DISSERTATION

IMPROVING SOIL PROPERTY PREDICTIONS FOR APPLICATIONS IN TAILINGS AND TERRAMECHANICS

Submitted by

Joseph R. Bindner

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2024

Doctoral Committee:

Advisor: Joseph Scalia
Co-Advisor: Rebecca Atadero

Christopher Bareither
Jeffrey Niemann
Jay Ham

ABSTRACT

IMPROVING SOIL PROPERTY PREDICTIONS FOR APPLICATIONS IN TAILINGS AND TERRAMECHANICS

Soil properties are used by engineers and scientists to better understand the state and behavior of soils. For example, soil properties can be used to estimate surficial soil strength for vehicle mobility models and can be used to better understand the engineering characteristics of mine waste (tailings) stored in tailings storage facilities. Soil and tailings properties often have high spatial variability and often require high resolution data for engineering analyses. Standard laboratory procedures are commonly used to determine soil properties but are often impractical for large spatial extents. While some existing soil data products provide estimates of surficial soil properties, the fidelity of soil data products is often poorly understood and insufficient for many applications. Additionally, some field tests used to estimate soil properties, such as the cone penetration test (CPT), rely on empirical correlations that cannot be used for some soils. There remains a need for procedures which improve the speed and accuracy of soil property estimates across large spatial extents. The objectives of this study are to (i) evaluate how surficial soil moisture and soil strength vary with soil and landscape attributes across a large spatial extent, (ii) explore the use of field-based hyperspectral sensing and machine learning for the prediction of surficial soil properties across a landscape, and (iii) assess the use of laboratory hyperspectral sensing and machine learning for the prediction of tailings properties for potential application in situ via direct push methods. Soil and landscape attributes were determined at sampling locations across a semi-arid foothills region and used to assess how soil moisture and soil strength vary

with soil and landscape attributes. Then, hyperspectral data were captured at select sampling locations and used to train and assess the performance of a convolutional neural network (CNN) for the predictions of soil properties. Finally, a diverse tailings-hyperspectral dataset was prepared in the lab and used to train and assess a CNN to provide proof of concepts for prediction of material properties relevant to TSF stability analyses.

ACKNOWLEDGEMENTS

I extend my gratitude for the opportunities afforded to me during my doctoral studies. I have had the pleasure of practicing research under mentors whose guidance has proven invaluable. I thank Dr. Joseph Scalia for his direction and support as my advisor and mentor, Dr. Becki Atadero, Dr. Chris Bareither, Dr. Jeffrey Niemann, and Dr. Jay Ham for their mentorship, counsel, and service as members of my committee. I extend my gratitude to Dr. John Ridley, Dr. Brian Schaible, and Dr. Kevin Harsh for their generous time and equipment contributions, and Colorado State University STRATA for facilitating use of study sites. I thank my colleagues at Colorado State University for their key contributions to this work. Finally, I thank my family and my spouse for their unwavering support.

Financial support for studies herein was provided by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-21-2-0252, the USDA National Institute for Food and Agriculture, Hatch grant 1000065-COL00797, the National Science Foundation (2312319), and the Tailings and Industrial Waste Engineering Center (TAILENG). The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office, the U.S. Government, TAILENG, or Colorado State University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Introduction

Spatially extensive information on soil properties is needed for a range of geotechnical and terramechanics applications including the prediction of vehicle mobility across terrain and for the design and management of tailings storage facilities (TSF). Specifically, assessing soil properties for vehicle mobility and tailings applications requires fine resolution data across large spatial extents. However, determining fine resolution soil properties across large spatial extents can be costly, time intensive, and may not be practical in many applications.

Vehicle mobility applications require fine resolution (5- to 30-m) estimates of surficial soil strength, which is informed by soil moisture, particle size distribution, soil plasticity, and other soil characteristics. Soil moisture estimates are available nearly globally at coarse resolutions (9-60-km) (Njoku et al., 2003; Kerr et al, 2010; O'Neill et al., 2010) from microwave satellites and various downscaling procedures can be used to obtain fine resolution soil moisture estimates (10- to 30-m) (Merlin et al., 2006; Coleman and Niemann, 2013; Peng et al., 2016; Sadeghi et al., 2017; Fang et al., 2018). Many soil moisture downscaling procedures rely on the relationship between soil moisture and topographic, soil, and vegetation attributes (Coleman and Niemann, 2013; Ranney et al, 2015; Cowley et. al, 2017; Deshon et al., 2020). However, these modeling procedures assume the relationship of soil moisture with topographic, vegetation, and soil attributes is constant throughout a downscaling extent. Studies have explored how soil moisture relationships vary spatially with soil and topographic attributes in small catchments (Mohanty et al., 2000; Western et al., 2004; Lin et al, 2006; Kim et al., 2007; Rosenbaum et al., 2012; Mascaro et al., 2011; Liang et al., 2017), however, there is a poor understanding of how soil moisture relationships vary within a coarse-resolution microwave remote sensing grid cell.

An improved understanding of spatial variability in these relationships may better inform soil moisture downscaling procedures and corresponding soil moisture estimates, therefore providing a higher quality input into soil strength and vehicle mobility models. The first objective of this study is to evaluate how soil moisture and soil strength relationships vary with topography, vegetation, and soil attributes within a microwave remote sensing grid cell.

Estimates of soil characteristics which help inform surficial soil strength are typically available from digital soil mapping (DSM) procedures, such as the Soil Survey Geographic Database (SSURGO) (Soil Survey Staff, 2023) and are used in soil moisture downscaling and soil strength modeling procedures (Coleman and Niemann, 2013; Ranney et al, 2015; Choi et al., 2019; Pauly, 2019; Bindner, 2020; Bullock, 2023). However, the fidelity of DSM data is often unreported and may negatively impact the performance of soil moisture and soil strength modeling procedures. Laboratory procedures can be used to characterize relevant soil attributes, although this is typically impractical given the resolution of soil properties needed for soil strength and vehicle mobility modeling. Hyperspectral data captured in controlled laboratory environments have shown strong promise for characterizing soil properties (Gupta et al., 2016; Tsakiridis et al., 2020; Knadel et al., 2021; Purushothaman et al., 2024). However, collection of laboratory hyperspectral data requires field soil sampling and soil processing which can be time intensive. Field captured hyperspectral data using instruments mounted on vehicles or drones may provide faster data collection, although there is little understanding if field captured hyperspectral data on undisturbed surficial soils under natural lighting conditions can be used for the prediction of soil properties. The second objective of this study is to explore how hyperspectral data captured in field environments can be used to predict soil properties.

In addition to aerial extents, characterization of soil strength with depth is critical for understanding the stability of tailings in TSFs (Morrison, 2022). While tailings behavior can be characterized using in situ testing such as the cone penetration test (CPT), not all relevant tailings properties are predicted and estimates of tailings properties from CPT rely on empirical relationships. Undisturbed samples can be used for the laboratory characterization of tailings properties. However, obtaining enough undisturbed samples for characterization at high spatial resolutions is generally impractical providing the spatial extent of many TSFs. Hyperspectral imagery on ex situ samples has shown promise for the prediction of tailings properties including moisture content, solids content, and particle size distribution (Entezari et al., 2022). However, the use of hyperspectral imagery for the prediction of degree of saturation, void ratio, porosity, and density is not well understood. The third objective of this study is to assess how hyperspectral imagery and machine learning can be used to predict tailings properties relevant to TSF analyses.

Methods and Materials

A field site in northern Colorado was identified for study objectives related to surficial soil moisture and strength for vehicle mobility applications (objectives one and two). Sampling locations were identified at multiple geographically separated regions (Regions A-D) for soil moisture and strength sampling (n=86) and hyperspectral imaging (n=48). Soil moisture and soil strength data were collected at sampling locations on 10 dates during May-August of 2022 and hyperspectral data were collected on three dates in August-October of 2023. Soils at sampling locations were characterized using laboratory procedures to determine particle size distribution parameters, soil plasticity (liquid limit, plastic limit, and plasticity index), and organic content. Pedotransfer functions were used to estimate soil porosity and hydraulic conductivity based on soil PSD. Topography and vegetation at the sampling locations were characterized using remote sensing products including the National Elevation Dataset (USGS, 2022) and Sentinel 2 satellite data (ESA, 2016). Hyperspectral data were collected using a broadband spectrometer (~300-1000 nm).

As part of the first objective of this study, correlation analysis was used to understand variability in the strength of soil moisture and soil strength relationships with topographic, vegetation, and soil attributes between regions. The regression slope of the relationships of soil moisture and soil strength with topography, vegetation, and soil were compared between regions to understand how these relationships deviate from average behavior.

To achieve the second study objective, hyperspectral data were split into training and testing datasets using conditioned Latin Hypercube Sampling (cLHS). The training dataset was used to cross validate a convolutional neural network (CNN). The model was then trained on the entirety of the training dataset for the prediction of percent sand, percent silt, percent clay, organic content, liquid limit, plastic limit, and soil moisture. The model performance was evaluated using the testing dataset.

To better understand the use of hyperspectral sensing for the characterization of tailings properties (viz. study objective three), tailings from a gold, silver, lead, and zinc producing mine were processed to obtain 100 artificial samples with unique particle size distributions. Each of the 100 samples were prepared to three unique moisture and density conditions and hyperspectral data were captured using a broadband spectrometer (350 – 2500 nm) and the properties of each sample were measured or calculated. The hyperspectral dataset was split into training and testing data. Training data were used to cross validate a CNN then train the model. Testing data were used to evaluate the performance of the model for the predication of percent sand, silt, clay, fines content, solids content, gravimetric moisture content, volumetric moisture content, saturation, void ratio, porosity, total density, and dry density.

Major Findings

This study highlights the potential of remote sensing, machine learning, and spatial analysis techniques for advancing surficial soil property prediction and tailings characterization essential to large-scale applications. Each chapter of this study was prepared as a stand-alone document.

In Chapter 1, soil moisture and soil strength correlations with topographic, vegetation, and soil attributes indicate the strength of the relationships can vary substantially within a microwave remote sensing grid cell. Soil moisture is strongly correlated to topography, vegetation, and soil in Region A, strongly correlated to vegetation and soil in Region C, and somewhat correlated to vegetation and soil in Region D. Correlations between soil moisture and topographic, vegetation, and soil attributes in Region B were generally weak. Soil strength has somewhat strong correlations with soil attributes in Region D but has poor correlations with soil and landscape attributes in all other regions. The regression slopes of the relationships between soil moisture and topographic, vegetation, and soil attributes are also observed to vary substantially within the microwave remote sensing grid cell. Region A has the greatest slopes for relationships between soil moisture and topography, and Regions A and C have greater slopes for relationships of soil moisture with vegetation and soil than Regions B and D. The variations in these relationships highlight the complexity of interactions between environmental factors and soil moisture and strength. Results indicate that soil moisture downscaling procedures, and associated soil strength modeling procedures, are likely to have suboptimal performance when constant parameters are used across a soil moisture downscaling extent.

In Chapter 2, field captured hyperspectral data and a CNN are successfully used to predict surficial soil properties. Soil property predictions from the CNN are compared with SSURGO predictions and results show that predictions from the CNN have lower errors than SSURGO predictions. Predictions of percent sand, silt, and clay using hyperspectral data have RMSEs of 10.8%, 7.5% and 8.5%, respectively, compared to SSURGO predictions of 11.0%, 10.4%, and 9.4% respectively. The CNN modeling procedure also outperformed SSURGO predictions of organic content, liquid limit, and plastic limit with CNN RMSEs of 6.3%, 9.8% and 10.7%, respectively, and SSURGO predictions with RMSEs of 10.5%, 16.5% and 17.7%, respectively. Additionally, CNN predictions better capture the variability in soil properties than SSURGO for percent sand, percent silt, percent clay, organic content, liquid limit, and plastic limit, with the CNN resulting in better coefficients of determination (COD) compared to SSURGO for all compared soil properties. Finally, although SSURGO does not provide estimates of soil moisture, the CNN yielded an RMSE of 0.04 cm^3/cm^3 and an COD of 0.66 for soil moisture predictions. The use of field captured hyperspectral imagery under uncontrolled lighting conditions shows promise for the characterization of undisturbed surficial soil samples.

In Chapter 3, hyperspectral data captured on tailings samples and a CNN are successfully used to predict tailings properties. Predictions from the CNN show promise for the estimation of tailings percent sand, percent silt, percent clay, fines content, solids content, gravimetric water content, volumetric water content, and saturation with RMSEs of 5.4%, 5.1%, 1.2%, 5.3%, 1.4%, 2.2%, 2.8%, and 8.3%, respectively Additionally, the prediction of void ratio, porosity, total density, and dry density show promise (RMSEs of 0.23, 0.05, 0.16 Mg/m^3, and 0.14 Mg/m^3 respectively), especially when samples have a high degree of saturation (RMSEs reduced to 0.01, 0.03, 0.07 Mg/m^3, and 0.07 Mg/m^3 for saturations greater than 50%, respectively). The results indicate hyperspectral sensing could serve as a method to rapidly characterize tailings properties relevant to geotechnical engineering.

1. SOIL MOISUTRE AND STRENGTH VARY WITH TOPOGRAPHY VEGETATION AND SOIL WITHIN A SMAP GRID CELL

1.1. Introduction

Understanding surficial soil strength across a landscape is important for assessing vehicle mobility for recreational, agricultural, and military purposes (Wulfohn et al., 1996; Priddy et al., 2012; Uusitalo et al., 2018; Schönauer et al., 2022). However, soil strength can be highly variable spatially and temporally due to variations in soil moisture and soil texture (percent gravel, sand, silt, and clay) (Fredlund, 2006). Therefore, fine-resolution (5- to 30-m) soil strength estimates are required.

For vehicle mobility applications, surficial soil strength is commonly quantified using the Rating Cone Index (RCI), which is calculated by multiplying the cone index (CI) by the remold index (RI) (Stevens et al., 2013). Cone index can be measured in the field by pressing a cone penetrometer into the soil profile at a constant rate and measuring the pressure required to advance the cone. The remold index quantifies the tendency of a soil to lose strength under vehicle loading and is determined on a remolded soil sample (Stevens et al., 2013). For coarse grained soils, the RI is 1, so RCI and CI are interchangeable.

Field measurement of RCI is impractical across large spatial extents, so the RCI is often inferred from soil composition and moisture information. For example, the Strength of Surface Soils (STRESS) model uses average soil strength and soil-water retention parameters for texture classifications to estimate fine resolution surficial soil strength (Pauly, 2019). Other soil strength modeling procedures include the North Atlantic Treaty Organization Reference Mobility Model (NRMM), which empirically estimates vehicle mobility based on vehicle, terrain, and environmental factors (Jurkat et al., 1975; Haley et al., 1979; Ahlvin and Haley; 1992).

McCollough et al. (2017) introduced the Next Generation NRMM (NG-NRMM), which implements complex terramechanics models (Choi et al., 2019) and requires a comprehensive understanding of soil properties. However, obtaining high fidelity fine resolution soil moisture and soil composition data can be difficult.

Soil moisture products are available nearly globally at resolutions ranging from 9- to 60-km grid cells from microwave satellites. Such products are generated by the Advanced Microwave Scanning Radiometer-EOS (AMSR-E) (Njoku et al., 2003), Soil Moisture Ocean Salinity (SMOS) (Kerr et al, 2010), and Soil Moisture Active Passive (SMAP) (O'Neill et al., 2010) instruments. However, these datasets do not have the fine spatial resolution needed for soil strength and vehicle mobility modeling. Various methods have been proposed to downscale coarse resolution soil moisture data. Coarse resolution microwave data has been downscaled to 1 km grid cells based on fine-resolution soil temperature and land-atmosphere interaction parameters (Merlin et al., 2006). Optical/thermal data are available at fine resolutions (tens of meters) and can be used to directly estimate fine resolution soil moisture based on vegetation indices from optical data and land-surface temperatures from thermal data (Sadeghi et al., 2017). Data from optical/thermal sources have also been used for soil moisture downscaling procedures (Peng et al., 2016; Fang et al., 2018). Other approaches for downscaling soil moisture rely on relationships between soil moisture and topographic, vegetation, and soil attributes. For example, the Equilibrium Moisture from Topography, Vegetation, and Soil (EMT+VS) model (Coleman and Niemann, 2013; Ranney et al, 2015; Cowley et. al, 2017; Deshon et al., 2020) can be used to downscale soil moisture to 10- to 30-m grid cells using fine-resolution topographic, vegetation, and soil data.

Some soil moisture downscaling methods that are informed by soil and landscape attributes currently assume the relationships between soil moisture and topography, vegetation, and soil composition are constant throughout a coarse resolution grid cell (Cowley et. al, 2017). Ali et al. (2010) studied the spatial relationships between soil moisture and topographic variables at different scales and found that for larger extents (0.54 ha – 1.4 ha), the dependence of soil moisture on topography is significant, but for smaller extents (0.02 ha – 0.54 ha), topography does not have strong correlations to soil moisture. Furthermore, soil moisture estimates can be inaccurate when model parameters are adjusted uniformly for the entire modeling extent (Mascaro et al., 2011), which indicates non-uniform relationships between soil moisture and soil and landscape attributes. Yu et al. (2018) conducted statistical analyses on the effects of micro-topography and vegetation type on soil moisture in a large gully (49 ha) and found that soil moisture has varying dependence on topographic attributes, depending on the location within the catchment. Additionally, Liang et al. (2017) studied the dependence of soil moisture on topographic, soil, and vegetation features at a site (0.16 ha) within a small headwater catchment and determined that soil moisture is most strongly correlated with different topographic, soil, and vegetation features within various portions of the catchment. While such studies have highlighted the variability in soil moisture dependencies within small spatial extents (Mohanty et al., 2000; Western et al., 2004; Lin et al, 2006; Kim et al., 2007; Rosenbaum et al., 2012), there is less understanding of how soil moisture and soil strength dependencies vary among different regions within the spatial extent of a microwave remote sensing grid cell. An improved understanding of how these dependencies vary within such extents may better inform soil moisture and associated soil strength modeling procedures.

The objective of this study is to examine the spatial variability in the dependence of both soil moisture and soil strength on topographic, vegetation, and soil composition attributes across a 9-km SMAP grid cell. Four geographically separated regions with varying soil, topography, and vegetation were selected within a 9-km region. Soil moisture and soil strength were measured at 86 sampling locations on 10 dates between May and August of 2022. Soil samples were collected at each sampling location and used to determine soil texture, plasticity, and organic content. Landscape attributes were derived from a digital elevation model (DEM), and vegetation indices were determined from multispectral satellite products. The Pearson correlation coefficient (r) was used to quantify the dependence of soil moisture and soil strength on soil and landscape attributes within each sampling region, and Pearson p-values were used to assess the statistical significance of the correlations. To assess how the significant relationships deviate from the average behavior, the relationship slopes in each region were compared to the slopes for the entire dataset.

1.2. Materials and Methods

1.2.1. Study Site

The study site (Maxwell Ranch) is approximately 50 km northwest of Fort Collins, Colorado in the Laramie Foothills, which mark the transition between the Great Plains and the Rocky Mountains. Maxwell Ranch includes 4000-ha of land, spanning approximately 9 km in both the north-south and east-west directions. Elevations at the site range from 1964 m to 2286 m with an average slope of 0.22 m/m. Soils are predominately coarse loams weathered from granite and (to a lesser extent) sandstone. Vegetation includes grasses, shrubs (primarily Antelope Bitterbrush and Mountain Mahogany), and pine forests (primarily Ponderosa and Piñon pine) (MRLC, 2019).

Four sampling regions (A, B, C, and D) were identified to capture the varying topography, vegetation, and soil conditions at the study site (Figure 1.1a). Region A, located in the northernmost portion of the study site, is characterized by forested hillslopes and a grassy valley bottom with a stream (Figure 1.1b). Region B consists of rolling hills and open-canopy Ponderosa pine forests (Figure 1.1c). Region C is comprised of rolling grassland hills with a grassy valley bottom containing a stream (Figure 1.1d). Region D is located near the southern boundary of the study site and is predominately vegetated by Mountain Mahagony shrubs and has several rock outcrops (Figure 1.1e). Within each region, sampling locations were selected to observe diverse soil, topography, and vegetation conditions. Regions A, B, C, and D contain 22, 22, 24, and 18 sampling locations, respectively.

The site is in a semi-arid continental climate and receives approximately 430 mm of annual precipitation, with most precipitation events occurring during the winter as snowfall. Soil moisture and strength sampling were conducted between May and August of 2022 to avoid the frozen soil and snow cover that can occur during winter. A private weather station (KCOLIVER 12) is located immediately west of the ranch. Figure 1.2 shows the daily precipitation during the sampling period and average soil moisture ($\bar{\theta}$) at the sampling locations (HydraProbe) and for the SMAP extent on each sampling date. Sampling dates were selected to capture a wide range of soil moisture conditions. Two major precipitation events occurred before sampling on 06-02 (MM-DD) and 07-27, and two dry-down periods occurred approximately from 06-01 to 07-05 and from 07-28 to 08-05 (Figure 1.2). The difference between HydraProbe average soil moisture and SMAP soil moisture was generally less than 0.05 cm^3/cm^3. The greatest difference is observed on 07-05 because a precipitation event occurred within the SMAP grid cell but not at the sampling regions.

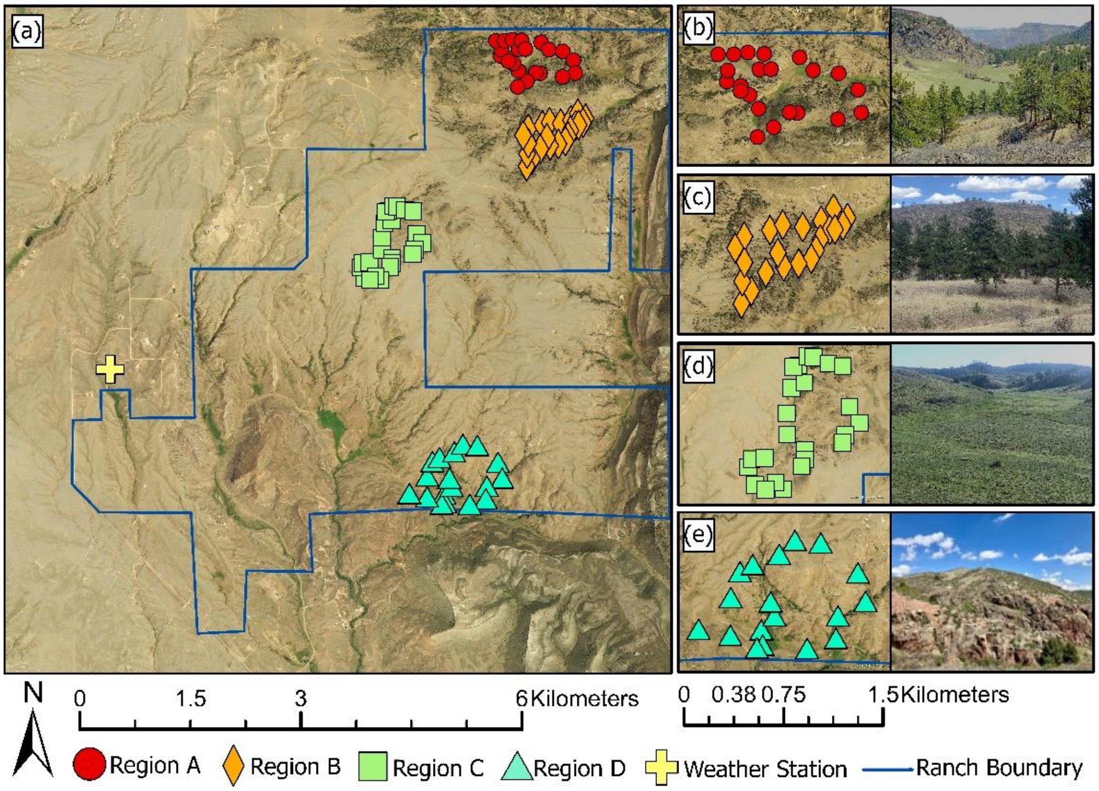

Figure 1.1 (a) Map of Maxwell Ranch including ranch boundary, study regions, weather station, and sampling locations (map extent corresponds to a SMAP 9-km grid cell); (b-e) maps of Regions A, B, C, and D with corresponding sampling locations and photos at each region.

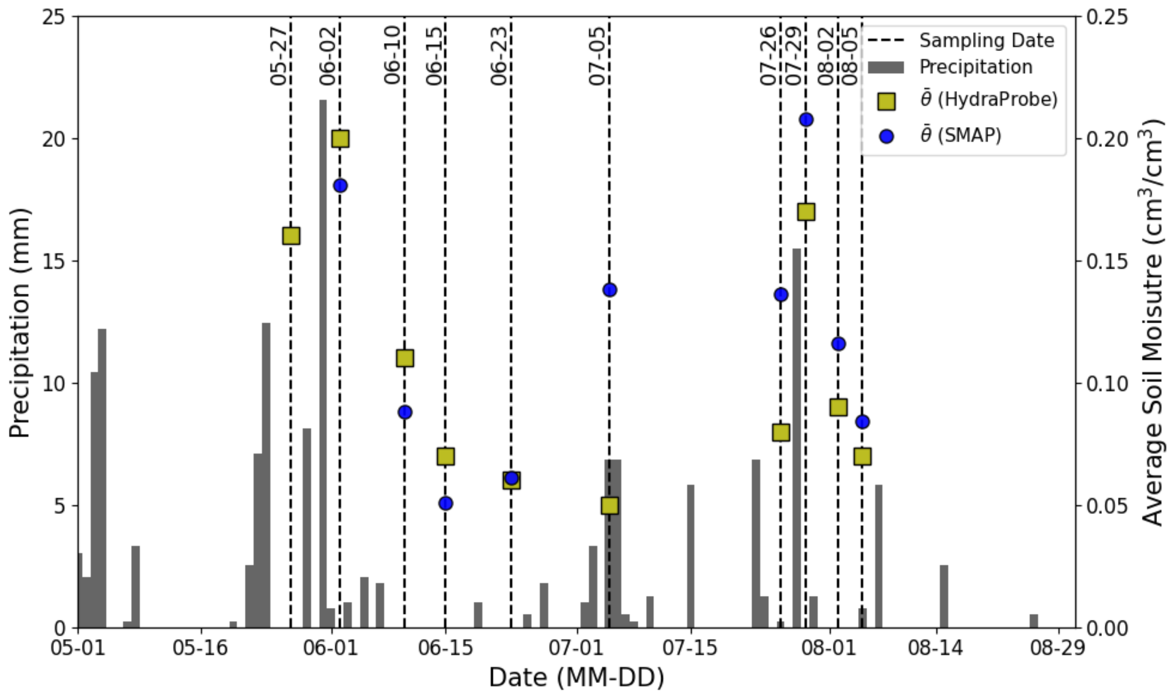

Figure 1.2. Daily precipitation, average soil moisture at sampling locations (HydraProbe), and SMAP soil moisture for sampling dates in 2022. SMAP soil moisture was determined using the Enhanced Level 3 Radiometer data product (Chan et al., 2018; O'Neill et al., 2021). SMAP data were unavailable for 05-27. Precipitation data were recorded by the KCOLIVER 12 weather station. Dashed vertical lines represent sampling dates.

1.2.2. Data Collection

1.2.2.1. Topography and Vegetation Characterization

Remote sensing data products were used to characterize topography and vegetation at the sampling locations and the 9-km region associated with the SMAP Enhanced Level 3 Radiometer product (Chan et al., 2018; O'Neill et al., 2021). The DEM was obtained at 10-m resolution from the National Elevation Dataset (USGS, 2022) and processed using the Terrain Analysis Using DEM tools in ArcGIS (TAUDEM) (Tarboton, 2008). Topographic attributes were selected based on prior use in soil moisture downscaling (Cowley et al., 2017; Eylander et al., 2023), including elevation, log of specific contributing area (log SCA), slope, cosine of aspect, and potential solar radiation index (PSRI) for the summer solstice, which is calculated using slope, aspect, latitude, and date of the summer solstice (Lee, 1964; Swift, 1976; Dingman, 2015).

The density of green vegetation cover was characterized using multispectral imagery at 10-m resolution from the Sentinel 2 satellite (ESA, 2016). Due to cloud cover on multiple dates during the sampling period, a single image from 05-26 was used to calculate the Normalized Difference Vegetation Index (NDVI) (Kriegler et al., 1969) and the Enhanced Vegetation Index (EVI) (Liu and Huete, 1995; Huete et al., 1997). NDVI is calculated using only the red and near-infrared bands, while EVI also uses the blue band.

Soil data were retrieved from the Soil Survey Geographic Database (SSURGO) and used to obtain U.S. Department of Agriculture (USDA) soil composition including percent sand, silt, and clay across the region (Soil Survey Staff, 2023).

1.2.2.2.Soil Moisture Measurement

Surface soil moisture (θ; cm^3/cm^3) measurements were recorded at each sampling location and date using a Stevens HydraProbe Soil Sensor (Stevens Water Monitoring Systems Inc., 2018). The HydraProbe has four 5.7 cm-long, 0.3 cm diameter measurement tines and uses dielectric impedance to determine soil moisture with an accuracy of ± 0.03 cm^3/cm^3 and precision of ± 0.001 cm^3/cm^3 (Stevens Water Monitoring Systems Inc., 2018). At each sampling location, three HydraProbe measurements were collected within approximately 1 m of the sampling location and averaged to determine the soil moisture of the sampling location.

1.2.2.3.Soil Strength Measurement

Soil strength was measured using a cone penetrometer with a 1 in^2 (6.45 cm^2) cone and load cell with 0.1 psi (0.7 kPa) accuracy (Humboldt Mfg., 2023). The penetrometer cone tip was advanced into the soil at a constant rate until the cone base reached 1 in (2.54 cm) below the soil surface. The cone penetrometer measures the force required to advance the cone into the soil, which can be divided by the cone area to produce CI. The RI was not measured for fine grained soils at the site (which are rare); therefore, the CI (kPa) is reported for strength measurements. Three cone penetrometer measurements were collected within approximately 1 m of each sampling location and averaged to determine the soil strength for that sampling location.

1.2.2.4.Soil Texture Sampling

Soil samples were collected at each sampling location for laboratory soil composition analysis. Three soil samples were collected within approximately 1 m of each sampling location and approximately within the top 5 cm of the soil surface. Vegetation and leaf litter were removed from the soil surface prior to sampling. The three soil samples collected at each sampling location were combined for laboratory analyses.

1.2.2.5.Soil Characterization

Sedimentation analysis was completed using the *PARIO Soil Particle Analyzer* to characterize the distribution of particles ranging in size from 2-63 μm according to the PARIO Classic Procedure (mass fraction detection error: ±0.03 g/g) (METER Group, 2017). Representative sedimentation specimens were obtained according to the moist specimen procurement procedure described in ASTM D7928 (ASTM International, 2021a). Data evaluation by the PARIO software is based on the integral suspension pressure method (Durner et al., 2017) and was used to determine percent sand, silt, and clay, after removal of gravel, based on the USDA particle size classifications. Porosity and saturated hydraulic conductivity K_s were estimated using USDA percent sand and clay in pedotransfer functions (Cosby et al., 1984). Percent sand, silt, and clay from the sedimentation analysis were also used to classify the soil according to the USDA soil classification system (USDA, 1975).

The distribution of soil particle sizes greater than 75 μm was characterized using mechanical sieve analysis, which was performed following ASTM D6913 (ASTM International, 2017c) standard. Representative sieve specimens were obtained according to the moist specimen procurement procedure, and the 25.4 mm, 19.0 mm, 9.5 mm, No. 4, No. 10, No. 20, No. 40, No. 60, No. 100, and No. 200 sieves were used. Data from the mechanical sieve analysis were processed following the ASTM D6913 (ASTM International, 2017c) standard to obtain particle-size distributions and related characteristics including the particle diameter corresponding to 10% mass finer (D_{10}), 30% mass finer (D_{30}), 60% mass finer (D_{60}), and coefficient of uniformity (C_u). The parameter C_u is defined as the ratio of D_{60} to D_{10} and provides information about the soil gradation (i.e., sorting) where greater values correspond to more well graded (poorly sorted) soils.

Soil plasticity testing is used to inform soil classification in the Unified Soil Classification System (USCS) for soils with more than 5% fines. Soil plasticity testing was conducted to obtain soil liquid limit (LL), plastic limit (PL), and plasticity index (PI) following ASTM D4318 (ASTM International, 2017b); samples were prepared using the dry preparation procedure. The LL was determined using the *Multipoint Method*, which typically has greater precision than the *One-Point Method* (ASTM International, 2017b). The PL was determined using the *Hand Rolling* procedure (ASTM International, 2017b). The USCS classification was determined following ASTM D2487 (ASTM International, 2017a) and based on the particle size distribution from mechanical sieve analysis as well as the LL and PI.

Soil organic content was also determined. Organic content was calculated using ASTM D2974 *Method A* (ASTM International, 2020) to the nearest percent of total dry mass.

1.3. Results

1.3.1. Remote Sensing Data

Figure 1.3 shows the histograms of the topographic, vegetation, and soil composition attributes at the sampling locations and across the SMAP grid cell. For most attributes, the distribution at the sampling locations generally resembles the distribution across the SMAP grid cell. However, the sampling locations have more points at high elevations and steeper slopes than the larger SMAP grid cell. Additionally, unlike the SMAP grid cell, the sampling locations do not contain any locations with high clay content according to the SSURGO data (Figure 1.3).

Figure 1.4 shows maps of topographic attributes, vegetation indices, and USDA percent sand, silt, and clay from SSURGO. The sampling points in Region D have the lowest average elevation (2067 m), while the points in Regions A, B, and C have similar average elevations (2163, 2172, and 2171 m, respectively). Regions A, B, C, and D have total reliefs of 85, 53, 69, and 72 m, respectively. The sampling points in Region B have a larger average log SCA (5.4 m) than the points in Regions A, B, and C (averages of 4.0, 4.2, and 3.8 m, respectively). The sampling points in Regions A and D have steeper slopes (averages of 0.28 and 0.20 m/m, respectively) than points in Regions B and C (averages of 0.17 and 0.16 m/m, respectively). Regions A and B tend to have more north- and south-facing sampling locations because the major valleys flow to the east. Regions C and D tend to have more east- and west-facing points because the main valleys flow to the south. The sampling points within all four regions tend to have similar ranges of PSRI values. Regions A and B tend to have higher NDVI values (averages of 0.22 and 0.22, respectively) than Regions C and D (averages of 0.20 and 0.18, respectively) due to greater forest cover. The EVI values follow similar patterns to the NDVI values. According to SSURGO, the soils are relatively uniform across the study site. However, Region

D is expected to have less sand and more silt than the other regions. Region C is expected to have the least silt, and Region A is expected to have the least clay.

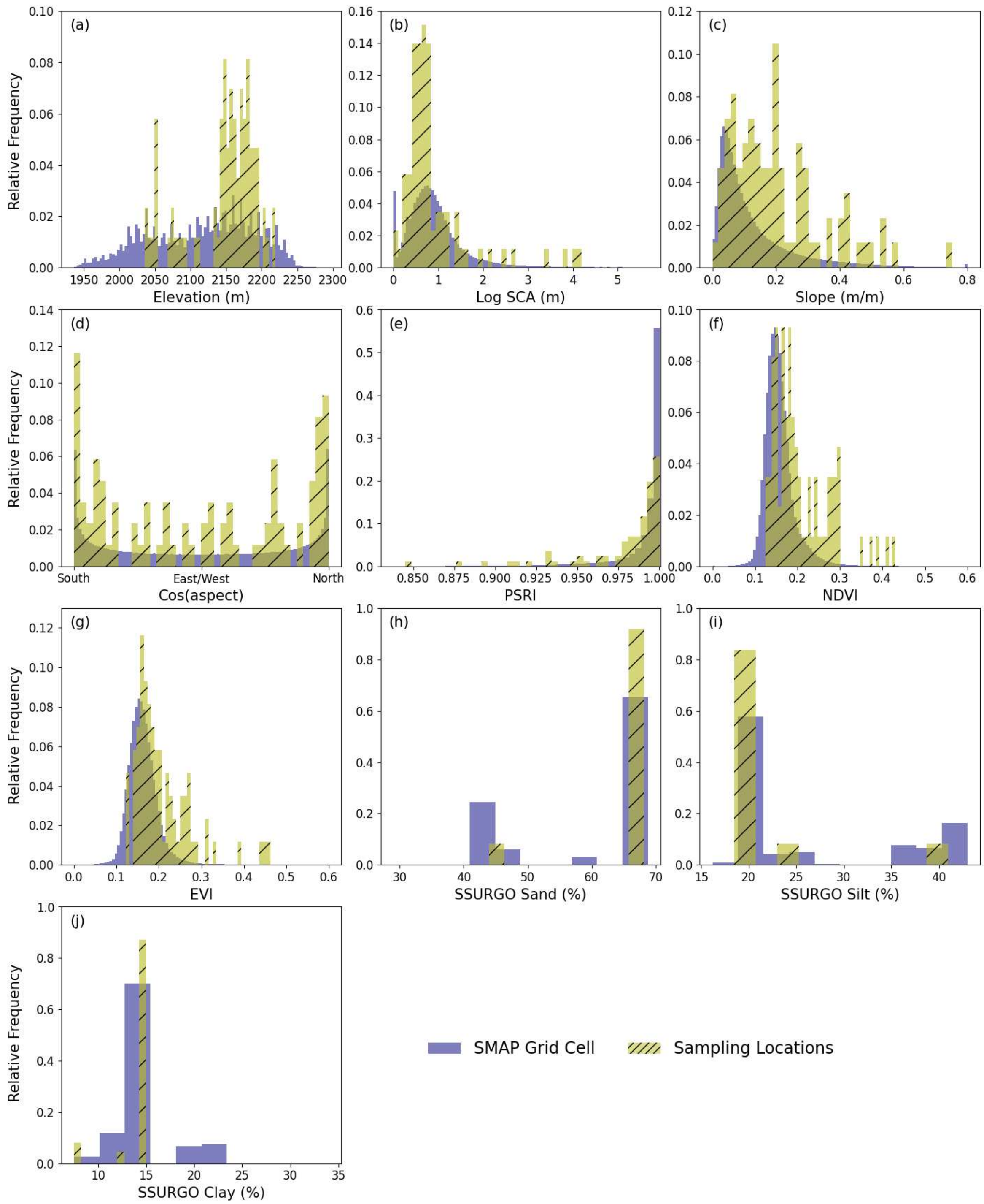

Figure 1.3. Histograms of the relative frequency of topographic, vegetation, and soil attributes for the SMAP grid cell (unhatched bars) and the sampling locations (hatched bars). Histograms include (a) elevation, (b) cos(aspect), (c) slope, (d) log SCA, (e) PSRI, (f) NDVI, (g) EVI, (h) SSURGO percent sand, (i) SSURGO percent silt, and (j) SSURGO percent clay.

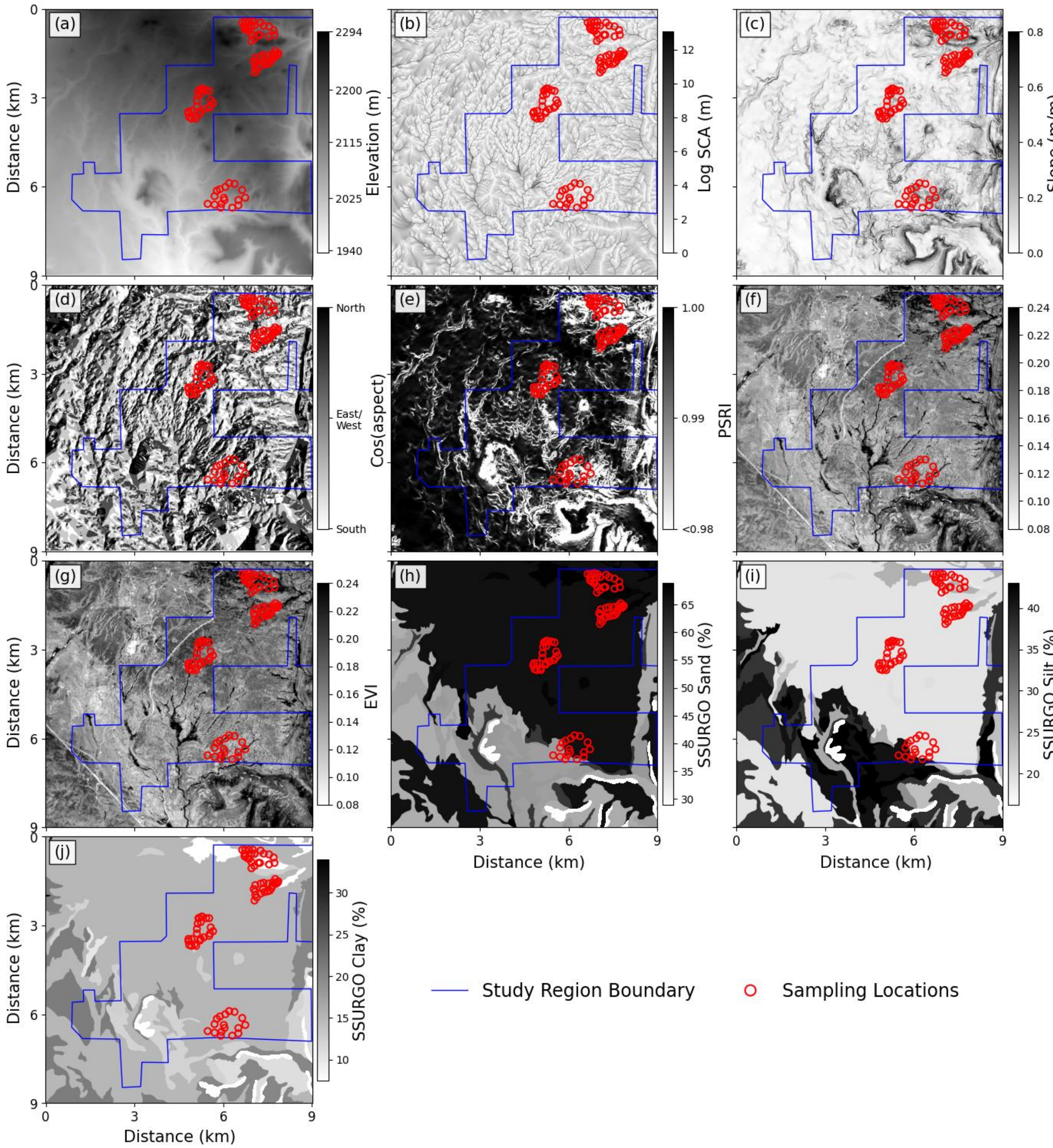

Figure 1.4. Maps of (a) elevation, (b) cos(aspect), (c) slope, (d) log SCA, (e) PSRI, (f) NDVI, (g) EVI, (h) SSURGO percent sand, (i) SSURGO percent silt, and (j) SSURGO percent clay along with sampling locations and ranch boundary.

1.3.2. Laboratory Soil Composition Data

Figure 1.5a displays the USDA texture classifications of the soils collected at the sampling locations. Most soils are classified as sand, loamy sand, sandy loam, sandy clay loam, and loam, with a few soils classifying as clay, clay loam, and sandy clay loam. Contrary to

SSURGO, the soils from Region D have the highest average USDA percent sand (75%), which is approximately 10% higher than Regions A, B, and C (averages of 66%, 66%, and 65%, respectively). Regions B and C do not contain any soil classified as USDA sand. The soils from Regions A, B, C, and D have similar average USDA percent silt (20%, 20%, 23%, and 19%, respectively). Region C has the most silt, which again contrasts with SSURGO. The soils from Regions A and B have the highest average clay content (13% and 14%, respectively) while Region C and D have lower mean clay contents (8% and 5%, respectively). Region D soils classified as either USDA sand, loamy sand, or sandy loam, with all soils containing less than 20% clay (Figure 1.5a).

As displayed in the soil plasticity chart (Figure 1.5b), the collected soils generally have LL below 50 and plot below the A-line, which distinguishes between silts and clays. The predominant fines classifications are low plasticity silt and low plasticity organic (ML or OL). Soils at Regions A and C have the highest average soil plasticity (LL $\approx$ 50%; PL $\approx$ 42%), compared to LL $\approx$ 44% and PL $\approx$ 39% at Regions B and D. Using the USCS soil classification system, 67% of the sampling locations are classified as silty sands (SM), and 16% are classified as gravels. A few sampling locations are classified as high plasticity silt (MH), poorly graded sand (SP), and silty clayey sand (SC-SM). Several samples were classified as peat according to the ASTM D2487 (ASTM International, 2017a) standard. Soil organic content ranges from 1% to 70% at the sampling locations. Regions B and C contain greater average soil organic content (14% and 15%, respectively) compared to Regions A and D (9% and 8%, respectively).

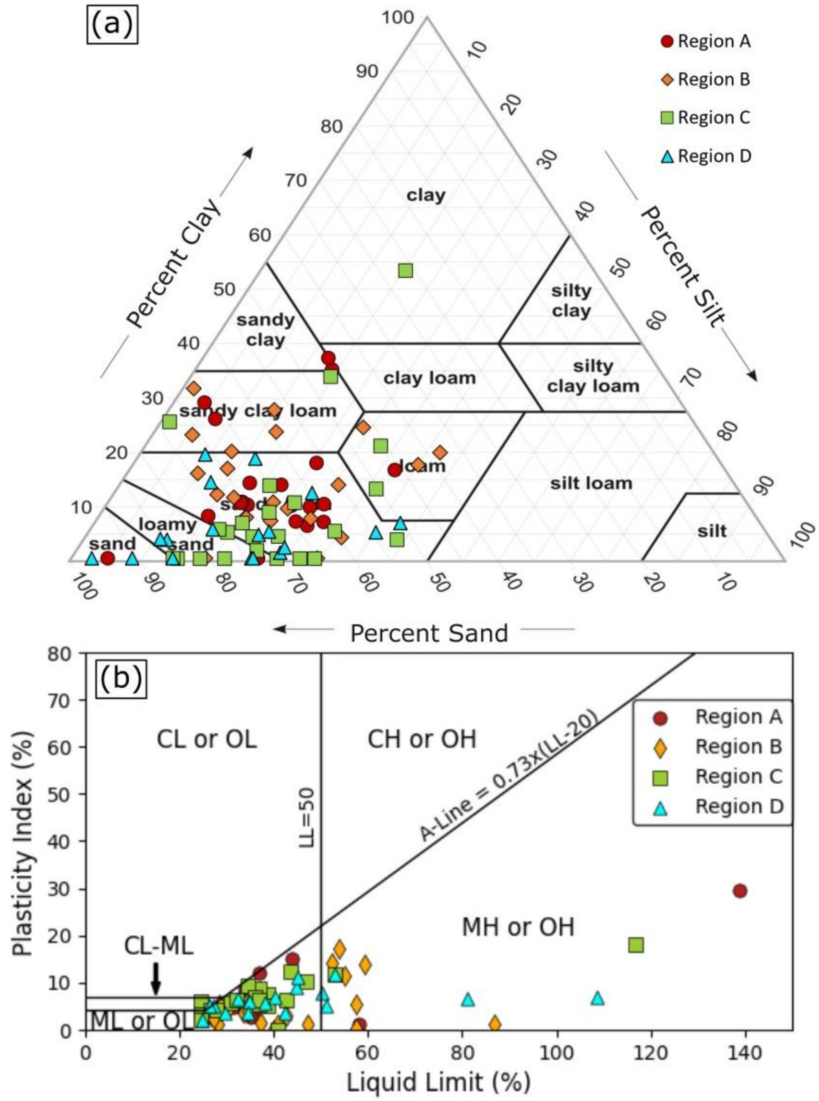

Figure 1.5. (a) USDA soil texture classification for all sampling locations; (b) Atterberg limit summary for all sampling locations where CL is low plasticity clay, ML is low plasticity silt, CH is high plasticity clay, MH is high plasticity silt, OL is low plasticity organic material, and OH is high plasticity organic material.

1.3.3. Soil Moisture and Strength Data

Soil moisture observations from a representative wet date and dry date are shown in Figure 1.6. Overall, the sampling locations in the valley bottoms are wet throughout the study period while the sampling locations in the upland areas exhibit more temporal variability. Regions A and C tend to have higher average soil moisture than Regions B and D.

Figure 1.7 shows the histogram of soil moisture for each sampling region and date. Most soil moisture measurements are less than 0.25 cm^3/cm^3, with only a few measurements exceeding that value. Regions A and C have more sampling locations with soil moisture greater than 0.25 cm^3/cm^3 than the other regions. Region D contains very few wet observations. Sampling locations with the highest soil moisture were generally classified as USCS peat, which can have high porosity (Rezanezhad et al., 2016).

Soil strength (CI) maps for a representative wet and dry date are shown in Figure 1.8. CI is highly variable among sampling locations and among dates at a given sampling location. Region C typically has greater CI values than the other regions, while Regions A, B, and D have more variability through time.

Figure 1.9 shows the histograms of CI for each region and sampling date. Regions A and B typically have similar distributions of soil strength, while Regions C and D are more dissimilar. On every date, the distribution of at least one region differs notably from the distributions in the other regions.

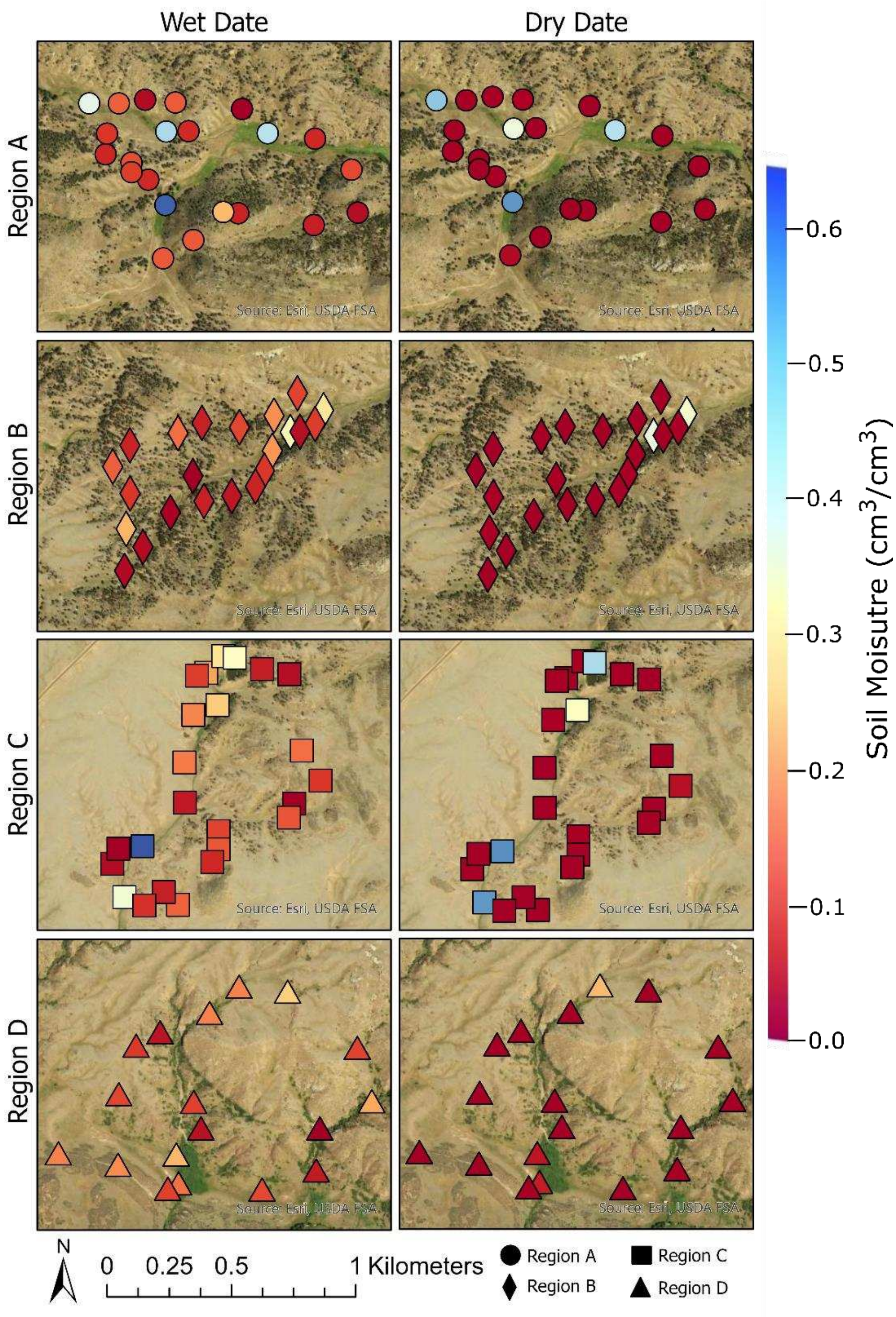

Figure 1.6. Maps of soil moisture for each region on a dry date (06-23) and a wet date (07-29). symbols represent sampling locations and infill colors represent measured soil moisture.

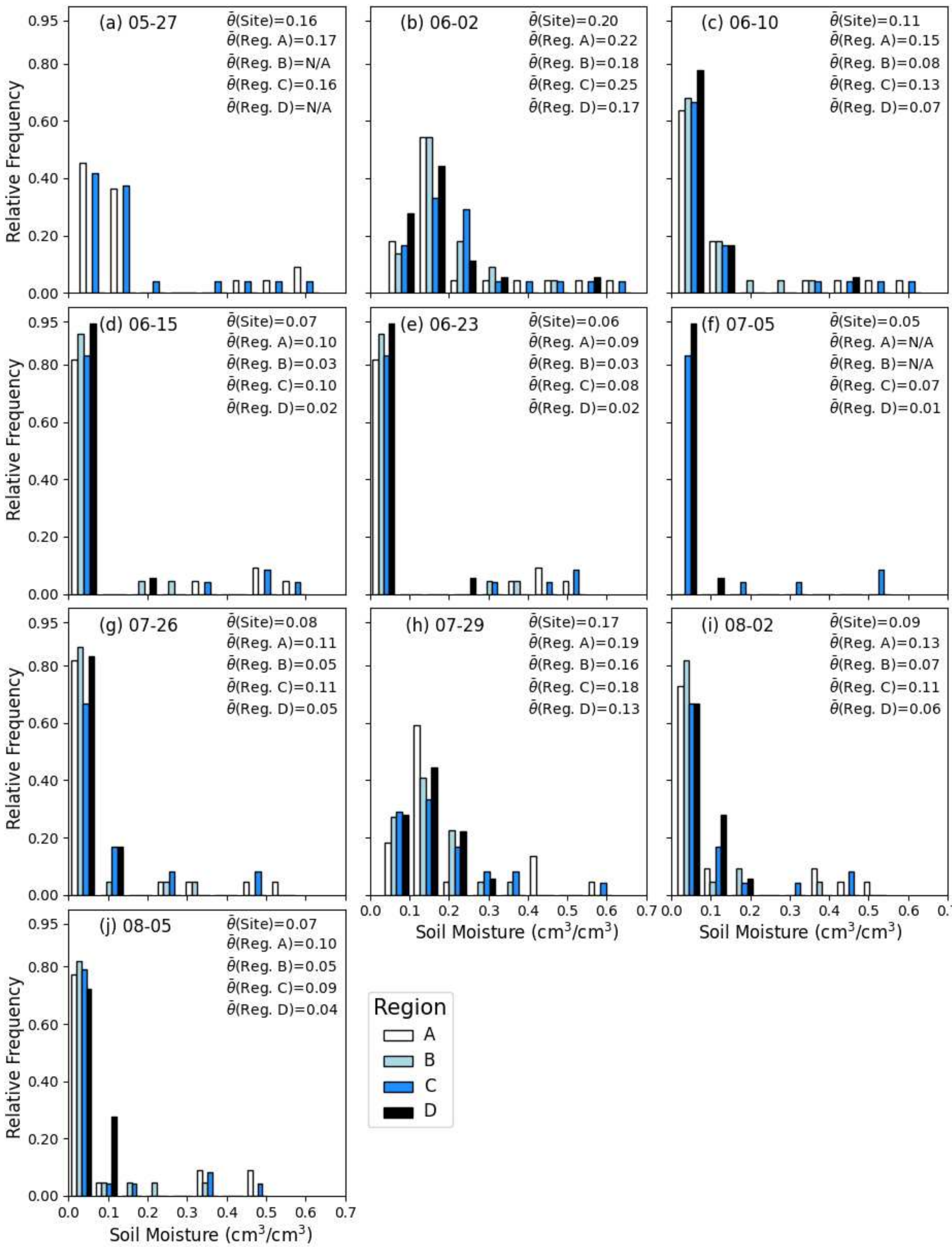

Figure 1.7. Histograms of the soil moisture measurements collected in each region for each sampling date. Spatial average soil moisture ($\bar{\theta}$) (cm^3/cm^3) for the entire study site and each study region are shown in the subplots. Soil moisture was not measured for all regions on 05-27 and 07-05 due to inclement weather.

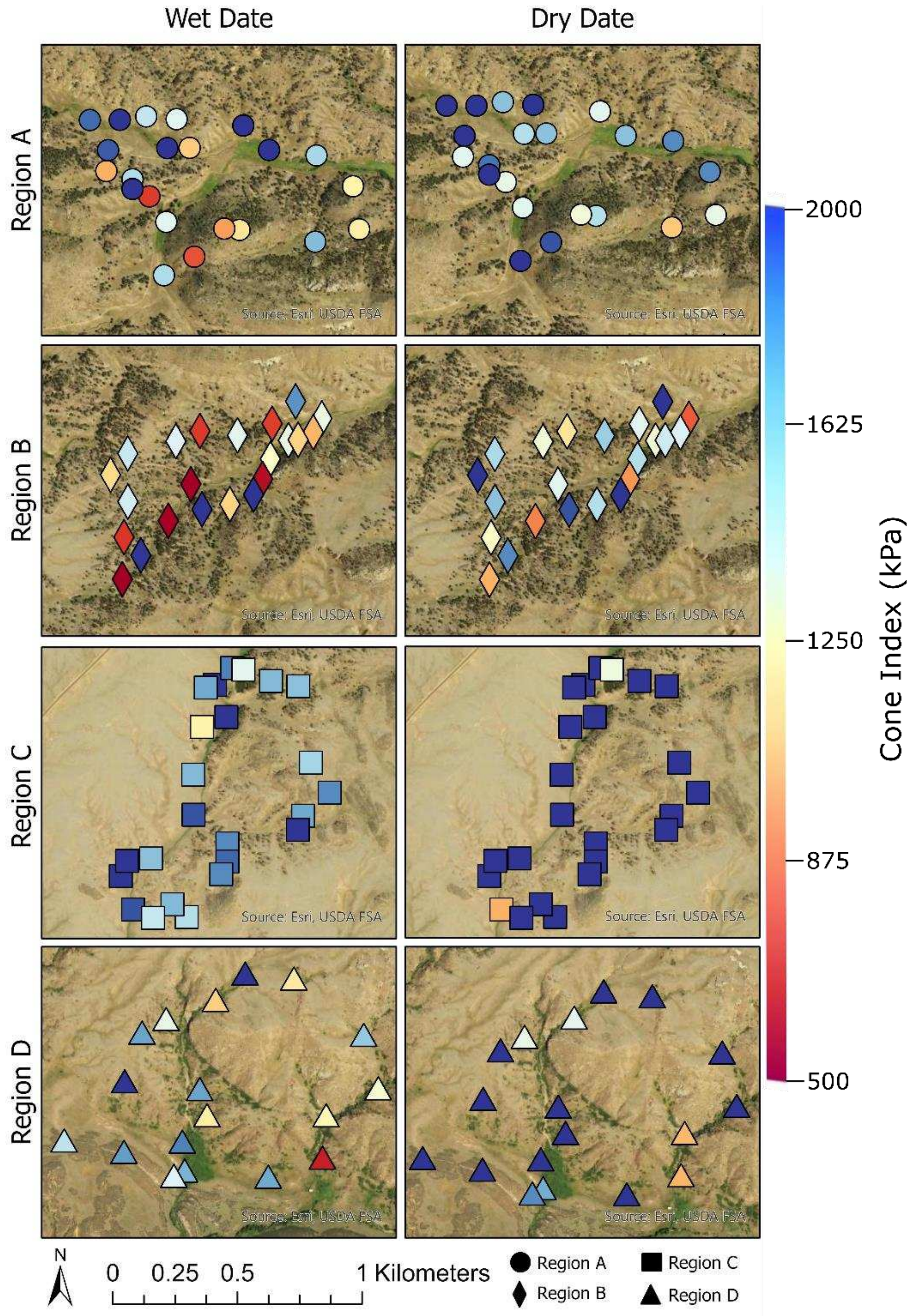

Figure 1.8. Maps of CI for each region on a dry date (06-23) and a wet date (07-29). Symbols represent sampling locations and infill color represents measured cone index (CI).

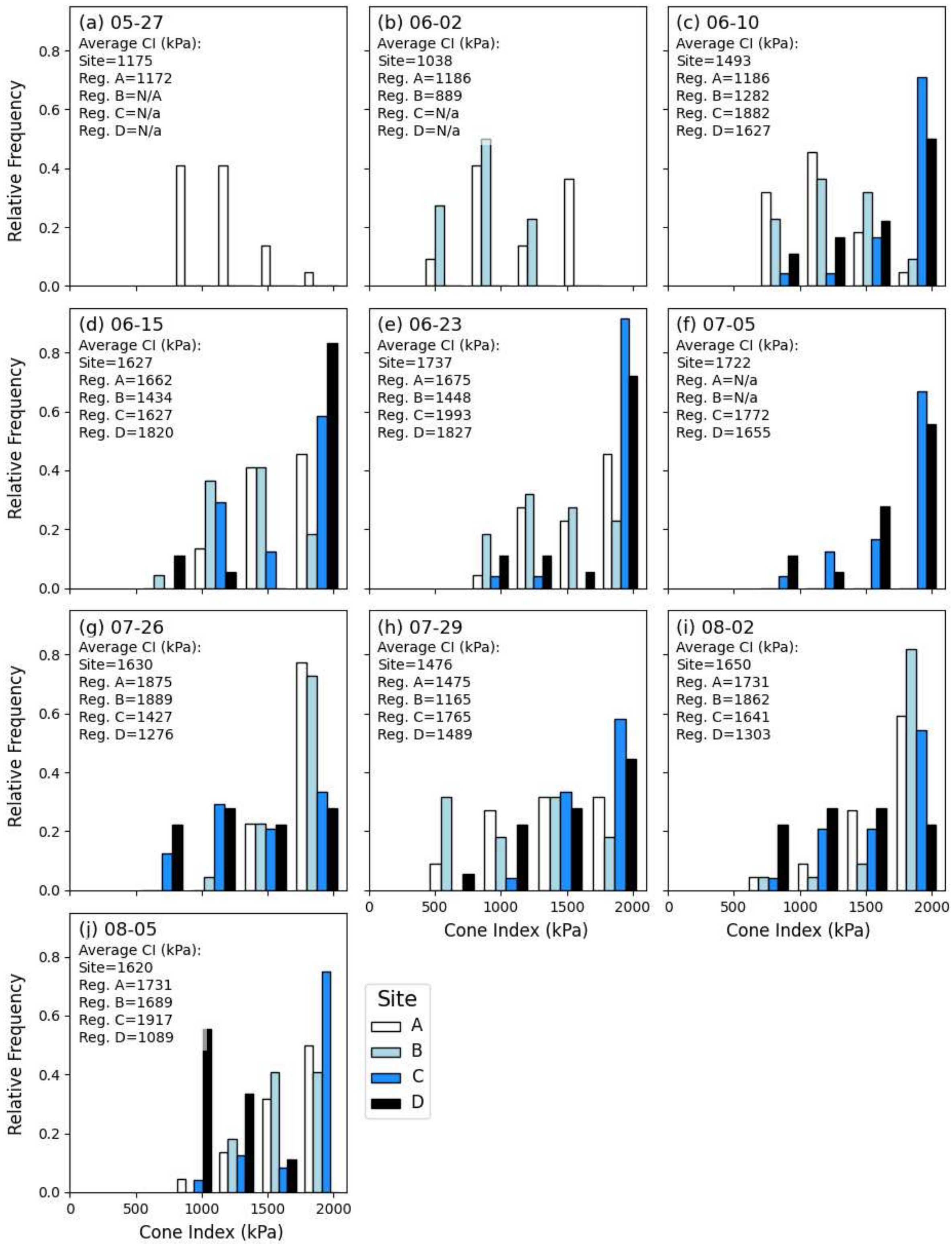

Figure 1.9. Histograms of cone index (CI) measurements collected at each location for each sampling date. Average CI for the entire study site and each study region are shown within the subplots. Not all region's CI was measured on 05-27, 06-02, and 07-05 due to inclement weather.

1.3.4. Relationships between Soil Moisture/Strength and Regional Attributes

Figure 1.10a displays the correlations between the vegetation indices (NDVI and EVI) and the topographic attributes for each region. In Region A, the vegetation indices are significantly correlated with log(SCA), slope, and cos(aspect). The correlations indicate that thicker vegetation occurs in valley bottoms, flatter slopes, and north-facing aspects. In Region C, the vegetation indices are again significantly correlated with slope but also with PSRI (thicker vegetation is occurring at locations receiving more summer sun). The vegetation indices in Regions B and D generally have weak correlations with the topographic attributes.

Figure 1.10b displays the correlations between the soil composition attributes and the topographic and vegetation attributes for each sampling region. Overall, the vegetation indices are stronger predictors of soil composition than the topographic indices. The correlations between the vegetation indices and soil attributes are stronger in Regions A and C and weaker in Regions B and D. EVI and NDVI typically have similar correlations to the soil attributes. The correlations between the topographic attributes and the soil attributes also vary among regions. Overall, they are strongest in Region A and weakest in Region C. In Region A, log SCA and cos(aspect) are the best predictors (among the topographic attributes) of soil composition. The correlations imply that finer soils tend to occur in valley bottoms and on north-facing slopes. In Regions B and D, slope and PSRI are the best predictors of the soil attributes with finer soils occurring on flatter slopes and slopes that receive more summer insolation.

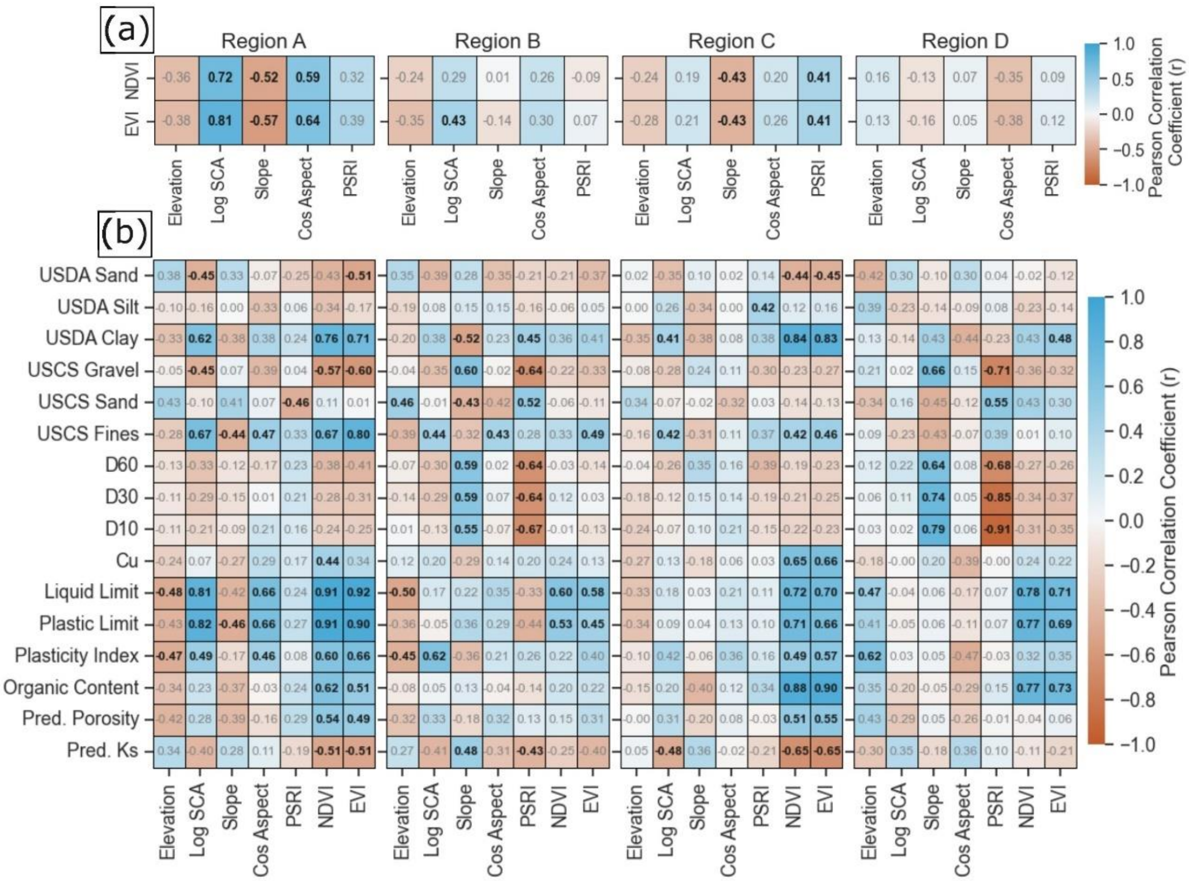

Figure 1.10. (a) Pearson correlations of vegetation indices with topographic attributes for each region and (b) Pearson correlations of soil attributes with topographic and vegetation indices for each region. Correlations displayed with bold black text represent statistically significant correlations based on the Pearson p-value ($p < 0.05$). Other correlations are displayed in grey text.

Figure 1.11 displays the correlations of soil moisture and soil strength to the topographic, vegetation, and soil composition attributes for each sampling region. Overall, soil moisture is predicted better by vegetation and soil composition than topography. However, the correlations between the vegetation indices and soil moisture vary among regions. They are stronger for Regions A and C and weaker for Regions B and D (similar to the previously noted correlations between vegetation and soil composition). NDVI and EVI have similar correlations with soil moisture, which again mimics the soil composition results and suggests that wetter conditions occur where vegetation is thicker.

The correlations between topography and soil moisture also vary among the regions (Figure 1.11). These correlations are strongest in Region A, moderate in Region B, and weak in Regions C and D. The most important topographic attributes for predicting soil moisture vary among regions. In Region A, log SCA, slope, and cos(aspect) are most correlated with soil moisture (again, similar to the earlier soil composition results). These correlations suggest that wetter conditions occur in valley bottoms, flatter slopes, and north-facing aspects. In Region B, log SCA and elevation are most correlated with soil moisture.

The correlations between the soil composition attributes and soil moisture also vary substantially among regions (Figure 1.11). They are generally stronger in Regions A and C and weaker in Regions B and D. The most important composition variables for predicting soil moisture vary among the regions. USDA percent sand and USCS fines are strongly correlated with moisture in all regions except Region D, suggesting that wetter conditions occur in soils with high percent fines (silt and clay). USDA percent clay is strongly correlated with moisture only in Regions A and C, indicating that clayey soils (which correspond with denser green vegetation in Regions A and C as noted earlier) are generally wetter. LL, PL, and organic content are strongly correlated with soil moisture for all regions except Region B which, again, supports the interdependence between soil moisture, vegetation, and soil composition characteristics.

Figure 1.11 also shows the correlations of soil strength (viz. CI) with the topographic, vegetation, and soil attributes. Overall, soil strength has weak correlations with the explanatory variables and the correlations are weaker for the topographic and vegetation attributes than the soil composition characteristics. In Region A, elevation has significant negative correlations with CI on several dates, but significant correlations with elevation are not observed in other regions. In Regions B and D, slope is negatively correlated with CI and PSRI is positively correlated with

CI on several dates, but the other regions have weak and inconsistent correlations with these variables. In Region C, cos(aspect) has the most frequent significant correlations with CI, but this tendency is not observed elsewhere. The vegetation indices have weak correlations with CI in all regions. The correlations between soil attributes and CI are generally strongest in Region D and weakest in Region A. The soil attributes that are most strongly correlated with CI vary by region. In Region D, USDA sand, USDA silt, USCS gravel, USCS fines, D_{60}, D_{30}, D_{10}, PI, and porosity all have significant correlations with CI on multiple dates. The results suggest that finer and more plastic soils typically have greater CI values.

Figure 1.12 examines the slopes of the relationships between both soil moisture and soil strength and each explanatory variable. More specifically, the figure shows the ratio of the slope calculated in each region to the slope calculated for the complete dataset for each statistically significant correlation that occurs in more than one region. The figure shows that the relationships between soil moisture and predictor variables also change among regions. For example, the slopes of the relationships between soil moisture and the vegetation indices (NDVI and EVI) are noticeably higher in Regions A and C than in Regions B and D. Region A and C generally have greater slopes for the relationships between soil moisture and soil composition parameters compared to Regions B and D.

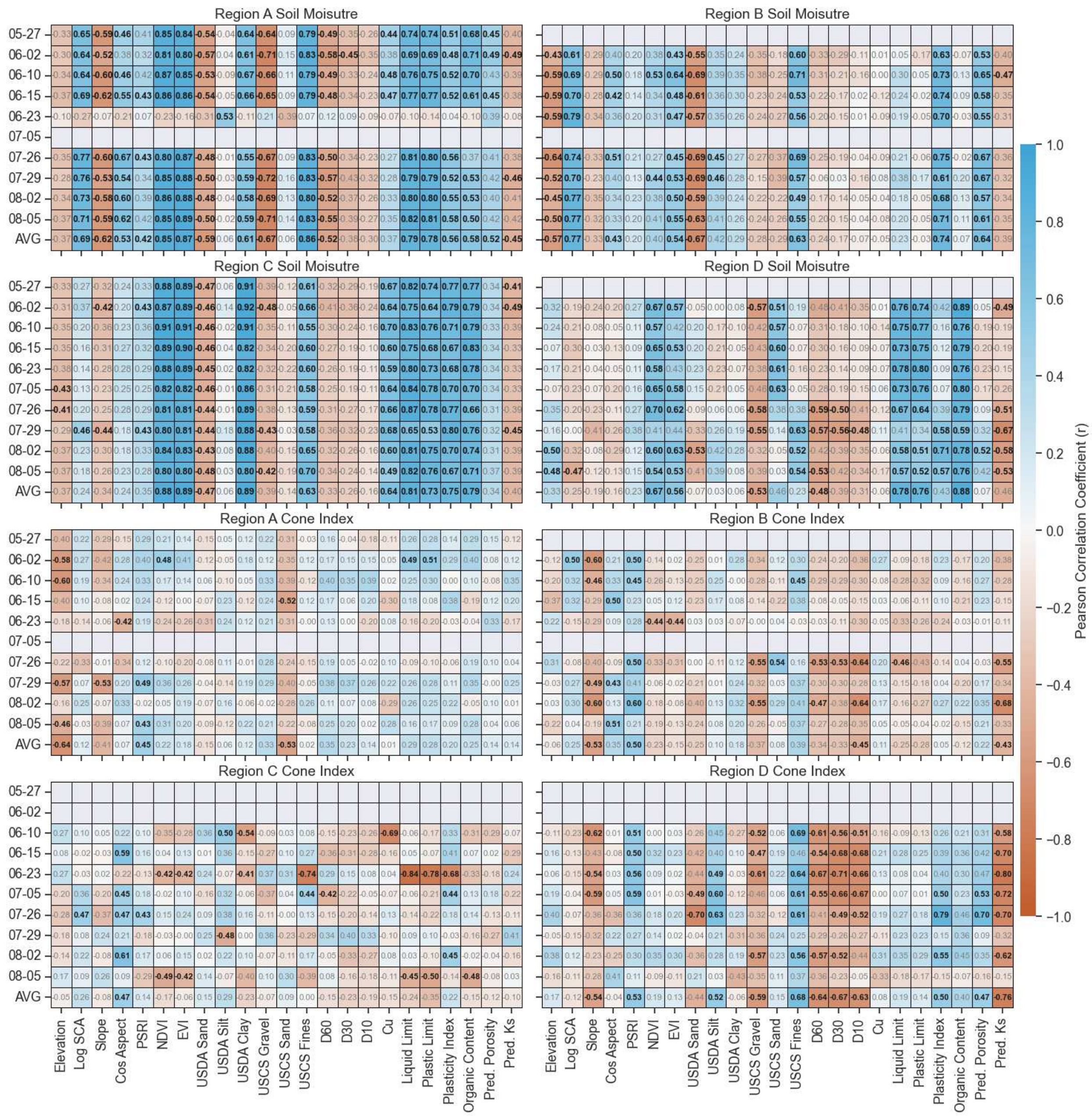

Figure 1.11. Pearson correlations of soil moisture and strength with topography, vegetation, and soil attributes for each region. Correlations displayed with bold black text represent statistically significant correlations based on the Pearson p-value ($p < 0.05$). Other correlations are displayed in grey text.

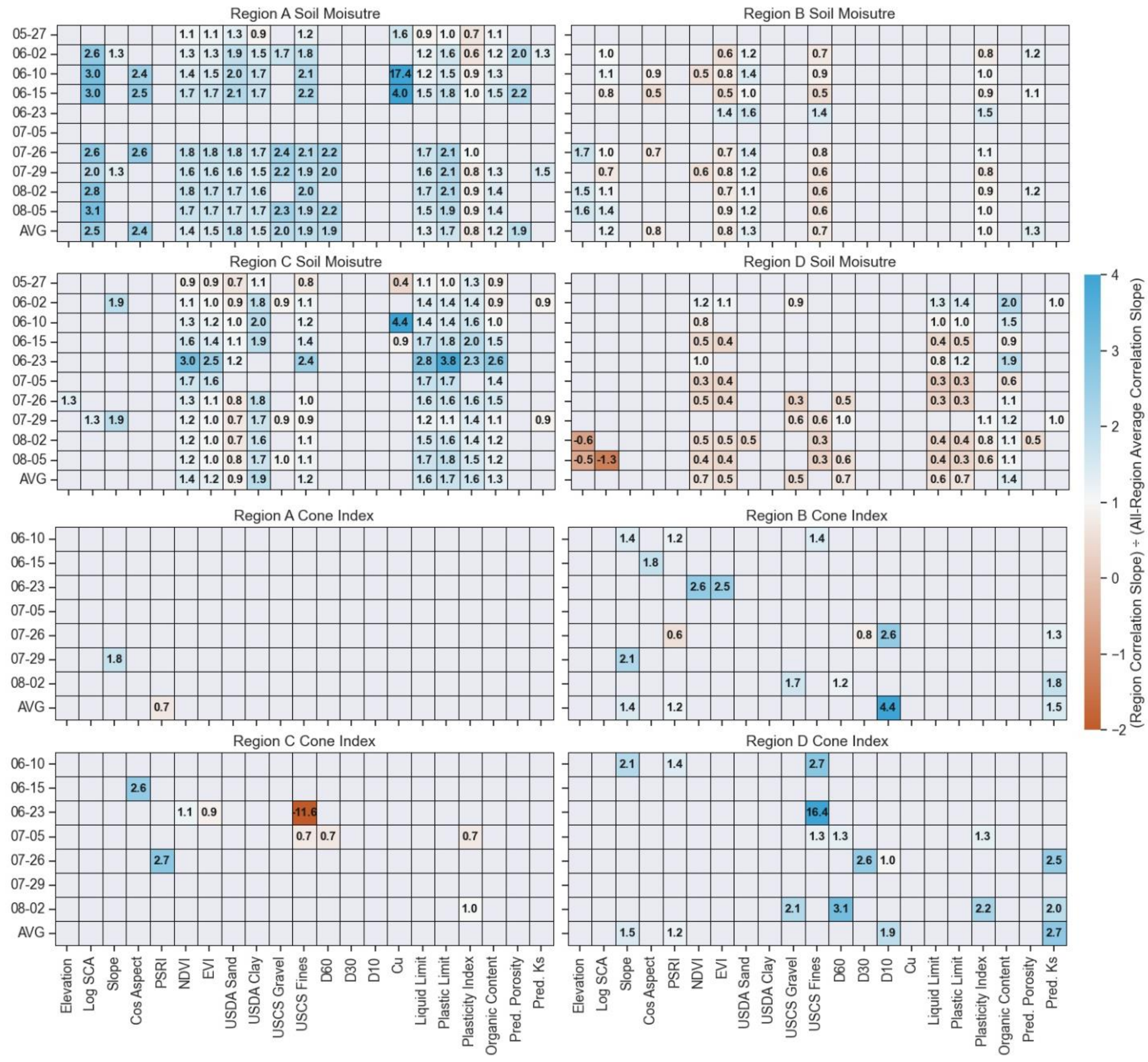

Figure 1.12. Ratio of the slope calculated in each region to the slope calculated for the complete dataset. Ratios are shown only for statistically significant correlations that appear in more than one region.

1.4. Discussion

The correlation analyses for both soil moisture and strength suggest the strengths of the correlations, and their respective slopes, vary by region. However, the signs of significant correlations are generally consistent between regions. When the dependence of soil moisture on topographic and soil attributes was studied at a smaller scale in a prior study, similar observations were made (Liang et al., 2017).

The attributes that are most strongly correlated with soil moisture vary by region. Soil moisture in Region A has strong correlations with topographic, vegetation, and soil attributes. Soil moisture in Regions B is somewhat correlated with topography and has weak correlations with vegetation and soil attributes. Soil moisture in Region C is most strongly associated with vegetation and parameters describing the quantity and behavior of fine-grained soils. In Region D, soil moisture is somewhat strongly correlated with vegetation and soil.

Soil moisture in Region A likely has strong correlations with topography due to the highly variable topography in the region. Similarly, the strong relationship between soil moisture and vegetation in Regions A and C is attributed to the diverse vegetation conditions in these regions. Regions A and C both have substantially greater standard deviations of NDVI and EVI than Regions B and D. Additionally, soil attributes in Regions A and C are strongly correlated with vegetation, which may reinforce soil moisture's correlation with vegetation in those regions. In particular, the soils in Regions A and C have relatively high fines content and have high mean LL, PL, and PI. Based on soil water retention and soil plasticity frameworks, soils with more fines and higher plasticity have higher affinity for water (Casagrande, 1948; Brooks and Corey, 1964; van Genuchten et al., 1991), which may result in denser vegetation (Moeslund et al., 2013). In Region D, the correlations of soil moisture with vegetation and soil attributes are again attributed to the relationships of soil fines content and plasticity with soil moisture and increased vegetation cover. Higher surficial moisture content in the presence of fines was also observed by Lin et al (2006).

Soil strength (viz. CI) generally has weaker correlations with topography, vegetation, and soil attributes than the correlations observed for soil moisture. The weak correlations between soil strength and topography may occur in part because topography has weak correlations with soil attributes, which are related to soil strength based on unsaturated soil strength frameworks (Fredlund et al., 1978; Lu et al., 2010). Correlations between soil strength and vegetation are mostly insignificant in all regions. Aside from a few sampling locations in grassy valley bottoms, root interference was not observed during sampling with the cone penetrometer. Soil strength is more strongly correlated with soil attributes than topography and vegetation. However, soil attributes only have consistent significant correlations with soil strength in Region D. Region D is characterized by low plasticity, coarse grained soils, and low moisture content, so the CI may be governed by non-cohesive materials with little to no strength contributions form suction stress (Fredlund et al., 1978; Lu et al., 2010). The low contribution of suction stress may also explain why Region D has the lowest average CI. The cone penetrometer has a maximum reading of 2070 kPa for the 2.54-cm cone. For many sampling locations in Regions A, B, and C, the maximum reading was reached before the cone penetrometer was fully inserted into the soil, resulting in incomplete distributions of CI, especially for dry sampling dates. Because Region D has the lowest average CI and the highest standard deviation, results from this region may better capture the variations of soil strength. While correlations of soil strength to topography, vegetation, and soil attributes are generally weak, the correlation strengths are observed to vary by region and the most influential attributes vary by region.

While the Pearson correlation framework was used to quantify relationships between variables, limitations of this method are acknowledged. Correlation analyses assume the relationships between variables are linear and may not capture nonlinear dependencies. Univariate correlations also do not capture cross-correlations explicitly, which can affect interpretation of the results. Even so, the present results can be used to identify the significant linear correlations. For strength measurements, the CI was reported up to 2070 kPa based on the cone penetrometer load cell calibration; consequently, soil strengths greater than this were not captured by the strength measurement method. The cone penetrometer operator sometimes varied by date and by region. While all cone penetrometer operators were trained, some inherent variability is expected. Additionally, field observations indicate the cone penetrometer can produce anomalously high readings if gravel and rocks are encountered during penetration.

1.5. Conclusions

The primary objective of this study was to examine spatial variability in the dependence of soil moisture and soil strength on topography, vegetation, and soil within a SMAP grid cell. Soil moisture and soil strength were measured on ten sampling dates in four geographically separated regions within a SMAP grid cell. The dependence of soil moisture and soil strength on topographic, vegetation, and soil characteristics was quantified using Pearson correlation coefficients, and the statistical significance of each correlation was assessed using the Pearson p-value. Assessed topographic attributes include elevation, log SCA, slope, cos(aspect), and PSRI. Vegetation indices include EVI and NDVI, and soil attributes include USDA percent sand, silt, and clay, USCS percent gravel, sand, and fines, D_{60}, D_{30}, D_{10}, C_u, LL, PL, PI, organic content, porosity, and K_s. Based on results of this study, the following conclusions are made:

1. The strength of the relationships between soil moisture and topographic, vegetation, and soil attributes can vary substantially within a SMAP grid cell. Soil moisture is strongly correlated to topography in Region A, weakly correlated to topography in Region B, but has few significant correlations with topography in Regions C and D. Soil moisture is strongly correlated to vegetation in Regions A and C, less correlated in Region D, and poorly correlated in Region B. Soil moisture is most correlated with soil attributes in Region A, followed by Regions C and D, with Region B having weak correlations.

2. The regression slopes of the relationships between soil moisture and topographic, vegetation, and soil attributes can also vary substantially within a SMAP grid cell. Region A has the greatest slopes for relationships between soil moisture and topography. Regions A and C have greater slopes for relationships of soil moisture with vegetation and soil than Regions B and D.
3. The strength of the relationships between soil strength and topographic, vegetation, and soil attributes can vary within a SMAP grid cell. Regions A, B, and D have weak correlations between soil strength and topography while Region C has few significant correlations. Region D is the only one with consistent significant correlations between soil strength and compositional attributes.

Although the four regions were identified based on physiographic attributes, additional studies are needed to identify and quantify the controlling factors. Future studies should consider what combinations of factors (e.g., geology, soil parent material, soil depth, rockiness, vegetation type, vegetation density, topographic variability) determine the number and extent of regions needed to characterize a given SMAP grid cell.

Overall, the variations in the relationships highlight the complex interactions between the soil moisture and strength and the environmental factors. The variations also indicate that using constant parameters for soil moisture downscaling across the extent of a coarse resolution grid cell is likely to lead to suboptimal performance. Furthermore, because soil strength is influenced by soil moisture, based on unsaturated strength modeling procedures (Fredlund et al., 1978; Lu et al., 2010; Pauly, 2019; Uusitalo et al., 2019), modeled soil strength predictions are expected to improve when soil moisture estimates from downscaling procedures consider the variability in the underlying relationships. Methods should be explored that cluster locations and infer separate parameters for those clusters. Machine learning methods that identify such clustering as part of their methodology are expected to outperform methods that implicitly assume consistent relationships across a coarse resolution grid cell.

2. FIELD HYPERSPECTRAL SENSING AND CONVOLUTIONAL NEURAL NETWORKS FOR PREDICTING SURFICIAL SOIL PROPERTIES

2.1. Introduction

Predictions of vehicle mobility on off-road terrain are important for military, agricultural, and recreational vehicle routing (Wulfohn et al., 1996; Priddy et al., 2012; Uusitalo et al., 2018; Schönauer et al., 2022). The strength of soils under vehicle loads are informed by soil properties including composition, plasticity, and moisture (Casagrande, 1948; Brooks and Corey, 1964; van Genuchten et al., 1991; Choi et al., 2019; Pauly, 2019; Bindner, 2020; Bullock, 2023). Soil characteristics including soil composition (percent sand, silt, and clay), organic content, and plasticity (liquid limit and plastic limit) are typically measured using laboratory procedures or estimated using digital soil maps (DSMs), such as the Soil Survey Geographic Database (SSURGO) (Soil Survey Staff, 2023) and POLARIS (Chaney et al., 2016). However, laboratory measurement of soil attributes can be time intensive, and the fidelity of DSM data are often unknown. Soil moisture can be measured using in situ moisture sensors or estimated from soil moisture downscaling procedures (Merlin et al., 2006; Coleman and Niemann, 2013; Peng et al., 2016; Sadeghi et al., 2017; Fang et al., 2018). However, field measurement of soil moisture using sensors may not be practical across large spatial extents and soil moisture downscaling procedures require field measurements to assess model performance. Methods that rapidly collect data related to soil properties may better support the characterization of soils across large spatial extents.

Hyperspectral sensors have been mounted on vehicles for collection of data to characterize terrain (Winkens et al., 2017) and could potentially support the high-speed characterization of soil properties. Hyperspectral sensing in the visible and near-infrared (300-

1000 nm) (VNIR) and short-wave-infrared (1000-2500 nm) (SWIR) portions of the electromagnetic spectrum has been shown to vary with soil properties (Clark et al., 1993; Jiang, 2016; Sadeghi et al., 2018), and hyperspectral data in the VNIR+SWIR range and machine learning have shown strong promise for the prediction of soil properties including soil composition, soil organic content, plasticity, and moisture content (Waiser et al., 2007; Waruru et al., 2014; Gupta et al., 2016; Tsakiridis et al., 2020; Knadel et al., 2021; Purushothaman et al., 2024). However, these studies conducted hyperspectral sensing on processed soils under engineered lighting conditions, which requires disruptive soil sampling and time intensive soil processing prior to hyperspectral sensing. Additionally, these studies used SWIR spectral detectors, which are expensive compared to VNIR detectors and may not be accessible for many practical applications. The use of VNIR detectors alone would greatly reduce the cost of hyperspectral instrumentation, with VNIR spectrometers generally costing hundreds to thousands of U.S. dollars and VNIR+SWIR spectrometers typically costing upward of 75,000 U.S. dollars. Additionally, the capture of hyperspectral data in situ would greatly improve the speed of hyperspectral data collection. However, there is currently a poor understanding of how VNIR hyperspectral sensing on undisturbed surficial soils under natural lighting conditions can be used to predict soil properties.

The objective of this study is to explore the use of field captured VNIR hyperspectral data for the prediction of soil properties. Sampling locations (n=48) were identified at a field site in northern Colorado and soil properties were measured using standard laboratory and field procedures. Hyperspectral data were captured at sampling locations on three field dates and used to compile a soil-property/hyperspectral dataset. The dataset was split into training and testing subsets using conditioned Latin Hypercube Sampling (cLHS) based on the hyperspectral data.

The training data were used to cross validate a CNN and the testing data were used to evaluate the performance of the trained model. Results from the trained CNN are compared to SSURGO reported values at sampling locations. Finally, the predictive accuracy of the CNN applied to field data is discussed and recommendations for future studies are provided.

2.2. Materials and Methods

2.2.1. Study Site

Maxwell Ranch (study site) is situated 53 km northwest of Fort Collins, Colorado, and encompasses 4000-ha. The study site has an average elevation of approximately 2,100 m and is located in the Larmie Foothills, which marks the transition between the Rocky Mountains and the Great Plains. The site exhibits a topographic relief of approximately 300 m, with an average slope of 0.22 m/m. Soils at the study site are predominately coarse grained and primarily derived from weathered granite (Soil Survey Staff, 2023).

Maxwell Ranch contains diverse vegetation, featuring scattered shrubs, short-grass prairies, and Ponderosa Pine trees. The study site has been subdivided into three regions for this study, labeled A, B, and C, to capture variability in soil and shading conditions (Figure 2.1a). Hyperspectral data were captured at 15, 16, and 17 sampling locations in Regions A, B, and C respectively. Sampling locations were selected as areas with exposed soil and minimal vegetation to reduce the influence of vegetation on hyperspectral data. Region A has high topographic relief, forested slopes, and a grassy valley bottom (Figure 2.1b). Region B is characterized by open canopy Ponderosa pine trees and relatively low topographic relief (Figure 2.1c). Region C features rolling grassland hills and prairies, with a grassy valley bottom (Figure 2.1d).

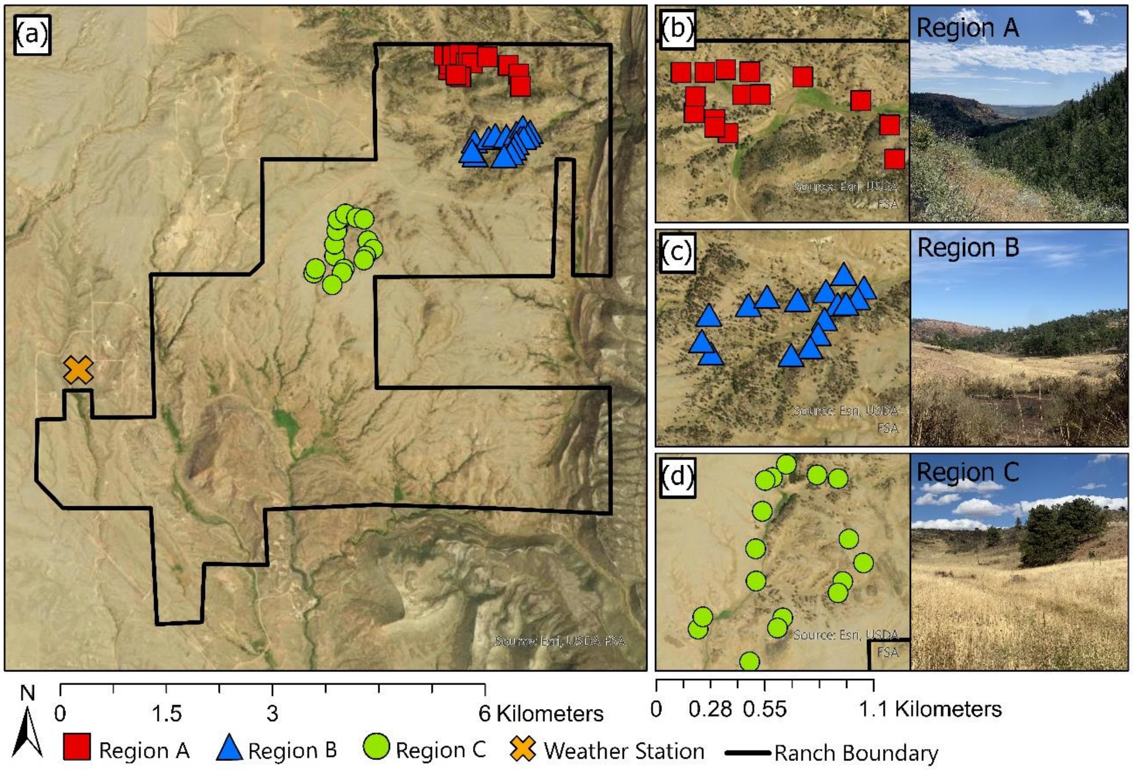

Figure 2.1. (a) Map of Maxwell Ranch including sampling locations, weather station, and ranch boundary, and (b-d) Regions A, B, and C sampling locations and photos captured at each region.

2.2.2. Soil Moisture Measurement

Soil moisture was determined at the sampling locations using a Stevens HydraProbe Soil Sensor (Stevens Water Monitoring Systems Inc., 2018), which measures the volumetric soil moisture (cm^3/cm^3) within the top 5 cm of soil. The HydraProbe sensor has a precision of ±0.001 (cm^3/cm^3) and an accuracy of ±0.03 (cm^3/cm^3) (Stevens Water Monitoring Systems Inc., 2018). Soil moisture measurements were collected within 1 m of each sampling location.

Figure 2.2 illustrates the precipitation during the sampling period (August 2023 – October 2023) and soil moisture at sampling locations on sampling dates. Sampling dates were selected to capture a variety of soil moisture conditions. Daily average temperatures ranged from 17°C to 21°C with no precipitation on sampling dates. Data collection at Region A was completed on August 3rd, which had the highest spatial average soil moisture among the sampling dates (0.19 cm^3/cm^3) due to precipitation events in the days prior to field sampling. Data were collected at Regions C and B on October 10th and 17th respectively, with spatial average soil moisture near 0.01 cm^3/cm^3 for both regions on these dates.

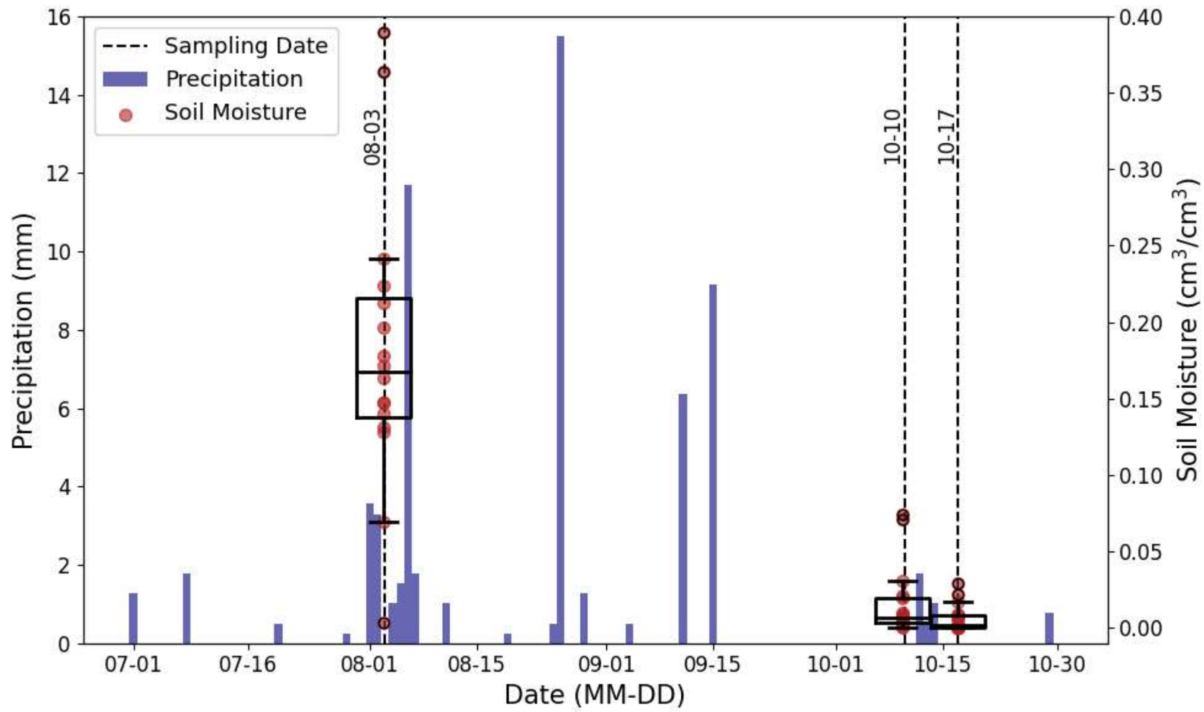

Figure 2.2. Daily precipitation during the 2023 sampling period with dashed vertical lines representing sampling dates and boxplots showing the soil moisture at sampling locations for each sampling date. Climate data were recorded by a nearby weather station (KCOLIVER 12).

2.2.3. Hyperspectral Sensing

Hyperspectral data in the VNIR portion of the electromagnetic spectrum were captured at each sampling location using a spectrometer. The spectrometer contains a charge coupled device linear image sensor for reflectance at wavelengths ranging from approximately 300-1000 nm with a resolution of approximately 0.2 nm (TOSHIBA Electronic Devices & Storage Corporation, 2018). Hyperspectral data were captured under natural lighting conditions which included sunny, partly cloudy, and cloudy conditions at shaded and unshaded sampling locations.

The spectrometer was attached to a mounting assembly, which maintained a sensing distance of 1 m above the soil surface (Figure 2.3a). Figure 2.3b displays the approximate extent of data captured by the spectrometer on the soil surface (~100 cm^2). A light insulating cap was placed over the spectrometer viewing window and dark hyperspectral data were captured at each sampling location to characterize the spectrometer's ambient detector noise (viz. dark scan). The light insulating cap was then removed and a square (15 cm × 15 cm) 2.5 cm thick sheet of polytetrafluoroethylene (PTFE) with an approximate VNIR reflectance of 99% was placed on the soil surface. Hyperspectral data were then collected for the PTFE sheet to obtain reference hyperspectral data (viz. PTFE scan), similar to the procedure suggested by Shaikh et al. (2001). The PTFE sheet was then removed, and hyperspectral data were collected for the soil surface (viz. soil scan). Spectrometer integration times describe the period over which the spectral detectors measure radiation before recording data and were adjusted based on lighting conditions at each site, to obtain data at 40-60% of sensor saturation.

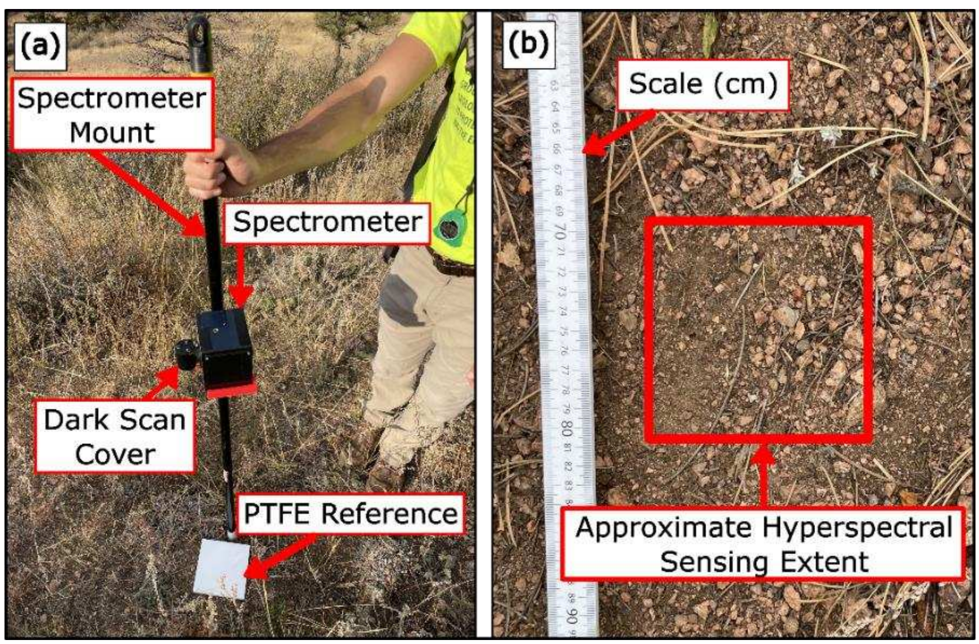

Figure 2.3. (a) Spectrometer assembly with PTFE reference, spectrometer mount maintaining 1 m elevation above ground surface, dark scan cover, and spectrometer and (b) extent of hyperspectral scan (approximately 100 cm^2).

2.2.4. Soil Characterization

Soil samples were collected within approximately 1 m of the hyperspectral sensing location and within the uppermost 5 cm of soil considering the surficial soil layer (0 – 15 cm) generally controls off-road vehicle trafficability (Wong, 2010). Surface debris and vegetation were removed prior to soil sampling. Laboratory analyses of soil samples included sedimentation analysis, plasticity testing, and measurement of organic content. United States Department of Agriculture (USDA) percent sand, percent silt, and percent clay were determined by sedimentation analysis using the *PARIO Soil Particle Analyzer,* PARIO Classic procedure (METER Group, 2017). The soil liquid limit (LL) and plastic limit (PL) were determined according to ASTM D4318 (ASTM International, 2017b) and organic content was measured according to ASTM D2974 (ASTM International, 2020).

The USDA classifications of sampled soils (USDA, 1975) are displayed in Figure 2.4a. Soils at the sampling locations are dominated by sands with most soils classifying as sandy loam. The results from soil plasticity testing are shown in Figure 2.4b. The fine-grained fraction of soils (less than No. 200 sieve) at sampling locations are dominated by low to high plasticity silts.

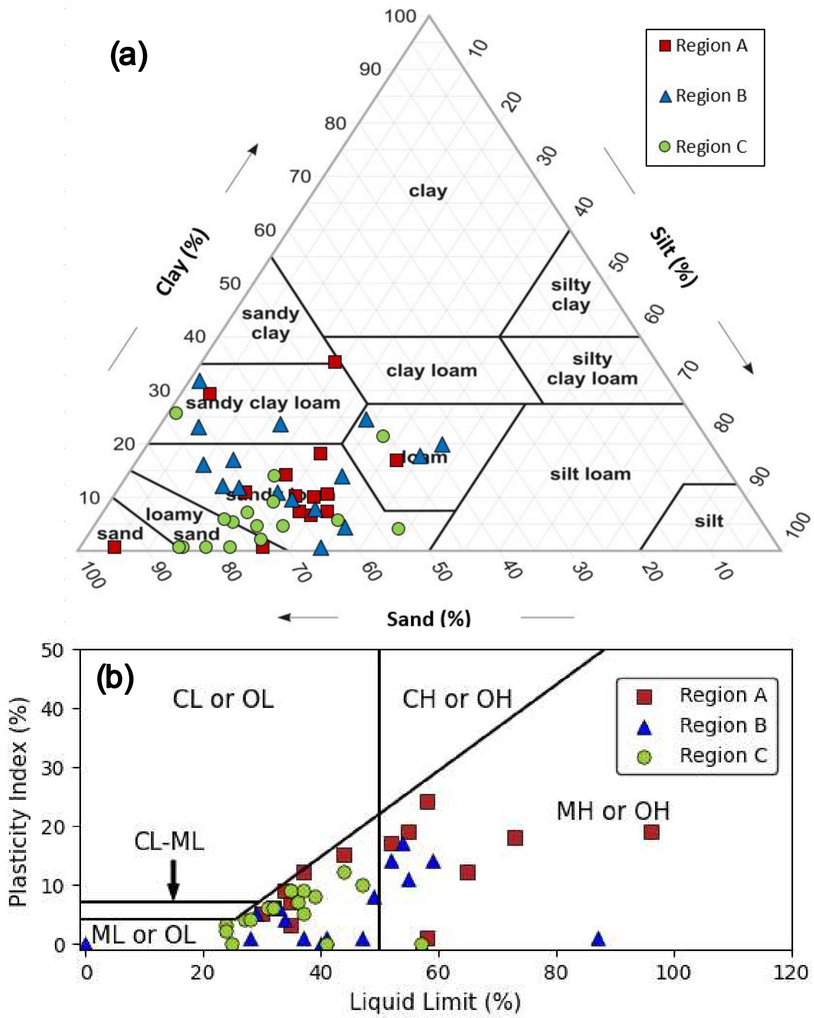

Figure 2.4. (a) USDA classification of soils at sampling locations and (b) the plasticity of soils at sampling locations. Point symbology varies by sampling region.

2.2.5. Hyperspectral Data Processing

Dark scans, PTFE scans, soil scans, and integration times were used to normalize hyperspectral data for each sampling location. All hyperspectral data were first divided by the respective integration time, then the relative intensity (RI) was calculated as:

$$RI = \frac{(Soil\ Scan - Dark\ Scan)}{PTFE\ scan} \times (reference\ reflectance) \qquad (3.1)$$

where the soil scan, dark scan, and PTFE scan are the integration time adjusted hyperspectral data for each 0.2 nm wavelength at each sampling location, and the reference reflectance is the approximate reflectance of the PTFE sheet. Relative intensity provides a metric to compare hyperspectral data in a single dataset but does not represent true spectral reflectance. The lighting conditions were uncontrolled and unmeasured at the sampling locations, therefore the true spectral reflectance of soils was not determined as part of this study and the relative intensities were used for machine learning. Two hyperspectral signals were removed from the dataset during hyperspectral processing due to signal anomalies.

2.2.6. Machine Learning

The dataset was partitioned into training and testing data using the conditioned Latin Hypercube Sampling algorithm (cLHS) (Minasny & McBratney, 2006) based on the relative intensities to replicate the multivariate distribution of the hyperspectral data in both the training and testing datasets. Python 3.10 was used for data processing and machine learning. The Keras 2.11.0 (Chollet, 2015), running on TenserFlow 2.11.0 (Abadi et al., 2018), was used to create, calibrate, and test the CNN.

Figure 2.5 displays the distribution of properties for soil samples present in training and testing datasets. A total of 16 sampling locations were extracted from the hyperspectral dataset to be used as testing data and the remaining sampling locations were used as training data (approximately 35% testing data and 65% training data). Soil properties belonging to samples in the testing dataset are generally within the upper and lower bounds for each property in the training dataset and both the training and testing datasets have a similar distribution of properties.

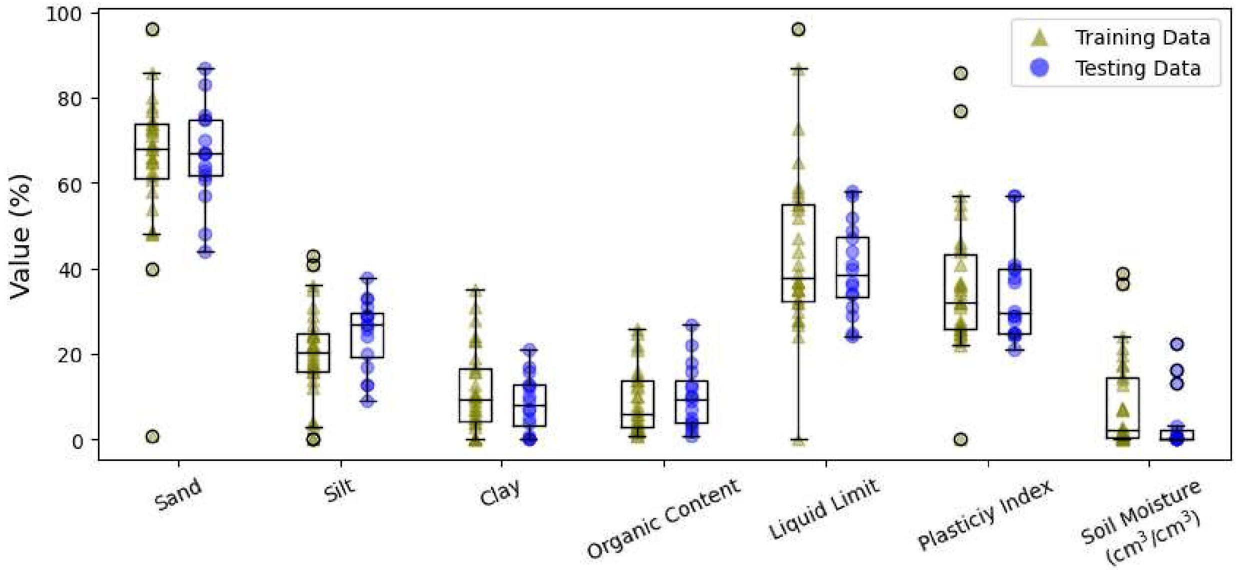

Figure 2.5. Boxplot displaying the values of soil properties for each sample used in CNN training and testing. The testing data is generally within the upper and lower bound of training values for each soil property and the training and testing data generally have similar sample property distributions.

While several machine learning procedures are commonly used in soil spectroscopy (Sjöström et al., 1983; Quinlan, 1992; Quinlan, 1993; Drucker et al., 1996; Ramirez-Lopez et al., 2013; Veres et al., 2015), CNNs were selected for machine learning based on the algorithms high predictive performance compared to other modeling procedures, as studied by Tsakiridis et al. (2020). Cross validation of CNNs helps to generalize the model and involves iteratively adjusting user-defined variables (viz. hyperparameters) and assessing model performance using subsets of the training data. To prepare data for cross validation, the training data were grouped into three spectrally similar clusters using the Fuzzy c-means algorithm (Bezdek et al., 1984), then samples from each cluster were randomly distributed into three groups. This was done to ensure cross validation groups contained spectrally comparable data. Cross validation was used to inform the final model architecture which consists of convolutional layers, batch normalization layers, max pooling layer, and dense (fully connected) layers.

Figure 2.6 illustrates the model architecture, activation functions, number of filters, and kernel size for each layer. Input data were zero-padded to preserve spatial dimensionality during convolution and L2 kernel regularizers (λ_2=1E-2) were used to prevent overfitting and improve model generalization (Säfken et al., 2020). Activation functions varied by layer and included Exponential Linear Unit (ELU), Rectified Linear Unit (RELU), Scaled Exponential Linear Unit (SELU), and Hyperbolic Tangent (Tanh) functions (Chollet, 2015). The bias term was added to several layers to increase model flexibility and better fit data with a non-zero mean. The model contained eight convolutional layers with varying kernel sizes followed by a flattening layer which converts the data into a one-dimensional vector. Flattened data were then passed through six fully connected layers for the identification of high-level relationships. The final dense layer then outputs predicted soil properties for each sample in the testing dataset.

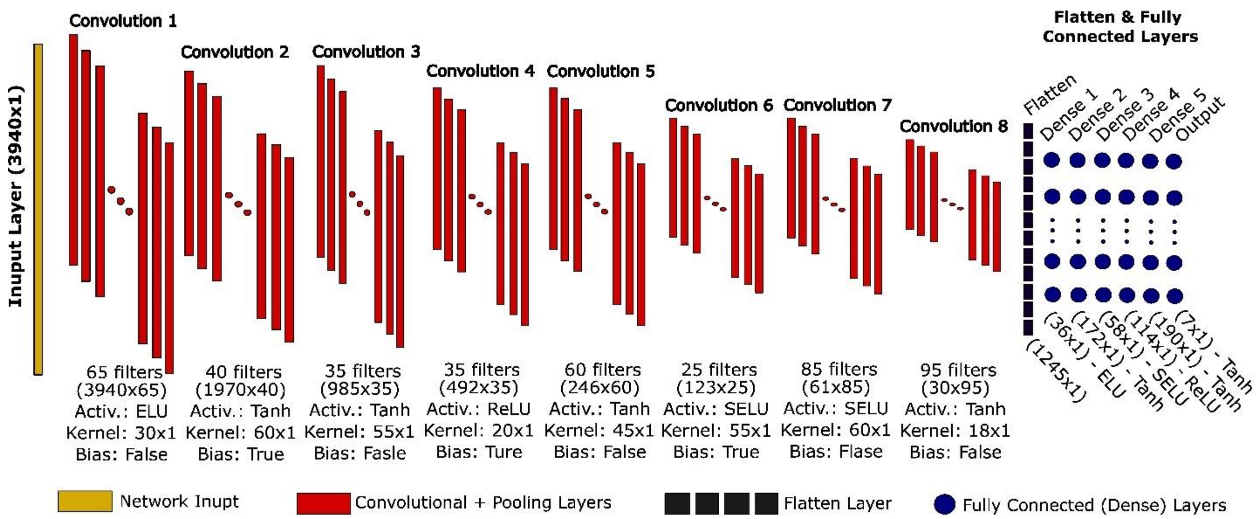

Figure 2.6. Graphical summary of the CNN model architecture including the input layer, convolutional layers, pooling layers, flatten layer, and fully connected layers. Below each convolutional layer is a list containing the number of filters per layer, the shape of the layer, the activation function of each layer, the kernel size for the layer, and the bias term. Each convolutional layer was accompanied by a pooling layer where the pooling size was two. Convolutional layers were followed by a flatten layer and six fully connected layers. Below each fully connected layer is the layer shape followed by the activation function.

2.3. Results

The accuracy of soil property predictions from the CNN using testing hyperspectral data and SSURGO are shown in Figure 2.7. The CNN predictions of percent sand, silt, and clay have root mean squared errors (RMSEs) between 7.5-10.8% and have lower errors compared to SSURGO estimates (RMSEs of 9.4-11.0%) (Figure 2.7a-c). The trendlines from CNN predictions of sand, silt, and clay have slopes closer to the 1:1 line and greater COD values (COD of 0.04, 0.12, and -0.8 respectively) compared to the SSURGO results (COD of 0.00, -0.68, and -1.17 respectively), demonstrating the CNN predictions better capture the trend of particle size distribution parameters. Predictions of organic content from the CNN have an a lower RMSE compared to SSURGO (6.3% compared to 10.5%, respectively) and better captures the trend of organic content compared to SSURGO (COD of 0.24 compared to -1.07, respectively) (Figure 2.7d). The CNN predictions of liquid limit and plastic limit have RMSEs of 9.8% and 10.7% respectively, compared to RMSEs of 16.5% and 17.5% for SSURGO respectively (Figure 2.7e-f). The liquid limit and plastic limit trends captured by the CNN (COD of 0.09 and -0.02, respectively) also performs better than SSURGO (COD of -1.60 and -1.76, respectively). Moisture content data are not available through SSURGO, however, the estimations of soil moisture using the CNN captures the variability of soil moisture and have an RMSE of 0.04 cm^3/cm^3 (COD of 0.66) (Figure 2.7g).

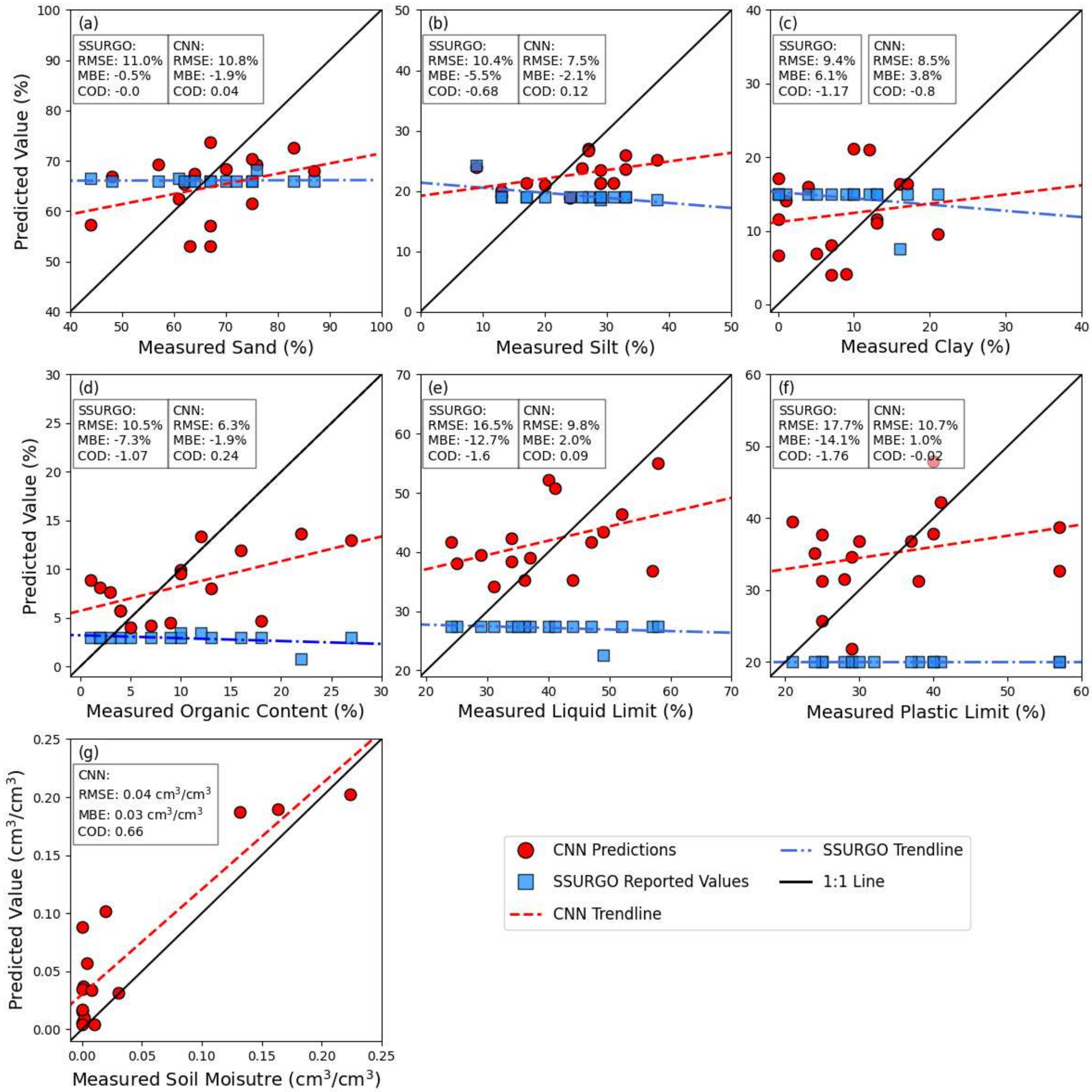

Figure 2.7. CNN and SSURGO soil property predictions for samples in the testing dataset plotted against laboratory measured values. Results include (a) percent sand, (b) percent silt, (c) percent clay, (d) percent organic content, (e) liquid limit, (f) plastic limit, and (g) soil moisture content. SSURGO does not include predictions of soil moisture. Root mean squared error (RMSE), mean bias error (MBE), and coefficients of determination (COD) are shown for all predictions.

2.4. Discussion

The use of CNNs and field VNIR hyperspectral data show promise for the prediction of soil properties. The CNN predicted soil characteristics have lower errors than SSURGO and

CNN predictions better capture the trend in soil property variability. While the CNN predictions are compared to SSURGO data, the SSURGO data was collected at scales ranging from 1:12,000 to 1:63,360 (Soil Survey Staff, 2023) while hyperspectral data were collected within 1 m of the point of interest and over a 100 cm^2 area.

Tsakiridis et al. (2020) used an extensive laboratory VNIR+SWIR hyperspectral dataset and CNNs to predict sand, silt, clay, and organic content for soils across the European Union and achieved predictive RMSEs of 11.95%, 9.33%, 4.80%, and 10.96% respectively, which are comparable to RMSEs achieved using the field VNIR hyperspectral data in this study (10.8%, 7.5%, 8.5%, and 6.3% respectively). However, Tsakiridis et al. (2020) achieved better COD values for predictions and their modeling procedure was used to estimate a greater range of compositions. Other studies which have predicted liquid limit and plastic limit using hyperspectral data and machine learning have achieved better RMSE (LL: 7-10%; PL: 3-5%) and COD values (LL:0.7-0.8; PL:0.45-0.69) compared to CNN RMSE (LL: 9.8%; PL:10.7%) and COD values (LL:0.09; PL:-0.02) from this study (Waruru, 2014; Gupta et al., 2019; Knadel et al., 2021). However, these studies captured hyperspectral data in controlled laboratory settings, had larger training datasets, and included SWIR data, which may contribute to improved performance. Zhu et al. (2010) used laboratory VNIR+SWIR hyperspectral sensing to predict surface soil moisture and achieved an RMSE of 0.065 cm^3/cm^3 and COD of 0.88, which are comparable to results from this study (RMSE of 0.04 cm^3/cm^3 and COD of 0.66).

Several factors introduce more complexity into field captured hyperspectral data including lighting conditions, surface vegetation, and surficial vegetation litter, despite attempts to minimize vegetation interference. Lighting conditions varied with weather, shading conditions from vegetation, and angle of solar radiation. Additionally, the intensity of radiation at several key wavelengths is reduced at the earth's surface due to radiation absorption by moisture in the atmosphere (Elachi and Van Zyl, 2006; Nolet et al., 2014). Despite variable lighting conditions, the results suggest the field captured VNIR hyperspectral data and CNNs show promise for the prediction of soil properties. Finally, this study did not include SWIR reflectance measurements, which is a substantial simplification relative to previous studies. Given the high cost of SWIR sensors compared to VNIR sensors, use of low cost VNIR may foster greater adoption.

While CNNs have been shown to perform well for the prediction of soil properties using hyperspectral data (Tsakiridis et al., 2020), the predictive performance of other models commonly used in lab-based soil spectroscopy were not explored in this study nor were data preprocessing procedures that are commonly used in laboratory soil spectroscopic studies (Tsakiridis et al., 2020). Additionally, other studies that used hyperspectral data to predict soil properties used larger datasets for the training and evaluation of predictive models (Waiser et al., 2007; Waruru et al., 2014; Gupta et al., 2016; Tsakiridis et al., 2020; Knadel et al., 2021; Purushothaman et al., 2024). While larger datasets are typically associated with improved model performance, studies have shown that for some machine learning problems, there is a threshold where additional training data does not improve model performance (Prusa et al., 2015; Zhu et al., 2016). Future studies should evaluate machine learning predictive performance for soil properties using larger field hyperspectral datasets. This study did not use SWIR hyperspectral data, which has been shown to capture differences in soil properties (Clark et al., 1993; Jiang, 2016; Sadeghi et al., 2018). The inclusion of SWIR data in machine learning procedures is expected to improve the prediction of soil properties. Finally, the relative intensity of hyperspectral data was used in machine learning and in-depth spectral analysis was not completed as part of this study. Obtaining true reflectance data and completing spectral analysis may support an understanding of how field VNIR spectral data vary from laboratory VNIR data.

2.5. Conclusions

The objective of this study was to use field VNIR hyperspectral data and a CNN to predict soil properties including percent sand, silt, clay, organic content, liquid limit, plastic limit, and soil moisture. A spectrometer was used to capture VNIR hyperspectral data under natural lighting conditions at 48 sampling locations in northern Colorado. Soil moisture from 0-5 cm was characterized in the field using a soil moisture sensor and soil characteristics were measured using standard laboratory procedures. Hyperspectral data were normalized using the hyperspectral detector integration time, soil scan, dark scan, and PTFE scan data. The conditioned Latin Hypercube Sampling algorithm was used to partition the data into training and testing datasets. The testing data were used for model cross validation and the cross validated model was trained and evaluated using the testing dataset. Results were compared to SSURGO predictions at sampling locations and the following conclusions are made:

1. Hyperspectral data captured under field conditions and CNNs show promise for the prediction of soil properties including percent sand, silt, clay, organic content, liquid limit, plastic limit, and soil moisture. For the prediction of percent sand, silt, clay, organic content, liquid limit, and plastic limit, the CNN had lower RMSE values (10.8%, 7.5%, 8.5%, 6.3%, 9.8%, and 10.7% respectively) compared to SSURGO (11.0%, 10.4%, 9.4%, 10.5%, 16.5%, and 17.7% respectively). Predictions of soil moisture are not provided by SSURGO. However, the CNN resulted in an RMSE of 0.04 cm^3/cm^3.
2. Soil properties predicted using field hyperspectral data and the CNN better captured the trend of soil properties compared to SSURGO at the study site. For predictions of percent sand, silt, clay, organic content, liquid limit, and plastic limit, the CNN resulted in COD

values of 0.04, 0.12, -0.80, 0.24, 0.09, and -0.02, respectively, compared to SSURGO COD values of 0.00, -0.68, -1.17, -1.07, -1.60, and -1.76, respectively.

Future studies should consider collecting a large soil-hyperspectral dataset for the prediction of soil properties and evaluate various modeling procedures. Additionally, future studies should attempt to quantify how laboratory hyperspectral data vary from field hyperspectral data.

3. USE OF HYPERSPECTRAL SENSING AND MACHINE LEARNING FOR THE EX SITU PREDICTION OF TAILINGS PROPERTIES

3.1. Introduction

Mineral commodities are essential to modern societies and are an important part of the United Nations Sustainable Development Goals (UNDP and UN Environment, 2018). Tailings are the byproduct of mineral commodity extraction and consist of a slurry mixture of ground rock, water, and chemical reagents (Oberle et al., 2020). Tailings are commonly stored in surface impoundments called Tailings Storage Facilities (TSF) which require large spatial extents (Vick 1990; Werner et al. 2020), with the largest facilities designed to store more than 1 billion m^3 of tailings (Oberle et al., 2020). In 2016, an estimated 8.8 billion tonnes of tailings were produced (Oberle et al., 2020), and the volume of waste material per unit of commodity is generally increasing due to decreasing ore grades (Mudd 2007; 2010). TSFs can pose a level of risk to downstream stakeholders and environments and recent failures have highlighted the complexities of TSF design (Morgenstern et al., 2016; Robertson et al., 2019; Williams, 2021). In response to these failures, the Global Tailings Review (GTR) released the Global Industry Standard on Tailings Management (GISTM) to provide guidance on the design and operation of TSFs that outline the need for ongoing characterization of tailings properties (GTR, 2020; ICMM, 2021). However, the scale of TSFs can make tailings characterization difficult at the resolutions needed for identification of features impactful to physical and chemical stability.

Tailings properties help inform geotechnical and geochemical analyses and include the particle size distribution (PSD) of tailings, moisture content, void ratio, and density. The PSD of tailings influences shear strength, hydraulic conductivity, volume change characteristics, dry density, and beach angle, and can inform geochemical and rheological characterization of the material (Morrison, 2022). Tailings moisture content influences hydraulic conductivity, shear strength, compaction behavior, and acid generation potential (Morrison, 2022). Geotechnical stability is improved at saturation levels below approximately 80% (Rodriguez et al., 2021) and oxygen diffusion decreases exponentially at saturation levels above approximately 85%, generally corresponding with a reduction of acid generation in sulfide tailings (Aachib et al. 2004). Parameters that describe void space, such as void ratio and porosity, can be used in consolidation modeling, can help inform tailings liquefaction potential, and provides information on potential water recovery (Morrison, 2022). Tailings density is important for estimating hydraulic conductivity, potential water recovery, and TSF storage capacity over time (Morrison, 2022). Characterization of these properties can include in situ and laboratory testing.

Sampling methods for laboratory characterization are classified as either disturbed or undisturbed. Disturbed samples are retrieved in bulk samples and do not preserve the in situ tailings condition and can be used for characterizing soil properties that are not a function of volume or fabric, such as particle size distribution and gravimetric moisture content. Undisturbed samples can be used to characterize tailings saturation, void ratio, porosity, and density and represent a point measurement. However, obtaining enough samples for adequate resolution of tailings properties throughout a TSF is usually not feasible. Additionally, undisturbed samples are generally only obtained for fine grained soils unless specialty methods such as ground freezing are used (Wride et al., 2000a; 2000b) and obtaining high quality samples may not always be feasible. Undisturbed samples are commonly retrieved using thin-walled tube sampling methods (ASTM International, 2015) and while some level of disturbance is inherent, undisturbed samples can be high quality providing proper sampling procedures (Di Buò et al., 2018).

The standard penetration test (SPT) and cone penetration test (CPT) are common in situ methods for characterizing tailings behavior. The SPT requires drilling to depth before the resistance of soil to penetration is measured (ASTM International, 2022). The resistance measured by the SPT is reported as the N-value (blows/foot) which can be used to make correlations with soil behavior. However, the process of drilling can disturb the tailings, creating reduced repeatability (Morrison, 2022). The CPT collects continuous cone tip resistance, sleeve friction, and pore water pressure in the tailings profile and can include geophysical tests such as seismic cone penetration testing (SCPT) (Tschuschke et al., 2020). Data collected from the CPT are empirically related to soil behavior type (SBT) (Robertson, 2016) which generally correlates with Unified Soil Classification System (USCS) classes (Molle, 2005). Interpretations of CPT data may not be appropriate for all types of materials and interpretations cannot be made for tailings or soils with microstructure (Robertson, 2016). Additionally, the CPT does not directly characterize all geotechnically relevant tailings properties including percent sand, silt, clay, fines content, solids content, moisture content, saturation, void ratio, and porosity.

An alternate method for material characterization, which has gained attention in recent literature, involves hyperspectral sensing for the prediction of soil and tailings properties. Hyperspectral data can be captured on disturbed or undisturbed samples in either laboratory or field settings (Hedley et al., 2015; Cho et al., 2017; Pei et al., 2018; Riese & Keller, 2019, Tsakiridis et al., 2020; Entezari et al., 2022) and contains information on the reflectance of a material at various wavelengths in the visible, near-infrared, and short-wave-infrared portions of the electromagnetic spectrum. Machine learning models have shown promise for the prediction of soil texture using laboratory hyperspectral data (Riese & Keller, 2019, Tsakiridis et al., 2020). Other studies have shown promise for the use of push probe-based hyperspectral sensing for the estimation of soil texture and in situ moisture content for near surface agricultural soils (Cho et al., 2017; Pei et al., 2018). Additionally, hyperspectral sensing has been used for the prediction of oil sands tailings solids content, fines content, and water content on samples tested ex situ (Entezari et al., 2022).

Several modeling procedures represent the state-of-art for predicting soil properties from hyperspectral data including the Partial Least Squares (PLS) regression algorithm (Sjöström et al., 1983), the Cubist algorithm (Quinlan, 1992; Quinlan, 1993), the Support Vetor Regression (SVR) algorithm (Drucker et al., 1996), the Spectrum-Based Learner (SBL) (Ramirez-Lopez et al., 2013), and 1-D Convolutional Neural Networks (CNN) (Veres et al., 2015). Tsakiridis et al. (2020) compared the performance of these modeling procedures for the prediction of soil properties from hyperspectral data and demonstrated that 1-D CNNs outperform other machine learning procedures. While hyperspectral sensing and machine learning have shown promise for the prediction of PSD and moisture content, less is understood on how hyperspectral data may be used to predict other geotechnically relevant tailings properties including saturation, void ratio, porosity, and density.

The objective of this study is to collect hyperspectral data for tailings with a range of properties and assess how hyperspectral data and CNNs can be used to predict geotechnically relevant tailings properties. Tailings from a gold, silver, lead, and zinc producing mine were processed to obtain 300 test specimens with unique PSDs, moisture contents, and densities. Hyperspectral data were collected on all test specimens in the laboratory using a handheld spectrometer. Specimen properties and hyperspectral data were separated into training and testing datasets for machine learning and a 1-D CNN was cross validated, trained, and tested to assess the use of hyperspectral imagery for the prediction of tailings properties including percent sand, percent silt, percent clay, fines content, solids content, gravimetric and volumetric moisture content, degree of saturation, void ratio, porosity, total density, and dry density.

3.2. Methods and Materials

Figure 3.1 outlines the sample preparation, hyperspectral sensing, and machine learning procedures. A whole tailings sample was processed to artificially prepare 300 test specimens with a range of PDSs, moisture conditions, and densities and a spectrometer was used to collect hyperspectral data for each test specimen. The hyperspectral data was then processed and spilt into training and testing data to cross validate, train, and test a CNN for the prediction of tailings properties. Each step outlined in Figure 3.1 is described in detail in the sections below.

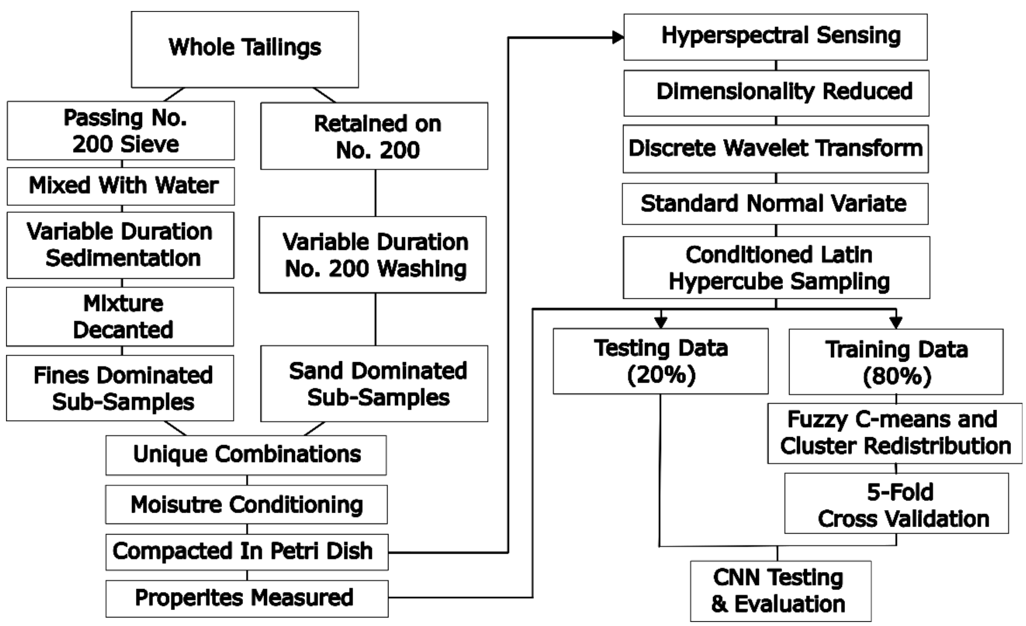

Figure 3.1. Flowchart displaying general procedures associated with material processing, test specimen preparation, hyperspectral data processing, and machine learning.

3.2.1. Test Specimens

A whole tailings sample from a gold, silver, lead, and zinc producing mine was processed and used to prepare 300 test specimens with varying geotechnical properties including variable percent sand, silt, clay, fines content, solids content, water content, saturation, void ratio, porosity, and density. The properties of the whole tailings sample are outlined in Table 3.1.

Plasticity testing was used to determine the whole tailings liquid limit (LL) and plasticity index (PI) (ASTM International, 2017b). Mechanical sieve analysis was used to determine the PSD for particles larger than 75 µm (ASTM International, 2017c) and the whole tailings was classified according to the Unified Soil Classification System (USCS) (ASTM International, 2017a). Sedimentation analysis was performed to determine the distribution of particles ranging in diameter from 2-63 µm using the *PARIO Soil Particle Analyzer* (METER Group, 2017). The standard Proctor compaction test was used to determine the maximum dry density and optimum water content according to ASTM International (2021b). Specific gravity of solids (G_s) was determined using ASTM D854 (ASTM International, 2023). Quantitative X-ray diffraction was performed to determine mineralogical composition of the whole tailings.

Table 3.1. Material properties of the whole tailings sample.

Method	Property	Sample Value
ASTM D4318	Liquid Limit, LL (%)	21
	Plasticity Index, PI (%)	1
ASTM D6913	Percent Gravel (4.75 mm - 76.2 mm)	0
	Percent Sand (75 µm - 4.75 mm)	46
	Percent Fines (< 75 µm)	54
ASTM D2487	USCS Classification	ML
METER PARIO Plus	Percent Silt (75 µm - 2 µm)	47
	Percent Clay (< 2 µm)	8
ASTM D698	Optimum Water Content, w_{opt} (%)	12
	Maximum Dry Density, ρ_d (Mg/m^3)	1.85
ASTM D854	Specific Gravity of Solids, G_s	2.62
X-ray Diffraction (% Composition)	Quartz	23
	Potassium Feldspar	27
	Calcite	4
	Dolomite	4
	Siderite	<0.5
	Magnetite	1
	Pyrite	7
	Kaolinite	0.5
	Chlorite	0.5
	Illite/Mica	31
	Mixed-Layered Illite/Smectite	2
	Percent Illite Layers in Illite /Smectite	10

Sieving, sedimentation, and decantation were used to separate the whole tailings sample into sub-samples with varying particle size distributions (referred to as tailings sub-samples) (Figure 3.1). The whole tailings sample was first processed over the No. 200 sieve (0.075 mm). The material retained on the No. 200 sieve was separated into three groups, each of which was washed over the No. 200 sieve for different durations to obtain sand tailings sub-samples with varying fines contents. The material passing the No. 200 sieve was separated into three groups then mixed with water and subject to variable-duration sedimentation, then decantated to obtain tailings sub-samples with various silt and clay contents. The particle size distribution for each processed tailings sub-sample was determined using the *PARIO Soil Particle Analyzer* (METER Group, 2017).

Portions of tailings sub-samples were combined using unique mass ratios to assemble 100 specimens with different PSDs (PSD specimens). Figure 3.2 illustrates the distribution of particles by mass for PSD specimens at various percentiles and for the whole tailings sample (original sample). The processed tailings sub-sample mass added to each PSD specimen was recorded and a weighted average was used to determine the PSD specimen percent sand, silt, and clay for each of the 100 PSD specimens. The mean particle diameter (D_{50}) of the whole tailings was 0.062 mm, the finest PSD specimen had a D_{50} of 0.008 mm and the coarsest PSD specimen had a D_{50} of 0.120 mm.

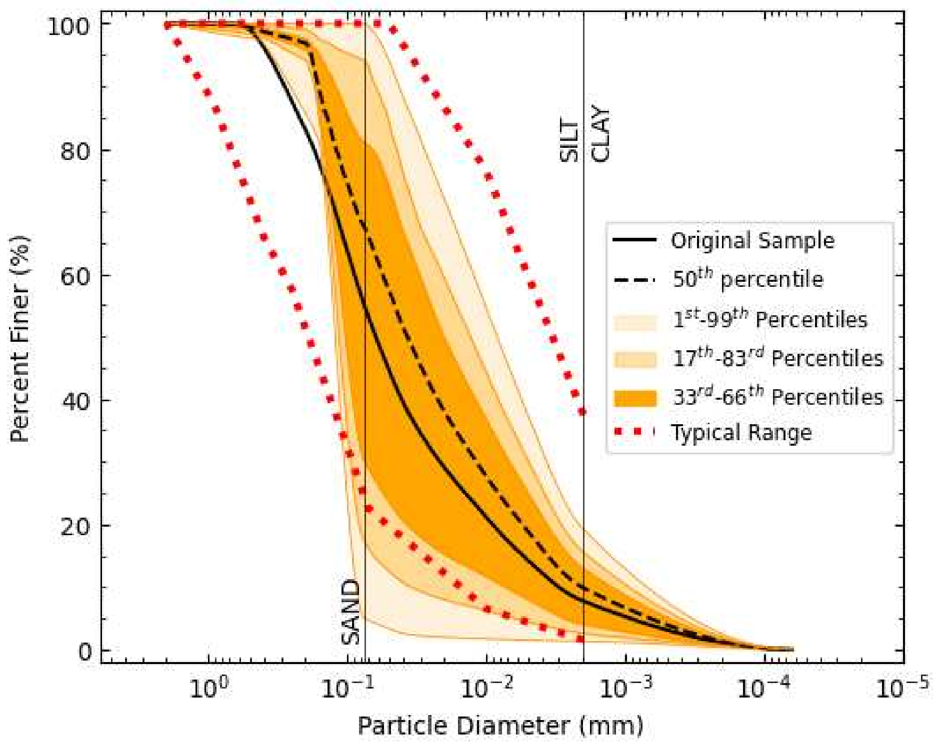

Figure 3.2. Particle size distribution of original tailings sample and PSD specimens. The 1st, 17th, 33rd, 66th, 83rd and 99th percentiles are shown for the 100 PSD specimens. Delineations for sand, silt, and clay sized particles are shown as solid vertical lines. Red dashed lines are the average upper and lower tailings bounds of mine tailings PSD from literature (after Gorakhki et al. 2019, adapted from Hamade 2017).

Each PSD specimen was prepared to three saturation and density conditions, resulting in a total of 300 test specimens. De-ionized water was added to each test specimen to achieve a target saturation, and the moisture was allowed to equilibrate for a minimum of 12 hours in an airtight container. The moisture equilibrated tailings were then compacted to a target density in a 50 mm inner diameter petri dish (15 mm height). The target saturation of each test specimen ranged from 0-100% and the target dry density was randomly distributed among test specimens and between 0.9-1.7 Mg/m^3.

The frequency distributions of measured or calculated properties among prepared test specimens are shown in Figure 3.3. The total mass and volume of each test specimen were recorded, and the gravimetric moisture content was measured according to ASTM D2216-19 (ASTM International, 2019). Specimen mass, volume, and gravimetric moisture content were used to calculate volumetric moisture content, solids content, saturation, void ratio, porosity, total density, and dry density via phase relationships. Sand, silt, clay, and fines content ranged from 0-99%, 2-84%, 1-20%, and 3-100%, respectively (Figure 3.3a-d). Solids content ranged from 65-100% (Figure 3.3e), covering a portion of the tailings continuum including thickened, paste, and filtered tailings (KCB, 2017). Gravimetric and volumetric moisture content ranged from 0-55% and 0-59%, respectively (Figure 3.3f-g) and saturation ranged from 0-100% (Figure 3.3h). Test specimen void ratios ranged from 0.5-2.0 (Figure 3.3i) and porosity from 0.34-0.67 (Figure 3.3j). For gold, silver, lead, and zinc tailings, studies have reported the in-place void ratio of slimes ranging from 0.6-1.0 (Kealy et al., 1974; Mabes et al., 1977; Blight and Steffen, 1979), 63% of test specimens had void ratios in this range. Measured total density ranged from 0.94-2.09 Mg/m^3 (Figure 3.3k) and dry density from 0.86-1.72 Mg/m^3 (Figure 3.3l). Kealy et al. (1974) and Mabes et al. (1977) reported the dry density for lead and zinc slimes to range from 1.28-1.81 Mg/m^3, respectively, 71% of test specimens had dry densities in this range. Test specimens with large void ratios and low dry densities that contained clods and were prepared to simulate conveyor deposition of filtered tailings.

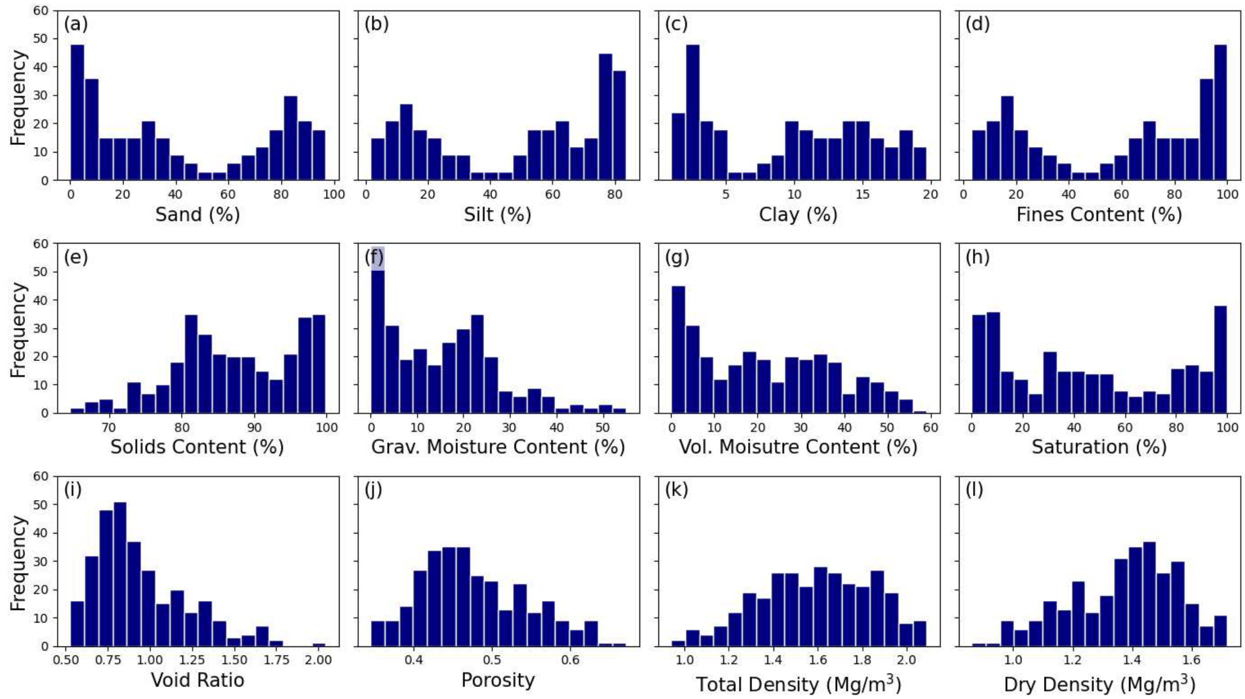

Figure 3.3. Distribution of properties among test specimens including (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

3.2.2. Hyperspectral Sensing

Hyperspectral scans were captured immediately after test specimen preparation. An *ASD TerraSpec Halo Mineral Identifier* (spectrometer) (Malvern Panalytical, 2018) was used to capture reflectance data from 350 – 2500 nm. Figure 3.4 shows a prepared test specimen and the spectrometer. The spectrometer contains three hyperspectral detector arrays with a resolution of 3 nm for the visible and near-infrared (VNIR) detector (350 – 1000 nm), 9.8 nm for the first short wave infrared (SWIR) detector (1001 – 1800 nm), and 8.1 nm for the second SWIR detector (1801 – 2500 nm) (Malvern Panalytical, 2018). The spectrometer has an internal halogen bulb for sample illumination and contains 386 spectral pixels, where each spectral pixel corresponds to a wavelength. The viewing window of the spectrometer was placed in direct contact with the tailings and hyperspectral data were captured at two locations on each test specimen. Hyperspectral data from the two locations were averaged to obtain one spectral signal for each test specimen.

Figure 3.4. Images of (a) prepared test specimen, (b) spectrometer, and (c) close up of the spectrometer in contact with a test specimen.

3.2.3. Principal Component Analysis

Principal component analysis (PCA) was used to assess how hyperspectral data relate to tailings properties. Each principal component represents a vector that is described by coefficients for each wavelength, referred to as loadings. The loadings describe how strongly each wavelength is associated with a given principal component. The explained variance ratio describes how well hyperspectral variability is captured by each principal component and principal component scores describe how significantly the hyperspectral data for each sample contributes to variability in the hyperspectral dataset. Principal component analysis was completed in Python 3 (van Rossum & Drake, 2009) using scikit-learn 1.4.0 (Pedregosa et al., 2011) and hyperspectral data from all test specimens. The explained variance ratio of each principal component was determined, and loadings were extracted for each wavelength. The sample scores associated with principal components one and two were determined and used to qualitatively visualize how hyperspectral data relate to tailings properties.

3.2.4. Machine Learning

To achieve similar resolutions for all hyperspectral detectors and reduce hyperspectral data dimensionality, the VNIR data was down sampled 3:1 to achieve a spectral resolution of 9 nm from 350 – 1000 nm, resulting in a total of 241 wavelengths. The discrete wavelet transform (DWT), which has been shown to improve performance in hyperspectral classification problems (Anand et al, 2021; Bruce et al, 2022), was used to further reduce hyperspectral dimensionality while preserving important spectral features. The PyWavelets package was used to perform the DWT using the haar wavelet (Lee et al., 2019). The DWT returns an approximation of the original spectrum that is comprised of wavelet decomposition coefficients that describe the original signal. The DWT yields a compressed signal with a down sampling of 2:1, resulting in 121 features.

The standard scaler transform was applied to the compressed signal to standardize the hyperspectral data by removing the mean and scaling to unit variance (Pedregosa et al., 2011). Scaling data to have a mean of zero and unit variance has been linked with improved model convergence and better compatibility with some objective functions (Pedregosa et al., 2011; Dangeti, 2017). Tailings properties were normalized between 0 and 1 to ensure all properties had similar magnitudes during training and testing.

Training and testing datasets were determined using conditioned Latin Hypercube Sampling (cLHS) (Minasny & McBratney, 2006) informed by compressed and scaled hyperspectral data. The cLHS algorithm was used to approximate the multivariate distribution of hyperspectral data in both the training and testing datasets, which has been associated with improved model performance (Althnian et al., 2021). The cLHS algorithm was run for 100,000 iterations to extract 20% of the data for testing (n=60) leaving 80% of the data for training (n=240). Figure 3.5 displays the distribution of test specimen properties in the training and testing datasets from cLHS. Results indicate sampling of testing data using cLHS based on hyperspectral data captures the multivariate distribution of test specimen properties. For each test specimen attribute, the distribution of testing data is within the limits of training data and has a similar distribution of data. The results from cLHS confirm that the CNN testing data is comparable to the training data which has been associated with improved model performance (Althnian et al., 2021).

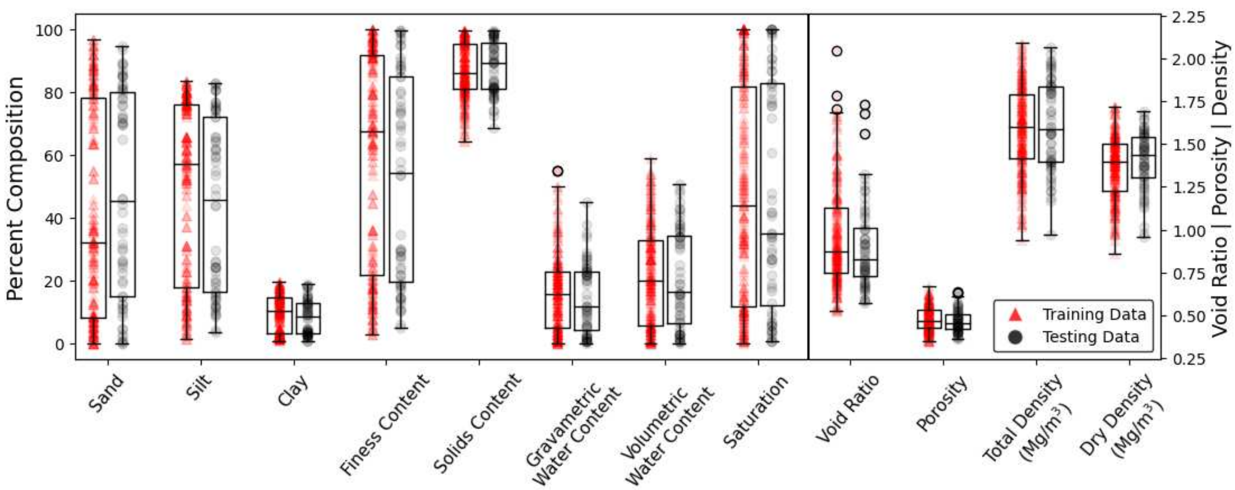

Figure 3.5. Boxplot displaying tailings attributes for test specimens in the training and testing data.

K-fold cross validation was used to select the CNN hyperparameters (viz. number of layers, number of neurons, activation function, learning rate, etc.). Cross validation helps to generalize a model, since hyperparameter optimization is informed by multiple validation datasets. Five-folds were selected for cross validation resulting in 20% of the training data contained in each fold (n=48). The fuzzy c-means algorithm (FCM) (Bezdek et al., 1984) was used to ensure each fold contained reflectance data that were representative of the training dataset, and the elbow method was used to determine the optimum number of clusters for the FCM algorithm. The FCM algorithm was executed for n=2 to n=10 clusters and the within-cluster sum of squares (WCSS) was plotted for each n-number of clusters. The point of diminishing reduction in WCSS (viz. the plot elbow) was identified at three clusters and the FCM algorithm was executed a final time to cluster hyperspectral data into three groups. The spectra associated with each of the three clusters, and corresponding test specimen data, were then randomly reassigned to five groups to be used for 5-fold cross validation.

The 5-fold training data was then used to optimize a 1-D CNN, using a different fold as validation data for each optimization. The Keras deep learning application program interface running on the TensorFlow machine learning platform was used to construct the CNN in Python 3 (Chollet, 2015; Abadi et al., 2015; van Rossum & Drake, 2009). The optimized CNN for each validation set was then used to inform the final model hyperparameters (Figure 3.6).

The final optimized CNN was trained using the entire training dataset. The testing dataset was used to assess the performance of the optimized and trained CNN. The root mean squared error (RMSE), mean bias error (MBE), coefficient of determination (COD), and the cumulative distribution function (CDF) were used to assess the performance of the CNN for the prediction of 12 tailings properties: percent sand, percent silt, percent clay, fines content, solids content, gravimetric moisture content, volumetric moisture content, saturation, void ratio, porosity, total density, and dry density.

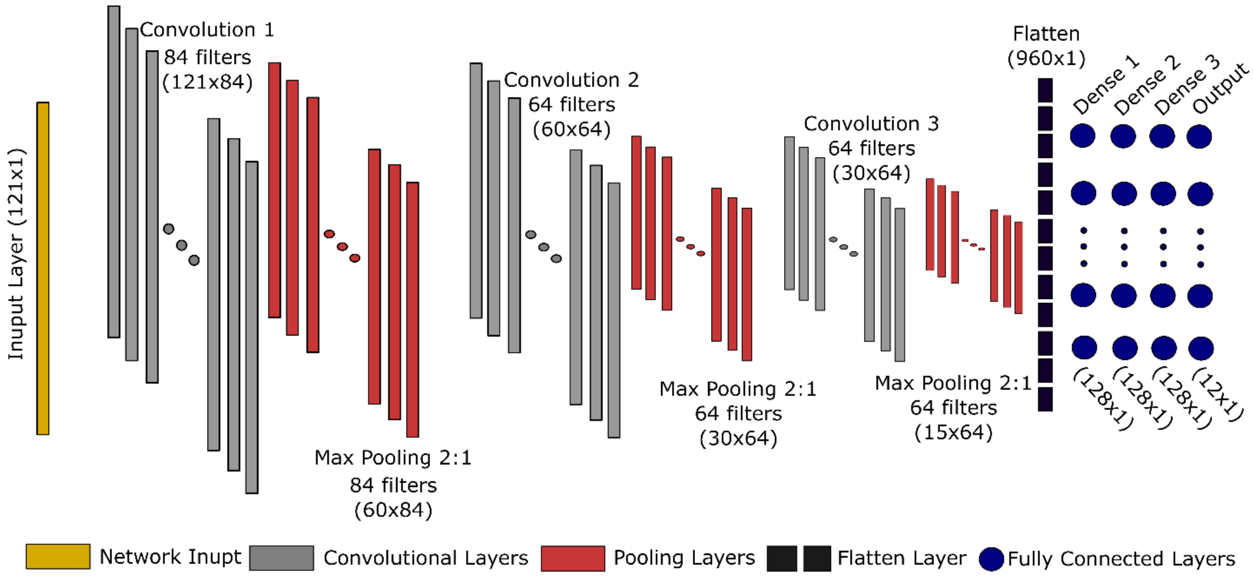

Figure 3.6. Architecture of the optimized CNN informed by cross validation. Values in parenthesis represent the output shape of each layer.

3.3. Results

3.3.1. Hyperspectral Data

The range of hyperspectral signals for test specimens is shown in Figure 3.7. The relative reflectance represents the spectrometer's measure of radiation reflected from the test specimen surface and generally ranges from 0 to 1, where 0 is complete absorption and 1 is total reflectance. The range of relative intensity for the test specimens generally ranges from 0 to 0.55. Hyperspectral signals with absorptive features (local minima) at 1450 and 1920 nm correspond to the presence of water (Jacquemoud and Ustin, 2003; van der Meer, 2004) and have lower relative reflectance at all wavelengths. The presence of a pronounced absorptive feature at 2200 nm is attributed to the presence of clay minerals containing hydroxyl groups including illite, chlorite, and kaolinite (Laukamp et al., 2021). The absorptive feature at 2200 nm is less pronounced for wetter samples, likely due to the presence of water masking the hydroxyl absorption feature (Demattê et al., 2010).

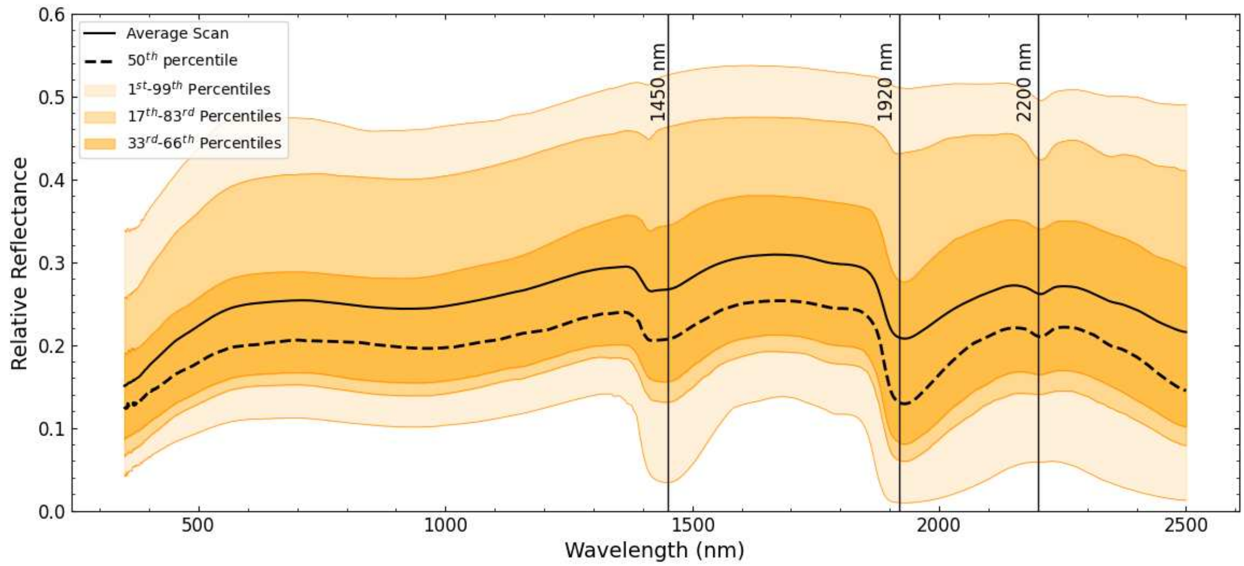

Figure 3.7. Percentile ranges for hyperspectral signals including the 1st to 99th percentile, 17th to 83rd percentile, and 33rd to 66th percentile.

The averaged hyperspectral signals for the highest and lowest 10% of each tailings property are shown in Figure 3.8. The least variation between the highest and lowest 10% signal is observed for PSD parameters including percent sand, silt, clay, and fines (Figure 3.8a-d). Studies have related increased particle size with decreased reflectance in the 350-2500 nm range (Clark et al., 1993; Sadeghi et al., 2018). However, this trend is only present at wavelengths greater than 1380 nm in this study (Figure 3.8a-d). Test specimens with high solids content have greater relative reflectance, attributed to the decreased presence of water at high solids content (Figure 3.8e). Specimens with high gravimetric water content, volumetric water content, saturation, and total density have lower relative reflectance and more pronounced features at water absorption bands, corresponding to the increased presence of water (Figure 3.8f-h and 4.8k). High dry density (low void ratio and porosity) corresponds with lower relative intensity (Figure 3.8i-j and 4.8l), which has been observed in previous studies on sandy materials (Bachman et al., 2014, Carson et al., 2015), however the trend may vary based on soil type (Demattê et al., 2010).

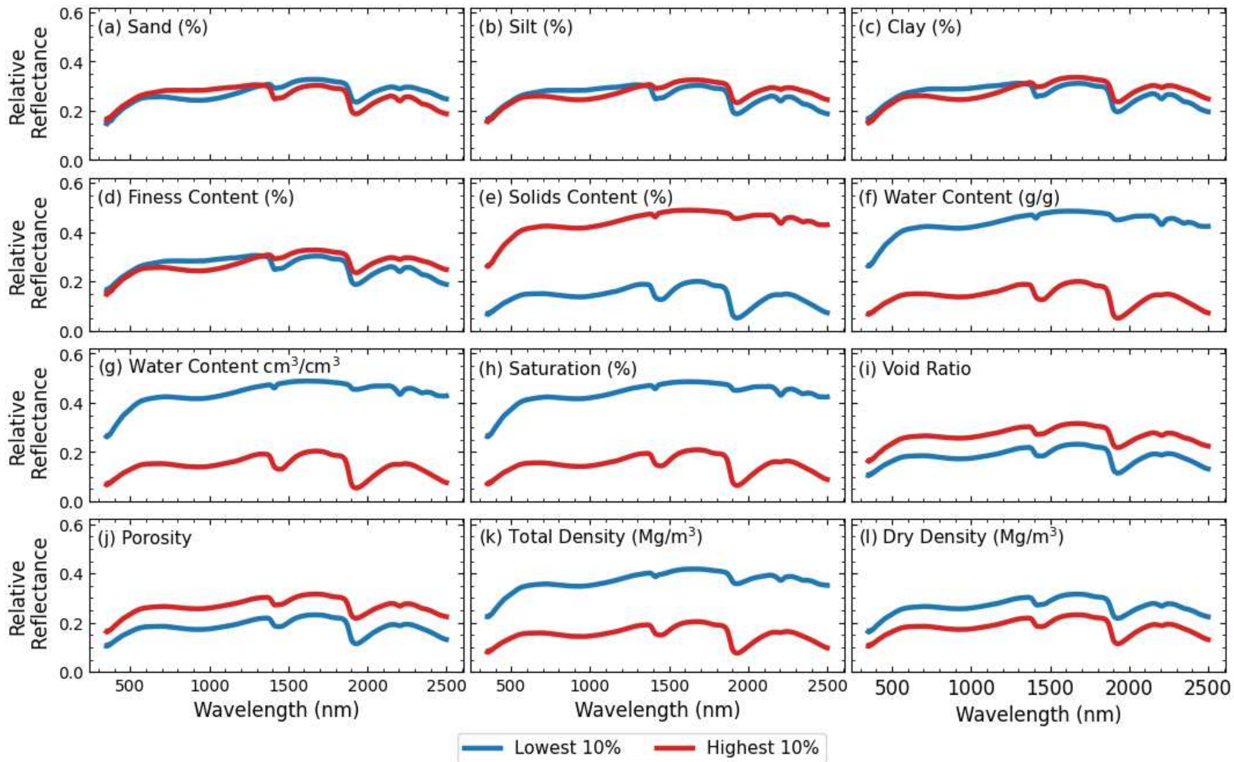

Figure 3.8. Average hyperspectral signal for the highest and lowest 10% of values for each property for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

3.3.2. Principal Component Analysis

Principal component analysis was completed to qualitatively visualize how hyperspectral data relates to tailings properties. The wavelength loadings from PCA for the first three principal components are displayed in Figure 3.9. Principal Component 1 (PC1) has an explained variance ratio of 98.2%, demonstrating that PC1 describes most of the variability in the hyperspectral dataset. Local maxima in the PC1 loadings curve occur at 1450 and 1920 nm, highlighting the significance of water in the variability of the hyperspectral dataset. The PC1 loadings exhibit minimal variation across all wavelengths, indicating the influence of overall signal position in the variability of the hyperspectral dataset, primarily attributed to the impact of water on hyperspectral signal position. Principal Component 2 (PC2) has an explained variance ratio of 1.4% with local maxima near 1450, 1890, 2020, 2200, and 2430 nm, likely cause by hydroxyl absorption feature near 1450 nm and the Al-OH combination absorption band at 2200 nm (Kokaly et al., 2017; Swayze, 1997). PC3 has an explained variance ratio of 0.2% and has local maxima loadings at 300 and 1120 nm.

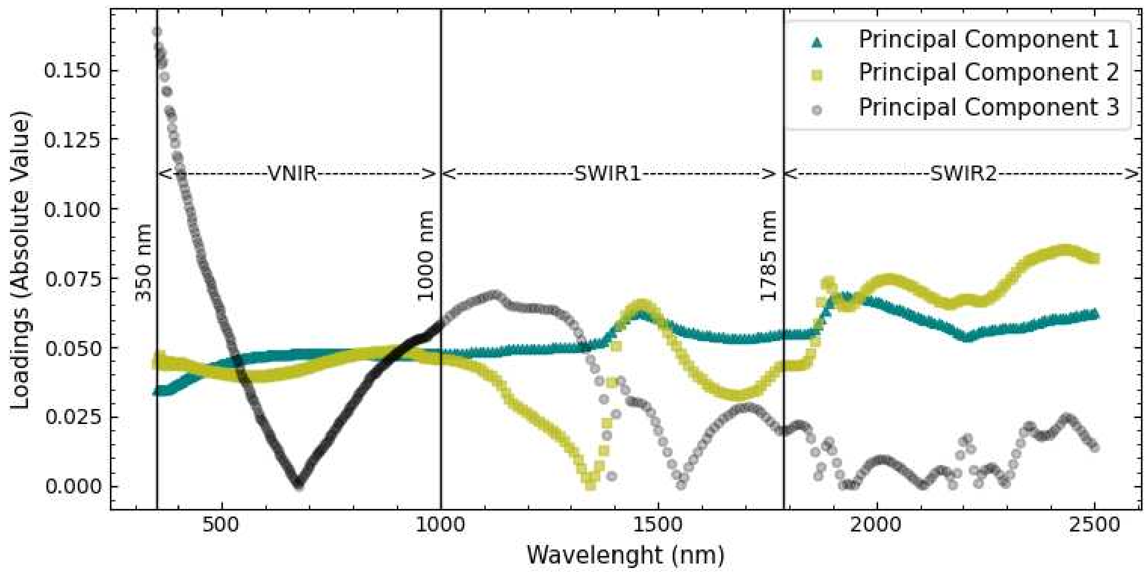

Figure 3.9. Absolute value of loadings at each wavelength for PC1, PC2, and PC3. Loadings describe the importance of features (wavelengths) in describing the variability of the dataset.

Figure 3.10 displays the test specimen scores for PC1 and PC2 from PCA, with the infill color intensity of each point representing the test specimen value for each property. The principal component score indicates the extent each test specimen aligns with the pattern of variation captured by a principal component. A greater positive score suggests strong positive correlation with the principal component while negative score indicates a negative correlation, with scores near zero having weak correlations with a principal component. PSD parameters including sand, silt, clay, and fines content are somewhat related to PC2, with higher sand content generally corresponding to higher PC2 scores (Figure 3.10a) and vice versa for silt, clay, and fines content (Figure 3.10b-d). Solids content, gravimetric moisture content, volumetric moisture content, saturation, and total density all have a strong relationship with PC1, providing the influence these properties have on spectral position and water absorption features (Figure 3.10e-h and 4.10k). No visibly obvious relationship is observed between PC1 or PC2 with void ratio, porosity, and dry density (Figure 3.10i-j and 4.10l). Results demonstrate PC1 is strongly associated with tailings moisture metrics and PC2 is associated with PSD parameters.

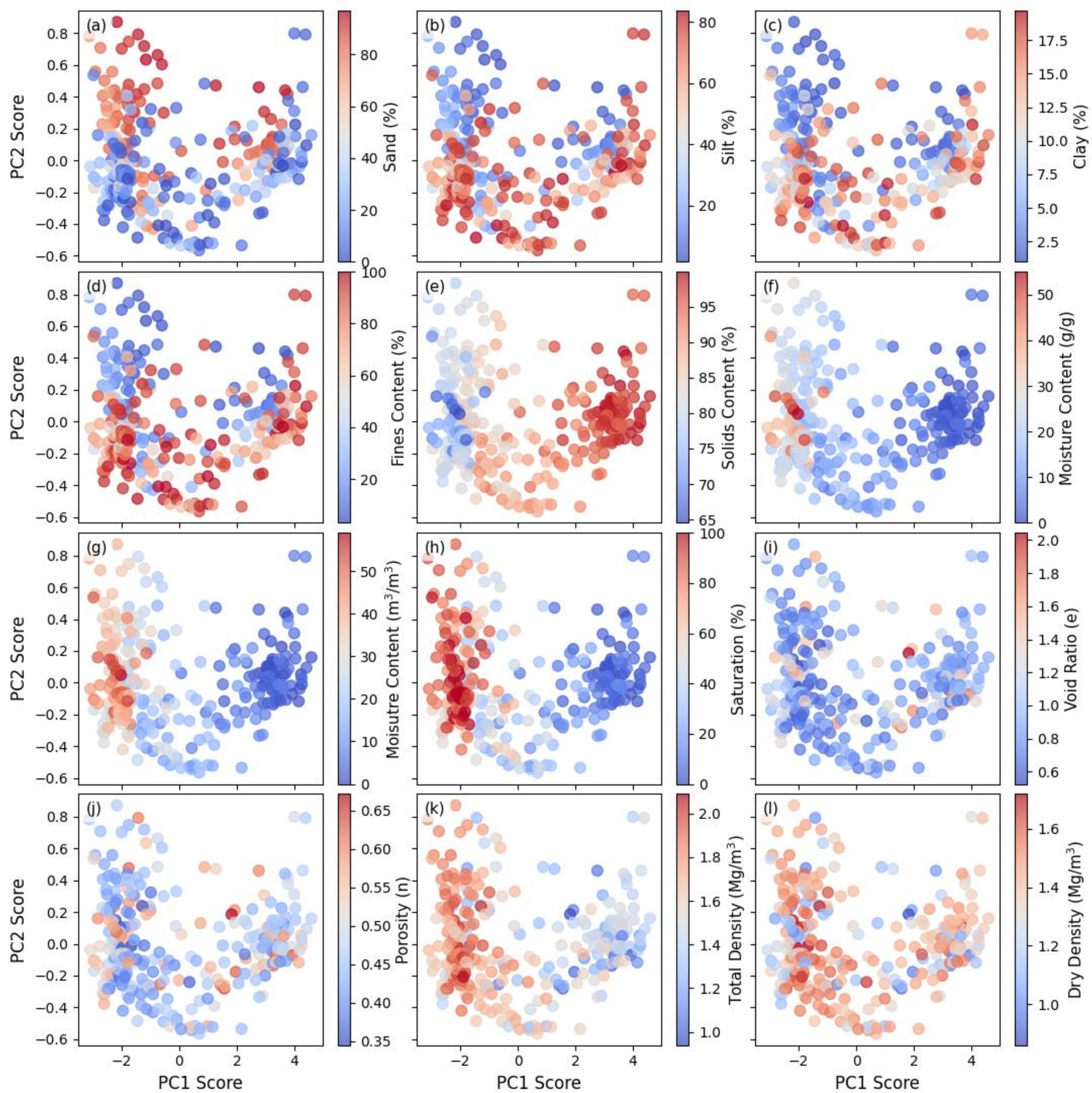

Figure 3.10. Tailings samples plotted in the principal component space for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density. Axis values correspond to the principal component score and point infill colors represent the property value for each test specimen.

3.3.3. Machine Learning

Tailings properties prediction accuracy using the trained CNN are shown in Figure 3.11. Predictions of sand, silt, clay, and fines content capture the trend for various PSDs, with COD values greater than 0.94 and RMSEs of 5.4%, 5.1%, 1.2%, and 5.3% respectively (Figure 3.11a-d). Predictions of solids content have a high COD of 0.97 and a low RMSE of 1.4% (Figure 3.11e). Properties describing tailings moisture, including gravimetric moisture content, volumetric moisture content, and degree of saturation, have a strong relationship between predicted and measured properties with RMSEs of 2.2%, 2.8% and 8.3% respectively and COD values greater than 0.94 (Figure 3.11f-h). Overall, the trend of void ratio and porosity is captured by the CNN predictions but have low COD values of 0.20 and 0.27 and RMSEs of 0.23 and 0.05 respectively (Figure 3.11i-j). Predictions of total and dry density resulted in RMSEs of 0.16 and 0.14 Mg/m^3 and COD values of 0.65 and 0.28 respectively (Figure 3.11k-l).

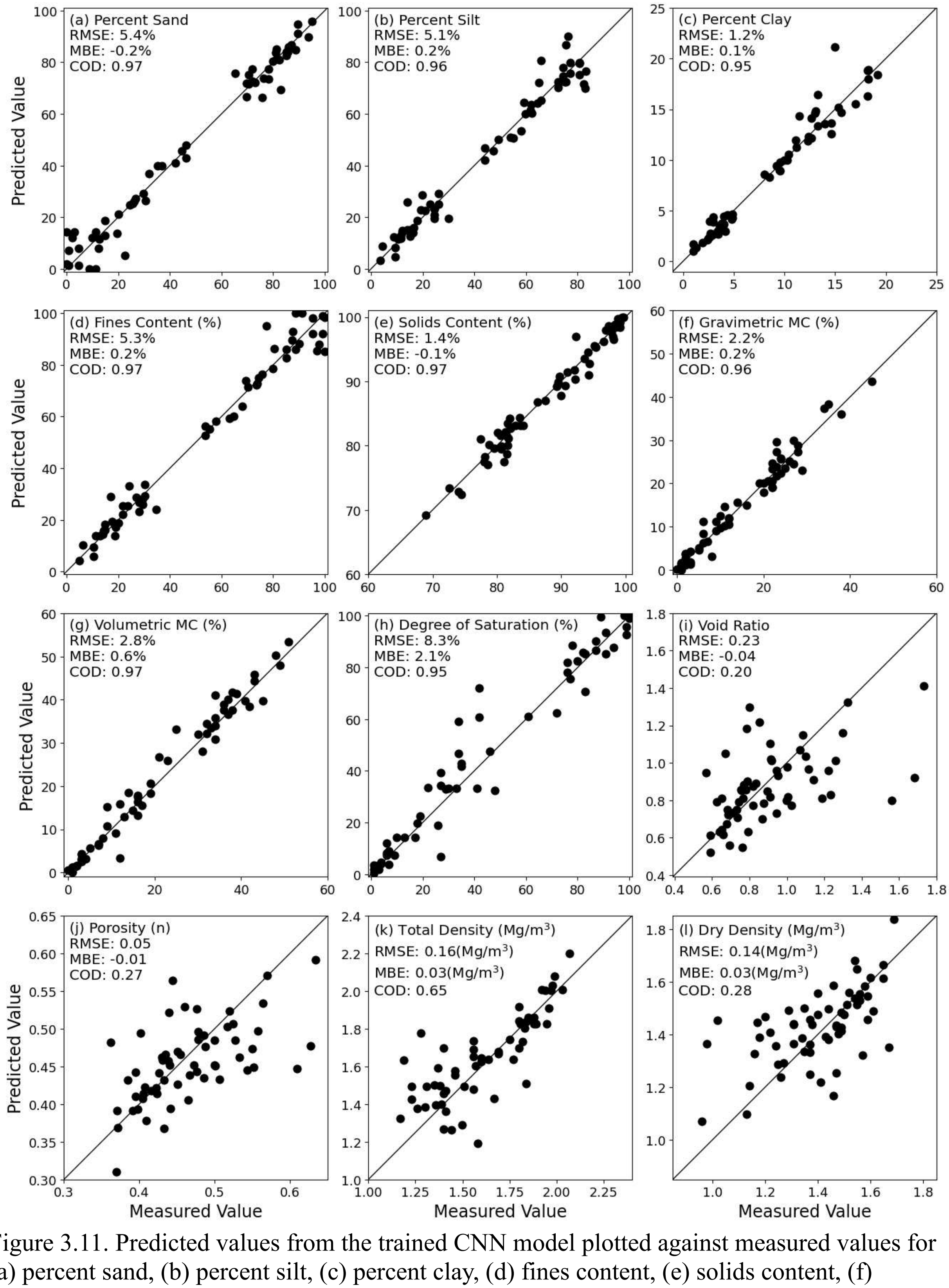

Figure 3.11. Predicted values from the trained CNN model plotted against measured values for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) degree of saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

The CDF and MBE for each predicted tailings property are displayed in Figure 3.12. The CDF is used to illustrate the probability that CNN prediction error is less than a given value and a positive MBE corresponds with overpredicted values. The CDF for PSD parameters including sand, silt, clay, and fines content shows that 80% of predictions have absolute errors less than approximately 5%, and 90% less than approximately 13% (Figure 3.12a-d). Predictions of percent clay have especially high confidence, with 90% of errors less than 2%. Bias for all PSD parameters are near zero and between -0.20% and 0.23%. Predictions of solids content, gravimetric moisture content, and volumetric moisture content have 80% of predictions with absolute errors less than 4% and 90% less than 6% absolute error (Figure 3.12e-g). Predictive bias of solids content, gravimetric moisture, and volumetric moisture content is between -0.7% and 0.1%. Predictions of saturation have a 90% probability that absolute errors are less than 13% and have a bias of -2.1% (Figure 3.12h). Predictions of void ratio and porosity have biases of 0.04 and 0.01 respectively and 90% of predictions have absolute errors less than 0.38 and 0.10 respectively (Figure 3.12i-j). The CDF for total and dry density perform similarly with 80% of predictions having an absolute error less than 0.19 Mg/m^3, 90% less than 0.27 Mg/m^3, and a bias of -0.03 Mg/m^3 for both total and dry density (Figure 3.12k-l).

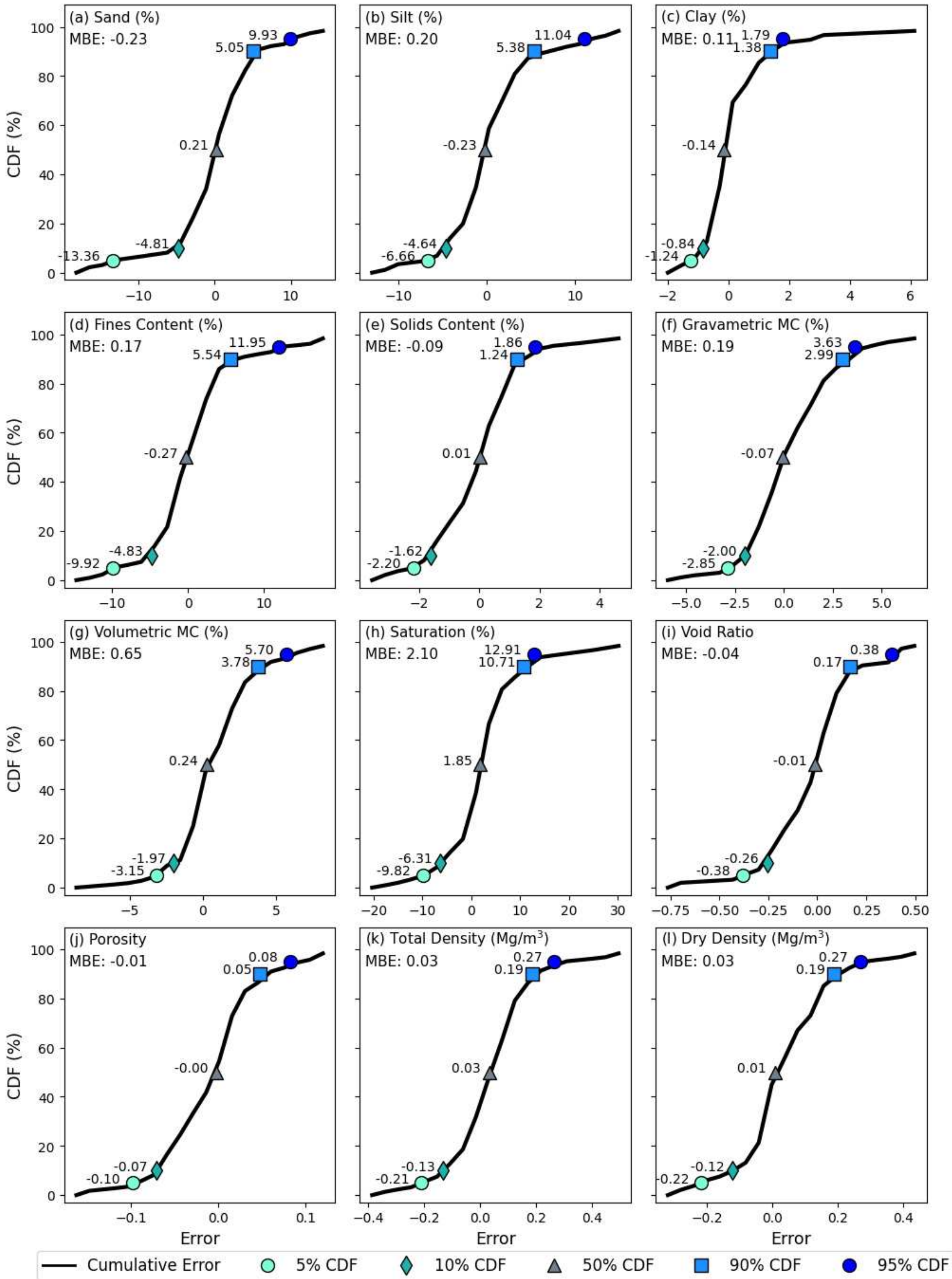

Figure 3.12. Cumulative distribution function (CDF) of the error for each predicted property and CDF values at the 5th, 10th, 50th, 90th, and 95th percentiles. Positive mean bias error (MBE) values correspond to overpredictions.

For some tailings properties, performance of the CNN is observed to vary with different levels of saturation (Table 3.2). Predictions of percent sand, silt, and clay, fines content, solids content, gravimetric water content and volumetric water content perform similarly for both high and low saturation conditions. However, for saturation, void ratio, porosity, total density, and dry density, lower errors are observed for samples with saturation above 50%. Predictions for samples with saturations greater than 50% have substantially greater COD values (0.76-0.80) than those with saturations less than 50% (0.09-0.18) for void ratio, porosity, total density, and dry density. Additionally, RMSE values are substantially reduced at saturations greater than 50% for predictions of saturation, void ration, porosity, total density, and dry density.

Table 3.2. CNN performance metrics for saturation, void ratio, porosity, total density, and dry density for test specimens with saturations from 0-50% and 50-100%.

		Saturation (%)	Void Ratio	Porosity	Total Density (Mg/m^3)	Dry Density (Mg/m^3)
0-50% Saturation	RMSE	9.73	0.29	0.07	0.19	0.18
	MBE	3.04	-0.06	-0.02	0.05	0.05
	COD	0.78	0.09	0.11	0.18	0.11
50-100% Saturation	RMSE	5.64	0.10	0.03	0.07	0.07
	MBE	0.71	0.00	0.00	0.00	0.00
	COD	0.79	0.77	0.76	0.80	0.80

Figure 3.13 displays the results from Shapley Additive Explanations (SHAP) analysis, which uses a game theory approach to explain the importance of various input features in machine learning models (Lundberg and Lee, 2017). Greater absolute SHAP scores at a given wavelength correspond with greater influence on the prediction of a given property. The prediction of soil properties from the CNN appears to be most influenced by wavelengths between 2100 nm and 2500 nm. Wavelengths from 2200 – 2350 nm appear to strongly influence the prediction of percent sand, silt, clay, and fines content. For the prediction of solids content,

gravimetric moisture content, volumetric moisture content, and degree of saturation, most wavelengths appear to have moderate influence on the prediction of these properties. Finally, the predictions of void ratio, porosity, total density, and dry density appear to be strongly influenced by wavelengths near 600 nm, 750 nm, 900 nm, 1400 nm, and 2200 nm.

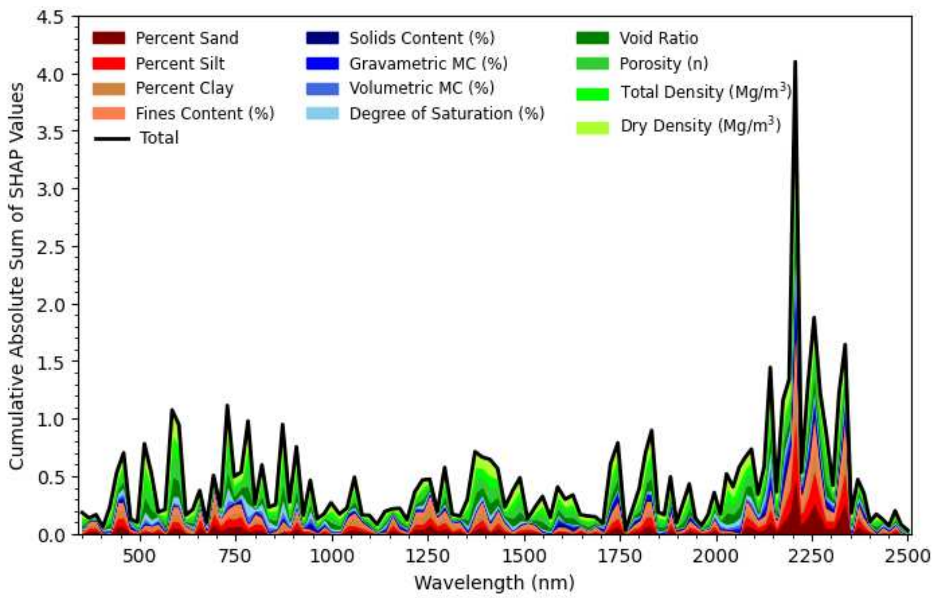

Figure 3.13. Cumulative absolute sum of SHAP values for the CNN input wavelengths. Each color represents a different soil property, and the solid black line represents the absolute sum of SHAP values for all predicted soil properties.

3.4. Discussion

The CNN results demonstrate that tailings properties related to PSD, moisture, and density can be estimated using hyperspectral data. Taskiridis et al. (2020) used hyperspectral data captured on approximately 18,000 natural soils across the European Union to predict percent sand silt and clay, and achieved RMSEs of 11.95%, 9.33%, and 4.80% respectively. CNN predictions of tailings percent sand, silt, and clay in this study have lower RMSEs compared to Taskiridis et al. (2020), which may be associated with reduced mineralogical complexity of tailings samples since tailings test specimens were prepared from a single whole tailings. The coefficients of determination for predicted percent sand, silt, and clay are greater than 0.95 demonstrating the CNN's ability to capture variability in PSD. Entezari et al., 2002 used hyperspectral data and machine learning to predict percent fines, solids content, and gravimetric moisture content of oil-sands tailings which resulted in CDFs similar to this study.

The CNN model has high predictive accuracy for both gravimetric and volumetric moisture content. High performance of moisture content predictions is associated with the strong correlation between soil moisture and spectral features (Diao et al, 2021). While the use of hyperspectral sensing for the prediction of tailings saturation has not previously been explored, results demonstrate strong promise of hyperspectral data and CNNs for the rapid measurement of tailings saturation.

Predictions of void ratio, porosity, total density, and dry density capture the general trend of tailings property variability. However, these predictions have low COD values compared to predictions of PSD and moisture content. Studies which have used hyperspectral data to predict total density have achieved results comparable to this study (Moreira et al., 2009; Katuwal et al., 2020; Haghi et al., 2021). Hyperspectral sensing of density conditions may be sensitive to sample preparation, given that the information depth of VNIR and SWIR soil spectroscopy is generally within the top 1 mm of the soil surface (Norouzi et al., 2021). Additionally, results from this study indicate density and saturation predictions have lower errors at high levels of saturation including at the critical saturation values associated with geotechnical and geochemical stability. Further studies are needed to identify contributing factors of improved performance at high saturations.

The use of hyperspectral sensing for the prediction of tailings properties has potential for the high-speed characterization of tailings properties in the field. Studies have demonstrated the use of visible light cameras on CPT probes to better characterize soil profiles (Ventola et al.,

2020) and hyperspectral sensing has been used in subsurface probes for characterization of agricultural soil properties (Cho et al., 2017; Pei et al., 2018). Additionally, tailings property predictive accuracy is expected to improve when data from multiple sensors are used as model inputs, such as data from hyperspectral sensing, current generation, and CPTs (tip resistance, sleeve friction, pore pressure, etc.) (Pei et al., 2018; Riese and Keller 2019). Future studies should explore the use of probe based hyperspectral sensing for the characterization of tailings properties in situ.

While results from this study show promise for the prediction of multiple tailings properties, the test specimens in this study did not include conventional slurry tailings. However, Entezari et al., (2020) demonstrated the accurate prediction of slurry tailings solids, water, and fines content. The tailings used in this study were from a single mine site and additional studies are needed to understand how this methodology applies to tailings from other mining facilities. Additionally, studies have demonstrated that larger datasets can correspond with improved model performance (Prusa et al., 2015; Zhu et al., 2016), therefore future studies should consider the impact of dataset size on the predictive accuracy of tailings properties.

3.5. Conclusions

The primary objective of this study was to assess the use of hyperspectral sensing for the prediction of tailings properties. A tailings from a gold, silver, lead, and zinc producing mine was processed to obtain 300 test specimens with a range of particle size distributions, moisture contents, and densities. Direct-contact hyperspectral sensing was conducted for each specimen and tailings properties were measured including percent sand, percent silt, percent clay, fines content, solids content, gravimetric moisture content, volumetric moisture content, saturation, void ratio, porosity, total density, and dry density. The tailings-hyperspectral data were partitioned into training and testing datasets. The training dataset was used to optimize and train a 1-D CNN for the simultaneous prediction of tailings properties. The testing dataset was then used to assess the performance of the CNN. Based on results of results of this study, the following conclusions are made:

1. Hyperspectral data shows promise for the prediction of tailings particle size distribution parameters including percent sand, percent silt, percent clay, and fines content. Predictions of PSD parameters resulted in high coefficient of determination (greater than 0.95) and low RMSE values (5.4%, 5.1%, 1.2%, and 5.3% respectively). The high predictive accuracy demonstrates the effectiveness of using CNNs and hyperspectral sensing for prediction of tailings properties related to particle size distribution.
2. Predictions of solids content, gravimetric water content, volumetric water content, and saturation show promise for the prediction of tailings moisture metrics. Solids content, gravimetric water content, and volumetric water content predictions have high coefficients of determination (greater than 0.96) with RMSE values below 3%. Saturation predictions result in a high coefficient of determination (0.95) and an RMSE near 8%. Predictions of saturation were observed to have lower errors at saturations greater than 50%, with RMSE reduced to 5.64%. (COD of 0.79). The strong relationship between hyperspectral signals and tailings moisture metrics coupled with the high predictive performance of the CNN demonstrates the effectiveness of using hyperspectral data for the prediction of tailings properties associated with moisture.
3. The predictions of tailings properties related to density including void ratio, porosity, total density, and dry density also show promise, with RMSEs of 0.23, 0.05, 0.16 Mg/m^3, and 0.14 Mg/m^3, respectively, and COD values of 0.20, 0.27, 0.65, and 0.28, respectively. Predictions of void ratio, porosity, total density, and dry density have lower

errors at saturations greater than 50%, with RMSEs reduced to 0.10, 0.03, 0.07 Mg/m^3, and 0.07 Mg/m^3 respectively and COD values greater than 0.77.

Hyperspectral sensing and machine learning demonstrate the potential to produce high-speed estimates of tailings properties. The inclusion of alternate data sources is expected to improve the predictive performance of tailings properties using machine learning. Futures studies may consider the use of this methodology for the prediction of tailings properties from different mine sites and explore the use of probe based hyperspectral sensing for tailings property prediction.

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jozefowicz, R., Jia, Y., Kaiser, L., & Kudlur, M. (2015). TensorFlow: Large-scale Machine Learning on Heterogeneous Systems. Software available from tensorflow.org.

Aachib, M., Mbonimpa, M., Aubertin, M. (2004). Measurement and Prediction of the Oxygen Diffusion Coefficient in Unsaturated Media, with Applications to Soil Covers. *Water, Air, and Soil Pollution*. 156: 163-193. http://doi.org/10.1023/B:WATE.0000036803.84061.e5

Ahlvin, R.B., Haley, P.W., 1992. NATO Reference Mobility Model, Edition II user's guide. GL-92-19. Vicksburg, MS: Waterways Experiment Station.

Ali, G., Roy, A., Legendre, P. (2010). Spatial Relationships Between Soil Moisture Patterns and Topographic Variables at Multiple Scales in a Humid Temperate Forested Catchment. *Water Resources Research*. 46 (10). https://doi.org/10.1029/2009WR008804.

Althnian, A., AlSaeed, D., Al-Baity, H., Samha, A., Bin Dris, A., Alzakari, N., Elwafa., A.A., and Kurdi, H. (2021). Impact of Dataset Size on Classification Performance: An Empirical Evaluation in the Medical Domain. *Applied Sciences*. 11(2). https://doi.org/10.3390/app11020796.

Anand, R., Veni, S., Aravinth, J. (2021). Robust Classification Technique for Hyperspectral Images Based on 3D-Discrete Wavelet Transform. *Remote Sensing*. (13)7. https://doi.org/10.3390/rs13071255.

ASTM International. (2015). *Standard Practice for Thin-Walled Tube sampling of Fine-Grained Soils for Geotechnical Purposes* (ASTMM D1587M-15).

ASTM International. (2017a). *Standard Practice for Classification of Soils for Engineering Purposes (Unified Soil Classification System)* (ASTM D2487-17e1).

ASTM International. (2017b). *Standard Test Methods for Liquid Limit, Plastic Limit, and Plasticity Index of Soils* (ASTM D4318-17e1).

ASTM International. (2017c). *Standard Test Methods for Particle-Size Distribution (Gradation) of Soils Using Sieve Analysis* (ASTM D6913/D6913M-17).

ASTM International. (2019). *Standard Test Methods for Laboratory Determination of Water (Moisture) Content of Soil and Rock by Mass* (ASTM D2216-19).

ASTM International. (2020). *Standard Test Methods for Determining the Water (Moisture) Content, Ash Content, and Organic Material of Peat and Other Organic Soils* (ASTM D2974-20e1).

ASTM International. (2021a). *Standard Test Methods for Particle-Size Distribution (Gradation) of Soils Using the Sedimentation (Hydrometer) Analysis* (ASTM D7928-21).

ASTM International. (2021b). *Standard Test Methods for Laboratory Compaction Chacteristics of Soil Using Standard Effort* (12.400 ft-lbf/ft3 (600kN-m/m3)) (ASTM D698-12).

ASTM Internation. (2022). *Standard Test Method for Standard Penetration Test (SPT) and Split-Barrel Sampling of Soils* (ASTM D1586/D1586M-18e1).

ASTM International. (2023). *Standard Test Methods for Specific Gravity of Soil Solids by the Water Displacement Method.* (ASTM D854-23).

Bachman, C., Philpot, W., Abelev, A., Korwan, D. (2014). Phase Angle Dependence of Sand Density Observable in Hyperspectral Reflectance. *Remote Sensing of Environment*. 150:53-65. https://doi.org/10.1016/j.rse.2014.03.024.

Bezdek, J.C., Ehrlich, R., Full, W. (1984). FCM: The Fuzzy c-Means Clustering Algorithm. *Computer Geoscience*. 10:191–203. https://doi.org/10.1016/0098-3004(84)90020-7

Bindner, J.R. (2020). Procedure for Measurement of Surficial Soil Strength via Bevameter. [Colorado State University]. https://mountainscholar.org/handle.

Blight, G., and Steffen, D. (1979). Geotechnics of Gold Mine Waste Disposal. In Current Geotechnical Practice in Mine Waste Disposal. New York: American Society of Civil Engineers. pp. 1–52.

Brooks, R.H, and A.T. Corey. (1964). Hydraulic Properties of Porous Media. Hydrology Papers. Colorado State University, Ft. Collins, CO.

Bruce, L.M., Koger, C.H., and Li, J. (2022). Dimensionality Reduction of Hyperspectral Data Using Discrete Wavelet Transform Feature Extraction. *IEEE Transactions on Geoscience and Remote Sensing*. 40(10). http://doi.org/10.1109/TGRS.2002.804721.

Bullock, M.D. (2023). Predicting Unsaturated Soil Strength for Mobility Assessments. [Colorado State University]. https://mountainscholar.org/handle.

Carson, T., Bachmann, C., and Salvaggio, C. (2015). Soil Signature Simulation of Complex Mixtures and Particle Size Distributions. *Optical Engineering*. 54(9). http://doi.org/10.1117/1.OE.54.9.094103.

Casagrande, A. (1948). Classification and Identification of Soils. *Transactions of the American Society of Civil Engineers.* 113(1). https://doi.org/10.1061/TACEAT.0006109.

Chaney, N., Wood, E., McBratney, A., Hempel, J., Nauman, T., Brungard, C., Odgers, N. (2016). POLARIS: A 30-Meter Probabilistic Soil Series Map of the Contiguous United States. *Geoderma*. 274:54-67. https://doi.org/10.1016/j.geoderma.2016.03.025.

Chan, S. K., Bindlish, R., O'Neill, P., Jackson, T., Njoku, E., Dunbar, S., Chaubell, J., Piepmeister, J., Yueh, S., Entekhabi, D., Colliander, A., Chen, F., Cosh, M.H., Caldwell, T., Walker, J., Berg, A., Mcnairn, H., Thibeault, M., Martínez-Fernández, J., Uldall, F., Seyfried, M., Bosch, D., Starks, P., Holifeld Collins, C., Prueger, J., vad der Velde, R., Asanuma, J., Palecki, M., Small, E.E., Zereda, M., Calvet, J., Crow, W.T., Kerr, Y. (2018). Development and Assessment of the SMAP Enhanced Passive Soil Moisture Product. *Remote sensing of environment*. 204:931-941. https://doi.org/10.1016/j.rse.2017.08.025.

Cho, Y., Sudduth, K.A., and Drummond, S.T. (2017). Profile Soil Property Estimation Using a Vis-NIR-EC-Force Probe. *Transactions of the ASABE*. http://doi.org/10.13031/trans.12049

Choi, K.K., Gaul, N., Jayakumar, P., and Wasfy, T.M. (2019). Framework of Reliability-Based Stochastic Mobility Map for Next Generation NATO Reference Mobility Model. *Journal of Computational and Nonlinear Dynamics*. 14(2). https://doi.org/10.1115/1.4041350.

Chollet, F. (2015). Keras. GitHub. https://github.com/fchollet/keras

Clark, R., Swayze, G., Gallagher, A., King, T., and Calvin, W. (1993). The U.S. Geological Survey, Digital Spectral Library: Version 1: 0.2 to 3.0 µm. U.S. Geological Survey, Open File Report 93-592.

Cosby, B.J., G.M. Hornberger, R.B. Clapp, and T.R. Ginn. (1984). A Statistical Exploration of the Relationships of Soil-Moisture Characteristics to the Physical Properties of Soils. *Water Resource Research.* 20:682-690. https://doi.org/10.1029/WR020i006p00682.

Cowley, G. S., Niemann, J. D., Green, T. R., Seyfried, M. S., Jones, A. S., and Grazaitis, P. J. (2017). Impacts of Precipitation and Potential Evapotranspiration Patterns on Downscaling Soil Moisture in Regions with Large Topographic Relief. *Water Resources Research.* 53(2):1553-1574. https://doi.org/10.1002/2016WR019907.

Dangeti, P. (2017). Statistics for Machine Learning: Techniques for Exploring Supervised, Unsupervised, and Reinforcement Learning Models with Python and R. Packt Publishing Ltd.

Demattê, J., Nanni, M., da Silva, A., de Melo Filho, J., Dos Santos, W., and Campos, R. (2010). Soil Density Evaluated by Spectral Reflectance as Evidence of Compaction Effects. *International Journal of Remote Sensing.* 31(2):403-422. https://doi.org/10.1080/01431160902893469.

Desilets, D., Zreda, M., and Ferre, P.A. (2010). Nature's Neutron Probe: Land Surface Hydrology at an Elusive Scale with Cosmic Rays. *Water Resources Research*. 46. W11505, https://doi.org/10.1029/2009WR008726.

Deshon, J., Niemann, J., Green, T., Jones, A., and Grazaitis, P. (2020). Stochastic Analysis and Probabilistic Downscaling of Soil Moisture in Small Catchments. *Journal of Hydrology.* 585. https://doi.org/10.1016/j.jhydrol.2020.124711.

Diao, W., Liu, G., Zhang, H., Hu, K., and Xiuliang, J. (2021). Influences of Soil Bulk Density and Texture on Estimation of Surface Soil Moisture Using Spectral Feature Parameters and an Artificial Neural Network Algorithm. *Agriculture*. 11(8). https://doi.org/10.3390/agriculture11080710.

Di Buò, B., Selänpää, J., Tapani Länsivaara, T., and D'Ignazio, M. (2018). Evaluation of Sample Quality from Different Sampling Methods in Finnish Soft Sensitive Clays. *Canadian Geotechnical Journal*. 56(8). https://doi.org/10.1139/cgj-2018-0066.

Dingman, S. L. (2015). Physical hydrology (pp. 571–577). Waveland Press.

Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A., and Vapnik, V. (1996). Support Vector Regression Machines. In Proceedings of the 9th International Conference on Neural Information Processing Systems NIPS'96. MIT Press, Cambridge, MA, USA. pp. 155–161.

Durner, W., S. C. Iden, and G. von Unold (2017), The Integral Suspension Pressure Method (ISP) for Precise Particle-Size Analysis by Gravitational Sedimentation, *Water Resource Res*earch. 53:33–48. https://doi.org/10.1002/2016WR019830.

[dataset] European Space Agency (ESA). (2016). Sentinel 2 MSI-Level 2A. Copernicus Open Access. Paris, France.

Elachi C, and Van Zyl J.J. (2006) Introduction to the Physics and Techniques of Remote Sensing. Volume 28. Wiley.com.

Entezari, I., McGowan, D., and Glavina, J. (2022). *Hyperspectral Imaging Technology for Oil Sands Tailings Characterization: Practical Aspects*. Tailings and Mine Waste. 26:1243-1252.

Eylander, J., Bieszczad, J., Ueckermann, M., Peters, J., Brooks, C., Audette, W., and Ekegren, M. (2023). Geospatial Weather Affected Terrain Conditions and Hazards (GeoWATCH) Description and Evaluation. *Environmental Modelling & Software*, 160, 105606. https://doi.org/10.1016/j.envsoft.2022.105606.

Fang, B., Lakshmi, V., Bindlish, R., and Jackson, T. (2018). AMSCOD Soil Moisture Downscaling Using Temperature and Vegetation Data. *Remote Sensing*. 10(10). https://doi.org/10.3390/rs10101575.

Ferran Parera, Núria M. Pinyol, and Eduardo E. Alonso. (2021). Massive, Continuous, and Non-Invasive Surface Measurement of Degree of Saturation by shortwave Infrared Images. *Canadian Geotechnical Journal*. 58(6): 749-762. https://doi.org/10.1139/cgj-2019-0051

Fredlund, D.G., Morgenstern, N.R., and Widger, R.A. (1978). The Shear Strength of Unsaturated Soils. *Canadian Geotechnical Journal.* 15(3): 313-321. https://doi.org/10.1139/t78- 029.

Fredlung, D. (2006). Unsaturated Soil Mechanics in Engineering Practice. *Journal of Geotechnical and Geoenvironmental Engineering.* 132(3). https://doi.org/10.1061/(ASCE)1090-0241(2006)132:3(286).

Gorakhki, M.R.H., Bareither, C.A., Scalia, J, and Jacobs, M. (2019). *Hydraulic Conductivity and Soil Water retention of Waste Rock and Tailings Mixtures*. Eighth International Conference on Case Histories in Geotechnical Engineering. https://doi.org/10.1061/9780784482148.005.

Grayson, R.B., Western, A.W., Walker, J.P. Kandel, D.G., Costelloe, J.F., and Wilson, D.J. (2006). Controls on Patterns of Soil Moisture in Arid and Semi-Arid Systems. In: Dryland Ecohydrology. Springer, Dordrecht, Netherlands. p. 109-127.

GTR (Global Tailings Review). (2020). Global Industry Standard on Tailings Management. https://globaltailingsreview.org/wp-content/uploads/2020/08/global-industry-standard_EN.pdf. Accessed January 2024.

Gupta, A., Das, B., Kumar, A., Chakraborty, P., and Mohanty, B. (2016). Rapid and Noninvasive Assessment of Atterberg Limits Using Diffuse Reflectance Spectroscopy. *Soil Science Society of America Journal.* 80(5):1283-1295. https://doi.org/10.2136/sssaj2015.11.0402.

Haghi, R., Pérez-Fernández, E., and Robertson A. (2021). Prediction of Various Soil Properties for a National Spatial Dataset of Scottish Soils Based on Four Different Chemometric Approaches: A Comparison of Near Infrared and Mid-Infrared Spectroscopy. *Geoderma*. 396. https://doi.org/10.1016/j.geoderma.2021.115071.

Haley, P.W., Jurkat, M.P., and Brady, P.M. (1979). NATO Reference Mobility Model, Edition I user's guide. Technical Report 12503. Warren, MI: Tank-Automotive Command.

Hamamatsu Photonics. (2021). InGaAs Linear Image Sensor [datasheet]. Hamamatsu Photonics. https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/ssd/g11508_etc_kmir1032e.pdf

Hamade, M. M. (2017). Undrained Shear Behavior of Mixed Mine Waste Rock and Tailings. MS Thesis. Civil and Environmental Engineering. [Colorado State University]. Fort Collins, Colorado.

Hawley, M. E., T. J. Jackson, and R. H. McCuen (1983), Surface Soil Moisture variation on Small Agricultural Satersheds. *Journal of Hydrology.* 62:179–200.

Hedley C., Poudier, P., Maddi, L. (2015). VNIR Soil Spectroscopy for Field Soil Analysis. *Communications in Soil Science and Play Analysis*. 46(1):104-121. https://doi.org/10.1080/00103624.2014.988582.

Huete, A.R. (1988). A Soil Adjusted Vegetation Index (SAVI). *Remote Sensing of the. Environment*. 25(3):295-309. https://doi.org/10.1016/0034-4257(88)90106-X.

Huete, A.R., Liu, H.Q., Batchily, K., and van Leeuwen, W. (1997). A Comparison of Vegetation Indices over a Global Set of TM Images for EOS-MODIS. *Remote Sensing of the Environment.* 59(3). https://doi.org/10.1016/S0034-4257(96)00112-5

Huete, A. R., Justice, C., and Van Leeuwen, W. (1999). MODIS Vegetation Index (MOD13) algorithm Theoretical Basis Document. Institute for Computational Earth System Science. University of California, Santa Barbara. 1(16).

Humboldt Mfg. (2023). Digital Proving Ring Penetrometer: Product Manual. https://www.humboldtmfg.com/manuals/H-4204D_man_0123.pdf

ICMM (International Council on Mining & Metals). (2021). Tailings Management: Good Practice Guide. ICMM. https://www.icmm.com. Accessed February 2024.

Jacquemoud, S., and Ustin, S.L. (2003) *Application of Radiative Transfer Models to Moisture Content Estimation and Burned Land Mapping*. Proceedings of the 4th International Workshop on Remote Sensing and GIS Applications to Forest Fire Management.

Jiang, Q., Chen, Y., Guo, L., Fei, T., and Qi, K. (2016). Estimating Soil Organic Carbon of Cropland Soil at Different Levels of Soil Moisture using VIS-NIR Spectroscopy. *Remote Sensing.* 8(9). https://doi.org/10.3390/rs8090755.

Jurkat, M.P., Nuttall, P.W., and Haley, P.W. (1975). The U.S. Army Mobility Model (AMM-75). WES Papers Presented at 5th International Conference International Society for Terrain-Vehicle Systems. 5-52.

Katuwal, S., Knadel, M., Norgaard, T., Moldrup, P., Greve, M., and de Jonge, L. (2020). Predicting the Dry Bulk Density of Soils Across Denmark: Comparison of Single-Parameter, Multi-Parameter, and Vis-NIR based Models. *Geoderma*. 361. https://doi.org/10.1016/j.geoderma.2019.114080

KCB (Klohn Crippen Berger). (2017). Study of Tailings Management Technologies. MEND Report 2.50.1. http://mend-nedem.org/wp-content/uploads/2.50.1Tailings_Management_TechnologiesL.pdf. Accessed December 2020.

Kealy, C., Busch, R., and McDonald, M. (1974). Seepage-Environmental Analysis of the Slime Zone of a Tailings Pond. RI 7939. Washington, DC: U.S. Bureau of Mines.

Kerr, Y, Wigneron, J., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M., Font, J., Reul, N., Gruhier, C., Juglea, S., Drinkwater, M., Hahne, A., Martin-Neira, M., and Mecklenburg, S. (2010). *The SMOS Mission: New Tool for Monitoring Key Elements of the Global Water Cycle*. Proceedings of the IEEE. 98(5):666-687. http://doi.org/10.1109/JPROC.2010.2043032.

Kim, S., Heysun, L., Woo, N., and Kim, J. (2007) Soil Moisture Monitoring on a Steep Hillside. *Hydrological Processes*. 21(21):2823-2964, http://doi.org/10.1002/hyp.6508.

Knadel, M., Rehman, H., Pouladi, N., Wollesen de Jonge, L., Moldrup, P., and Arthur, E. (2021). Estimating Atterberg Limits of Soils from Reflectance Spectroscopy and Pedotransfer Functions. *Geoderma.* 402. https://doi.org/10.1016/j.geoderma.2021.115300

Kokaly, R.F., Clark, R.N., Swayze, G.A., Livo, K.E., Hoefen, T.M., Pearson, N.C., Wise, R.A., Benzel, W.M., Lowers, H.A., Driscoll, R.L., and Klein, A.J. (2017). USGS Spectral Library Version 7: U.S. Geological Survey Data Series 1035, 61 p., https://doi.org/10.3133/ds1035.

Kriegler FJ, Malila WA, Nalepka RF, and Richardson, W. (1969). Preprocessing Transformations and Their Effect on Multispectral Recognition. *Remote Sensing of the Environment.* 6:97–132.

Laukamp, C., Rodger, A., LeGras, M., Lampinen, H., Lau, I., Pejcic, B., Stromberg, J., Francis, N., and Ramanaidou, E. (2021). Mineral Physiochemistry Underlying Feature-Based Extraction of Mineral Abundance and Composition from Shortwave, Mid and Thermal Infrared Reflectance Spectra. *Minerals*. 11(4). https://doi.org/10.3390/min11040347.

Lee, G., Gommers, R., Waselewski, F., Wohlfahrt, K., and O'Leary, A. (2019). PyWavelets: A Python Package for Wavelet Analysis. *The Journal of Open-Source Software*. 4(36):1237. https://doi.org/10321105/joss.01237.

Lee, R. (1964). Potential Insolation as a Topoclimatic Characteristic of Drainage Basins. *Hydrological Sciences Journal.* 9(1):27-41. https://doi.org/10.1080/02626666409493652.

Lin, H., Kogelmann, W., Walker, C., and Burns, M. (2006). Soil Moisture Patterns in a Forested Catchment: A Hydropedological Perspective. *Geoderma*. 131(3-4):345-368. https://doi.org/10.1016/j.geoderma.2005.03.013.

Lu, N., Godt, J., Wu, D. (2010). A Closed-Form Equation for Effective Stress in Variably Saturated Soil. *Water Resources Research.* 46. https://doi.org/10.1029/2009WR008646.

Liang W., Li, S., and Hung, F. (2017). Analysis of the Contributions of Topographic, Soil, and Vegetation Features on the Spatial Distributions of Surface Soil Moisture in a Steep Natural Forested Headwater Catchment. *Hydrological Processes*. 32(22):3796-3809. https://doi.org/10.1002/hyp.11290.

Liston, R. A. (1973). Combined Normal and Tangential Loading of Soil. Michigan Technological University.

Liu, H.Q., and Huete, A. (1995). A Feedback Based Modification of the NDVI to Minimize Canopy Background and Atmospheric Noise. *IEEE Transactions on Geoscience and Remote Sensing.* 33(2). https://ieeexplore.ieee.org/abstract/document/8746027.

Lundberg, S., Lee, S. (2017). *A Unified Approach to Interpreting Model Predictions*. Advances in Neural Information Processing Systems 30. https://proceedings.neurips.cc/paper_files/paper/2017

Mabes, D., Hardcastle, J., and Williams, R. (1977). *Physical properties of Pb-Zn mine-process wastes*. In Proceedings of the Conference on Geotechnical Practice for Disposal of Solid Waste Materials. New York: American Society of Civil Engineers. pp. 103–117.

Malvern Panalytical. (2018). TerraSpec Halo User Manual (ASD Document 600073). Malvern Panalytical. https://www.malvernpanalytical.com/

Mascaro G., Vivoni, E., and Deidda, R. (2011). Soil Moisture Downscaling Across Climate Regions and its Emergent Properties. *Journal of Geophysical Research*. 116(D22). https://doi.org/10.1029/2011JD016231.

McCoullough, M., Jayakumar, P., Dasch, J., and Gorsich, D. (2017). The Next Generation NATO Reference mobility model development. *Journal of Terramechanics*. 73:49-60. https://doi.org/10.1016/j.jterra.2017.06.002.

Merlin, O., Chebouni, A.G., Kerr, Y.K., and Goodrich, D.C. (2006). A Downscaling Method for Distributing Surface Soil Moisture within a Microwave Pixel: Application to the Monsoon '90 data. *Remote Sensing of the Environment.* 101(3): 379-389. https://doi.org/10.1016/j.rse.2006.01.004.

METER Group. (2017). PARIO User Manual V2017/12. METER Group AG. http://library.metergroup.com.

Minasny, B., and McBratney, A.B. (2006). A Conditioned Latin Hypercube Method for Sampling in the Presence of Ancillary Information. *Computer Geoscience.* 32:1378–1388. https://doi.org/10.1016/j.cageo.2005.12.009

Moeslund, J., Arge, L., Bocher, P., Dalgaard, T., Odgaard, M., Nygaard, B., and Svenning, J. (2013). Topographically Controlled Soil Moisture is the Primary Driver of Local Vegetation Patterns Across a Lowland Region. *Ecosphere*. 4(7):1-26. https://doi.org/10.1890/ES13-00134.1.

Mohanty, B. P., Famiglietti, J. S., and Skaggs, T. H. (2000). Evolution of Soil Moisture Spatial Structure in a Mixed Vegetation Pixel During the Southern Great Plains 1997 (SGP97) Hydrology Experiment. *Water Resources Research.* 36(12):3675–3686. https://doi.org/10.1029/2000WR900258.

Molle, J. (2005). The Accuracy of the Interpretation of CPT-Based Soil Classification Methods in Soft Soils. MSc Thesis, Section for Engineering Geology, Department of Applied Earth Sciences, Delf University of Technology, Report No. 242, Report AES/IG/05-25, December.

Moreira, C.S., Burnet, D., Verneyre, L., Sá, S.M.O., Galdos, M.V., Cerri, C.C., and Bernoux, M. (2009). Near Infrared Spectroscopy for Soil Bulk Density Assessment. *European Journal of Soil Science*. 60(5):785-791. https://doi.org/10.1111/j.1365-2389.2009.01170.x.

Morgenstern, N., Vick, S., Viotti, C., and Watts, B. (2016). Report on the Immediate Causes of the Failure of the Fundão Dam. Resolution Copper Project and Land Exchange Environmental Impact Statement. https://www.resolutionmineeis.us/documents/fundao-2016.

Morrison, K. F. (2022). Tailings Management Handbook: A Lifecycle Approach. Society for Mining, Metallurgy & Exploration, Incorporated.

Mudd, G.M. (2007). Global Trends in Gold Mining: Towards Quantifying Environmental and Resource Sustainability. *Resources Policy*. 32(1-2):42-56. https://doi.org/10.1016/j.resourpol.2007.05.002.

Mudd, G M. (2010). The Environmental Sustainability of Mining in Australia: Key Mega-Trends and Looming Constraints. *Resources Policy.* 35(2):98-115. https://doi.org/10.1016/j.resourpol.2009.12.001.

Multi-Resolution Land Characteristics Consortium (MRLC). (2019). National Land Cover Database (NLCD) 2019 Land Cover (CONUS). U.S. Geologic Survey. https://www.mrlc.gov/data/nlcd-2019-land-cover-conus

Njoku, E.G., Jackson, T.J., Lakshmi, V., Chan, T.K., and Nghiem, S.V. (2003). Soil Moisture Retrieval from AMSR-E. *IEEE Transactions on Geoscience and Remote Sensing*. 41(2): 215-229. https://doi.org /10.1109/Tgrs.2002.808243.

Nolet C., Ate, P., Roosjen, P., Bartholomeus, H., and Ruessink, G. (2014). Measuring and Modeling the Effect of Surface Moisture on the Spectral Reflectance of Coastal Beach Sand. *PLoS ONE.* 9(11). https://doi.org/10.1371/journal.pone.0112151.

Norouzi, S., Sadeghi, M., Liaghat, A., Tuller, M., Jones, S., and Ebrahimian, H. (2021). Information Depth of NIR/SWIR Soil Reflectance Spectroscopy. *Remote Sensing of Environment*. 256. https://doi.org/10.1016/j.rse.2021.112315.

Oberle, B., Brereton, D., and Mihaylova, A. (2020). Towards Zero Harm: A Compendium of Papers Prepared for the Global Tailings Review. St Gallen, Switzerland: Global Tailings Review. https://globaltailingsreview.org/.

O'Neill, P. E., S. Chan, E. G. Njoku, T. Jackson, R. Bindlish, J. Chaubell, and A. Colliander. (2021). SMAP Enhanced L3 Radiometer Global and Polar Grid Daily 9 km EASE-Grid Soil Moisture, Version 5 [Dataset]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/4DQ54OUIJ9DL.

O'Neill, P. Entekhabi, D., Njoku, E., and Kellogg, K. (2010). *The NASA Soil Moisture Active Passive (SMAP) mission: Overview.* IEEE International Geoscience and Remote Sensing Symposium. https://doi.org/10.1109/IGARSS.2010.5652291.

Pauly, M. J. (2019). Modeling and Field Evaluation of Unsaturated Soil Strength for Vehicle Mobility [Colorado State University]. https://mountainscholar.org/handle.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passo, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É. (2011). Scikit-Learn: Machine Learning in

Python. *Journal of Machine Learning Research*. 12(85):2825-2830. https://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html.

Pei, X., Sudduth, K., Veum, K., and Minzan, L. (2018). Improving In-Situ Estimation of Soil Profile Properties Using a Multi-Sensor Probe. *Sesnors*. 19(1011). http://doi.org/10.3390/s19051011.

Peng, J., Loew, A., Zhang, S., Wang, J., and Niesel, J. (2016). Spatial Downscaling of Satellite Soil Moisture Data Using a Vegetation Temperature Condition Index. *IEEE Transactions on Geoscience and Remote Sensing.* 54(1): 558-566. https://doi.org/10.1109/TGRS.2015.2462074.

Priddy, J., Berney, E., and Peters, J. (2012). Effects of Near-Surface Hydrology on Soil Strength and Mobility. *Geological Society, London, Special Publications*. 362:301-320. https://doi.org/10.1144/SP362.17.

Prusa, J., Khoshgoftarr, T., and Seliya, N. (2015). *The Effect of Dataset Size on Training Tweet Sentiment Classifiers.* International Conference on Machine Learning Applications. https://doi.org/10.1109/ICMLA.2015.22

Purushothaman, N., Reddy, N., Majeed, I., Chowdhury, P., and Das, B. (2024). Diffuse Reflectance Spectroscopy and Digital Soil Mapping for Assessing Soil-Associated Off-Road Vehicle Mobility Risk. *Journal of Environmental Management.* 35. https://doi.org/10.1016/j.jenvman.2023.119874.

Quinlan, J.R. (1992). *Learning with Continuous Classes*. Proceedings of the 5th Australian Joint Conference on Artificial Intelligence. https://doi.org/10.1142/9789814536271

Quinlan, J.R. (1993). *Combining Instance-Based and Model-Based Learning*. Machine Learning: Proceedings of the Tenth International Conference.

Ramirez-Lopez, L., Behrens, T., Schmidt, K., Stevens, A., Demattê, J.A.M., and Scholten, T. (2013). The Spectrum-Based Learner: a New Local Approach for Modeling Soil Vis-NIR Spectra of Complex Datasets. *Geoderma*. https://doi.org/10.1016/j.geoderma.2012.12.014.

Ranney, K.J., Niemann, J.D., Lehman, B.M., Green, T.R., and Jones, A.S. (2015). A Method to Downscale Soil Moisture to Fine Resolutions using Topographic, Vegetation and Soil Data. *Advances in Water Resources.* 76:81–96. https://doi.org/10.1016/j.advwatres.2014.12.003.

Riese, F.M., and Keller, S. (2019). Soil Texture Classification with 1D Convolutional Neural Networks Based on Hyperspectral Data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. IV-2 (W5). https://doi.org/10.5194/isprs-annals-IV-2-W5-615-2019.

Rezanezhad, F., Price, J. S., Quinton, W. L., Lennartz, B., Milojevic, T., and Van Cappellen, P. (2016). Structure of Peat Soils and Implications for Water Storage, Flow and Solute Transport: A Review Update for Geochemists. *Chemical Geology*. 429:75-84. https://doi.org/10.1016/j.chemgeo.2016.03.010.

Robertson, P. (2016). Cone Penetration Test (CPT)-Based Soil Behaviour Type (SBT) Classification System – an update. *Canadian Geotechnical Journal*. 53(12). https://doi.org/10.1139/cgj-2016-0044.

Robertson, P., Melo, L., Williams, D., and Wilson, G. (2019). Report of the Expert Panel on the Technical Causes of the Failure of Feijão Dam I. http://www.b1technicalinvestigation.com/.

Rodríguez, R., Muñoz-Moreno, A., Caparrós, A., García-García, C., Brime-Barrios, A., Arranz-González, J., Rodríguez-Gómez, V., Fernández-Naranjo, F., Alcolea, A. (2021). How to Prevent Flow Failures in Tailings Dams. *Mine Water and Environment*. 40:83-112. http://doi.org/10.1007/s10230-021-00752-8

Rosenbaum, U., Bogena, H. R., Herbst, M., Huisman, J.A., Peterson, T.J., Weuthen, A., Western, A.W., and Vereecken, H. (2012). Seasonal and Event Dynamics of Spatial Moisture Patterns at the Small Catchment Scale. *Water Resources Research.* 48. https://doi.org/10.1029/2011WR011518.

Sadeghi, M., Babaeian, E., Tuller, M., and Jones, S. B. (2017). The Optical Trapezoid Model: A Novel Approach to Remote Sensing of Soil Moisture Applied to Sentinel-2 and Landsat-8 74 observations. *Remote sensing of environment*. 198:52-68. https://doi.org/10.1016/j.rse.2017.05.041.

Sadeghi, M., Babaeian, E., Tuller, M., and Jones, S. (2018). Particle Size Effects on Soil Reflectance Explained by an Analytical Radiative Transfer Model. *Remote Sensing of Environment*. 210:375-386. https://doi.org/10.1016/j.rse.2018.03.028.

Säfken, B., Silbersdorff, A., & Weisser, C. (2020). *Learning deep perspectives on deep learning algorithms and Artificial Intelligence Benjamin Säfken, Alexander Silbersdorff, Christopher Weisser (eds.)*. Universitätsverlag Göttingen.

Schönauer, M., Prinz, R., Väätäinen, K., Astrup, R., Pszenny, D., Lindeman, H., and Jaeger, D. (2022). Spatio-Temporal Prediction of Soil Moisture using Soil Maps, Topographic Indices, and SMAP Retrievals. *International Journal of Applied Earth Observations and Geoinformation*. 108. https://doi.org/10.1016/j.jag.2022.102730.

Shaikh, M., Jafarzadeh, K., Thörnberg, B., and Casselgren, J. (2021). Calibration of a Hyper-Spectral Imaging System Using a Low-Cost Reference. *Sensors.* 21(11). https://doi.org/10.3390/s21113738.

Sjöström, M., Wold, S., Lindberg, W., Persson, J., and Martines, H. (1983). The Multivariate Calibration Problem in Chemistry Solved by Partial Least-Squares Method in Latent Variables. *Analytica Chimica Acta*. (150):61-170. https://doi.org/10.1016/S0003-2670(00)85460-4.

Soil Survey Staff, Natural Resources Conservation Service, United Stated Department of Agriculture. Web Soil Survey. http://websoilsurvey.nrcs.usda.gov/. (Accessed July 2023)

Stevens, M., Towne, B., Mason, G., Priddy, J., Osorio, J., and Barella, C. (2013). Procedure for One-Pass Vehicle Cone Index Determination for Acquisition Support. (ERDC/GSL SR-13-2). US Army Corps of Engineers. https://discover.dtic.mil/technical-reports/.

Stevens Water Monitoring System, Inc. (2018) *Comprehensive Stevens HydraProbe User's Manual.* https://www.stevenswater.com/resources/documentation/hydraprobe.

Swayze, G.A. (1997). The Hydrothermal and Structral History of the Cuprite Mining District, Southwester Nevada: An Integrated Geological and Geophysical Approach (Publication No. 9725795) [Doctoral dissertation, University of Colorado at Boulder]. ProQuest Dissertations & Theses Global.

Swift Jr, L. W. (1976). Algorithm for Solar Radiation on Mountain Slopes. *Water Resources Research.* 12(1):108-112. https://doi.org/10.1029/WR012i001p00108

Tarboton, D. T. G. (2008, September). *Terrain Analysis Using Digital Elevation Models (TAUDEM)*. David Tarboton: Hydrology Research Group-Terrain Analysis Using Digital Elevation Models (Taudem). https://hydrology.usu.edu/taudem/taudem3.1/.

TOSHIBA Electronic Devices & Storage Corporation. (2018). TOSHIBA CCD Linear Image Sensor [datasheet]. TOSHIBA Electronic Devices & Storage Corporation. https://toshiba.semicon-storage.com/info/TCD1304DG_Web_Datasheet_en_20190108.pdf?did=13709&prodName=TCD1304DG

Tsakiridis, N., Keramaris, K., Theocharis, J., and Zalidis, G. (2020). Simultaneous Prediction of Soil Properties for VNIR-SWIR Spectra using a Localized Multi-Channel 1-D Convolutional Neural Network. *Geoderma.* 367. https://doi.org/10.1016/j.geoderma.2020.114208.

Tschuschke, W., Gogolik, S., Wróżyńska, M., Kroll, M., and Stefanek, P. (2020). The Application of the Seismic Cone Penetration Test (SCPTU) in Tailings Water Conditions Monitoring. *Water*. 12(737). http://doi.org/10.3390/w12030737.

United Nations Development Programme (UNDP) and UN Environment. (2018). Managing Mining for Sustainable Development: A Sourcebook. Bangkok: United Nations Development Programme.

United States Department of Agriculture. (1975). Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys. Washington, D.C.: Government Printing Office.

Uusitalo J., Ala-Ilomäki J., Lindeman H., Toivio J., and Siren M. (2019). Modelling Soil Moisture–Soil Strength Relationship of Fine-Grained Upland Forest Soils. *Silva Fennica*. 53(1). https://doi.org/10.14214/sf.10050

[dataset] U.S. Geological Survey (USGS), 2022. USGS 1/3 Arc Second n41w106 20221118: U.S. Geological Survey.

van der Meer, F. (2004). Analysis of Spectral Absorption Features in Hyperspectral Imagery. *International Journal of Applied Earth Observation and Geoinformation*. 5(1), 55-68. https://doi.org/10.1016/j.jag.2003.09.001

van Genuchten, M.Th., F.J. Leij, and S.R. Yates. (1991). The RETC Code for Quantifying the Hydraulic Functions of Unsaturated Soils. U.S. Salinity Lab, U.S. Dept. of Agriculture, Agricultural Research Service. Riverside, CA.

van Rossum, G., and Drake, F. L. (2009). Python 3 Reference Manual. Scotts Valley, CA: CreateSpace.

Ventola, A., Dolling, and R., Hryciw, R. (2020). Image Analysis and Hardware Development for the Vision Cone Penetrometer (VisCPT). Geo-Congress 2020. https://doi.org/10.1061/9780784482803.068.

Veres, M., Lacey, G., and Taylor, G.W. (2015). *Deep Learning Architectures for Soil Property Prediction*. 12th Conference on Computer and Robot Vision. https://doi.org/10.1109/CRV.2015.15.

Vick, S. G. (1990). Planning, Design, and Analysis of Tailings Dams [B]. http://doi.org/10.14288/1.0394902.

Waiser, T., Morrgan, C., Brown, D., Hallmark, B. (2007). In Situ Characterization of Soil Clay Content with Visible Near-Infrared Diffuse Reflectance Spectroscopy. *Soil Science Society of America.* 71(2):398-396. https://doi.org/10.2136/sssaj2006.0211

Waruru, B., Shepherd, K., Ndegwa, G., Kamoni, P., and Sila, A. (2014). Rapid Estimation of Soil Engineering Properties using Diffuse Reflectance Near Infrared Spectroscopy. *Biosystems Engineering.* 121. https://doi.org/10.1016/j.biosystemseng.2014.03.003.

Werner, T., Mudd, G., Schipper, A., Huijbregts, M., Taneja, L., and Northey, S. (2020). Global-Scale Remote Sensing of Mine Areas and Analysis of Factors Explaining their Extent. *Global Environmental Change*. 60. https://doi.org/10.1016/j.gloenvcha.2019.102007.

Western, A., Zhou, S., Grayson, R., McMahon, T., Bloschl, G., and Wilson., D. (2004). Spatial Correlation of Soil Moisture in Small Catchments and its Relationship to Dominant Spatial Hydrological Processes. *Journal of Hydrology*. 256(1-4):113-134. https://doi.org/10.1016/j.jhydrol.2003.09.014.

Williams, D. (2021). Lessons from Tailings Dam Failures – Where to Go from Here?. *Minerals*. 11(8). https://doi.org/10.3390/min11080853.

Winkens, C., Sattler, F., and Palus, D. (2017). *Hyperspectral Terrain Classification for Ground Vehicles.* Proceedings of the 12^{th} International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. https://doi.org/ 10.5220/0006275404170424.

Wong, J.Y. (2010). Terramechanics and Off-Road Vehicle Engineering. 2^{nd} Edition. Elsiver, Oxford, U.K.

Wride, C., Hofmann, D., Sego, D., Plews, H.D., Konrad, J.M., Biggar, K.W., Robertson, P.K., and Monahad, P.A. (2000a). Ground Sampling Program at the CANLEX Test Sites. *Canadian Geotechnical Journal*. 37(3):530–542. https://doi.org/10.1139/t00-045.

Wride, C., Robertson, P., Biggar, K., Campanella, R.G., Hofmann, B.A., Hughes, J.M., Kupper, A., and Woeller, D.J. (2000b). Interpretation of In Situ Test Results from the CANLEX Sites. *Canadian Geotechnical Journal*. 37(3)1505–529. https://doi.org/10.1139/t00-044.

Wulfsohn, D., Adams, B., and Fredlund, D. (1996). Application of Unsaturated Soil Mechanics for Agricultural Conditions. *Canadian Agricultural Engineering*. 38(3):173-181. Retrieved from: https://library.csbe-scgab.ca/.

Yu, B., Liu, G., Liu, Q., Feng, J., Wang, X., Han, G., and Haung, C. (2018). Effects of Micro-Topography and Vegetation Type on Soil Moisture in a Large Gully on the Loess Plateau of China. *Hydrology Research*. 46 (4):1255-1270. https://doi.org/10.2166/nh.2017.023.

Zhu, X., Vondrick, C., Fowlkes, C., and Remanan, D. (2016). Do We Need More Training Data?. *International Journal of Computer Vision.* 119:76-92. https://doi.org/10.1007/s11263-015-0812-2.

Zhu, Y., Weindorf, D., Chakraborty, S., Beatrix, H., Johnson, S., and Bakr, N. (2010). Characterizing Surface Soil Water Within Field Portable Diffuse Reflectance Spectroscopy. *Journal of Hydrology*. 291(1-2):133-140. https://doi.org/10.1016/j.jhydrol.2010.07.014.

APPENDIX

A1. Supplemental Figures for Chapter 1.

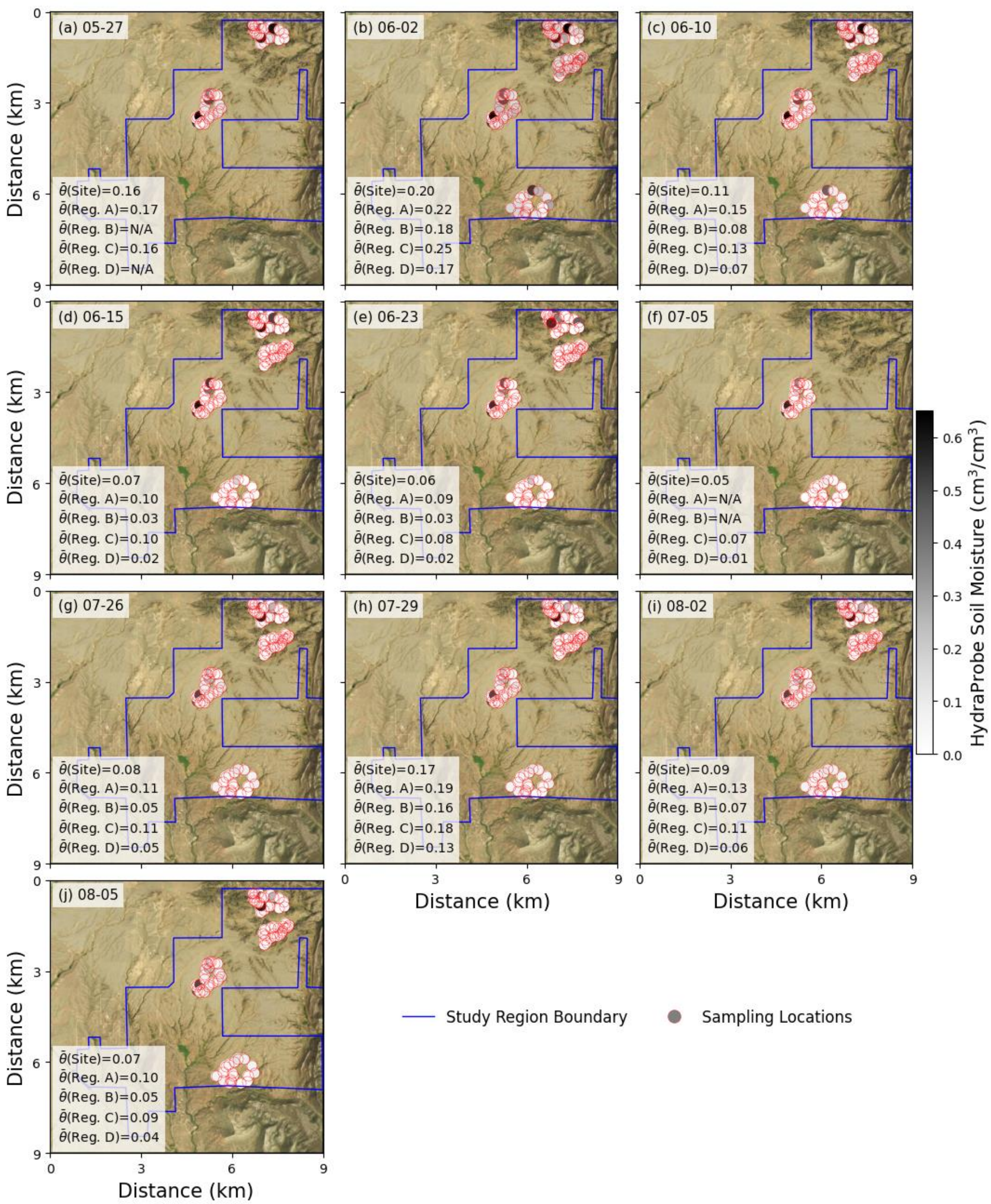

Figure A.1 Maps of soil moisture for each sampling date. Circles represent sampling locations and infill colors represent measured soil moisture. Spatial average soil moistures ($\bar{\theta}$) (cm^3/cm^3) for the entire study site and each study region are shown in the subplots.

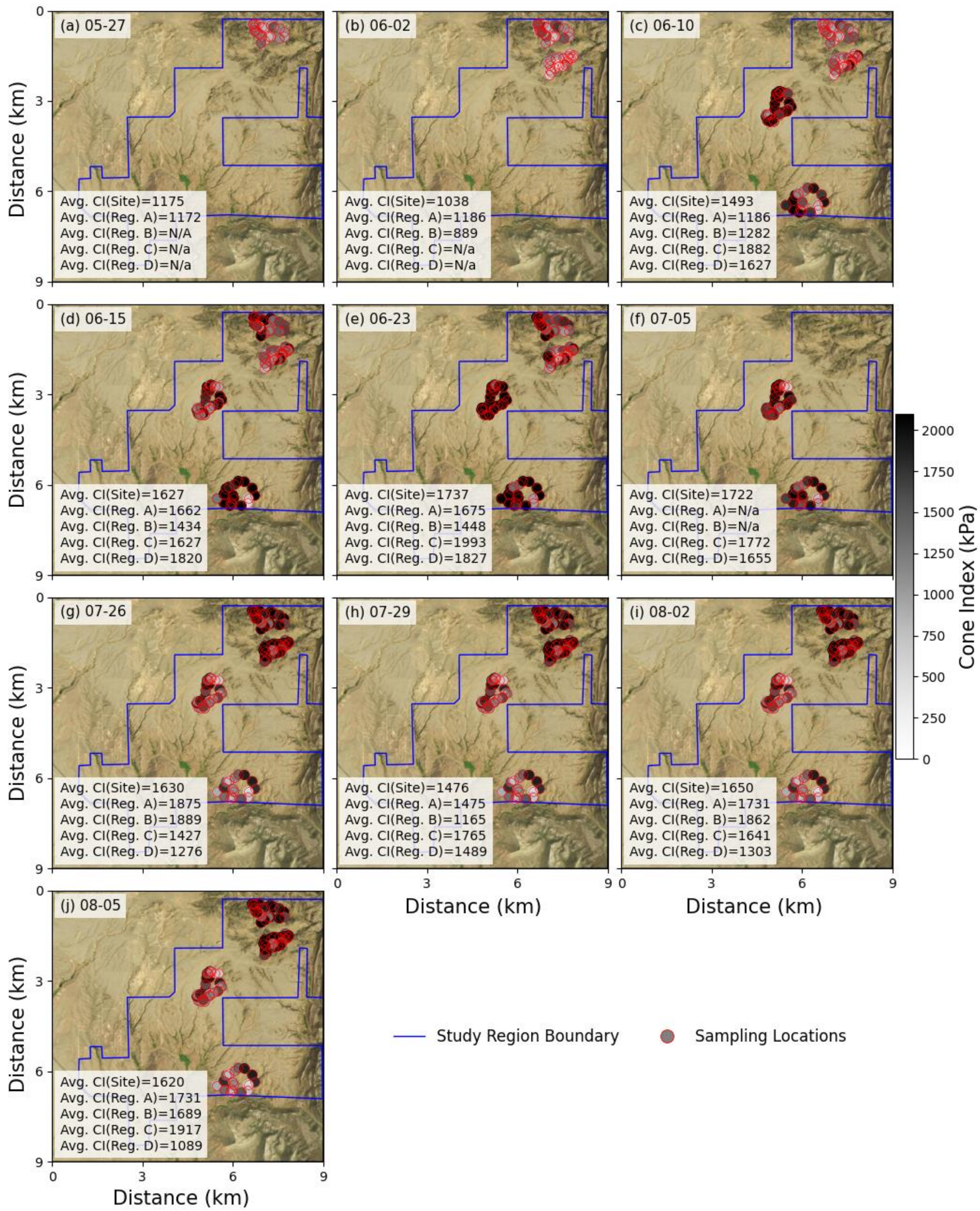

Figure A.2. Maps of soil strength for each sampling date. Circles represent sampling locations and infill color represents measured cone index (CI). Average cone indices for the entire study site and each study region are shown within the subplots.

Figure A.3. Photo of cone penetrometer during insertion into the surficial soil as a proxy measurement for surficial soil strength. From Bindner (2020).

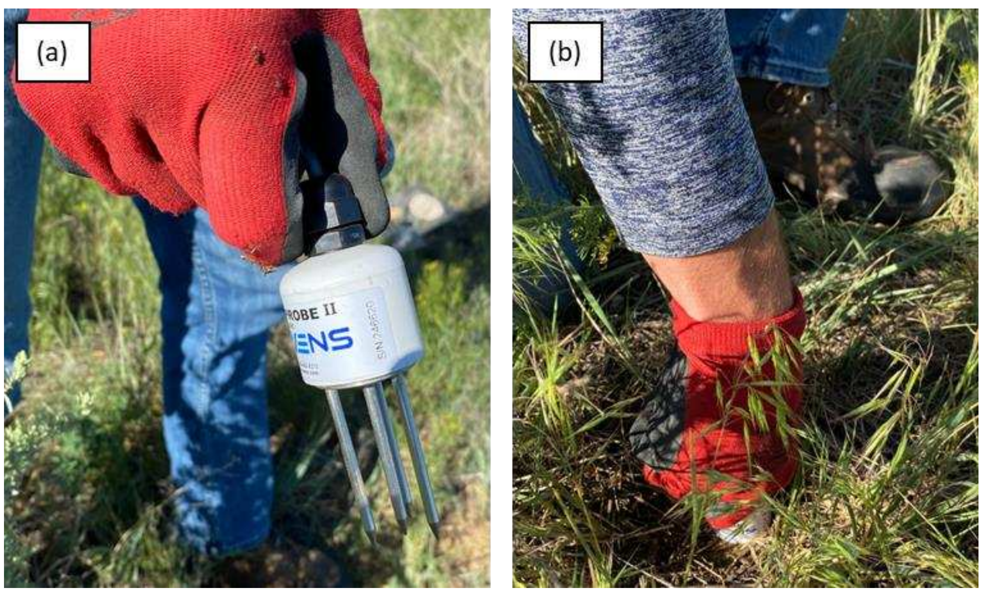

Figure A.4. Photos of (a) HydraProbe soil moisture sensor used to measure surficial soil moisture to a depth of 5 cm and (b) HydraProbe soil moisture sensor inserted into the soil.

A3. Supplemental Figures for Chapter 3

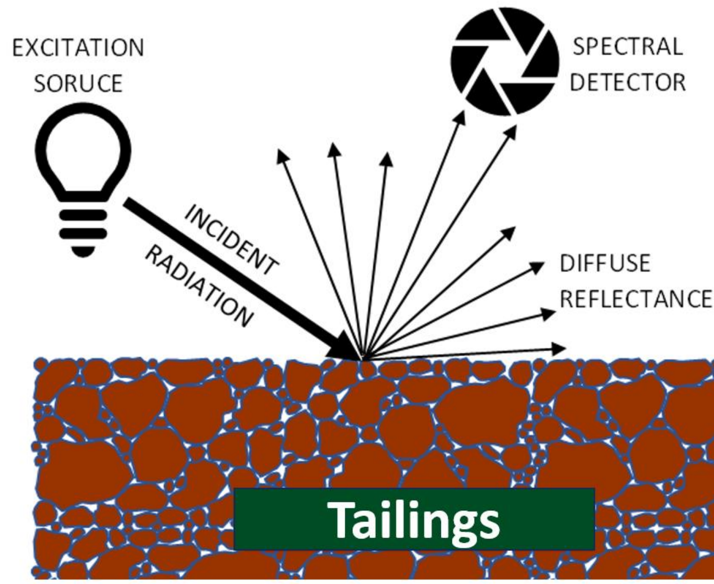

Figure A.5. Simplified illustration of primary interactions and equipment involved in hyperspectral sensing (also known as diffuse reflectance spectroscopy (DRS)).

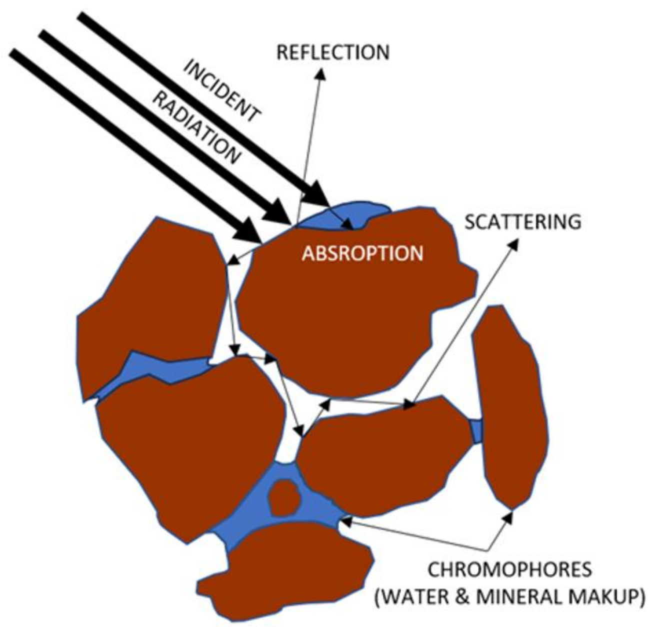

Figure A.6. Simplified illustration of interactions between incident radiation and soil chromophores (viz. material properties), resulting in radiation being either reflected, absorbed, or scattered, in this example.

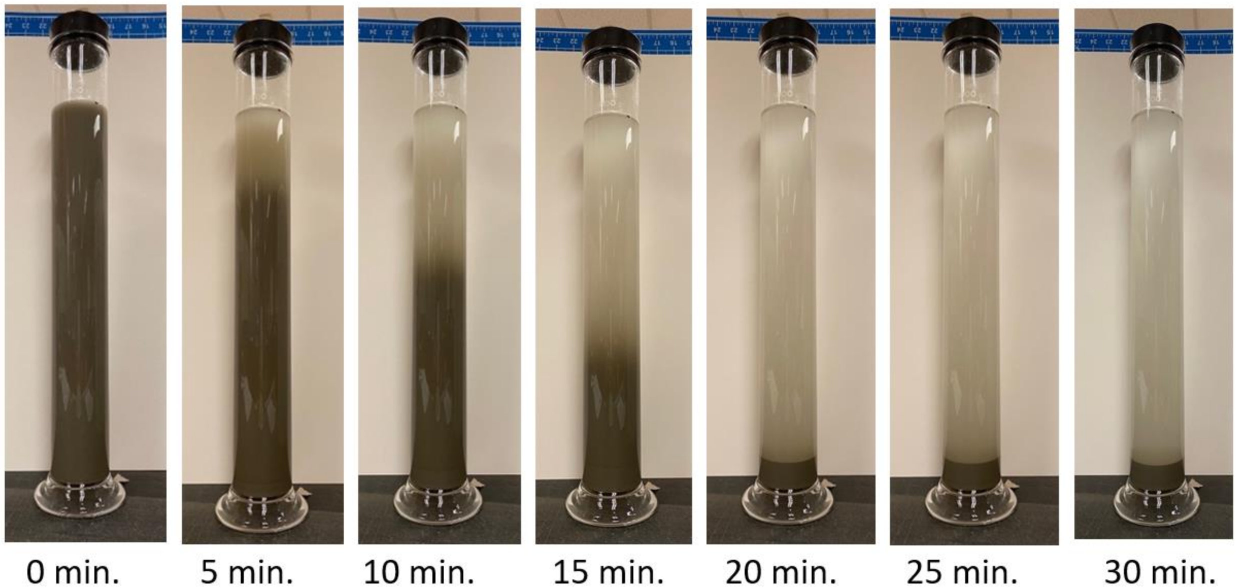

Figure A.7. Whole tailings sample used for sedimentation analyses illustrating the settling of most particles within the first 30 minutes after agitation.

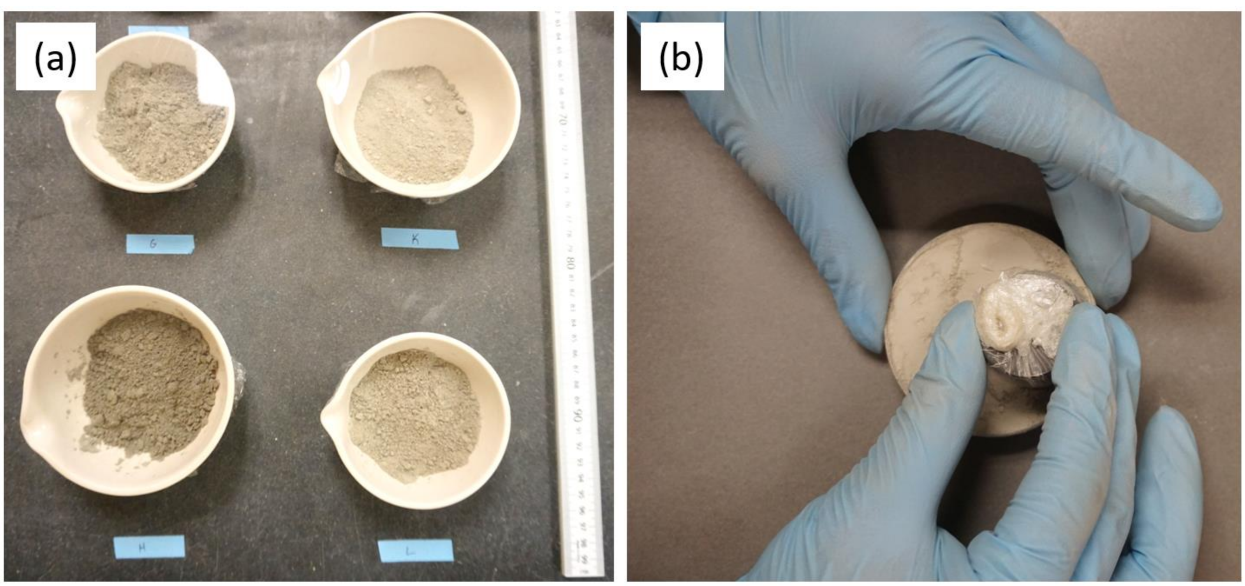

Figure A.8. (a) Test specimens during moisture equilibration procedure contained in ceramic dishes sealed with plastic wrap and (b) photo of moisture equilibrated test specimen during compaction into petri dish (compacted in three lifts) to achieve target density.

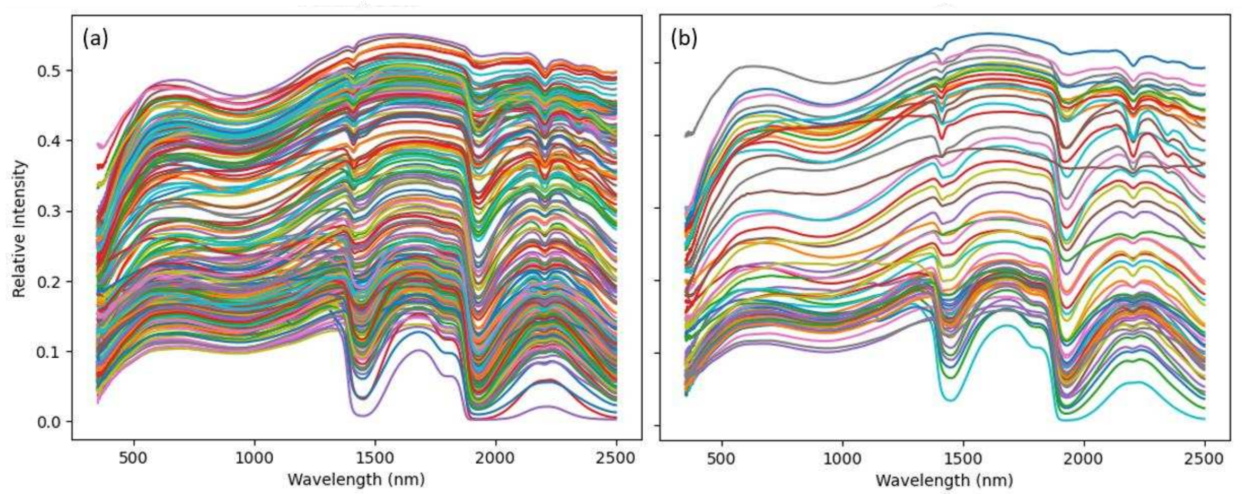

Figure A.9. Hyperspectral signals in (a) the training dataset (n = 240) and (b) the testing dataset (n = 60) from conditioned Latin Hypercube Sampling (cLHS) based on hyperspectral data. The range and general distribution density of hyperspectral signals in both the training and testing are visually similar, indicating the cLHS algorithm successfully replicated the multivariate distribution of hyperspectral data in both datasets.

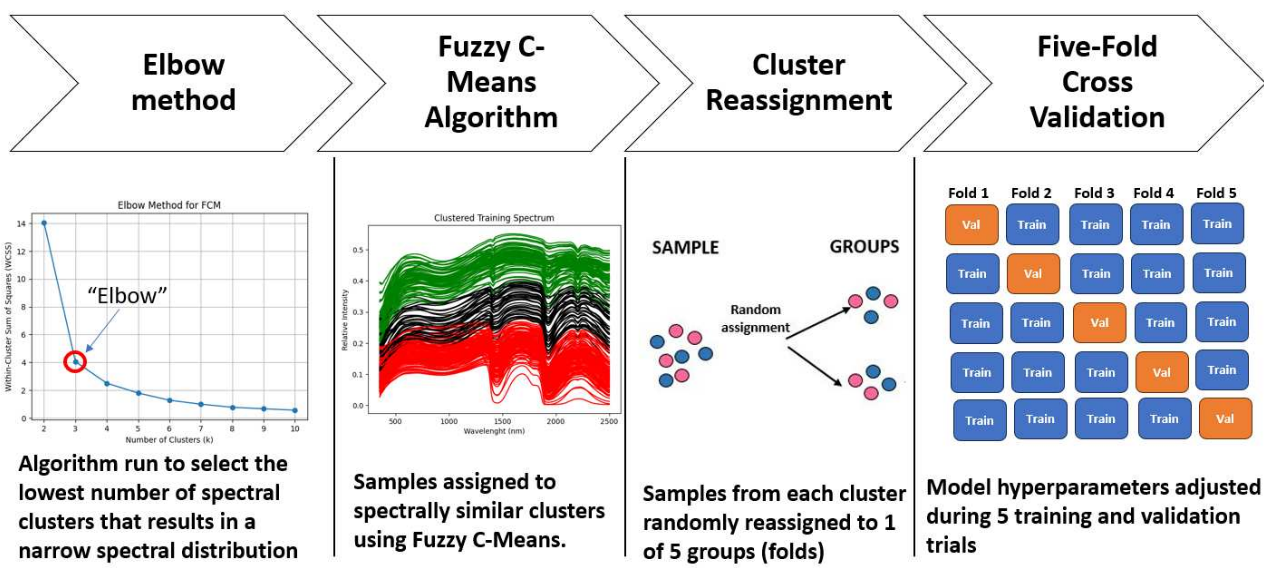

Figure A.10. Simplified flowchart illustrating the Elbow Method results, the Fuzzy C-Means Algorithm results, graphic illustrating cluster reassignment, and graphic illustrating the general procedure associated with five-fold cross validation.

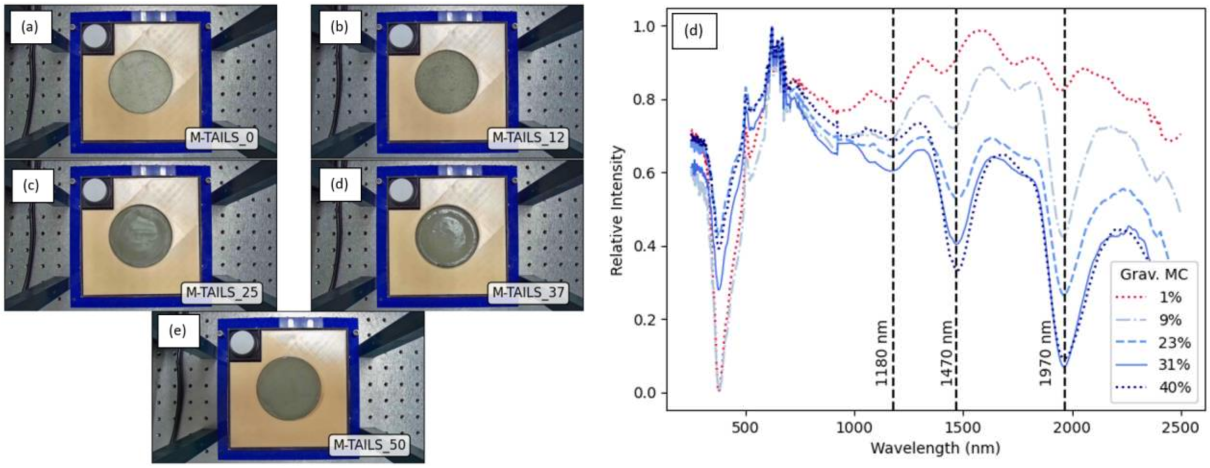

Figure A.11. Photographs of tailings from a gold mine (M-Tailings; different than tailings studied in Chapter 3) prepared to various target gravametric moisure contents including (a) 0%, (b) 12%, (c) 25%, (d) 37%, and (e) 50% and later used in hyperspectral sensing. (d) Hyperspectral signals of the M-Tailings at different moisutre contents and water absorption bands located near 1180nm, 1470 nm, and 1970 nm. Hyperspectral data were captured using a custom laboratory spectrometer (different than spectrometers discussed in Chapter 2 and Chapter 3). This preliminary analysis was completed to better understand spectral signal of tailings at various gravametic mosture contents. Density was not controlled or measured in this demonstration.

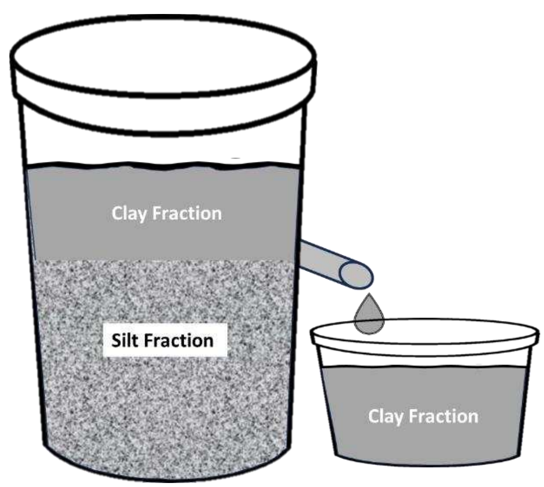

Figure A.12. Simplified schematic of decantation process for separation of tailings silt and clay sized particles.

Figure A.13. Photographs of (a) wet sieving setup used to achieve whole tailings sub-samples with varying sand contents and (b) buckets used to mix the fine fraction of tailings (viz. passing No. 200 sieve) with water and allowed to settle for general separation of silt and clay sized particles. After the settling period, the bucket valves were opened, and clay dominated portions were drained. Both clay and silt dominated portions were then dried in forced-air ovens at 110°C.

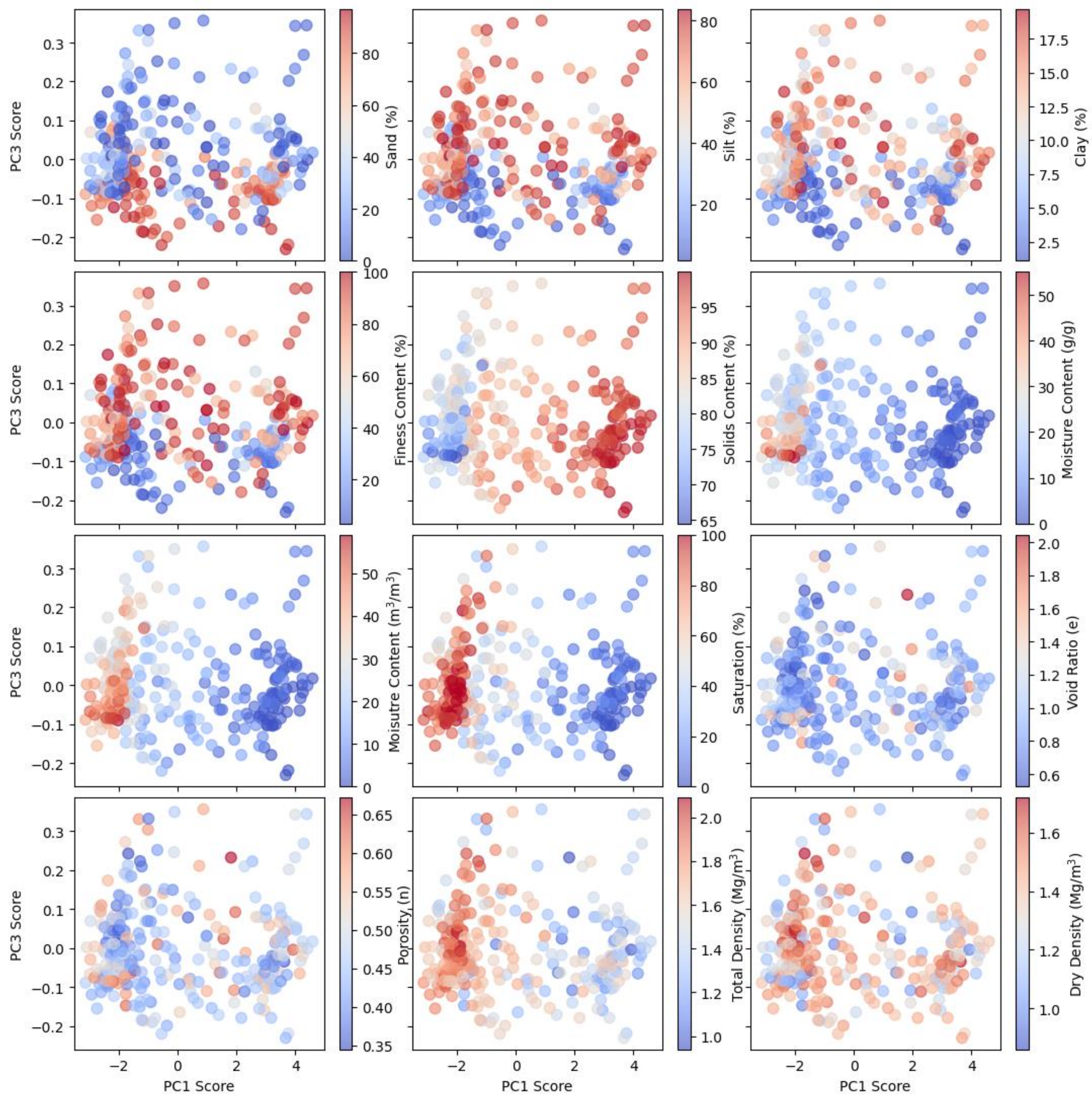

Figure A.14. Tailings samples plotted in the principal component 1 vs. principal component 3 space for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density. Axis values correspond to the principal component score and point infill colors represent the property value for each test specimen.

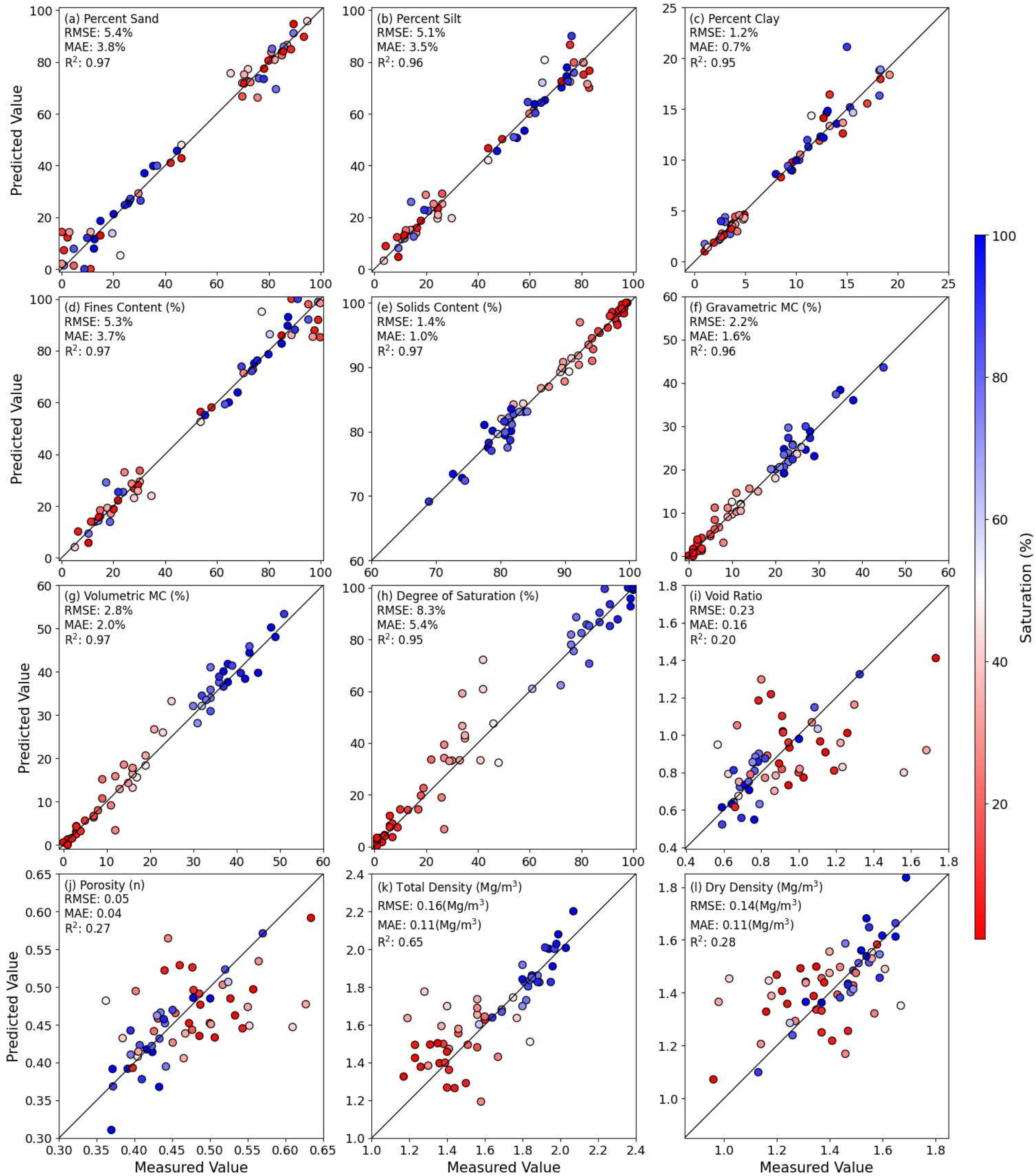

Figure A.15. Predictions values from the CNN with infill colors displaying the degree of saturation for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

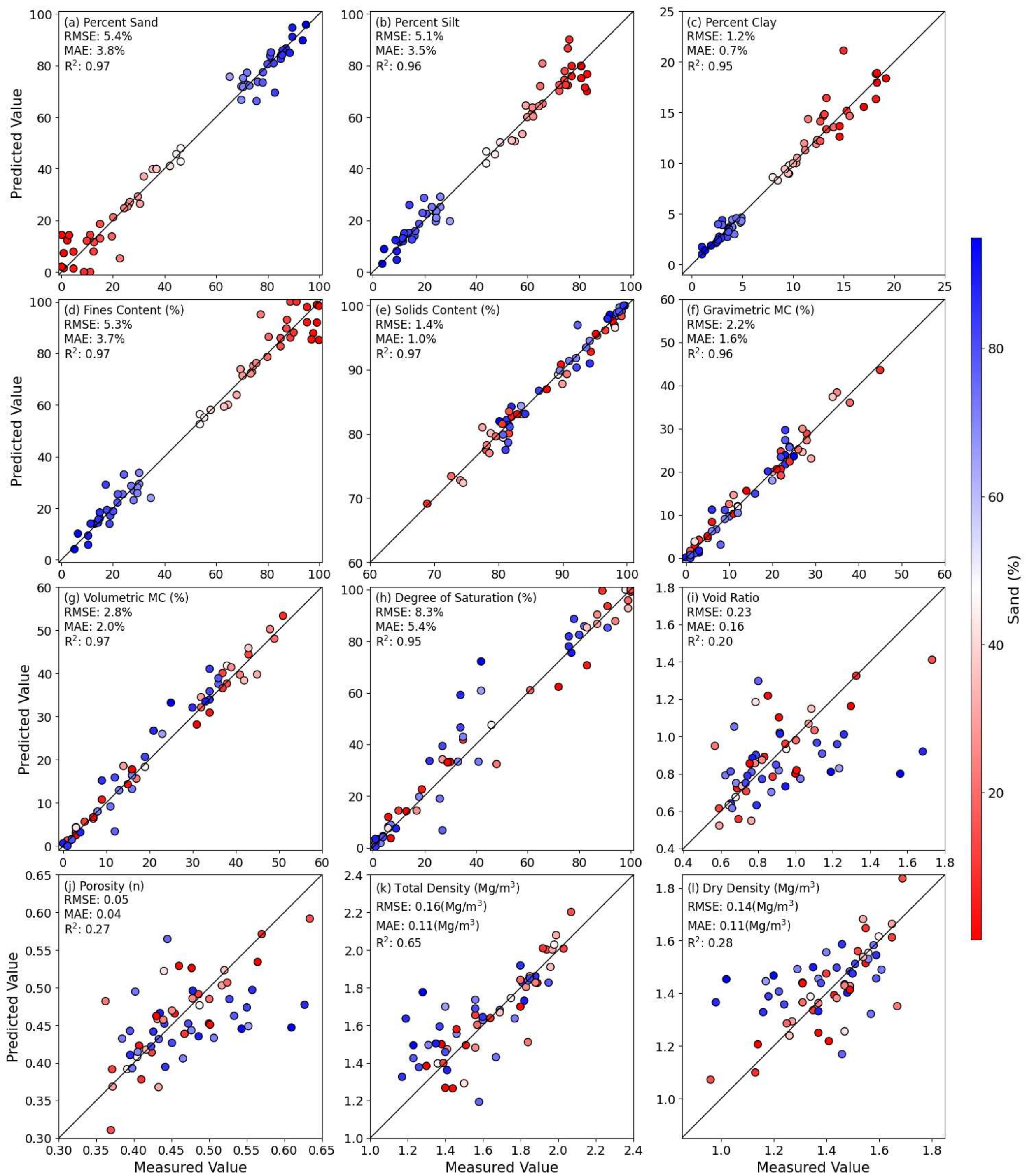

Figure A.16. Predictions values from the CNN with infill colors displaying the percent sand for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

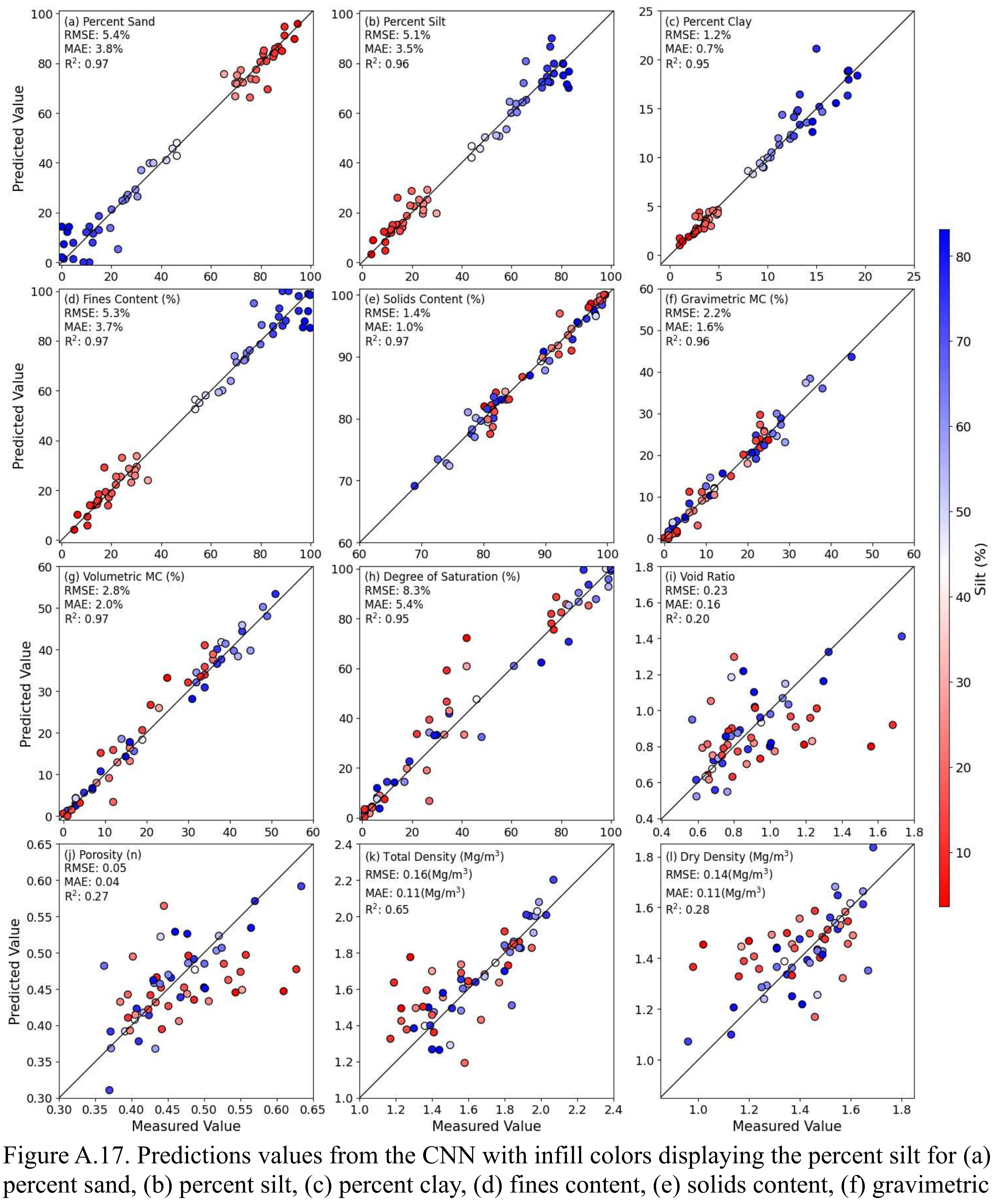

Figure A.17. Predictions values from the CNN with infill colors displaying the percent silt for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

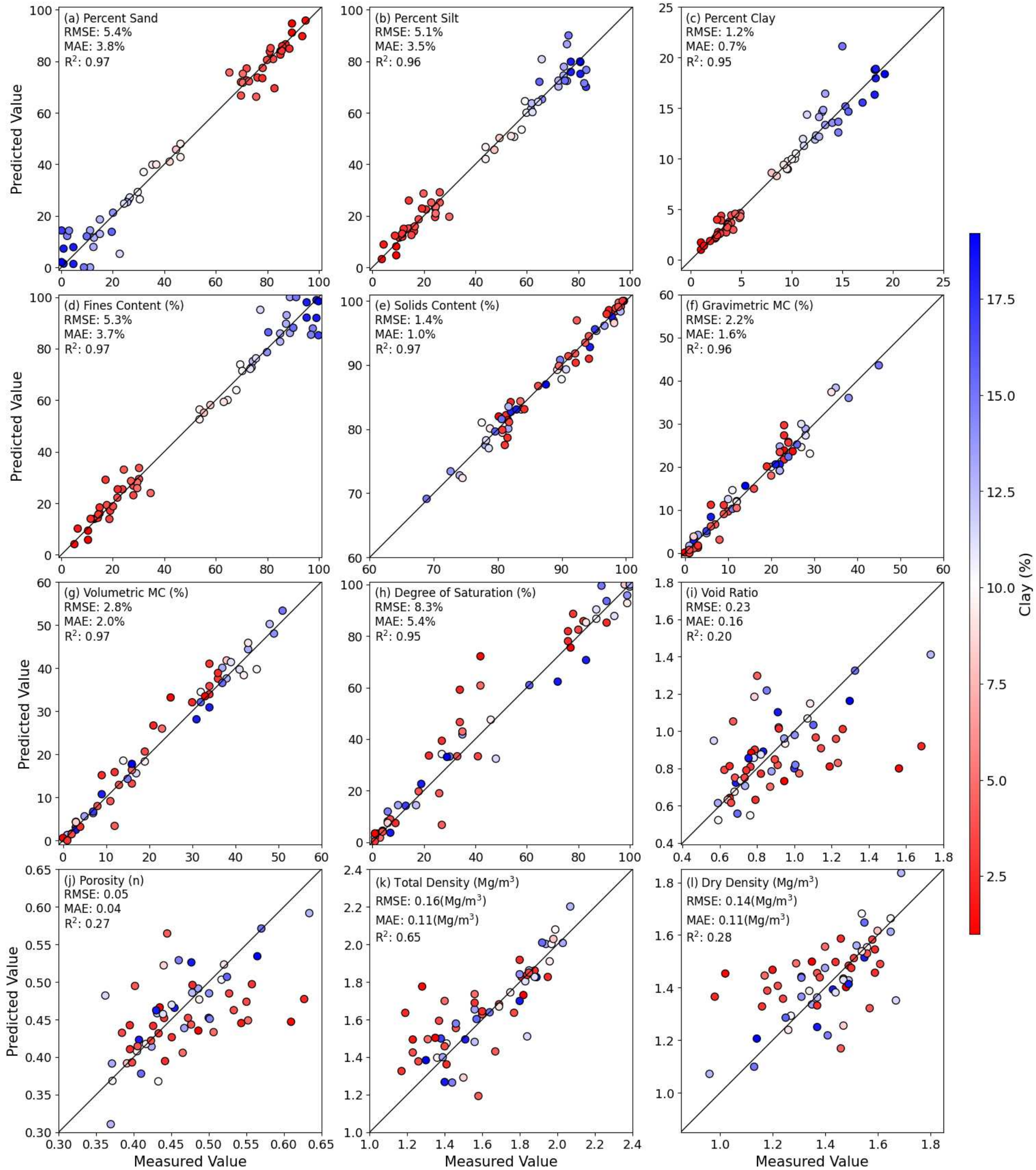

Figure A.18 Predictions values from the CNN with infill colors displaying the percent clay for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

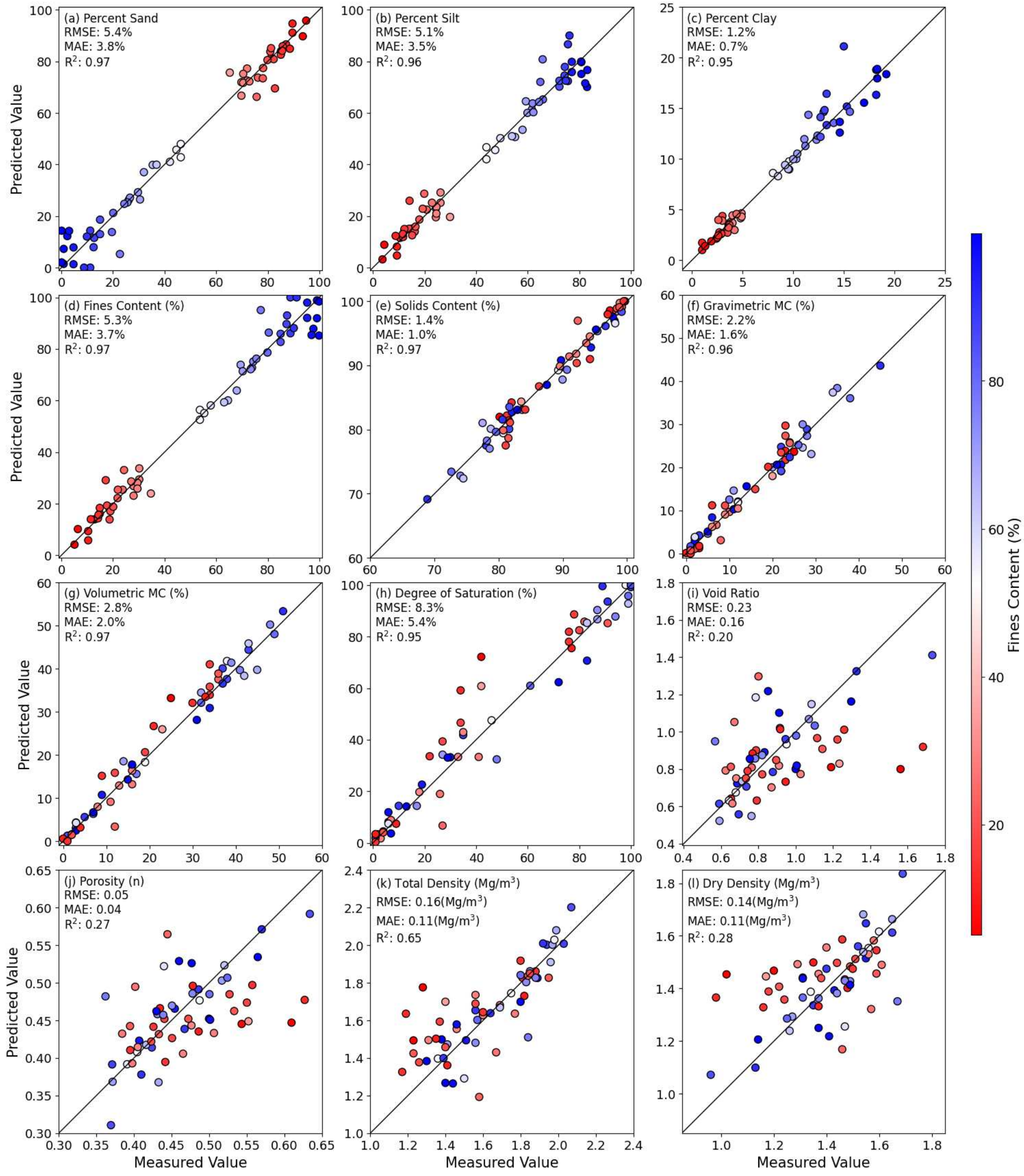

Figure A.19. Predictions values from the CNN with infill colors displaying the fines content for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

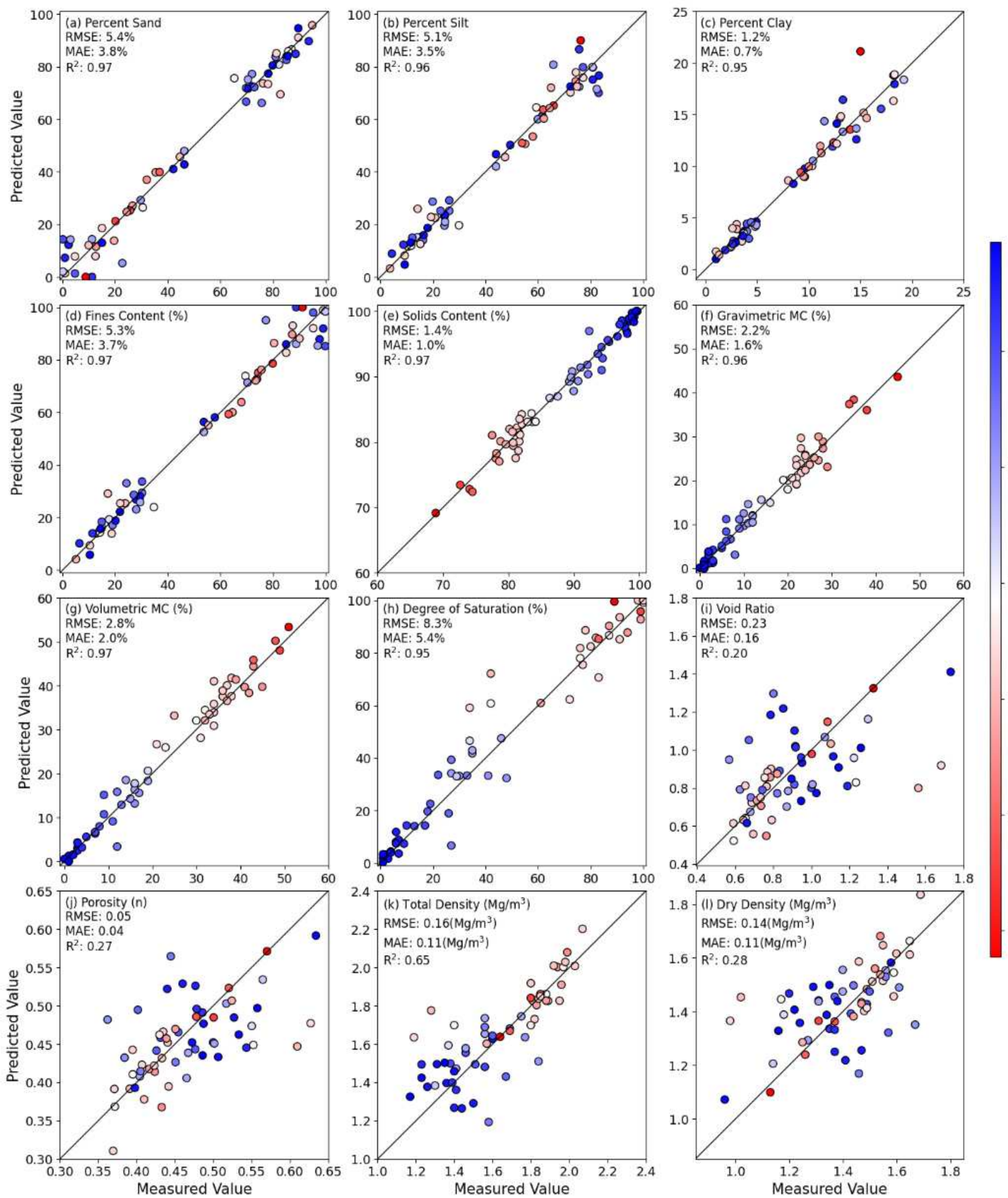

Figure A.20. Predictions values from the CNN with infill colors displaying the solids content for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

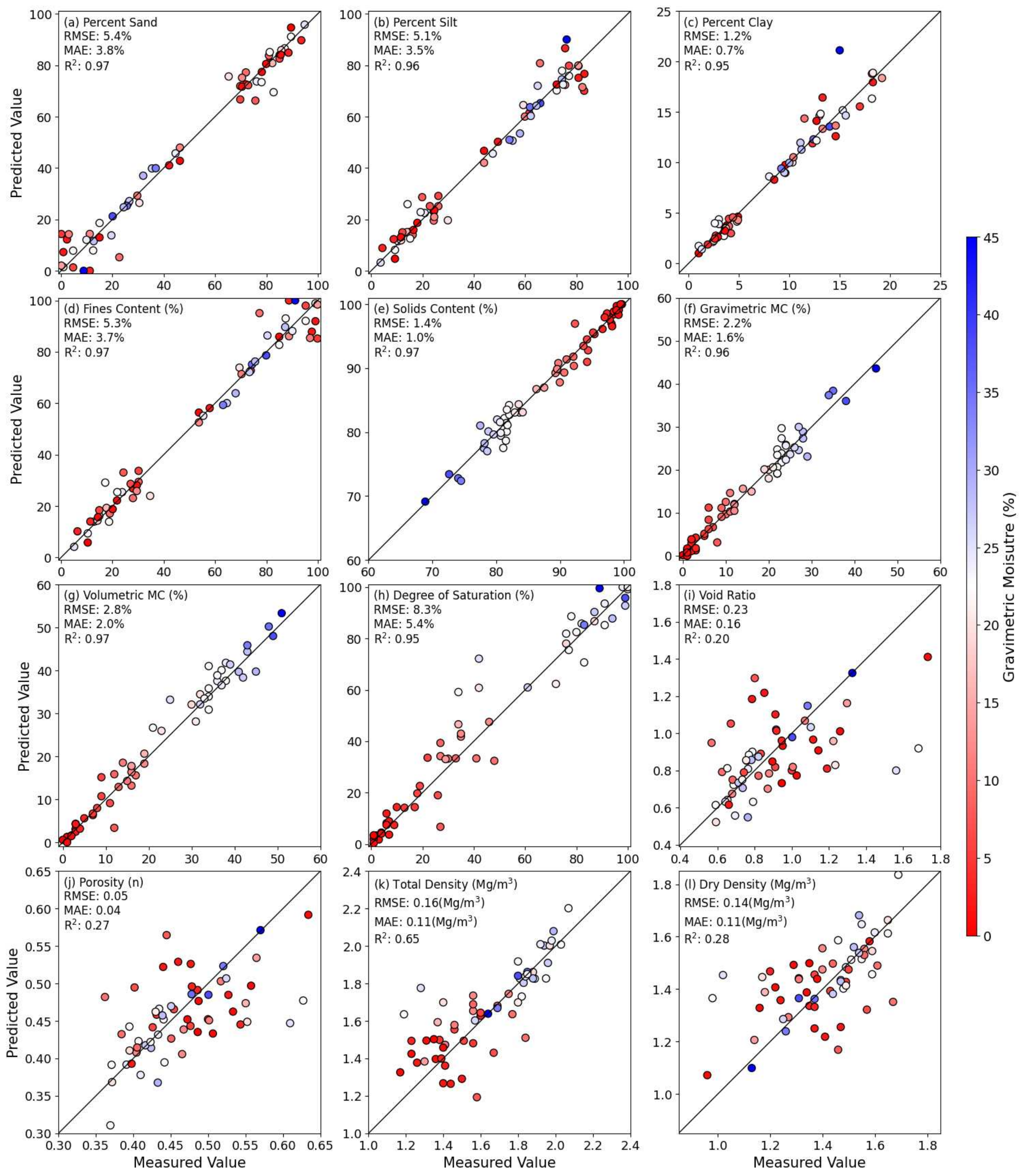

Figure A.21. Predictions values from the CNN with infill colors displaying the gravimetric moisture content for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

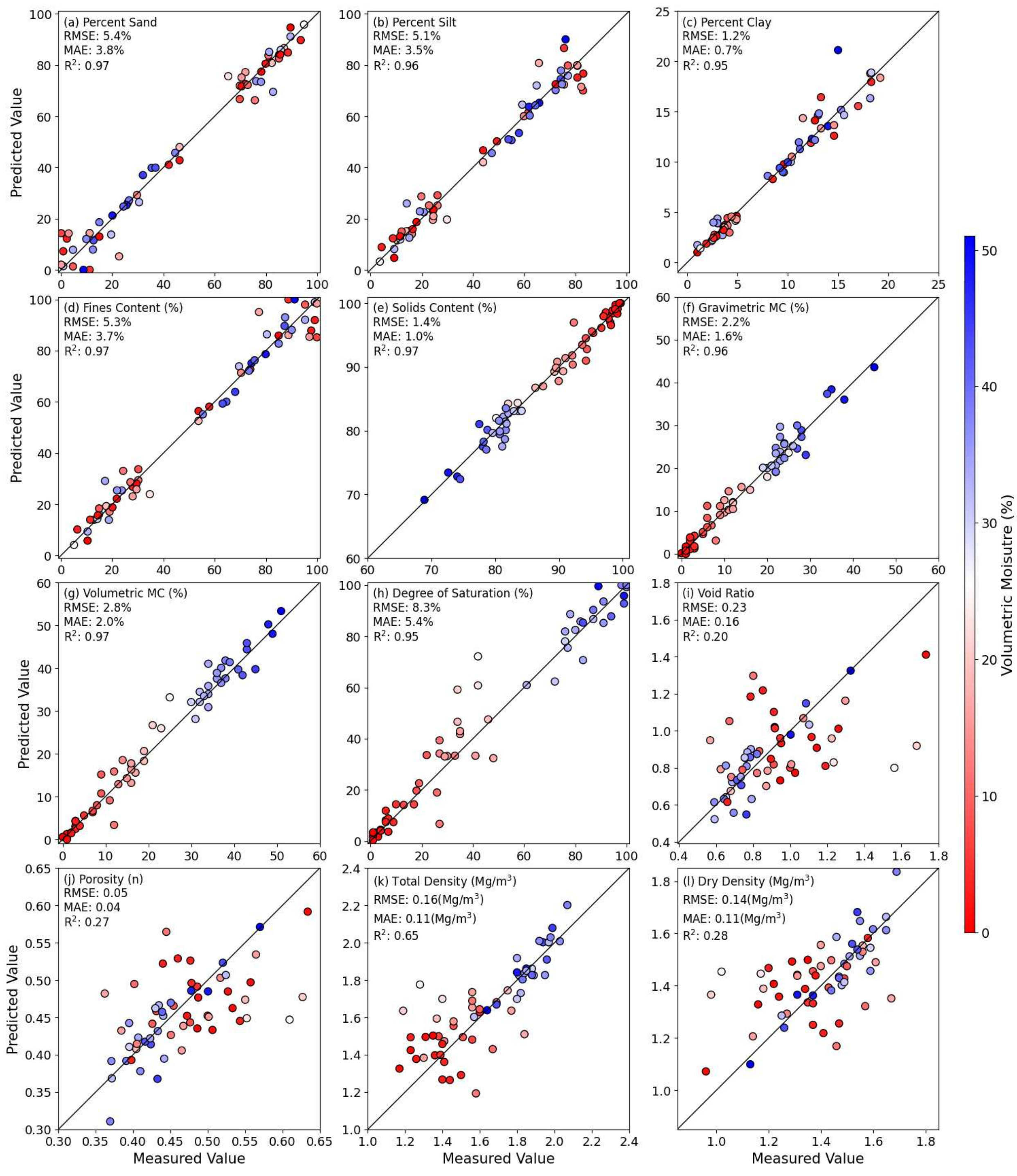

Figure A.22. Predictions values from the CNN with infill colors displaying the volumetric moisture content for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

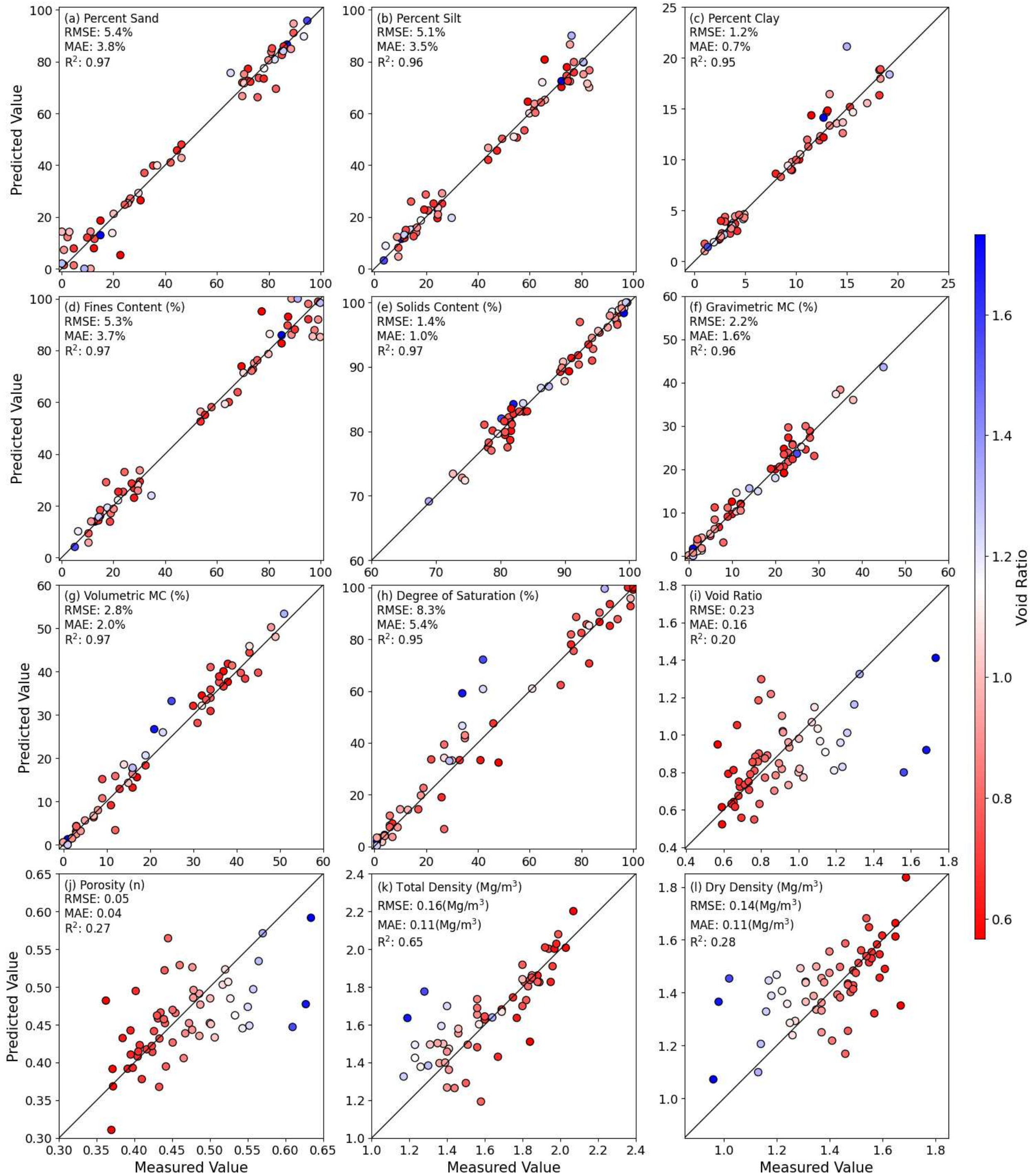

Figure A.23. Predictions values from the CNN with infill colors displaying the void ratio for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

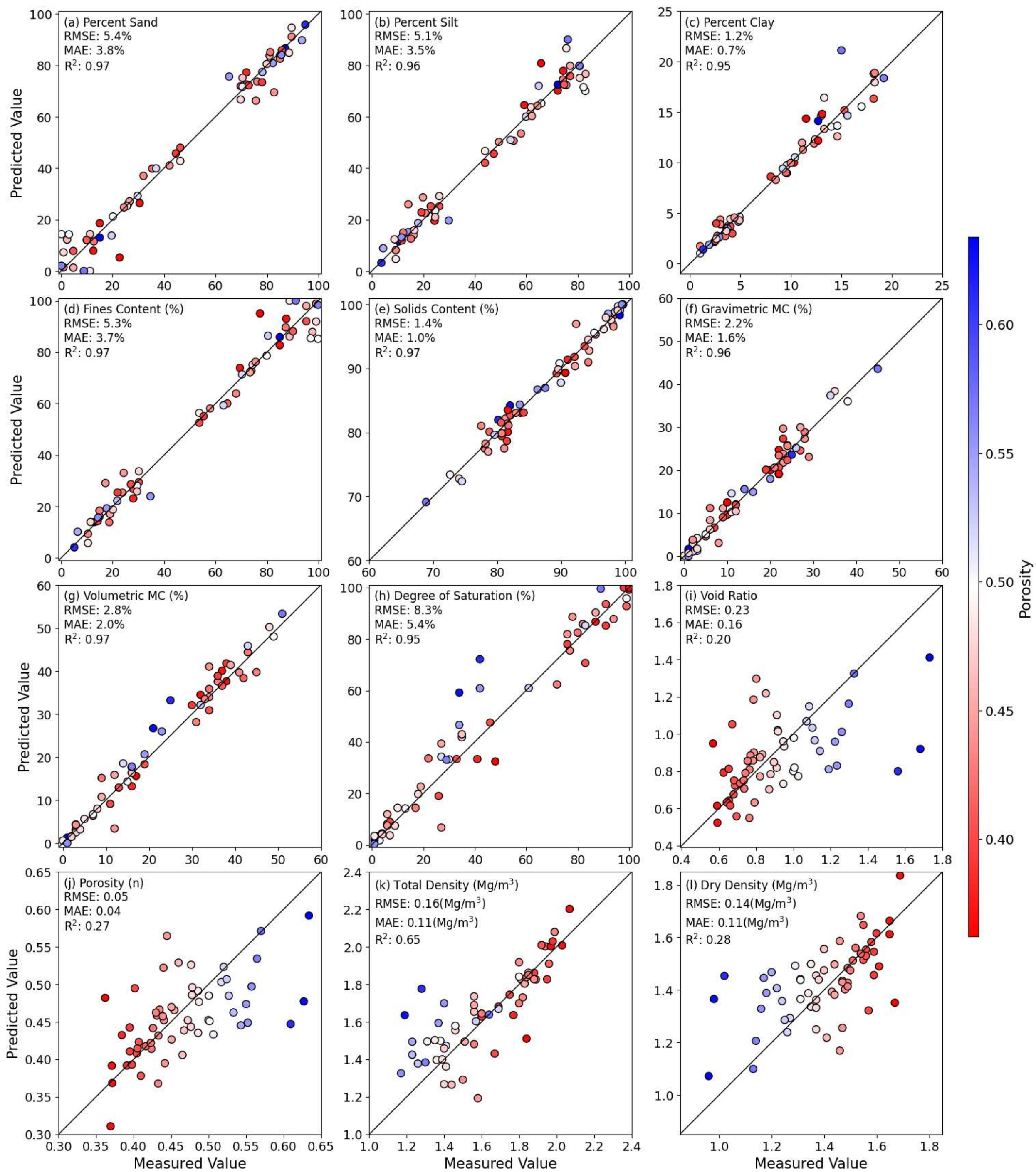

Figure A.24. Predictions values from the CNN with infill colors displaying the porosity for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

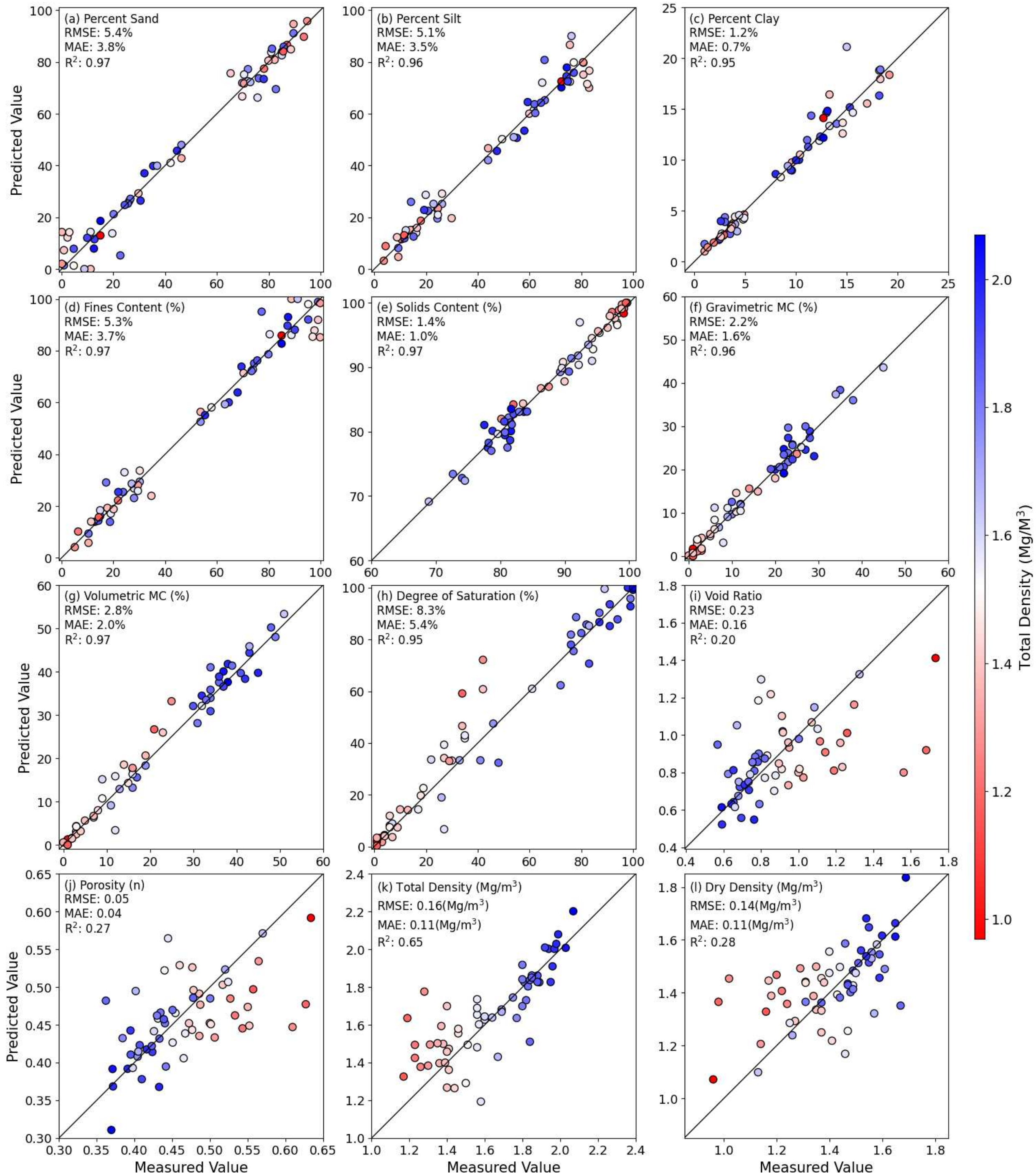

Figure A.25. Predictions values from the CNN with infill colors displaying the total density for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

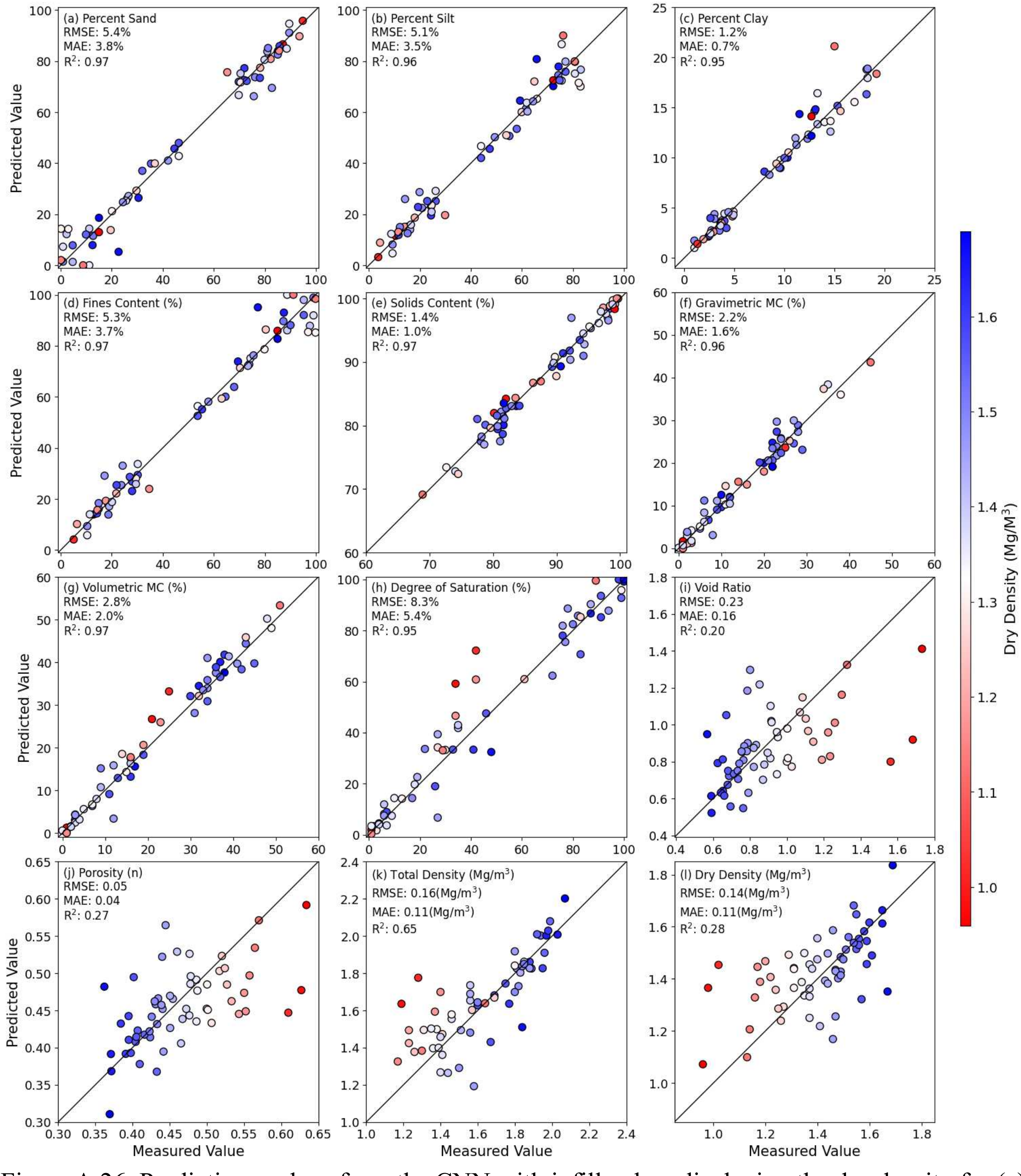

Figure A.26. Predictions values from the CNN with infill colors displaying the dry density for (a) percent sand, (b) percent silt, (c) percent clay, (d) fines content, (e) solids content, (f) gravimetric moisture content, (g) volumetric moisture content, (h) saturation, (i) void ratio, (j) porosity, (k) total density, and (l) dry density.

LIST OF ABBREVIATIONS

AMSR-E	Advanced Microwave Scanning Radiometer-EOS	DEM	Digital Elevation Model
ASTM	American Society for Testing and Materials	DRS	Diffuse Reflectance Spectroscopy
CDF	Cumulative Distribution Function	DSM	Digital Soil Map
CH	High Plasticity Clay	DWT	Discrete Wavelet Transform
CI	Cone Index	ELU	Exponential Linear Unit
CL	Low Plasticity Clay	EMT+VS	Equilibrium Moisture from Topography, Vegetation, and Soil
cLHS	Conditioned Latin Hypercube Sampling	ESA	European Space Agency
cm	Centimeter	EVI	Enhanced Vegetation Index
CNN	Convolutional Neural Network	FCM	Fuzzy c-Means Algorithm
CPT	Cone Penetration Test	GISTM	Global Industry Standard on Tailings Management
C_u	Coefficient of Uniformity	G_s	Specific gravity of Solids
D_{10}	Diameter at 10 percent mass finer	GTR	Global Tailings Review
D_{30}	Diameter at 30 percent mass finer	ha	Hectare
D_{50}	Diameter at 50 percent mass finer	KCB	Klon Crippen Berger
D_{60}	Diameter at 60 percent mass finer	kPa	Kilopascal

K_s	Saturated Hydraulic Conductivity	PC2	Second Principal Component
LL	Liquid Limit	PC3	Third Principal Component
m	Meter	PCA	Principal Component Analysis
MBE	Mean Bias Error	PI	Plasticity Index
Mg/m^3	Megagram per Cubic Meter	PL	Plastic Limit; PI, Plasticity Index
MH	High Plasticity Silt	PLS	Partial Least Squares
ML	Low Plasticity Silt	PSD	Particle Size Distribution
mm	Millimeter	PSRI	Potential Solar Radiation Index
MRLC	Multi-Resolution Land Characteristics Consortium	PTFE	Polytetrafluoroethylene
NDVI	Normalized Difference Vegetation Index	r	Pearson Correlation Coefficient
NRMM	North Atlantic Treaty Organization Reference Mobility Model	COD	Coefficient of Determination
OH	High Plasticity Organic Material	RCI	Rating Cone Index
OL	Low Plasticity Organic Material	RELU	Rectified Linear Unit
p	Pearson Statistical Significance	RI	Remold Index
PC1	First Principal Component	RI	Relative Intensity

RMSE	Root Mean Squared Error	Tanh	Hyperbolic Tangent
SBT	Soil Behavior Type	TAUDEM	Terrain Analysis Using Digital Elevation Model
SCA	Specific Contributing Area	TSF	Tailings Storage Facility
SCPT	Seismic Cone Penetration Test	UN	United Nations
SC-SM	Silty Clayey Sand	UNDP	United Nations Development Programme
SELU	Scaled Exponential Linear Unit	USCS	Unified Soil Classification System
SHAP	Shapley Additive Explanations	USDA	United States Department of Agriculture
SM	Silty Sand	USGS	United States Geologic Survey
SMAP	Soil Moisture Active Passive	VNIR	Visible & Near-Infrared
SMOS	Soil Moisture Ocean Salinity	WCSS	Within-Cluster Sum of Squares
SP	Poorly Graded Sand	w_{opt}	Optimum Water Content
SPT	Standard Penetration Test	θ	Surface soil moisture (cm^3/cm^3)
SSURGO	Soil Survey Geographic Database	λ_2	L2 Kernel Regularization Parameter
STRESS	Strength of Surface Soils	μm	Micro-meter
SVR	Support Vector Regression	ρ_d	Dry Density
SWIR	Short Wave Infrared		