank indicated no contamination was present if the pXRF read about 50% Si and 50% LE (oxygen). As shown in Fig. 3.3, LE and Si concentrations mirror each other to reflect the composition of the quartz blank. Most readings reflect LE and Si in a typical range between 490,000 and 510,000 ppm, but for some readings, LE appears to be much higher, while Si is lower. A likely explanation for this is that the water deposited onto the blank by the water wipe was not completely evaporated. As a result, the LE concentration was overestimated which in turn drove down the Si concentration. Even a very small film of water left on the blank can lead to this uptick in LE concentrations (Michael Hull, personal communication, 15 March 2022).

Every 10 sample scans, NIST 2711a and 2706 SRS materials were scanned. These standard reference soils were used to monitor for precision and any possible drift rather than percent recovery. The % recovery method, while useful for many QC plans, is inadequate when it comes to monitoring calibration accuracy and utilization of the reference check sample for pXRF related uses. For elements with small concentrations like Se, fractions may skew the percent recovery statistics and cause large swings. For example, if the instrument reads 1 ppm and then 3 ppm, a denominator of 2 ppm (certified value) would cause the % recovery to swing from 50% under recovery to 50% over recovery.

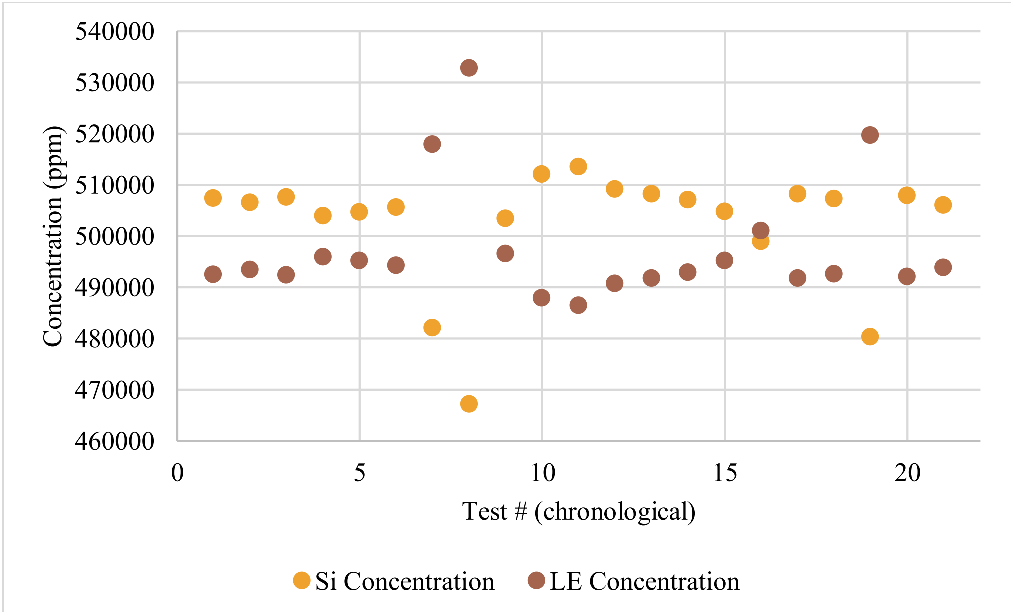

Figure 3.3: Silicon and LE concentrations of quartz blank over time. The concentrations of Si and LE track each other to make up 100% of the blank. LE can be seen to deviate to higher concentrations, likely due to some residual water present on the blank.

The Vanta pXRF is calibrated using a wide range of geochemical, mineral, and soil samples to ensure it performs well over a large concentration range. Looking at a single reading for a certain element and comparing that value to the certified reference value could lead one to believe the instrument is functioning poorly— also called the single point fallacy. An example of this fallacy can be demonstrated with the following scenario. A pXRF reading underreports arsenic for the 2711a standard (Fig. 3.4), so looking at this point alone (one analyte in one standard), might lead the user to believe the instrument's calibration is off. However, adjusting the calibration to be perfect on the arsenic calibration for NIST 2711a, would actually make it worse over the long range (M. Hull, personal communication, 24, June 2021). A look at the larger dataset with several standards reveals that the instrument demonstrates reliable accuracy (R^2 = 0.99) over a range of diverse soil types (Fig 3.5). Thus, while precision can be monitored with a single check sample, accuracy for the instrument can only be evaluated with multiple, multi-element, large range reference samples.

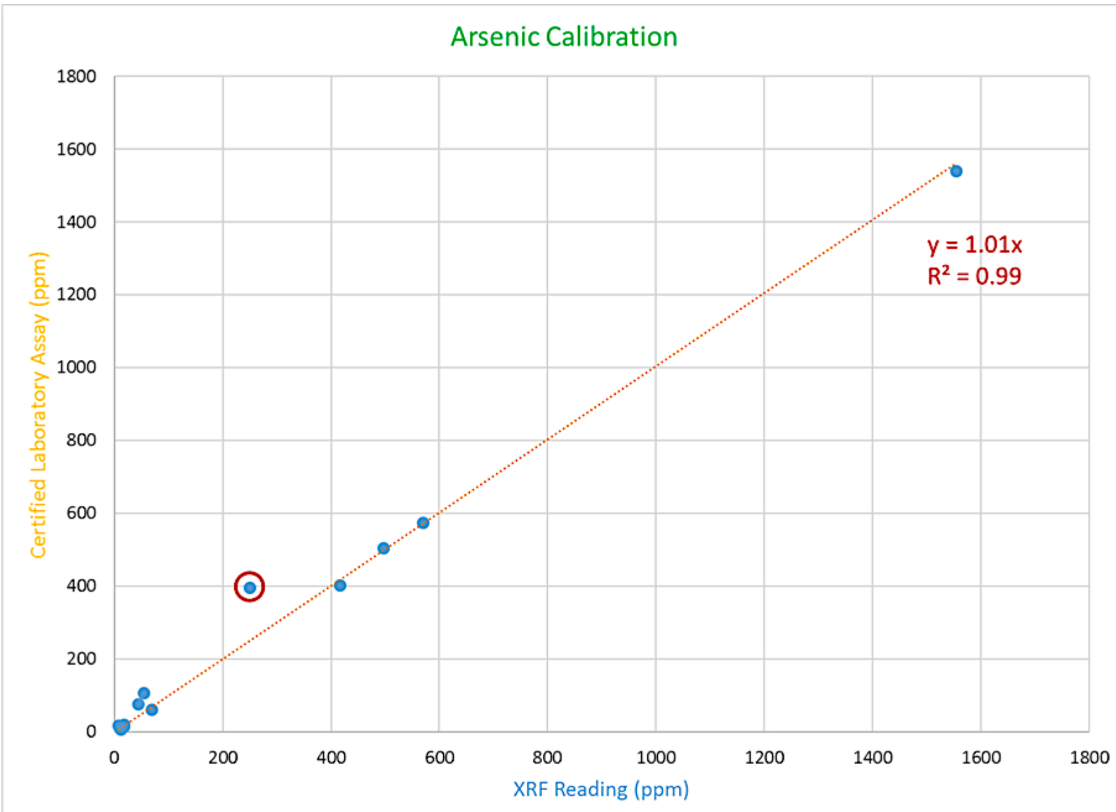

Figure 3.4: Vanta concentration readings for arsenic for a range of test samples. The point circled in red represents the As reading for the NIST 2711a standard and the dotted orange line represents the overall calibration curve (Image from OLYMPUS Scientific Solutions).

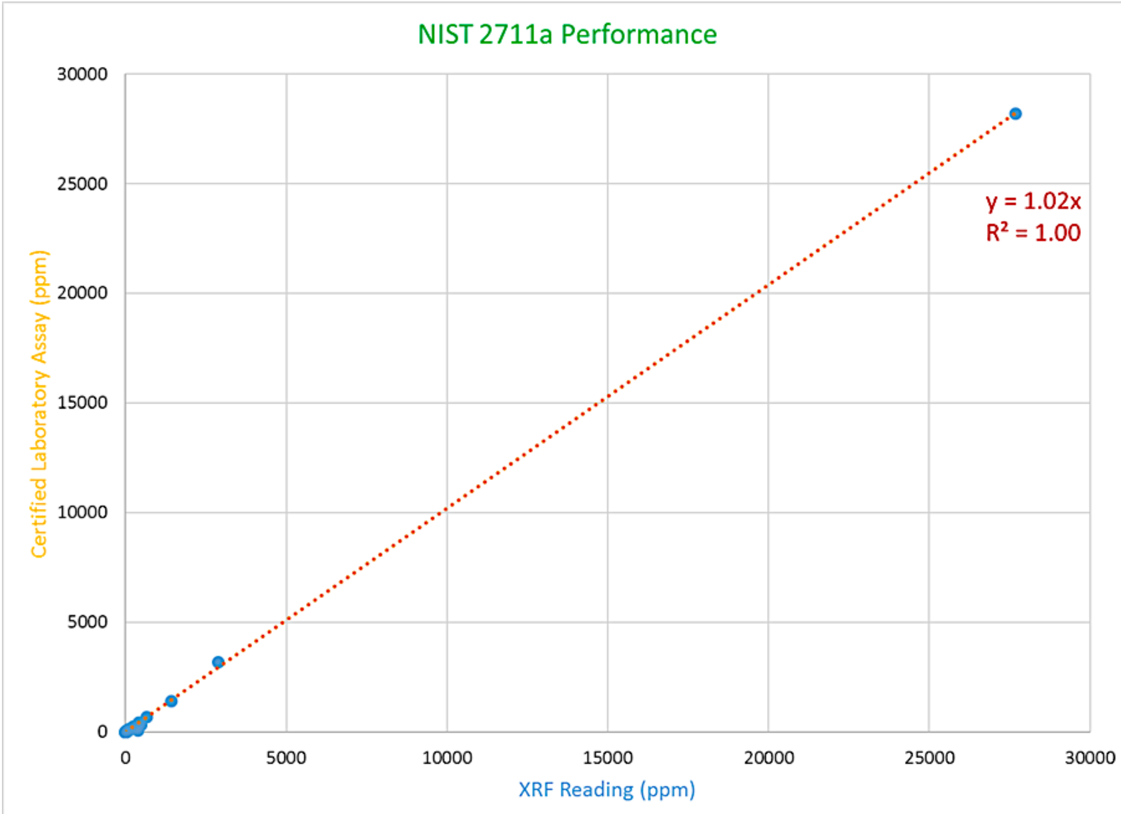

Figure 3.5: Calibration results for 2711a using the Vanta analyzer, showing that the instrument's calibration is reliable overall (Image from OLYMPUS Scientific Solutions).

To monitor the precision of the reference sample readings the measured values for several elements were plotted over several sessions of readings. The exported .csv of scans presented the concentrations for measured elements in ppm as well as the 1 standard deviation error value (sigma; σ**)** of the analyte concentration that was unique to that particular scan. In order to examine the readings of 2711a and 2706 standards over time, the concentrations of a few elements (Pb, Zn, and Ni) were plotted over the duration of the experiment, as shown in Figures 3.6-3.11. Lead, zinc, and nickel were chosen as the elements to graph because they are 'Beam 1' elements, meaning that they are detected in the first of two sequential beams by the pXRF, and are generally more stable than the lighter elements detected via Beam 2. The test numbers are chronological and span from 10/22/20 to 1/31/22 for 2711a, with 109 total readings, and from 7/27/21 to 1/31/22 for

2706, with 53 total readings. Fewer test numbers exist for 2706 because it was acquired part-way through this experiment. To create these figures, the 1σ error provided by the instrument for each measurement was averaged across all readings and then tripled to find the 3σ value. This value was then added to the average concentration to find the upper bound (+3σ) and subtracted from the average concentration to find the lower bound (−3σ).

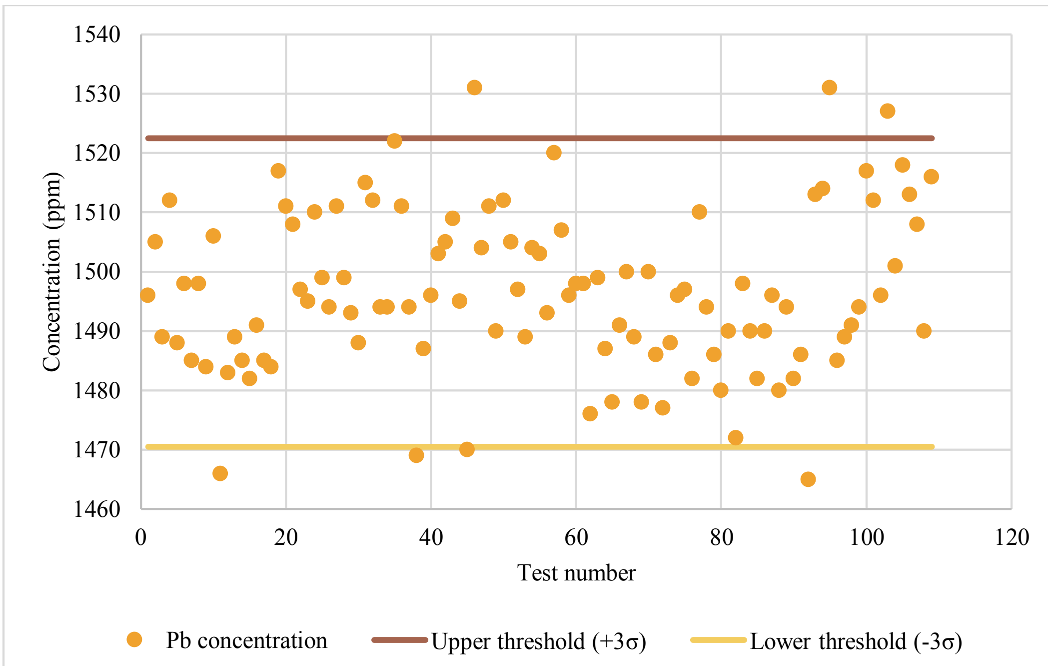

Figure 3.6: 2711a Pb readings over time. 93.6% of 2711a's Pb readings points lie within the average 3 standard deviation bounds of the average Pb concentration.

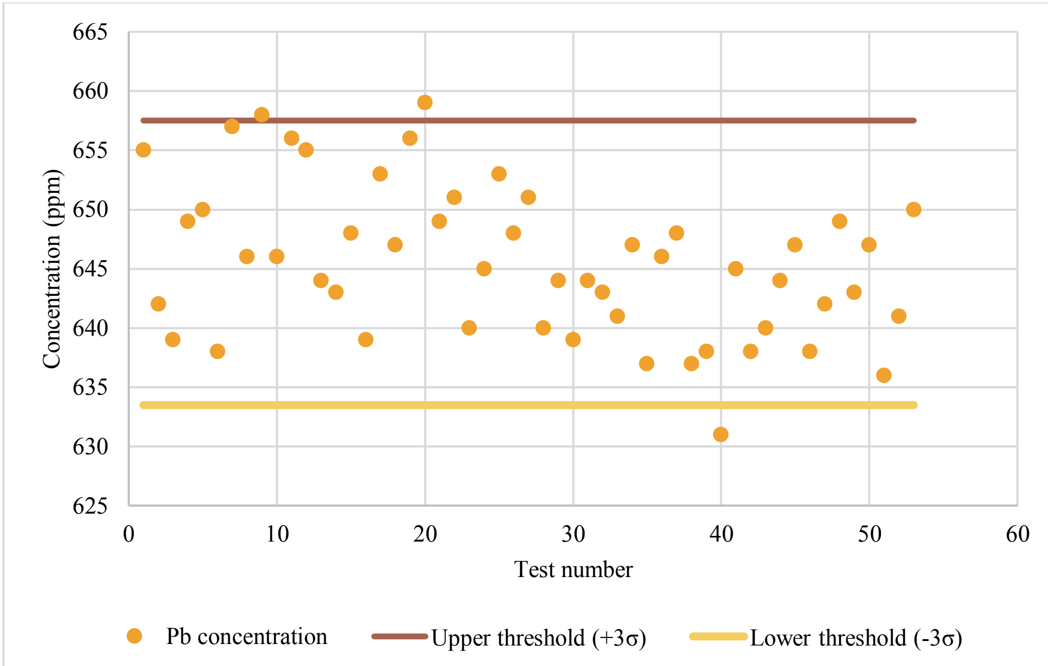

Figure 3.7: 2706 Pb readings over time. 94.3% of 2706's Pb readings lie within the average 3 standard deviation bounds of the average Pb concentration.

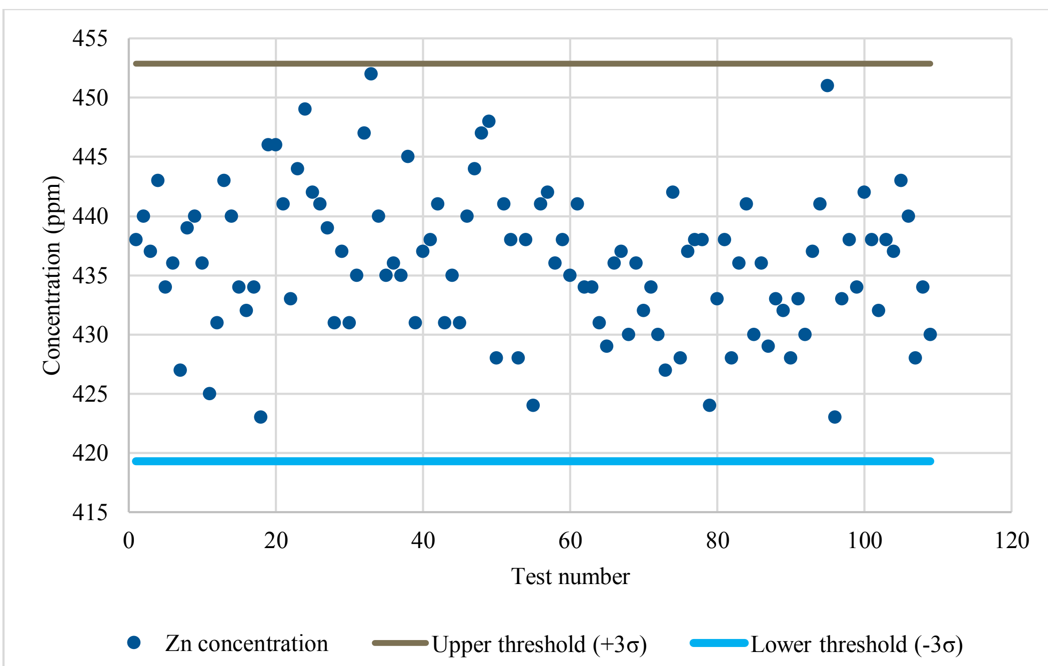

Figure 3.8: 2711a Zn readings over time. 100% of 2711a's Zn readings lie within the average 3 standard deviation bounds of the average Zn concentration.

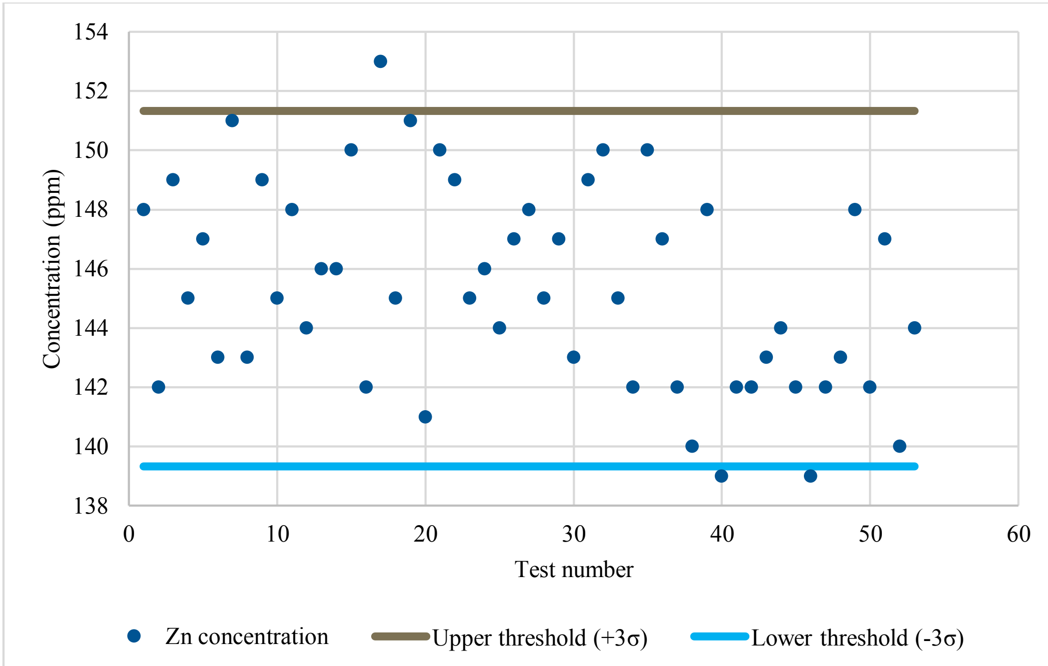

Figure 3.9: 2706 Zn readings over time. 94.3% of 2706's Zn readings lie within the average 3 standard deviation bounds of the average Zn concentration.

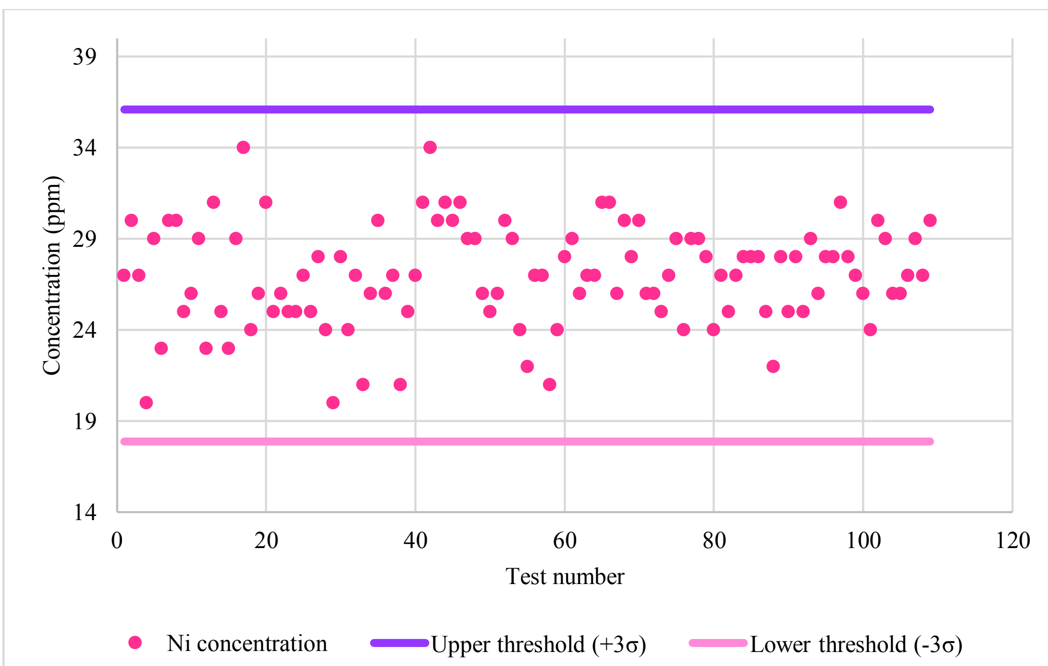

Figure 3.10: 2711a Ni readings over time. 100% of 2711a's Ni readings lie within the average 3 standard deviation bounds of the average Ni concentration.

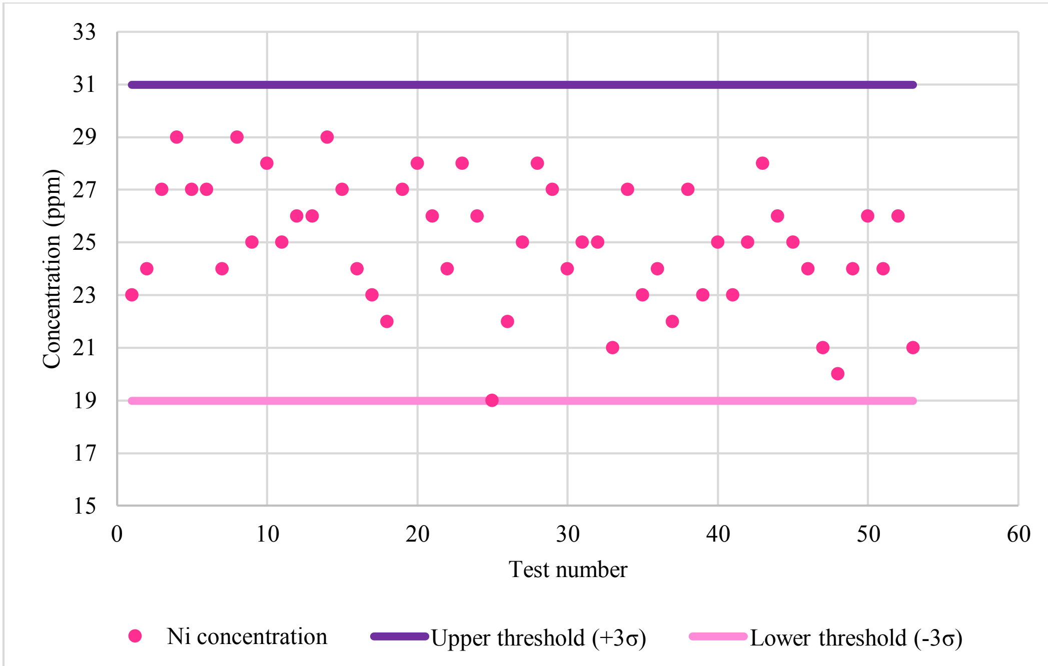

Figure 3.11: 2706 Ni readings over time. 100% of 2706's Ni readings lie within the average 3 standard deviation bounds of the average Ni concentration.

The distance covered by the three standard deviations above and below the average concentration should include 99.7% of the measured values, which is demonstrated by figures 3.8, 3.10 and 3.11, but not by figures 3.6, 3.7, and 3.9. A possible explanation for points outside these bounds could be due to the fact that since each measurement has an associated 1 sigma error which depends on the particular test, averaging across over a year of measurements may have decorrelated the errors from their associated measurements. However, most importantly, none of the graphs show drift in any one direction, which would indicate a problem with the instrument. From this data, we are able to trust that the instrument calibration is reliable overall, and the instrument is functioning properly.

Additional QC monitoring to assess method precision was performed by inspecting the relative standard deviation (RSD) of elemental concentrations for both standards over the duration of the experiment. According to Method 6200 (USEPA, 2007), for measurement values to be considered adequately precise, the RSD should be <20%, with the exception of chromium which should be <30%. The RSD were calculated using Eq. 3.13 and are shown in Table 3.4. All analytes fall within the acceptable RSD range except for P (20.68%) and Sn (20.87%) for 2706, and Sb (22.02%) and Th (25.80%) for 2711a.

Equation 3.13

$$RSD = (\text{Standard deviation}/\text{Average concentration}) \times 100$$

Table 3.4: The relative standard deviations of each analyte.

	Relative Standard Deviation (%)	
Element	**2706**	**2711a**
Mg	-	10.07
Al	2.44	1.77
Si	1.81	1.57
P	20.68	6.15
Si	2.27	2.63
K	1.75	0.74
Ca	4.65	1.09
Ti	2.53	3.09
V	14.73	10.01
Cr	18.43	18.37
Mn	3.54	3.06
Fe	0.99	0.99
Ni	9.57	10.29
Cu	2.89	2.93
Zn	2.35	1.39
As	-	9.89
Rb	1.70	1.22
Sr	1.32	0.91
Y	4.53	3.43
Zr	3.62	2.38
Nb	9.12	6.19
Mo	17.99	-
Cd	-	10.05
Sn	20.87	-
Sb	5.07	22.02
Hg	-	15.74
Pb	0.99	0.91
Bi	-	-
Th	12.67	25.80

Dashed cells represent no certified value exists for that element on the Certificate of Analysis sheet or it was detected less than 10% of the time by the analyzer.

3.6 Data Preprocessing

The dataset was prepared for modeling by first removing some elements which had a high percentage of <LOD readings (Co, Se, Mo, Ag, Cd, Sn, Sb, W, Hg, Bi, Th, U). This left 19 elements subject to regression (Mg, Al, Si, P, S, K, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, As, Rb, Sr, Y, Zr, Nb, Pb). Within these 19 elements, missing (<LOD)

elemental concentrations for Mg (90 values), P (69 values), S (23 values), Ca (5 values), Cr (9 values), As (16 values), and Nb (69 values) were imputed. Imputation was conducted based on logical rules— the detection threshold for the scans that was registered by the analyzer provided a basis for the range of values from which the missing value could be in. To fill in missing values for each element, the associated 1σ error for the <LOD readings were multiplied by three to find the average limit of detection for the missing values of that analyte, since the LOD is typically represented as the 3σ error (Rousseau, 2001). The average limit of detection and average standard deviation of the 1σ errors was used to impute the missing values along a normal distribution curve (Dennis Sun, personal communication, 20 Jan. 2022; Nichols, 2018). Because the missing data mechanism was known, the strategy of imputation described above did not rely on particularly strong assumptions and missing data could be accounted for, without the need to throw away many observations which can lead to biased estimates (Gelman and Hill, 2007). The imputed concentrations can be found in Appendix D. Data processing, analysis, model building, and graphing was accomplished using R Studio (Version 1.3.1093 PBC, 2009-2020).

3.7 Testing existing models

Selected existing research using pXRF analysis to predict soil properties of interest was outlined in Chapter 2 of this report. The models themselves (in the case that our dataset contained the same variables), the indicated variables with generated coefficients, or the author's modeling processes were mimicked using our dataset for each property to see if these models or modeling approaches could produce reasonable estimates. Because regression model equations were being applied for predicative purposes and not inference purposes, it was not necessary to check that typical linear regression assumptions were met. Multiple linear regression models were evaluated for pH, texture, and CEC, but no linear models existed for SOC, TN, and C:N ratio. Thus, for these three properties, RF modeling techniques set forth by Towett et al. (2015) were mimicked. Model performance was assessed using R^2, RMSE, RPD, and RPIQ. A number of metrics for measuring model performance were considered in order to strengthen model comparisons and interpretations.

3.7.1 pH

To evaluate and formulate models suitable for predicting soil pH, laboratory determined values of soil reaction (pH) measured in deionized water were used as the target value with averaged elemental concentrations as the predictors. First, Eq. 2.2 developed by Sharma et al. (2014) and discussed in Chapter 2 was applied to the entire dataset and evaluated for its performance. Of the initial 480 samples, 3 samples were excluded from analysis due to missing pH values, leaving n=478. For Eq 2.3 derived from author's dataset B, scanning was operated in Soil Mode, which detects elements not detected in Geochem mode. For this reason, Eq. 2.3 could not be validated with our dataset. After poor performance from applying Eq. 2.2 as is, a new regression model was built using the same variables found by Sharma et al. (2014), but with coefficients generated specifically for our dataset. The dataset of 478 observations was split into 80% training and 20% testing/validation sub-datasets. Using the lm() function, pH was set as the x variable and the log values of Al, Si, Mn, Fe, K, Ca, and Zn concentrations were used as predictors for the training dataset. The resulting Eq. 4.1 when applied to the test dataset, improved model performance, but was still an inadequate predictive model.

3.7.2 Soil texture

The model developed by Zhu et al., (2011) and summarized in Chapter 2, could not be validated with the California soils dataset, due to the absence of Co and Ba concentrations from our scans (as a result of the different modes of operation). In lieu of applying the variables and coefficients used by authors, the same methodology was applied for deriving correlations between soil texture and the studied elements. To achieve this, observations without lab verified texture data were eliminated, leaving n= 358. The modeling dataset was then split into 2/3 modeling and 1/3 validation sub-datasets. Backward stepwise multiple linear regression was conducted on the training dataset using the stepwise () function and specifying an entry significance of 0.5, exit significance of 0.1, and 15 maximum steps. AIC was indicated as the selection criterion for keeping elements in the model. The sand and clay percentages were individually set as the dependent variables, with all elemental concentrations listed as predictors. To find silt percentage of the entire dataset, clay and sand contents were subtracted from 100. Zhu et al., (2011) did not logarithmically transform elemental concentrations, so the concentrations were left in their original form for this regression analysis.

3.7.3 CEC

The model equation (Eq. 2.4) created by Sharma et al., (2014) to predict CEC from pXRF analysis was applied to our entire dataset to evaluate its performance. After poor performance from applying Eq. 2.4 as is, a new equation was built using the same elemental variables found by Sharma et al. (2014), but with coefficients generated specifically for our dataset. Our dataset was split into 80% training and 20% testing/validation sub-datasets. Using the lm() function, CEC was set as the x variable and the concentrations of Ca, Ti, V, Cr, Fe, Cu, Sr, and Zr were used as predictors for the training dataset. The resulting Eq. 4.5 when applied to the test dataset, improved model metrics slightly, but was still an inadequate predictive model.

3.7.4 Soil organic carbon, total nitrogen, C:N ratio

The random forest modeling process followed by Towett et al., (2015) and summarized in Chapter 2 was applied to the SOC data. It should be noted that while this study used conventional benchtop XRF and not pXRF, they harness the same technology, and it has been established that with proper preprocessing of samples and QC protocols the two methods have been shown to correlate very well (Goff et al., 2020; Laperche and Lemière 2021). Using the randomForest library in R, regression computations were performed and validated using out-of-bag (OOB) validation. To corroborate the mean square error (MSE) calculated on a 1/3 OOB validation set, the MSE was compared to a 50% random hold out sample. A similar MSE from both methods substantiated the OOB process, and an RF model was developed using the entire sample set. The authors building criteria of number of trees built (ntree = 200) was specified but the criteria of number of variables tried at each split (mtry=50) could not be replicated because the CA modeling dataset only had 22 variables/elements. Therefore, mtry was set to equal 22.

3.8 Multiple linear regression model building

Multiple linear regression models were also constructed from scratch using the 'tidyverse package' (Wickham et al., 2019). For modeling using this method, all elemental concentrations were log transformed to fix right skewed distributions (present in 18/22 elements) and improve predictions. Instead of dividing the dataset into a single train/test set to evaluate model performance, 10-fold cross validation on a training set

(75% subset) was used as well as an unseen test 'hold-out' set (25% subset). This resampling approach trained a model using a portion of the data from each fold as training data and measured the accuracy of the developed model on the remaining part of the data. Variables were selected to be in the final model based on their significance ($p < 0.05$) following the 10-fold cross-validation step. Then, the refined model was applied to the unseen test set, where model metrics (R^2, RMSE, RPD, and RPIQ) were determined. Coefficient significance was then determined using the final model on the entire dataset once at the end. The results of the MLR model that performed the best on the test set was displayed graphically, with the 1:1 line as a black line and the line of best fit as a blue line.

The main intention of MLR modeling in this study was for prediction, which does not require regression assumptions to be met. However, these assumptions must be met for inference purposes, such as interpreting variable coefficients and reporting their significance. Thus, for all properties modeled, four residual plots (residuals vs fitted, normal Q-Q, scale-location, and residuals vs leverage) were created to check the assumptions of linearity, homoscedasticity, normality of residuals, and to indicate any influential points. Residual values of the log transformed elemental concentrations used to create MLR models were regressed against raw values for each of the seven soil properties investigated. These plots can be found in Appendix C.

3.9 Algorithmic modeling using random forest

Random forest models were also created in an attempt to further improve predictions. Using tidymodels in RStudio, raw elemental concentrations were used as predictors and soil properties of interest were used as the target variable. 10-fold cross-validation on the training set (75%) was used to determine hyperparameters for the random forest model. Tuning indicated an optimal mtry (# of predictors randomly sampled at each split) of 5, min_n (minimum number of data points required to further split the node) of 2, and trees (number of trees in the forest) as 100. In effect, this created many 'deep' trees fit on fewer variables. The RF model was then applied to the test set, from which model performance metrics were derived.

3.10 Grouping predictive models by land type and characterization method

To see if the addition of the categorical variables land type and sample set would improve predictions for each property, MLR models built using only elemental data were compared to MLR models using elemental data as well as categorical data. The models were built using the entire dataset, with 10-fold CV to pick significant coefficients and model results (R^2 and RMSE) were reported on the unseen 'folds' (no unseen test set was employed for this application). Categorical data was transformed into nominal data by using the step_dummy() function which mapped the land type and sample set to a sequence of 0/1 indicator variables, so they could be used as regression predictors. A categorical variable with n levels would require n-1 dummy variables to represent the categories (an n^{th} dummy variable would be redundant and carry no new information). For instance, to discover how helpful land type and sample set were for predicting pH, a model with pH set as the dependent variable and all elements as predictors was compared to a model with pH as the dependent variable and all elemental predictors as well as 9 columns filled with either a 0 or 1 to represent the 6 land types and 5 sample sets.

In an attempt to further refine models, additional MLR models were constructed within land type categories and laboratory methodologies. The aforementioned methods of 10-fold CV on a 75% train set to pick significant variables and model testing/metric reporting on the 25% test dataset were carried out. Land type categories distinguishable from the entire modeling dataset and large enough for modeling were forest (n=166), grassland (n= 81) and marine terrace (n = 159) environments. For all other land types, regression was unreasonable due to the small sample sizes (n ≤ 39). Separate models were also created by grouping samples based on the methodology used to determine each property. For instance, pH was determined using 1:1 DI to soil method for the LA Urban, SPR/LHBC Mollisols, NRCS Chico, and UC Merced soils, so these samples were grouped together. By contrast, the pH of marine terrace samples was determined using a saturated paste method, so these samples comprised their own group. In some instances, all samples in a certain land type were also characterized using a separate method, so making another model for these same samples twice would be redundant. For example, the marine terrace soils were the only samples characterized for CEC using method S - 10.10 (Soil, Plant, and Water Reference Methods for the Western Region), so making another model for this method would be redundant. For SOC, N, and C:N ratio, dry combustion was the only method used, so samples could not be divided based on methodology for these properties.

Chapter 4

RESULTS

4.1 pH

4.1.1 Descriptive statistics

Summary statistics for pH values can be found in Table 4.1. A boxplot for each sample set (Fig. 4.1) and for all samples in the dataset (Fig. 4.2) show the spread of pH values. The Tukey outlier test revealed a bottom threshold of 3.160 and upper threshold of 9.355, indicating 1 outlier: LA Urban Plot 2 (pH: 10.38).

Table 4.1: pH summary statistics.

	Sample set					
	SPR/LHBC	NRCS Chico	LA Urban	UC Merced	Marine terrace	Total dataset
Samples	218	58	39	4	159	478
Minimum	4.8	4.70	3.28	5.18	4.50	3.28
Median	6.31	6.40	6.39	5.56	5.80	6.20
Mean	6.37	6.52	6.48	5.60	6.00	6.29
Maximum	8.02	8.80	10.38	6.09	7.40	10.38

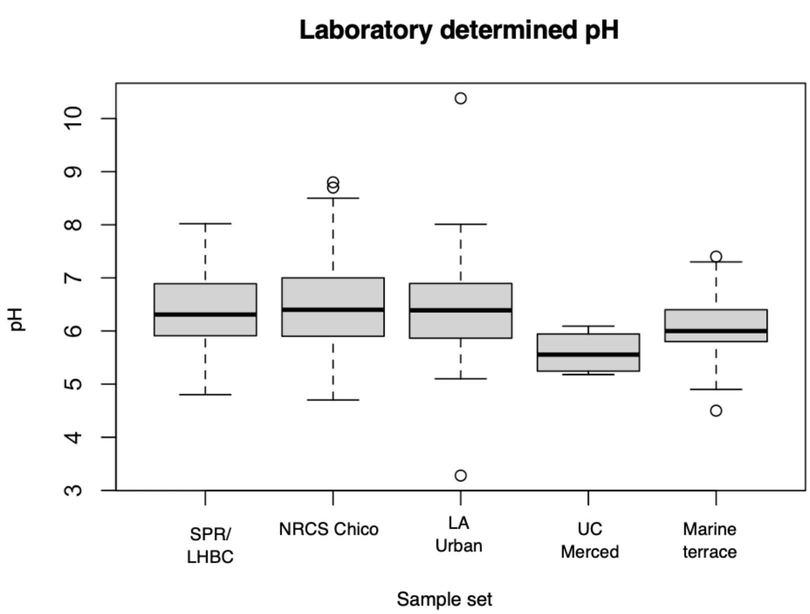

Figure 4.1: pH by sample set.

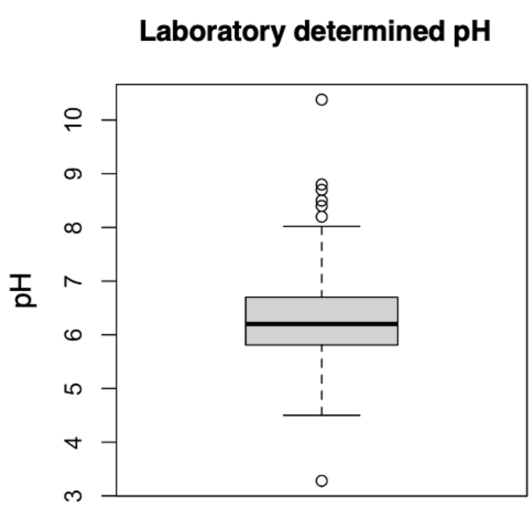

Figure 4.2: pH values for entire dataset.

4.1.2 Data models

Applying Eq. 2.2 to our entire dataset exhibited poor predictive performance, producing an $R^2 = 0.0109$, RMSE = 1.21, RPD = 0.5918825, and RPIQ = 0.732941. Using the same variables as Eq. 2.2 but with coefficients derived from our own dataset Eq. 4.1 was created. When applied to the validation sub-dataset, Eq. 4.1 produced an $R^2 = 0.153$, RMSE = 0.666, RPD = 1.045788, and RPIQ = 1.324972.

Using the tidyverse approach to build a linear regression model, the results of 10-fold CV on a 75% training set showed P, S, Ca, V, Fe, Ni, Zn, Sr, and Y to be significant so they were chosen for the final model (Eq. 4.2). When applied to the test set, this model produced a further improved $R^2 = 0.532$, RMSE = 0.489, RPD = 1.455922, and RPIQ = 1.732563 (Fig. 4.3).

Equation 4.1

$$pH = -19.8528 + 1.4055 * \log(Al) + 2.4755 * \log(Si) + 0.1693 * \log(Mn) + 1.0162 * \log(Fe) - 0.5235 * \log(K) + 0.6103 * \log(Ca) + 0.5010 * \log(Zn)$$

Equation 4.2

$$pH = 5.4066 - 0.4374 * \log(P) - 0.4995 * \log(S) + 1.0886 * \log(Ca) + 1.1805 * \log(V) - 0.6077 * \log(Fe) + 0.6122 * \log(Ni) + 1.0926 * \log(Zn) - 1.0331 * \log(Sr) - 0.8818 * \log(Y)$$

Table 4.2: Model parameters for pH regression models.

Variable (Logged)	**Eq. 2.2 Sharma et al. (2014)**	**Eq. 4.1 Generated coefficients**	**Eq. 4.2 10-fold CV method**
Constant	9.7164	-19.8528^{***}	5.4066^{***}
Al	−5.9247	1.4055^{**}	
Si	1.8491	2.4755^{***}	
Mn	−2.0419	0.1693	
Fe	1.9212	1.0162^{**}	-0.6077^{*}
K	2.3906	−0.5235	
Ca	0.4396	0.6103^{***}	1.0886^{***}
Zn	0.6689	0.5010^{*}	1.0926^{***}
P			-0.4374^{***}
Ni			0.6122^{***}
S			-0.4995^{***}
Sr			-1.0331^{***}
Y			-0.8818^{***}
V			1.1805^{**}
Sample #	478	478	478
R^2	0.0109	0.153	0.532
RMSE	1.21	0.666	0.489
RPD	0.5918825	1.045788	1.455922
RPIQ	0.732941	1.324972	1.732563

The model performance metrics were calculated using the entire dataset for Eq. 2.1, the 20% test sub-dataset for Eq. 4.1, and the 25% test subset for Eq. 4.2. Significance codes (p-values): '***' 0.001 '**'
0.01 '*' 0.05

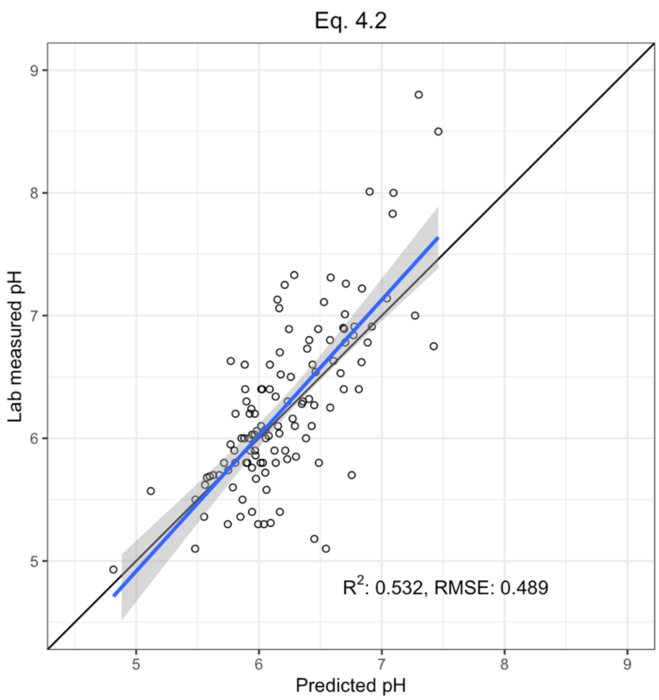

Figure 4.3: Eq. 4.2 for pH applied to the holdout set.

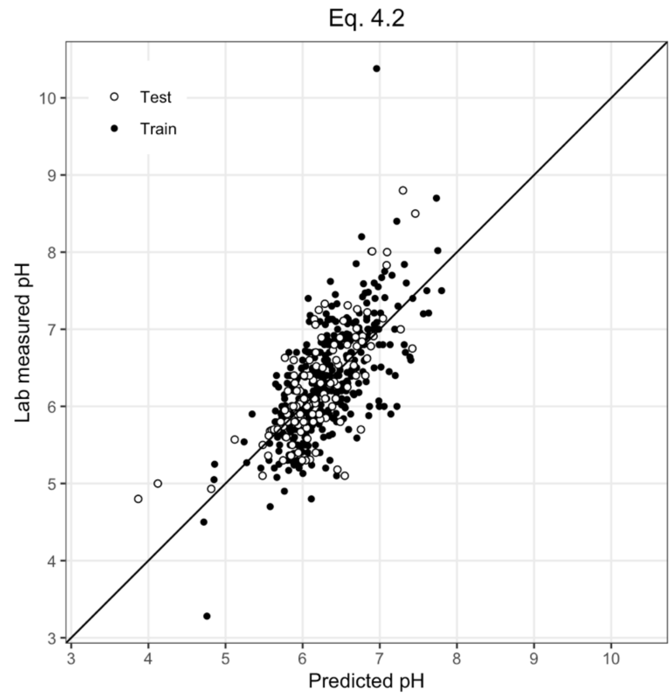

Figure 4.4: Eq. 4.2 for pH applied to the train and test set.

4.1.3 Algorithmic modeling

Implementing a random forest model for predicating pH revealed an $R^2 = 0.485$, RMSE $= 0.490$, RPD $= 1.377041$, and RPIQ $= 1.547162$ when applied to the test sub-dataset (Fig 4.5).

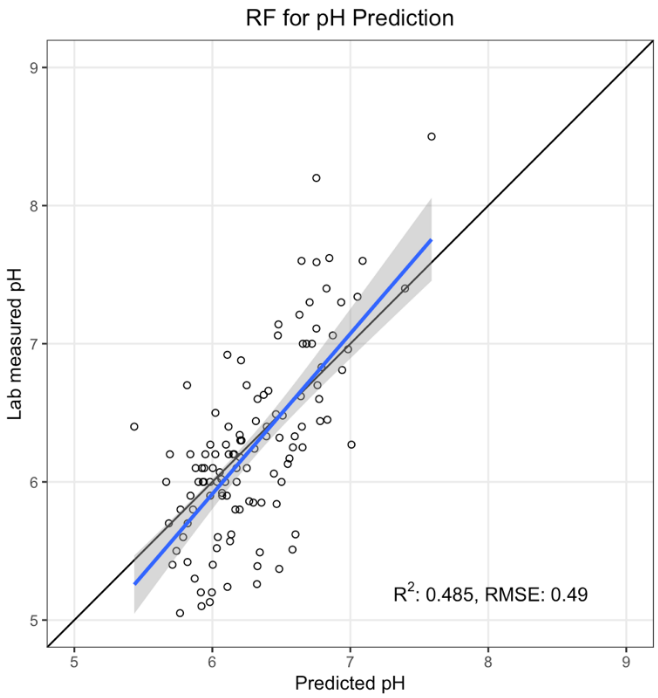

Figure 4.5: RF modeling for pH on the holdout set.

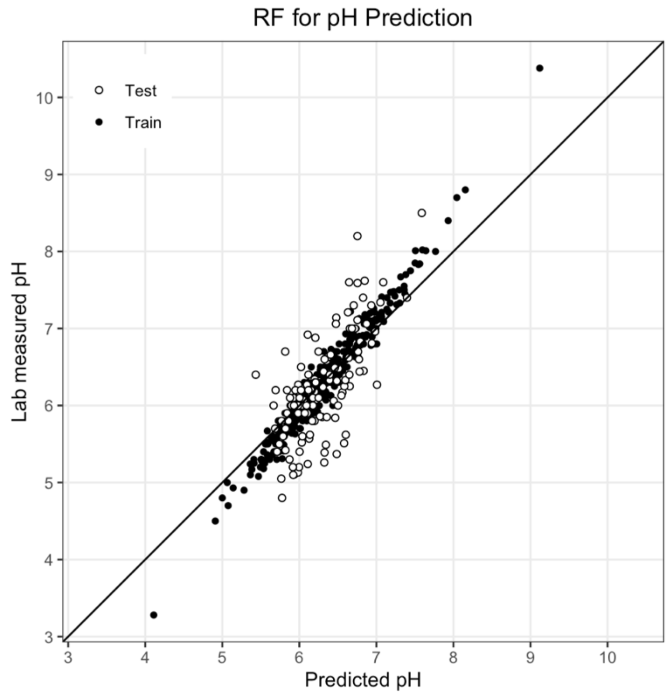

Figure 4.6: RF modeling for pH on the train and test set.

4.2 Texture

4.2.1 Descriptive statistics

Soil textures for the samples used this study were plotted on a soil texture triangle using the 'soiltexture' package (v1.5.1, Moeys, 2018) (Fig. 4.7) and are shown in Table 4.3.

Table 4.3: Texture classes of 358 characterized samples.

Texture class	S	LS	SL	SCL	SC	L	CL	C	SiC	SiCL	SiL	Si
Sample #	10	42	107	31	2	83	58	6	2	11	6	0

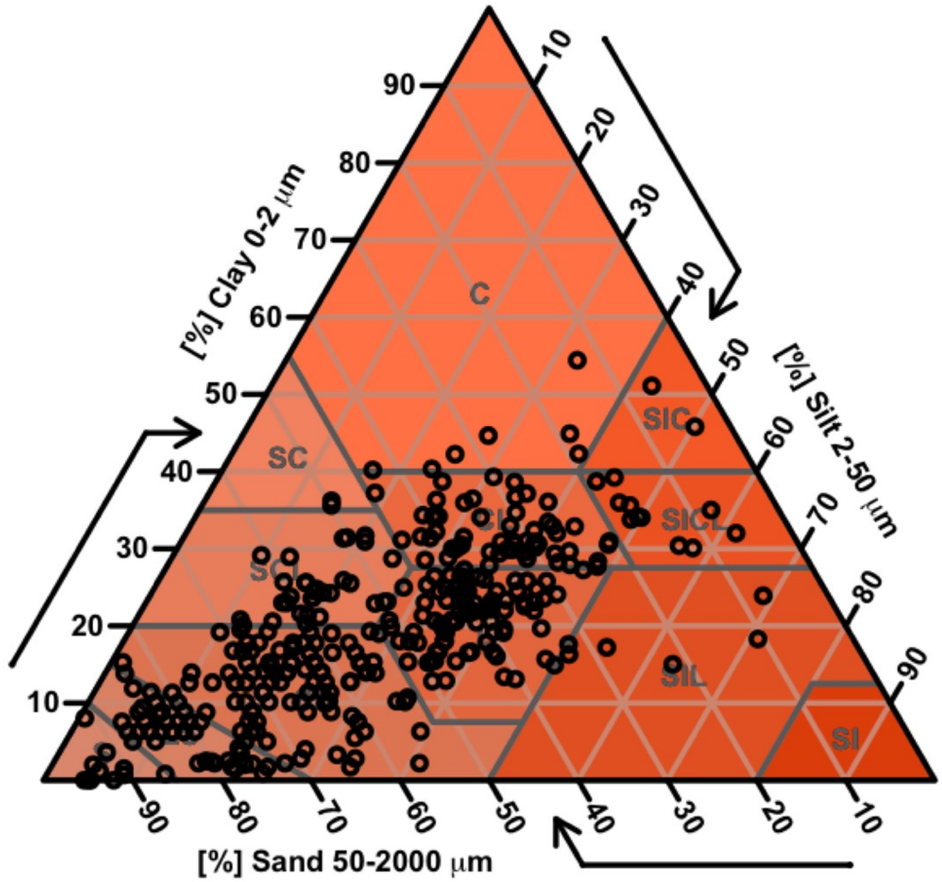

Figure 4.7: The texture classifications for 358 samples plotted on USDA-NRCS texture triangle.

4.2.2 Data models

Following the methodology of creating backward stepwise regression models from untransformed elemental concentrations set forth by Zhu et al. (2011), Eq. 4.3 and 4.4 were developed for sand and clay percentage. Eq. 4.3 produced an $R^2 = 0.599$, RMSE = 14.2, RPD = 1.582784, and RPIQ = 2.359735 while Eq. 4.4 produced an $R^2 = 0.575$, RMSE = 7.23, RPD = 1.534052, and RPIQ = 2.40448 for the 33% validation sub-dataset. Predicted silt percentages were found by subtracting the predicted sand and clay content from 100%. Applying these two models to the total dataset of 358 samples resulted in 3 samples with a negative value for a textural separate, so these observations were removed prior to determining predicted texture class. For the remaining 355 samples with predicted and actual texture classes, the correct texture class was predicted 54% of the time. Predicted textures obtained via Eq. 4.3 showed no strong tendency to over or underestimate sand contents (51% vs 49% of the time) while Eq. 4.4 exhibited a marginal tendency to overestimate clay contents (55% of the time).

When using the tidyverse approach to build a linear regression model, the results of 10-fold CV on a 75% training set for predicting sand contents showed 12 variables (Mg, Al, P, S, K, Fe, Ni, Cu, As, Rb, Sr, Zr) to be significant so they were used for the final model (Eq. 4.5). When applied to the test set, this model produced a further improved R^2 = 0.616, RMSE = 12.3, RPD = 1.600894, and RPIQ = 2.589406 (Fig. 4.8). Applying this approach for predicting clay contents indicated 13 significant variables used for the final model (Eq. 4.6). Applied to the test set, model metrics were as follows: R^2 = 0.599, RMSE = 6.83, RPD = 1.586135, and RPIQ = 2.294147 (Fig. 4.10). When these two models were applied to the entire dataset, 8 values were negative and had to be excluded prior to texture class determination. The remaining 350 samples with a predicted texture class matched the actual texture class 55% of the time. Sand and clay contents obtained via Eq. 4.5 and 4.6 showed a slight tendency to overestimate these values (53/54% of the time). The coefficients found by Zhu et al. (2011) for Louisiana and New Mexico soils is compared with Eq. 4.3, 4.4, 4.5, and 4.6 in Table 4.4. As shown in Table 4.5, while Zhu et al. (2011) found that more Fe implied a higher sand and clay content and more Rb implied more clay and less sand, models produced following their process showed the same pattern only for clay. Eq. 4.3 and 4.5 contained a negative coefficient for Fe, which made its weight negative.

Equation 4.3

$$\text{Sand \%} = 63.46884 - 0.09080045 * (Zr) - 0.1518520 * (Ni) + 0.0007474723 * (Al) - 0.001509740 * (Fe) + 0.02812715 * (Sr) + 0.001212738 * (Mg) - 0.3078903 * (Rb) + 0.001328229 * (P) + 1.021068 * (Nb) + 0.03346892 * (Cr) - 0.00007750474 * (Si) + 0.001105562 * (K)$$

Equation 4.4

$$\text{Clay \%} = 15.83067 - 0.001951922 * (K) + 0.3827918 * (Rb) + 0.00003171078 * (Si) - 0.0002763703 * (Al) + 0.0006723170 * (Fe) - 0.0001804027 * (Ca) - 0.6357696 * (Nb) - 0.00045686710 * (Mg) + 0.06044283 * (V) - 0.002901094 * (Mn)$$

Equation 4.5

$$\text{Sand \%} = -166.194 + 18.044 * \log(Mg) + 76.115 * \log(Al) + 9.024 * \log(P) - 4.988 * \log(S) + 24.110 * \log(K) - 40.852 * \log(Fe) - 18.098 * \log(Ni) - 12.396 * \log(Cu) - 14.755 * \log(As) - 54.038 * \log(Rb) + 31.098 * \log(Sr) - 29.866 * \log(Zr)$$

Equation 4.6

$$\text{Clay \%} = -81.893 - 3.986 * \log(Mg) - 32.472 * \log(Al) + 32.186 * \log(Si) - 4.666 * \log(P) + 6.319 * \log(S) - 32.344 * \log(K) - 5.189 * \log(Ca) - 3.395 * \log(Mn) + 39.162 * \log(Fe) + 8.437 * \log(Ni) - 5.446 * \log(Cu) + 7.350 * \log(As) + 40.218 * \log(Rb)$$

Table 4.4: Model parameters for sand and clay % regression models.

Variable	**Eq. 4.3: sand Zhu et al. methods**	**Eq. 4.4: clay Zhu et al. methods**	Variable (Logged)	**Eq. 4.5: sand 10-fold CV method**	**Eq. 4.6: clay 10-fold CV method**
Constant	63.47^{***}	15.83^{*}	Constant	-166.19^{*}	−81.89
Al	$7.47 \times 10^{-4***}$	$-2.76 \times 10^{-4***}$	Al	76.12^{***}	-32.47^{***}
Fe	$-1.51 \times 10^{-3***}$	$6.72 \times 10^{-4***}$	Fe	-40.85^{**}	39.16^{***}
Sr	$2.81 \times 10^{-2*}$		Sr	-31.10^{***}	
Rb	-0.31^{**}	0.383^{***}	Rb	-54.04^{***}	40.22^{***}
K	$1.11 \times 10^{-3*}$	-1.95×10^{-3} ***	K	24.11^{*}	-32.34^{***}
Si	$-7.75 \times 10^{-5*}$	3.17×10^{-5}	Si		32.19^{**}
Mn		-2.90×10^{-3}	Mn		−3.40
Cu			Cu	-12.40^{*}	-5.45^{*}
Zn			Zn		
As			As	-14.76^{**}	7.35^{*}
Ni	-0.15^{**}		Ni	-18.10^{**}	8.44^{**}
Ca		$-1.80 \times 10^{-4*}$	Ca		-5.19^{**}
V		$6.04 \times 10^{-2*}$	V		
Zr	$-9.08 \times 10^{-2***}$		Zr	-29.87^{***}	
Mg	$1.21 \times 10^{-3***}$	$-4.57 \times 10^{-4**}$	Mg	18.04^{***}	−3.986
P	1.33×10^{-3}		P	9.02^{**}	-4.67^{**}
Cr	$3.35 \times 10^{-2*}$		Cr		
Nb	1.02^{*}	-0.64^{**}	Nb		
S			S	−4.99	6.32^{**}
Sample #	358	358	Sample #	358	358
R^2	0.599	0.575	R^2	0.616	0.599
RMSE	14.2	7.23	RMSE	12.3	6.83
RPD	1.583	1.534	RPD	1.601	1.586
RPIQ	2.360	2.404	RPIQ	2.589	2.294

Model evaluation metrics were calculated using the 1/3 test set for Eq. 4.3 and 4.4 and using the 1/4 test set for Eq. 4.5 and 4.6. Significance codes (p-values): '***' 0.001 '**' 0.01 '*' 0.05

Table 4.5: Weights for Fe/Rb coefficients found by Zhu et al. (2011) and developed equations.

	Weights for selected coefficients							
	Louisiana sand	Capulin sand	Eq. 4.3 sand	Eq. 4.5 sand	Louisiana clay	Capulin clay	Eq. 4.4 clay	Eq. 4.6 clay
Fe	18.1	11.9	−44.8	-182.7	29.6	11.5	19.9	175.1
Rb	−38.3	−24.1	−21.0	-99.1	29.7	13.1	26.1	73.7

Weights were calculated as a function of the variable coefficient and the average concentration of that element.

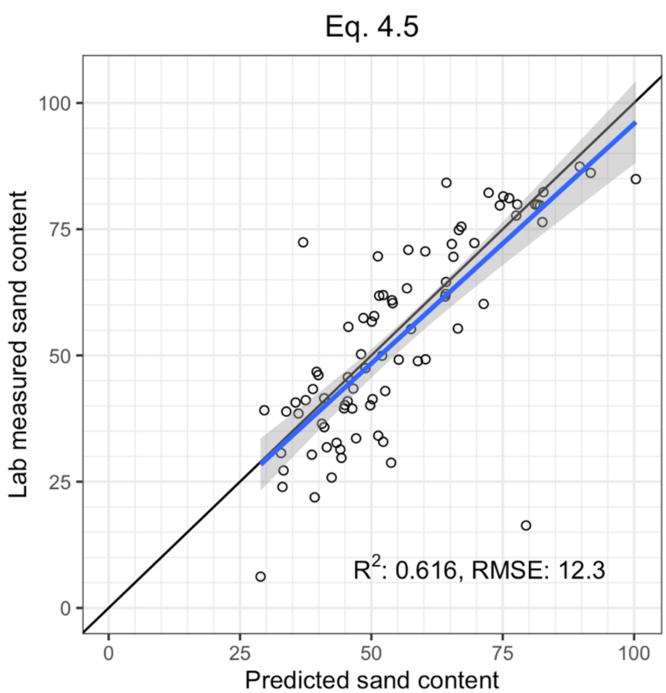

Figure 4.8: Eq. 4.5 for sand % applied to the holdout set.

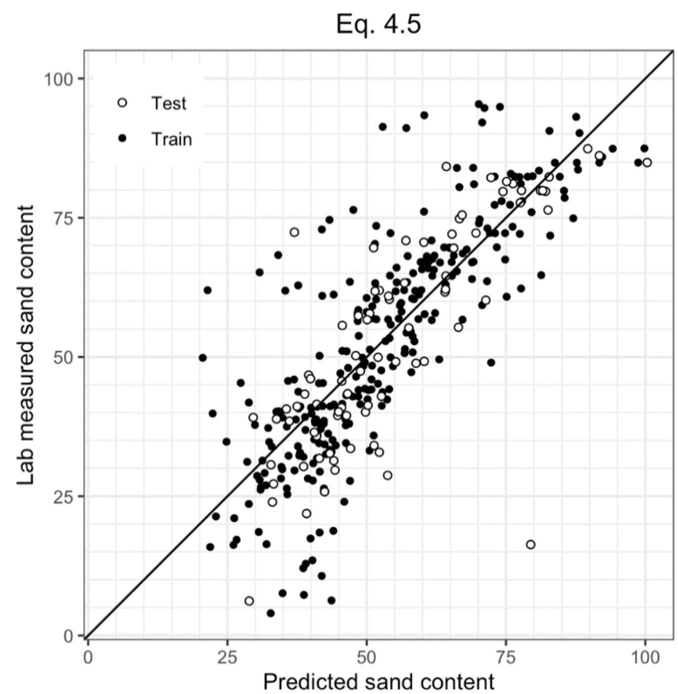

Figure 4.9: Eq. 4.5 for sand % applied to the train and test set.

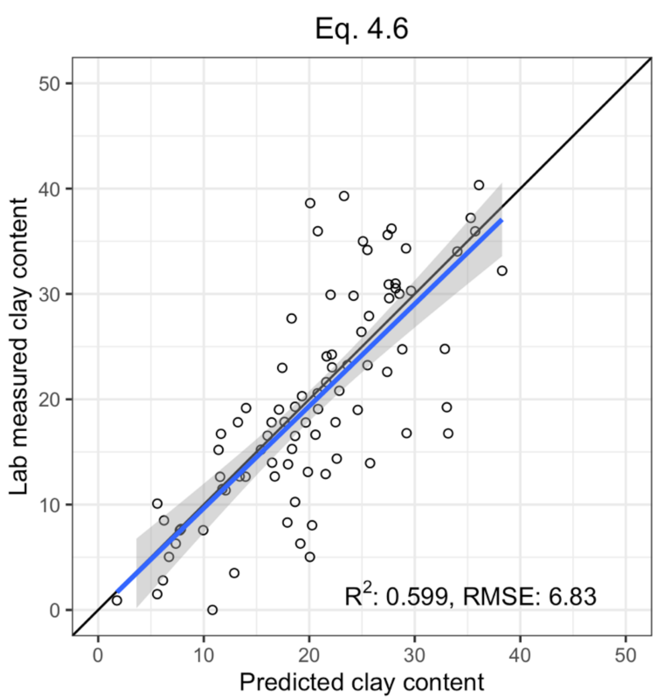

Figure 4.10: Eq. 4.6 for clay % applied to the holdout set.

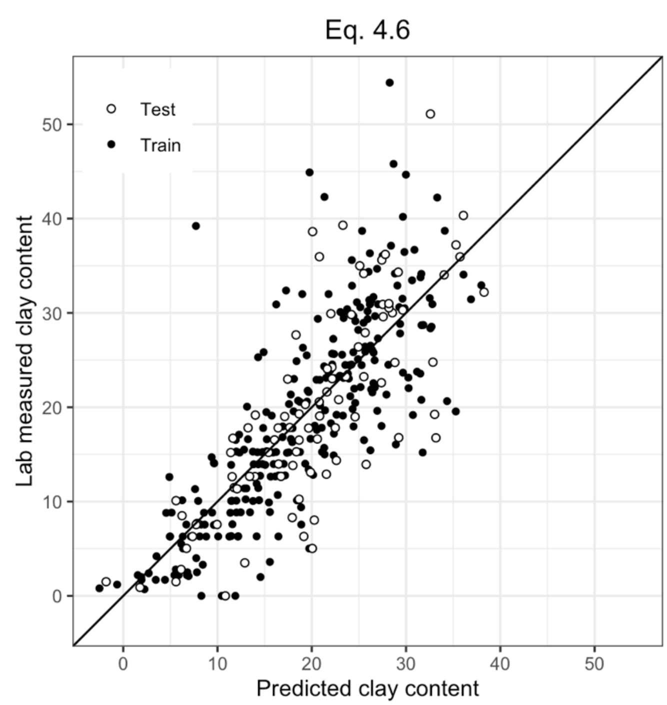

Figure 4.11: Eq. 4.6 for clay % applied to the train and test set.

4.2.3 Algorithmic modeling

Random forest modeling to predict sand contents revealed an $R^2 = 0.658$, RMSE = 10.8, RPD = 1.697111, and RPIQ = 2.448834 when applied to the validation subset (Fig. 4.12). For the prediction of clay contents, model metrics were improved to an $R^2 = 0.625$, RMSE = 6.06, RPD = 1.612843, and RPIQ = 2.539862 (Fig. 4.14). When these two models were applied to the entire modeling dataset, only one value needed to be excluded prior to texture class determination for negative values. The correct texture class was predicted 72% of the time, with sand contents underestimated 51% of the time and clay contents overestimated 54% of the time.

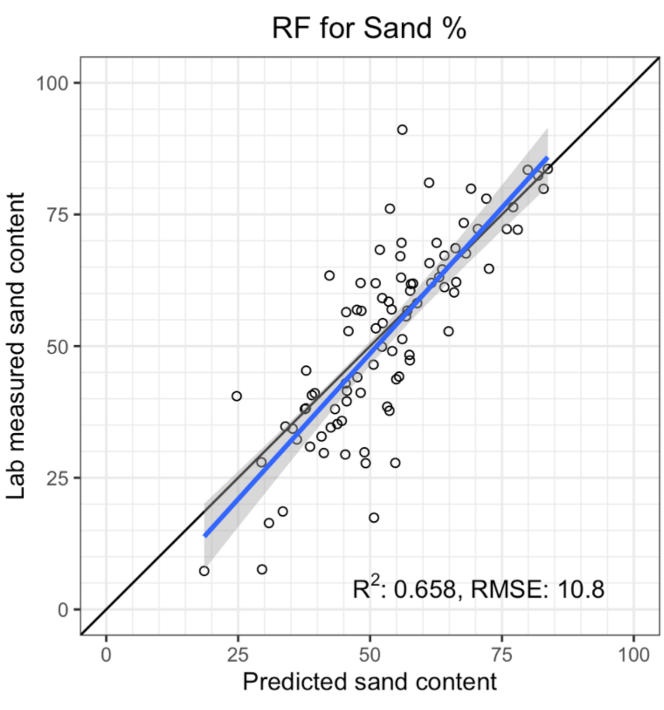

Figure 4.12: RF modeling for sand % on the holdout set.

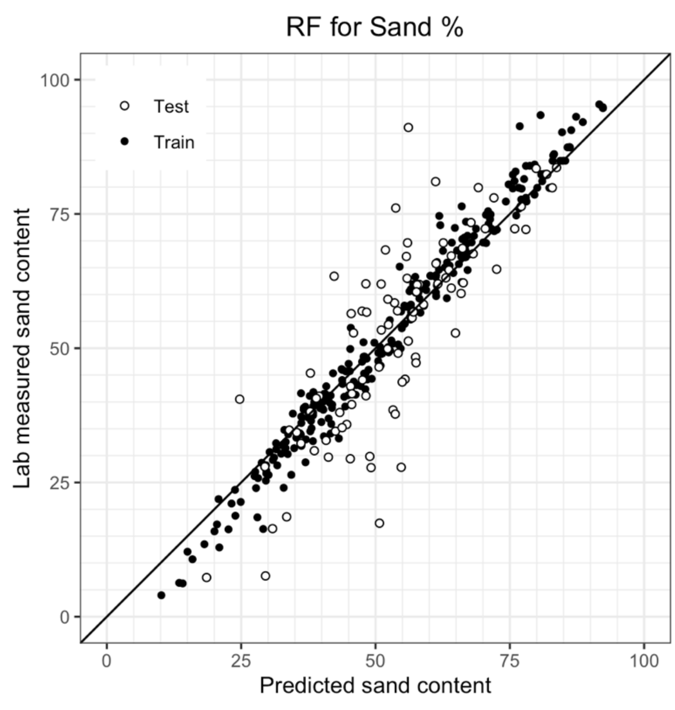

Figure 4.13: RF modeling for sand % on the train and test set.

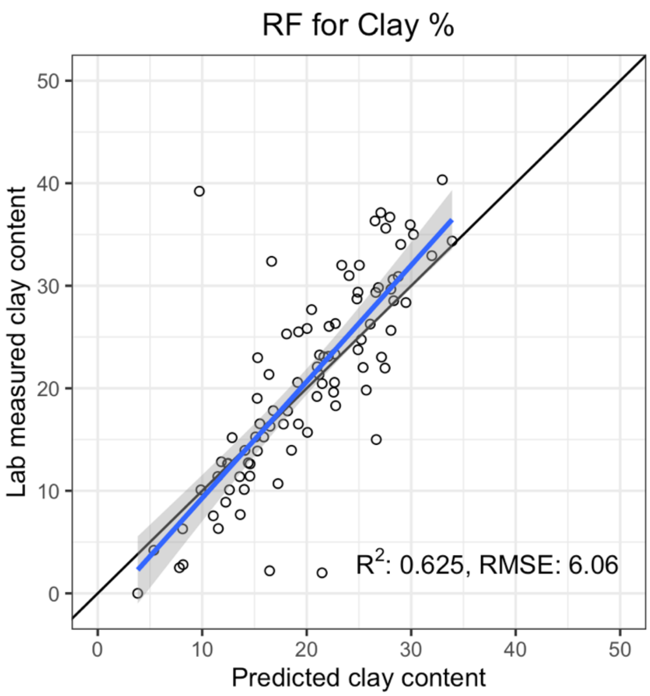

Figure 4.14: RF modeling for clay % on the holdout set.

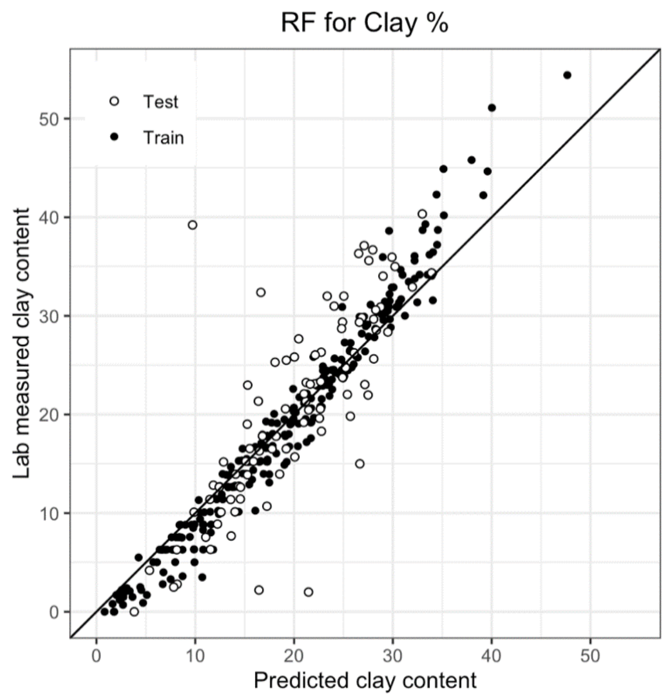

Figure 4.15: RF modeling for clay % on the train and test set.

4.3 CEC

4.3.1 Descriptive statistics

474 samples had measured CEC values and were viable for modeling testing. Summary statistics for CEC values can be found in Table 4.6. A boxplot for each sample set (Fig. 4.16) and for all samples in the dataset (Fig. 4.17) show the spread of CEC values. The Tukey outlier test revealed an upper threshold of 68.8, indicating 1 outlier in the CEC values: SPR/LHBR 34 (CEC: 74.57 cmolc/kg soil).

Table 4.6: CEC summary statistics.

	Sample set					
	SPR/LHBC Mollisols	NRCS Chico	LA Urban	UC Merced	Marine terrace	Total dataset
Samples	218	56	37	4	159	474
Minimum	5.36	0.40	5.118	3.059	2.50	0.40
Median	23.25	13.40	16.059	4.441	10.80	16.14
Mean	26.14	15.78	18.243	7.647	11.24	19.15
Maximum	74.57	41.90	51.647	18.647	32.20	74.57

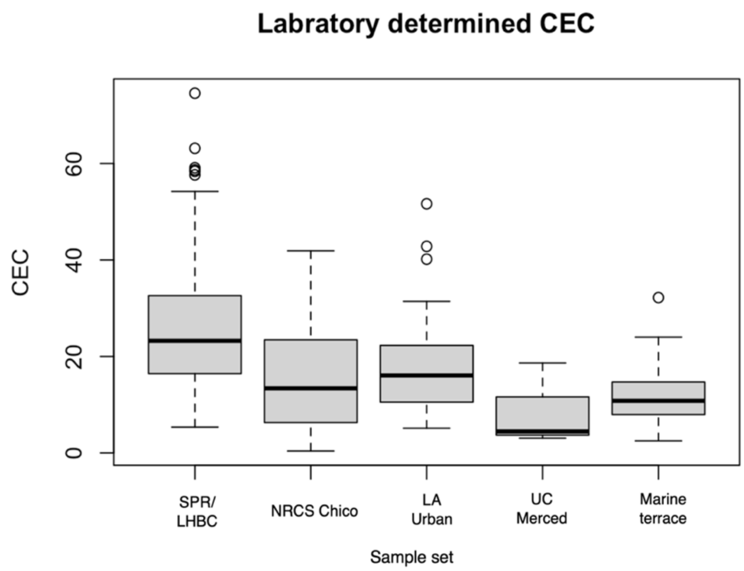

Figure 4.16: CEC by sample set.

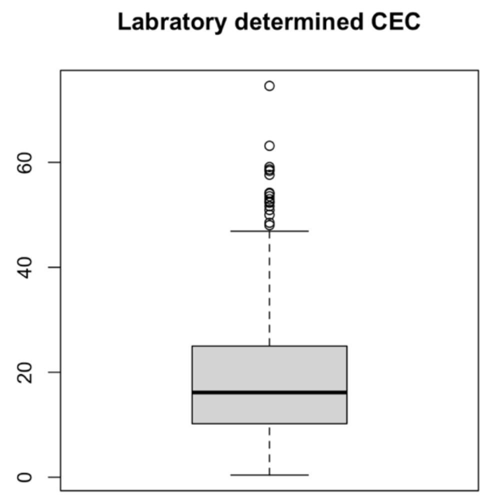

Figure 4.17: CEC for entire dataset.

4.3.2 Data models

Applying Eq. 2.4 to the entire dataset showed no predictive capacity (R^2 = 0.0006, RMSE = 20.9, RPD = 0.5913471, and RPIQ = 0.7008987) and predicted 13 negative CEC values. Generating new coefficients with the variables from Eq. 2.4 using an 80% train dataset produced Eq. 4.7. When applied to the validation sub-dataset, somewhat improved metrics were observed: R^2 = 0.138, RMSE = 11.5, RPD = 1.079546, and RPIQ = 0.9894607.

Applying the tidyverse approach to MLR model building revealed 12 significant variables (Mg, Al, Si, K, Ti, Cr, Cu, Zn, As, Rb, Y, Pb) from 10-fold CV on a 75% training set for predicting CEC. When applied to the test set, Eq. 4.8 produced a further improved $R^2 = 0.761$, RMSE = 6.88, RPD = 1.985829, and RPIQ = 2.078793 (Fig. 4.18). The coefficients found by Sharma et al., (2011) for Louisiana and New Mexico soils is compared with Eq. 4.7 and 4.8 in Table 4.7.

Equation 4.7

$$CEC = 35.83 + 0.0003427 * (Ca) – 0.0007776 * (Ti) – 0.2516 * (V) – 0.04001 * (Cr) + 0.0004214 * (Fe) + 0.07008 * (Cu) – 0.04428 * (Sr) + 0.0315 * (Zr)$$

Equation 4.8

$$CEC = 820.779 – 10.182 * \log(Mg) – 70.462 * \log(Al) – 71.186 * \log(Si) – 25.764 * \log(K) – 15.310 * \log(Ti) – 6.097 * \log(Cr) + 9.246 * \log(Cu) + 15.486 * \log(Fe) + 9.268 * \log(Zn) + 12.062 * \log(As) + 32.262 * \log(Rb) – 7.940 * \log(Sr) – 6.288 * \log(Y) – 6.529 * \log(Pb)$$

Table 4.7: Model parameters for CEC regression models.

Variable	**Eq. 2.4 Sharma et al. (2015)**	**Eq. 4.7 Generated coefficients**	Variable (Logged)	**Eq. 4.8 10-fold CV method**
Constant	17.2507	35.83***	Constant	820.779***
Ca	−0.00036514	0.0003427**	Ca	
Ti	−0.0034957	−0.0007776	Ti	−15.310*
V	0.070977	−0.2516***	V	
Cr	0.070991	-0.04001***	Cr	−6.097**
Fe	0.00059759	0.0004214 ***	Fe	15.486
Cu	0.1479	0.07008*	Cu	9.246**
Sr	−0.062096	−0.04428***	Sr	–7.940*
Zr	0.0056551	0.0315*	Zr	
Al			Al	–70.462***
Si			Si	–71.186***
K			K	–25.764 ***
Zn			Zn	9.268**
As			As	12.062***
Rb			Rb	32.262***
Mg			Mg	–10.182***
Y			Y	−6.288*
Pb			Pb	−6.529***
Sample #	474	474	Sample #	474
R^2	0.0006	0.138	R^2	0.761
RMSE	20.9	11.5	RMSE	6.88
RPD	0.5913471	1.079546	RPD	1.985829
RPIQ	0.7008987	0.9894607	RPIQ	2.078793

Model performance metrics are from the entire dataset for Eq. 2.4 and for the 20/25% validation sets for Eq. 4.7 and 4.8. Significance codes (p-values): '***' 0.001 '**' 0.01 '*' 0.05

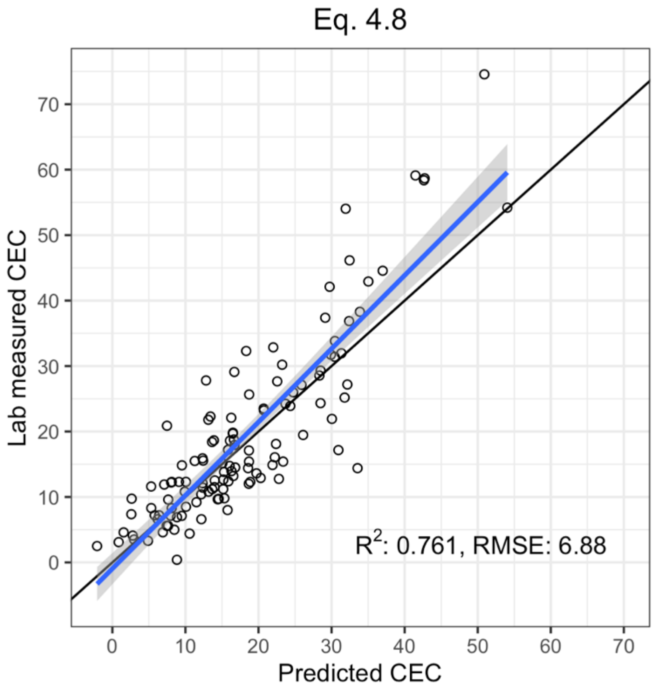

Figure 4.18: Eq. 4.8 for CEC applied to the holdout set.

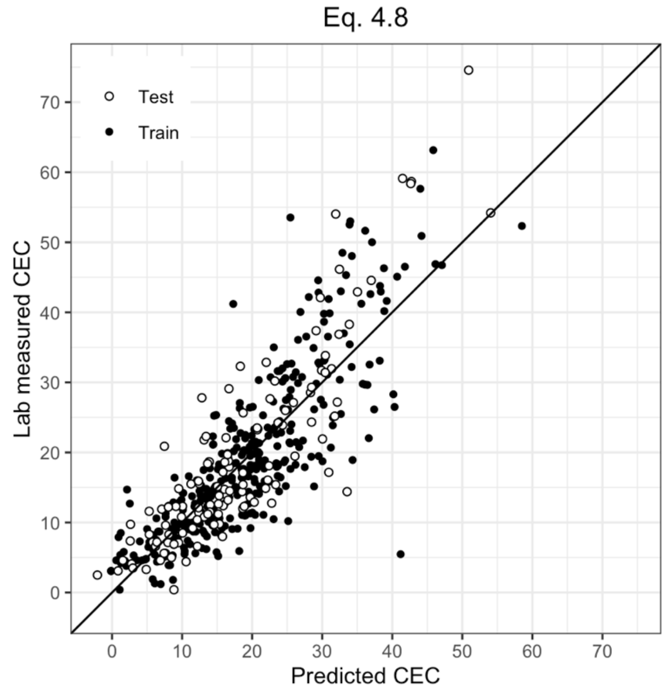

Figure 4.19: Eq. 4.8 for CEC applied to the train and test set.

4.3.3 Algorithmic modeling

A random forest model created with the training subset resulted in an $R^2 = 0.788$, RMSE = 6.79, RPD = 2.009571, and RPIQ = 2.554648 when applied to the validation set (Fig. 4.20).

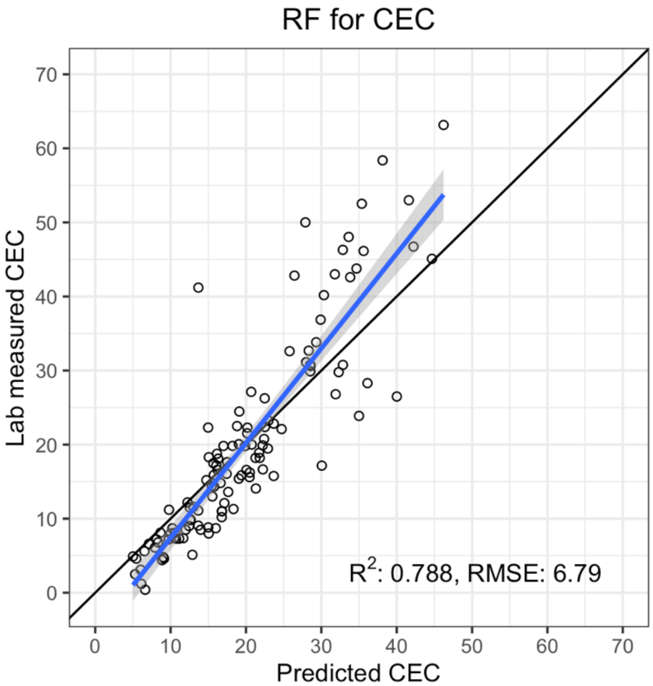

Figure 4.20: RF modeling for CEC on the holdout set.

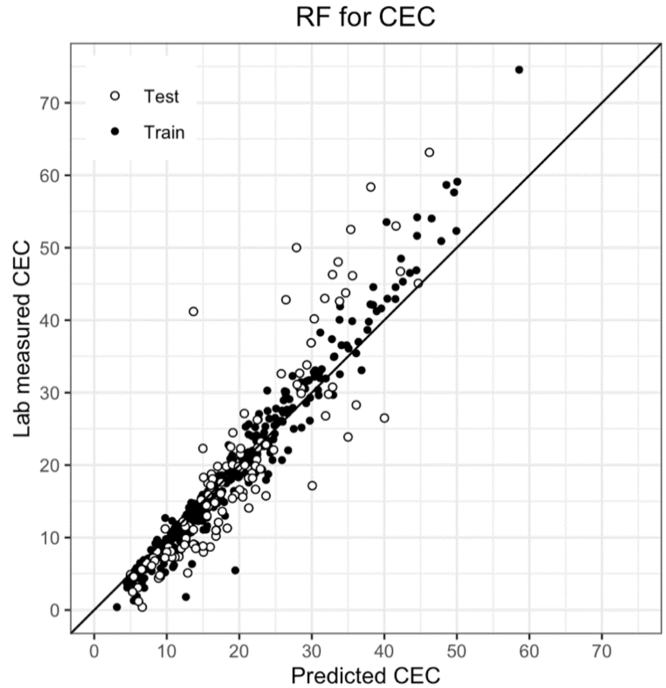

Figure 4.21: RF modeling for CEC on the train and test set.

4.4 SOC, TN, and C:N ratio

4.4.1 Descriptive statistics

After omitting those samples without lab data for the property of interest, 475 samples had measured SOC and N values, and 479 had C:N values. Boxplots for all samples in the dataset (Fig. 4.22-4.27) show the spread of N, SOC, and C:N values for the entire dataset and grouped by sample set. Summary statistics for these properties' values can be found in Table 4.8. For total nitrogen, the Tukey outlier test indicated two outliers: 0.86% and 0.757%. There were also four outliers for SOC: 11.902%, 12.446%, 14.08%, and 11.45%, and 13 outliers for C:N ratio: 27, 28, 28.404, 29, 31, 32, 34, 45.378, 56.9, 81, 85, 99, and 167.

Table 4.8: Summary statistics for SOC, N, and C:N.

	Soil organic carbon (%)	Total nitrogen (%)	C:N ratio
Samples	475	475	479
Minimum	0.02	0.01	1
Median	1.7305	0.1424	11.75
Mean	2.3562	0.17774	13.35
Maximum	14.08	0.86	167

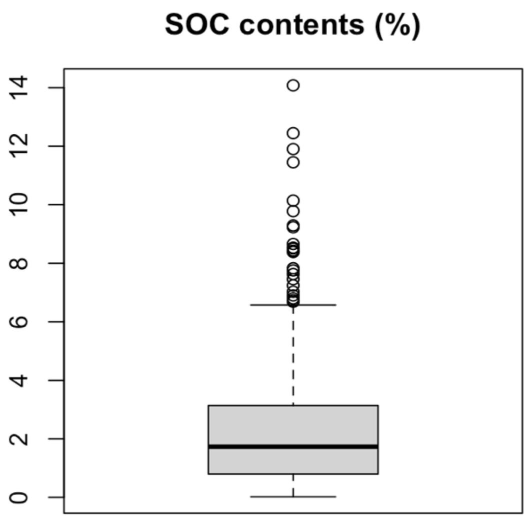

Figure 4.22: SOC % for entire dataset.

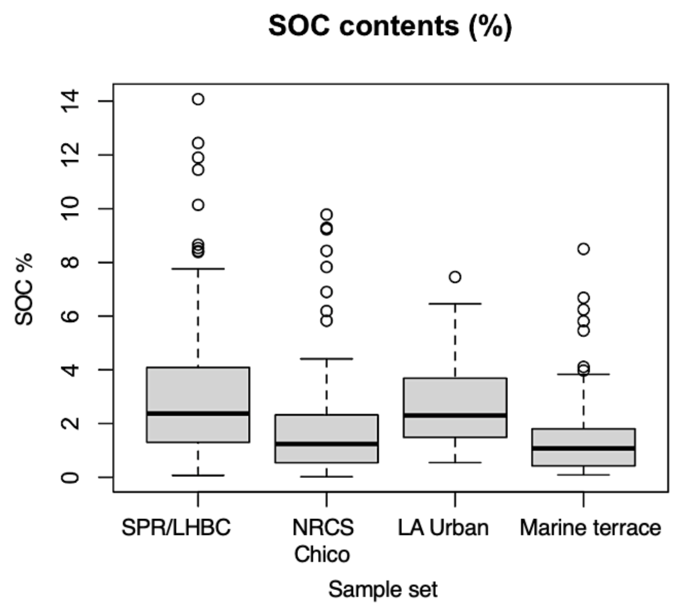

Figure 4.23: SOC % by sample set.

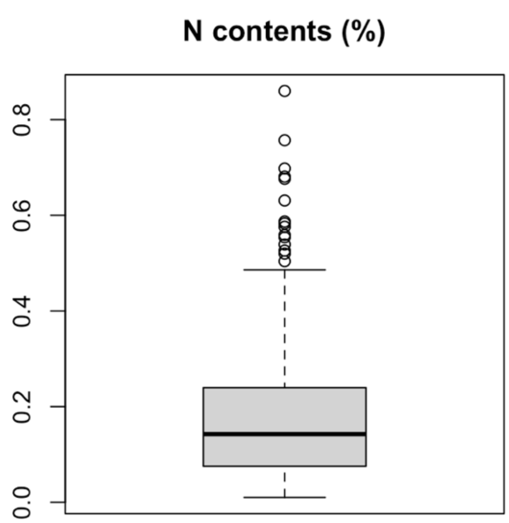

Figure 4.24: TN % for entire sample set.

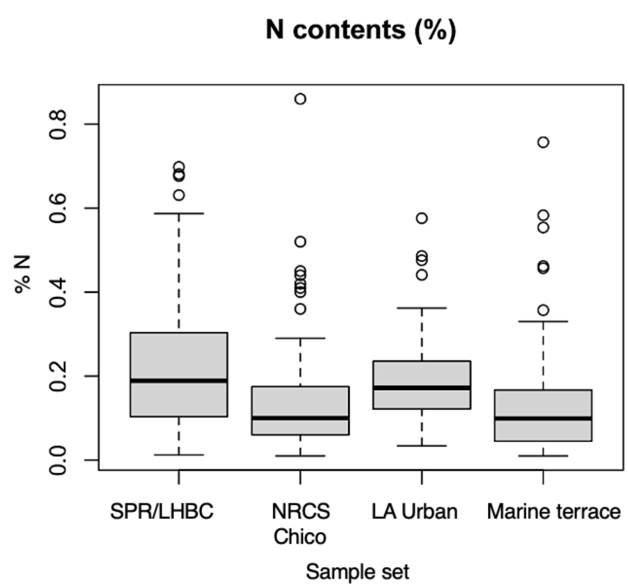

Figure 4.25: TN % for each sample set.

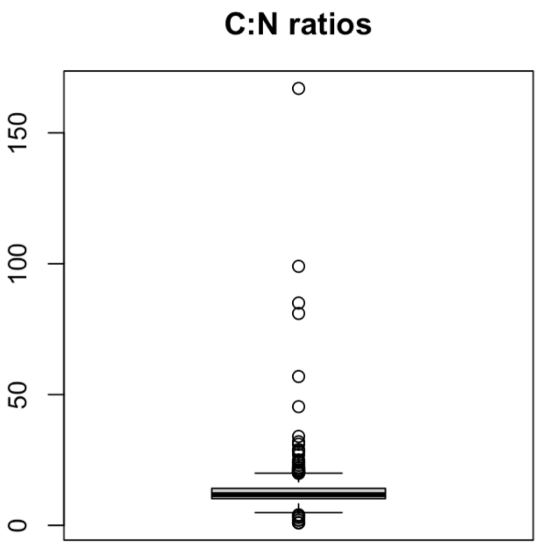

Figure 4.26: SOC to TN % for entire dataset.

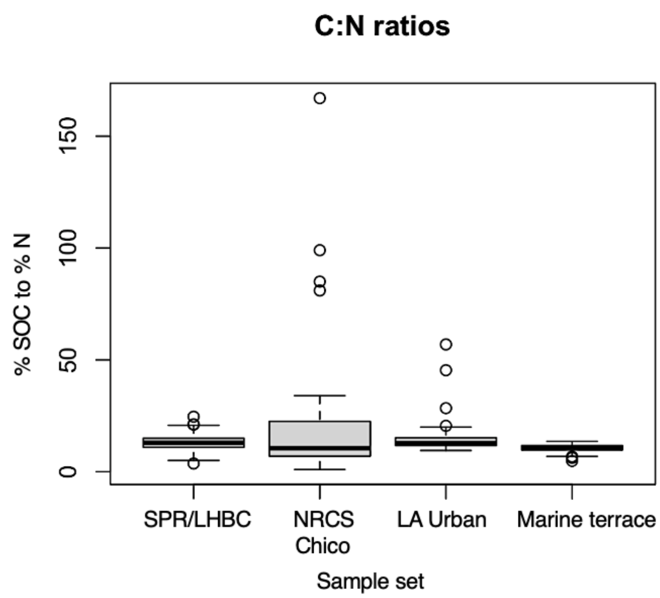

Figure 4.27: SOC to TN % for entire dataset.

4.4.2 Data models

Presently, there is no readily accessible published literature which describes using pXRF derived elemental data to produce regression models for predicting SOC, N or C:N ratio. Despite this, an attempt was made here to see if MLR modeling can produce acceptable estimates of these properties. Using the 10-fold CV on the train set indicated 12 significant elements for predicting SOC. When applied to the test set, this model (Eq. 4.9) produced an R^2 = 0.719, RMSE = 1.01, RPD = 1.767523, and RPIQ = 1.835098 (Fig. 4.28). Applying this same process for total N content produced Eq. 4.10 for total N, and Eq. 4.11 for C:N ratio. When applied to the test set, Eq. 4.10 produced four negative values with an R^2 = 0.699, RMSE = 0.0753, RPD = 1.793764, and RPIQ = 2.081165 (Fig. 4.30) while Eq. 4.11 showed poor performance, producing three negative values and an R^2 = 0.00308, RMSE = 7.02, RPD = 0.5893667, and RPIQ = 0.4986284 (Fig. 4.32). A summary of the model coefficients and their significance can be found in Table 4.9.

Equation 4.9

$$SOC\ \% = 204.4158 - 1.6341 * \log(Mg) - 13.7128 * \log(Al) - 22.8982 * \log(Si) + 1.1454 * \log(S) - 1.1736 * \log(Ca) - 3.8801 * \log(Ti) - 0.8717 * \log(Cr) + 1.8537 * \log(Mn) + 1.8565 * \log(Cu) - 1.5105 * \log(Y) + 3.0772 * \log(Zr) - 0.9256 * \log(Pb)$$

Equation 4.10

$$N\ \% = 9.53996 - 0.09183 * \log(Mg) - 0.81023 * \log(Al) - 0.84295 * \log(Si) + 0.04199 * \log(P) + 0.07529 * \log(S) - 0.15990 * \log(Ti) - 0.20653 * \log(V) + 0.06524 * \log(Mn) + 0.11225 * \log(Cu) - 0.20783 * \log(Sr) + 0.12302 * \log(Zr)$$

Equation 4.11

$$C{:}N = -62.765 + 15.938 * \log(Al) - 17.086 * \log(K) + 6.621 * \log(Ca) + 23.885 * \log(Cu) - 23.672 * \log(Zn) - 12.785 * \log(As) + 13.612 * \log(Rb) + 17.253 * \log(Sr)$$

Table 4.9: Model parameters for SOC %, TN % and C:N regression models.

Variable (Logged)	**Eq. 4.9 SOC (%)**	**Eq. 4.10 TN (%)**	**Eq. 4.11 C:N**
Constant	204.4158***	9.53996***	−62.765
Mg	−1.6341***	−0.09183***	
Al	−13.7128***	−0.81023***	15.938
Si	−22.8982***	−0.84295***	
K			−17.086
Ca	−1.1736***		6.621**
Ti	−3.8801***	−0.15990*	
Mn	1.8537***	0.06524**	
S	1.1454***	0.07529***	
Rb			13.612
Y	−1.5105***		
Cu	1.8565***	0.11225***	23.885
Zr	3.0772***	0.12302**	
Cr	−0.8717**		
Pb	−0.9256***		
P		0.04199*	
V		−0.20653**	
Sr		−0.20783***	17.253
As			−12.785
Zn			−23.672
Sample number	475	475	479
R^2	0.719	0.699	0.00308
RMSE	1.01	0.0753	7.02
RPD	1.767523	1.793764	0.5893667
RPIQ	1.835098	2.081165	0.4986284

Model performance metrics are from the 25% validation sets for SOC and TN model equations. CN variables were not tested for their significance due to not meeting residual assumptions. Significance codes (p-values): '***' 0.001 '**' 0.01 '*' 0.05

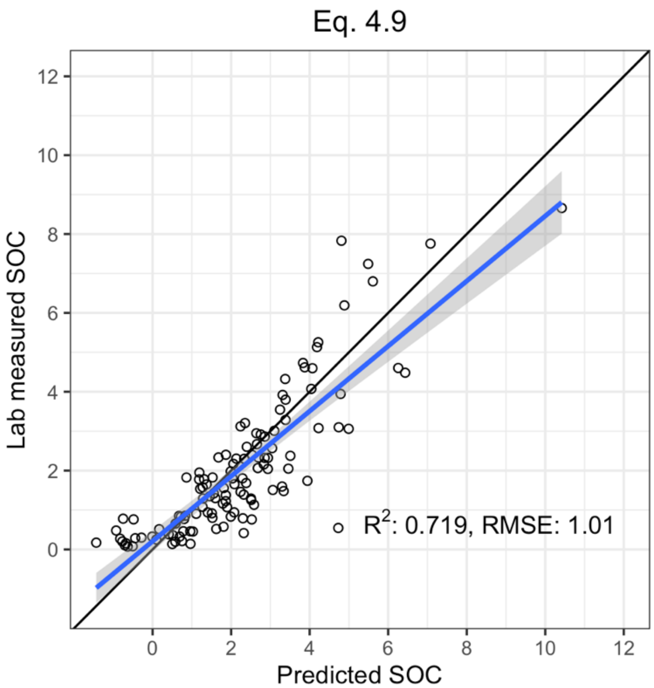

Figure 4.28: Eq. 4.9 for SOC % applied to the holdout set.

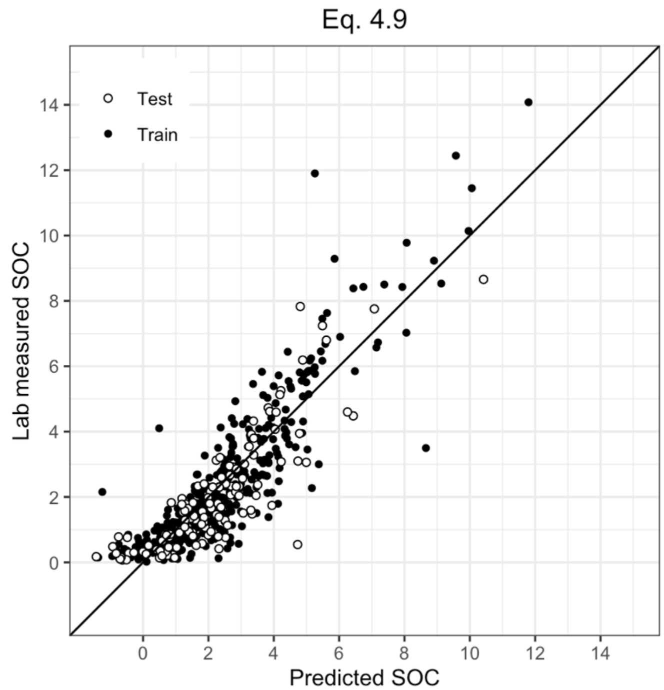

Figure 4.29: Eq. 4.9 for SOC % applied to the test and train set.

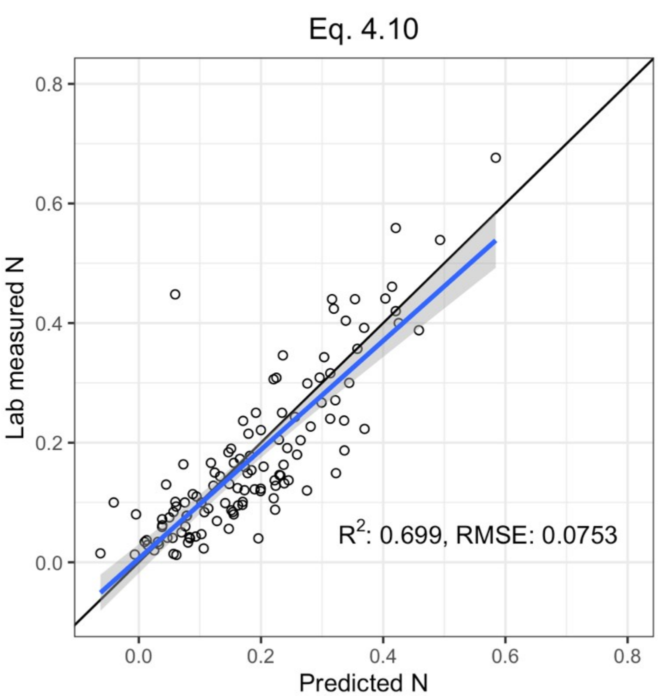

Figure 4.30: Eq. 4.10 for TN % applied to the holdout set.

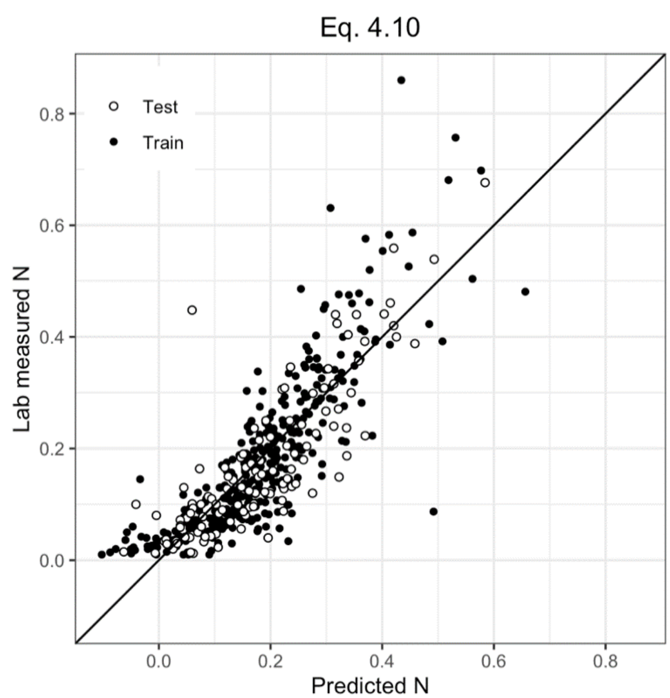

Figure 4.31: Eq. 4.10 for TN % applied to the train and test set.

4.4.3 Existing RF methodology

Following the process outlined by Towett et al. (2015) using randomForest applied to SOC data, the MSE determined on a 50% hold out sample of the data was compared to the MSE found by OOB validation. The OOB errors (1.172955) were found to be only 6.4% lower than for the 50% hold out sample (1.252541). Similarly, Towett et al. (2015) found OOB errors to be only 10% lower than for the 50% hold out sample, which authors used as validation of the OOB error calculation process and justification for reporting model validation metrics from the entire modeling dataset. When the resultant RF model was applied to the entire dataset, model results indicated $R^2 = 0.748$, RMSE = 1.08, RPD = 1.995045 and RPIQ = 2.155392 (Fig. 4.31). Thus, when compared to the results found by Towett et al., (2015) for SOC, a higher R2 (0.75 vs 0.68) but also higher RMSE (1.1 vs 0.7) was obtained.

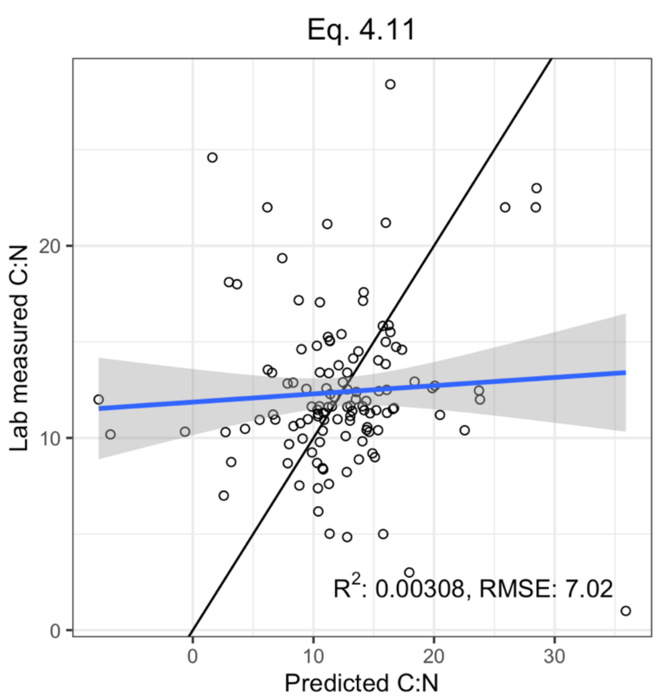

Figure 4.32: Eq. 4.11 for SOC to TN % applied to the holdout set.

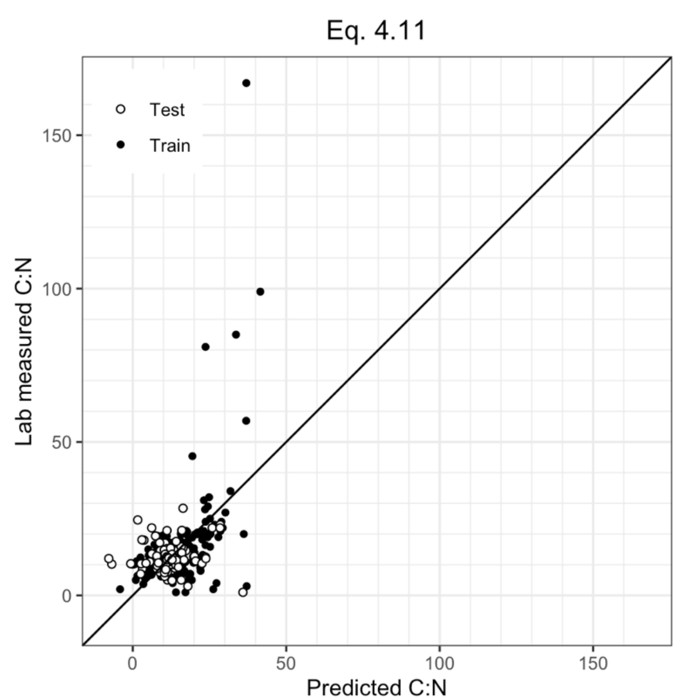

Figure 4.33: Eq. 4.11 for SOC to TN % applied to the train and test set.

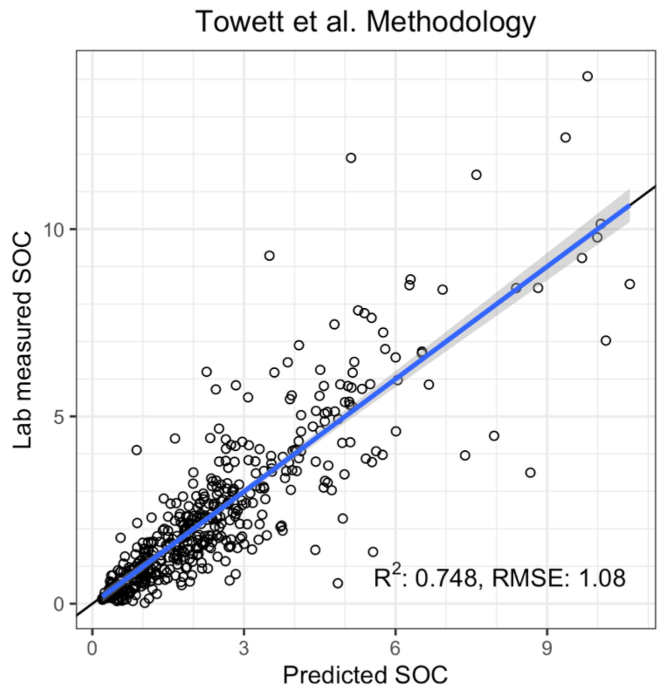

Figure 4.34: Towett et al. (2015) random forest modeling methodology applied to the entire dataset.

4.4.4 Algorithmic modeling

Applying random forest modeling to predict SOC with use of the tidyverse package resulted in an R^2 = 0.735, RMSE = 1.14, RPD = 1.877131, and RPIQ = 2.335256 when applied to the test set (Fig. 4.32). For total nitrogen, this technique yielded an R^2 = 0.782, RMSE = 0.0615, RPD = 2.041243, and RPIQ = 2.960288 (Fig. 4.34), and for C:N ratio, results showed R^2 = 0.373, RMSE = 9.04, RPD = 1.263024, and RPIQ = 0.4686274 (Fig. 4.36).

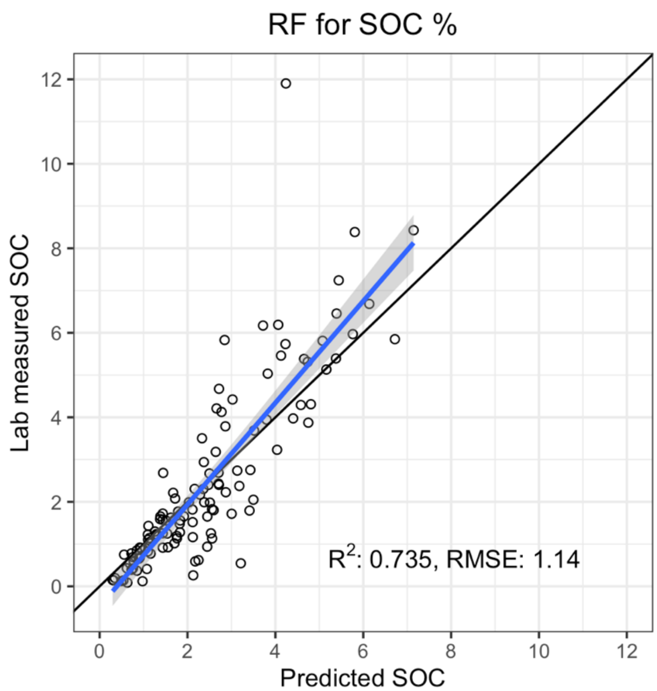

Figure 4.35: RF modeling for SOC % on the holdout set.

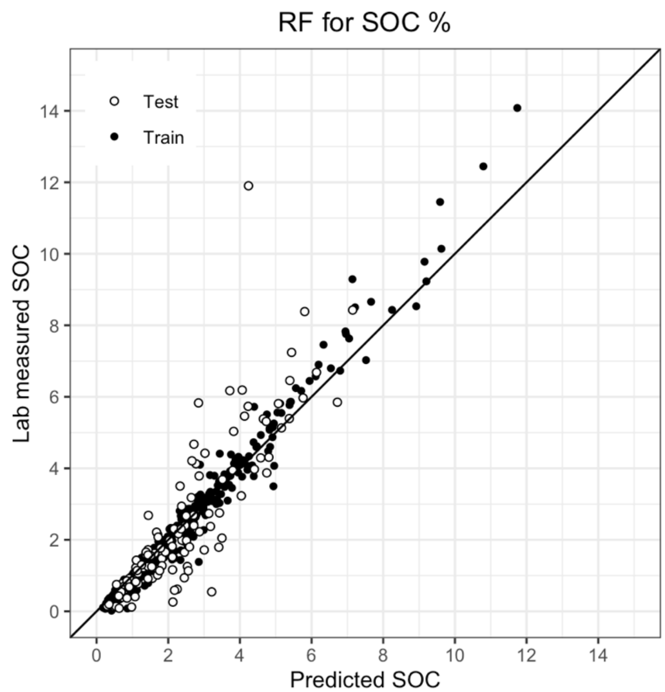

Figure 4.36: RF modeling for SOC % on the train and test set.

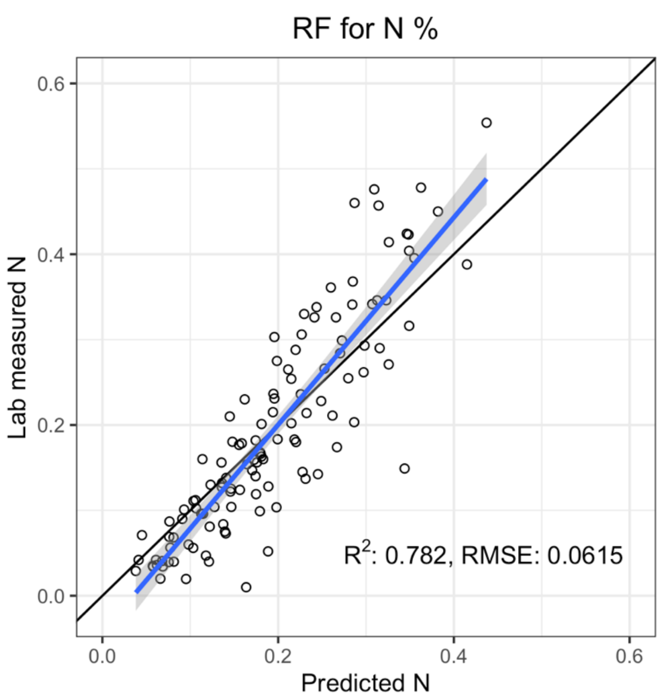

Figure 4.37: RF modeling for TN % on the holdout set.

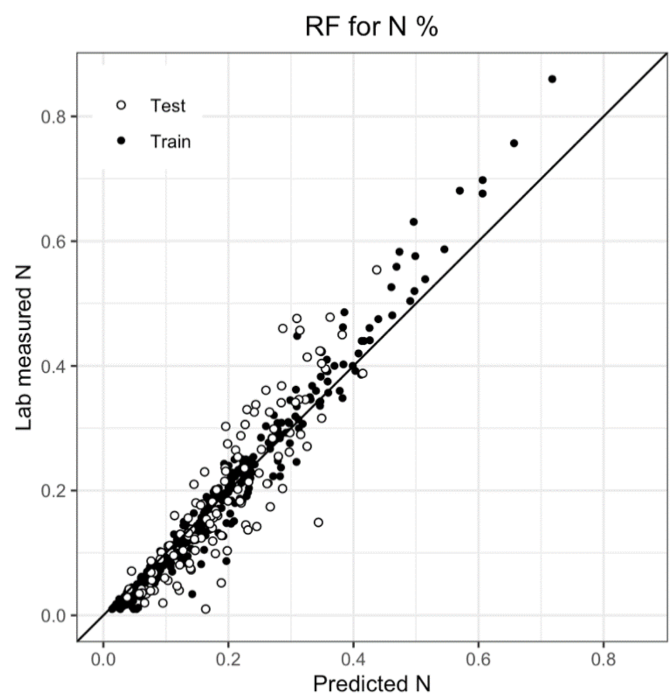

Figure 4.38: RF modeling for TN % on the train and test set.

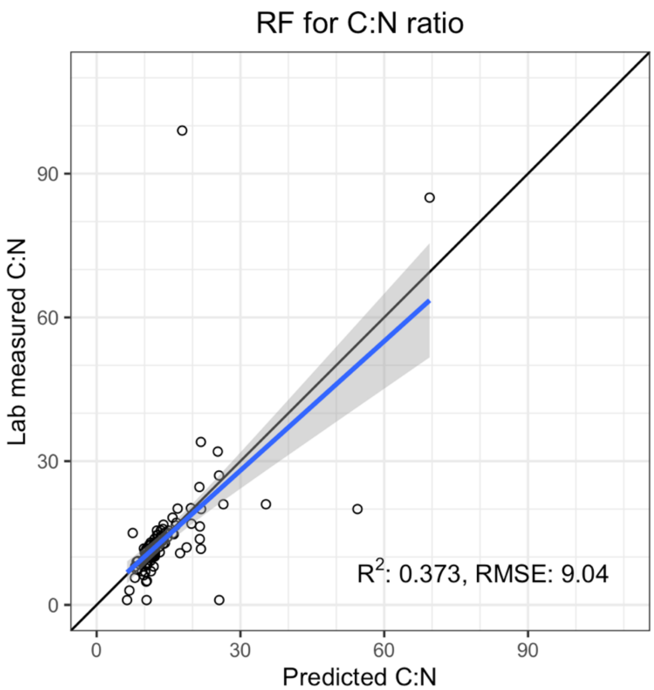

Figure 4.39: RF modeling for C:N on the holdout set.

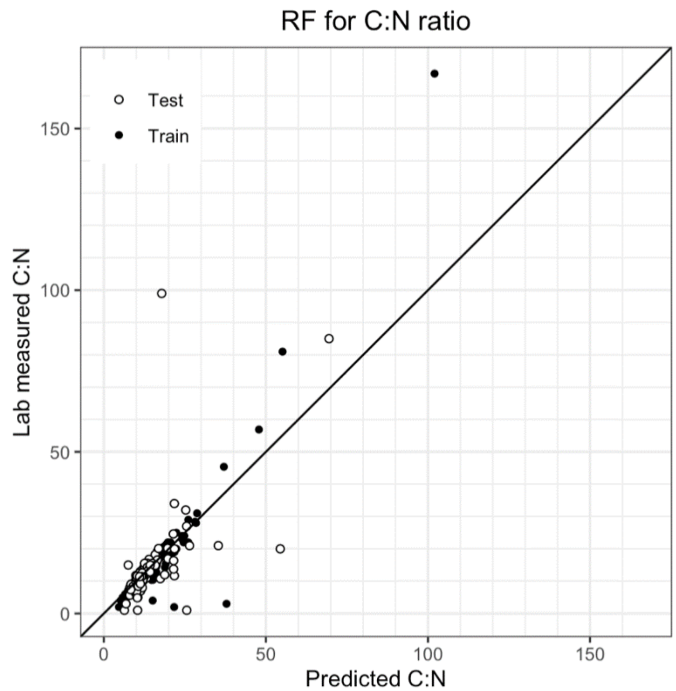

Figure 4.40: RF modeling for C:N on the train and test set.

4.5 Significance of land type and characterization methods on predictions

4.5.1 pH

The addition of land type and sample set as predictor variables improved pH metrics from $R^2 = 0.479$/RMSE = 0.520 to $R^2 = 0.547$/RMSE = 0.492. The results for models within land type and methods are shown in Table 4.10. Model equations for the groupings within each property can be found in Appendix E.

Table 4.10: pH MLR model metrics differentiated by land type and lab method.

	Complete dataset		Land Type			Methodology	
Metric	MLR model	RF model	Forest	Grassland	Marine terrace	1:1	Saturated paste
Sample #	478	478	166	81	159	319	159
R^2	0.532	0.485	**0.554**	**0.721**	0.463	0.488	-
RMSE	0.489	0.490	**0.392**	**0.346**	**0.396**	0.531	-
RPD	1.456	1.377	**1.492**	**1.832**	1.330	1.405	-
RPIQ	1.733	1.547	**2.182**	**2.484**	1.580	**1.791**	-

Bolded text indicates metric was improved beyond MLR/RF modeling using complete dataset. Marine terrace and saturated paste groupings represent the same group of samples.

4.5.2 Texture: sand

The addition of land type and sample set as predictor variables improved sand metrics from R^2 = 0.614/RMSE = 12.6 to R^2 = 0.644/RMSE = 12.1. The results for models within land type and methods are shown in Table 4.11.

Table 4.11: Sand content MLR model metrics differentiated by land type and lab method.

	Complete dataset		Land Type			Methodology	
Metric	MLR model	RF model	Forest	Grassland	Marine terrace	Hydrometer	Pipette
Sample #	358	358	86	41	159	298	60
R^2	0.616	0.658	**0.672**	**0.895**	**0.884**	0.462	0.348
RMSE	12.3	10.8	13.8	**5.56**	**5.71**	13.3	24.5
RPD	1.601	1.697	1.523	**2.985**	**2.905**	1.239	**1.197**
RPIQ	2.589	2.449	**2.634**	**4.945**	**4.812**	1.902	1.820

Bolded text indicates metric was improved beyond MLR/RF modeling using complete dataset.

4.5.3 Texture: clay

Including land type and sample set variables when predicting clay content improved metrics from R^2 = 0.586/RMSE = 6.72 to R^2 = 0.639/RMSE = 6.28. Resultant models within land type and methodology are shown in Table 4.12.

Table 4.12: Clay content MLR model metrics differentiated by land type and lab method.

	Complete dataset		Land Type			Methodology	
Metric	MLR model	RF model	Forest	Grassland	Marine terrace	Hydrometer	Pipette
Sample #	358	358	86	41	159	298	60
R^2	0.599	0.625	**0.714**	0.101	**0.812**	0.624	**0.631**
RMSE	6.83	6.06	6.60	9.25	**3.55**	**6.05**	10.1
RPD	1.586	1.613	**1.802**	0.559	**2.194**	1.594	1.143
RPIQ	2.294	2.540	**3.425**	0.743	**2.760**	2.180	1.486

Bolded text indicates metric was improved beyond MLR/RF modeling using complete dataset.

4.5.4 CEC

Inclusion of land type and sample set variables as predictor variables for CEC improved metrics from $R^2 = 0.694$/RMSE = 6.80 to $R^2 = 0.721$/RMSE = 6.50. Resultant models within land type and methodology are shown in Table 4.13.

Table 4.13: CEC MLR model metrics differentiated by land type and lab method.

	Complete dataset		Land Type			Methodology			
Metric	MLR model	RF model	Forest	Grass land	Marine terrace	Ammonia absorbance	UN-FAO	CEC 7	S - 10.10
Sample #	474	474	164	81	159	218	41	56	159
R^2	0.761	0.788	**0.819**	0.517	0.653	0.689	0.238	0.646	-
RMSE	6.88	6.79	7.52	**5.44**	**2.57**	7.38	12.8	8.54	-
RPD	1.986	2.010	**2.328**	1.083	1.717	1.777	1.071	1.141	-
RPIQ	2.079	2.555	2.497	1.139	**2.721**	2.248	0.801	1.562	-

Bolded text indicates metric was improved beyond MLR/RF modeling using complete dataset. Marine terrace and S - 10.10 groupings represent the same group of samples.

4.5.5 SOC content

The additional categorical variables of land type and sample set for SOC prediction improved metrics: from $R^2 = 0.766$/RMSE = 1.03 to $R^2 = 0.800$/ RMSE = 0.901. Models created within land type for SOC are displayed in Table 4.14.

Table 4.14: SOC % MLR model metrics differentiated by land type.

	Complete dataset		Land Type		
Metric	MLR model	RF model	Forest	Grassland	Marine terrace
Sample #	472	472	168	81	159
R^2	0.719	0.735	**0.815**	0.661	**0.821**
RMSE	1.01	1.14	1.3	1.01	**0.738**
RPD	1.768	1.877	**2.204**	1.638	**2.367**
RPIQ	1.835	2.335	**2.432**	1.845	2.239

Bolded text indicates metric was improved beyond MLR/RF modeling using complete dataset.

4.5.6 TN content

Including land type and sample set variables when predicting TN content improved metrics from $R^2 = 0.689$/RMSE = 0.0767 to $R^2 = 0.717$/RMSE = 0.0724. Resultant models within land type and methodology are shown in Table 4.15.

Table 4.15: TN % MLR model metrics differentiated by land type.

	Complete dataset		Land Type		
Metric	MLR model	RF model	Forest	Grassland	Marine terrace
Sample #	475	475	165	81	159
R^2	0.699	0.782	0.738	**0.842**	0.762
RMSE	0.075	0.062	0.084	**0.057**	**0.0544**
RPD	1.794	2.041	1.973	**2.456**	1.920
RPIQ	2.081	2.960	2.307	**4.435**	2.697

Bolded text indicates metric was improved beyond MLR/RF modeling using complete dataset.

4.5.7 C:N ratio

Inclusion of land type and sample set variables as predictor variables for C:N ratio improved metrics from $R^2 = 0.274$/RMSE = 7.62 to $R^2 = 0.402$/RMSE = 7.2. Resultant models within land type and methodology are shown in Table 4.16.

Table 4.16: C:N ratio MLR model metrics differentiated by land type.

	Complete dataset		Land Type		
Metric	MLR model	RF model	Forest	Grassland	Marine terrace
Sample #	472	472	168	81	159
R^2	0.003	0.373	0.273	0.192	0.342
RMSE	7.02	9.04	**6.46**	**2.03**	**1.04**
RPD	0.589	1.263	0.705	1.097	1.240
RPIQ	0.499	0.469	**0.711**	**1.067**	**1.177**

Bolded text indicates metric was improved beyond MLR/RF modeling using complete dataset.

Chapter 5
DISCUSSION

The objectives of this study were to assess the accuracy of existing data models and model building approaches to predict pH, texture, CEC, SOC%, TN% and C:N ratio from pXRF elemental data for 480 California soils. As expected, existing data models were inadequate in predicting California soil properties, likely due to the fact that elemental coefficients will be specific to the sample set used to calibrate the model. While certain covariates can be seen to be important for specific properties, the exact weight of these coefficients will vary by sample set, making large range data models unrealistic. Multiple linear regression and random forest models were also constructed specific to the California soils dataset, the results of which can be seen in summary Table 5.1. Overall, RF models tended to produce better estimates when compared to MLR models. Variable importance plots from RF models can be created to uncover the relative importance of predictors while MLR models can give an estimate of absolute importance through variable coefficients. However, the multicollinearity of many important soil elements, wherein some elements are associated with each other, (which was not investigated in this study) can obfuscate beta coefficients given by MLR model equations. Grouping sample sets by land type and lab characterization approach showed no clear improvement in data models but may have an effect for larger sample sets where RF modeling can be used. Considering that most existing literature in which pXRF is used to predict soil properties has only a single characterizing lab body for the 'lab truth' measurements, the model predictions for this study were decent, and are expected to improve with more contained sampling areas used to calibrate models.

Table 5.1: Test set model metrics for each property investigated.

	Model outcomes							
	MLR				RF			
	R^2	RMSE*	RPD	RPIQ	R^2	RMSE*	RPD	RPIQ
pH	0.532	0.489	1.456	1.733	0.485	0.490	1.377	1.547
Sand %	0.616	12.3	1.601	2.589	0.658	10.8	1.697	2.449
Clay %	0.599	6.83	1.586	2.294	0.625	6.06	1.613	2.540
CEC	0.761	6.88	1.986	2.079	0.788	6.79	2.010	2.555
SOC %	0.719	1.01	1.768	1.835	0.735	1.14	1.877	2.335
TN %	0.699	0.0753	1.794	2.081	0.782	0.062	2.041	2.960
C:N	0.003	7.02	0.589	0.499	0.373	9.04	1.263	0.469

Red indicates poor predictive power, yellow indicates fair models with the potential for improvement, and green represents stable and accurate models. The ranges for red, yellow, and green categorization were as follows: R^2: <0.6, 0.6 - 0.8, >0.8 (Malley et al., 2004), RPD: <1.4, 1.4-2, >2 (Chang et al., 2001), RPIQ: ≥ 1.5 , ≥ 1.9, ≥ 3.0 (Veum et al., 2015).
*No established range for poor, fair, and good RMSE values have been established for the properties. Depending on the intended purpose, acceptable levels of error may vary.

5.1 RF models tended to outperform MLR models

This study investigated and built MLR data models as well as RF algorithmic models. While many different algorithmic modeling approaches exist, RF was chosen as the machine learning algorithm to compare predictions from data models. The reasons for this decision were that RF handles complex data types well without the need to scale or normalize predictor variables (Towett et al., 2015). Since RF is a tree-based model which relies on rules to make predictions, the need to pre-process features is eliminated, unlike many other types of machine learning algorithms which assume normality or compute distance. In the past, the use of machine learning for any regression or classification problem may have been discouraged by a steep learning curve of various statistical software programs. However, modern statistics packages such as tidymodels in R makes it relatively straightforward to adapt machine learning approaches. Both RF and MLR models can also be used to determine variance importance, via variable importance plots and coefficient values that indicate the influence of certain elements for predicting a given property. These are helpful where relationships between elements and the properties of interest may help explain the underlying processes which connect the chemistry of the sample to its observable characteristics. For six out of the seven soil properties investigated, RF models improved predictive estimates. The only property to have worse estimates from the RF model was pH, which may indicate that this feature is best predicted through a linear relationship. Thus, when prediction accuracy is the priority of modeling efforts, RF models will almost always be most appropriate, likely due to the non-linear nature of this type of modeling.

5.2 Complications for MLR model interpretation

Despite prediction being the main goal of MLR modeling, residual plots were created (Appendix D) to see if interpretations surrounding coefficient values and their significance could be carried out. Although some of the residual's vs fitted and scale-location plots appear to follow a non-horizontal line, the studentized Breusch-Pagan test produced a p-value <0.05 for each property, to rejecting the null hypothesis of homoscedasticity and meet the assumption of random variance. The normal Q-Q plots displaying the distribution of values shows a normal distribution for pH, CEC, SOC, and TN. For sand and clay, however, the Q-Q plots appear to have a light tail, which indicates that compared to the normal distribution, more data is located at the extremes of the distribution and less is located in the center of the distribution. This may suggest that stronger transformations such as square root transformations may be needed to meet the assumption of normal distribution. The Q-Q plot for CN has a heavy tailed distribution, which indicates more extreme values are present than would be expected in a normal distribution. Additionally, the residuals vs leverage plot showed highly influential points for CN only, which can alter regression outcomes. A high number of outlier data points for CN ratio likely contributed to the non-normal distribution and high influence points. In summary, pH, SOC, and TN showed random variance and normal distribution of residuals, while clay and sand had slightly non-normal distributions, and CN was even more non-normal. Thus, since regression assumptions were not met universally across the board, coefficient significance and interpretation of significant variables and beta coefficients, especially for texture, should be taken with caution. No attempt was made to assign variable significance or interpretation for MLR CN models. Furthermore, the independence assumption of the datapoints may be invalid due to the fact that multiple samples from the same point locations (at different depths) may show spatial autocorrelation. In addition, multicollinearity of elemental concentrations was not controlled for, which can weaken the reliability of statistical inferences.

Summarily, predictive power was the main goal of the study, and model performance metrics (R^2, RMSE, RPD, RPIQ) are still valid without these assumptions. Any interpretations of the MLR models were best attempts to connect sample composition to physical and chemical properties and should be taken with discretion.

5.3 Data models outperform machine learning for pH

The existing model equation created for predicting pH performed considerably worse (RMSE=1.21) compared to the findings of Sharma et al., (2014) who found an RMSE = 0.822. The model was noticeably improved by finding coefficients from the CA soils dataset (RMSE = 0.666) with the same variables as the model equation. An MLR approach using 10-fold CV further improved predictions to RMSE = 0.489, signifying an impressive ability to predict pH within half a unit through a linear relationship. In fact, the MLR model represented by Eq. 4.2 outperformed the RF model in all four of the calculated model metrics. Contrarily, Wan et al. (2019) found estimates of pH to be improved by non-linear modeling of pH from pXRF spectra (PLSR vs SVMR) and attributed the finding to non-linear relationships between pH and the chemical composition.

Deriving a relationship between elemental makeup and pH is not surprising, since it is well known that the availability of plant nutrients in the soil depends upon soil reaction. Since the proportion of basic cations (Ca, Mg, and K) to H^+ ions directly affects the pH, it was expected that these elements would be significant. However, only Ca was included in Eq. 4.2. Calcium may have been an important element for pH prediction due to calcium containing amendments such as gypsum in agricultural soils, shell meal from the marine terrace soils, and a high quantity of base cations in Mollisols. The heavy metals Ni and Zn may indicate pH level because these cationic metals are more highly soluble at low pH levels (USDA-NRCS, 2000). In addition, P is directly affected by pH, reaching with Ca in alkaline environments to form soluble compounds and reacting with Fe in acidic environments to form soluble compounds (Snyder, 2014). Given the relationship between the presence of aluminum and exchangeable soil acidity (Weil and Brady, 2017), it was surprising that Al was not indicated as a significant variable for Eq. 2.4.

For some applications, a prediction of pH within half a unit may be adequate. For purposes such as these, creating a MLR or RF model calibrated to the specific dataset may give the advantage of getting rough estimates of pH quickly. Libohova et al., (2018) found that uncertainty in soil pH measurement can vary widely, between an error of 0.06

for measurement methods and up to 1.3 for database attribution of pH using spatial interpolation. Thus, for pH determination, the amount of acceptable error should be determined beforehand. If an error up to above 1 unit is acceptable, digital soils maps which predict pH using polygon-based aggregation and spatial interpolation rules such as the US Soil Survey Geographic (SSURGO) and General Soil Map (STATSGO2) can be a convenient resource (Libohova et al., 2018). However, if a large collection of datapoints is needed within budget and time constraints, indirect pH measurement with pXRF analysis is a viable option. Considering that pH can be quickly measured in the field using a miniaturized pH meter to indicate pH within 0.01 units (Weil and Brady, 2017), direct measurements of this property may make more sense than use of predictive models.

5.4 Soil texture: clay lends itself to better predictions than sand

To assess the suitability for predicting soil texture from pXRF analysis, sand and clay % were estimated through backward stepwise models created following the methods of Zhu et al. (2011). This method gave inadequate predictions of these contents (RMSE = 14.2/7.2). Despite its pervasive use, the shortcomings of stepwise regression as a variable selection technique are evident. For instance, explanatory variables unrelated to the dependent variable may just happen to show significance while variables without a causal effect on the outcome may not register as significant— resulting in an overfit model that performs poorly on unseen data (Smith, 2018). Additional issues with stepwise regression include artificially high R^2 values and low p-values, falsely high regression coefficients (in absolute value), and low biased standard errors for regression coefficients leading to inaccurately narrow confidence intervals for predicted values (Harrell, 2015). Despite the drawbacks of this approach and performing the worst of the three methods attempted, predictions made both the MLR method with 10-fold CV and the RF modeling approach did not appreciably improve estimates. Only a slight improvement was seen with the MLR approach compared to the stepwise approach (RMSE reduction of 1.9% for sand and 0.4% for clay). There was also only a slight prediction improvement between the MLR approach and the RF approach (further RMSE reduction of 1.5% for sand and 0.77% for clay). The RF model that gave the best result still had an RMSE of 10.8 for sand and 6.06 for clay— representing less accurate estimates than those obtained by Zhu et al. (2011) for sand (RMSE: 5.53-6.26%) and clay (RMSE: 2.66-2.68%) of Louisiana and New Mexico soils. However, correct prediction of the soil texture class was improved from of 55% (MLR) to 72% (RF modeling). For many pertinent applications of texture including surface runoff class and infiltration, designation of the right textural class may be all the information that is needed. Where within-lab error for traditional sedimentation methodologies (hydrometer and pipette) have relatively low error rates (0-6%), these methods are still inherently biased and oftentimes based on untrue assumptions (Salley et al., 2018). In addition, the time and cost of PSA analyses is a drawback to traditional lab texture analysis. By contrast, the texture by feel method is rapid and has the advantage of being completed in the field. However, in a study investigating the accuracy of these estimates, Salley et al., (2018) found that amongst professional soil scientists from the NCSS-SCD (National Cooperative Soil Survey soil characterization database), the correct texture class was predicted 66% of the time. Broadening the definition of the correct texture class to include adjacent textural classes, authors found that accuracy was increased to 91% for professionals. By comparison, the

RF models produced for texture could predict the correct texture class or adjacent class 98.6% of the time. Thus, pXRF estimates of sand and clay may be a good intergrade between higher accuracy lab measurements and more subjective texture by feel measurements in the field.

Both our MLR (Eq. 4.6) and RF model estimates of clay content produced lower RMSE values (6.83 and 6.06 respectively) than GLM, SVM, and RF modeling approaches (RMSE = 9.84, 7.11, 7.68) taken to predict clay of subsuperficial horizons from pXRF data by Silva et al. (2020). However, authors did achieve a lower RMSE for sand contents from GLM and RF modeling (11.92, 8.53) when compared to our RF model (RMSE = 10.8). Compared to the findings of Duda et al. (2017) both our MLR and RF models outperformed their SVR model findings for sand content (for R^2, RPD and RPIQ but not for RMSE). Our MLR and RF models for clay content outperformed their SVR model for clay % within all measured metrics.

Consistently superior predictions for clay than sand is likely a result of the inability of the pXRF to detect light elements correlated to quartz-derived sand (SiO_2) in comparison to the heavier elements associated with the more highly weathered clay fraction. Both Fe and Al were highly significant ($p<0.001$) coefficients for clay prediction, which suggests that oxides in the soil may provide the basis for detecting this textural fraction. Like Zhu et al. (2011), this study showed a strong relationship between Fe/Rb and clay contents, with an Fe coefficient weight of 175.1 and Rb weight of 73.7 for the 10-fold CV MLR model (Eq. 4.6). This finding lends more support for the possibility of a 'unified pXRF clay model' as referenced by Zhu et al. (2011). On the other hand, sand contents showed a negative correlation with Fe in this study, which contrasts with the positive relationship found by Zhu et al. (2011).

Models created in this study may have had improved success in relating soil minerology to textural separates given a more even distribution of samples over the 12 textural classes. A predominance of samples with a sandy texture may have caused the models to be better calibrated to associate sand with elemental concentrations. Additional samples in clay, silt loam, silt, silty clay loam, and silty clay categories could have the effect of improved clay and silt predictions.

5.5 CEC models gave reasonably good estimates

The MLR model to predict CEC developed by Sharma et al. (2015) (Eq. 2.4) could not produce useful estimates for CEC when applied to the California soils dataset (RMSE = 20.9). Estimates were noticeably improved by generating our own coefficients in Eq. 4.7 (RMSE = 11.5) and even more so by creating an MLR model with 10-fold CV to pick coefficients as in Eq. 4.8 (RMSE = 6.88). Random forest modeling further improved model metrics, but only marginally (RMSE = 6.79). Since CEC can vary widely from just a few cmolc/kg in low organic matter sandy soils to up to 100 cmolc/kg in fine textured organic soils, a RMSE of under 7 cmolc/kg as given by the developed models could be helpful for getting quick rough estimates of CEC to decipher spatial variation. While sandy soils can easily be discerned from highly organic soils, soils with a more intergrade composition may benefit from pXRF analysis to detect ballpark approximations of CEC.

Compared to the validation results of Li et al. (2018) (R^2 = 0.60 and RMSE = 8.07) who used RF modeling to predict CEC of compost from pXRF analysis, both our

MLR and RF were able to achieve a higher R^2 and lower RMSE. It's likely that the preprocessing techniques used by Li et al. (2018) obscured the relationship between elemental concentrations and CEC. Authors used recovery percentages of 2711a SRS to apply a correction factor to raw elemental concentrations, rather than using it to check for RSD and stability over time. Our model results for both MLR and RF models achieved better R^2 and RPIQ but worse RMSE values compared to PLSR models predicting CEC from pXRF data as found by Wan et al. (2020) ($R^2 = 0.50$, RMSE = 5.30, and RPIQ = 0.82), suggesting that other machine learning models may work even better for predicting some soil properties.

For CEC prediction from MLR, concentrations of exchangeable cations were expected to be significant coefficients in a model equation. Of the main exchangeable cations, Ca was used in Eq. 2.4 and 4.8, and Mg, K, Na, and Al were strongly significant ($p<0.001$) for Eq. 4.8. Since CEC depends on soil pH, clay content, type of clay, and organic matter, teasing out direct relationships between pure elemental data and CEC becomes less clear, especially with sample sets characterized in different ways. Sharma et al. (2015) created their model using only agricultural soils, and in effect, calibrated their model with high organic matter samples. It can also be assumed that these agricultural soils were at one time amended, which could affect the elemental analysis when compared to unmanaged soils. In addition, the CA soils dataset exhibited a good spread of CEC values, likely due to the contrasting land types represented (from sandy marine terrace environments to forested Mollisols). Where the highest CEC value used in building Eq. 2.4 was <40 cmolc/kg soil, the CA soils dataset contained 40 values between 40-75 cmolc/kg soil.

Another method for indirect predictions of CEC involves the use of pedotransfer functions (PTFs), in which basic known soil properties such as particle size distribution are used to predict unknown soil features which are cumbersome to measure directly. Khodaverdiloo et al., (2018) used PTFs to correlate CEC with clay, OC, and d_g (geometric mean particle diameter) on a training set of Iranian soil samples. The reliability of the developed PTFs was evaluated on an unseen test set and produced accuracies ranging from R^2: 0.48-0.72 depending on the size of the calibration set and the derived PTF. The R^2 values found in this study (MLR: 0.76/RF: 0.79) for CEC prediction are improved from the PTFs estimates found by Khodaverdiloo et al., (2018), which may represent that elemental covariates are better for predicting CEC and can be attained more easily when compared to other laboratory derived soil properties like clay content or organic carbon content.

5.6 Good predictions for N, SOC models show some potential, and C:N models are poor

The MLR model for predicting SOC produced the best RMSE (1.01) of the three SOC models, outperforming both the RF model (RMSE = 1.14) and RF methodology used by Towett et al. (2015) (RMSE = 1.08). However, both RF approaches improved the three other performance metrics when compared to the MLR approach. Towett et al., (2015) was able to achieve an RMSE of 0.7 for predicting SOC from TXRF using an immense dataset of 700 samples. For their RF model, authors chose the hyperparameters of mtry = 50 and ntree = 200 somewhat arbitrarily, and also reported OOB errors from the entire dataset rather than from an unseen hold-out set. This technique of bootstrapping means that as samples were used for model building/validation, they were returned to the data pool to be used again (sampling with replacement). While a valid statistical approach, model performance on an unseen test set may have given a worse prediction that more accurately represents how the model will perform on new data. The method used in this study used 10-fold CV to pick hyperparameters that were shown to reduce OOB error, and then tested model performance on a holdout set. Since SOC content in soils is usually low in typical soils (≤3%) an error >1% is not likely to be helpful in distinguishing SOC levels for management needs or tracking SOC pools across time.

Interestingly, all 12 elements used in the SOC regression model (Eq. 4.9) were highly significant ($p<0.001$), indicating that a number of elemental covariates are important in illustrating the relationship between soil chemistry and SOC. Si had the strongest negative correlation to SOC (coefficient of -22.9), which may be due to the fact that very sandy soils with a high proportion of quartz (SiO_2) tend to have less organic matter than soils with finer textures. Al was also highly negatively correlated to SOC contents (coefficient of -13.7), despite the fact that aluminum bearing minerals are thought to protect and stabilize SOC (Hall and Thompson, 2022). However, the form of Al in the soil (free-metal cations, poorly crystalline, organically complexed phases) and its behavior (sorption to mineral surfaces vs downward leaching) is highly dependent on environmental conditions and solution characteristics (McLean and Bledsoe 1992). For instance, in a study comparing the association between pedogenic Al at four different forest sites, Porras et al., 2017 found site specific factors to be a major component in the relationship between SOC stability and Al content. For instance, authors saw that at low pH levels, organo-metal complexes were less stable and can be negatively correlated with SOC turnover times (Porras et al., 2017). Zr, on the other hand showed a positive correlation (coefficient of 3.1) with SOC. Since Zr is relatively resistant to weathering and its amount tends to increase as weathering progresses (Stockmann et al., 2016), it could be possible that more developed and mature soils may have elevated levels of Zr, in addition to larger stores of SOC which accumulated over time.

Predictions for nitrogen content were quite good, achieving an RMSE of 0.075/0.062 from MLR/RF modeling approaches. One way to conceptualize how these errors compare to typical N totals is through the lens of C:N ratios of SOM. This ratio typically ranges between 8 and 15, with the lower end being more representative of agricultural soils and the higher end being more representative of natural ecosystem soils. With typical ranges of 0.5-5% SOC in California soils, a rough range for TN would be between 0.06-0.33%. The range of TN in this dataset was 0.01-0.86%, representing a good spread of values. Thus, the error achieved by MLR/RF models would probably only be acceptable for soils with high SOM including those in forest and grassland ecosystems, but less useful for low OM soils, including many agricultural soils, which typically have around 0.1% TN (W. Horwath, personal communication, 27 July 2022). No logical relationships were able to be identified between the variables used in Eq. 4.10 and TN content. This could be due to the fact that nitrogen content in is a highly dynamic soil property which changes throughout the season, and samples for this study were collected at different time periods and at different depths in the profile. These confounding factors which may have obscured the relationship between the chemistry of the soils and TN contents.

For three out of the four metrics, our RF model outperformed the SVR model for total nitrogen created by Duda et al. (2017). This result is significant given the importance of N for plant health and productivity. Without enough nitrogen, plants can become stunted, and vigor is dramatically reduced. On the other hand, nitrogen oversupply can increase disease susceptibility and worsen crop quality (Weil and Brady, 2017). Additionally, the production of nitrogen fertilizer contributes a large fraction of the fossil fuels used by agriculture (Woods et al., 2010). Clearly, knowledge about the nitrogen content of soils at any given time is essential for sustainable land management. The option to curtail conventional dry combustion methods to determine TN via pXRF modeling could allow for nitrogen contents to be estimated with a high spatial density. Similar to Eq. 4.9 for SOC, Si and Al were negatively correlated with TN contents, as shown in Eq. 4.11.

C:N ratio was not able to be effectively predicted in this study. This could be a result of the large range of C:N values from 1-167, and 13 outlier values. In addition, samples coming from different characterizing entities had varied conventions for reporting C:N ratio values. For instance, KSSL primary characterization data rounds C:N to a whole number, whereas results from the CN analyzer used for the SPR/LHBC Mollisols, marine terrace, and LA Urban samples provided a more precise number with several decimal places included. Furthermore, CN may not lend itself to adequate predictions from elemental data due to different and uncorrelated proxies for detecting C and N contents individually. In other words, since C and N may have different error frameworks, a combination of these uncorrelated errors magnifies the overall prediction error beyond what would be useful for prediction purposes. Considering that C:N ratio is easily determined from SOC and TN content and provides only a proportion of these elements rather than their actual quantities, future research may benefit from focusing on models to predict SOC and TN directly rather than their ratio.

5.7 Significance of land-type and characterization methods on predictions

To understand how the addition of categorical variables impacted predictions for each property, model predictions on unseen folds using only elemental predictors were compared to model predictions which also included the dummy variables land type and sample set. The inclusion of these categories resulted in an improved R^2 and RMSE for all properties. Since R^2 will always improve with more predictor variables, RMSE is the better metric to reference in this case. Relatively small reductions in RMSE at a large statistical cost (9 extra predictors) revealed that the additional variables may not add enough predictive power to justify their inclusion.

Beyond using land type and sample set as predictor variables, models for each property were also constructed within three land types and within the same methodologies. The interpretation of the results of grouped models was complicated by improvement in some metrics and not others. A model was considered to be improved (beyond the MLR and RF models constructed with all samples) if the majority (3 or more) of the model performance metrics were improved. Where RMSE and R^2 were seen to fluctuate drastically with a change in the seed (ensuring the same split of the data for 10-fold CV and the train/test split), RPD and RPIQ were more stable with varying seeds. This may indicate that these metrics are a better indication of model performance on future datasets, whereas RMSE and R^2 are mostly relevant to the current dataset.

Only the groupings within land type were seen to improve predictions, which may indicate that method of characterization is not a significant factor for elemental predictions and supports the assumption that various lab methods for the same property can be compared. However, consistent characterization across the entire sample set is expected to improve overall model MLR and RF model estimates as a result of less overall error in lab truth data. In other words, inter-lab variability errors would not also be compounded with errors associated with the methods and modeling errors. For land type groupings, estimates of pH were improved within the forest and grassland subsets, sand estimates were improved within the grassland and marine terrace subsets, clay and SOC % were improved within the forest and marine terrace groups, TN % was improved only in the grassland group, and estimates for CEC and C:N were not improved by any grouping.

Given these results, it is difficult to differentiate any hard and fast ‘rules’ for if and how separating samples by land type influences predictions. An improvement in predictive power shown by some of the equations is attributable to the fact that model equations derived from subsets of the total modeling dataset were allowed to have different slopes and only incorporate variables significant to that subset. By including all significant variables in the all-inclusive models, overall error was increased due to the presence of variables that were only significant for certain samples. Considerable improvements in predictions for some properties may suggest that some soil characteristics are more clearly derived from elemental data within a certain type of soil. However, the small sample sizes used to build some of the models (as small as $n = 41$) could be to blame for limited predictive power and higher uncertainty in MLR. These relatively small sub-datasets made RF modeling impractical. However, since RF models were seen to generally improve soil property predictions with the entire dataset, it’s likely that with enough samples to use RF, grouped model estimates would be further improved.

While the forest land type group had samples from multiple sample sets, grassland soils all came from the SPR/LHBC Mollisol dataset and all marine terrace samples came from the marine terrace dataset. Thus, it cannot be said with certainty if improvements for predictions in these categories may be attributed to the soil's origin or simply inter-lab variability from the characterizing party. Additionally, even standard methods will be performed differently by specific labs given the available equipment, financial/time constraints, and technician experience. Therefore, to uncover if differences in land type /categorization are truly important for predictions, assessing the inter-lab variability of models would also be necessary. For instance, if samples within the same land type were all subject to the UN-FAO method for determining CEC, grouping models by characterizing body would help unveil if this was a prominent deciding factor in model efficacy. These results may point to the need for site specific calibration of predictive models which are calibrated using a single lab's methodology for the 'true' values.

Another important consideration in this vein is that while conventional lab analyses may be viewed as the definitive "truth," error is still present in conventional techniques, even when performed with high caliber equipment in reputable laboratories. For instance, when ICP analyses is compared to pXRF analysis, a poor correlation is typically associated with the alternative analysis despite the fact that traditional lab bench equipment also includes errors (Crumbling et al., 2010). Existing studies that use pXRF to make predictive models about soil characteristics attempt to draw correlations between common wet chemistry measures of soil properties and pXRF measurements, but if the measured values deviate too far from the actual values, it becomes difficult to assess the accuracy of pXRF based models.

5.8 California soils dataset

The strength of the predictive models developed from the California soils sample set may have been limited by the consortia of characterizing entities and their respective techniques. This study relied on the assumption that pre-characterized samples had accurate values for the properties of interest. However, even within the same methods, labs may follow different protocols or be limited with their time and financial resources. Existing models for pXRF prediction did not indicate different methods used to characterize any single property, which could explain their lower RMSE values. This difference could be considered a disadvantage due to the fact that more inter-lab uncertainty existed for lab measurements but could also be regarded as a strength of the study, because the developed models may have a wider scope of applicability.

The sample set for this study was 480, which while sizable, could not capture all the diverse ecosystems and soil types in the state of California. Marine terrace, grassland, and forested ecosystems made up the bulk of the samples in this study. To achieve a more robust sample set and improve the models for the uses for which they are most likely destined, more soils from agricultural origins would be beneficial.

5.9 pXRF analysis: areas for improvement

A pXRF scan time of 60 seconds (two 30 second beams) was used to analyze the samples in this study due to precedent from other studies (Sharma et al., 2014) and for ease of sample analysis. For this technology to meet the promise of truly high sample throughput, a rapid one-minute scan time is appealing. However, some elements were close to their LOD for that test time, resulting in a fraction of those elements falling out of their limit of detection. For example, Mg was <LOD for ~19% of scans and P and Nb had <LOD readings for ~14% of scans. This signifies that the 60 second scan time was not long enough (could not capture a high enough number of counts) to cause an atom in these elements to fluoresce and become differentiated from the background noise. Data imputation, a common statistical technique to account for missing data, was performed prior to model building to preserve as many observations as possible. The imputation method used in this study was based on the limit of detection for those scans where a '<LOD' concentration was recorded. The associated 1 sigma error provided by the pXRF output allowed for the LOD to be determined and was then used as the upper bound of a normal distribution curve from which values were imputed. This technique was based off knowing something about the missing datapoint and was preferable over deleting all observations missing any elemental readings. Of course, it would be best to have the concentration as reported by the analyzer, so to ensure most elements are reliably captured with each scan, a longer scan time (90-120 seconds) is recommended for future studies. This would have the effect of lower limits of detection and in turn capture elemental concentrations more consistently. While a two-minute scan is significantly longer than a one-minute scan per sample, the additional counts contribute to a more complete and accurate chemistry result in what is still a very quick timeframe. While some applications may only need a 20 second test time to be fit for purpose, when building predictive models, it is expected that a longer scan time will pay off for the increase in element detection, needed for robust modeling.

Some existing models could not be assessed by the CA soils dataset due to missing elements as a result of using Soil Mode of operation. Where Soil Mode uses Compton normalization, a computationally straightforward method, it's unrealistic to meet the assumptions of dilute samples and no interelement interferences, of which this method relies on. Further, Mg, Al, and Si cannot be detected in this mode, which is a considerable drawback when reflecting on the frequency and significance of these elements in the models developed for this study. It is likewise recommended that future studies use GeoChem mode with a FP calibration to determine many important elemental concentrations making up the total chemistry of a sample that are not detected in Soil Mode.

5.10 Applicability of modeling

Singular models formulated from hundreds of samples across the diverse state of California showed fair performance, which is expected to improve with more targeted calibrations. However, aspirations to use pXRF analysis for consistent and reliable estimates of the soil properties explored in this study throughout an open-ended geographic range seems implausible. To ensure repeatability of results, certain sample preparation, scan times, and preprocessing conventions would need to be followed. Since these steps, as well as model building, are generally beyond the capabilities of typical landowners, a practical application of these models could be use by soil testing laboratories. It is not uncommon for a testing facility to offer indirect, calculated measurements, such as CEC from base cation summation and %OC from SOM%. These laboratories could similarly offer pXRF estimations of soil properties for a much lower cost, because many pXRF samples could be scanned rapidly with minimal sample preparation and no advanced training or traditional reagents/laboratory equipment needed. The deployable use for this technology could then be increasing the number of samples that can be quickly and reasonably characterized, after calibrating for the specific sample range using laboratory data. Basic sample preparation (air-dry, sieved, ground) is recommended to obtain the most accurate chemistry results. For some properties that are more easily determined via direct methods, it's probably most logical to measure them via existing approaches, rather than indirectly with pXRF analysis. For instance, pH is easily measured with the use of just deionized water and an inexpensive pH probe and can be performed in the field.

Despite the advanced capabilities of pXRF devices, the fact that they are more widespread now than ever before and are popular in geological exploration and ore identification purposes, they are still relatively obscure as a tool for soils characterization (apart from heavy metal detection). It might be impractical for every farm to own a pXRF due to the high startup costs of purchasing the analyzer, especially when it could be reasoned that those funds could go directly towards lab measurements. However, for a sizable operation where soil properties are highly dynamic, a high-density number of pXRF characterized samples (even at a lower accuracy) may better serve a land manager's goals than a handful of very precise lab measured samples.

Chapter 6

CONCLUSIONS

In the face of impending climactic shifts amid continued population growth, soils will be tasked with maintaining crucial ecosystem services and supporting intensified crop production. Understanding the characteristics of soils on a spatial and temporal basis is compulsory to ensuring soils are managed productively and to abate the negative consequences of land-use shifts from wildland to constructed environments. Obtaining a baseline picture of soil health is required for tracking changes over time which allows for the benefits or drawbacks of certain management practices to be quantified.

To curtail current time and cost intensive analytical techniques for soil characterization, indirect measurements of soil properties with proximal sensors have been widely explored. This study aimed to assess existing model equations and model building techniques to determine their applicability for California soils. Multiple linear regression models were constructed by including those elements which showed a significant relationship to the target variable, and model performance was tested on a holdout set. This study also aimed to leverage the random forest machine learning algorithm to go beyond data modeling and uncover underlying relationships between soil properties and elemental data.

Both the MLR and RF models formed via 10-fold CV tuning methods were able to achieve an $RMSE<0.5$ for prediction of pH. The high density throughput sampling made possible by pXRF measurements can overcome its analytical uncertainty to give a higher level of confidence than that provided by sparse lab measurements (Lemière, 2018). Thus, pXRF estimations can be more fit for purpose than lab characterization for some applications, such as when the level of accuracy needed for pH measurements is between coarse resolution digital maps (RMSE≤1.3) and very accurate pH meter measurements (RMSE≥0.06) (Libohova et al., 2019). However, since portable pH meters can achieve quicker and more accurate measurements compared to indirect pXRF scanning and model building, where only a few point measurements are needed for pH, predictive models for pH would be extraneous.

As noticed by Zhu et al. (2011) and Wang et al. (2013) Rb and Fe were consistently significant in predicting clay contents. Also similar to Zhu et al. (2011), Zr was consistently significant in sand prediction, which authors attributed to zircon, a mineral resistant to weathering. Patterns indicating the significance of certain elements for sand and clay prediction in conjunction with only moderately improved metrics from RF modeling, indicates that linear modeling may be the best option for predicting sand and clay contents. However, correct texture class prediction was improved from 55% with MLR models to 72% with the use of RF modeling— so, if textural class is more important than specific percent of sand and clay, RF modeling would be the better option.

Despite relatively high RMSE values for CEC prediction from MLR and RF models (6.88 and 6.79, respectively), good RPD (1.986/2.010) and RPIQ (2.079/2.555) values point to stable predictive capacity of the developed models. Considering four different methods were used in laboratory CEC determinations, these predictions are quite good and would be expected to improve further with consistent lab characterization methods.

Surprisingly good estimates for SOC and TN contents were achieved from MLR models, and predictions were further improved with RF modeling. These results are significant in symbolizing the ability to correlate heavier elements with light elements that are extremely important in soil health. Inaptly high RMSE values for SOC % by the modeling techniques used in this study would likely improve with a site-specific calibration to where indirect SOC tracking over time could be possible. By providing more ways to assess SOC in a timely fashion, carbon sequestration driven goals and timelines can be projected with increased accuracy. Poor predictive power of C:N ratio was likely a combined result of many outliers, different rounding protocols, and indirect inference of carbonate presence for some samples.

To further improve predictive power of models, pXRF may be used in tandem with other sensors, which has shown good success for many important soil properties (Swetha and Chakraborty, 2021; Wang et al., 2015; Wan et al., 2019). Combining sensors that can be used estimate the organic fraction of soils (Vis-NIR, color sensors) with pXRF analysis providing information about the inorganic fraction may be necessary to achieve higher accuracy predictions of SOC. Other approaches to linear modeling (PLSR, GLM) and machine learning models (SVM) outside those used in this study have also shown success for predicting various properties. Using the tidyverse package to create and tune models makes it possible to easily change the type of model being run, and thus many different modeling techniques can be used with relative simplicity.

It is recommended that future studies use longer scan times (at least 90 seconds) to achieve a robust modeling dataset, minimize confounding errors by constraining the number of different characterizing methods and bodies, and report results on unseen data to candidly represent how the model performs. Reporting several model evaluation metrics as was done in this study, helps to interpret results and compare models.

This study identified key relationships between elements of interest and soil properties and expressed those relationships both linearly and non-linearly. Model results indicate good prospects for the use of indirect ex-site pXRF estimates where a reasonable level of error is established and accepted. This study used simple sample preparation and scanned samples via pXRF according to the current best practices. Estimates for soil properties based off of in-situ scans would be expected to be much poorer due to interfering factors which are difficult to control for in the field. The challenge of detecting clear relationships between chemical composition and soil features and creating dependably reliable and accurate models for those features points to the incredibly complex and heterogenous nature of soils. Understanding more about the soils underfoot in California and beyond will be key for confronting imminent ecological challenges.

REFERENCES

Acree, A., D.C. Weindorf, L. Paulette, N. van Gestel, S. Chakraborty, et al. 2020. Soil classification in Romanian catenas via advanced proximal sensors. Geoderma 377: 114587. doi: 10.1016/J.geoderma.2020.114587.

Afzal, A., A. Aabid, A. Khan, S. Afghan Khan, U. Rajak, et al. 2020. Response surface analysis, clustering, and random forest regression of pressure in suddenly expanded high-speed aerodynamic flows. Aerosp. Sci. Technol. 107. doi: 10.1016/J.AST.2020.106318.

Aldabaa, A.A.A., D.C. Weindorf, S. Chakraborty, A. Sharma, and B. Li. 2015. Combination of proximal and remote sensing methods for rapid soil salinity quantification. Geoderma 239: 34–46. doi: 10.1016/j.geoderma.2014.09.011.

Allard, D. 2016. WE-H-204-01: William D. Coolidge, Inventor of the Modern X-Ray Tube. Med. Phys. 43(6Part42): 3838–3839. doi: 10.1118/1.4957971.

Anderson, C., Field, C., McCarty, P. L., & Mach, K. 2017) Forests Contribute to California's Climate Change Goals. Stanford Woods Institute for the Environment. Available at http://www.jstor.org/stable/resrep37180

Andrade, R., S.H.G. Silva, D.C. Weindorf, S. Chakraborty, W.M. Faria, et al. 2020. Assessing models for prediction of some soil chemical properties from portable X- ray fluorescence (pXRF) spectrometry data in Brazilian Coastal Plains. Geoderma 357: 113957. doi: 10.1016/J.GEODERMA.2019.113957.

Appel, C., and C. Stubler. 2018. SS 423 Soil and Water Chemistry Laboratory Manual. 11th ed. California State Polytechnic University of San Luis Obispo.

Auffhammer, Maximilan. 2014. Estimating Impacts of Climate Change on California's Most Important Crops. ARE Update 18(1): 16-19. University of California Giannini Foundation of Agricultural Economics. Available at https://giannini.ucop.edu/filer/file/1453327771/16951/

Beckhoff, B., B. Kanngießer, N. Langhoff, R. Wedell, and H. Wolff, editors. 2006. Handbook of Practical X-Ray Fluorescence Analysis. Springer, Berlin.

Bedsworth, Louise, Dan Cayan, Guido Franco, Leah Fisher, Sonya Ziaja. 2018. Statewide Summary Report. California's Fourth Climate Change Assessment. California Governor's Office of Planning and Research, Scripps Institution of Oceanography, California Energy Commission, California Public Utilities Commission. Publication number: SUMCCCA4-2018-013.

Bellon-Maurel, V., E. Fernandez-Ahumada, B. Palagos, J.M. Roger, and A. McBratney. 2010. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. TrAC - Trends Anal. Chem. 29(9): 1073–1081. doi: 10.1016/J.TRAC.2010.05.006.

Benedet, L., W.M. Faria, S.H.G. Silva, M. Mancini, J.A.M. Demattê, et al. 2020a. Soil texture prediction using portable X-ray fluorescence spectrometry and visible near- infrared diffuse reflectance spectroscopy. Geoderma 376. doi: 10.1016/j.geoderma.2020.114553.

Benedet, L., W.M. Faria, S.H.G. Silva, M. Mancini, L.R.G. Guilherme, et al. 2020b. Soil subgroup prediction via portable X-ray fluorescence and visible near-infrared spectroscopy. Geoderma 365: 114212. doi: 10.1016/j.geoderma.2020.114212.

Bos, J.F.F.P., H.F.M. ten Berge, J. Verhagen, and M.K. van Ittersum. 2017. Trade-offs in soil fertility management on arable farms. Agric. Syst. 157: 292–302. doi: 10.1016/J.AGSY.2016.09.013.

Bosco, G.L. 2013. Development and application of portable, hand-held X-ray fluorescence spectrometers. TrAC - Trends in Analytical Chemistry. Elsevier B.V. p. 121–134

Breiman, L. 2001. Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). https://doi.org/10.1214/ss/1009213726 16(3): 199–231. doi: 10.1214/SS/1009213726.

Brouwer, P. 2010. Theory of XRF - Getting acquainted with the principles. 3rd ed. PANalytical, Almelo.

Brus, D.J., B. Kempen, and G.B.M. Heuvelink. 2011. Sampling for validation of digital soil maps. Eur. J. Soil Sci. 62(3): 394–407. doi: 10.1111/J.1365-2389.2011.01364.X.

Calibration standards. The XRF standards of the iron and Steel Institute of Japan. 1982. X-Ray Spectrom. 11(1): 36–39. doi: 10.1002/XRS.1300110111.

California Department of Finance, Demographic Research Unit. 2021a. Report P-1A: Total Population Projections, California, 2010-2060 (Baseline 2019 Population Projections; Vintage 2020 Release). Sacramento, California.

California Department of Finance. 2021b. State Population Dips to 39.5 Million per New State Demographic Report. Available at https://www.dof.ca.gov/Forecasting/Demographics/Estimates/e-1/documents/E-1_2021PressRelease.pdf

Callesen, I., H. Keck, and T.J. Andersen. 2018. Particle size distribution in soils and marine sediments by laser diffraction using Malvern Mastersizer 2000—method uncertainty including the effect of hydrogen peroxide pretreatment. J. Soils Sediments 18(7): 2500–2510. doi: 10.1007/S11368-018-1965-8/FIGURES/3.

Carr, R., A.E. Chaosheng, Z. Ae, N.M. Ae, and M. Harder. 2008. Identification and mapping of heavy metal pollution in soils of a sports ground in Galway City, Ireland, using a portable XRF analyser and GIS. Environ. Geochem. Health 30:45–52. doi: 10.1007/s10653-007-9106-0.

Castaldi, F., F. Pelosi, S. Pascucci, and R. Casa. 2017. Assessing the potential of images from unmanned aerial vehicles (UAV) to support herbicide patch spraying in maize. Precis. Agric. 18(1): 76–94. doi: 10.1007/s11119-016-9468-3.

[CDFA] California Department of Food and Agriculture. 2019. California Agricultural Statistics Review 2019-2020.

Chai, S., Z. Wang, B. Zhang, L. Cui, and R. Chai. 2020. Wireless Sensor Networks. Springer Singapore..

Chang, C.-W., D.A. Laird, M.J. Mausbach, and C.R. Hurburgh. 2001. Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. Soil Sci. Soc. Am. J. 65(2): 480–490. doi: 10.2136/SSSAJ2001.652480X.

Chen, Z., Y. Xu, G. Lei, Y. Liu, J. Liu, et al. 2021. A general framework and practical procedure for improving pxrf measurement accuracy with integrating moisture content and organic matter content parameters. Sci. Reports 2021 111 11(1): 1–10. doi: 10.1038/s41598-021-85045-4.

Clark, B. 2021. Characteristics and management implications of mollic soils in forest versus grassland settings in central California. Master's thesis, California Polytechnic State University, San Luis Obispo, CA.

[CRS] Congressional Research Service. 2021. Wildfire Statistics. IF10244. Available at https://sgp.fas.org/crs/misc/IF10244.pdf

Crumbling, D., S. Dyment, and B. Johnson. 2008. X-ray Fluorescence (XRF) Seminar Series. Available at https://clu-in.org/conf/tio/xrf_080408/

D'Amore, D. and E. Kane. 2016. Climate Change and Forest Soil Carbon. U.S. Department of Agriculture, Forest Service, Climate Change Resource Center. Available at www.fs.usda.gov/ccrc/topics/forest-soil-carbon

Declercq, Y., N. Delbecque, J. De Grave, P. De Smedt, P. Finke, et al. 2019. A comprehensive study of three different portable XRF scanners to assess the soil geochemistry of an extensive sample dataset. Remote Sens. 11(21). doi: 10.3390/RS11212490.

Deryng, D (editor). 2020. Climate change and agriculture. 1st ed. Burleigh Dodds Science Publishing, London.

Duda, B.M., D.C. Weindorf, S. Chakraborty, B. Li, T. Man, et al. 2017. Soil characterization across catenas via advanced proximal sensors. Geoderma 298: 78–91. doi: 10.1016/J.GEODERMA.2017.03.01

FAO and UNEP. 2020. The State of the World's Forests 2020. Forests, biodiversity and people. Rome.

Ferro, V., and S. Mirabile. 2009. Comparing Particle Size Distribution Analysis by Sedimentation and Laser Diffraction Method. J. Agric. Eng. 40(2): 35. doi: 10.4081/JAE.2009.2.35.

Figueroa, M., and C. Pope. 2017. Root System Water Consumption Pattern Identification on Time Series Data. Sensors 17(6): 1410. doi: 10.3390/s17061410.

Fischlin, A., G.F. Midgley, J.T. Price, R. Leemans, B. Gopal, C. Turley, M.D.A. Rounsevell, O.P. Dube, J. Tarazona, A.A. Velichko. 2007. Ecosystems, their properties, goods, and services. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson, Eds., Cambridge University Press, Cambridge, 211-272.

Food and Agriculture Organization of the United Nations. 2009. Feeding the world in 2050. Available at https://www.fao.org/fileadmin/templates/wsfs/Summit/WSFS_Issues_papers/WSFS_feeding_E.pdf

Francaviglia, R., L. Ledda, and R. Farina. 2018. Organic Carbon and Ecosystem Services in Agricultural Soils of the Mediterranean Basin. : 183–210. doi: 10.1007/978-3- 319-90309-5_6.

Friedlander, Gerhat; Kennedy, Joseph W.; Macias, Edward S.; Miller, Julian Malcolm. 1981. Nuclear and Radiochemistry, 3rd Ed. Chapter 9. New York: Wiley.

Garratt, M.P.D., R. Bommarco, D. Kleijn, E. Martin, S.R. Mortimer, et al. 2018. Enhancing Soil Organic Matter as a Route to the Ecological Intensification of European Arable Systems. Ecosystems 21(7): 1404–1415. doi: 10.1007/S10021-018-0228-2/FIGURES/3.

Gauquelin, T., G. Michon, R. Joffre, R. Duponnois, D. Génin, et al. 2018. Mediterranean forests, land use and climate change: a social-ecological perspective. Reg. Environ. Chang. 18(3): 623–636. doi: 10.1007/S10113-016-0994-3/TABLES/1.

Gavlak, R., D. Horneck, and R. Miller. 2005. Plant, soil and water reference methods for the Western Region. Western Regional Extension Publication (WREP) 125, WERA-103 Technical Committee.

Ge, L., W. Lai, and Y. Lin. 2005. Influence of and correction for moisture in rocks, soils and sediments on in situ XRF analysis. X-Ray Spectrom. 34(1): 28–34. doi: 10.1002/xrs.782.

Gea-Izquierdo, G., S. Gennet, and J.W. Bartolome. 2007. Assessing plant-nutrient relationships in highly invaded Californian grasslands using non-normal probability distributions. Appl. Veg. Sci. 10(3): 343–350. doi: 10.1111/J.1654-109X.2007.TB00433.X.

Gelman, A., and J. Hill. 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press (CUP).

Glanzman, R.K., and L.G. Closs. Field Portable X-Ray Fluorescence Geochemical Analysis-Its Contribution to Onsite Real-time Project Evaluation.

Goodale, N., D.G. Bailey, G.T. Jones, C. Prescott, E. Scholz, et al. 2012. pXRF: a study of inter-instrument performance. J. Archaeol. Sci. 39(4): 875–883. doi: 10.1016/J.JAS.2011.10.014.

Gozukara, G., M. Acar, E. Ozlu, O. Dengiz, A.E. Hartemink, et al. 2022. A soil quality index using Vis-NIR and pXRF spectra of a soil profile. Catena 211: 105954. doi: 10.1016/J.CATENA.2021.105954.

Gozukara, G., Y. Zhang, and A.E. Hartemink. 2021. Using vis-NIR and pXRF data to distinguish soil parent materials – An example using 136 pedons from Wisconsin, USA. Geoderma 396: 115091. doi: 10.1016/J.GEODERMA.2021.115091.

Grimmond, S. 2007. Urbanization and global environmental change: local effects of urban warming. Geogr. J. 173(1): 83–88. doi: 10.1111/J.1475-4959.2007.232_3.X.

Guimarães, D., T.M. Cleaver, S.F. Martin, and P.J. Parsons. 2015. Radioisotope-based XRF instrumentation for determination of lead in paint: An assessment of the current accuracy and reliability of portable analyzers used in New York State. Anal. Methods 7(1): 366–374. doi: 10.1039/c4ay00819g.

Hall, S.J., and A. Thompson. 2022. What do relationships between extractable metals and soil organic carbon concentrations mean? Soil Sci. Soc. Am. J. 86(2): 195–208. doi: 10.1002/SAJ2.20343.

Harrell, F.E. 2015. Regression Modeling Strategies. 2nd ed. Springer, London.

Horta, A., B. Malone, U. Stockmann, B. Minasny, T.F.A. Bishop, et al. 2015. Potential of integrated field spectroscopy and spatial analysis for enhanced assessment of soil contamination: A prospective review. Geoderma 241–242: 180–209. doi: 10.1016/j.geoderma.2014.11.024.

House, JI., Brovkin, V., Betts, R., Constanza, R., Silva Dias, MA., Holland, E., le Que're, C., Kim Phat, N., Riebesell, U., & Scholes, M. 2006. Climate and Air Quality. In R. Hassan, R. Scholes, & N. Ash (Eds.), Ecosystems and Human Wellbeing: Current State and Trends: Findings of the Conditions and Trend Working Group of the Millennium Ecosystem Assessment (918pp ed., Vol. 1, pp. 355 - 390). Island Press, Washington DC.

Huntsinger, L., and S. Barry. 2021. Grazing in California's Mediterranean Multi-Firescapes. Front. Sustain. Food Syst. 5: 291. doi: 10.3389/FSUFS.2021.715366/BIBTEX.

IPCC. 2018. Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [V. Masson-Delmotte, P. Zhai, H. O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J. B. R. Matthews, Y. Chen, X. Zhou, M. I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, T. Waterfield (eds.)]. In Press.

IPCC. 2021. Climate Change 2021. The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou (eds.)]. Cambridge University Press. In Press.

Jackson, L., Van R. Haden, Stephen M. Wheeler, Allan D. Hollander, Josh Perlman, Toby O'Geen, Vishal K. Mehta, Victoria Clark, John Williams, and Ann Thrupp (University of California, Davis). 2012. Vulnerability and Adaptation to Climate Change in California Agriculture. California Energy Commission. Publication number: CEC-500- 2012-031.

Jaremko, D., and D. Kalembasa. 2014. A comparison of methods for the determination of cation exchange capacity of soils. Ecol. Chem. Eng. S 21(3): 487–498. doi: 10.2478/ECES-2014-0036.

Jia, G., E. Shevliakova, P. Artaxo, N. De Noblet-Ducoudré, R. Houghton, et al. 2019. Land-climate interactions: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems.

Kabir, M.H. 2013. Particle Induced X-ray Emission (PIXE): A Tool of Qualitative Elemental Analysis for Biological Sample. Rajshahi Univ. J. Sci. 39: 1–10. doi: 10.3329/RUJS.V39I0.16538.

Kalnicky, D.J., and R. Singhvi. 2001. Field portable XRF analysis of environmental samples. J. Hazard. Mater. 83(1–2): 93–122. doi: 10.1016/S0304-3894(00)00330-7.

Keng, N. 2015a. EDXRF: How Does It Work? Thermo Fish. Sci. Available at https://www.thermofisher.com/blog/metals/edxrf-how-does-it-work/

Keng, N. 2015b. WDXRF: How Does It Work? Thermo Fish. Sci. Available at https://www.thermofisher.com/blog/metals/wdxrf-how-does-it-work/

Kerr, A., J. Dialesandro, K. Steenwerth, N. Lopez-Brody, and E. Elias. 2018. Vulnerability of California specialty crops to projected mid-century temperature changes. Clim. Change 148(3): 419–436. doi: 10.1007/S10584-017-2011-3/FIGURES/4.

Khan, R., I. Ali, M. Zakarya, M. Ahmad, M. Imran, et al. 2018. Technology-Assisted Decision Support System for Efficient Water Utilization: A Real-Time Testbed for Irrigation Using Wireless Sensor Networks. IEEE Access 6: 25686–25697. doi: 10.1109/ACCESS.2018.2836185.

Kilbride, C., J. Poole, and T.R. Hutchings. 2006. A comparison of Cu, Pb, As, Cd, Zn, Fe, Ni and Mn determined by acid extraction/ICP–OES and ex situ field portable X- ray fluorescence analyses. Environ. Pollut. 143(1): 16–23. doi: 10.1016/J.ENVPOL.2005.11.013.

Klausmeyer, K. R., M. R. Shaw, J. B. MacKenzie, and D. R. Cameron. 2011. Landscape-scale indicators of biodiversity's vulnerability to climate change. Ecosphere 2(8):art88. doi:10.1890/ES11-00044.1

Kome, G.K., R.K. Enang, B.P.K. Yerima, and M.G.R. Lontsi. 2018. Models relating soil pH measurements in H_2O, KCl and $CaCl_2$, for volcanic ash soils of Cameroon. Geoderma Reg. 14. doi: 10.1016/j.geodrs.2018.e00185.

Kramar, U. 2017. X-ray fluorescence spectrometers. Encyclopedia of Spectroscopy and Spectrometry. Elsevier. p. 695–706

Lal, R. 2016. Soil health and carbon management. Food Energy Secur. 5(4): 212–222. doi: 10.1002/FES3.96.

Laperche, V., and B. Lemière. 2020. Possible Pitfalls in the Analysis of Minerals and Loose Materials by Portable XRF, and How to Overcome Them. Miner. 2021, Vol. 11, Page 33 11(1): 33. doi: 10.3390/MIN11010033.

Lemière, B. 2018. A review of pXRF (field portable X-ray fluorescence) applications for applied geochemistry. J. Geochemical Explor. 188: 350–363. doi: 10.1016/j.gexplo.2018.02.006.

Li, B., S. Chakraborty, M.F.G. Sosa, N.Y.O. Kusi, and D.C. Weindorf. 2018. Compost Cation Exchange Capacity via Portable X-Ray Fluorescence (PXRF) Spectrometry. https://doi.org/10.1080/1065657X.2018.1522280 26(4): 271–278. doi: 10.1080/1065657X.2018.1522280.

Li, S., V. Dao, M. Kumar, P. Nguyen, and T. Banerjee. 2022. Mapping the wildland-urban interface in California using remote sensing data. Sci. Reports 2022 121 12(1): 1–12. doi: 10.1038/s41598-022-09707-7.

Libohova, Z., C. Seybold, K. Adhikari, S. Wills, D. Beaudette, et al. 2019. The anatomy of uncertainty for soil pH measurements and predictions: Implications for modellers and practitioners. Eur. J. Soil Sci. 70(1): 185–199. doi: 10.1111/EJSS.12770.

Liu, Y., C. Wang, C. Xiao, K. Shang, Y. Zhang, et al. 2021. Prediction of multiple soil fertility parameters using VisNIR spectroscopy and PXRF spectrometry. Soil Sci. Soc. Am. J. 85(3): 591–605. doi: 10.1002/SAJ2.20223.

Lobell, D.B., C.B. Field, K.N. Cahill, and C. Bonfils. 2006. Impacts of future climate change on California perennial crop yields: Model projections with climate and crop uncertainties. Agric. For. Meteorol. 141(2–4): 208–218. doi: 10.1016/J.AGRFORMET.2006.10.006.

Longoni, A., and C. Fiorini. X-Ray Detectors and Signal Processing. In: Beckhoff, B., Kanngießer, B., Langhoff, N., Wedell, R., and Wolff, H., editors, Handbook of Practical X-Ray Fluorescence Analysis. Springer-Verlag Berlin Heidelberg. p. 203– 262

Lucas-Tooth, J., and C. Pyne. 1963. The Accurate Determination of Major Constituents by X-Ray Fluorescent Analysis in the Presence of Large Interelement Effects. Adv. X-ray Anal. 7: 523–541. doi: 10.1154/S0376030800002780.

Magdoff, F., and H. van Es. 2021. Amount of Organic Matter in Soils. Building Soils for Better Crops. Sustainable Agriculture Research and Education. Available at https://www.sare.org/publications/building-soils-for-better-crops/amount-of-organic-matter-in-soils

Malley, D.F., P.D. Martin, and E. Ben-Dor. 2004. Application in Analysis of Soils. In: Roberts, C.A., Workman, J.J., and Reeves III, J.B., editors, Near-infrared spectroscopy in agriculture. American Society of Agronomy, Inc. Crop Science Society of America, Inc. Soil Science Society of America, Inc., Madison, WI. p. 729–784

Mantler, M. 2006. Monte Carlo Methods. Handbook of Practical X-Ray Fluorescence Analysis. 1st ed. Springer Berlin Heidelberg. p. 394–399

Markowicz, A. 2011. An overview of quantification methods in energy-dispersive X-ray fluorescence analysis.

McLean, J. E. and B. E. Bledsoe. 1992. Behavior of Metal in Soils (EPA/540/S-92/018). U.S. Environmental Protection Agency, Washington, D.C., 1992.

Medellín-Azuara, J., Sumner, D. A., Yolanda Pan, Q., Lee, H., Espinoza, V., Bell, A., Lund, J. R. (2018). Economic and Environmental Implications of California Crop and Livestock, Adaptation to Climate Change. California's Fourth Climate Change Assessment, California Natural Resources Agency. Publication number: CCCA4- CNRA-2018-018.

Miller, R.O., R. Gavlak, and D. Horneck. 2013. Soil, Plant, and Water Reference Methods for the Western Region, 4th Ed.

Mölders, N. 2012. Land-use and land-cover change: Impact on climate and air quality. Springer, Fairbanks, AK.

National Cooperative Soil Survey. National Cooperative Soil Survey Soil Characterization Database. Available at http://ncsslabdatasmart.sc.egov.usda.gov/

[NIOSH] National Institute for Occupational Safety and Health. 1998. NIOSH Manual of Analytical Methods, Method No. 7702: Lead by field portable XRF, Fourth Edition, Cincinnati, OH. Available at https://www.cdc.gov/niosh/docs/2003-154/pdfs/7702.pdf

Nichols, D.P. 2018. RPubs - Random Imputation. Available at https://rpubs.com/nichols16/411765

Nummi, E. 2015. Going to the Source: X-ray Tubes. Available at https://www.thermofisher.com/blog/mining/going-to-the-source-x-ray-tubes/

[OEHHA] Office of Environmental Health Hazard Assessment, California Environmental Protection Agency. 2018. Indicators of Climate Change in California. Available at https://oehha.ca.gov/media/downloads/climate-change/report/2018caindicatorsreportmay2018.pdf

O'Rourke, S.M., U. Stockmann, N.M. Holden, A.B. McBratney, and B. Minasny. 2016. An assessment of model averaging to improve predictive power of portable vis-NIR and XRF for the determination of agronomic soil properties. Geoderma 279: 31–44. doi: 10.1016/J.GEODERMA.2016.05.005.

Panda, A. 2018. Transformational adaptation of agricultural systems to climate change. Wiley Interdisciplinary Reviews: Climate Change, 9(4), e520. Available at https://doi.org/10.1002/WCC.520

Pitt, J.L., F.M. Provin, F.D. Hons, and Waskom J.S. 2003. Use of a Total Carbon/Nitrogen Analyzer for the Determination of Organic and Inorganic Carbon in Soil, Manure, and Composts. (S08-pitt925747-poster).

Porras, R.C., C.E. Hicks Pries, K.J. McFarlane, P.J. Hanson, and M.S. Torn. 2017. Association with pedogenic iron and aluminum: effects on soil organic carbon storage and stability in four temperate forest soils. Biogeochemistry 133(3): 333–345. doi: 10.1007/S10533-017-0337-6/TABLES/5.

Potts, P.J. 1999. Portable X-ray Fluorescence Analysis XA9953280. Vienna.

Potts, P.J., and M. West, editors. 2008. Portable X-ray Fluorescence Spectrometry Capabilities for In Situ Analysis 19000. The Royal Society of Chemistry, Cambridge.

Potts, P.J., O. Williams-Thorpe, and P.C. Webb. 1997. The bulk analysis of silicate rocks by portable X-ray fluorescence: Effect of sample mineralogy in relation to the size of the excited volume. Geostand. Newsl. 21(1): 29–41. doi: 10.1111/J.1751- 908X.1997.TB00529.X.

Radu, T., and D. Diamond. 2009. Comparison of soil pollution concentrations determined using AAS and portable XRF techniques. J. Hazard. Mater. 171(1–3): 1168–1171. doi: 10.1016/j.jhazmat.2009.06.062.

Ravansari, R., and L.D. Lemke. 2018. Portable X-ray fluorescence trace metal measurement in organic rich soils: pXRF response as a function of organic matter fraction. Geoderma 319: 175–184. doi: 10.1016/J.GEODERMA.2018.01.011.

Ravansari, R., S.C. Wilson, and M. Tighe. 2020. Portable X-ray fluorescence for environmental assessment of soils: Not just a point and shoot method. Environ. Int. 134: 105250. doi: 10.1016/j.envint.2019.105250.

Rawal, A., S. Chakraborty, B. Li, K. Lewis, M. Godoy, et al. 2019. Determination of base saturation percentage in agricultural soils via portable X-ray fluorescence spectrometer. Geoderma 338: 375–382. doi: 10.1016/J.GEODERMA.2018.12.032.

Ray, D.K., N.D. Mueller, P.C. West, and J.A. Foley. 2013. Yield Trends Are Insufficient to Double Global Crop Production by 2050. PLoS One 8(6): e66428. doi: 10.1371/JOURNAL.PONE.0066428.

Root, T. L., Hall, K. R., Herzog, M. P., & Howell, C. A. (Eds.). 2015. Biodiversity in a changing climate : Linking science and management in conservation. University of California Press.

Rouillon, M., M.P. Taylor, and C. Dong. 2017. Reducing risk and increasing confidence of decision making at a lower cost: In-situ pXRF assessment of metal-contaminated sites. Environ. Pollut. 229: 780–789. doi: 10.1016/J.ENVPOL.2017.06.020.

Rumpel, C., F. Amiraslani, L.-S. Koutika, P. Smith, D. Whitehead, et al. 2018. Put more carbon in soils to meet Paris climate pledges. Nature (564): 32–34. doi: https://doi.org/10.1038/d41586-018-07587-4.

Ruxton, B.P. 2015. Measures of the Degree of Chemical Weathering of Rocks. https://doi.org/10.1086/627357 76(5): 518–527. doi: 10.1086/627357.

Salley, S.W., J.E. Herrick, C. V. Holmes, J.W. Karl, M.R. Levi, et al. 2018. A Comparison of Soil Texture-by-Feel Estimates: Implications for the Citizen Soil Scientist. Soil Sci. Soc. Am. J. 82(6): 1526–1537. doi: 10.2136/SSSAJ2018.04.0137.

Sarala, P. 2016. Comparison of different portable XRF methods for determining till geochemistry. Geochemistry Explor. Environ. Anal. 16(3–4): 181–192. doi: 10.1144/GEOCHEM2012-162.

Sarala, P., A. Taivalkoski, J.V. Sarala, and P. Taivalkoski. 2015. Portable XRF: An Advanced on-site Analysis Method in Till Geochemical Exploration. Novel technologies for greenfield exploration Edited by P., Sarala. Geol. Surv. Finland, Spec. Pap. 57: 63–86.

Sauer, D., G. Schellmann, and K. Stahr. 2007. A soil chronosequence in the semi-arid environment of Patagonia (Argentina). CATENA 71(3): 382–393. doi: 10.1016/J.CATENA.2007.03.010.

Schlesinger, W.H., and E.S. Bernhardt. 2013. The Lithosphere. Biogeochemistry. Elsevier. p. 93–133

Scholze, F. et al. 2006. X-Ray Detectors and XRF Detection Channels. In: Beckhoff, B., Kanngießer, h.B., Langhoff, N., Wedell, R., Wolff, H. (eds) Handbook of Practical X-Ray Fluorescence Analysis. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-36722-2_4

Shand, C.A., and R. Wendler. 2014. Portable X-ray fluorescence analysis of mineral and organic soils and the influence of organic matter. J. Geochemical Explor. 143: 31–42. doi: 10.1016/J.GEXPLO.2014.03.005.

Sharma, A., D.C. Weindorf, T. Man, A.A.A. Aldabaa, and S. Chakraborty. 2014. Characterizing soils via portable X-ray fluorescence spectrometer: 3. Soil reaction (pH). Geoderma 232–234: 141–147. doi: 10.1016/j.geoderma.2014.05.005.

Sharma, A., D.C. Weindorf, D.D. Wang, and S. Chakraborty. 2015. Characterizing soils via portable X-ray fluorescence spectrometer: 4. Cation exchange capacity (CEC). Geoderma 239: 130–134. doi: 10.1016/j.geoderma.2014.10.001.

Shahzad, A., Ullah, S., Dar, A.A. et al. 2021. Nexus on climate change: agriculture and possible solution to cope future climate change stresses. Environ Sci Pollut Res 28, 14211–14232. https://doi-org.ezproxy.lib.calpoly.edu/10.1007/ s11356-021-12649-8

Shefsky, S., 1997. Comparing field portable X-ray fluorescence (XRF) to laboratory analysis of heavy metals in soil. In: International Symposium of Field Screening Methods for Hazardous Wastes and Toxic Chemicals, Las Vegas, Nevada, USA, January 29–31, 1997.

Shuttleworth, E.L., M.G. Evans, S.M. Hutchinson, and J.J. Rothwell. 2014. Assessment of lead contamination in peatlands using field portable XRF. Water. Air. Soil Pollut. 225(2). doi: 10.1007/s11270-013-1844-2.

Silva, F.M., D.C. Weindorf, S.H.G. Silva, E.A. Silva, B.T. Ribeiro, et al. 2019. Tropical Soil Toposequence Characterization via pXRF Spectrometry. Soil Sci. Soc. Am. J. 83(4): 1153–1166. doi: 10.2136/SSSAJ2018.12.0498.

Silva, S.H.G., D.C. Weindorf, L.C. Pinto, W.M. Faria, F.W. Acerbi Junior, et al. 2020. Soil texture prediction in tropical soils: A portable X-ray fluorescence spectrometry approach. Geoderma 362: 114136. doi: 10.1016/J.GEODERMA.2019.114136.

Singh, S., S. Ambegaokar, K.S. Champawat, A. Gupta, and S. Sharma. 2015. Time series analysis of clustering high dimensional data in precision agriculture. ICIIECS 2015 - 2015 IEEE International Conference on Innovations in Information, Embedded and Communication Systems. Institute of Electrical and Electronics Engineers Inc.

Smith, G. 2018. Step away from stepwise. J. Big Data 5(1): 1–12. doi: 10.1186/S40537-018-0143-6/FIGURES/1.

Snyder, C. 2014. Soil pH Management. Efficient Fertilizer Use Manual. Available at https://dokumen.tips/documents/soil-ph-management-by-dr-cliff-snyder-amendpdf-efficient-fertilizer-use.html?page=1

Soil Survey Staff. 2004. Soil Survey Laboratory Methods Manual, Soil Survey Investigations Report No. 42, Version 4.0.

Soil Survey Staff. 2011. Soil Survey Laboratory Information Manual. Soil Survey Investigations Report No. 45, Version 2.0. R. Burt (ed.). U.S. Department of Agriculture, Natural Resources Conservation Service.

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. 2014a. Keys to Soil Taxonomy. 12th ed. Natural Resource Conservation Service. U.S. Department of Agriculture.

Soil Survey Staff. 2014b. Kellogg soil survey laboratory methods manual – soil survey investigations. Report No. 42. Version 5.0. Natural Resource Conservation Service, National Soil Survey Center. Lincoln, NE

Sparks, D.L. 2003. Ion Exchange Processes. Environ. Soil Chem.: 187–205. doi: 10.1016/B978-012656446-4/50006-2.

Srbinovska, M., C. Gavrovski, V. Dimcev, A. Krkoleva, and V. Borozan. 2015. Environmental parameters monitoring in precision agriculture using wireless sensor networks. J. Clean. Prod. 88: 297–307. doi: 10.1016/J.JCLEPRO.2014.04.036.

Steiner, A.E., R.M. Conrey, and J.A. Wolff. 2017. PXRF calibrations for volcanic rocks and the application of in-field analysis to the geosciences. Chem. Geol. 453: 35–54. doi: 10.1016/j.chemgeo.2017.01.023.

Stockmann, U., S.R. Cattle, B. Minasny, and A.B. McBratney. 2016. Utilizing portable X-ray fluorescence spectrometry for in-field investigation of pedogenesis. Catena 139: 220–231. doi: 10.1016/j.catena.2016.01.007.

Sumner, M.E., and W.P. Miller. 1996. Cation exchange capacity and exchange coefficients. p. 1201–1229. In D.L. Sparks (ed.) Methods of soil analysis. Part 3. Chemical methods. No. 5. ASA and SSSA, Madison, WI.

Svirejeva-Hopkins, A., H.J. Schellnhuber, and V.L. Pomaz. 2004. Urbanised territories as a specific component of the Global Carbon Cycle. Ecol. Modell. 173(2–3): 295–312. doi: 10.1016/J.ECOLMODEL.2003.09.022.

Swetha, R.K., and S. Chakraborty. 2021. Combination of soil texture with Nix color sensor can improve soil organic carbon prediction. Geoderma 382: 114775. doi: 10.1016/J.GEODERMA.2020.114775.

Tilman, D., C. Balzer, J. Hill, and B.L. Befort. 2011. Global food demand and the sustainable intensification of agriculture. Proc. Natl. Acad. Sci. U. S. A. 108(50): 20260–20264. doi: 10.1073/PNAS.1116437108/-/DCSUPPLEMENTAL.

Tomich, T.P., S.B. Brodt, R.A. Dahlgren, and K.M. Scow, editors. 2016. The California nitrogen assessment : challenges and solutions for people, agriculture, and the environment. 1st ed. University of California Press, Oakland.

Towett, E.K., K.D. Shepherd, A. Sila, E. Aynekulu, and G. Cadisch. 2015. Mid-Infrared and Total X-Ray Fluorescence Spectroscopy Complementarity for Assessment of Soil Properties. Soil Sci. Soc. Am. J. 79(5): 1375–1385. doi: 10.2136/sssaj2014.11.0458.

Trivedi, P., B.P. Singh, and B.K. Singh. 2018. Soil Carbon: Introduction, Importance, Status, Threat, and Mitigation. In: Singh, B.K., editor, Soil Carbon Storage: Modulators, Mechanisms and Modeling. Academic Press. p. 1–28

Trumbore, S. 2009. Radiocarbon and soil carbon dynamics. Annu. Rev. Earth Planet. Sci. 37: 47–66. doi: 10.1146/annurev.earth.36.031207.124300.

UC Agricultural Issues Center. 2009. The Measure of California Agriculture, 2000. https://aic.ucdavis.edu/publications/moca-the-measure-of-california-agriculture/the- measure-of-california-agriculture-2000/

Ultraviolet/Visible/Near Infrared Spectroscopy (UV/VIS/NIR). 2022. Eurofins Scientific. https://www.eag.com/techniques/spectroscopy/uv-vis-spectroscopy/

[USDA-NRCS] U.S. Department of Agriculture, Natural Resources Conservation Service. Heavy Metal Soil Contamination. 2000. Auburn, AL. https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcs142p2_053279.pdf

U.S. Environmental Protection Agency, 1996a. Method 3050B: Acid Digestion of Sediments, Sludges, and Soils. In SW-846 Pt 1; Office 404 of Solid and Hazardous Wastes, USEPA, Cincinnati, OH. Available at http://www.epa.gov/osw/hazard/testmethods/sw846/pdfs/3050b.pdf

U.S. Environmental Protection Agency, 1996b. Method 3052: Microwave Assisted Acid Digestion of Siliceous and Organically Based Matrices. In: Test Methods For Evaluating Solid Waste, USEPA, Washington, DC. Available at http://www.epa.gov/wastes/hazard/testmethods/sw846/pdfs/3052.pdf

U.S. Environmental Protection Agency [USEPA] 2007. EPA Method 6200 Field Portable X-Ray Fluorescence Spectrometry for the Determination of Elemental Concentrations in Soil and Sediment. Available at https://www.epa.gov/hw-sw846/sw-846-test-method-6200-field-portable-x-ray-fluorescence-spectrometry- determination

Vahmani, P., F. Sun, A. Hall, and G. Ban-Weiss. 2016. Investigating the climate impacts of urbanization and the potential for cool roofs to counter future climate change in Southern California. Environ. Res. Lett. 11(12): 124027. doi: 10.1088/1748-9326/11/12/124027.

Van Bodegom, Arend Jan, Herman Savenije and Marieke Wit (eds). 2009. Forests and Climate Change: adaptation and mitigation. Tropenbos International, Wageningen, The Netherlands.

Verheye, W.H., and T.G. Boyadgiev. 1997. Evaluating the land use potential of gypsiferous soils from field pedogenic characteristics. Soil Use Manag. 13(2): 97–103. doi: 10.1111/J.1475-2743.1997.TB00565.X.

Veum, K.S., K.A. Sudduth, R.J. Kremer, and N.R. Kitchen. 2015. Estimating a Soil Quality Index with VNIR Reflectance Spectroscopy. Soil Sci. Soc. Am. J. 79(2): 637–649. doi: 10.2136/SSSAJ2014.09.039.Wade, A.M., D.D. Richter, C.B. Craft, N.Y. Bao, P.R. Heine, et al. 2021. Urban-Soil Pedogenesis Drives Contrasting Legacies of Lead from Paint and Gasoline in City Soil. Environ. Sci. Technol 55: 7981–7989. doi: 10.1021/acs.est.1c00546.

Wan, M., M. Qu, W. Hu, W. Li, C. Zhang, et al. 2019. Estimation of soil pH using PXRF spectrometry and Vis-NIR spectroscopy for rapid environmental risk assessment of soil heavy metals. Process Saf. Environ. Prot. 132: 73–81. doi: 10.1016/J.PSEP.2019.09.025.

Wan, M., W. Hu, M. Qu, W. Li, C. Zhang, et al. 2020. Rapid estimation of soil cation exchange capacity through sensor data fusion of portable XRF spectrometry and Vis-NIR spectroscopy. Geoderma 363: 114163. doi: 10.1016/j.geoderma.2019.114163.

Wang, D., S. Chakraborty, D.C. Weindorf, B. Li, A. Sharma, et al. 2015. Synthesized use of VisNIR DRS and PXRF for soil characterization: Total carbon and total nitrogen. Geoderma 243–244: 157–167. doi: 10.1016/J.GEODERMA.2014.12.011.

Weil, R.R., and N.C. Brady. 2017. The Nature and Properties of Soils. Fifteenth. Pearson, Columbus.

Weindorf, D.C., N. Bakr, and Y. Zhu. 2014a. Advances in portable X-ray fluorescence (PXRF) for environmental, pedological, and agronomic applications. Advances in Agronomy. Academic Press Inc. p. 1–45

Weindorf, D.C., N. Bakr, Y. Zhu, A. Mcwhirt, C.L. Ping, et al. 2014b. Influence of Ice on Soil Elemental Characterization via Portable X-Ray Fluorescence Spectrometry. Pedosphere 24(1): 1–12. doi: 10.1016/S1002-0160(13)60076-4.

Weindorf, D.C., S. Chakraborty, B. Li, S. Deb, A. Singh, et al. 2018. Compost salinity assessment via portable X-ray fluorescence (PXRF) spectrometry. Waste Manag. 78: 158–163. doi: 10.1016/J.WASMAN.2018.05.044.

Weindorf, D.C., and S. Chakraborty. 2020. Portable X-ray fluorescence spectrometry analysis of soils. Soil Sci. Soc. Am. J. 84(5): 1384–1392. doi: 10.1002/saj2.20151.

Weindorf, D.C., S. Chakraborty, A. Abdalsatar, A. Aldabaa, L. Paulette, et al. 2015. Lithologic Discontinuity Assessment in Soils via Portable X-ray Fluorescence Spectrometry and Visible Near-Infrared Diffuse Reflectance Spectroscopy. Soil Sci. Soc. Am. J. 79(6): 1704–1716. doi: 10.2136/sssaj2015.04.0160.

Weindorf, D.C., J. Herrero, C. Castañeda, N. Bakr, and S. Swanhart. 2013. Direct Soil Gypsum Quantification via Portable X-Ray Fluorescence Spectrometry. Soil Sci. Soc. Am. J. 77(6): 2071–2077. doi: 10.2136/sssaj2013.05.0170.

Weindorf, D.C., Y. Zhu, B. Haggard, J. Lofton, S. Chakraborty, et al. 2012. Enhanced Pedon Horizonation Using Portable X-ray Fluorescence Spectrometry. Soil Sci. Soc. Am. J. 76(2): 522–531. doi: 10.2136/sssaj2011.0174.

Weindorf D.C., Zhu Y, McDaniel P, Valerio M, Lynn L, Michaelson G, Clark M, Ping CL (2012) Characterizing soils via portable X-ray fluorescence spectrometer: 2. Spodic and Albic horizons. Geoderma 189-190, 268-277. doi:10.1016/j.geoderma.2012.06.034

Wilson, T.S., B.M. Sleeter, and D. Richard Cameron. 2016. Future land-use related water demand in California. Environ. Res. Lett. 11(5). doi: 10.1088/1748-9326/11/5/054018.

Wolfgong, W.J. 2016. Chemical analysis techniques for failure analysis: Part 1, common instrumental methods. Handbook of Materials Failure Analysis with Case Studies from the Aerospace and Automotive Industries. Elsevier Inc. p. 279–307

Woods, J., A. Williams, J.K. Hughes, M. Black, and R. Murphy. 2010. Energy and the food system. Philos. Trans. R. Soc. B Biol. Sci. 365(1554): 2991. doi: 10.1098/RSTB.2010.0172.

[WHO] World Health Organization. 2021. Lead poisoning. Available at https://www.who.int/news- room/fact-sheets/detail/lead-poisoning-and-health

Xu, D., S. Chen, H. Xu, N. Wang, Y. Zhou, et al. 2020. Data fusion for the measurement of potentially toxic elements in soil using portable spectrometers. Environ. Pollut. 263: 114649. doi: 10.1016/j.envpol.2020.114649.

Yuan, Z., H. Chang, S. Zhou, Z. Zhang, Q. Cheng, et al. 2021. In situ monitoring of elemental losses and gains during weathering using the spatial element patterns obtained by portable XRF. J. Geochemical Explor. 229: 106842. doi: 10.1016/J.GEXPLO.2021.106842.

Zhang, Y., and A.E. Hartemink. 2019. Soil horizon delineation using vis-NIR and pXRF data. CATENA 180: 298–308. doi: 10.1016/J.CATENA.2019.05.001.

Zhang, Y., and A.E. Hartemink. 2020. Data fusion of vis–NIR and PXRF spectra to predict soil physical and chemical properties. Eur. J. Soil Sci. 71(3): 316–333. doi: 10.1111/EJSS.12875.

Zilberman, D. and S. Kaplan. 2014. An Overview of California's Agricultural Adaptation to Climate Change. ARE Update 18(1): 16-19. University of California Giannini Foundation of Agricultural Economics. Available at https://giannini.ucop.edu/filer/file/1453327771/16951/

APPENDICES

Appendix A: Laboratory data

Table A.1: LA Urban laboratory soils data.

Sample ID	Land use	Texture class	Sand %	Silt %	Clay %	N %	SOC %	C:N	pH	CEC (cmolc/kg soil)
LA Plot 11	Urban	SL	65.9	25.2	8.8	0.148	2.537	14.27	6.10	9.82
LA Plot 115	Urban	SiCL	16.3	44.5	39.2	0.088	0.938	12.47	6.33	9.06
LA Plot 116	Urban	L	48.9	32.0	19.2	0.184	2.886	13.99	6.72	-
LA Plot 12	Urban	SL	56.8	29.2	14.0	0.095	1.596	19.93	7.62	13.82
LA Plot 120	Urban	SL	54.5	39.1	6.3	0.117	1.482	16.37	5.62	14.24
LA Plot 124	Urban	SL	62.0	25.3	12.7	0.242	3.811	11.23	5.84	17.65
LA Plot 125	Urban	SL	54.4	35.5	10.1	0.132	1.383	11.83	7.33	17.59
LA Plot 134	Urban	L	51.0	30.9	18.0	0.197	2.122	13.20	7.11	18.88
LA Plot 151	Urban	SL	63.0	25.5	11.5	0.228	2.483	12.44	6.47	20.06
LA Plot 154	Urban	SL	53.7	27.0	19.3	0.229	2.042	9.48	5.40	16.06
LA Plot 16	Urban	L	49.2	35.5	15.2	0.335	4.387	11.11	6.39	27.06
LA Plot 169	Urban	L	34.1	40.1	25.8	0.576	7.457	10.35	6.53	42.82
LA Plot 171	Urban	L	49.0	35.7	15.3	0.197	3.182	12.50	7.18	12.29
LA Plot 172	Urban	SL	60.8	32.9	6.3	0.128	1.448	13.84	6.13	8.82
LA Plot 176	Urban	SL	72.4	22.6	5.0	0.134	2.590	15.83	5.26	10.53
LA Plot 185	Urban	SL	63.3	27.9	8.9	0.105	1.436	12.72	6.13	6.35
LA Plot 189	Urban	SL	60.9	31.5	7.6	0.067	0.790	15.39	5.10	5.76
LA Plot 198	Urban	SL	73.5	20.2	6.3	0.266	3.687	11.51	5.37	15.76
LA Plot 2	Urban	SL	74.6	19.0	6.3	0.087	3.498	56.90	10.38	5.47
LA Plot 202	Urban	LS	81.0	11.4	7.6	0.215	2.693	12.72	6.70	18.76
LA Plot 204	Urban	L	51.4	33.2	15.3	0.140	1.741	13.21	6.89	9.12
LA Plot 207	Urban	SL	74.8	20.1	5.0	0.034	0.545	28.40	3.28	20.88
LA Plot 21	Urban	L	50.0	37.2	12.8	0.078	0.919	12.34	7.85	22.76
LA Plot 31	Urban	CL	28.8	38.9	32.4	0.082	1.133	17.58	7.67	12.76
LA Plot 34	Urban	SL	55.3	29.4	15.3	0.155	1.757	12.15	6.54	9.65
LA Plot 35	Urban	SL	66.1	23.9	10.1	0.300	6.172	11.67	6.29	16.65
LA Plot 4	Urban	LS	81.1	12.6	6.3	0.052	1.507	45.38	6.29	5.12
LA Plot 41	Urban	SL	69.6	20.3	10.1	0.362	4.674	12.48	6.89	29.82
LA Plot 46	Urban	L	48.3	36.2	15.5	0.147	2.375	20.50	6.90	12.65
LA Plot 48	Urban	SL	61.9	24.1	14.0	0.174	1.708	10.90	6.35	14.88
LA Plot 57	Urban	SL	61.9	21.6	16.5	0.172	1.793	11.66	7.75	40.18
LA Plot 6	Urban	SiCL	17.4	46.6	36.0	0.416	6.456	13.75	5.88	31.41
LA Plot 68	Urban	L	47.3	36.0	16.7	0.486	5.721	10.79	6.02	25.24
LA Plot 74	Urban	L	47.6	35.8	16.6	0.299	4.422	15.20	5.59	27.12
LA Plot 84	Urban	SL	62.9	26.9	10.3	0.183	2.635	14.99	5.64	19.47
LA Plot 87	Urban	SCL	49.9	27.0	23.1	0.211	-	-	6.51	-
LA Plot 9	Urban	SL	56.6	29.4	14.0	0.127	1.485	13.85	8.01	12.53
LA Plot 91	Urban	SL	62.0	22.8	15.2	0.476	3.780	10.60	6.63	51.65
LA Plot 97	Urban	L	50.0	32.1	18.0	0.163	2.227	12.89	5.85	22.29

Table A.2: SPR/LHBC Mollisols laboratory soils data.

Sample ID	Land use	Texture class	Sand %	Silt %	Clay %	N %	SOC %	C:N	pH	CEC (cmolc/kg soil)
SPR/LHBC 1	Forest	-	-	-	-	0.346	5.734	16.57	6.63	32.61
SPR/LHBC 2	Forest	-	-	-	-	0.631	11.902	18.86	6.44	53.54
SPR/LHBC 3	Grassland	-	-	-	-	0.243	2.644	10.88	6.15	18.30
SPR/LHBC 4	Forest	-	-	-	-	0.681	12.446	18.28	6.53	59.11
SPR/LHBC 5	Grassland	-	-	-	-	0.303	3.122	10.30	5.85	21.34
SPR/LHBC 6	Forest	-	-	-	-	0.145	2.049	14.13	6.79	16.58
SPR/LHBC 7	Forest	-	-	-	-	0.559	8.385	15.00	6.62	39.86
SPR/LHBC 8	Grassland	-	-	-	-	0.229	2.569	11.22	6.47	15.16
SPR/LHBC 9	Grassland	-	-	-	-	0.111	1.235	11.13	6.17	15.20
SPR/LHBC 10	Grassland	-	-	-	-	0.128	1.301	10.17	6.78	14.39
SPR/LHBC 11	Grassland	CL	41.8	26.6	31.6	0.065	0.761	11.65	5.93	19.83
SPR/LHBC 12	Forest	CL	39.1	33.6	27.3	0.461	7.631	16.56	6.78	46.51
SPR/LHBC 13	Forest	CL	33.9	32.0	34.1	0.080	1.105	13.78	6.45	15.85
SPR/LHBC 14	Forest	SL	83.9	2.2	13.8	0.039	0.771	19.56	6.48	10.73
SPR/LHBC 15	Grassland	L	39.5	34.2	26.3	0.236	2.947	12.46	5.86	21.16
SPR/LHBC 16	Forest	CL	24.0	43.2	32.9	0.165	2.401	14.59	6.88	18.78
SPR/LHBC 17	Grassland	C	12.9	32.7	54.4	0.093	1.293	13.87	6.32	15.29
SPR/LHBC 18	Grassland	CL	39.2	26.6	34.2	0.157	2.323	14.80	6.25	21.44
SPR/LHBC 19	Grassland	CL	38.6	27.3	34.2	0.111	1.722	15.51	6.28	19.68
SPR/LHBC 20	Grassland	CL	40.9	27.7	31.4	0.068	0.916	13.41	6.06	17.74
SPR/LHBC 21	Grassland	CL	38.9	26.8	34.4	0.089	1.348	15.15	6.3	19.13
SPR/LHBC 22	Grassland	-	-	-	-	0.085	0.791	9.30	7.16	17.49
SPR/LHBC 23	Grassland	-	-	-	-	0.056	0.415	7.41	6.91	14.81
SPR/LHBC 24	Grassland	-	-	-	-	0.156	1.558	9.99	5.91	18.25
SPR/LHBC 25	Grassland	-	-	-	-	0.202	2.398	11.87	6	27.48
SPR/LHBC 26	Forest	-	-	-	-	0.539	8.532	15.83	6.45	45.09
SPR/LHBC 27	Grassland	-	-	-	-	0.107	1.105	10.33	6.18	15.15
SPR/LHBC 28	Grassland	-	-	-	-	0.243	2.684	11.05	6.3	15.14
SPR/LHBC 29	Grassland	CL	40.3	25.4	34.3	0.102	1.591	15.53	6.2	18.40
SPR/LHBC 30	Forest	SCL	61.2	15.8	23.0	0.237	4.602	19.42	5.49	25.00
SPR/LHBC 31	Grassland	CL	39.1	30.9	30.0	0.104	1.480	14.23	6.28	18.88
SPR/LHBC 32	Grassland	CL	40.2	31.0	28.8	0.121	1.508	12.51	5.84	18.08
SPR/LHBC 33	Forest	CL	21.4	48.1	30.5	0.321	5.508	17.16	6.62	27.53
SPR/LHBC 34	Forest	-	-	-	-	0.698	14.080	20.17	6.84	74.57
SPR/LHBC 35	Forest	SL	65.2	15.6	19.2	0.088	0.716	8.15	7.45	18.74
SPR/LHBC 36	Forest	-	-	-	-	0.084	0.762	9.10	7.31	18.92
SPR/LHBC 37	Forest	SL	68.3	14.9	16.8	0.086	0.576	6.68	7.33	19.48
SPR/LHBC 38	Forest	-	-	-	-	0.121	1.099	9.10	7.12	21.47
SPR/LHBC 39	Forest	SCL	59.7	19.0	21.3	0.142	2.121	14.90	5.39	16.43
SPR/LHBC 40	Forest	SCL	76.4	9.6	13.9	0.052	0.940	18.11	6.24	14.43
SPR/LHBC 41	Forest	CL	34.8	29.3	36.0	0.075	1.167	15.51	6.32	21.28
SPR/LHBC 42	Forest	LS	84.0	4.1	11.9	0.031	0.494	15.92	6.7	8.00
SPR/LHBC 43	Forest	LS	84.2	6.4	9.4	0.023	0.568	24.59	6.55	10.20
SPR/LHBC 44	Forest	LS	91.3	0.6	8.0	0.012	0.260	21.13	6.92	5.92
SPR/LHBC 45	Forest	CL	29.2	40.2	30.6	0.144	1.998	13.90	6.93	27.25
SPR/LHBC 46	Forest	CL	28.2	43.6	28.2	0.225	3.206	14.25	7.08	27.66
SPR/LHBC 47	Forest	CL	26.7	39.9	33.5	0.084	0.856	10.21	7.13	30.12
SPR/LHBC 48	Grassland	L	43.8	35.4	20.8	0.414	4.338	10.47	4.93	20.69
SPR/LHBC 49	Grassland	CL	31.4	39.9	28.7	0.303	3.268	10.77	5.69	22.01
SPR/LHBC 50	Grassland	SiCL	18.6	42.7	38.7	0.164	1.608	9.81	5.9	21.79
SPR/LHBC 51	Grassland	L	45.7	33.6	20.6	0.391	3.944	10.08	4.8	33.09
SPR/LHBC 52	Grassland	CL	31.2	35.9	32.9	0.448	4.102	9.16	5.57	18.99
SPR/LHBC 53	Forest	L	45.4	35.1	19.6	0.084	0.763	9.10	5.54	23.08
SPR/LHBC 54	Forest	CL	27.9	42.7	29.4	0.180	2.213	12.27	7.47	24.49
SPR/LHBC 55	Forest	CL	37.3	40.7	22.0	0.122	1.389	11.37	5.57	20.88
SPR/LHBC 56	Forest	SCL	58.4	20.4	21.2	0.114	1.480	12.99	7.06	25.34
SPR/LHBC 57	Forest	L	26.2	44.2	29.7	0.166	1.933	11.63	7.34	22.39
SPR/LHBC 58	Forest	L	47.1	27.9	25.0	0.156	2.238	14.32	7.13	30.33
SPR/LHBC 59	Forest	L	40.0	33.6	26.4	0.159	1.972	12.39	7.26	31.45
SPR/LHBC 60	Forest	L	34.7	42.7	22.5	0.177	2.157	12.21	6.85	26.25
SPR/LHBC 61	Forest	L	41.1	33.8	25.1	0.189	2.401	12.72	7.59	29.89
SPR/LHBC 62	Forest	L	32.1	44.4	23.6	0.299	4.323	14.45	6.86	30.55
SPR/LHBC 63	Forest	CL	34.1	36.5	29.4	0.255	3.884	15.25	8.01	30.76
SPR/LHBC 64	Forest	L	38.0	35.6	26.3	0.342	5.854	17.14	7.83	46.14
SPR/LHBC 65	Grassland	CL	38.2	31.5	30.3	0.201	1.657	8.24	5.62	32.86
SPR/LHBC 66	Grassland	SCL	46.8	31.9	21.3	0.587	6.168	10.51	5.08	26.50
SPR/LHBC 67	Grassland	CL	37.1	31.4	31.5	0.213	1.817	8.53	5.65	31.64
SPR/LHBC 68	Grassland	CL	36.1	36.1	27.8	0.250	2.192	8.78	5.73	31.62
SPR/LHBC 69	Grassland	CL	26.4	40.7	32.9	0.309	3.078	9.97	5.92	30.38
SPR/LHBC 70	Forest	L	30.7	43.6	25.8	0.173	2.329	13.44	6.41	19.89
SPR/LHBC 71	Forest	L	34.0	44.1	22.0	0.189	2.306	12.18	7.06	20.33
SPR/LHBC 72	Forest	L	34.3	46.0	19.7	0.360	6.444	17.90	6.87	36.08
SPR/LHBC 73	Forest	CL	27.0	45.1	27.9	0.188	2.319	12.33	7.6	28.94
SPR/LHBC 74	Forest	L	41.1	32.4	26.5	0.235	3.626	15.40	6.96	32.68
SPR/LHBC 75	Forest	L	30.3	45.6	24.1	0.309	4.134	13.40	7.55	37.38

SPR/LHBC 76	Forest	L	45.7	31.3	23.0	0.248	5.254	21.20	6.93	39.79
SPR/LHBC 77	Forest	L	42.9	30.7	26.4	0.101	1.070	10.60	6.63	22.61
SPR/LHBC 78	Forest	CL	29.7	40.1	30.2	0.316	5.392	17.06	5.92	38.66
SPR/LHBC 79	Forest	L	40.1	35.4	24.5	0.091	0.924	10.10	6.6	26.52
SPR/LHBC 80	Forest	CL	24.0	48.3	27.7	0.308	4.931	16.01	6.66	22.10
SPR/LHBC 81	Forest	L	34.5	40.9	24.5	0.145	2.154	14.86	5.69	8.50
SPR/LHBC 82	Grassland	CL	23.6	48.0	28.4	0.092	0.976	10.67	6.28	16.22
SPR/LHBC 83	Forest	CL	29.4	39.7	30.9	0.395	6.799	17.19	5.41	33.09
SPR/LHBC 84	Grassland	SCL	60.4	19.3	20.3	0.034	0.211	6.16	6.25	5.50
SPR/LHBC 85	Grassland	CL	27.2	35.7	37.1	0.234	3.140	13.45	6.18	27.91
SPR/LHBC 86	Grassland	CL	25.3	38.6	36.1	0.383	5.118	13.37	5.81	32.82
SPR/LHBC 87	Grassland	SCL	48.1	20.3	31.7	0.041	0.270	6.63	5.67	13.72
SPR/LHBC 88	Grassland	CL	28.7	34.6	36.7	0.179	2.343	13.12	5.9	23.55
SPR/LHBC 89	Grassland	SCL	60.3	18.1	21.6	0.040	0.285	7.15	5.36	7.36
SPR/LHBC 90	Forest	CL	31.8	37.3	30.9	0.341	5.131	15.04	5.51	35.43
SPR/LHBC 91	Grassland	SL	52.8	28.2	19.0	0.020	0.073	3.69	5.05	5.36
SPR/LHBC 92	Forest	CL	30.9	39.4	29.7	0.271	3.873	14.30	6.83	37.01
SPR/LHBC 93	Grassland	SCL	55.7	20.1	24.2	0.056	0.424	7.53	6.16	9.06
SPR/LHBC 94	Forest	CL	25.8	46.9	27.2	0.166	1.825	10.98	6.34	16.40
SPR/LHBC 95	Grassland	SCL	51.3	25.8	22.9	0.101	1.036	10.26	6.04	9.74
SPR/LHBC 96	Grassland	SCL	50.2	26.6	23.1	0.184	2.143	11.65	5.72	11.42
SPR/LHBC 97	Grassland	CL	31.8	37.6	30.6	0.078	0.638	8.23	5.52	14.20
SPR/LHBC 98	Grassland	CL	37.8	31.7	30.5	0.073	0.675	9.25	5.36	14.07
SPR/LHBC 99	Forest	CL	27.8	33.6	38.6	0.676	11.451	16.93	7.14	52.32
SPR/LHBC 100	Grassland	SL	68.1	12.4	19.5	0.166	1.825	10.98	5.25	7.11
SPR/LHBC 101	Grassland	SCL	67.6	12.4	20.1	0.084	0.832	9.94	5.92	7.13
SPR/LHBC 102	Grassland	SCL	61.0	9.9	29.1	0.057	0.384	6.79	6.46	9.06
SPR/LHBC 103	Forest	CL	33.0	37.1	29.9	0.348	5.969	17.14	6.9	42.92
SPR/LHBC 104	Forest	L	51.1	29.1	19.8	0.104	1.255	12.11	6.52	23.56
SPR/LHBC 105	Grassland	CL	21.1	48.0	30.9	0.125	1.557	12.43	6.38	18.65
SPR/LHBC 106	Forest	CL	29.6	35.7	34.7	0.214	3.000	14.02	6.45	22.04
SPR/LHBC 107	Grassland	SiCL	17.2	49.0	33.8	0.191	2.603	13.61	6.52	23.90
SPR/LHBC 108	Grassland	SiCL	15.9	50.1	34.0	0.265	3.787	14.29	6.47	27.18
SPR/LHBC 109	Forest	CL	32.6	37.6	29.8	0.183	2.329	12.70	7.25	30.77
SPR/LHBC 110	Forest	CL	31.4	39.2	29.5	0.262	3.767	14.39	7.13	31.95
SPR/LHBC 111	Forest	CL	39.9	27.9	32.2	0.128	1.652	12.91	7.22	29.79
SPR/LHBC 112	Forest	L	40.9	35.3	23.8	0.220	3.138	14.25	5.97	25.99
SPR/LHBC 113	Grassland	SiCL	16.3	49.6	34.1	0.402	5.548	13.79	5.95	32.55
SPR/LHBC 114	Forest	CL	35.1	35.3	29.6	0.203	2.753	13.54	6.27	36.54
SPR/LHBC 115	Grassland	SCL	50.7	18.0	31.3	0.034	0.192	5.65	5.24	13.25
SPR/LHBC 116	Forest	CL	29.8	30.9	39.3	0.526	7.757	14.74	7.41	46.73
SPR/LHBC 117	Grassland	-	-	-	-	0.121	1.225	10.12	6.59	15.21
SPR/LHBC 118	Forest	-	-	-	-	0.054	0.661	12.24	7.15	12.87
SPR/LHBC 119	Grassland	-	-	-	-	0.053	0.337	6.36	7.01	12.02
SPR/LHBC 120	Forest	-	-	-	-	0.246	3.958	16.09	6.66	17.94
SPR/LHBC 121	Forest	-	-	-	-	0.083	1.192	14.36	6.94	12.29
SPR/LHBC 122	Grassland	-	-	-	-	0.093	1.018	10.95	7.1	17.94
SPR/LHBC 123	Grassland	-	-	-	-	0.127	1.375	10.83	6.4	13.94
SPR/LHBC 124	Forest	-	-	-	-	0.107	2.131	19.92	7.13	19.51
SPR/LHBC 125	Forest	-	-	-	-	0.096	1.989	20.72	6.81	19.82
SPR/LHBC 126	Forest	-	-	-	-	0.114	1.703	14.94	6.63	15.55
SPR/LHBC 127	Grassland	-	-	-	-	0.350	4.229	12.08	6.07	23.38
SPR/LHBC 128	Grassland	-	-	-	-	0.123	1.292	10.50	6.85	17.67
SPR/LHBC 129	Grassland	-	-	-	-	0.239	2.689	11.25	6.25	16.02
SPR/LHBC 130	Grassland	-	-	-	-	0.075	0.376	5.02	7.06	14.15
SPR/LHBC 131	Grassland	-	-	-	-	0.102	1.319	12.93	6.75	19.61
SPR/LHBC 132	Forest	-	-	-	-	0.386	7.026	18.20	6.89	42.60
SPR/LHBC 133	Grassland	-	-	-	-	0.092	0.937	10.18	6.25	13.99
SPR/LHBC 134	Grassland	-	-	-	-	0.072	0.846	11.76	7.21	17.49
SPR/LHBC 135	Forest	-	-	-	-	0.037	0.511	13.81	6.89	8.71
SPR/LHBC 136	Forest	-	-	-	-	0.064	0.905	14.14	7	14.33
SPR/LHBC 137	Grassland	-	-	-	-	0.052	0.321	6.18	7.09	14.83
SPR/LHBC 138	Forest	-	-	-	-	0.400	8.427	21.07	6.91	48.04
SPR/LHBC 139	Forest	-	-	-	-	0.056	0.660	11.79	6.78	12.31
SPR/LHBC 140	Forest	-	-	-	-	0.159	2.776	17.46	7.48	17.63
SPR/LHBC 141	Forest	-	-	-	-	0.081	1.468	18.12	7.42	13.76
SPR/LHBC 142	Forest	-	-	-	-	0.172	2.333	13.56	7.21	21.89
SPR/LHBC 143	Forest	-	-	-	-	0.196	2.916	14.88	6.8	21.53
SPR/LHBC 144	Forest	-	-	-	-	0.075	0.753	10.04	7.4	12.39
SPR/LHBC 145	Grassland	-	-	-	-	0.131	1.650	12.60	6.64	24.47
SPR/LHBC 146	Forest	-	-	-	-	0.184	1.981	10.77	6.02	30.73
SPR/LHBC 147	Forest	-	-	-	-	0.392	5.146	13.13	5.75	41.23
SPR/LHBC 148	Forest	-	-	-	-	0.121	1.235	10.21	6.24	13.90
SPR/LHBC 149	Forest	-	-	-	-	0.346	3.919	11.33	6.01	23.23
SPR/LHBC 150	Forest	-	-	-	-	0.275	3.075	11.18	6.03	25.66

SPR/LHBC 151	Forest	-	-	-	-	0.236	2.430	10.30	6.06	30.62
SPR/LHBC 152	Forest	-	-	-	-	0.254	3.120	12.28	5.77	33.83
SPR/LHBC 153	Forest	-	-	-	-	0.229	2.301	10.05	5.58	31.12
SPR/LHBC 154	Forest	-	-	-	-	0.284	4.085	14.38	5.85	33.24
SPR/LHBC 155	Grassland	-	-	-	-	0.246	2.685	10.91	5.94	22.38
SPR/LHBC 156	Grassland	-	-	-	-	0.187	1.989	10.64	6.15	20.16
SPR/LHBC 157	Grassland	-	-	-	-	0.440	5.078	11.54	5.63	23.87
SPR/LHBC 158	Forest	-	-	-	-	0.307	4.871	15.87	6.17	45.32
SPR/LHBC 159	Forest	-	-	-	-	0.202	2.738	13.55	6.6	17.16
SPR/LHBC 160	Forest	-	-	-	-	0.187	3.279	17.53	6.1	58.67
SPR/LHBC 161	Forest	-	-	-	-	0.223	4.310	19.33	6.27	50.91
SPR/LHBC 162	Grassland	-	-	-	-	0.338	3.794	11.22	5.64	24.14
SPR/LHBC 163	Forest	-	-	-	-	0.388	5.849	15.07	6.73	63.15
SPR/LHBC 164	Grassland	-	-	-	-	0.310	3.610	11.65	5.8	25.18
SPR/LHBC 165	Forest	-	-	-	-	0.180	2.284	12.69	6.63	36.87
SPR/LHBC 166	Forest	-	-	-	-	0.306	4.127	13.49	5.85	32.59
SPR/LHBC 167	Forest	-	-	-	-	0.151	2.084	13.80	5.17	46.29
SPR/LHBC 168	Forest	-	-	-	-	0.168	2.250	13.39	5.68	21.93
SPR/LHBC 169	Grassland	-	-	-	-	0.345	4.241	12.29	5.67	43.00
SPR/LHBC 170	Forest	-	-	-	-	0.236	3.009	12.75	5.92	25.49
SPR/LHBC 171	Forest	-	-	-	-	0.072	0.891	12.38	5.13	25.30
SPR/LHBC 172	Grassland	-	-	-	-	0.227	2.760	12.16	5.74	50.01
SPR/LHBC 173	Forest	-	-	-	-	0.103	1.504	14.60	5.52	12.77
SPR/LHBC 174	Forest	-	-	-	-	0.184	3.080	16.74	6.6	21.35
SPR/LHBC 175	Forest	-	-	-	-	0.084	1.626	19.36	6.44	13.61
SPR/LHBC 176	Forest	-	-	-	-	0.183	2.787	15.23	5.28	14.35
SPR/LHBC 177	Forest	-	-	-	-	0.178	1.958	11.00	5.86	14.50
SPR/LHBC 178	Forest	-	-	-	-	0.481	8.659	18.00	6.02	54.20
SPR/LHBC 179	Forest	-	-	-	-	0.392	6.730	17.17	6.21	57.64
SPR/LHBC 180	Forest	-	-	-	-	0.088	0.918	10.43	5.92	14.12
SPR/LHBC 181	Forest	-	-	-	-	0.218	2.990	13.72	5.76	14.73
SPR/LHBC 182	Forest	-	-	-	-	0.034	0.330	9.70	7.84	7.74
SPR/LHBC 183	Forest	-	-	-	-	0.191	3.030	15.86	7.12	11.07
SPR/LHBC 184	Forest	-	-	-	-	0.168	1.841	10.96	5.49	18.46
SPR/LHBC 185	Forest	-	-	-	-	0.057	1.017	17.84	8.02	10.47
SPR/LHBC 186	Forest	-	-	-	-	0.223	4.482	20.10	7	34.91
SPR/LHBC 187	Forest	-	-	-	-	0.206	2.665	12.94	6.33	24.33
SPR/LHBC 188	Forest	-	-	-	-	0.504	10.141	20.12	6.7	46.89
SPR/LHBC 189	Forest	-	-	-	-	0.222	2.859	12.88	5.24	20.76
SPR/LHBC 190	Forest	-	-	-	-	0.198	2.808	14.18	5.95	24.23
SPR/LHBC 191	Forest	-	-	-	-	0.251	3.504	13.96	6.07	30.29
SPR/LHBC 192	Grassland	-	-	-	-	0.254	3.183	12.53	6.01	22.83
SPR/LHBC 193	Forest	-	-	-	-	0.375	5.561	14.83	7.21	44.57
SPR/LHBC 194	Forest	-	-	-	-	0.361	4.729	13.10	5.83	21.69
SPR/LHBC 195	Forest	-	-	-	-	0.290	4.072	14.04	5.95	19.39
SPR/LHBC 196	Forest	-	-	-	-	0.228	2.870	12.59	5.92	26.12
SPR/LHBC 197	Forest	-	-	-	-	0.368	5.769	15.68	5.98	52.52
SPR/LHBC 198	Forest	-	-	-	-	0.285	4.210	14.77	6.92	40.05
SPR/LHBC 199	Grassland	-	-	-	-	0.346	4.618	13.35	5.86	28.54
SPR/LHBC 200	Forest	-	-	-	-	0.404	5.861	14.51	7.02	54.03
SPR/LHBC 201	Forest	-	-	-	-	0.293	4.084	13.94	7.07	42.19
SPR/LHBC 202	Grassland	-	-	-	-	0.160	2.064	12.90	6	23.26
SPR/LHBC 203	Forest	-	-	-	-	0.149	2.274	15.26	6.49	41.62
SPR/LHBC 204	Forest	-	-	-	-	0.240	3.524	14.68	6.24	38.29
SPR/LHBC 205	Forest	-	-	-	-	0.120	1.561	13.01	5.91	26.15
SPR/LHBC 206	Forest	-	-	-	-	0.423	6.574	15.54	6.27	58.38
SPR/LHBC 207	Forest	-	-	-	-	0.478	7.242	15.15	7.21	53.00
SPR/LHBC 208	Forest	-	-	-	-	0.336	5.382	16.02	6.8	48.50
SPR/LHBC 209	Grassland	-	-	-	-	0.343	4.292	12.51	6.28	26.14
SPR/LHBC 210	Grassland	-	-	-	-	0.475	5.811	12.23	6.15	43.77
SPR/LHBC 211	Grassland	-	-	-	-	0.424	5.312	12.53	6.09	29.64
SPR/LHBC 212	Forest	-	-	-	-	0.212	3.099	14.62	6.92	29.68
SPR/LHBC 213	Grassland	-	-	-	-	0.205	1.793	8.75	5.25	31.72
SPR/LHBC 214	Grassland	-	-	-	-	0.460	5.034	10.94	5.42	42.96
SPR/LHBC 215	Grassland	-	-	-	-	0.368	3.792	10.30	5.27	28.30
SPR/LHBC 216	Forest	-	-	-	-	0.221	2.793	12.64	6.03	36.54
SPR/LHBC 217	Forest	-	-	-	-	0.284	3.565	12.55	6.36	42.11
SPR/LHBC 218	Forest	-	-	-	-	0.341	4.596	13.48	6.7	44.56

Table A.3: Marine terrace laboratory soils data.

Sample ID	Land use	Texture class	Sand %	Silt %	Clay %	N %	SOC %	C:N	pH	CEC (cmolc/kg soil)
Marine Terrace 1	Marine terrace	SL	59.1	28.1	12.8	0.267	3.039	11.38	6.40	15.90
Marine Terrace 2	Marine terrace	SL	56.8	28.0	15.3	0.154	1.736	11.27	5.80	13.80
Marine Terrace 3	Marine terrace	SL	63.3	24.1	12.7	0.117	1.278	10.92	6.20	11.50
Marine Terrace 4	Marine terrace	SL	64.6	20.2	15.2	0.056	0.629	11.23	6.40	8.10
Marine Terrace 5	Marine terrace	SL	61.8	24.2	14.0	0.032	0.228	7.13	6.70	7.20
Marine Terrace 6	Marine terrace	L	48.4	38.7	12.9	0.457	5.460	11.95	6.40	18.20
Marine Terrace 7	Marine terrace	L	46.0	36.0	18.0	0.231	2.395	10.37	5.90	15.40
Marine Terrace 8	Marine terrace	L	41.2	38.4	20.5	0.152	1.497	9.85	6.10	15.90
Marine Terrace 9	Marine terrace	L	46.1	35.9	18.0	0.122	1.120	9.18	6.20	16.60
Marine Terrace 10	Marine terrace	SL	63.4	22.2	14.4	0.036	0.249	6.92	6.30	11.20
Marine Terrace 11	Marine terrace	L	41.3	42.1	16.6	0.292	3.545	12.14	6.40	22.30
Marine Terrace 12	Marine terrace	L	44.1	35.6	20.3	0.174	1.807	10.39	5.70	9.80
Marine Terrace 13	Marine terrace	L	41.5	36.9	21.6	0.125	1.292	10.34	5.80	9.80
Marine Terrace 14	Marine terrace	L	41.5	36.9	21.6	0.138	1.420	10.29	6.00	11.20
Marine Terrace 15	Marine terrace	L	40.2	36.9	22.9	0.112	1.255	11.21	6.40	10.00
Marine Terrace 16	Marine terrace	LS	82.4	11.3	6.3	0.103	1.298	12.60	6.40	6.90
Marine Terrace 17	Marine terrace	L	39.8	41.0	19.2	0.235	2.656	11.30	5.80	15.60
Marine Terrace 18	Marine terrace	L	41.2	35.8	23.0	0.182	2.025	11.13	6.10	13.90
Marine Terrace 19	Marine terrace	L	36.5	40.2	23.3	0.179	2.093	11.69	6.20	12.50
Marine Terrace 20	Marine terrace	L	32.3	43.0	24.7	0.104	0.915	8.80	6.50	12.10
Marine Terrace 21	Marine terrace	SL	64.5	24.1	11.4	0.233	2.862	12.28	6.20	8.30
Marine Terrace 22	Marine terrace	SL	64.5	24.1	11.4	0.149	1.664	11.17	5.60	8.30
Marine Terrace 23	Marine terrace	SL	69.7	18.9	11.4	0.074	0.743	10.04	5.90	5.30
Marine Terrace 24	Marine terrace	SL	77.3	13.9	8.8	0.042	0.462	11.00	6.30	4.90
Marine Terrace 25	Marine terrace	SL	79.9	8.8	11.3	0.014	0.143	10.21	6.60	3.90
Marine Terrace 26	Marine terrace	LS	78.6	13.9	7.6	0.116	1.576	13.59	6.30	8.30
Marine Terrace 27	Marine terrace	LS	82.4	8.8	8.8	0.071	0.830	11.69	5.60	6.50
Marine Terrace 28	Marine terrace	LS	84.9	7.6	7.6	0.029	0.363	12.52	5.80	4.50
Marine Terrace 29	Marine terrace	LS	79.7	11.4	8.9	0.015	0.153	10.20	6.40	4.20
Marine Terrace 30	Marine terrace	LS	82.3	8.8	8.8	0.013	0.138	10.62	6.80	3.30
Marine Terrace 31	Marine terrace	L	49.2	35.6	15.3	0.314	3.739	11.91	6.30	12.90
Marine Terrace 32	Marine terrace	SL	56.7	26.7	16.5	0.171	1.803	10.54	6.00	10.10
Marine Terrace 33	Marine terrace	SL	56.9	27.9	15.2	0.113	1.248	11.04	6.10	10.00
Marine Terrace 34	Marine terrace	SL	58.2	22.8	19.0	0.096	1.064	11.08	6.50	9.80
Marine Terrace 35	Marine terrace	SL	72.3	17.7	10.1	0.033	0.357	10.82	6.70	4.50
Marine Terrace 36	Marine terrace	L	33.3	43.6	23.1	0.293	3.354	11.45	6.50	18.20
Marine Terrace 37	Marine terrace	LS	81.1	12.6	6.3	0.063	0.769	12.21	5.80	5.60
Marine Terrace 38	Marine terrace	LS	82.3	7.6	10.1	0.043	0.494	11.49	5.60	5.30
Marine Terrace 39	Marine terrace	SL	79.9	10.1	10.1	0.025	0.300	12.00	6.20	4.60
Marine Terrace 40	Marine terrace	LS	82.4	7.5	10.1	0.018	0.189	10.50	6.40	4.10
Marine Terrace 41	Marine terrace	SL	59.3	24.2	16.5	0.137	1.685	12.30	6.60	9.20
Marine Terrace 42	Marine terrace	SL	60.6	21.6	17.8	0.136	1.543	11.35	5.90	11.00
Marine Terrace 43	Marine terrace	SL	61.9	22.8	15.2	0.124	1.322	10.66	6.10	10.30
Marine Terrace 44	Marine terrace	SL	69.6	16.5	13.9	0.083	0.923	11.12	6.20	8.00
Marine Terrace 45	Marine terrace	SL	72.2	13.9	13.9	0.042	0.431	10.26	6.40	5.40
Marine Terrace 46	Marine terrace	LS	84.9	8.8	6.3	0.081	1.045	12.90	6.80	5.90
Marine Terrace 47	Marine terrace	LS	79.8	12.6	7.6	0.069	0.808	11.71	5.90	5.10
Marine Terrace 48	Marine terrace	LS	84.9	6.3	8.8	0.026	0.334	12.85	6.20	5.80
Marine Terrace 49	Marine terrace	LS	87.4	5.0	7.6	0.018	0.196	10.89	6.50	4.60
Marine Terrace 50	Marine terrace	LS	87.4	6.3	6.3	0.01	0.105	10.50	6.60	4.60
Marine Terrace 51	Marine terrace	SL	72.1	19.0	8.9	0.204	2.246	11.01	6.20	11.40
Marine Terrace 52	Marine terrace	SL	72.1	15.2	12.7	0.147	1.523	10.36	5.50	9.70
Marine Terrace 53	Marine terrace	SL	74.7	12.6	12.6	0.095	1.082	11.39	5.90	12.30
Marine Terrace 54	Marine terrace	LS	82.2	6.4	11.4	0.069	0.801	11.61	6.20	9.20
Marine Terrace 55	Marine terrace	LS	82.4	11.3	6.3	0.023	0.272	11.83	6.30	4.70
Marine Terrace 56	Marine terrace	LS	86.1	7.6	6.3	0.128	1.518	11.86	6.60	7.40
Marine Terrace 57	Marine terrace	LS	84.9	10.1	5.0	0.056	0.682	12.18	6.00	5.00
Marine Terrace 58	Marine terrace	LS	83.6	10.1	6.3	0.04	0.455	11.38	6.10	3.90
Marine Terrace 59	Marine terrace	LS	87.4	7.5	5.0	0.014	0.150	10.71	6.30	3.10
Marine Terrace 60	Marine terrace	LS	84.9	8.8	6.3	0.012	0.131	10.92	6.40	12.70
Marine Terrace 61	Marine terrace	L	41.3	40.8	17.9	0.278	3.215	11.56	6.40	15.20
Marine Terrace 62	Marine terrace	L	44.1	39.4	16.5	0.173	1.826	10.55	5.80	10.80
Marine Terrace 63	Marine terrace	L	45.3	35.6	19.1	0.127	1.391	10.95	6.30	11.00
Marine Terrace 64	Marine terrace	SL	55.8	30.3	13.9	0.072	0.709	9.85	6.40	6.90
Marine Terrace 65	Marine terrace	SL	72.2	20.2	7.6	0.026	0.221	8.50	6.60	3.50
Marine Terrace 66	Marine terrace	SL	55.2	34.5	10.2	0.357	3.974	11.13	6.50	14.70
Marine Terrace 67	Marine terrace	SL	59.3	28.0	12.7	0.212	2.322	10.95	6.00	10.90
Marine Terrace 68	Marine terrace	SL	56.8	25.4	17.8	0.119	1.321	11.10	6.40	10.60
Marine Terrace 69	Marine terrace	SL	56.7	28.0	15.3	0.118	-	-	6.50	9.50
Marine Terrace 70	Marine terrace	SL	64.6	20.2	15.2	0.06	0.680	11.33	6.70	6.80
Marine Terrace 71	Marine terrace	SL	72.3	18.9	8.8	0.112	1.415	12.63	6.40	5.60
Marine Terrace 72	Marine terrace	SL	62.2	26.5	11.4	0.099	1.143	11.55	5.80	6.60
Marine Terrace 73	Marine terrace	SL	67.2	22.7	10.1	0.075	0.721	9.61	5.90	6.10
Marine Terrace 74	Marine terrace	LS	79.9	13.8	6.3	0.021	0.173	8.24	6.20	3.10
Marine Terrace 75	Marine terrace	LS	82.5	10.0	7.5	0.013	0.087	6.69	6.50	2.50

Marine Terrace 76	Marine terrace	L	47.5	35.9	16.7	0.33	3.832	11.61	6.30	18.30
Marine Terrace 77	Marine terrace	L	49.1	33.1	17.8	0.175	1.776	10.15	5.40	12.60
Marine Terrace 78	Marine terrace	SL	61.8	20.4	17.8	0.087	0.857	9.85	5.90	9.10
Marine Terrace 79	Marine terrace	SCL	57.7	16.7	25.6	0.05	0.372	7.44	5.40	10.00
Marine Terrace 80	Marine terrace	SCL	57.0	18.3	24.8	0.041	0.251	6.12	4.50	8.50
Marine Terrace 81	Marine terrace	SL	67.1	20.3	12.7	0.19	2.304	12.13	5.90	11.30
Marine Terrace 82	Marine terrace	SL	67.1	17.7	15.2	0.087	1.086	12.48	5.20	10.80
Marine Terrace 83	Marine terrace	SL	67.0	19.0	13.9	0.073	0.802	10.99	5.90	8.60
Marine Terrace 84	Marine terrace	SL	65.7	20.3	14.0	0.042	0.377	8.98	6.40	7.20
Marine Terrace 85	Marine terrace	SCL	64.0	15.4	20.6	0.03	0.224	7.47	6.80	10.40
Marine Terrace 86	Marine terrace	SL	69.6	19.0	11.4	0.143	1.802	12.60	5.60	8.20
Marine Terrace 87	Marine terrace	SL	69.7	19.0	11.4	0.137	1.717	12.53	5.30	9.00
Marine Terrace 88	Marine terrace	SL	72.2	15.1	12.6	0.099	1.137	11.48	5.70	8.70
Marine Terrace 89	Marine terrace	SL	67.0	19.1	14.0	0.045	0.521	11.58	6.20	9.20
Marine Terrace 90	Marine terrace	SCL	67.5	11.7	20.8	0.036	0.291	8.08	6.70	14.80
Marine Terrace 91	Marine terrace	SL	60.5	20.4	19.1	0.276	3.231	11.71	6.70	16.00
Marine Terrace 92	Marine terrace	SL	63.1	19.1	17.8	0.146	1.802	12.34	5.50	12.20
Marine Terrace 93	Marine terrace	SL	52.8	21.7	25.5	0.077	0.824	10.70	6.40	11.80
Marine Terrace 94	Marine terrace	SL	42.9	16.9	40.2	0.045	0.428	9.51	7.10	18.60
Marine Terrace 95	Marine terrace	SL	57.9	13.2	29.0	0.035	-	-	7.20	14.70
Marine Terrace 96	Marine terrace	SL	65.7	20.4	14.0	0.277	3.285	11.86	5.60	12.90
Marine Terrace 97	Marine terrace	SL	66.9	15.3	17.8	0.14	1.599	11.42	5.30	9.80
Marine Terrace 98	Marine terrace	SCL	60.2	14.1	25.7	0.071	0.750	10.56	6.00	13.20
Marine Terrace 99	Marine terrace	SL	70.6	10.2	19.2	0.023	0.196	8.52	7.10	8.60
Marine Terrace 100	Marine terrace	SL	70.3	12.9	16.8	0.015	0.105	7.00	7.30	8.40
Marine Terrace 101	Marine terrace	CL	38.9	31.2	29.9	0.326	3.799	11.65	5.70	20.80
Marine Terrace 102	Marine terrace	CL	37.7	25.9	36.3	0.144	1.377	9.56	6.00	18.50
Marine Terrace 103	Marine terrace	CL	33.6	30.0	36.5	0.084	0.769	9.15	6.70	18.20
Marine Terrace 104	Marine terrace	C	36.2	23.4	40.3	0.062	0.543	8.76	5.70	18.10
Marine Terrace 105	Marine terrace	CL	44.2	18.6	37.2	0.037	0.289	7.81	4.90	14.60
Marine Terrace 106	Marine terrace	L	50.8	28.5	20.7	0.583	6.684	11.46	5.60	17.90
Marine Terrace 107	Marine terrace	L	46.0	38.6	15.4	0.149	1.626	10.91	5.50	17.10
Marine Terrace 108	Marine terrace	SCL	53.4	20.7	25.9	0.062	0.596	9.61	5.70	11.50
Marine Terrace 109	Marine terrace	SL	69.3	14.1	16.6	0.023	0.202	8.78	5.10	8.10
Marine Terrace 110	Marine terrace	SL	83.5	1.3	15.3	0.025	0.173	6.92	5.10	7.20
Marine Terrace 111	Marine terrace	L	45.2	35.2	19.6	0.554	6.243	11.27	5.80	20.90
Marine Terrace 112	Marine terrace	L	43.3	32.2	24.5	0.192	2.132	11.10	5.80	16.60
Marine Terrace 113	Marine terrace	L	43.1	33.6	23.3	0.113	1.192	10.55	6.20	15.50
Marine Terrace 114	Marine terrace	C	27.8	27.6	44.7	0.062	0.469	7.56	6.40	18.40
Marine Terrace 115	Marine terrace	C	32.7	25.1	42.2	0.047	0.326	6.94	5.30	16.40
Marine Terrace 116	Marine terrace	L	42.4	32.7	24.9	0.757	8.503	11.23	5.50	32.20
Marine Terrace 117	Marine terrace	CL	44.2	24.6	31.1	0.11	1.074	9.76	5.70	16.60
Marine Terrace 118	Marine terrace	SCL	48.5	20.6	30.9	0.078	0.763	9.78	5.80	15.90
Marine Terrace 119	Marine terrace	CL	42.9	28.5	28.5	0.047	0.391	8.32	5.80	15.90
Marine Terrace 120	Marine terrace	SCL	46.5	24.8	28.7	0.032	0.156	4.88	5.30	14.50
Marine Terrace 121	Marine terrace	SL	73.1	19.2	7.7	0.204	2.356	11.55	5.90	10.60
Marine Terrace 122	Marine terrace	SL	73.4	16.5	10.1	0.099	1.163	11.75	5.50	7.30
Marine Terrace 123	Marine terrace	SL	77.3	10.1	12.6	0.073	0.767	10.51	5.90	7.90
Marine Terrace 124	Marine terrace	SL	69.5	15.2	15.2	0.069	0.656	9.51	6.20	7.50
Marine Terrace 125	Marine terrace	SC	49.9	14.5	35.6	0.042	0.326	7.76	6.80	10.80
Marine Terrace 126	Marine terrace	SL	62.0	22.8	15.2	0.228	2.671	11.71	6.00	11.90
Marine Terrace 127	Marine terrace	SL	70.9	16.4	12.6	0.133	1.519	11.42	5.40	8.70
Marine Terrace 128	Marine terrace	SL	70.9	16.4	12.6	0.076	0.806	10.61	5.70	7.40
Marine Terrace 129	Marine terrace	SCL	57.4	18.1	24.5	0.05	0.420	8.40	6.30	9.90
Marine Terrace 130	Marine terrace	SCL	57.9	18.4	23.7	0.036	0.314	8.72	6.80	11.30
Marine Terrace 131	Marine terrace	SL	65.1	18.1	16.8	0.288	3.343	11.61	5.70	14.30
Marine Terrace 132	Marine terrace	SCL	61.7	15.3	23.0	0.14	1.643	11.74	5.30	12.30
Marine Terrace 133	Marine terrace	SCL	57.8	16.6	25.6	0.076	0.774	10.18	6.70	12.70
Marine Terrace 134	Marine terrace	SC	49.6	14.2	36.2	0.039	-	-	7.10	11.60
Marine Terrace 135	Marine terrace	SCL	60.4	15.8	23.8	0.014	0.137	9.79	6.10	10.60
Marine Terrace 136	Marine terrace	SL	68.1	19.1	12.8	0.319	3.479	10.91	5.80	12.70
Marine Terrace 137	Marine terrace	SL	67.0	17.8	15.2	0.163	1.725	10.58	5.30	10.00
Marine Terrace 138	Marine terrace	SL	68.2	14.0	17.8	0.096	0.933	9.72	5.80	9.60
Marine Terrace 139	Marine terrace	SCL	50.2	18.3	31.4	0.058	0.593	10.22	6.40	14.40
Marine Terrace 140	Marine terrace	SCL	58.1	18.3	23.6	0.017	0.124	7.29	7.40	11.10
Marine Terrace 141	Marine terrace	L	43.4	36.0	20.6	0.282	3.452	12.24	5.80	18.60
Marine Terrace 142	Marine terrace	L	42.4	35.8	21.8	0.204	2.571	12.60	5.70	14.50
Marine Terrace 143	Marine terrace	SCL	56.4	19.2	24.3	0.132	1.564	11.85	6.00	12.10
Marine Terrace 144	Marine terrace	L	38.8	42.1	19.1	0.084	1.071	12.75	6.00	8.70
Marine Terrace 145	Marine terrace	L	43.7	39.7	16.6	0.044	0.369	8.39	6.00	7.30
Marine Terrace 146	Marine terrace	L	39.5	39.9	20.6	0.326	4.108	12.60	5.80	17.60
Marine Terrace 147	Marine terrace	L	34.5	42.4	23.1	0.218	2.672	12.26	5.70	15.40
Marine Terrace 148	Marine terrace	L	36.9	38.6	24.5	0.163	2.077	12.74	6.20	16.40
Marine Terrace 149	Marine terrace	CL	30.3	38.7	31.0	0.141	1.655	11.74	6.00	13.70
Marine Terrace 150	Marine terrace	L	41.5	36.4	22.1	0.059	0.514	8.71	5.80	9.70
Marine Terrace 151	Marine terrace	L	32.8	41.3	25.8	0.285	3.537	12.41	6.00	18.10
Marine Terrace 152	Marine terrace	L	38.0	38.7	23.2	0.25	2.940	11.76	5.80	17.10
Marine Terrace 153	Marine terrace	L	37.5	36.5	26.0	0.191	2.299	12.04	6.10	18.60
Marine Terrace 154	Marine terrace	L	38.5	41.9	19.6	0.105	1.013	9.65	6.00	15.10
Marine Terrace 155	Marine terrace	L	41.1	42.8	16.1	0.051	0.377	7.39	6.00	11.20
Marine Terrace 156	Marine terrace	L	38.8	39.1	22.1	0.462	5.812	12.58	5.20	24.00
Marine Terrace 157	Marine terrace	L	38.0	38.7	23.2	0.195	2.322	11.91	5.20	17.20
Marine Terrace 158	Marine terrace	L	40.7	34.8	24.5	0.136	1.538	11.31	5.90	16.30
Marine Terrace 159	Marine terrace	L	40.8	37.3	21.9	0.087	0.978	11.24	6.10	12.30

Table A.4: NRCS Chico laboratory soils data.

Sample ID	Land use	Texture class	Sand %	Silt %	Clay %	N %	SOC %	C:N	pH	CEC (cmolc/kg soil)
NRCS Chico 1	Forest	LS	78	19.5	2.5	0.25	5.83	24	4.8	13.2
NRCS Chico 2	Forest	SL	65.4	31.3	3.3	0.15	4.41	29	5.8	17.6
NRCS Chico 3	Forest	LS	77.7	20.8	1.5	0.07	2.17	32	5.7	6.4
NRCS Chico 4	Forest	SL	68.6	27.4	4	0.07	1.64	22	6.2	9.9
NRCS Chico 5	Forest	S	92.1	7.9	0	0.09	1.76	20	5.9	5.6
NRCS Chico 6	Forest	S	94.9	5.1	0	0.04	0.12	3	6.7	0.4
NRCS Chico 7	Forest	S	95.4	4.6	0	0.01	0.59	85	6.4	1.3
NRCS Chico 8	Forest	S	94.7	5.3	0	tr	0.33	167	6.6	1.2
NRCS Chico 9	Forest	LS	76.1	20.4	3.5	0.01	0.59	99	6.7	13
NRCS Chico 10	Forest	SL	69.7	27.5	2.8	0.36	9.78	27	5.9	29.3
NRCS Chico 11	Forest	SL	63.6	33.6	2.8	0.18	6.19	34	6.4	20.3
NRCS Chico 12	Forest	LS	74	24.3	1.7	0.42	9.23	22	5.8	26.8
NRCS Chico 13	Forest	SL	56.7	41.1	2.2	0.44	8.43	19	6	32
NRCS Chico 14	Forest	L	44.3	30.4	25.3	0.41	9.29	22	4.7	41.9
NRCS Chico 15	Forest	CL	35.9	25.4	38.7	0.10	1.22	12	5	41.2
NRCS Chico 16	Forest	LS	81.5	16.3	2.2	0.06	1.77	31	5.2	-
NRCS Chico 17	Forest	SL	62.3	32.2	5.5	0.03	0.62	21	6.1	-
NRCS Chico 18	Forest	LS	80.5	17	2.5	0.12	3.06	25	5.4	9.2
NRCS Chico 19	Forest	LS	74.9	24.2	0.9	0.10	1.95	20	6	8.2
NRCS Chico 20	Forest	SL	64.7	33.6	1.7	0.06	0.75	12	6	4.6
NRCS Chico 21	Forest	LS	79.9	17.7	2.4	0.13	2.67	20	-	-
NRCS Chico 22	Forest	LS	79.7	18.1	2.2	0.09	1.80	19	-	-
NRCS Chico 23	Forest	S	90.6	8.7	0.7	0.08	2.34	28	5.3	4.4
NRCS Chico 24	Forest	S	85.9	13.3	0.8	0.03	0.12	4	6.1	0.4
NRCS Chico 25	Forest	LS	75.5	22.4	2.1	0.02	0.02	1	6.4	1.9
NRCS Chico 26	Forest	LS	76.4	21.6	2	0.11	2.30	22	5.5	9.7
NRCS Chico 27	Forest	LS	78	20.3	1.7	0.06	1.46	24	5.9	6.6
NRCS Chico 28	Forest	S	93.1	5.7	1.2	0.02	0.46	21	6.1	4.4
NRCS Chico 29	Forest	LS	76	19.8	4.2	0.29	6.90	23	5.5	19.8
NRCS Chico 30	Forest	SL	71.8	26	2.2	0.01	0.97	81	6.1	4.8
NRCS Chico 31	Forest	S	90.2	8.3	1.5	0.02	0.41	23	6.5	3.8
NRCS Chico 32	Bay Delta	L	35.8	48.5	15.7	0.52	4.07	8	5.9	20.7
NRCS Chico 33	Bay Delta	L	35.2	49.9	14.9	0.09	0.78	8	7.3	11.6
NRCS Chico 34	Bay Delta	SiL	32.3	50.1	17.6	0.05	0.48	9	7.5	12.3
NRCS Chico 35	Bay Delta	L	38.5	38.9	22.6	0.45	3.54	8	6	23.3
NRCS Chico 36	Bay Delta	L	45.3	34.5	20.2	0.06	0.55	9	7.5	13
NRCS Chico 37	Bay Delta	CL	33.2	35.9	30.9	0.86	7.83	9	7	35
NRCS Chico 38	Bay Delta	C	18.5	36.6	44.9	0.06	0.96	15	8.2	27.8
NRCS Chico 39	Bay Delta	SICL	16.4	48	35.6	0.07	0.38	4	8.4	29.1
NRCS Chico 40	Bay Delta	SIL	28.2	54.6	17.2	0.03	0.07	1	8.5	26.3
NRCS Chico 41	Bay Delta	L	41.6	45	13.4	0.16	1.43	9	6.8	14.9
NRCS Chico 42	Bay Delta	L	40.5	46.4	13.1	0.11	0.77	7	6.8	13.6
NRCS Chico 43	Bay Delta	S	91.1	5.3	3.6	0.01	0.15	13	7.2	5.2
NRCS Chico 44	Bay Delta	SL	61.2	30.5	8.3	0.40	2.69	7	5.7	11.1
NRCS Chico 45	Bay Delta	SL	72.9	20.8	6.3	0.08	0.39	5	7	6.2
NRCS Chico 46	Bay Delta	S	93.4	4.6	2	0.04	0.08	2	7.6	1.8
NRCS Chico 47	Bay Delta	SL	63.4	26.7	9.9	0.17	1.20	7	6	11.3
NRCS Chico 48	Bay Delta	SL	53.8	35.5	10.7	0.10	0.53	5	6.8	11
NRCS Chico 49	Bay Delta	SiL	7.3	68.8	23.9	0.23	1.95	8	6.5	20
NRCS Chico 50	Bay Delta	SiL	21.9	63.1	15	0.13	0.85	6	7	16.4
NRCS Chico 51	Bay Delta	SiCL	13.5	56.1	30.4	0.21	1.78	8	7	23.5
NRCS Chico 52	Bay Delta	SiCL	12.1	57.8	30.1	0.13	0.69	5	7.4	22.5
NRCS Chico 53	Bay Delta	SiCL	6.3	61.7	32	0.24	1.98	8	6.3	23.4
NRCS Chico 54	Bay Delta	SiC	4	50.2	45.8	0.15	0.52	3	8	30.2
NRCS Chico 55	Bay Delta	SiL	10.7	71	18.3	0.13	0.60	1	8.8	14.4
NRCS Chico 56	Bay Delta	SiCL	7.6	57.4	35	0.19	1.65	8	6	23.5
NRCS Chico 57	Bay Delta	SiC	6.2	42.7	51.1	0.11	0.60	5	6.8	32.3
NRCS Chico 58	Bay Delta	SiL	32.9	50.8	16.3	0.07	0.10	2	8.7	17.7
NRCS Chico 59	Agriculture	CL	26.4	41.6	32	0.17	1.40	9	7	21.1
NRCS Chico 60	Agriculture	C	18.8	38.9	42.3	0.06	0.50	8	7.7	26.4

Table A.5: UC Merced laboratory soils data.

Sample ID	Land use	Texture class	Sand %	Silt %	Clay %	N %	SOC %	C:N	pH	CEC (cmolc/kg soil)
Atwater	Agriculture	LS	82.9	8.6	8.5	-	-	-	6.09	3.06
Bear Creek	Agriculture	SL	63.5	19.4	17.1	-	-	-	5.31	18.65
Alamo	Agriculture	SL	67.1	18.2	14.7	-	-	-	5.18	4.29
San Juaquin	Agriculture	SL	68.2	17.7	14.1	-	-	-	5.8	4.59

Appendix B: pXRF elemental data

Table B.1: LA Urban pXRF data.

	Concentration (ppm)																						
Sample ID	**Mg**	**Al**	**Si**	**P**	**S**	**K**	**Ca**	**Ti**	**V**	**Cr**	**Mn**	**Fe**	**Ni**	**Cu**	**Zn**	**As**	**Rb**	**Sr**	**Y**	**Zr**	**Nb**	**Pb**	**LE**
LA Plot 11	7695.75	69825.75	246372.75	3776.5	1291.75	19385.75	17560	4476.75	103.5	47	608	20307.25	27.25	63.75	295	<LOD	83.75	376.75	21	284.5	10.25	109	607313.75
LA Plot 115	23211	75123.5	196312	5912.5	428	14719.25	35774.75	9787.75	102.25	59	1171.25	69610.75	38	49.25	157	<LOD	43.5	608.5	39	162	12.5	66	566641.5
LA Plot 116	10832.75	64018.25	216216.5	2116	1267.5	17556	19477	4427.5	112.5	68	533	29797.75	27	41	103	6.67	79.75	362	21.25	167.5	11.25	22	632755
LA Plot 12	13944.75	73065.75	215921.5	681.25	1936.75	17860	27430.75	4900.5	118.5	55	688.75	33519	32.5	37.75	108.75	7	76	422.5	19.25	174.75	8.75	30.5	608986.5
LA Plot 120	14979.5	101351.5	198708.25	606.75	411	18816	14827.25	6669	121.5	70	820.75	46211	22.25	39.5	112	5.5	98.5	395.25	17.25	318.5	11.25	26	595413.75
LA Plot 124	16455	72351.75	204299.25	3104.75	1486.25	17632.25	25741.5	5241.5	111	77	902	40650.75	33.25	97	405.5	14.67	82.75	529.25	20	144.5	12.25	147.5	610458.25
LA Plot 125	11906.25	56648.5	220428.5	1133	826.75	16820.5	24031	3930.5	115	76	390.5	25485	37.75	33.25	91.25	6	76	358.25	20.75	159.75	11.25	34.25	637426.75
LA Plot 134	10615	65317.5	218257.75	1552	1438.5	18568.25	20852.75	4288	106.75	73	443.25	29626.25	44.5	52.5	149.75	8	102.5	282.25	19.5	127.75	6.25	76.5	627991.75
LA Plot 151	10057	64162.25	209478.25	1277.5	1444.25	16449.5	17670.75	4246.25	101.25	53.5	493.75	26550	22.25	51.75	284	<LOD	78	452	15.5	154	8.25	192.75	646784.5
LA Plot 154	14110.75	71048.5	215742	3690.5	2218	18162.5	23702.75	6113.5	101.5	59.25	771.25	44166	31.75	56.25	394.75	14	74.25	511	24	156	8.25	79.75	598763.75
LA Plot 16	11015.5	70111.25	204843.25	1867	1102.5	15076.5	19600	4938.25	102.25	64	732.25	36820.25	37.25	58.25	421.5	<LOD	75	425.5	23.5	150.5	11.75	309	632230.25
LA Plot 169	9392.5	57197.75	202309.75	3662.5	2754	17246	23351.25	4726.75	101.25	96	536.5	29080.5	31.75	83.5	317.5	9.5	77.5	439	17.75	159	7	161.75	648244.25
LA Plot 171	20175	70463	194424	5063.75	1426.5	14791	35514.75	8261.25	111	76.67	1026.75	57669.75	33.75	51.25	267.75	9	52	619.75	29.25	129.25	9.5	99.5	589715.25
LA Plot 172	18273	73679.25	200486.25	5078.5	1096	15432.5	31692.75	7919.5	124.25	79.75	904	51359.25	35.25	57.75	215.75	11	54.5	660	24.75	115	6.5	102.5	592593.75
LA Plot 176	17091.5	93616.5	203538.5	1581.75	489	21758.25	18782.75	8442.5	129.5	87.5	759.25	51224.25	23	57.25	130.25	7	143	360	48	616.5	27.75	44	581011
LA Plot 185	14984	76785.25	187629.75	4561.75	771.75	14722	26146.75	14542	118.75	70.33	1237.5	68592.75	34	91.5	302.75	<LOD	54	515	42	240.5	20.5	168	588386
LA Plot 189	26048	75042.5	207508	1452.25	471.25	23372.75	21989	9011.5	153	52.33	1077.75	59393.25	34.5	43	133.75	7	99.25	452.25	21.5	225.75	14.5	25.5	573387.25
LA Plot 198	16612.75	63570.75	187262.5	4066.25	2131	14381	29773.25	10847.5	109.25	70	927.25	53188.5	30.75	73.5	203.75	22	56.25	577.5	32	195.75	13.5	351	615539
LA Plot 2	11915.5	53819	180177.25	734.5	7164.75	14327.5	63364.25	3519	111.25	68.5	655.5	24405.5	40.75	320.25	217.75	8	67.5	444.75	15.5	116.5	5.67	74.25	638421.25
LA Plot 202	10769.75	68943.5	219731.5	1515.75	1593.75	18082.5	19638	4342	112.25	55.5	508.5	27399.75	25.25	30.5	79.5	5	76.75	463.25	16.5	170	7.75	25.75	626438.25
LA Plot 204	21249.25	68258.5	192358.5	2324	979	18329.25	28301	7723.25	110.5	70.5	923.5	49223.75	39	53.5	151.5	5.75	76.5	481.5	22.25	216.25	15.5	39.25	609046.25
LA Plot 207	7833.5	60902	242792.25	524.25	28743.25	21217	7873.5	4911	104.75	<LOD	311.25	18614.25	16	11	32	4.33	78.75	359.75	13.75	171.25	6.75	13.5	605474.5
LA Plot 21	12183.75	76523.25	214402.25	322.25	320	13719.25	20270	5460.25	123	56	836.75	40974	34.5	37.75	135.75	7	77.75	461.75	25.25	212.75	12.25	31	613782.5
LA Plot 31	15397.5	66837.5	214125.75	1144.5	957	15703	32341.75	4878	110.25	91.5	476.75	33075.75	39.75	53.75	269.25	8.67	80.5	360.25	18.75	143.5	7	78.75	613802.5
LA Plot 34	16208.5	70761	210140.5	3125.5	1723.5	18733.25	30245.5	5692.25	116	61.75	749.75	38864	31.75	60	323	<LOD	77.75	527	21.5	166	9.75	173	602186
LA Plot 35	12006.75	60786	194598.75	3621.5	3230.5	18060.75	25184.5	4946.25	109.5	69.5	610.25	33371	32.25	82.25	584.5	<LOD	78.75	534.75	21	196.25	8.75	691.25	641199
LA Plot 4	10389	68262.75	231385.25	844.75	3280.25	19821.75	12806.75	3440	99	<LOD	430.25	20115.75	21.25	18.25	60.5	<LOD	81.5	412.75	13.25	141	6	34.5	628337.25
LA Plot 41	15086	64885.25	194025.5	2003.25	1421.25	15034.25	32389	4711	96.75	57.25	720	37477.75	32	58	227.25	19	73.75	443.25	19.75	157	8.5	167	630881
LA Plot 46	11166	73043	217242.75	1238.75	1981.25	17097.25	20549.25	4970.75	122.5	66	650	35067.25	30.75	121.25	530.5	<LOD	76.75	520.25	23	240.75	10.25	546.25	614723.25
LA Plot 48	12352	72468.75	207804	1983.25	1047.75	14882.75	19681.5	5598.5	96.75	80.75	650	36574.25	23.5	62.25	259.5	16.75	73	450.25	21.25	144	7.5	183.75	625538.25
LA Plot 57	10014.5	49450.5	225229	1262.75	283.5	8837.25	35434.75	4072.25	118.5	<LOD	439	34431	56	62.25	129.75	9	65.5	285.5	26	136	10.75	31.75	629615.75
LA Plot 6	16076.75	56523.75	190498.75	2070.5	3669.25	16368.5	18388	5359.25	114.75	85.25	623.5	36678	46.25	123.25	813	<LOD	87.75	323.5	24.5	155.5	9.5	768.75	651187.25
LA Plot 68	10064.75	72163.5	204041.75	1980	1607.75	19471.5	16242.75	4976.25	120	68.75	726	33503.75	28.25	56.75	229	<LOD	88.75	389	23.25	226.75	12.25	164	633815.5
LA Plot 74	7770.5	65954.5	227943	1824.5	1448.25	17065.5	17262.75	4196.25	102	59.33	490.25	24194	32.25	73	280.5	9	80.5	314	15	150.25	5.75	84.25	630653.25
LA Plot 84	6863.75	86415.5	193846.25	407.5	547.75	20413.75	7783	4058.25	89.5	55.5	503.25	30601	26.25	25.5	91.5	10	111	104	26.5	161.75	8.25	18	647858
LA Plot 87	7407.75	85246.25	216801	529	655.25	16944.75	8734.5	4914.25	98.75	84.5	886.25	41597.75	57.75	72	491.25	85	88.25	208.25	26.5	159	11.75	61.5	614828
LA Plot 9	8191.5	69891.25	244354.75	1323.75	1280.25	18576	16901.75	4025.5	146.25	57.75	510	22542.5	30.75	59.75	284.5	<LOD	84	339.25	16.75	163.75	6.75	152.25	611060.5
LA Plot 91	10328.75	53490.75	225079.5	1322	601	13337.25	17183	5075.25	110	64	453.25	40757.5	68.25	68.25	216.75	9.5	92.75	207.75	30.75	156.25	12.25	101	631210.5
LA Plot 97	10999	64924.5	235679.25	1301.5	1598.5	17047.5	13690.75	4449.25	104.5	63.67	516.75	28097.75	33.5	49.75	148.5	<LOD	80	357	22.5	183.25	10.25	239.25	620418.75

Table B.2: SPR/LHBC Mollisols pXRF data.

	Concentration (ppm)																						
Sample ID	**Mg**	**Al**	**Si**	**P**	**S**	**K**	**Ca**	**Ti**	**V**	**Cr**	**Mn**	**Fe**	**Ni**	**Cu**	**Zn**	**As**	**Rb**	**Sr**	**Y**	**Zr**	**Nb**	**Pb**	**LE**
SPR/LHBC 1	22074.5	83796.5	167491.5	1514	696	17339.5	13334.5	5462.5	105.5	164	1162.5	50301	142	54	115	8.5	82	178	26.5	185.5	10	11.5	635745.5
SPR/LHBC 2	20520	82264.5	163196.5	1626.5	649	17258	13890.5	6153	105	179.5	1118	49483.5	128.5	69.5	113	7.5	78	177.5	25	183.5	8.5	11	642746
SPR/LHBC 3	24008	86401.5	184911.5	765	457.5	17596	3692.5	5943	118	253	1148	50938.5	222	67	111.5	11	85.5	92.5	27	161	8.5	11.5	622964
SPR/LHBC 4	16640.5	64046	128179	1610.5	1265.5	13656	24116	4513.5	83	124.5	1195	40825.5	94.5	51.5	119	6.5	67	203	18.5	144	5.5	15	703005.5
SPR/LHBC 5	25932	93508.5	211699	803.5	540	20497.5	1876.5	6218.5	123.5	223	858	46327	174.5	53	112.5	22.5	118	81.5	24	232	16	20	590539
SPR/LHBC 6	23156.5	87814.5	184371	1612.5	520.5	17342	8879	6141.5	101	176.5	963.5	50288.5	155	57.5	99.5	8.5	83	160.5	29	202	11	11	617816
SPR/LHBC 7	26733.5	71800.5	153446	1038.5	944.5	15672	14585	4923.5	82	255.5	1411.5	46112.5	206	61.5	123	7	83.5	137	24	145.5	9	21.5	662177
SPR/LHBC 8	26601.5	91199	189555	548	448.5	19057.5	3038	6091	108	272.5	1000	50175	241.5	64.5	108.5	12	91.5	76.5	31.5	171	9	10.5	611090
SPR/LHBC 9	25873	90378.5	207171.5	541.5	209.5	19000	1853	6683.5	125.5	245	926.5	48454.5	195	56	114	26	123.5	82.5	26.5	215.5	17.5	18	597663.5
SPR/LHBC 10	27773.5	99969	204895	465	278	19162	2896.5	6651	118.5	336.5	951.5	54889	255	83.5	113	17	102	82	38	201	10.5	8.5	580703
SPR/LHBC 11	3657.5	73046.5	272474.5	<LOD	<LOD	9082	1208	6778	101.5	183.5	1637.5	34347.5	33	22	51.5	7	63	154	11.5	226	7.5	14	596890.5
SPR/LHBC 12	4873	48194.5	253709	924.5	1035	13023.5	17278	3553.5	67.5	54	884.5	26202.5	45	41	223.5	7	83	187.5	15	121	10	14.5	629475
SPR/LHBC 13	4498	71744	250399	<LOD	96	10330.5	5975	5872	103.5	157	151.5	43493	36.5	17.5	66.5	7.5	73	217.5	6	161.5	7.5	14.5	608822
SPR/LHBC 14	7609	72397	262564.5	620	229	15229.5	11030.5	4613	82.5	51.5	510.5	30285	18	12.5	78	7	77.5	333.5	15.5	166.5	6.5	9	594052.5
SPR/LHBC 15	<LOD	60680	278364.5	326.5	438	11654.5	3946	5372.5	89	124	861	30615.5	29.5	21.5	52	4	67	190	8.5	170.5	6.5	17	606953
SPR/LHBC 16	<LOD	54459.5	303684	<LOD	299.5	14926	8123	5606.5	91	167	361.5	21868	16.5	24	38.5	6	86	164.5	8.5	236	9.5	14.5	589806.5
SPR/LHBC 17	<LOD	68170.5	288714.5	114.5	161.5	10649	3435	6411	97.5	131	1051.5	31250	34	21.5	52	6.5	66	190.5	11	186	7	13	589220
SPR/LHBC 18	<LOD	63166	258014.5	266	182	9617.5	5092.5	5787.5	109	99.5	1205	29916	34	18	64	9	70.5	204.5	10	155.5	6.5	8.5	625964.5
SPR/LHBC 19	<LOD	66916.5	271602.5	235	189.5	10251	4867.5	6146.5	94	110.5	1309	30756	34	25	58	6.5	68	212.5	10.5	220	9	12	606868
SPR/LHBC 20	4771	70825	275517	<LOD	61	9661	2048	6513.5	104.5	150	1211.5	32057.5	32	18.5	55	6	63	171.5	10.5	215	7.5	11.5	598868
SPR/LHBC 21	3559	67506	284512	116	66.5	10567.5	3403.5	6093	104	215	878.5	30484	30.5	28	54.5	8	65.5	196	12	221.5	9	12	593640.5
SPR/LHBC 22	23482	94210.5	205197.5	388.5	143	16418	3085.5	6638	107.5	290	1018	54385.5	239	77.5	111.5	13	86	96.5	32	173	11	9	593788.5
SPR/LHBC 23	22249.5	87676.5	196414	431.5	149.5	18217.5	3953	6605	122	281.5	1118.5	54181.5	235	74.5	111	12	90.5	98.5	36.5	176.5	13.5	12	607738.5
SPR/LHBC 24	22467	87190	204252	584	290	19865.5	1377.5	6047.5	120.5	247.5	873.5	47304	174	55	112	23	119.5	79	24	219.5	17	17.5	608542
SPR/LHBC 25	30594.5	80853	191189	676.5	300	12479.5	8825	6524.5	108.5	413	1173	56395.5	322.5	74	100	9.5	71	98	22	143.5	8	6.5	609606.5
SPR/LHBC 26	19076	62241	134819	1662.5	1221.5	14145.5	19170	4719	102	162.5	1186.5	40898	126.5	49.5	117.5	7	70.5	178.5	19	146.5	6	12	699850
SPR/LHBC 27	19495.5	82241.5	200355	470.5	159	18816.5	1855	6272	115	223.5	895	47840.5	170.5	53	109	24.5	109	74	24	201.5	18.5	18	620444.5
SPR/LHBC 28	30139	94861	214628.5	567.5	332.5	18850	2496	6332	102.5	262	951	47962	208	57	103	12	89.5	79.5	25.5	198	11.5	9	581716.5
SPR/LHBC 29	4489	69380	279032	<LOD	182	10226	3995	6251	119	117.5	821.5	32996.5	36.5	30	60.5	6	69.5	203	12	203	8	13.5	593991.5
SPR/LHBC 30	9031	56882	193852	1142.5	1126	13468.5	14742.5	4321	64	<LOD	558.5	26423	19	13.5	96	6	74.5	289.5	9.5	159	4	13.5	677705.5
SPR/LHBC 31	<LOD	61689	260118.5	176.5	247.5	9830.5	3841	5896.5	112	135.5	1058	31774	29.5	25	53.5	6.5	64.5	181.5	11	185.5	8	11.5	624544
SPR/LHBC 32	<LOD	60223.5	266334	159.5	154.5	9651.5	3569.5	5642	88.5	170.5	863.5	30676.5	27.5	22	58	6.5	65	187.5	9	166.5	6.5	14	621903
SPR/LHBC 33	<LOD	44216	252343	104	419.5	13351.5	10549.5	5385	71.5	104.5	400	23006	23.5	21.5	44.5	6.5	81.5	174	10	248	13.5	15.5	649460
SPR/LHBC 34	17960	58815.5	112508	6690.5	1179.5	13577	36188.5	4379.5	116.5	135	1569.5	40334	120.5	79	161.5	7	73	330	18.5	154	6.5	13.5	705569.5
SPR/LHBC 35	4785	58548.5	302882	575	497.5	11889.5	5027.5	3833	66.5	70	358	29892	71.5	39	162	7.5	72.5	177	18.5	122.5	6.5	6.5	580884
SPR/LHBC 36	5306.5	54979	294630.5	400	432	12072	5299.5	3603	64	83.5	425	28063	88	42	168.5	10	73.5	187	19	95.5	7.5	5	593928.5
SPR/LHBC 37	5372	55156.5	306314	431	440.5	12254.5	4480	3748.5	65	75.5	315.5	29344	56.5	36.5	137.5	8	72	171.5	20	96.5	8.5	7	581370.5
SPR/LHBC 38	7794	66240.5	281375	470	465.5	15434.5	6039	5171	68	66.5	345.5	37082	48	41	162.5	8.5	95.5	163	25.5	121.5	13.5	9	578748.5
SPR/LHBC 39	4614	64013	236651.5	1088	706	14805	7989	4678.5	81.5	49	362.5	23471	21.5	24	76	6	81	277.5	12.5	149	7	11	643154
SPR/LHBC 40	14040.5	80736	214861	897.5	361	18496.5	10011	7175	93	<LOD	819	46984.5	24	17	131	14	108.5	259.5	16.5	195	5	12.5	604733.5
SPR/LHBC 41	3330	61287	279753	<LOD	115	11909	5699	5524.5	96	141.5	204	38626.5	34	20.5	54	8.5	76.5	183	7	187	8.5	11.5	594382.5
SPR/LHBC 42	6636.5	68042	255785.5	808.5	296	15105.5	10593.5	4175.5	79.5	59	433	26522.5	16.5	17	71.5	6	71.5	320.5	10	143.5	5	11	610782
SPR/LHBC 43	8039	69612	206562.5	681.5	214	16268	12456.5	5523	80.5	<LOD	622.5	35025.5	17	9	93	12	81.5	338	11.5	136.5	5.5	10	644200.5
SPR/LHBC 44	11429	86478	242577.5	1287.5	214	14782.5	12908	6088	103.5	59	2304.5	45713	28	27	119	14	84	297	17.5	198	6	12	575245
SPR/LHBC 45	3791	63622.5	292712.5	<LOD	179.5	13027	4726.5	4295.5	80	107.5	287.5	27965.5	24.5	25	49	10	77.5	148.5	9	161	6.5	11.5	590580
SPR/LHBC 46	5282	63061	281579.5	314	293.5	13666	7840	4263	77	98	609	28043	28.5	25	67.5	7	80	170.5	7.5	156.5	6	12	596955
SPR/LHBC 47	3845.5	68859.5	308146.5	<LOD	107.5	11916.5	2087.5	4182	68.5	97	168	30224	20.5	36	44.5	11	76	125	7.5	161.5	6.5	8	569802
SPR/LHBC 48	<LOD	58746.5	285702	1300.5	840.5	9348.5	614	2992.5	52	80	276	24789	16.5	28	51	8	67.5	81.5	8	124.5	6	11	614855.5
SPR/LHBC 49	4371	67556	292303	245	518.5	13529.5	1697.5	4247	96	89.5	271.5	26705	25.5	24	45.5	9	81.5	126.5	16.5	190	5	11.5	587829.5
SPR/LHBC 50	4649	74658.5	308541.5	129.5	205.5	13481.5	482.5	4657.5	89	162	282	29660.5	26.5	27.5	49	7.5	84.5	122	15.5	203	6.5	12	564769.5
SPR/LHBC 51	<LOD	60584	298161.5	1329	839.5	9851.5	120	3110.5	39	108	237	23501	13.5	30.5	52.5	7.5	68.5	81	8	124	5.5	12	601772
SPR/LHBC 52	3680.5	77366	287909	<LOD	123.5	12479.5	<LOD	4454	77.5	151	56	33219	20.5	28.5	38.5	10.5	86.5	109	9.5	181	7.5	9.5	579970
SPR/LHBC 53	<LOD	54651.5	336009	<LOD	132	12410.5	445	4268	57.5	108	131	24461	20	22	39	8.5	70.5	124	7	158	10.5	9.5	566921
SPR/LHBC 54	3460	68261.5	294200	410	156	16500.5	5344.5	4899.5	96	120.5	663.5	29923	40.5	38.5	93.5	10	92	181	9	173.5	9.5	11	577037.5
SPR/LHBC 55	<LOD	55744	337607.5	<LOD	115.5	15030	1314	4677	81	92.5	292	24961.5	19	23	41	7.5	84	144	8	174	9	11	559566
SPR/LHBC 56	5389.5	65596	278225	277.5	299.5	11728.5	8059	4103	90.5	202	198.5	27954	41.5	32.5	51	6.5	61.5	298	12.5	158.5	5	9.5	597203
SPR/LHBC 57	4158	67443.5	296287.5	325.5	163.5	16450	4379	4880.5	99.5	108	606.5	28999.5	44.5	33.5	88	9	91	175.5	8	186.5	8.5	11	575444
SPR/LHBC 58	6174.5	63517.5	280859	231	341	11346.5	7234.5	3745.5	82	101	394	29418.5	40	30	71	7.5	63	248	13.5	133	4	8.5	595938
SPR/LHBC 59	3911	59170	290581.5	125	199.5	14525	6723	4245	68.5	73.5	410.5	24623	26	32.5	71	8.5	82.5	196	8.5	155.5	9.5	11	596698.5
SPR/LHBC 60	<LOD	52012.5	312491.5	186.5	246.5	14659.5	5333.5	4450.5	83.5	107	531.5	23814.5	18	27.5	48.5	7	83.5	166.5	8	170	11.5	11	585517.5
SPR/LHBC 61	4396	61717.5	297614	335.5	325	14574.5	8419.5	4056	98	105.5	502.5	24866.5	26.5	34	73	10.5	84.5	211	6.5	156	7.5	7	582373
SPR/LHBC 62	<LOD	52207	295586	450.5	473.5	14216	8852	4066	80.5	108.5	840	23294.5	17	20.5	63.5	8.5	80	175.5	10	180	7.5	13	599251
SPR/LHBC 63	4587	59347	281710	734	515	13752.5	14998	4139	104	106	606	22648	29	37.5	84.5	8	82	251	6.5	148.5	5	8.5	596093
SPR/LHBC 64	<LOD	52210	244640.5	899.5	554	12555.5	20263	3714	107	132	813.5	22862.5	33	33	105	6.5	76.5	270	6.5	132	6	12.5	640568.5
SPR/LHBC 65	5741.5	65962.5	296078	495.5	327.5	11684	1829.5	3688.5	61.5	103.5	222.5	29965.5	26.5	37	64.5	7	73.5	145.5	12.5	133	8	10.5	583321.5
SPR/LHBC 66	<LOD	50613.5	266874	1608	1400	8854.5	3204.5	2773.5	49.5	55.5	372.5	21112	15.5	25.5	83	8	67.5	93.5	7	96	5	15	642671.5
SPR/LHBC 67	4253	66034	292036.5	468	330	11724.5	2183.5	3563	69.5	89.5	245.5	28243	23.5	23	61	9	72	143	10.5	129.5	5.5	8.5	592398
SPR/LHBC 68	4564	67375.5	294747	472	476.5	11969	2871	3687	77	83	236	29642.5	23	34	70	8.5	73	151	11.5	122.5	6	11	583290
SPR/LHBC 69	4573	56009	258413.5	471.5	538	12410	3608	3491.5	67.5	68.5	279	27809.5	29.5	28	70.5	7.5	73	151	10.5	127.5	7	10	631746.5
SPR/LHBC 70	<LOD	63507	291433	298.5	214	17693	3259	4944	84	115.5	465	27877	33	29	53.5	10	87	180.5	7.5	184.5	11	11.5	589497
SPR/LHBC 71	4679	63878.5	283741.5	672.5	264	17679	5322	5091.5	93	122	623.5	29507	35.5	27	58.5	9	88.5	197.5	12.5	184.5	11	11.5	590030
SPR/LHBC 72	5312	56141.5	250965.5	941.5	516.5	15966	11388	4623.5	87	88	961.5	26738.5	33	22.5	76	8	82	232.5	8	168	9	15	628272.5
SPR/LHBC 73	4818	66246.5	278279	435	163	15902.5	5710	4673	112.5	112.5	790	31313.5	44	37	103	8.5	91.5	185	9.5	180	8.5	10.5	593158
SPR/LHBC 74	6071.5	59915	284543.5	339.5	397	13202.5	7696.5	4173.5	97.5	70.5	681	27353	38	30	88.5	7.5	76.5	207.5	9	129	6	9	594859.5
SPR/LHBC 75	4847.5	62857	257875	1191.5	385	15073.5	11006	4656	122.5	110	1231	29297	39	38	122	7.5	88.5	210	8	158.5	10	11.5	610655.5

SPR/LHBC 76	5186.5	55563	250112	459	585	10695	13722	3710.5	87	82.5	481	27194.5	31	27.5	74.5	7.5	61.5	250	9	127	4	10	631511
SPR/LHBC 77	6029.5	62177	282761.5	1222.5	144.5	12043	4862	3464.5	77	101.5	690.5	33236.5	26.5	25.5	62.5	9	74	213	19	131.5	5	9.5	592613
SPR/LHBC 78	6752.5	54433	253802.5	1128.5	820.5	11044	11583.5	3180.5	59.5	102.5	476.5	27610.5	43.5	27	97	10.5	62.5	252.5	17	132	5	11.5	628348
SPR/LHBC 79	6581	61568	293714.5	751	142.5	12309	5017	3582	70	96	296.5	27221.5	27	24.5	59.5	8	73.5	219.5	17.5	157.5	4.5	8	588052
SPR/LHBC 80	<LOD	50831.5	294457.5	678.5	654	14843	8803.5	5209.5	93.5	208.5	511	19388.5	15	26	67	5	80.5	199.5	8.5	217.5	10.5	26	603665.5
SPR/LHBC 81	<LOD	61135	359529.5	<LOD	108	16899.5	2077.5	6360.5	108.5	257.5	209.5	18688	15.5	15	33.5	6	81.5	180	10	256	8.5	8.5	533995
SPR/LHBC 82	4094	60534.5	322563.5	197	247.5	14737.5	2255	5164.5	88	139	363.5	30301.5	19	28	46.5	13	81	171	9.5	206	11.5	10.5	558814.5
SPR/LHBC 83	4352	49019.5	234860.5	1106	858.5	10070	10665.5	3140.5	64.5	82	423	27133.5	34	30.5	96	8.5	60.5	241	17	111	4	12.5	659787.5
SPR/LHBC 84	3187	75434	277136	<LOD	77	12686	1143.5	4379.5	81	166.5	166	21140	22	13	28	6	67.5	136	6	174	3	8	605578
SPR/LHBC 85	4228	59621	300226	515.5	408	14949	5347	4147.5	81.5	103	466.5	24076	27	30.5	61.5	10	98.5	200.5	13.5	180	7.5	8.5	585187.5
SPR/LHBC 86	4594	54572	282132	662.5	643	13898	6643	3828	83	106.5	501	23092	25.5	26.5	74	8.5	92	193	12.5	161	5.5	8	608640
SPR/LHBC 87	4753	77760.5	251835	<LOD	92	9283.5	1280	3554	66	163	73.5	27335.5	24	11.5	34	7.5	51.5	120	9.5	187	4	8	625711
SPR/LHBC 88	4143	60221	300564	308	255.5	14758.5	2785	4199	86	122.5	392.5	26576	28	27.5	59.5	8.5	95	176	21	189	7.5	10	584968
SPR/LHBC 89	<LOD	85809.5	281454.5	<LOD	131	11481	586	4477	81.5	180.5	147	26168.5	32	14	30.5	5.5	66.5	131	8	198	4.5	12	588968
SPR/LHBC 90	5034.5	54011.5	257005	1084	731.5	10737.5	9753	3199.5	62.5	94	490.5	27616	46	32.5	107	10.5	64.5	255	18.5	113	4	9	629513.5
SPR/LHBC 91	<LOD	75950.5	328119	<LOD	54	14020.5	<LOD	4763	74.5	208	108.5	18481	21.5	13.5	20	<LOD	69.5	136.5	5.5	216	7	10.5	557744.5
SPR/LHBC 92	4290	53626.5	260052	1412	648	11757.5	12721.5	3244.5	79	83	680	26088.5	43	33.5	95	9.5	69	231.5	19	122	6.5	11	624678
SPR/LHBC 93	4484	73182.5	250923.5	<LOD	88	12297.5	1503.5	4071.5	75	137.5	302.5	24593	24	14	36	5.5	71	133	5	194	4	8.5	630082.5
SPR/LHBC 94	<LOD	55457	325577.5	407.5	351	16385.5	4911.5	5957	74	186	488	20306	21.5	19.5	60.5	<LOD	87	186.5	10	253.5	11	24	569226
SPR/LHBC 95	<LOD	75126.5	283291.5	145	244.5	13636	2870.5	4664	96.5	332	489.5	23712	29	15	34.5	3.5	73	155.5	9.5	195.5	5.5	10	594923.5
SPR/LHBC 96	3773	71432	272334.5	255	578	12955	3308	4092.5	105	128	470.5	22088.5	22	15.5	36	4.5	72	151	7.5	190.5	4	10.5	607957.5
SPR/LHBC 97	4123.5	64052	341608.5	<LOD	166.5	15303	1805	4693.5	72	134.5	228	21751.5	19	27	45.5	7	83.5	172	18	212	7.5	10	545461
SPR/LHBC 98	4838	58125	291948.5	<LOD	183	13905.5	1896	4623	82.5	144	173.5	25687.5	21.5	25	42	10	76.5	159	12.5	239	7.5	9	600208.5
SPR/LHBC 99	5040	45691.5	177337.5	1593.5	1381	7326	35353	2595.5	56.5	43.5	476.5	22016	31.5	33.5	101	7.5	58.5	189.5	4	105	4	9.5	700534.5
SPR/LHBC 100	<LOD	57267	289311.5	365	691.5	12233.5	3532	3697.5	77.5	261	278	14561.5	14	14.5	24.5	6	61	138.5	7	194	4.5	10	617253
SPR/LHBC 101	<LOD	62482.5	295002	164	259	13243	2902.5	2846	73.5	111	266	13682	18	9.5	29.5	4	63	136.5	7.5	183	3	8	608509
SPR/LHBC 102	4179	79892	261234.5	<LOD	113.5	10853	2609	4010	83.5	365.5	109.5	26922	24	13	36	7	64	139.5	9.5	287	6	9	611123
SPR/LHBC 103	5267.5	53710	241528	1755	673.5	10823	17656.5	3312	77	95	1187.5	25478.5	57	35.5	134.5	7.5	69	265.5	15.5	117.5	4.5	11.5	637700.5
SPR/LHBC 104	4594	62660.5	293616.5	592	397.5	12950.5	5847.5	3236.5	71.5	97.5	412	25587.5	27	26.5	59.5	9	71.5	211	10	149.5	3	7.5	589364
SPR/LHBC 105	4320	61544.5	307639.5	193.5	224	14494.5	3106.5	5113.5	86	140	414.5	32254	22	28.5	45	11.5	85.5	173.5	9.5	221	10	10.5	571997
SPR/LHBC 106	5334.5	55854	235498	1558.5	341	11930.5	7939	3451	91.5	106.5	1732	38592.5	44	36	128.5	8.5	73.5	236.5	24.5	125.5	5	10	636880.5
SPR/LHBC 107	4562	62921.5	303083.5	334	279	14114.5	4531	4873	97.5	150	610	31449.5	21	34	48	11.5	95.5	188	8	229.5	7	10	572332.5
SPR/LHBC 108	3231	61259	288600.5	580	435	14017.5	7301	4871	108	113	716.5	31069.5	24	31	56	13.5	109	191	7.5	197	8	9.5	588667
SPR/LHBC 109	4397.5	60760.5	282592	1152.5	366.5	12858	8572	3772	87	112.5	788	27287	40	34	90	9.5	76	241.5	18	136	7	9	596594
SPR/LHBC 110	6238	56971.5	260541.5	1801	525	12488	12893	3627	89	74.5	1014	28097	47.5	37	116.5	9.5	75.5	269	18.5	127.5	6	9.5	614924
SPR/LHBC 111	4478	55438.5	324124.5	234	311.5	9890.5	7452	3197	52.5	83	97.5	29763	23	24	58	12.5	75.5	86	4.5	116	5.5	7.5	564455
SPR/LHBC 112	3942	54714	270523	722.5	490	12231.5	7264	3240	74	130.5	527	22116	28	23.5	69	6	70	202	15.5	145.5	3.5	12	625420.5
SPR/LHBC 113	4778	57539	268084.5	733	675	13251.5	8683.5	4453.5	95	128	785	29929.5	18	39.5	69.5	12.5	108.5	187	7	170.5	6	10	612625.5
SPR/LHBC 114	5139	61897.5	281584	1305.5	630.5	11342	8589	3388	96.5	103	437	29530	46.5	37	101	11	67.5	243	18.5	112.5	4	8	595295
SPR/LHBC 115	3613	76815.5	256984.5	<LOD	60	7845	952	3528	77	244.5	79	27484.5	23	7	30	4.5	48	121	8	144.5	5	9	623743
SPR/LHBC 116	5986	49414.5	214037	1162.5	840	9580	23632	3235	71	65.5	367.5	25210	33.5	37.5	93	6.5	71	161	5	130	5.5	9	668841
SPR/LHBC 117	27729.5	92753.5	205418.5	486.5	196	19300	2628	6525	113	286.5	994	50401.5	229	75.5	113	12	93.5	78	30.5	178.5	15	10	592332
SPR/LHBC 118	22700	99811	204137	740	170	16997.5	5160	8191.5	117.5	156	929.5	53323.5	113.5	52.5	103.5	6.5	76	159.5	27	242	11	11.5	586757.5
SPR/LHBC 119	29787.5	102670	206738	359.5	71.5	21420	900	7318.5	126.5	270.5	820.5	55257	219	82	125	16.5	106	60.5	35	209.5	13.5	9	573377.5
SPR/LHBC 120	22680.5	85581	166707	1450.5	776.5	17384	14326.5	6186.5	93.5	162.5	1182.5	49877	130.5	63	112.5	8	79.5	179	24	190.5	9.5	11.5	632770.5
SPR/LHBC 121	25810	98640	201851.5	1180	297	17176	6915	6139.5	114.5	199.5	974.5	50622	142.5	57	102.5	8	78.5	159.5	26	250	9.5	11	589219
SPR/LHBC 122	29872	102249.5	205007.5	435	222	19361.5	2684	6571.5	106.5	306.5	944.5	54225.5	255.5	95	112	13.5	98	81.5	36.5	191.5	11	11	577093.5
SPR/LHBC 123	28666	94099.5	208774.5	543.5	276	19872	2754	6246	108.5	271.5	969	49245	218	67	107	10.5	94	76	30	176	12	9.5	587374.5
SPR/LHBC 124	20481	89346	172486.5	3755	343.5	17812.5	12559	6680	131	165	1093	53869	123	59.5	108	6.5	79	205.5	22	203.5	10	10.5	620443.5
SPR/LHBC 125	22874.5	93729.5	189590.5	2756.5	322	17929.5	10366.5	6593.5	107	154	1127	55067	119.5	56.5	105	8.5	78	194	21.5	217	12	10	598550.5
SPR/LHBC 126	25341	98145	200817	906.5	381	19852	7776.5	5934.5	102.5	194	1127.5	52346.5	147	62.5	110.5	8	84.5	161.5	26.5	212	9	12	586231.5
SPR/LHBC 127	32883.5	79260.5	195216	879.5	654.5	13003	9506.5	6497	108.5	416	1113	54247.5	309.5	70.5	96	8	65.5	97.5	24	145	7.5	7	605364
SPR/LHBC 128	25646.5	90061.5	202631.5	719.5	359.5	18674	3360.5	6020.5	110.5	275	1077	52374.5	236	69	114	11	85.5	96.5	29.5	170	9.5	10	597854
SPR/LHBC 129	22254	99666	212149	609	539.5	20035	3783	5210.5	131.5	186	1011	45124.5	138.5	59	95.5	9	89.5	87.5	38	176.5	9.5	13.5	588583.5
SPR/LHBC 130	24041	104888.5	200033	228.5	124.5	20644.5	2779	6176.5	125	229	938.5	59844.5	212.5	75.5	136.5	14.5	125	139.5	27	163.5	12.5	15	579026.5
SPR/LHBC 131	36751	85021.5	204635	554	151	12416.5	9038	7355.5	119	440.5	1179	59185	340.5	88.5	100.5	9	71	96.5	25.5	160	9	10.5	582243.5
SPR/LHBC 132	22614	71542.5	139920.5	5849.5	871.5	15856.5	25178.5	4759	118.5	165	1406	45719.5	140	76	150	7	80.5	275.5	21.5	163	7.5	11	665054
SPR/LHBC 133	30141	99029	252825.5	368.5	117	17237	<LOD	6918	111.5	264	774.5	45899	196	56.5	110	23.5	103.5	65	23.5	241	13.5	14.5	545456
SPR/LHBC 134	40869.5	92185	220118.5	364.5	51	12571	8048	7276.5	111.5	520	1152	60840.5	349	86.5	96.5	9	71	94	26	166	7.5	9.5	554996
SPR/LHBC 135	31002	92571	228473	680.5	58	16613.5	3554	5375.5	122.5	213	792	46840	145	39.5	82.5	6.5	73	147.5	20	207.5	6.5	9.5	572952.5
SPR/LHBC 136	26980	93821.5	216660.5	533	150.5	17037.5	6196	5245	112.5	235	969	50358	169.5	54	102	9	80	130	23.5	185.5	8	9.5	580920.5
SPR/LHBC 137	29627.5	104758.5	228426.5	255	83	17161	2253.5	7003	113	309	1061.5	55395	241	75	103	11.5	84.5	87.5	28	169	9.5	13	552764
SPR/LHBC 138	21414	68577.5	141658.5	947.5	1058.5	14909	22025.5	4427.5	91	112	1020.5	40480	116.5	46	95.5	7	64	183.5	20	146	5	10	682584.5
SPR/LHBC 139	23451.5	97974	208503.5	691	203.5	18572.5	5350	6808	113	191.5	866.5	52991.5	133	42	98	8	79	162	31.5	200	12.5	6	583506.5
SPR/LHBC 140	28697.5	92127.5	194365.5	727	447.5	18641.5	10880	5797	112	181.5	1106	49793.5	147	53.5	108.5	8	79.5	158.5	24	168.5	8	9	596346
SPR/LHBC 141	25747	92195.5	199018	735	226	17772	8154	6042	128.5	146	1016.5	55334	121.5	44.5	104.5	6	70.5	180.5	24	215.5	9.5	10.5	592697.5
SPR/LHBC 142	36687	95444.5	198130	750	347	18922	6454.5	6006	121.5	279	1137.5	53881	240.5	75.5	115.5	12	98.5	109.5	29	157	8.5	13.5	580981.5
SPR/LHBC 143	25059	87430.5	195268	884.5	478.5	16821.5	9655	5041	106	195.5	1009.5	48046.5	159.5	50	103	10	84.5	136.5	23.5	150.5	10	10.5	609258
SPR/LHBC 144	30176	95925	220204	310.5	119.5	22525.5	3853	5174	99.5	231	925.5	48946	182	66.5	100	10.5	120.5	97	27.5	154.5	9.5	12.5	570729.5
SPR/LHBC 145	38850.5	87208.5	208894.5	594.5	178	12091	9530	7042.5	117.5	444.5	1235.5	59955	348.5	84	93	8	70	98	23	156.5	8	9.5	572948.5
SPR/LHBC 146	4713.5	66069.5	266182	450	412.5	13374.5	3711.5	3991.5	78	139.5	293.5	29567	25	21	62.5	8	76.5	172	7.5	136.5	6.5	11.5	610489.5
SPR/LHBC 147	3374	53065	263861.5	544.5	610.5	13136.5	5609	3483.5	50.5	65.5	413.5	24234.5	25	29.5	74	9	71.5	150	7.5	132.5	9	9.5	632717.5
SPR/LHBC 148	3432	67701.5	295204.5	148	257	14906	2767	4349.5	94.5	125	205	21091	14	18.5	33	6.5	79.5	171.5	8.5	151.5	6.5	10.5	589211
SPR/LHBC 149	5236	60599	263012.5	629	736.5	14314	5860	3966.5	85	149	413.5	21193.5	17	22	48	7	85	188	10	165	6	12	625856
SPR/LHBC 150	<LOD	64451.5	269308.5	438	480.5	14375	5064.5	4211	78.5	114.5	322	21912	19	18.5	41	6.5	85.5	187.5	8	159	5.5	12	618700.5

SPR/LHBC 151	6851	65133	260546.5	499	508.5	13212.5	4590.5	4089.5	58.5	113	284.5	30396.5	31	22.5	63	10	75	169	8.5	129.5	6	8	613192.5
SPR/LHBC 152	4958.5	57577.5	280790	417	405	13366.5	3387	3412	67	97.5	392.5	24462.5	23	27.5	69.5	10.5	74.5	137	6.5	122	5.5	8	610182
SPR/LHBC 153	4312	55206	281992.5	351	318.5	13735.5	2774.5	3626	63	94	347.5	24203.5	20	28.5	66.5	10	75.5	129.5	8	116.5	8	9	614653
SPR/LHBC 154	5497.5	59194.5	237646.5	514	591.5	12570	5599.5	3667	76	92.5	399.5	30248.5	29	21.5	77	7	72	158.5	8	124.5	7.5	13.5	643385.5
SPR/LHBC 155	3673	72807	265760.5	670	558.5	13337	4280.5	3986.5	85	111	533.5	25393	22	28	50.5	7.5	83.5	200	14	178.5	4.5	9.5	610042
SPR/LHBC 156	4052	72341	265571.5	445	363.5	13401.5	3846.5	4165	93	173.5	522.5	25548	21.5	24.5	50	8.5	80.5	200	15.5	203	5.5	10	610876
SPR/LHBC 157	6002	57323	249236.5	728	1044	11739.5	5563	3733	55.5	106	487.5	26875	28.5	31.5	77	8	67.5	152.5	7	128.5	5	10.5	639591
SPR/LHBC 158	6005	54297.5	264231	695.5	614	12875.5	8813.5	3663	70	90.5	441	29189	24	27	94	9	75.5	175.5	7.5	133	6.5	9.5	618454.5
SPR/LHBC 159	<LOD	55731	277286	587.5	380	13568	6437	3839	63	97	364.5	29620.5	26.5	26	78.5	11	77.5	168	8.5	130	8.5	7.5	611473.5
SPR/LHBC 160	4196	57921.5	246802.5	302.5	636.5	7732	7111.5	3148.5	56.5	86	125	32193	24	45	64	14.5	59	115.5	8.5	100	6	11.5	641318.5
SPR/LHBC 161	5248.5	54261	231515	389.5	896.5	8117	9780.5	3091	56	72.5	159	30739	24	36	78	12.5	60	126	7.5	104.5	6.5	14.5	655180.5
SPR/LHBC 162	4464	71042.5	265620.5	690.5	694	13315	4488.5	4301	92.5	92	527	23444	22	24	50.5	6.5	81.5	200	13	173	5.5	12	612871
SPR/LHBC 163	5725.5	47495.5	229242	1020	947	8936.5	14373	2740.5	48.5	61	420	25801.5	23.5	32.5	86	10	57	162	6.5	84.5	4.5	10	662741.5
SPR/LHBC 164	<LOD	60946.5	259820.5	681.5	794	12424.5	4945.5	3795.5	72	101.5	454	29592.5	31.5	32.5	74	8	71.5	159	7	131	5.5	12	625839
SPR/LHBC 165	4627	56619	282702.5	599	339	13334	5325	3951.5	54	96	332	29978.5	24.5	27.5	75.5	12	79	162.5	7.5	132.5	7	8	601507
SPR/LHBC 166	4771	61043.5	282194.5	379.5	559	10379	2867	4170	74.5	119	158	33870.5	20.5	33	45	16	66	131	8.5	148	8	12	598925
SPR/LHBC 167	4657.5	65500	269560.5	347.5	730	8142	2678	3325	47.5	142.5	71.5	35792	21	54.5	63.5	20.5	64	91	7	102	4.5	13.5	608542.5
SPR/LHBC 168	4929	58048.5	286465	140	320.5	11331.5	1078	4316	77.5	91.5	125.5	32292	17.5	28.5	39	15.5	65	123	7.5	151.5	6.5	7.5	600322.5
SPR/LHBC 169	5827.5	62919	260593	294.5	667.5	8055	3368.5	3915	59	152.5	130.5	34475	19.5	34	50	16.5	55	112.5	7.5	135	5	9.5	619097.5
SPR/LHBC 170	3819.5	59607	277915.5	259.5	436	10740.5	2299	4305	72	100	132.5	32861	16	36	40.5	16	68	130.5	6.5	165.5	6.5	11.5	606953
SPR/LHBC 171	3722.5	100900.5	201168.5	<LOD	172	5852	619.5	4036.5	67.5	135.5	70.5	41577.5	29	9	36.5	8	47	98.5	5.5	154.5	<LOD	10.5	641264.5
SPR/LHBC 172	6187.5	61972	246149	193	430.5	7225.5	2231.5	3725.5	52.5	92	93.5	36036.5	20	35	43.5	16.5	53	99.5	7.5	135	4	10	635185.5
SPR/LHBC 173	<LOD	72383.5	257114.5	<LOD	292	9816	2222	3452	71.5	122.5	167.5	22175.5	17.5	15	25	5.5	58.5	119	4.5	156.5	<LOD	8	631774.5
SPR/LHBC 174	3543	52681	286971.5	485.5	426	15189	8691.5	3908	110	96.5	523.5	18432.5	15.5	19	56.5	3	76	185	5.5	168	6.5	10	610155
SPR/LHBC 175	<LOD	53207	310575	169	183.5	15855	4683.5	4523	85	110	253	17656	12	15	42.5	<LOD	74.5	161.5	4.5	169.5	6	10	592201
SPR/LHBC 176	3766	65503.5	247530	343	612.5	10468	3846	3870.5	80.5	89	301	21867.5	21.5	10	28.5	5	61.5	129	7.5	130	<LOD	10	643192
SPR/LHBC 177	3972	65629.5	285434	333.5	372.5	14293	3679.5	3764.5	76	171.5	468.5	19355	20	17	39.5	4.5	80.5	141.5	6	155	5	10	603952
SPR/LHBC 178	4554	36460	190186	779	1100.5	8563	13192.5	2297.5	43	64.5	631.5	17158	13.5	16.5	84.5	5	52	118	5	80	<LOD	8.5	724590
SPR/LHBC 179	4820	43513.5	232667	856.5	860.5	9761.5	11325	3001.5	44.5	74	519	21022	16.5	24.5	86	5	62	123.5	4.5	93.5	4.5	9	673517.5
SPR/LHBC 180	<LOD	68074	302144.5	92	196.5	14306.5	2042	4597	89	191.5	236	20748	20.5	15	30	7.5	77.5	143.5	7.5	221	6.5	9	586791
SPR/LHBC 181	<LOD	64511.5	277871.5	561.5	736.5	14563.5	3690.5	4004	91.5	179.5	523.5	19852	15	16	45	5	82	142.5	6.5	151	5	10	612938
SPR/LHBC 182	30536	92947	228205	189	72.5	21265	4473	4977	117.5	230	925.5	45021	188.5	44.5	107	10.5	83	113.5	25.5	178.5	10	11.5	570270
SPR/LHBC 183	25984	82742.5	181852	679.5	600	18788.5	13529	5665	95.5	240	876.5	47895.5	160.5	55	113	7	79	166	25.5	195	11.5	11	620227.5
SPR/LHBC 184	<LOD	54526.5	302349.5	286.5	275	16207.5	2238	4583.5	86	122	296.5	27884	21.5	21.5	52	8	89	152.5	8	163.5	10.5	12	590601
SPR/LHBC 185	53585	79628.5	186132.5	495	159.5	18409	16308.5	6271	109	425.5	1095.5	60260.5	371	93.5	121.5	11	80.5	98	26.5	148	10.5	8	576150
SPR/LHBC 186	21660	70300.5	152106.5	833.5	785	16893	18204.5	5048.5	95	199.5	760	43626.5	157	45	107	6.5	72	178.5	24	180.5	10	11.5	668694
SPR/LHBC 187	3871	60474.5	256249.5	488.5	309.5	16044	6169	4742.5	99.5	222.5	610	29401	25.5	31.5	74	9.5	93.5	172.5	7.5	174.5	10.5	13.5	622644
SPR/LHBC 188	17538	59678	126430	1078.5	1223.5	13281.5	21905	3903.5	73.5	146.5	986	36105	137	45.5	93.5	7.5	74	159	19	124	7.5	11	716974
SPR/LHBC 189	3530	55322.5	295471.5	320.5	370	15642	3548	4754	84	106	348.5	27429	19	23.5	55.5	9	89	156.5	7.5	179.5	10	10.5	594280
SPR/LHBC 190	4234.5	64521.5	254058.5	393	262	9328	4107.5	3114	68.5	96.5	396.5	22225	25.5	18	43	6	70.5	148.5	20.5	97	<LOD	9	636741.5
SPR/LHBC 191	3903	63119	261322.5	503.5	443.5	8959	7342.5	3053.5	76	172	644	21603.5	25.5	17.5	54	6	69.5	161	18.5	82.5	<LOD	9.5	628404
SPR/LHBC 192	3885.5	59112	272786.5	454	439	13533.5	5007	4193	75	160	401.5	24844	23	26.5	50	8.5	86.5	173	17.5	172	8.5	10.5	614530
SPR/LHBC 193	5583	60142	240285.5	643.5	617.5	15600.5	12340.5	4628	124	126	1238	27708.5	34.5	32.5	105.5	7	89.5	190.5	8	166.5	8	13.5	630307
SPR/LHBC 194	4330.5	53852.5	280845.5	468	558	15126	6334.5	4306.5	78.5	144	531	26179	19.5	24	56.5	8.5	85	154	7.5	183.5	7.5	11	606689.5
SPR/LHBC 195	4160	59731.5	251021.5	610.5	714	9499	8267	3436	77.5	107.5	834.5	21788.5	26	25	64.5	6	69.5	167.5	11.5	94	<LOD	11.5	641350.5
SPR/LHBC 196	4874	62425	273840.5	217	324.5	12778	3948.5	4151.5	65.5	98	288.5	27187	18.5	18.5	48	8.5	76	130	5.5	138	5.5	9	609331
SPR/LHBC 197	4830	55601.5	242523.5	370.5	581.5	12139	9161	3650	72.5	79.5	866	25957	20.5	22	63	9	73	154	6	125	6.5	12.5	643677.5
SPR/LHBC 198	5175	65569.5	259299	606.5	456	15562	9023	4511.5	106	144	968.5	28696	29.5	33	90.5	8.5	91	184.5	6	184.5	8	11.5	609231
SPR/LHBC 199	5147	55288	261423	461.5	594.5	12939	5581.5	3887	81	94.5	460	23900	24	26	55.5	8	81.5	166.5	14	138.5	6	10.5	632187.5
SPR/LHBC 200	4287	54717.5	253039	1289	666.5	13486.5	13762	4332.5	85.5	92	1012	24203.5	30	35	115	5.5	81.5	194.5	8.5	155.5	8	12.5	630523.5
SPR/LHBC 201	5574	57501	273410.5	1057.5	516.5	14288.5	10792.5	4699	83.5	119.5	828	26929	29.5	31	96	8.5	84.5	183.5	9	170.5	9.5	9.5	606355
SPR/LHBC 202	4299	58653	281120.5	225.5	313	13682	3259	4260	80.5	99	264.5	23407.5	22.5	27	49.5	8	80	163.5	17	200	7	10	611902.5
SPR/LHBC 203	4093	51232.5	258169	783.5	1105.5	11038.5	5867	3593	67.5	94	254.5	32239	35	29	75	10.5	67.5	227	6.5	121.5	6.5	6.5	630869
SPR/LHBC 204	4391	45819	278870	158.5	448.5	10568	6010	3188	55	91	269	21014	16.5	19.5	40.5	7	66.5	161.5	5.5	124.5	5	6	628665.5
SPR/LHBC 205	4465	60911.5	285809	<LOD	230	12179	1683.5	3688.5	67	92	165.5	24586.5	14	19	48.5	9.5	80	123.5	4.5	148.5	8	6.5	605703.5
SPR/LHBC 206	3296	41965.5	249067.5	422	786.5	9418.5	12147.5	2916.5	37.5	80	448	20755.5	19.5	25.5	46.5	6.5	58.5	177.5	4	106	4.5	8	659851
SPR/LHBC 207	4602	52909	244813.5	1235.5	760.5	13213.5	16218.5	4129	78.5	96	1015.5	23778	28.5	29.5	123	6.5	79	208.5	8.5	151.5	6.5	11.5	636490.5
SPR/LHBC 208	3739	51420	268196.5	724.5	854	11945.5	10525	3830	77	88	443	25732.5	28	24	72.5	7.5	74.5	174	8	151.5	6	12.5	623729.5
SPR/LHBC 209	4325.5	54181.5	269260.5	743	594	11471.5	7435.5	3768.5	79	86.5	510.5	28017.5	19	29.5	55	9.5	94.5	166.5	6	141.5	6.5	10.5	618988.5
SPR/LHBC 210	5465	53147	265086.5	888.5	834.5	11915	8244	3734.5	73.5	93	607.5	25662	15	36	73.5	10	92.5	169.5	6.5	139.5	5.5	9.5	623691
SPR/LHBC 211	<LOD	53173	272723.5	888	758.5	12353	8147.5	3694	76.5	94	590.5	26453	20.5	29	67	10.5	96	173.5	7.5	148.5	6	11	620476.5
SPR/LHBC 212	4602	49511	263819.5	535	621.5	11777	7130.5	3623.5	66	93	294	26204	25.5	26	68.5	9.5	69	162	7	140.5	7.5	8.5	633499
SPR/LHBC 213	4329.5	59682	302335.5	313.5	327	11072.5	<LOD	3619.5	56.5	104	82	28512	21	32	43	10.5	75	86.5	7	125.5	7.5	9.5	589138
SPR/LHBC 214	4992	52883.5	283987	609	754	10262.5	2824	3365	47	107	173	28586.5	17	26	52.5	13	74	99	5.5	118.5	6.5	6.5	613472
SPR/LHBC 215	3610	54318	272114	670	597	10044.5	909.5	3274.5	48	105	108.5	28568	18	26	48.5	12	72	86.5	5	112	6.5	7	625231
SPR/LHBC 216	5485.5	61386	276795.5	450.5	594	10470.5	2869.5	3567.5	81.5	73.5	140.5	27847	19.5	28.5	53	13	64	146	5.5	127.5	5.5	8	609769
SPR/LHBC 217	4832.5	59427.5	270732.5	503	668.5	10698.5	4755	3544	57.5	131.5	160.5	27631	15	27.5	61	11.5	63	150	6	110	4	10.5	616398
SPR/LHBC 218	5076.5	55489.5	253569	493.5	824	10413	7488	3606.5	67.5	67	185	28570.5	22	32	78.5	10.5	65.5	158	6.5	112	6	11	633645

Table B.3: Marine terrace pXRF data.

	Concentration (ppm)																						
Sample ID	**Mg**	**Al**	**Si**	**P**	**S**	**K**	**Ca**	**Ti**	**V**	**Cr**	**Mn**	**Fe**	**Ni**	**Cu**	**Zn**	**As**	**Rb**	**Sr**	**Y**	**Zr**	**Nb**	**Pb**	**LE**
Terrace 1	2848	55983.5	276424	827	770	13370.5	6848.5	3208.5	74.5	105	324.5	17292.5	15.5	18	34.5	5	72.5	152	8.5	147	3	9.5	622877.5
Terrace 2	3383	62531	298972	595.5	484.5	14418	5176	3709.5	80.5	132	275.5	19079	18	16.5	32	5	75.5	147.5	8.5	158.5	2.5	9.5	590676.5
Terrace 3	2883.5	63260.5	300117.5	529	376.5	14393	5068	3250.5	79.5	125	279.5	17760	19.5	18.5	31.5	5	73	146.5	10.5	160	2	13	591389
Terrace 4	2868	67392.5	311085	319.5	288.5	15024.5	4580	3661	85.5	126.5	321	18826	18.5	17	30.5	5.5	73	150.5	14	137.5	2	10	574948
Terrace 5	4431	65085	327205.5	<LOD	<LOD	13058.5	3092.5	3698.5	86.5	114	162	18150	18.5	16.5	26	5	61	144.5	11.5	143	2.5	8	564469.5
Terrace 6	<LOD	53349	286619	790.5	895	12149.5	5531.5	3401.5	55.5	102.5	219.5	20336	16	19	39.5	6.5	65	152.5	7	175.5	6	10	616050.5
Terrace 7	<LOD	56797	306927.5	394	535.5	12141.5	3973	3540.5	69	113	196.5	22789	15.5	19	33.5	7.5	66	151	7.5	166	5.5	11	592036
Terrace 8	3366	63080	319454.5	334.5	424.5	13136.5	3614	3555	83	136	218.5	21963	17.5	20.5	34.5	7	71	155.5	6.5	207.5	5	11	571764
Terrace 9	3126	59923.5	302589.5	253.5	292	12737.5	3721	3938.5	69	114	222.5	22197.5	16.5	21.5	33	7	71	154	6.5	195	6	12.5	590292
Terrace 10	3983	67926.5	320764.5	<LOD	71	9675.5	1283	3209.5	75.5	158	116.5	19970	16	16	26	6	60.5	131	3	120.5	2	8	572403.5
Terrace 11	<LOD	58165	288486.5	1102.5	723	15054.5	6929.5	3785.5	79.5	133	392.5	19029.5	14	18.5	36	4.5	73	178	7	212.5	4	17	605550
Terrace 12	<LOD	63475.5	314173	506.5	412.5	15185.5	5055	4056	82.5	158.5	377.5	18579.5	15.5	18.5	32	6.5	75.5	175	7.5	210	5	11	577370
Terrace 13	3316	66327.5	314304	416.5	311	14499	4567.5	4164.5	84	121.5	392.5	20398	17.5	19.5	37	6.5	77	169	7.5	186.5	5.5	10.5	570549
Terrace 14	2653	65982.5	301456.5	431	275	14748.5	5411	4081	82.5	200.5	473	22199	19.5	20.5	39	7	83.5	179	8.5	175	7	12	582771
Terrace 15	3808	69096.5	298015	331	219	14357.5	5330.5	4294	89	135	584	22228.5	20	20.5	35	7	84.5	178	8	206	5	9.5	582830
Terrace 16	2566	61592.5	300192	1138	526	14683	4664.5	1968	68.5	59	218.5	11555.5	22	7.5	26	3	60	151	3	60.5	<LOD	8	601695.5
Terrace 17	2955	59441.5	290033.5	639.5	481	14506	6861.5	4106	77.5	112.5	433	21759	18.5	22	40.5	7	85.5	181.5	7.5	198	5.5	11	599487.5
Terrace 18	2825	64739	301759.5	631.5	385.5	14416	6973.5	4131.5	92	139.5	423	22192	18.5	22.5	41	7.5	86	185	9	179	5.5	11	580714.5
Terrace 19	3004	65579.5	298716.5	452.5	298.5	14964	6227.5	4086.5	87.5	123.5	430	23672	20.5	23.5	43	7	88	184	7	194	6.5	10.5	583269.5
Terrace 20	3833	66138.5	321795	92	224.5	13219.5	3651.5	4175.5	76	199.5	224.5	25561.5	19.5	24.5	37	8	76.5	165.5	7	189.5	7.5	11	560256
Terrace 21	2553	61193.5	290030.5	1253.5	906.5	14336.5	6222.5	3268.5	77	86	361	15712.5	16	12	31.5	4.5	66.5	139.5	5	108.5	<LOD	8	604873
Terrace 22	<LOD	61338.5	299970	809	584	15587.5	5426	3348	65	144.5	341	15931	16	13	30.5	4.5	68.5	154	6.5	151	2.5	9.5	595992
Terrace 23	2650	65557	298611.5	648.5	334.5	14154	4357	3317.5	79.5	264	356	15757	18	13	29	4.5	65	143	4.5	124.5	<LOD	10	593487.5
Terrace 24	3307	67399	304801.5	554	177	13880.5	4289.5	2925	76.5	89.5	293	13392.5	21.5	8.5	23	3.5	64.5	137.5	3.5	94.5	<LOD	7.5	590088.5
Terrace 25	2351	71387.5	299394	229.5	<LOD	14450	2492.5	2308.5	69	70.5	185	11538.5	19	7	19.5	4	61.5	133	2	68	<LOD	6.5	596366
Terrace 26	4310	65575.5	296008	1003	606.5	14144	4639	2324	65.5	108	236	12622	20.5	6	28.5	3	59	134	3	76.5	<LOD	8	600162.5
Terrace 27	<LOD	72000	296729.5	911	552.5	13923	4668.5	2712.5	74	79.5	280	14778	26	7.5	27.5	4	62.5	148	4.5	85	<LOD	8	592906.5
Terrace 28	2954.5	74236.5	295037	496.5	205.5	15232.5	4580	2947	81.5	152	247	14113	27.5	7	25	4	63	161.5	3.5	62	<LOD	9	589345
Terrace 29	<LOD	119085	258231	241	<LOD	10109	2685	2511.5	68.5	94.5	828.5	19178	48.5	7	26	6.5	56	155.5	3	67	<LOD	10	586575
Terrace 30	<LOD	108531.5	252255	200.5	<LOD	8929.5	2752	2316.5	60	65.5	586	18358.5	47.5	10	23.5	7	52	152.5	4.5	93.5	<LOD	8	605536.5
Terrace 31	3285	59152	271364.5	1328	848	13727.5	7322.5	3571.5	78.5	116.5	495	18828.5	16.5	18	45	4.5	76	161.5	8	136.5	3	10	621032.5
Terrace 32	3549	64850.5	293624.5	819	546	14647.5	6200.5	3615	80.5	165	467.5	18255	19.5	18	38.5	5.5	76.5	159	6.5	163.5	3	10.5	594442
Terrace 33	<LOD	67417.5	288371.5	814.5	365	14963	6066	4109	97.5	153.5	542	20644	20	21	44.5	6	80	168	9	184.5	5	12.5	595893.5
Terrace 34	2770	66624.5	292105.5	736.5	304	13996.5	5591.5	3717	81.5	131.5	531.5	19444.5	21	15.5	37.5	5.5	77	151.5	7	113	3	9	593514.5
Terrace 35	2828.5	66254	305954	379.5	171.5	14696	4310	3550	87.5	93	400	16552	19.5	12	27.5	3	68.5	145.5	7	87.5	2.5	8.5	584330.5
Terrace 36	3845	60249	291956	750.5	590	14687	7347	4006	79	120	453.5	22812	18.5	22	45	7	84	184	8.5	187.5	5.5	12.5	594444
Terrace 37	2829.5	69701	289240	1002	477.5	14579.5	5022.5	2711.5	75.5	133	238.5	14858.5	22.5	7.5	30	5	58.5	161.5	4	89.5	<LOD	7.5	598735
Terrace 38	3448.5	73964	304788.5	1011.5	314	13692	4203.5	2497.5	71.5	133.5	226	12883.5	25	9	27	3.5	56.5	137.5	3	92	<LOD	7	582388.5
Terrace 39	3897	77476	296376	809.5	224	15115	4607	2996	77.5	86	235.5	15093.5	30.5	10	27	4.5	62.5	153	3.5	73.5	4	9	582618
Terrace 40	4046	81229	301261	149	102	13881	4473	2727	80	96.5	193.5	16286	35.5	10	29.5	5.5	58	165	4.5	97	<LOD	8.5	575052.5
Terrace 41	<LOD	61187.5	274504	1245	670	14231	6279	3131.5	76.5	114.5	444.5	17126.5	19	16	37.5	4	70.5	131	5.5	70.5	<LOD	8.5	620620.5
Terrace 42	3392	70870	291289	999	575	14455	6272.5	3766.5	81.5	217.5	570.5	21491	25.5	22	50.5	6	84.5	179	10.5	158.5	4	11	587155
Terrace 43	2944	72909.5	299505.5	1047.5	443	14075	5704.5	3845	80.5	116	534.5	19840.5	25.5	19.5	44.5	6.5	78	150	8	147.5	2	9	579920.5
Terrace 44	3001	76474	278338.5	1016.5	292	12001	5636	3242.5	86.5	99.5	462.5	19001	24	15.5	35.5	5	71	145.5	8	94	<LOD	8.5	599928
Terrace 45	3339	76407.5	300394.5	478	165.5	11774	4001	3182.5	77	105.5	323.5	16499.5	20	12	25.5	4.5	61.5	141.5	6	92	<LOD	8	584535
Terrace 46	2591.5	70363	278320	1038.5	616	15730.5	4671	2074.5	77	52.5	206.5	12721	24	7	24.5	4.5	60	151.5	3.5	47.5	<LOD	8	611196
Terrace 47	<LOD	71517.5	294877	759.5	525.5	14985.5	3325.5	2025.5	68	73	208	12708	21	7.5	23	4.5	58	145	4	58	<LOD	7	598586
Terrace 48	<LOD	75001	291071	559.5	215	14800.5	2457	1984.5	85.5	66.5	179.5	11678	24.5	5.5	19	4	56.5	152.5	2.5	43.5	<LOD	7	601570.5
Terrace 49	<LOD	90762	274339	386.5	84	10708.5	2046	1965	68	45.5	165.5	13274.5	29	6.5	19.5	4.5	54.5	151.5	3	59.5	<LOD	7	605805
Terrace 50	<LOD	86904	287313	132.5	45	13655	2233	1763	75.5	43	145	12623.5	33.5	6	18.5	6	60.5	159	2	47	<LOD	6.5	594737.5
Terrace 51	4289	62026.5	260085.5	1340	987	11462	7294.5	2766.5	65	73	362.5	16389	28	12.5	39.5	5.5	62.5	156.5	10	89.5	<LOD	6.5	634583
Terrace 52	<LOD	59350	278607	1248	504	12249.5	3998	2185.5	73.5	82	282.5	14424	23	9.5	34.5	4	66.5	150.5	9.5	68	<LOD	8	626604.5
Terrace 53	2961	71694	279667	1257.5	472	11005	4606	2864.5	81	88.5	366	16779	26	12.5	36.5	5	66	167	11	78	<LOD	8.5	607727.5
Terrace 54	3120.5	71417	279157.5	857.5	313.5	11685	4479	2527.5	82	137.5	551	16608	29.5	10.5	34.5	6	63	153	13	62.5	<LOD	9	608673.5
Terrace 55	<LOD	64283.5	314294	134	68	11693.5	4074.5	2923.5	82	136.5	387.5	14354.5	18.5	7	22.5	5	57	160	9	67.5	<LOD	7.5	587198
Terrace 56	3018	58922	291739	1267	777	14596.5	4345.5	1772.5	66.5	80	212	9818	18	6	23	3	59	134.5	2.5	47	<LOD	8	614581.5
Terrace 57	<LOD	62915.5	290855	1023	472.5	14496.5	3526.5	2037	64.5	56	213.5	11577.5	20	6.5	23.5	3	60.5	131	3	43	<LOD	11.5	612453.5
Terrace 58	2589	65443.5	304342	931.5	356	15125.5	2752	2166.5	73.5	53	192	11034	20	5.5	20.5	3.5	67	141	2.5	51.5	<LOD	10	595904
Terrace 59	<LOD	69587	313533	296.5	86	16038	2288	1986	83	62	133	10460.5	22	4	19.5	4.5	67	155	2	44	<LOD	9	585100.5
Terrace 60	<LOD	70173.5	303042	239.5	56	12622.5	2348.5	2020.5	70.5	61	131.5	11512	19.5	5	15	3	55	148	2	45.5	<LOD	7	597434.5
Terrace 61	3024	61916.5	292682	992.5	672	14720.5	6601.5	4139	86.5	148	481.5	20424	16.5	21.5	45	4.5	79	178	7	185	3.5	17.5	595051.5
Terrace 62	3873	64980	300799.5	662	443.5	14786	5460	4698	89	200.5	482	21346	20	21.5	42	5	78.5	171.5	8.5	204	4	16.5	583525.5
Terrace 63	2474	66071	295596	636.5	331.5	15405	6007	4121.5	90	118	532	22150.5	22	22.5	46.5	7	86	174	8	183	6	11	587134
Terrace 64	<LOD	66614.5	311662.5	337	201.5	16319.5	4653.5	3950	89.5	127.5	425.5	18506	18.5	16.5	31	4	79.5	156	6	167	4.5	10.5	576607.5
Terrace 65	<LOD	62330.5	324193.5	<LOD	157	15348.5	3339	3192.5	82	88.5	196.5	13802.5	14	9	22.5	3.5	61	142.5	4.5	103	<LOD	7	576889
Terrace 66	<LOD	52254	257919	1438	992	14767.5	8293	3487.5	66	234.5	452	17384.5	16	15.5	44	3.5	71	162	7.5	149.5	3.5	15	642219.5
Terrace 67	3324	62815	290501.5	1069	678	14747.5	6870	3403.5	78	100	410	18290	18	18	39.5	4	74	157.5	7	167	2	10.5	597201
Terrace 68	3455	62947	287468.5	823.5	374.5	14422.5	5997.5	3920.5	85	144	519.5	20338.5	20.5	17.5	43.5	5.5	79	162	9	166.5	5	13.5	600704.5
Terrace 69	3217.5	69260.5	285341	941	365	14696	7080.5	3912.5	86.5	151	572.5	21498	26.5	20	48.5	6.5	82.5	159	10.5	139	4	9.5	592359.5
Terrace 70	2651	69356	299812.5	485	189	15273.5	5431	3665	80	96.5	479	17557	22.5	15.5	32	4	74	149.5	9	111.5	2	10.5	584485.5
Terrace 71	<LOD	62788	307104	640.5	424.5	16137.5	4109	3282	81.5	88	343.5	14370.5	15	12	27.5	4	68.5	143.5	4.5	114	2	8.5	590222
Terrace 72	<LOD	65513	312563.5	516	397.5	16109.5	3718.5	3377.5	80.5	105	356	15149.5	18	11.5	28	3.5	70	146.5	5	143	3	10.5	581664.5
Terrace 73	2572	68972	310838	358	233.5	17395.5	3854	3487	84	190	385	16221.5	18.5	14.5	30.5	4	76	152	5.5	138.5	2.5	10	576233.5
Terrace 74	<LOD	67493	315904.5	<LOD	54	16477	2760	3248.5	85.5	101	261.5	12981.5	16.5	7.5	20	4	70	142.5	3.5	101.5	<LOD	6.5	580274
Terrace 75	<LOD	67925.5	322208	<LOD	<LOD	17346.5	1724	2477.5	84	80.5	149	10556	16	7	17	2.5	73	136.5	4	85.5	<LOD	10	577081.5

Terrace 76	3309	59761	294788	611	701.5	13675.5	4796.5	3456	67.5	146	346.5	18543.5	23	16	39.5	5.5	68.5	152.5	9.5	168.5	3	10	600945.5
Terrace 77	<LOD	63701.5	311262	234	408	14233.5	3482.5	3768	88	99	313	19401.5	23.5	18.5	35.5	5.5	72	158	10.5	173.5	4.5	13	582488
Terrace 78	3640	68938.5	314989	<LOD	202	13349	2974.5	3466	81	109.5	224.5	20314	23	13.5	30.5	6	68	151	8.5	144.5	3	12	573060
Terrace 79	4253.5	88591	278630.5	<LOD	49	10454	857	3461	83	97.5	116.5	26101	26	13.5	32	6	59.5	127.5	5.5	116	<LOD	11.5	586885.5
Terrace 80	2709	78138	291130.5	<LOD	227	10076.5	232.5	3586	70	297	99	25794	21	14.5	30.5	7.5	60	127	9.5	102	3.5	13	588594
Terrace 81	<LOD	62302	283581	985.5	618	10230	7175.5	3589.5	74	235.5	406.5	19353	23	11.5	34	5	45	169	6.5	177.5	<LOD	9.5	610953.5
Terrace 82	2611	70749.5	290098	746	425	10208	6529.5	4016.5	77	176	484.5	21179.5	24	11	31.5	6.5	46.5	167.5	8.5	167.5	<LOD	9	593516
Terrace 83	<LOD	68478	269483.5	831	250.5	8939.5	6515	3866	80	321.5	611	22059	29	10.5	37	5.5	50	165	10	180.5	2.5	10	618053
Terrace 84	2372	78257.5	280435.5	275	109	8781.5	5199	4025.5	77	238	594.5	21541	24	11.5	30	5.5	45	156	8	175	<LOD	8	598860
Terrace 85	3048	91105	235995	<LOD	<LOD	4939.5	5311	3617	71	150.5	366.5	25232	37.5	10.5	34	5.5	36	178.5	8.5	97	2	7	631260.5
Terrace 86	<LOD	61961.5	292533	1311	746.5	9772	6995.5	3793.5	65.5	193.5	387	17422	20.5	12.5	34.5	5	46.5	155	6	172	<LOD	9	604344
Terrace 87	<LOD	59024	286425.5	1130.5	662	10805	7355.5	3365.5	71.5	195.5	379	16981.5	20	9.5	34	4.5	45.5	150.5	6.5	150.5	3	8	613164
Terrace 88	<LOD	62646.5	292453.5	890.5	404.5	10441	6437.5	3814	77.5	349.5	368.5	17340.5	17.5	13	32	4.5	51	161.5	5.5	220	<LOD	8.5	604246
Terrace 89	2633	72451.5	262218	496	128	8550	5401.5	3456	78.5	215	410.5	23070.5	29.5	9	28	4.5	42.5	164.5	6	100.5	2	10.5	621799
Terrace 90	<LOD	99006	219915.5	<LOD	83.5	4509.5	4551	3370.5	76.5	98	176	32571.5	46.5	11	35	6	35.5	174	6.5	62.5	<LOD	8.5	635242.5
Terrace 91	<LOD	54476	260226.5	1266	919	8974.5	6628	3286.5	59	187.5	236	18610	19	10	36	5	42.5	145.5	5.5	213	2.5	10.5	644628.5
Terrace 92	<LOD	64740	279380.5	1317	658.5	10333.5	6235.5	3431	61	174	236.5	20325	20.5	11	33	4	45.5	143	5.5	217	<LOD	9.5	612606.5
Terrace 93	<LOD	72573.5	270680.5	938.5	291.5	8063.5	5574.5	3937.5	75.5	316.5	216.5	25049.5	23	9	34.5	5	45	132	5.5	170.5	<LOD	10.5	611831.5
Terrace 94	3038	84052.5	228277	<LOD	<LOD	4852	3652	3789.5	60	213.5	85.5	38159.5	37	6.5	29	6.5	32.5	117.5	6.5	136	5	8	634941
Terrace 95	4087.5	85547	245828.5	<LOD	<LOD	4487.5	3885.5	4105.5	74.5	251	105.5	32687.5	35	6	30	3	30.5	146	7	182.5	2.5	20	618416.5
Terrace 96	<LOD	60981.5	259312.5	1300	986.5	7552	6491.5	3055	63.5	319	232	20861	20	11	30.5	5	40.5	146	5	209.5	<LOD	7.5	638356.5
Terrace 97	2842	69417.5	258788	1026.5	555	7787.5	5986	4070.5	72	257	245	22219.5	22.5	8.5	30	4.5	39	148.5	5.5	212	<LOD	9	627663
Terrace 98	<LOD	85757	237737.5	531.5	208	5486.5	3386.5	3211	72	121	177	29776	30	8	30	5.5	37	130	5.5	93.5	<LOD	8	633172
Terrace 99	3499	96042	222676.5	<LOD	<LOD	4057.5	2844	3509.5	71.5	134.5	196.5	33124.5	36	7	28.5	6	35	143.5	8	94	<LOD	7.5	635218
Terrace 100	4394	91420	208358.5	<LOD	<LOD	4464.5	4945.5	4472	89.5	314.5	289.5	38223.5	41	11.5	32.5	7	34	203	14.5	146.5	2	8	642517.5
Terrace 101	2932.5	59772.5	272459.5	518	906.5	7353	4352.5	3451.5	66.5	178	161.5	20287	14.5	15.5	35.5	4.5	48.5	128	8	217.5	2	11	627060
Terrace 102	4227	64058	274746.5	<LOD	305	7514.5	3141	3245	75.5	190.5	112	23027.5	17	14	31	5.5	51	127	8	194	2.5	9.5	618882
Terrace 103	4050	69548	288079.5	<LOD	236	6403	2851.5	3824.5	106.5	140	111	26711	21	14	33.5	6	47	142	7.5	175	4	10.5	597466
Terrace 104	3241	67258	299917	<LOD	265	5478.5	1875	3825.5	116.5	133.5	86.5	26854	22	14	30.5	8	46.5	146	7.5	145	3	7	592129.5
Terrace 105	4574	66357	296627	<LOD	320.5	6001.5	1269.5	3617.5	92.5	124.5	88.5	25924	28.5	12.5	28.5	6.5	43	151.5	7	104.5	4	9	596883
Terrace 106	3191	47893	234357	940.5	1483	7964	6822	2792.5	60	136	151.5	17699	16.5	11	38.5	3.5	43.5	141	8	118	<LOD	8.5	677703.5
Terrace 107	3191	59907.5	275837.5	136	453	7715	4200	3028.5	67.5	115	118	20893.5	20	10	32	5.5	45.5	141.5	8.5	133.5	2.5	8.5	623988.5
Terrace 108	3626	67186.5	278511	<LOD	94.5	7380	3638.5	3219	64.5	183.5	95	24736.5	25	9.5	27.5	7	43	155.5	8	132.5	2	8	610833
Terrace 109	2407	68279	291510	<LOD	74.5	9514	4491.5	3428.5	95	104.5	105	17807.5	28	7.5	24.5	4	48.5	208.5	8	108	<LOD	9	602930
Terrace 110	2884.5	77716.5	278924.5	<LOD	<LOD	8340.5	1776.5	2547	68	126	79.5	16038	24.5	7	21.5	3	59.5	161	12	46	<LOD	7	611143.5
Terrace 111	3473.5	51641.5	259603	1064.5	1353	9609.5	6439	3181.5	56	140	306.5	18006	12.5	13.5	42.5	3.5	56.5	132.5	7	185	2	10.5	644644
Terrace 112	2422	57949.5	284760.5	612.5	553.5	10180.5	4516	3917.5	68.5	191.5	214	21203.5	16	13	31	7	58	139.5	9	283	4.5	10	614040.5
Terrace 113	3659.5	63993.5	299597.5	256.5	274.5	10363.5	4142.5	3852	78.5	201	183.5	22161	15	12	30	6	59	140	8.5	245	3	11.5	590693.5
Terrace 114	2950	70590	274448	<LOD	86	5838	1863.5	3543.5	75	145	69	27766.5	20.5	11	31	4.5	45	118	9.5	160	6	27	612150
Terrace 115	3284	70556.5	286541	<LOD	121.5	5662	756	3541	51	144	63.5	27326	18.5	15	29.5	6	43	114	5.5	164.5	5.5	16	601517
Terrace 116	3581	45122	218193	1250	1875	7255.5	6620	2706	43.5	171	209	17468.5	13	11.5	42.5	4	43.5	125	7	143.5	<LOD	7	696884
Terrace 117	3408	64606.5	294441.5	<LOD	287.5	8116.5	3660.5	3180	69	252.5	116	22758	16	10	28.5	5.5	47.5	140	9	158.5	2.5	11	598661
Terrace 118	3827	65981.5	303290.5	<LOD	148	7663	3248	3481	78	147.5	92	23747.5	22	12	30	6.5	46	140.5	10.5	141.5	3	9	587862
Terrace 119	4075.5	63864	305485.5	<LOD	423	7586	3035	3961.5	159.5	196	132.5	25502	34.5	11.5	28.5	8	47	162	11	128.5	4.5	8	585124.5
Terrace 120	4129	68183	291953	<LOD	167.5	7947	3814.5	3652.5	97.5	182	273.5	27902.5	59	13.5	37	7.5	53.5	188	26	96	4.5	9	591196
Terrace 121	<LOD	61104.5	280579.5	1637.5	879	12479	6455.5	3749	75	249	348.5	16274.5	14.5	11	32.5	3.5	51	131.5	4.5	156	<LOD	10.5	615738.5
Terrace 122	<LOD	60635.5	290346	1395	510.5	11063.5	5334	3029.5	63.5	259.5	326.5	15616.5	16.5	12	32.5	3	46.5	123.5	6	146	<LOD	10	611014
Terrace 123	<LOD	68923.5	284081.5	1705	365.5	11532.5	5420.5	4116.5	85.5	268	352.5	16784.5	18.5	13	33	3	49	126	6	196.5	2	9.5	605893.5
Terrace 124	<LOD	76651	267723.5	1106.5	180	9580	4230.5	3643.5	72	215.5	283.5	24794.5	22.5	11	30.5	5.5	49	123	7	182	2.5	9.5	611068
Terrace 125	<LOD	111455	217132	<LOD	47	4677	1913.5	4798	85	191	189.5	41557	43	11.5	31	7.5	37	108	5.5	171	3.5	9	617516
Terrace 126	<LOD	65653	285659.5	1671.5	936.5	11715	6356	3942	84.5	220	362.5	18327.5	19.5	13.5	37.5	3.5	51.5	138	6.5	247	2	9.5	604529.5
Terrace 127	<LOD	67192.5	291443.5	1458	662.5	10597	5626.5	4115.5	76.5	366.5	384	18144	18.5	13.5	35.5	4.5	47.5	126.5	6	191.5	<LOD	8	599463.5
Terrace 128	<LOD	70702.5	293524.5	1225.5	323.5	10933	4764	4242	80.5	252	388.5	19097	22.5	11	36.5	5	51	130	7	181.5	2	8.5	593999
Terrace 129	2578	106668	228093.5	257.5	84	7911	2175	4531	92.5	277.5	240	35818.5	35	10	32	8.5	50	121	6	198	2	9.5	612073
Terrace 130	2660	115514.5	198132.5	<LOD	62	3927.5	2206	4838	84	219.5	232.5	43781.5	48	11.5	41	10	37	121	9.5	138.5	4	7.5	629235.5
Terrace 131	3335	66431	239808	1232	967.5	7677	4923	2934	59	195	161	20705.5	18.5	11	33	5.5	40.5	126.5	6	192	<LOD	6.5	652787
Terrace 132	3173	85903.5	238475	867	607.5	7036	3645.5	3693	70.5	267.5	157	25892	24	10	34	4.5	38.5	116.5	6	152	2	10	631386
Terrace 133	2809	97320	244536.5	669	326.5	6761	3937	4171	76	227	182.5	30991	31	13	39	7	44.5	125	6.5	218.5	2.5	10	608884
Terrace 134	3456	96672	205635.5	<LOD	64.5	4014	2311	3667	70	161.5	91	35518.5	33.5	7.5	30	7	32.5	114	6	73	3	7.5	649740.5
Terrace 135	3933	95677.5	222585.5	<LOD	344.5	4678	2320	3592	81.5	140.5	425	38692.5	35.5	10.5	25.5	10.5	34.5	146.5	7.5	77.5	<LOD	9.5	629127
Terrace 136	3061	60650.5	244227.5	1569.5	1154.5	7932.5	7354.5	2917.5	57.5	337.5	247	20296	19.5	11	34.5	4	43	120.5	4.5	127.5	<LOD	9	651334.5
Terrace 137	<LOD	68681.5	260048	1430.5	638.5	9710	4452	3202.5	65.5	337	227	20984.5	19.5	12	34	5.5	47.5	128.5	5.5	175	<LOD	9	629774.5
Terrace 138	<LOD	79410.5	242633	1368.5	398	8229.5	4018.5	3991	71	284	234	25340.5	24	10.5	33.5	5	43	113	5.5	185	2.5	10	633574
Terrace 139	3417	99673.5	183928	497.5	103.5	3784	2759.5	3697	74	134	311	57988	43	9.5	31.5	21.5	36.5	110	5.5	103.5	<LOD	11	644955
Terrace 140	3688	95997	200836.5	<LOD	<LOD	3933	3152	4120	90	192.5	405.5	41150	42	9	29	11.5	33	134	8.5	122.5	3	6.5	647868.5
Terrace 141	<LOD	44251.5	266753.5	544	660	9920.5	5587.5	3482.5	67.5	161	296.5	16981.5	14.5	13	33.5	4.5	60	149	8	202	4	10	650787.5
Terrace 142	2416	55168	304922.5	465.5	481.5	10525	4976	3577.5	67.5	144	282	17937	10.5	13	31.5	5	59	141.5	9.5	190.5	2.5	9	599760.5
Terrace 143	<LOD	59365.5	309135.5	234.5	320	10918	5153	3695.5	86.5	196	266.5	19190.5	14.5	14.5	30.5	6.5	64	145	9	222.5	3.5	8.5	590906.5
Terrace 144	2534	54053.5	334255.5	<LOD	134	12281	3504.5	4287	71.5	182.5	141.5	17706.5	11	13	24	7	65	155.5	10.5	248.5	5.5	11	571554
Terrace 145	<LOD	55460	350247.5	<LOD	66.5	12171.5	1451	4093.5	74.5	157	117	16620	12	10	23	5.5	59.5	147	7	233.5	5.5	9	559019
Terrace 146	<LOD	50891	290095	608.5	686.5	11229.5	6842.5	3449	63	179	382.5	17755.5	14	14.5	34.5	5	63.5	154.5	9	228.5	5	12.5	617266.5
Terrace 147	2724	56648.5	307156	326.5	468.5	10940.5	5475.5	3797	79	153.5	355.5	18669	15.5	15.5	33.5	7	66	152.5	9.5	207.5	5	8.5	594036
Terrace 148	3447	60155.5	308235	90	329.5	11012.5	5248	3801	82	229.5	354	20358.5	15.5	16.5	32	6	68.5	156	10.5	217.5	4	10.5	587875
Terrace 149	3657	60694.5	291131	<LOD	226	11393.5	3693.5	3897.5	74	155.5	247.5	24724	19.5	17.5	31	8	72	148	10.5	221	6.5	9	601386
Terrace 150	3819.5	61556.5	330725.5	<LOD	55.5	11083.5	831	3977	73	204.5	97.5	20427	14	14	25	7.5	62	132	7	258	6	9	566604.5
Terrace 151	2552.5	56212.5	294219.5	670.5	593.5	12056	5786.5	3615	68.5	128.5	352.5	20833.5	17.5	18	39	7	73	151.5	11.5	204.5	6	11	602360
Terrace 152	3214	58137	299787	435.5	489	11688.5	4999	3951	70.5	142	322.5	20578.5	13.5	17	34.5	7.5	70.5	145.5	11	206.5	4.5	8.5	595651.5
Terrace 153	2778	58453.5	299717	286	360.5	12229	4703.5	3734.5	84.5	184	312.5	22429.5	18	20	36	8	73.5	145.5	10.5	183	5.5	10	595595.5
Terrace 154	4177	59875.5	320166	<LOD	162.5	11264	2194	3642.5	69.5	135	140	22649	15.5	16.5	33.5	9	65.5	125.5	7.5	176.5	6	10.5	577143.5
Terrace 155	3007.5	57743	341170	<LOD	73.5	11192.5	822	4045.5	69	472.5	138.5	21615	15.5	15.5	31	9	60.5	127	5.5	209.5	6	8.5	559151.5
Terrace 156	2835	52659	277465.5	748	984.5	10456.5	4457	3441.5	61	142	491	21452.5	16	17.5	40.5	6	64.5	135.5	11.5	204	5.5	10.5	625707.5
Terrace 157	3128	61715	315051.5	521.5	459.5	11708.5	3413.5	3769.5	79.5	151.5	263	21861	17	18	36	6.5	71	143	11.5	195.5	3.5	10	577351
Terrace 158	3554.5	59852	312082	276	267	11581.5	3400	4098	78	208.5	219.5	22712	19	17	35.5	8.5	72	147.5	11.5	235	6	12	581092.5
Terrace 159	4671	58600.5	302986.5	<LOD	129.5	10920	3017.5	3748	72.5	150	155.5	23829	18	18	30.5	9	64	136	10	230.5	6	8.5	593525.5

Table B.4: NRCS Chico pXRF data.

	Concentration (ppm)																						
Sample ID	**Mg**	**Al**	**Si**	**P**	**S**	**K**	**Ca**	**Ti**	**V**	**Cr**	**Mn**	**Fe**	**Ni**	**Cu**	**Zn**	**As**	**Rb**	**Sr**	**Y**	**Zr**	**Nb**	**Pb**	**LE**
NRCS 1	5561	73031	228101	1340.5	558	15179.5	19267.5	3405	80	34.5	688.5	26017	26.5	27.5	56	5	59	350	11.5	117.5	7	13.5	628838.5
NRCS 2	4495	94168	226277	1020	356.5	14448.5	11932.5	3874	81.5	33	516.5	32356	21.5	36.5	60	5.5	62	268.5	13	173.5	6.5	10.5	609791
NRCS 3	8670.5	73960	236217.5	1455.5	277	14367.5	26513	3641.5	95	106.5	1104	30226	40	32	71	4.5	54.5	372.5	13	104.5	7	15.5	602646
NRCS 4	<LOD	103197.5	217855	317	174.5	13377	12847	4034	87.5	56.5	492.5	32483	32.5	42	59	5	57.5	299.5	16	142.5	9.5	14	614392.5
NRCS 5	20867	68222	205946	859.5	187.5	8575	43876	4412	88.5	282	939.5	46078	146.5	60.5	76	2.5	35	374.5	17	86.5	6	10	598850
NRCS 6	24235	76464	218438	608.5	<LOD	8069.5	52704.5	4348	81.5	314.5	897.5	48438	144	74	69.5	4	29.5	364	15.5	83	4.5	6.5	564608.5
NRCS 7	24380.5	76699	220203	640.5	<LOD	9638.5	43189.5	4402.5	95.5	362	1216	46064	161.5	56.5	77	3.5	37.5	368.5	16.5	90.5	6	10.5	572271
NRCS 8	25540.5	80798	220203.5	423.5	<LOD	8985	41142	4544	98	294	1020	46270.5	153.5	65	70.5	3.5	38.5	373.5	16.5	89	5	7.5	569848
NRCS 9	7332.5	144005.5	161996.5	746.5	145.5	3779	14367.5	7582	107	140	695	60154	102	108	83.5	5	28.5	263.5	28.5	173	9.5	14	598130.5
NRCS 10	9108	88367	121870.5	3430	1054.5	6184	23037	6515.5	90.5	101	1463.5	47406	68	53.5	113.5	4	31	394	14	121.5	8	17.5	690544.5
NRCS 11	6762.5	105840.5	147198	3064.5	445.5	7168	20723.5	8432.5	106.5	77	978	58102	59.5	70	122	5	31.5	443.5	17	167	15.5	14.5	640157.5
NRCS 12	7203.5	83478	125324	2778	1075.5	7694	22500	4966	91	105	1654	43701.5	78	53	105.5	2.5	33.5	281.5	12.5	104	7	11	698740.5
NRCS 13	6672.5	112182.5	127701	2494	923.5	5423	16520.5	6171	91.5	113.5	1173.5	55980	93.5	55.5	100.5	4	25	221.5	15.5	130	10	8	663887
NRCS 14	5258	78214	183941.5	1195.5	1332	12466	6785	3677	76	31	573	25659.5	15.5	22.5	47	22.5	60	121	15	149.5	6.5	26	680308.5
NRCS 15	4778	104928.5	243487	353.5	222	22545.5	<LOD	4844	105	30.5	72.5	21812.5	10	13.5	28	53	109	135	16	200	8.5	4.5	596222.5
NRCS 16	3715	75945.5	272522	461	249	17495.5	18546.5	3305.5	77.5	34	450.5	22272.5	23	21	43	5	60	365.5	12	135	8.5	12.5	584229.5
NRCS 17	3821	91828.5	277333.5	871	411.5	17580	17283.5	3505.5	91	<LOD	527	26347	19	21.5	51.5	5	62.5	375	13.5	172	8.5	13	559651.5
NRCS 18	4735	65624	216259	797.5	614.5	14482	20030	3129.5	70	50.5	588	25331.5	33	29	47	4.5	52.5	327	11	127	7	13.5	647639
NRCS 19	<LOD	106947	223068.5	1034	543.5	13800	12884.5	3209	76	27	521	24392	14	16.5	44	4.5	57	325.5	11	159.5	8.5	11.5	612854
NRCS 20	4013.5	94955.5	275991.5	304.5	182.5	17006	18049	3480	90.5	<LOD	593	25877.5	17	19.5	51.5	4	62.5	370	10	149	7.5	11.5	558750
NRCS 21	5070	83079	237007.5	852	272.5	15686.5	18968.5	3446.5	106.5	46	760	23084	21.5	26	49	5.5	62.5	370	14	147.5	8	14.5	610926
NRCS 22	<LOD	96525	229259	631	239	13974	16217	4163.5	94	28.5	606.5	29120	19	26.5	53.5	6.5	57	361.5	20	172.5	9.5	15.5	608391
NRCS 23	3325	66235	260968	503.5	383	19067	21391.5	2650	88.5	30	523.5	19386	17.5	20	43	3.5	60	369	10	88.5	5.5	18	606489
NRCS 24	2494	72516.5	291366.5	<LOD	<LOD	17954	21193	2258	88	<LOD	388.5	13372	9	13.5	27.5	3	64.5	414.5	9.5	103	6	13	578944.5
NRCS 25	7523.5	84302.5	264217	388.5	<LOD	15380.5	36300.5	4395.5	104.5	50	595	30971	37.5	52	53	5.5	62.5	486.5	19.5	191	9	13	554836.5
NRCS 26	5514	74110.5	260196.5	1174	219	17211	20754.5	3012.5	84.5	27	576.5	21830	18.5	23.5	43	4.5	65	376	12.5	127.5	7	11.5	594609.5
NRCS 27	4607	75271.5	248095.5	605	155	16539.5	20550	3195.5	81	31	560	23383.5	20	32.5	45.5	4.5	61	387	13	125.5	7.5	12.5	606219.5
NRCS 28	5790.5	79137.5	266251.5	531	171	15895.5	20560	3022	91.5	35	551	20383	16.5	30	39	6	55.5	411	13	109	6	11	586879
NRCS 29	4341	58632.5	209590.5	854	1002.5	14480.5	26146	2675.5	81	29.5	669	19871.5	15.5	20	38	2.5	58.5	378	11.5	103	6.5	16.5	663144
NRCS 30	5878	80981	275882.5	418.5	137	18035.5	19642	3068.5	92	35	571	21225	17	23	42.5	6.5	63	377	13.5	129.5	7	11.5	573357
NRCS 31	4753.5	88189	278531	255.5	49.5	17543.5	16352	3081	103	27	527.5	23990	14.5	26	42	4	58	336.5	13	109.5	7	11	565976.5
NRCS 32	10766.5	56964	252496.5	1360	1324.5	7516	10770	4466.5	61.5	430.5	621	30148	152.5	49.5	109.5	4.5	33	117	16	126.5	2	13	622437
NRCS 33	14346.5	71188.5	289399.5	254	164	7577	7963.5	5026.5	83.5	597	684.5	35196.5	191	44	75.5	6	33	104.5	19	135	2.5	9	566882
NRCS 34	14976	75141	275882.5	271	52	9798	8846	5346.5	90.5	610.5	844.5	42842	262	55	90.5	7	38.5	115	20.5	132.5	4	6	564554.5
NRCS 35	17916	60686.5	239470.5	813.5	980.5	8264	12772.5	4322	85	884.5	747	41301	242	56	124.5	6	38.5	135	16.5	102	3.5	7.5	611016
NRCS 36	20185.5	68895.5	270638.5	282	60	8008.5	13340	4614	91	1193.5	907.5	45259.5	286.5	53	84	7.5	31	112.5	18	101.5	4.5	5	565845.5
NRCS 37	10833.5	56101	226435.5	3183	1629	9666.5	14104	3464.5	60	389.5	609	32442.5	137	55	123.5	4	35.5	150	15	100	<LOD	8.5	640442.5
NRCS 38	14248.5	78830	243461	757.5	321.5	11372	7040	4556	86	328	475.5	45632.5	153	55	74	9.5	37	91	19	114.5	4	7.5	592323.5
NRCS 39	16597	77746	230842.5	<LOD	134.5	8487.5	7634.5	4671.5	77.5	229.5	603.5	47903	191.5	52.5	71	10	34	92.5	23.5	128.5	4.5	7.5	604456
NRCS 40	22891	74988	237953	<LOD	48	8186	7972	4644.5	91.5	217	841.5	48450.5	188.5	60.5	87.5	9	39	95	27	128	5	8	593092.5
NRCS 41	18968	78168	255489	581.5	243.5	11043	11059.5	4783	93	181	783	43892	110.5	63	128	10.5	44	163	19.5	142.5	4	9.5	574012.5
NRCS 42	17291	74472	251340	542.5	150.5	10576.5	10413.5	4653.5	96.5	184.5	783	43950	119.5	62	121	9.5	44	156	21	125	6.5	9.5	584871.5
NRCS 43	16785.5	70114	263845	223.5	66	10782	12062.5	4033.5	86.5	177.5	631	38066	92	49.5	143	9	39	197.5	15	93	4.5	9	582508.5
NRCS 44	22149.5	70297	214858	1209.5	1102.5	12762.5	6794.5	4456	83.5	203	691.5	39517	113	42.5	114.5	7.5	42.5	69	15	101	3	7.5	625347.5
NRCS 45	23468	76922.5	251652.5	551	80	12925	3250.5	4675	96.5	257.5	643	42605.5	136.5	47.5	97	8	42	50	16	104	3.5	6	582352
NRCS 46	23557	58324.5	316218.5	215	<LOD	6449	3258.5	2977	62.5	335	510.5	30222	89	23	59.5	5	18	31.5	8.5	51	<LOD	3	557563
NRCS 47	19794.5	78804	252707	751.5	359	10918.5	12968.5	4540	100	337	697.5	40801.5	99	88.5	139.5	9	40.5	203	17.5	114	<LOD	8	576486
NRCS 48	18998	76985	243390.5	467.5	108	10118.5	11692.5	4700	102	184.5	733.5	42967.5	105	96.5	142	11	42	193.5	18.5	123.5	4.5	8.5	588800
NRCS 49	18328	78819	242094.5	715	273	12967	9365.5	5119	93	131	862.5	49726	122.5	77.5	139	11.5	52	138	24	124.5	7.5	13	580789.5
NRCS 50	19062	78480.5	243360.5	600	96	11616	10102.5	5117	100	163	839	49183	139	68	115.5	11	52	148	23	137	7	9	580564
NRCS 51	17757.5	80881.5	234989	813.5	269.5	11809	10354	4914.5	99.5	149.5	952	51910	123.5	73.5	135.5	10.5	56	156.5	23.5	112	6	10.5	584393.5
NRCS 52	18916.5	84061.5	240116.5	290	55	10359.5	9684.5	5069.5	110.5	173.5	949.5	53324	136	72	118	13	51	158.5	23.5	124	7	9.5	576171
NRCS 53	12844.5	74002	237901	698.5	324	12578	9325.5	4898.5	90.5	141	916	44405	115	66	107	9	45	157	23	119.5	7.5	14	601211.5
NRCS 54	17279	78902	230464	<LOD	<LOD	7655	7997	4988	109.5	173	793.5	49171.5	127	69	99.5	11	46	144	23.5	126.5	6.5	8	601805.5
NRCS 55	23306.5	73739.5	241772.5	509	<LOD	10296.5	17183	5241.5	108	163.5	840	47623.5	140	65	111	11.5	46.5	179	22.5	129.5	8.5	8.5	578492.5
NRCS 56	10367.5	73999.5	261756.5	287	245.5	8391	7051	5293	97.5	157.5	921	40926.5	110	59	89	9.5	40.5	135	22	144.5	7	9.5	589871
NRCS 57	15085.5	83806	255598	<LOD	<LOD	5950.5	5487	5285.5	100.5	171	804.5	45843	125	61	83.5	9	43.5	122.5	22	151	3.5	8.5	581222.5
NRCS 58	31291	70226.5	207316.5	238	<LOD	8785	50847	4333.5	106.5	151.5	696.5	45150	135.5	61.5	91.5	9	40.5	272	18	106	4.5	7.5	580109
NRCS 59	11644	80429	234119.5	365.5	306	12633.5	14434	5543.5	102	161.5	969.5	45656	103.5	68.5	92.5	10.5	55	243.5	22	132.5	7.5	24.5	592876
NRCS 60	14401	89202.5	213591	<LOD	57	9000.5	11442	4933	87.5	148.5	715.5	55940.5	118.5	66.5	90.5	9.5	48	183.5	21	118.5	6	12.5	599831.5

Table B.5: UC Merced pXRF data.

	Concentration (ppm)																						
Sample ID	**Mg**	**Al**	**Si**	**P**	**S**	**K**	**Ca**	**Ti**	**V**	**Cr**	**Mn**	**Fe**	**Ni**	**Cu**	**Zn**	**As**	**Rb**	**Sr**	**Y**	**Zr**	**Nb**	**Pb**	**LE**
Merced Atwater	3924	85553	292602.5	246.5	314	20724.5	12389.5	3016.5	98	63	420	18868.5	17.5	17.5	50	3.5	80.5	276	9	136	3.5	19	561156.5
Merced Alamo	4981	87580.5	269126.5	155.5	169	18641.5	6895	3845.5	96	45	518.5	19174.5	24	16.5	39	3.5	89	218.5	12.5	134	5	19	590701.5
Merced Bear Creek	<LOD	78412.5	264760	1096.5	1031.5	13968	6267.5	4354.5	89	53.5	635	22216	28.5	46	62.5	4	67.5	131	11.5	232	6	18	606508.5
Merced San Juaquin	3038.5	82249.5	290488.5	130	211	20270	8813	3580.5	98	37.5	506	18439	21	18	41.5	2.5	92	241.5	14	193.5	8.5	18.5	571486.5

Appendix C: Regression diagnostic plots

pH

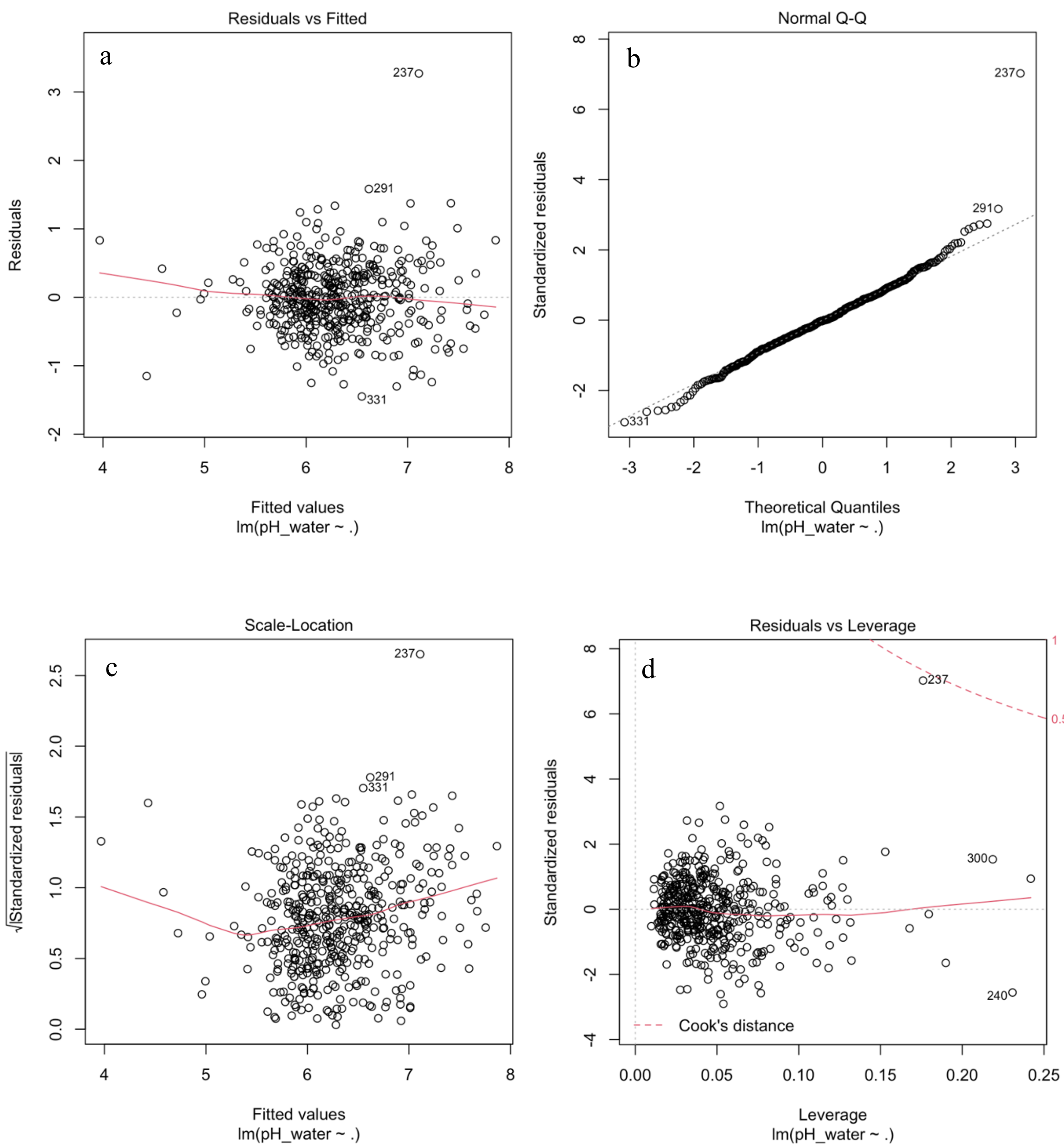

Figure C.1: Regression diagnostic plots for pH (a) Residual vs Fitted graph (b) Normal Q-Q plot (c) Scale-Location plot (d) Residuals vs Leverage plot

Sand %

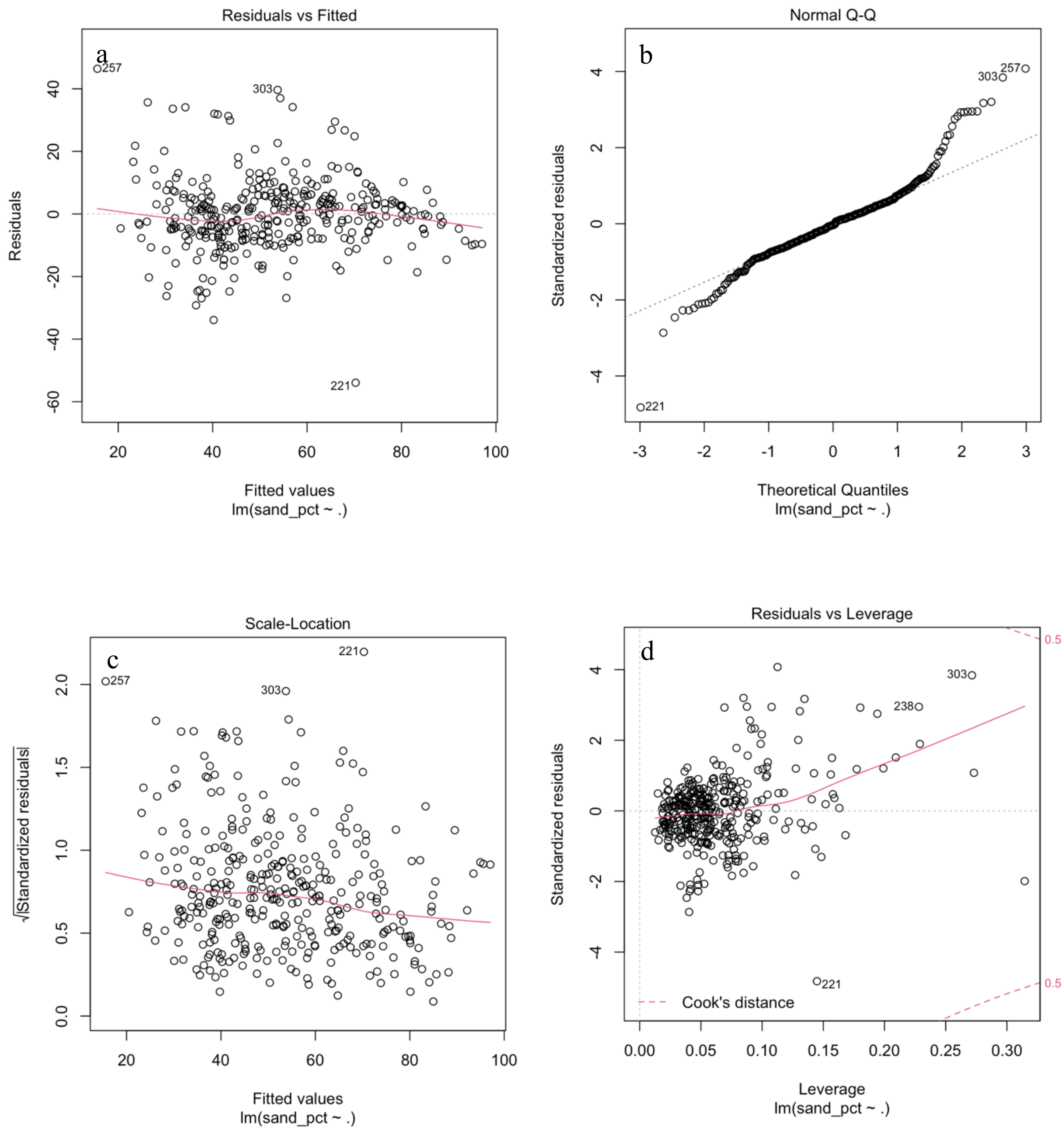

Figure C.2: Regression diagnostic plots for sand % (a) Residual vs Fitted graph (b) Normal Q-Q plot (c) Scale-Location plot (d) Residuals vs Leverage plot

Clay %

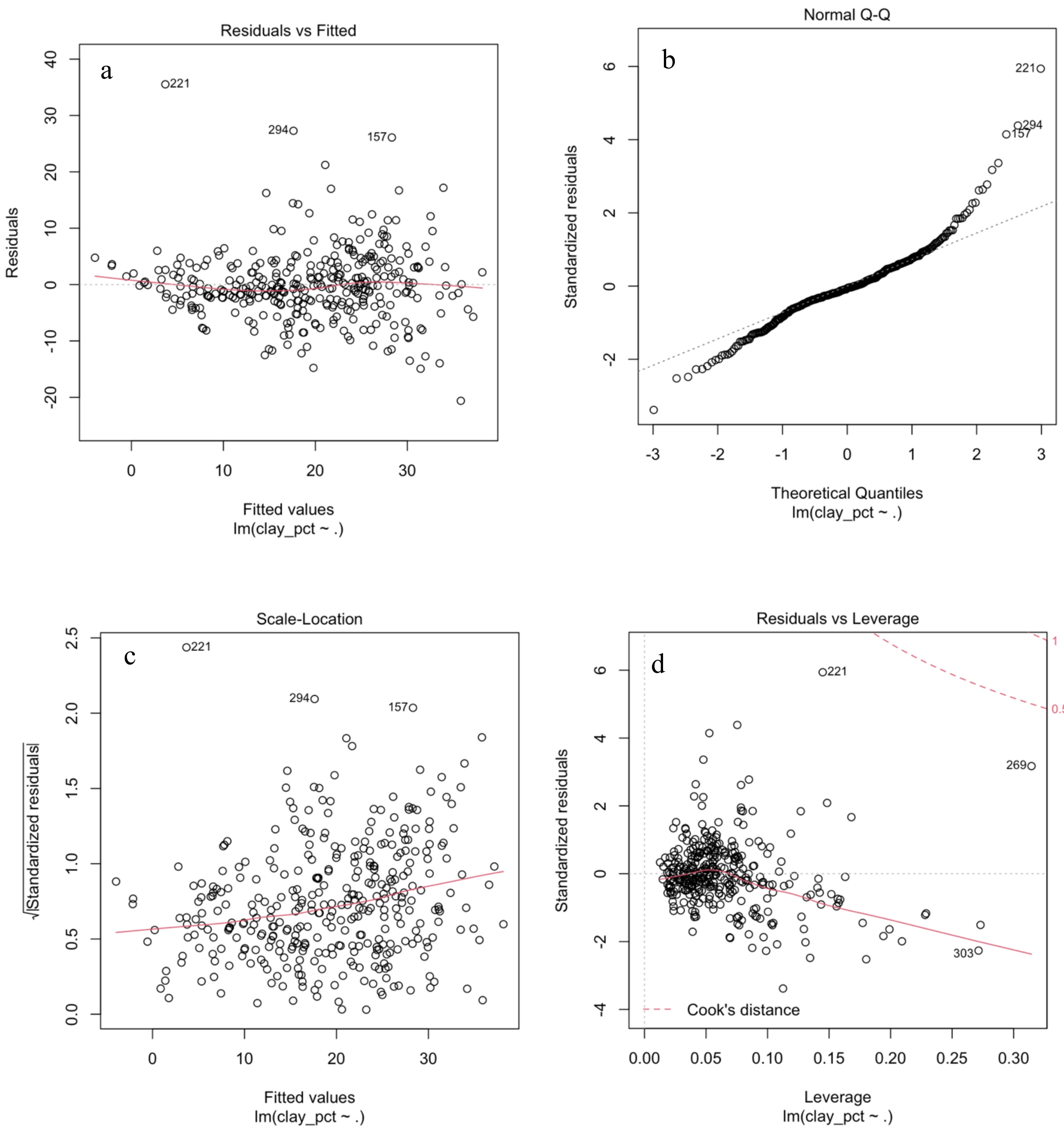

Figure C.3: Regression diagnostic plots for clay % (a) Residual vs Fitted graph (b) Normal Q-Q plot (c) Scale-Location plot (d) Residuals vs Leverage plot

CEC

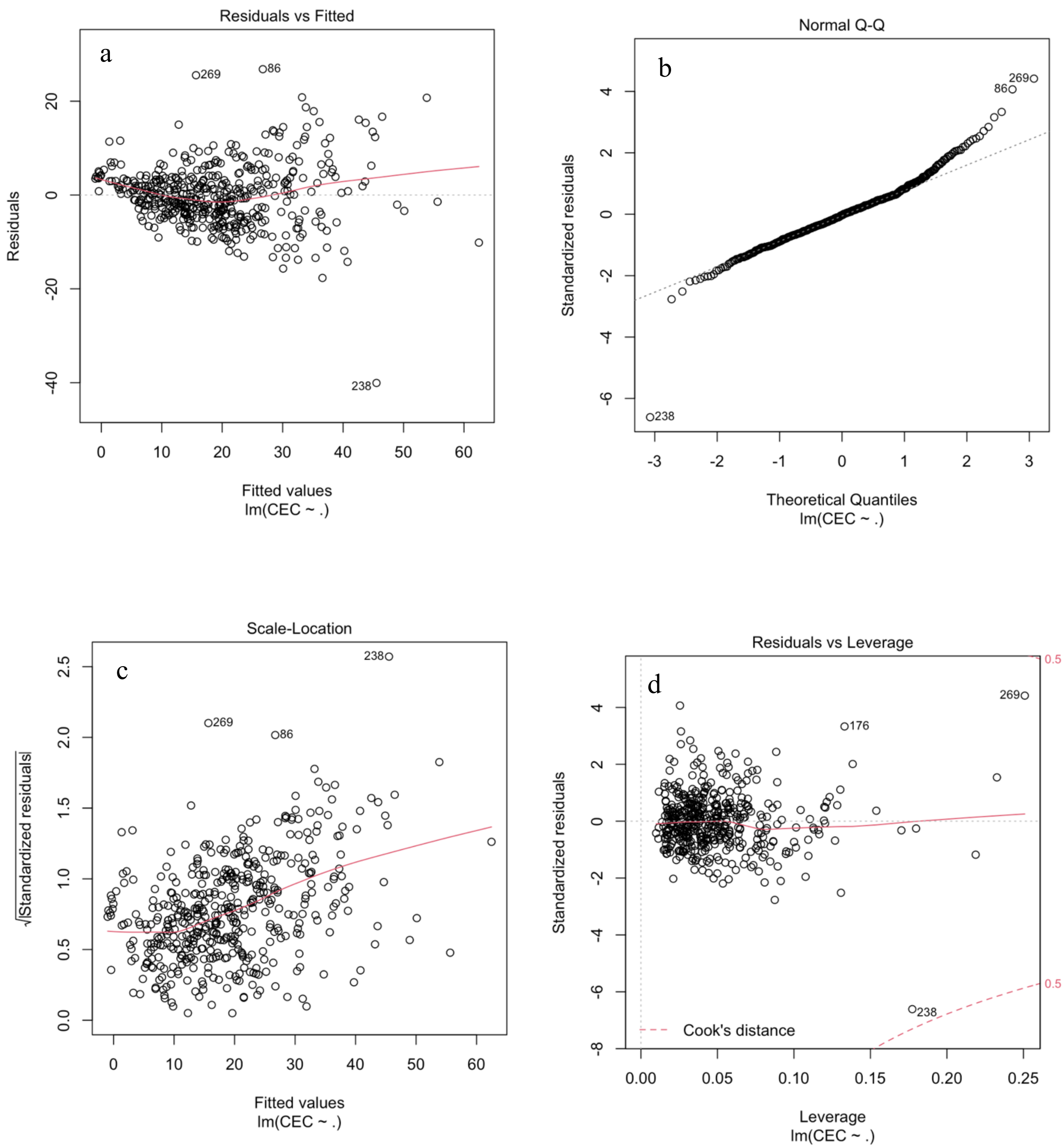

Figure C.4: Regression diagnostic plots for CEC (a) Residual vs Fitted graph (b) Normal Q-Q plot (c) Scale-Location plot (d) Residuals vs Leverage plot

SOC %

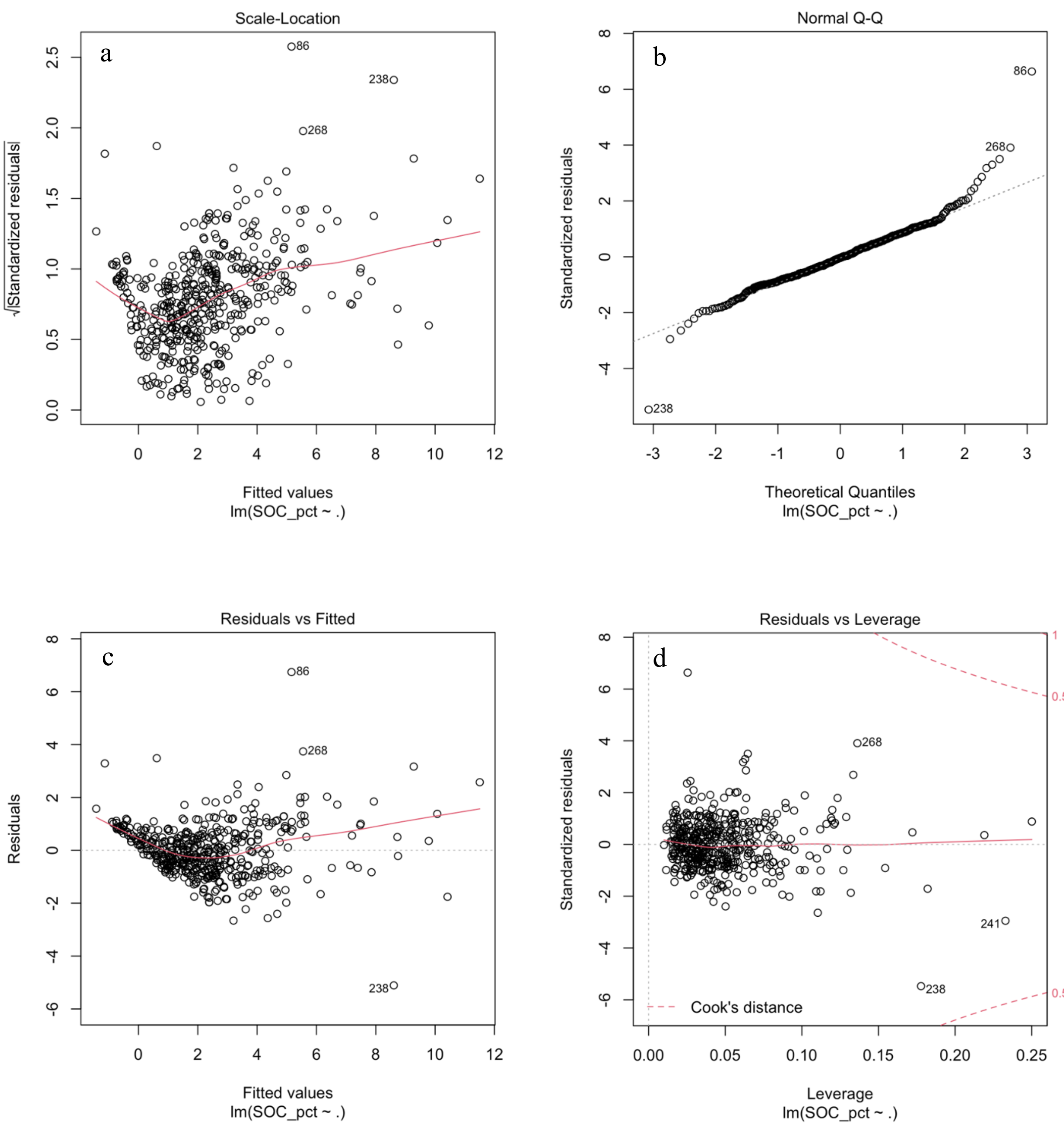

Figure C.5: Regression diagnostic plots for SOC % (a) Residual vs Fitted graph (b) Normal Q-Q plot (c) Scale-Location plot (d) Residuals vs Leverage plot

TN %

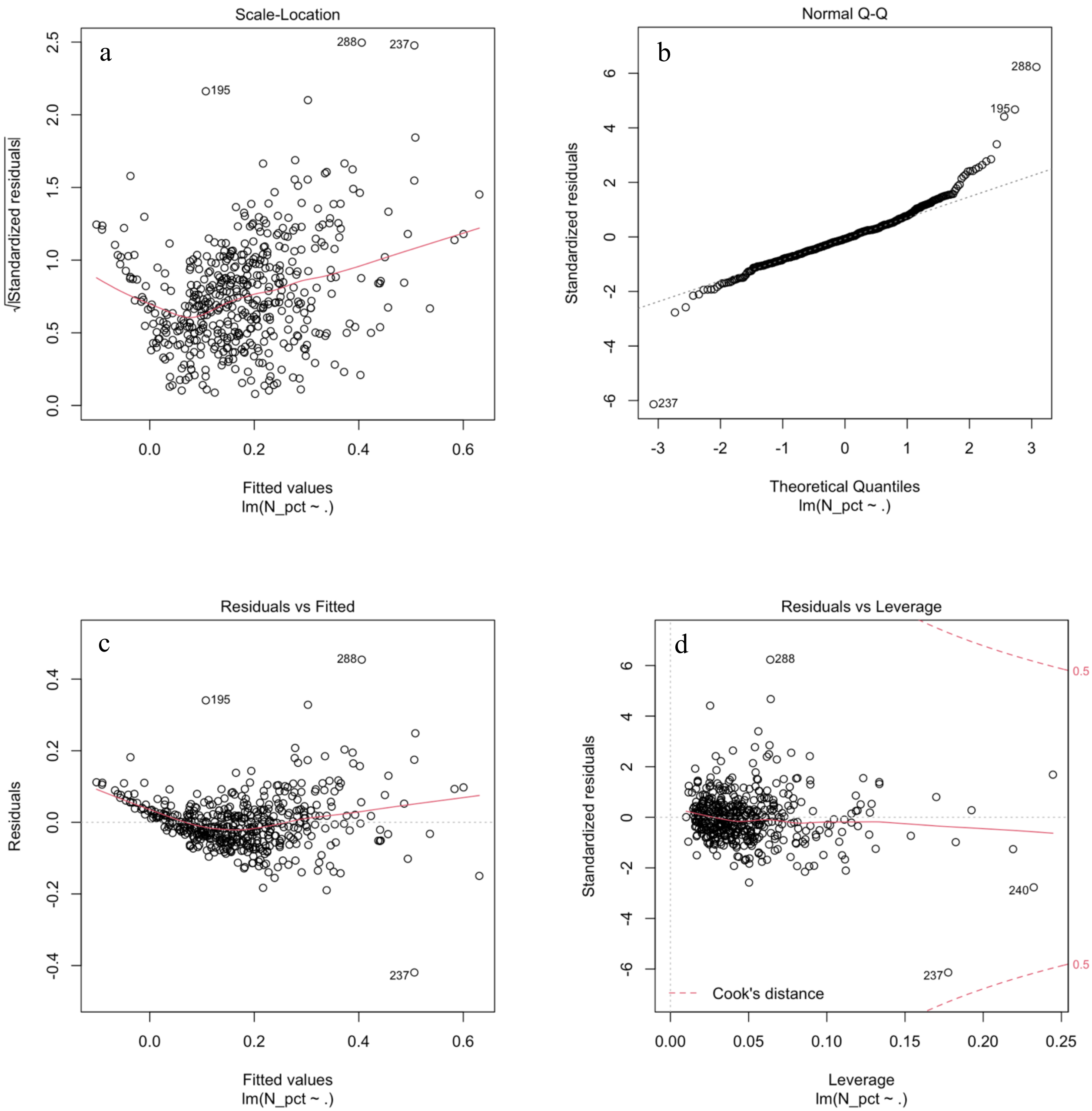

Figure C.6: Regression diagnostic plots for TN % (a) Residual vs Fitted graph (b) Normal Q-Q plot (c) Scale-Location plot (d) Residuals vs Leverage plot

CN ratio

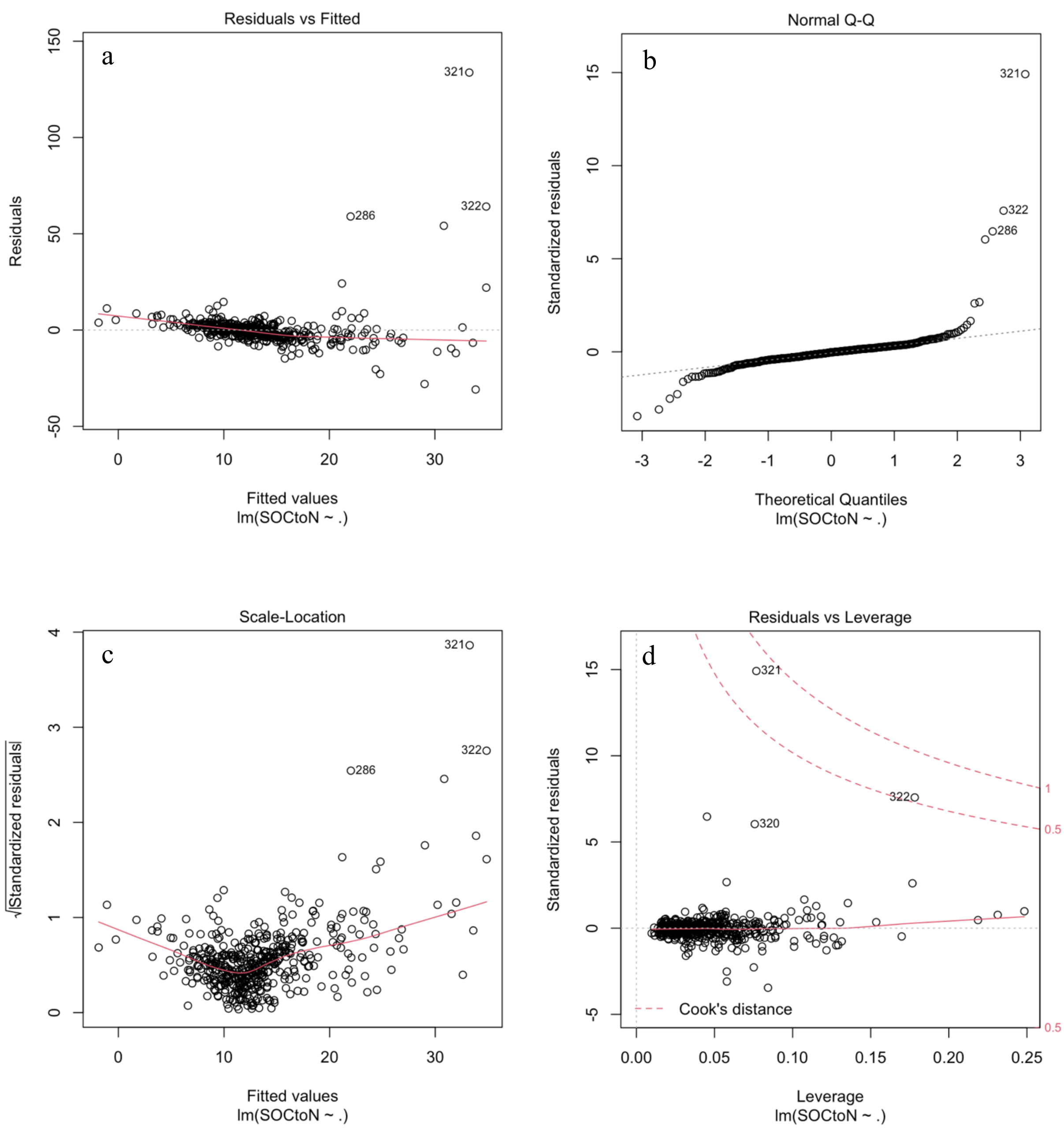

Figure C.7: Regression diagnostic plots for CN ratio (a) Residual vs Fitted graph (b) Normal Q-Q plot (c) Scale-Location plot (d) Residuals vs Leverage plot

Appendix D: Imputed analyte concentrations

Magnesium

Table D.1: Mg normal distribution curve parameters.

Average 1σ error of <LOD readings	3071.27 ppm
Average LOD (1σ error * 3)	9213.82 ppm
Mean concentration	4606.91 ppm
Standard deviation	1535.64 ppm

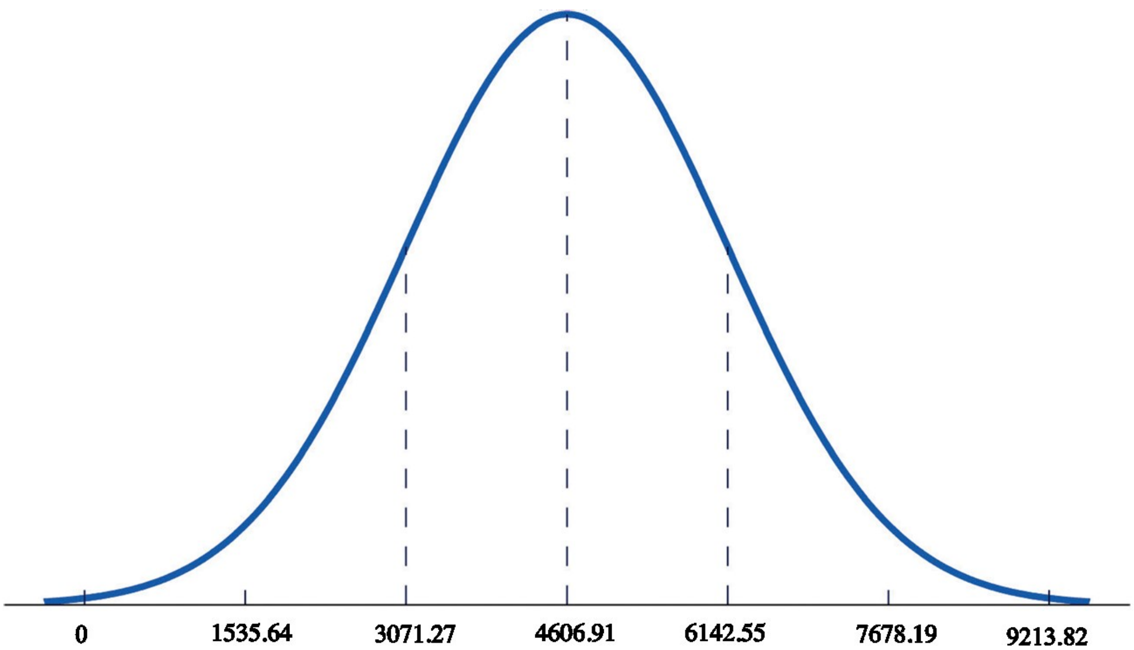

Figure D.1: Normal distribution curve for Mg concentration imputation.

Imputed concentrations in ppm (n = 90)

5458	6495	2517	4431	5559	4594	4096	5477	5033	4967	2670	3726
4177	4970	1757	5685	4836	5909	4236	7086	6770	4150	5065	6373
7333	4046	3931	5010	2366	4018	3282	3887	5026	3316	4380	6363
4895	6304	4309	7427	1494	3798	3306	3492	5767	5876	3242	4239
6361	2928	6751	5156	2984	4187	4761	3035	4071	6885	6163	2938
5245	5316	4761	3002	3489	3677	7049	1631	3768	5681	3193	4513
3813	1677	4771	2799	7287	3995	5286	5298	4490	5064	2773	1540
6739	4480	5210	2944	7067	6149						

Phosphorous

Table D.2: P normal distribution curve parameters.

Average 1σ error of <LOD readings	103.05 ppm
Average LOD (1σ error * 3)	309.14 ppm
Mean concentration	154.57 ppm
Standard deviation	51.52 ppm

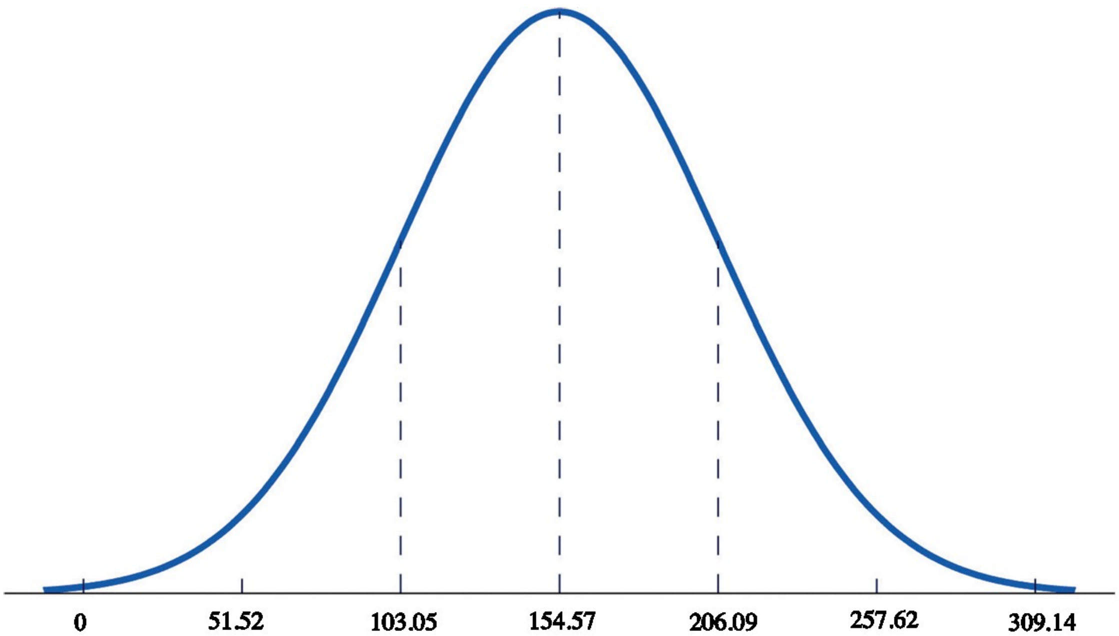

Figure D.2: Normal distribution curve for P concentration imputation.

Imputed concentrations in ppm (n = 69)

104	198	135	143	121	178	101	114	175	242	76	233
271	131	72	129	124	44	95	228	234	95	130	169
65	135	81	162	228	206	196	194	87	212	151	99
141	115	131	176	206	190	125	122	96	85	174	61
174	86	141	223	119	188	92	247	207	145	225	162
114	20	165	186	327	109	215	105	159			

Sulfur

Table D.3: S normal distribution curve parameters.

Average 1σ error of <LOD readings	101.03 ppm
Average LOD (1σ error * 3)	303.09 ppm
Mean concentration	151.54 ppm
Standard deviation	50.51 ppm

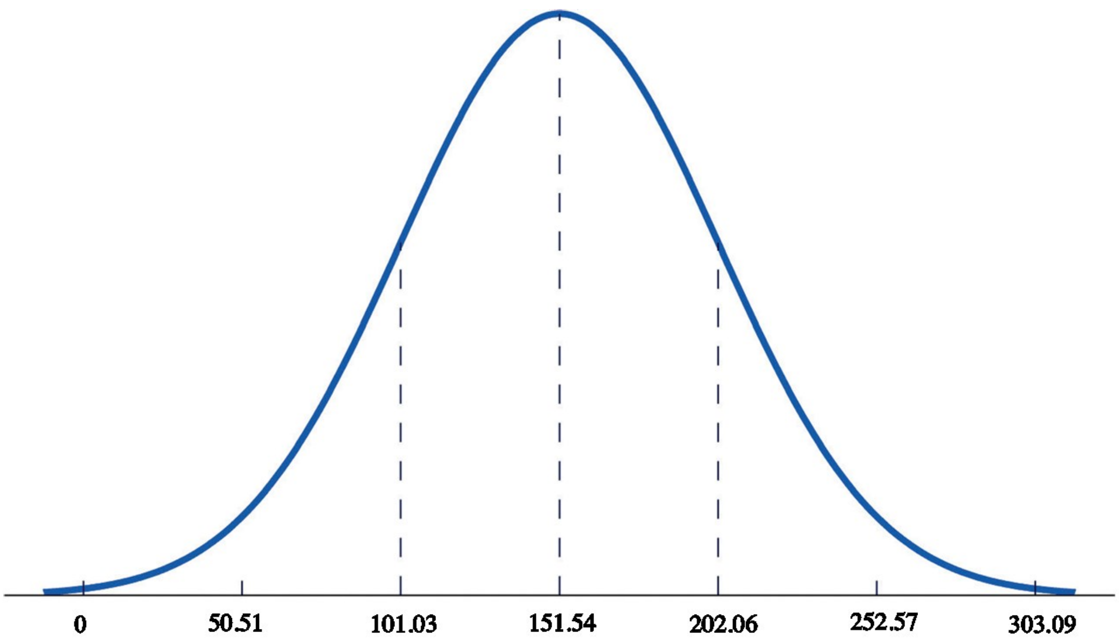

Figure D.3: Normal distribution curve for S concentration imputation.

Imputed concentrations in ppm (n = 23)

155	144	107	115	189	174	202	198	123	272	111	218
147	195	235	139	149	134	196	201	113	122	145	

Calcium

Table D.4: Ca normal distribution curve parameters.

Average 1σ error of <LOD readings	149.27 ppm
Average LOD (1σ error * 3)	447.80 ppm
Mean concentration	223.90 ppm
Standard deviation	74.63 ppm

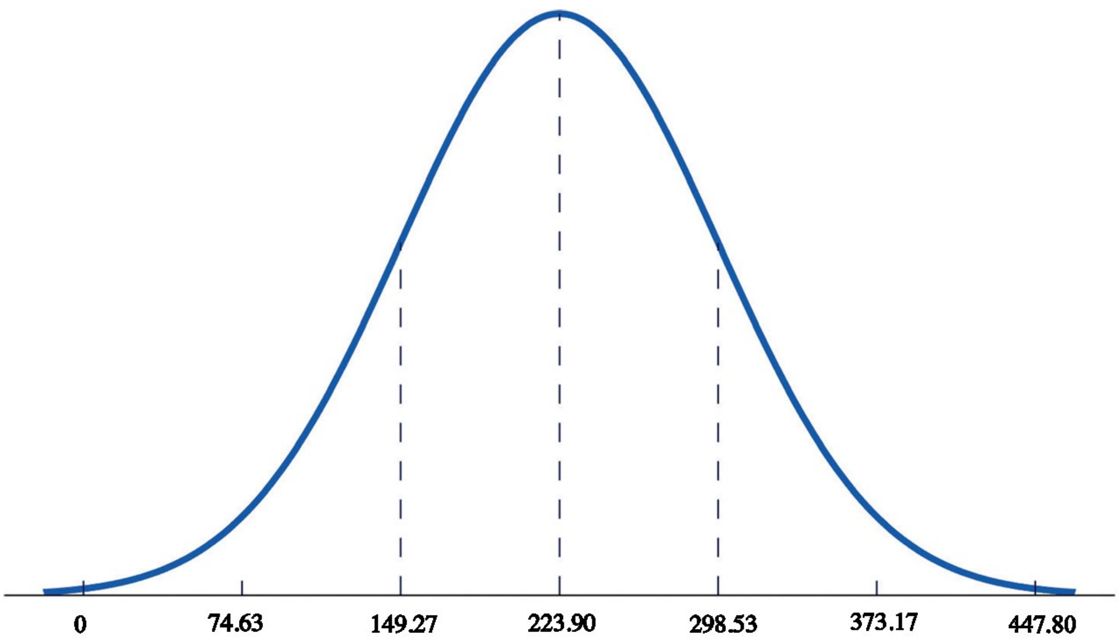

Figure D.4: Normal distribution curve for Ca concentration imputation.

<u>Imputed concentrations in ppm (n = 5)</u>

128 130 240 295 136

Chromium

Table D.5: Cr normal distribution curve parameters.

Average 1σ error of <LOD readings	26.35 ppm
Average LOD (1σ error * 3)	79.06 ppm
Mean concentration	39.59 ppm
Standard deviation	13.18 ppm

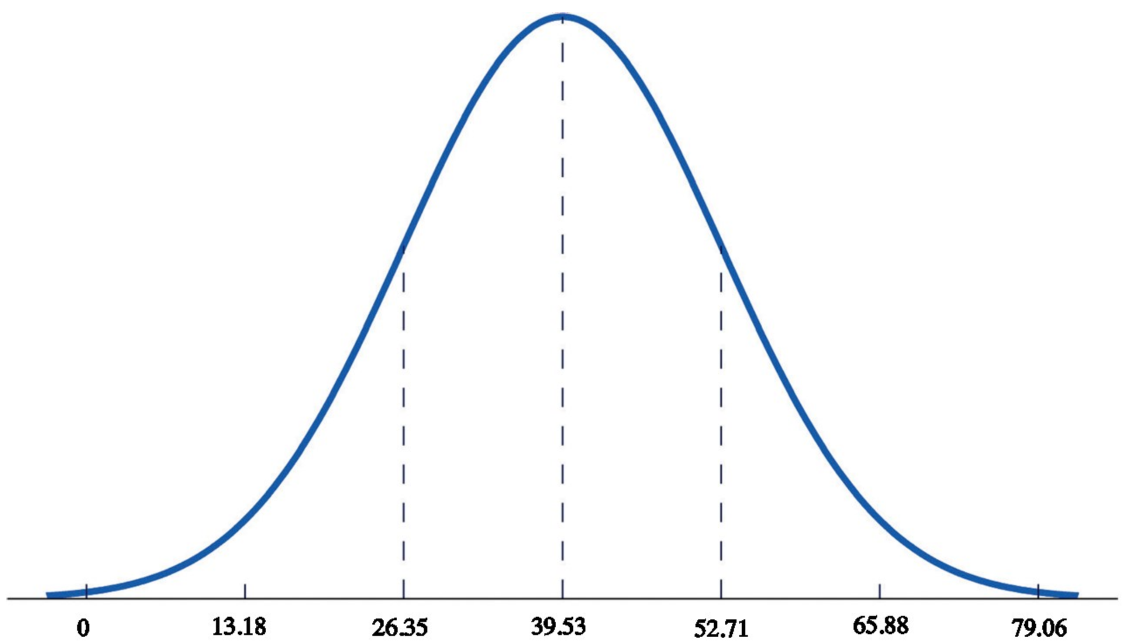

Figure D.5: Normal distribution curve for Cr concentration imputation.

<u>Imputed concentrations in ppm (n = 9)</u>

36 33 34 29 47 59 46 43 49

Arsenic

Table D.6: As normal distribution curve parameters.

Average 1σ error of <LOD readings	3.87 ppm
Average LOD (1σ error * 3)	11.60 ppm
Mean concentration	5.80 ppm
Standard deviation	1.93 ppm

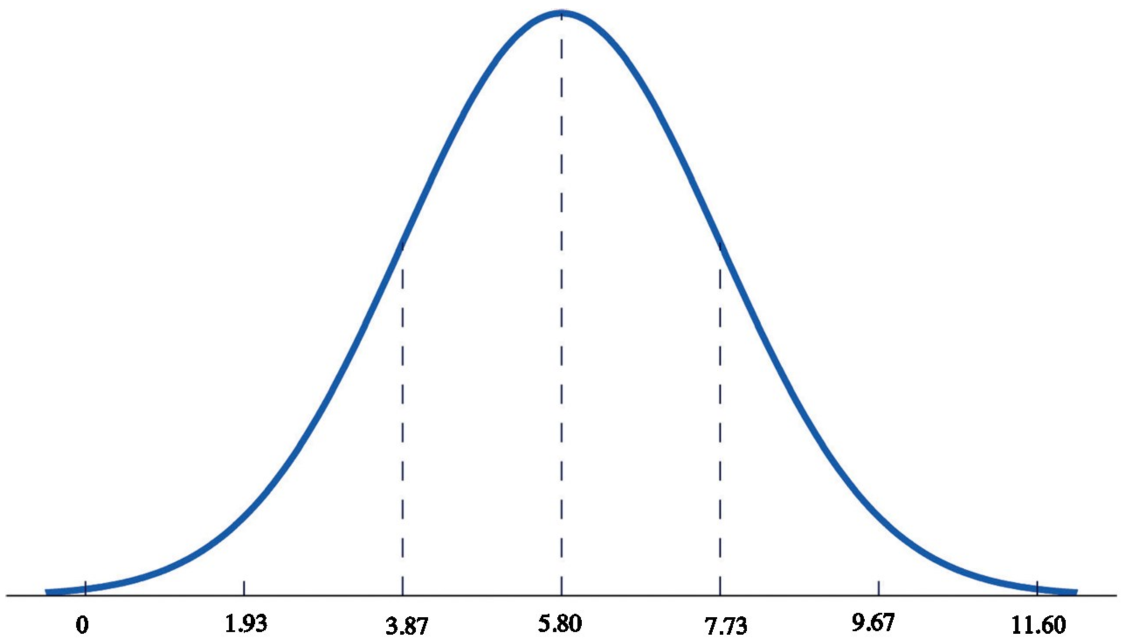

Figure D.6: Normal distribution curve for As concentration imputation.

Imputed concentrations in ppm (n = 16)

10 6 7 5 5 5 4 8 5 6 4 4
5 3 4 6

Niobium

Table D.7: Nb normal distribution curve parameters.

Average 1σ error of <LOD readings	2.81 ppm
Average LOD (1σ error * 3)	8.44 ppm
Mean concentration	4.22 ppm
Standard deviation	1.41 ppm

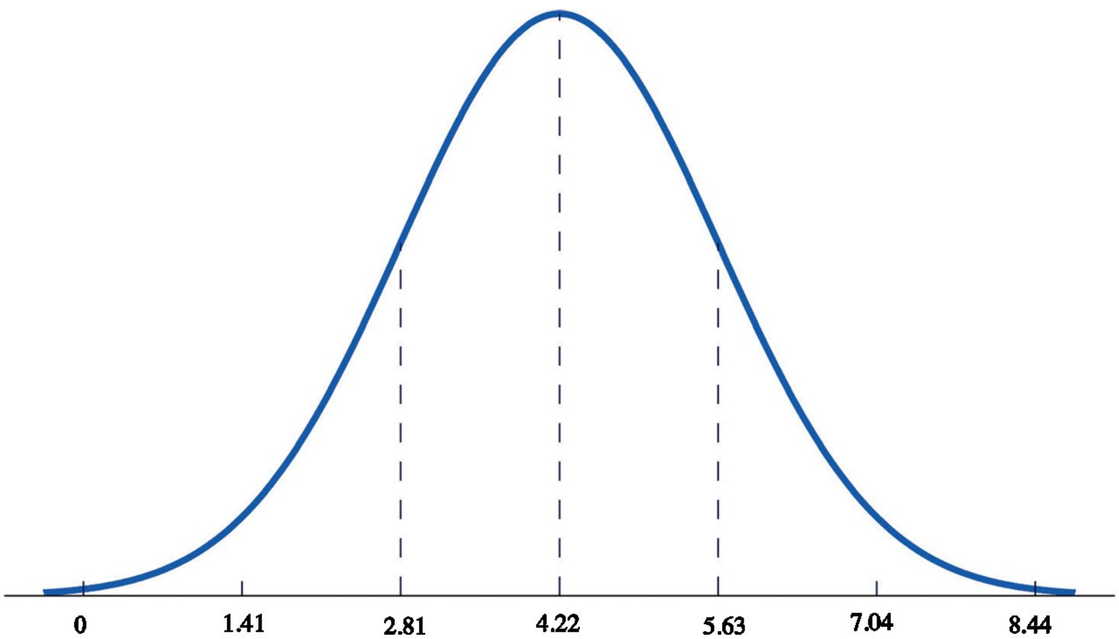

Figure D.7: Normal distribution curve for Nb concentration imputation.

<u>Imputed concentrations in ppm (n = 69)</u>

6	6	6	4	5	5	6	5	5	5	4	4
4	3	4	2	4	4	4	6	3	4	6	5
2	3	3	6	4	5	6	3	3	5	5	3
5	4	6	4	4	7	4	5	2	3	3	5
3	4	6	7	1	6	6	6	7	2	1	4
4	4	4	5	3	4	3	4	4			

Appendix E: MLR models grouped by land type and methodology

pH

Predictions within land type

Forest: (n=164) $R^2 = 0.554$, RMSE = 0.392, RPD = 1.491989, RPIQ = 2.181967

pH (forest) = 12.4764 − 0.9890 * log(S) − 2.9041 * log(K) + 1.2929 * log(Ca) + 1.2405 * log(V) − 2.0563 * log(Fe) + 0.9099 * log(Cu) + 1.5726 * log(Zn) + 3.4226 * log(Rb) − 1.3226 * log(Sr) + 1.9926 * log(Zr) − 1.5661 * log(Pb)

Grassland: (n=81) $R^2 = 0.721$, RMSE = 0.346, RPD = 1.832114, RPIQ = 2.484338

pH (grassland) = 0.3702 − 0.5780 * log(S) + 0.4813 * log(Ca) + 1.2014 * log(Ti) + 0.3710 * log(Mn)

Marine terrace: (n=159) $R^2 = 0.463$, RMSE = 0.396, RPD = 1.330089, RPIQ = 1.579797

pH (marine terrace) = 4.86504 − 0.70310 * log(P) + 0.72776 * log(Ca) + 0.09614 * log(V) − 0.25164 * log(Cr) + 0.54533 * log(Mn) − 0.62438 * log(Y)

Predictions within methodologies

1:1 DI to soil: (n=319) $R^2 = 0.488$, RMSE = 0.531, RPD = 1.405067, RPIQ= 1.790669

pH (1:1) = 9.1443 − 2.0023 * log(Al) − 0.6190 * log(P) − 0.5009 * log(S) + 0.6448 * log(Ca) + 1.2714* log(V) + 1.0910* log(Cu) + 1.7927* log(Zn) − 0.5594* log(Y) + 0.7178 * log(Zr) − 0.9327 * log(Pb)

Saturated paste: (same model as pH for marine terrace)

Sand %

Predictions within land type

Forest: (n=86) $R^2 = 0.672$, RMSE = 13.8, RPD = 1.523129, RPIQ = 2.633648

Sand % (forest) = $-1102.86 + 31.20 * \log(Mg) + 67.21 * \log(Al) + 114.93 * \log(Si) + 90.09 * \log(Ti) - 46.67 * \log(Cr) - 49.31 * \log(As) - 23.054 * \log(Zr) + 43.62 * \log(Sr) - 80.97 * \log(Zr) - 29.62 * \log(Pb)$

Grassland: (n=41) $R^2 = 0.386$, RMSE = 10.1, RPD = 1.140051, RPIQ = 1.351508

Sand % (grassland) = $590.72 - 74.78 * \log(Fe) - 117.84 * \log(Rb)$

Marine terrace: (n=159) $R^2 = 0.895$, RMSE = 5.56, RPD = 2.985065, RPIQ= 4.94454

Sand % (marine terrace) = $-3.438 + 6.782 * \log(Mg) + 110.230 * \log(Al) - 59.377 * \log(Si) + 6.566 * \log(P) + 69.368 * \log(K) + 10.794 * \log(Cr) - 65.167 * \log(Fe) - 19.689 * \log(Zn) - 77.312 * \log(Rb) + 15.831 * \log(Y) - 21.735 * \log(Zr) - 11.772 * \log(Pb)$

Predictions within methodologies

Hydrometer: (n= 298) $R^2 = 0.462$, RMSE = 13.3, RPD = 1.238532, RPIQ = 1.902471

Sand % (hydrometer) = $73.594 + 115.730 * \log(Al) - 76.393 * \log(Si) + 6.591 * \log(P) + 41.964 * \log(K) - 69.658 * \log(Fe) - 30.785 * \log(Cu) - 34.578 * \log(Rb) + 19.141 * \log(Sr) + 10.272 * \log(Y)$

Pipette: (n= 60) $R^2 = 0.348$, RMSE = 24.5, RPD = 1.19692, RPIQ = 1.819865

Sand % (pipette) = $-41.1663 + 0.1271 * \log(Mg) - 36.6582 * \log(Si) + 87.7636 * \log(K) - 79.0289 * \log(As)$

Clay %

<u>Predictions within land type</u>

Forest: (n= 86) R^2 = 0.714, RMSE = 6.60, RPD = 1.801769, RPIQ = 3.425242

Clay % (forest) = 312.642 − 34.958 * log(Al) − 22.128 * log(Si) − 47.058 * log(K) + 8.385 * log(Cr) + 22.251 * log(As) + 76.126 * log(Rb) + 6.630 * log(Zr)

Grassland: (n=41) R^2 = 0.101, RMSE = 9.25, RPD = 0.558785, RPIQ = 0.7429938

Clay % (grassland) = −1053.14 + 242.05 * log(Si) − 20.20 * log(P) − 147.94 * log(K) + 19.28 * log(Ca) + 72.46 * log(Ti) − 71.76 * log(V) − 10.87 * log(Mn) + 145.83 * log(Rb) − 27.05 * log(Nb)

Marine terrace: (n= 159) R^2 = 0.812, RMSE = 3.55, RPD = 2.193866, RPIQ = 2.759737

Clay % (marine terrace) = −125.565 + 37.201 * log(Si) − 50.809 * log(K) − 7.667 * log(Mn) + 18.487 * log(Fe) + 46.287 * log(Rb) − 11.529 * log(Sr) + 7.591 * log(Zr) + 11.613 * log(Pb)

<u>Predictions within methodologies</u>

Hydrometer: (n= 298) R^2 = 0.624, RMSE = 6.05, RPD = 1.59405, RPIQ = 2.180083

Clay % (hydrometer) = −63.433 − 36.985 * log(Al) + 49.569 * log(Si) − 55.291 * log(K) + 24.823 * log(Fe) + 12.253 * log(Ni) + 42.641 * log(Rb) + 7.279 * log(Zr)

Pipette: (n=60) R^2 = 0.631, RMSE = 10.1, RPD = 1.143407, RPIQ = 1.485874

Clay % (pipette) = −68.354 − 28.919 * log(Mg) + 8.771 * log(S) − 112.427 * log(Ti) + 125.696 * log (Fe) − 38.959 * log(Zn) + 38.438 * log(As) + 79.109 * log(Y) − 54.210 * log(Nb)

CEC

Predictions within land type

Forest: (n= 164) R^2 = 0.819, RMSE = 7.52, RPD = 2.328453, RPIQ = 2.496853

CEC (forest) = 960.22 − 16.74 * log(Mg) − 70.88 * log(Al) −101.23 * log(Si) − 12.07 * log(Ti) + 17.00 * log(Cu) + 13.68 * log(Zn) + 32.03 * log (As) − 17.18 * log(Y)

Grassland: (n= 81) R^2 = 0.517, RMSE = 5.44, RPD = 1.082659, RPIQ = 1.139399

CEC (grassland) = 448.11 − 116.72 * log(K) − 31.79 * log(Fe) + 36.44 * log(Cu) + 84.84 * log(Rb) − 17.99 * log(Pb)

Marine terrace: (n=159) R^2 = 0.653, RMSE = 2.57, RPD = 1.716937, RPIQ = 2.721496

CEC (marine terrace) = 282.558 − 43.197 * log(Al) − 24.900 * log(K) − 24.897 * log(Ti) − 4.258 * log(Cr) + 16.979 * log(Fe) + 9.954 * log(Zn) + 16.208 * log(Rb) + 8.594 * log(Zr)

Predictions within methodologies

Ammonia gas absorbance: (n= 218) R^2 = 0.689, RMSE = 7.38, RPD = 1.777052, RPIQ = 2.248444

CEC (absorbance) = 866.41 − 17.42 * log(Mg) − 57.19 * log(Al) − 84.50 * log(Si) − 26.04 * log(Ti) + 32.08 * log(Cu) + 11.87 * log(Zn) − 15.45 * log(Y)

UN-FAO CEC: (n= 41) R^2 = 0.238, RMSE = 12.8, RPD = 1.071226, RPIQ = 0.800528

CEC (UN-FAO) = 31.82 − 43.53 * log(Ti) + 30.10 * log(Fe) + 11.97 * log(As)

CEC7: (n= 56) R^2 = 0.646, RMSE = 8.54, RPD = 1.140609, RPIQ = 1.562391

CEC (CEC7) = 1118.46 − 42.04 * log(Mg) − 58.67 * log(Al) − 117.73 * log(Si) − 21.35 * log(Ca) − 90.29 * log(Ti) + 26.11 * log(Mn) + 41.93 * log(Fe) + 27.19 * log(As) + 21.73 * log(Sr) + 90.34 * log(Y) − 39.18 * log(Nb) −18.16 * log(Pb)

S - 10.10: (n=159) (same model as CEC for marine terrace)

SOC content

<u>Predictions within land type</u>

Forest: (n= 168) $R^2 = 0.815$, RMSE = 1.3, RPD = 2.203682, RPIQ = 2.432288

$$SOC\ \%\ (forest) = 200.254 - 11.670 * \log(Al) - 22.149 * \log(Si) + 2.337 * \log(S) - 6.716 * \log(Fe) + 1.790 * \log(Cu)$$

Grassland: (n=81) $R^2 = 0.661$, RMSE = 1.01, RPD = 1.638437, RPIQ = 1.844786

$$SOC\ \%\ (grassland) = -21.705 - 9.916 * \log(Al) + 2.863 * \log(P) + 1.672 * \log(S) + 16.823 * \log(Ti) - 3.322 * \log(Mn) - 1.310 * \log(Ni) - 1.986 * \log(Cu) + 5.537 * \log(Zn) + 4.028 * \log(Rb) - 5.465 * \log(Nb)$$

Marine terrace: (n= 159) $R^2 = 0.821$, RMSE = 0.738, RPD = 2.36662, RPIQ = 2.239114

$$SOC\ \%\ (marine\ terrace) = 123.814 - 11.556 * \log(Al) - 13.193 * \log(Si) + 1.002 * \log(S) - 1.772 * \log(Cr) + 2.439 * \log(Zn) + 1.439 * \log(Zr)$$

TN content

Predictions within land type

Forest: (n= 165) $R^2 = 0.738$, RMSE = 0.0840, RPD = 1.973068, RPIQ = 2.306534

TN % (forest) = 9.0712 − 0.5387 * log(Al) − 0.8676 * log(Si) + 0.1171 * log(S) − 0.2682 * log(K) + 0.1413 * log(Mn) − 0.4349 * log(Fe) + 0.1308 * log(Cu) + 0.1861 * log(As) − 0.1302 * log(Y) + 0.2561 * log(Zr)

Grassland: (n=81) $R^2 = 0.842$, RMSE = 0.0568, RPD = 2.456222, RPIQ = 4.434534

TN % (grassland) = −2.8321 + 0.2104 * log(P) + 0.1217 * log(S) + 0.6091* log(Ti) − 0.1660 * log(Mn) − 0.2159 * log(Ni) + 0.4400 * log(Zn) + 0.1190 * log(Rb) − 0.3014 * log(Nb)

Marine terrace: (n= 159) $R^2 = 0.762$, RMSE = 0.0544, RPD = 1.919689, RPIQ = 2.69714

TN % (marine terrace) = 10.99877 − 0.97424 * log(Al) − 1.09490 * log(Si) − 0.05386 * log(P) + 0.10879 * log(S) − 0.03612 * log(Ca) − 0.37419 * log(V) − 0.13065 * log(Cr) + 0.41760 * log(Zn) − 0.09371* log(Y) + 0.11565 * log(Zr)

CN ratio

<u>Predictions within land type</u>

Forest: (n= 168) $R^2 = 0.273$, RMSE = 6.46, RPD = 0.7049186, RPIQ= 0.7112989

C:N (forest) = 100.97 + 25.32 * log(Mg) − 54.53 * log(K) − 23.30 * log(Zn) + 38.45 * log(Sr)

Grassland: (n= 81) $R^2 = 0.192$, RMSE = 2.03, RPD = 1.096776, RPIQ= 1.067139

C:N (grassland): 36.184 − 8.783 * log(Al) + 2.957 * log(S) − 4.100 * log(K) + 10.370 * log(Zr) + 3.552 * log(Pb)

Marine terrace: (n= 159) $R^2 = 0.342$, RMSE = 1.04, RPD = 1.239946, RPIQ= 1.176866

C:N (marine terrace) = 3.369 + 1.019 * log(S) + 4.091 * log(Ca) − 5.359 * log(Ni) − 4.323 * log(Cu) + 2.406