Canadian Geotechnical Journal

Research Article

Ex situ hyperspectral sensing and machine learning for tailings characterization

Joseph Bindner, Christopher Bareither, and Joseph Scalia, IV

Civil and Environmental Engineering, Colorado State University, Fort Collins, CO, US

Corresponding author: **Joseph Bindner** (email: Joseph.Bindner@conetec.com)

Abstract

Understanding tailings properties at high spatial resolutions is needed for many geotechnical analyses. While tailings properties can be characterized using undisturbed samples and in situ tests, the current methods face limitations regarding adequate spatial resolution and comprehensive property assessment. This study explores the use of hyperspectral sensing and convolutional neural networks for the simultaneous prediction of 12 tailings properties, including percent sand, silt, clay, fines content, solids content, gravimetric moisture content, volumetric moisture content, saturation, void ratio, porosity, total density, and dry density. Tailings from a precious metal mine were used to prepare samples with diverse properties and hyperspectral data were captured. The tailings-hyperspectral dataset was then split into training and testing subsets, a convolutional neural network was optimized and trained, and the model's performance was assessed using the testing data. The prediction of particle size distribution metrics and moisture metrics have root mean squared errors below 8% with coefficients of determinations above 0.95. Predictions of void ratio, porosity, total density, and dry density have poorer performance than other properties. However, density predictions have lower errors for samples with high saturation. Results show promise for the rapid characterization of tailings properties using hyperspectral data.

Key words: convolutional neural networks, soil properties, soil spectroscopy, reflectance spectroscopy, tailings properties

Résumé

La compréhension des propriétés des résidus à des résolutions spatiales élevées est nécessaire pour de nombreuses analyses géotechniques. Bien que les propriétés des résidus puissent être caractérisées à l'aide d'échantillons non perturbés et de tests in situ, les méthodes actuelles rencontrent des limitations en termes de résolution spatiale ad. Cette étude explore l'utilisation de la télédétection hyperspectrale et des réseaux neuronaux convolutifs pour la prédiction simultanée de douze propriétés des résidus, incluant le pourcentage de sable, de limon, d'argile, la teneur en fines, la teneur en solides, la teneur en humidité gravimétrique, la teneur en humidité volumétrique, la saturation, le rapport de vide, la porosité, la densité totale et la densité sèche. Les résidus d'une mine de métaux précieux ont été utilisés pour préparer des échantillons aux propriétés variées et des données hyperspectrales ont été capturées. Le jeu de données résidus-hyperspectral a ensuite été divisé en sous-ensembles d'entraînement et de test, un réseau neuronal convolutif a été optimisé et entraîné, et la performance du modèle a été évaluée en utilisant les données de test. La prédiction des mesures de distribution granulométrique et des mesures d'humidité a des erreurs quadratiques moyennes inférieures à 8% avec des coefficients de détermination supérieurs à 0,95. Les prévisions du rapport des vides, de la porosité, de la densité totale et de la densité sèche ont un rendement inférieur à celui des autres propriétés. Cependant, les prédictions de densité présentent des erreurs plus faibles pour les échantillons avec une saturation élevée. Les résultats montrent un potentiel pour la caractérisation rapide des propriétés des résidus à l'aide de données hyperspectrales.

Mots-clés : réseaux neuronaux convolutifs, propriétés du sol, spectroscopie des sols, spectroscopie de réflectance, propriétés des résidus miniers

1. Introduction

Mineral commodities are essential to modern society and are an important part of achieving 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 (TSFs) that 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 to provide guidance on tailings management that outlines 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 that impact 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% (Rodríguez 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 (Jefferies and Been 2016; Morrison 2022). Tailings density is important for estimating hydraulic conductivity, potential water recovery, and TSF storage capacity over time (Agapito and Bareither 2018; 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 that do not preserve the in situ tailings condition and are used for characterizing soil properties that are not a function of volume or fabric, such as PSD 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. 2000*a*, 2000*b*) 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).

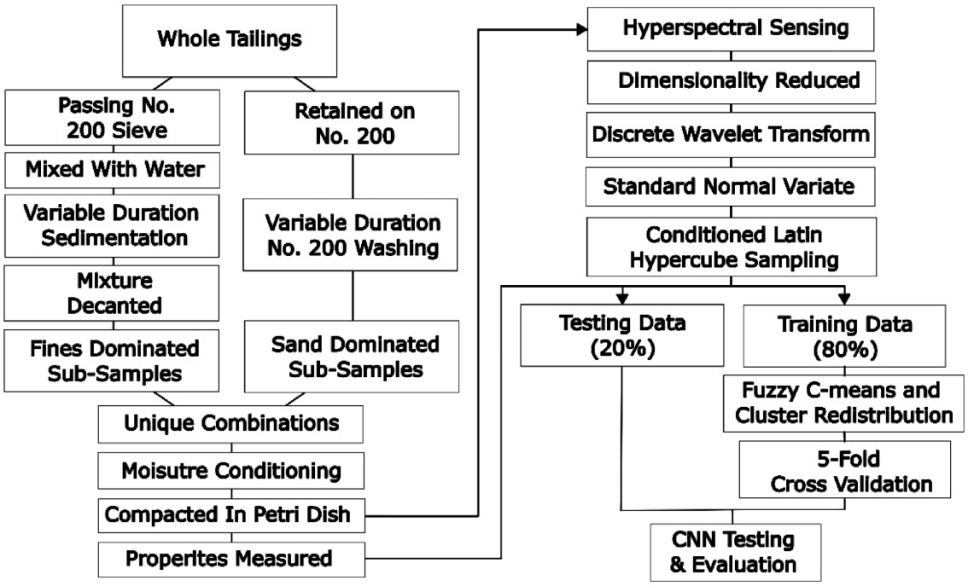
The standard penetration test (SPT) and piezocone penetration test (CPTu) are common in situ methods for characterizing tailings behavior. The SPT requires drilling to depth before the resistance of soil to penetration is measured (ASTM International 2022). The resistance measured by the SPT is reported 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 (Mayne et al. 2009; Morrison 2022). The CPTu collects near 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 (Tschuschke et al. 2020). Data collected from the CPTu are empirically related to soil behavior type (Robertson 2016), which generally correlates with Unified Soil Classification System (USCS) classes (Molle 2005). Interpretations of CPTu 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 CPTu 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 and 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 and 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 regression algorithm (Sjöström et al. 1983), the Cubist algorithm (Quinlan 1992, 1993), the support vector regression algorithm (Drucker et al. 1996), the spectrum-based learner (Ramirez-Lopez et al. 2013), and 1D convolutional neural networks (CNNs) (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 1D 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 was 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

Fig. 1. Flowchart displaying general procedures associated with material processing, test specimen preparation, hyperspectral data processing, and machine learning.



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 compiled into a dataset (Bindner et al. 2025), then separated into training and testing datasets for machine learning. A 1D CNN was cross validated, trained, and tested to assess the use of hyperspectral data 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.

2. Methods and materials

The sample preparation, hyperspectral sensing, and machine learning procedures are outlined in Fig. 1. A whole tailings sample was processed to artificially prepare 300 test specimens with a range of PSDs, moisture conditions, and densities and a spectrometer was used to collect hyperspectral data for each test specimen. The hyperspectral data (Bindner et al. 2025) were 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 Fig. 1 is described in detail in the sections below.

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, percent silt, percent clay-sized particles, fines content, solids content, water content, saturation, void ratio, porosity, and density. The properties of the whole tailings sample are outlined in Table 1. Plasticity testing was used to determine the whole tailings liquid limit and plasticity index (ASTM International 2017*b*). Mechanical sieve analysis was used to determine the PSD for particles larger than 75 μm (ASTM International 2017*c*) and the whole tailings was classified according to the USCS (ASTM International 2017*a*). Sedimentation analysis was performed to determine the distribution of particles ranging in diameter from 2 to 63 μm using the *PARIO Soil Particle Analyzer* (mass fraction detection error of ±0.005 g/g) (METER Group 2017). The standard Proctor compaction test was used to determine the maximum dry density and optimum water content according to ASTM International (2021). Specific gravity of solids (G_s) was determined using the American Society for Testing and Materials (ASTM) standard D854 (ASTM International 2023). Quantitative X-ray diffraction was performed to determine mineralogical composition of the whole tailings.

Sieving, sedimentation, and decantation were used to separate the whole tailings sample into sub-samples with varying PSDs (referred to as tailings sub-samples). The whole tailings sample was first dried, then clods and agglomerates were disaggregated using a pestle and mortar before processing the tailings 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, mixed with water, subjected to variable-duration sedimentation, and then decantated to obtain tailings sub-samples with various silt and clay contents. The PSD for each of the six 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). The PSD of each PSD specimen was determined by recording the mass of each tailings sub-sample

Table 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 (4.75–76.2 mm)	0
	Percent sand (75 μm to 4.75 mm)	46
	Percent fines (<75 μm)	54
ASTM D2487	USCS classification	ML
METER PARIO Plus	Percent silt (75–2 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

added to a PSD specimen, then calculating a weighted average of particle sizes by mass. Although the distribution of particle sizes of each PSD specimen was not directly measured, the calculated PSD is assumed to be generally representative of the measured PSD in this study. The PSD curves for PSD specimens and the whole tailings sample (original sample) are shown in Fig. 2. 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.

Each PSD specimen was prepared to three saturation and density conditions, resulting in a total of 300 test specimens. Deionized water was added to each test specimen to achieve a target saturation, and the moisture was allowed to equilibrate for a minimum of 12 h 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% to 100% and the target dry density was randomly distributed among test specimens ranging between 0.9 and 1.7 Mg/m^3.

The frequency distributions of measured or calculated properties among prepared test specimens are shown in Fig. 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% to 99%, 2% to 84%, 1% to 20%, and 3% to 100%, respectively (Figs. 3*a*–3*d*). Solids content ranged from 65% to 100% (Fig. 3*e*), covering a portion of the tailings continuum including thickened, paste, and filtered tailings (KCB 2017). Gravimetric and volumetric moisture content ranged from 0% to 55% and 0% to 59%, respectively (Figs. 3*f* and 3*g*) and saturation ranged from 0% to 100% (Fig. 3*h*). Test specimen void ratios ranged from 0.5 to 2.0 (Fig. 3*i*) and porosity ranged from 0.34 to 0.67 (Fig. 3*j*). For gold, silver, lead, and zinc tailings, studies have reported the in-place void ratio of slimes ranging from 0.6 to 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 to 2.09 Mg/m^3 (Fig. 3*k*) and dry density from 0.86 to 1.72 Mg/m^3 (Fig. 3*l*). Kealy et al. (1974) and Mabes et al. (1977) reported the dry density for lead and zinc slimes to range from 1.28 to 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 were prepared to simulate conveyor deposition of filtered tailings.

2.2. Hyperspectral sensing

Hyperspectral scans were captured immediately after test specimen preparation. An *ASD TerraSpec Halo Mineral Identifier* (spectrometer) was used to capture reflectance data from 350

Fig. 2. Particle size distribution (PSD) 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).

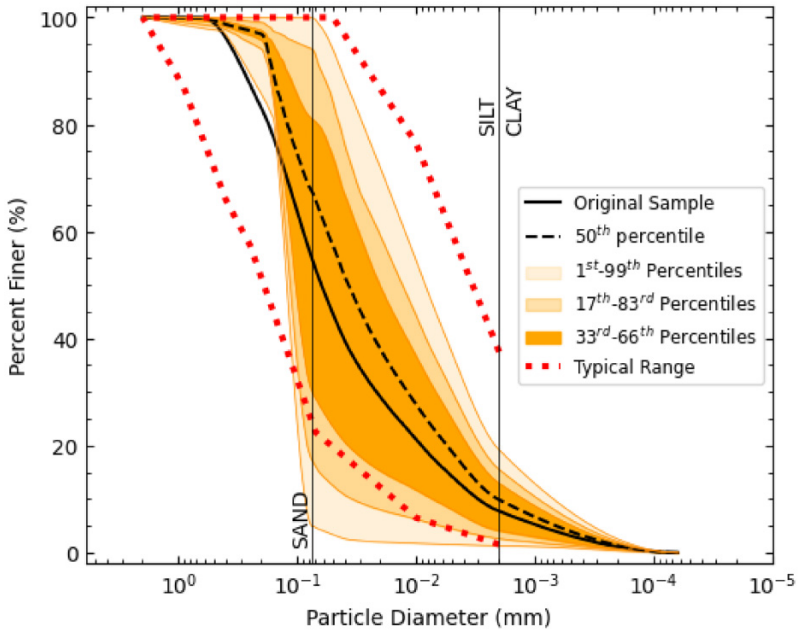
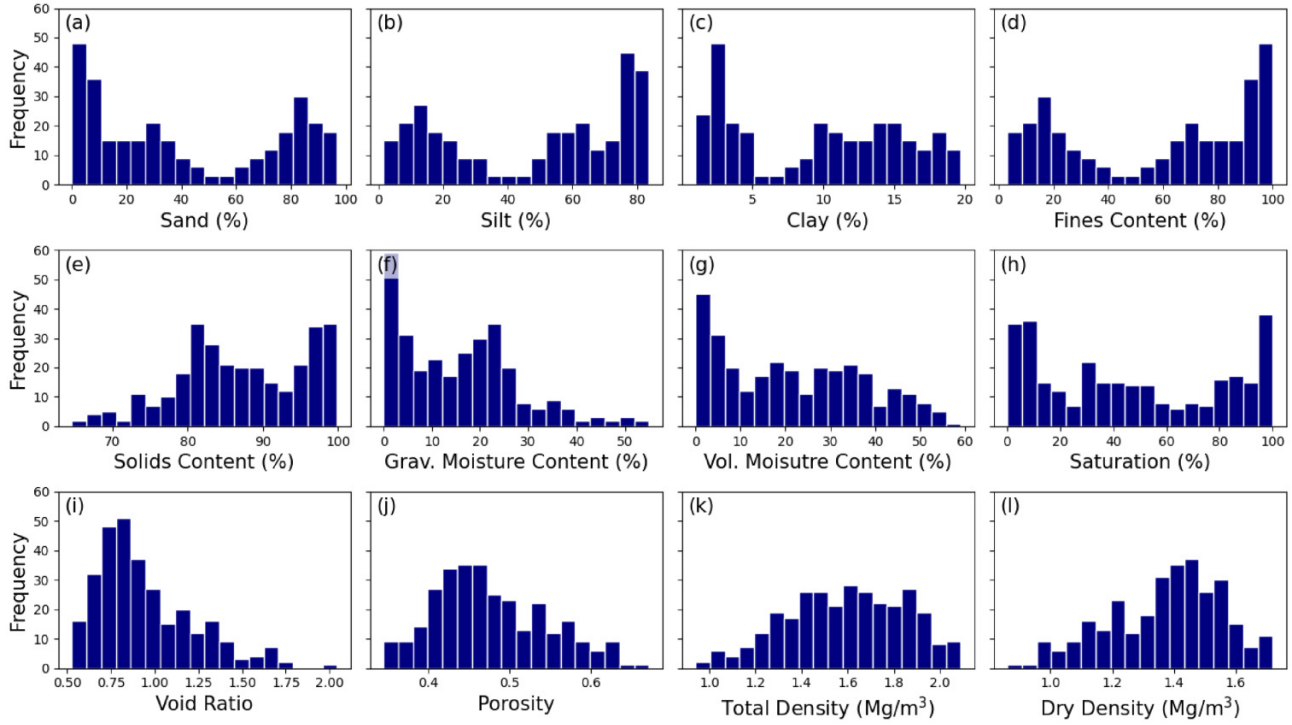
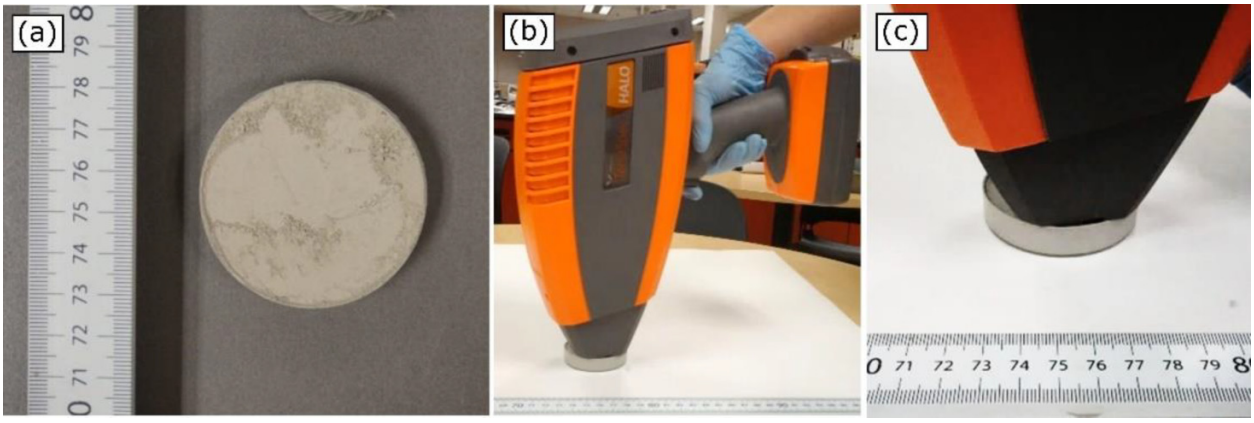


Fig. 3. Distribution of properties among test specimens including (*a*) percent sand, (*b*) percent silt, (*c*) percent clay, (*d*) fines content, (*e*) solids content, (*f*) gravimetric moisture content, (*g*) volumetric moisture content, (*h*) saturation, (*i*) void ratio, (*j*) porosity, (*k*) total density, and (*l*) dry density.



to 2500 nm with a wavelength reproducibility of ±0.1 nm (Malvern Panalytical 2018). A prepared test specimen and the spectrometer are shown in Fig. 4. 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)

Fig. 4. Images of (*a*) a prepared test specimen, (*b*) the TerraSpec Halo spectrometer, and (*c*) close up of the spectrometer in contact with a test specimen.



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 approximately 0.8 cm^2 locations for each test specimen. Hyperspectral data from the two locations were compared for similarity, then averaged to obtain one spectral signal for each test specimen.

2.3. Machine learning

To achieve similar resolutions for all hyperspectral detectors and reduce hyperspectral data dimensionality, the VNIR data were downsampled 3:1 to achieve a spectral resolution of 9 nm from 350 to 1000 nm, resulting in reflectance reported at 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 downsampling 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.

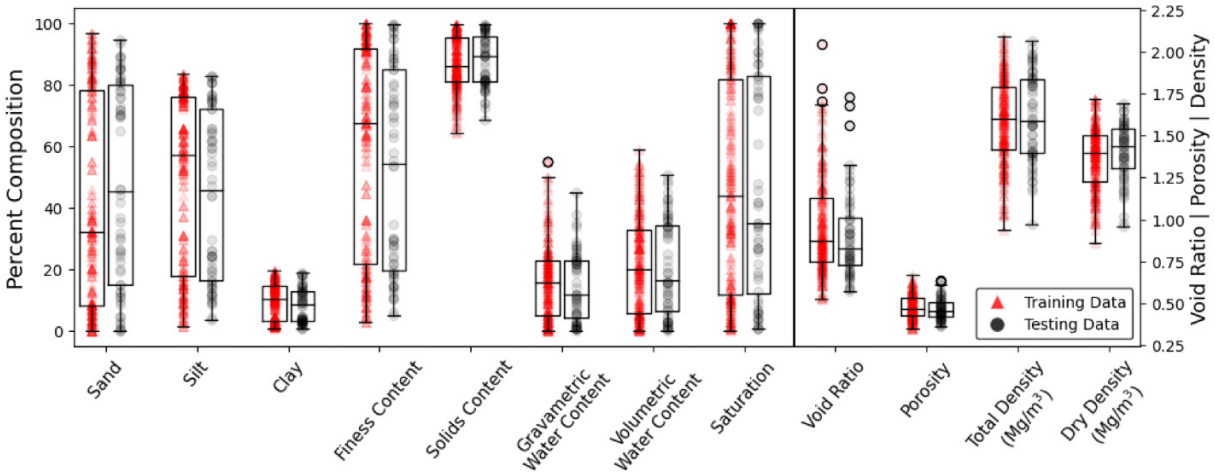
Training and testing datasets were determined using conditioned Latin Hypercube Sampling (cLHS) (Minasny and McBratney 2006) informed by compressed and scaled hyperspectral data. The cLHS algorithm was used to approximate the multivariate distribution of hyperspectral data in both the training and testing 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 ($n = 240$). The distribution of test specimen properties in the training and testing datasets from cLHS is displayed in Fig. 5. Results indicate that 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 properties.

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 ($n = 48$).

To ensure each fold had a similar range and distribution of hyperspectral reflectance data, the fuzzy c-means algorithm (FCM) and elbow method were used (Bezdek et al. 1984). The FCM was first used to cluster similar hyperspectral spectra together for $n = 2$ to $n = 10$ clusters. The within-cluster sum of squares (WCSS) was then calculated and plotted for each *n*-number of clusters. The WCSS decreases as the number of clusters increases, since less diverse data are contained with each cluster. 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 hyperspectral 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 were then used to optimize a 1D 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 (van Rossum and Drake 2009; Abadi et al. 2015; Chollet 2015). The optimized

Fig. 5. Boxplot displaying tailings attributes for test specimens in the training and testing data.



CNN for each validation set was then used to inform the final model hyperparameters.

The optimized hyperparameters and final model architecture are shown in Fig. 6. The input layer consists of a one-dimensional array containing the 121 wavelength approximations from the DWT. The input data are then passed to the first convolutional layer with 84 filters (neurons), where general patterns in the hyperspectral data are identified. The max pooling layer then reduces the spatial dimensionality of the output from convolutional layer one. This is done to lessen the computational load on the next layer. This process is repeated for convolutional layers two and three which both have 64 filters. As data pass through the neural network, more detailed patters in the hyperspectral data are recognized. The output of convolution three is then flattened and passed to the dense, or fully connected, layers for correlation between the target values and the convolved input data. The output of the model contains 12 features, each of which is a floating-point value representing the 12 soil properties predicted in this study. The 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, coefficient of determination (R^2), and the cumulative distribution function 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.

3. Results

3.1. Hyperspectral data

The range of hyperspectral signals for test specimens is shown in Fig. 7. The relative reflectance represents the measured intensity of radiation reflected from the test specimen surface and ranges from 0 to 1, where 0 is complete absorption and 1 is complete 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).

3.2. Machine learning

Tailings property prediction accuracies using the trained CNN are shown in Fig. 8. Predictions of sand, silt, clay, and fines content capture the trend for various PSDs, with R^2 values greater than 0.94 and RMSEs of 5.4%, 5.1%, 1.2%, and 5.3%, respectively (Figs. 8*a*–8*d*). Predictions of solids content have a high R^2 of 0.97 and a low RMSE of 1.4% (Fig. 8*e*). 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 R^2 values greater than 0.94 (Figs. 8*f*–8*h*). Overall, the trend of void ratio and porosity is captured by the CNN predictions but have low R^2 values of 0.20 and 0.27 and RMSEs of 0.23 and 0.05, respectively (Figs. 8*i* and 8*j*). Predictions of total and dry density resulted in RMSEs of 0.16 and 0.14 Mg/m^3 and R^2 values of 0.65 and 0.28, respectively (Figs. 8*k* and 8*l*).

Histograms of errors for each predicted tailings property are shown in Fig. 9. Positive errors represent the overprediction of a given property by the CNN and negative errors represent underpredictions. The red line on each plot represents the kernel density estimate (KDE), which uses a continuous density curve to represent the general shape of the error histograms. The shape of the KDE curves appear to have near normal distributions with some having negative or positive skewness.

Fig. 6. Architecture of the optimized convolutional neural network informed by cross validation. Values in parenthesis represent the output shape of each layer.

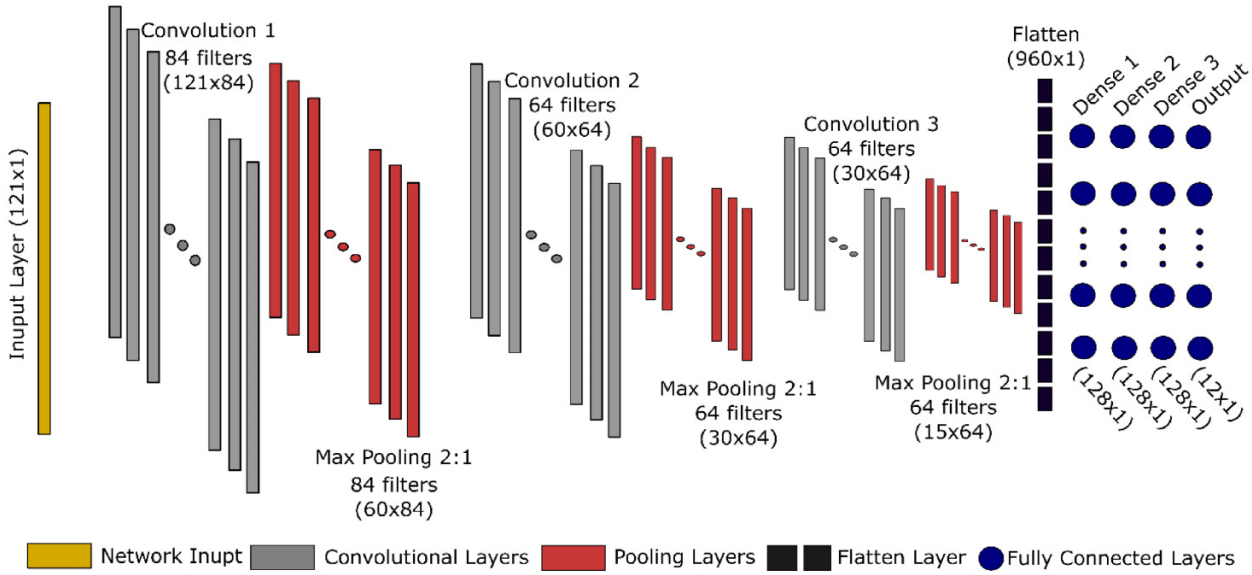
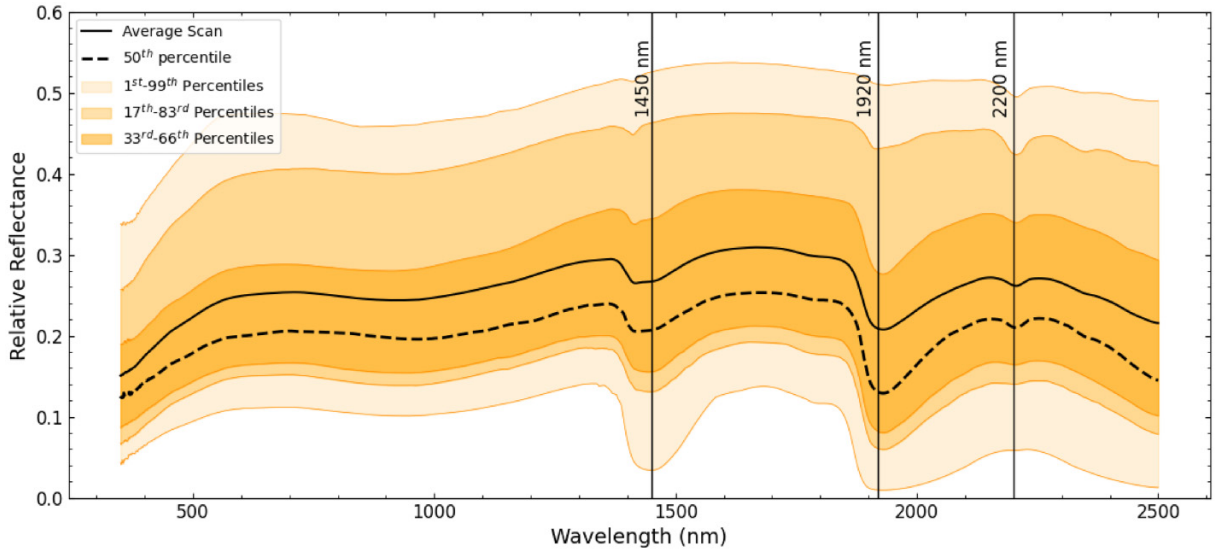


Fig. 7. Percentile ranges for hyperspectral signals including the 1st to 99th percentile, 17th to 83rd percentile, and 33rd to 66th percentile.



For some tailings properties, performance of the CNN is observed to vary with different levels of saturation (Table 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 R^2 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 ratio, porosity, total density, and dry density.

The results from Shapley Additive Explanations (SHAP) analysis are shown in Fig. 10, in which SHAP 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 and 2500 nm. Wavelengths from 2200 to 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.

Fig. 8. Predicted values from the trained convolutional neural network 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. RMSE, root mean squared error; MBE, mean bias error.

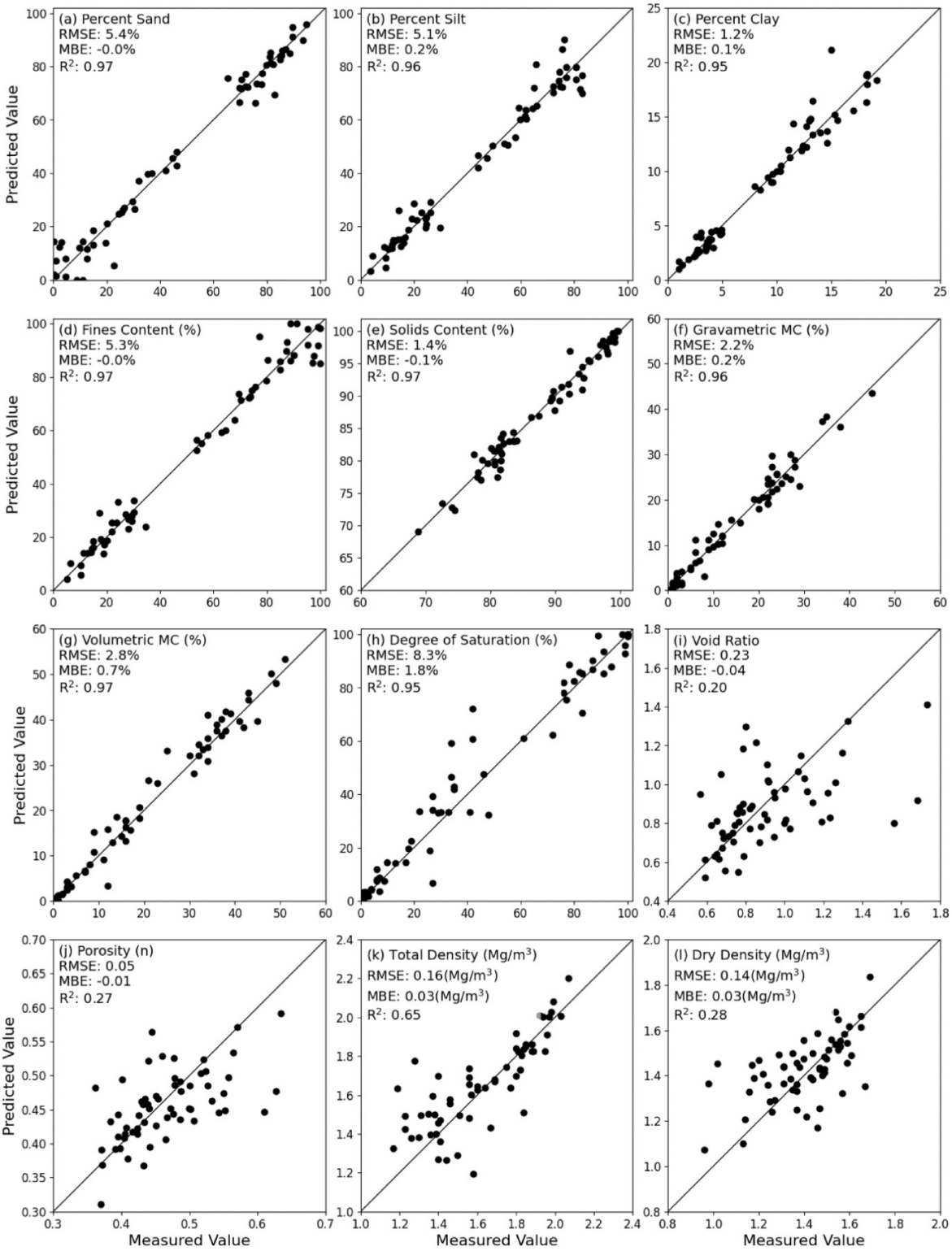


Fig. 9. Histograms of convolutional neural network prediction errors and the kernel density estimate of errors for (*a*) percent sand, (*b*) percent silt, (*c*) percent clay, (*d*) fines content, (*e*) solids content, (*f*) gravimetric moisture content, (*g*) volumetric moisture content, (*h*) degree of saturation, (*i*) void ratio, (*j*) porosity, (*k*) total density, and (*l*) dry density.

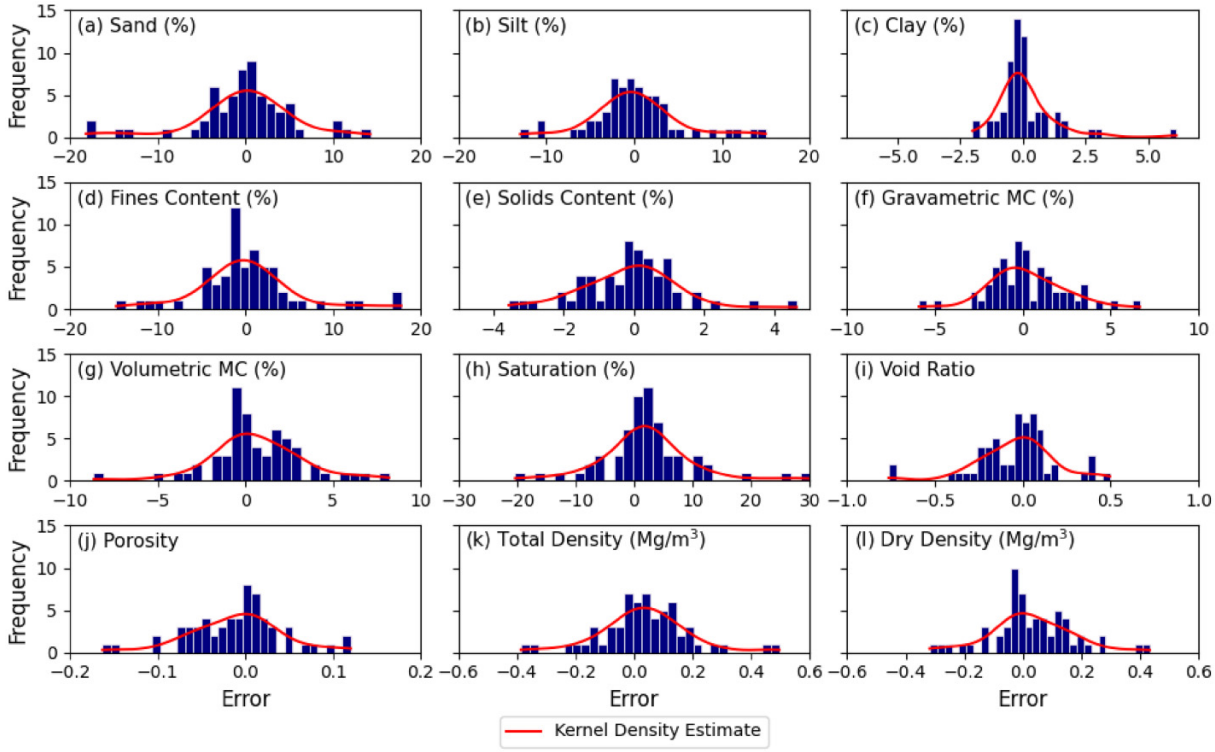


Table 2. Convolutional neural network performance metrics for saturation, void ratio, porosity, total density, and dry density for test specimens with saturations from 0% to 50% and 50% to 100%.

		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
	R^2	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
	R^2	0.79	0.77	0.76	0.80	0.80

Note: RMSE, root mean squared error; MBE, mean bias error.

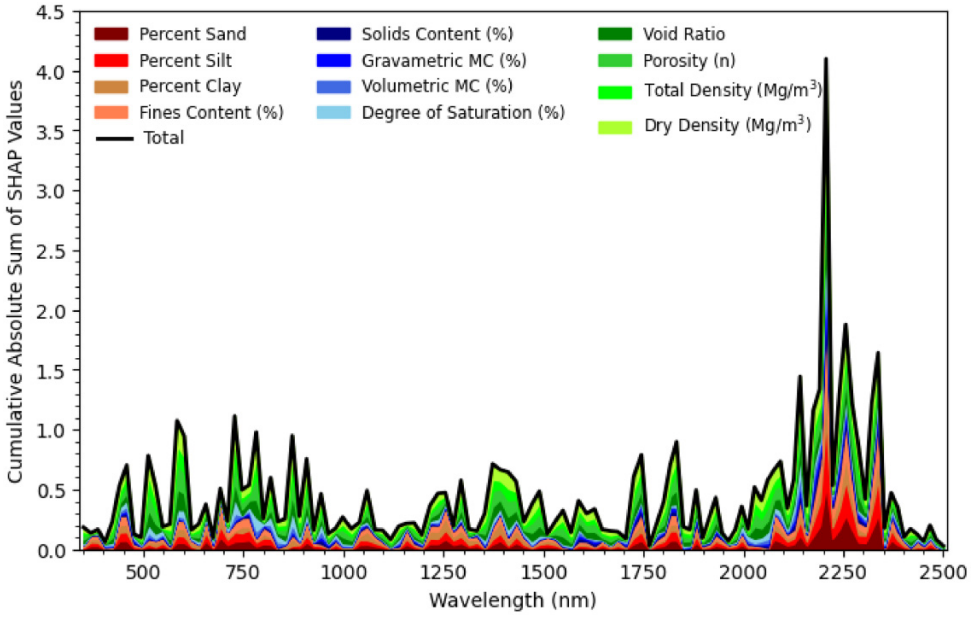
Finally, the predictions of void ratio, porosity, total density, and dry density appear to be strongly influenced by wavelengths near 600, 750, 900, 1400, and 2200 nm.

4. Discussion

The CNN model results demonstrate that the estimation of tailings properties related to PSD, moisture, and density using hyperspectral sensing shows promise for tailings characterization. Taskiridis et al. (2020) used hyperspectral reflectance measured for 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 model 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 ability of the CNN model to capture variability in PSD. Entezari et al. (2022) used hyperspectral data and machine learning to predict percent fines, solids content, and gravimetric moisture content of oil-sands tailings which resulted in errors 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 that hyperspectral data

Fig. 10. Cumulative absolute sum of Shapley Additive Explanations (SHAP) values for the convolutional neural network 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.



and CNNs show promise 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 R^2 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 laboratory hyperspectral data for the rapid characterization of tailings may benefit practitioners by providing near real-time estimates of geotechnical properties and reducing the time and labor associated with traditional laboratory testing. This has been demonstrated for the prediction of oil-sands tailings properties by Entezari et al. (2018, 2019, 2022, 2024). The application of this methodology, however, requires that the model is trained on data that properly represent the type and range of materials that the model is intended to be applied to. Practitioners should consider the investment required to develop similar models for their applications. Although the creation of these models can require front-end investment, studies show promise for the use of transfer learning to leverage models developed on large datasets to supplement the calibration of new models for relatively small datasets (Liu et al. 2018). Future studies should consider how transfer learning can be applied to similar methodologies to add efficiencies to the training process.

The use of hyperspectral sensing for the prediction of tailings properties also has potential for the high-speed characterization of tailings properties in the field. Studies have demonstrated the use of visible light cameras on CPTu probes to better characterize soil profiles (Ventola et al. 2020; Entezari et al. 2024) and hyperspectral sensing has been used in subsurface probes for characterization of agricultural soil properties (Cho et al. 2017; Pei et al. 2018; Entezari et al. 2024). 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 and the CPTu (tip resistance, sleeve friction, pore pressure, etc.) (Pei et al. 2018; Riese and Keller 2019; Entezari et al. 2024). 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. In addition, 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. The geotechnic laboratory reference methods in this study were selected for their common use in practice and standardized procedures, however, there may be some level of expected error in the measured reference values, future studies should consider how error in reference measurements propagate though model training and influence model predictions.

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 and complexity on the predictive accuracy of tailings properties. This study did not include assessments of variability in spectral response with changes in specific tailings properties. However, the relationship between spectral response and moisture content is well documented (Jacquemoud and Ustin 2003; van der Meer 2004; Diao et al. 2021) and other studies have related increased particle size with decreased reflectance in the 350–2500 nm range for natural soils (Clark et al. 1993; Sadeghi et al. 2018). High dry density (low void ratio and porosity) has also been observed to correspond with lower reflectance for sandy materials (Bachman et al. 2014; Carson et al. 2015). However, this trend may vary based on soil type (Demattê et al. 2010). Future studies should further investigate how variations in tailings properties directly impact spectral response.

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 PSDs, 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 1D 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 this study, the following conclusions are made:

1. Hyperspectral data show promise for the prediction of tailings PSD 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 PSD.
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%. (R^2 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 R^2 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 R^2 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.

Acknowledgements

The authors gratefully acknowledge the Tailings and Industrial Waste Engineering Foundation (TAILENG) for providing funding for this work. The authors thank Holly Ho, Kaylee Romero, Celie Brockett, and Alec Shields for their support of laboratory testing. Finally, the authors extend gratitude to Dr. John Ridley for supplying instrumentation critical to this work.

Article information

History dates

Received: 24 October 2024
Accepted: 7 August 2025
Accepted manuscript online: 14 August 2025
Version of record online: 24 September 2025

Copyright

Data availability

The complete tailings hyperspectral dataset can be found in the repository associated with this article (https://doi.org/10.5061/dryad.v15dv4278) (Bindner et al. 2025). Code generated from this study is available from the corresponding author upon reasonable request.

Author information

Author ORCIDs

Joseph Bindner https://orcid.org/0009-0001-4231-4679
Christopher Bareither https://orcid.org/0000-0003-3427-4344

Author notes

Joseph Scalia served as Editorial Board Member at the time of manuscript review and acceptance; peer review and editorial

decisions regarding this manuscript were handled by another editorial board member.

Author contributions

Conceptualization: JB, CB, JS
Data curation: JB
Formal analysis: JB
Funding acquisition: CB, JS
Investigation: JB
Methodology: JB, JS
Project administration: CB, JS
Resources: CB, JS
Supervision: CB, JS
Software: JB
Visualization: JB
Writing – original draft: JB
Writing – review & editing: CB, JS

Competing interests

The authors declare there are no competing interests.

Funding information

This research was funded by the Tailings and Industrial Waste Engineering Foundation (TAILENG).

References

Aachib, M., Mbonimpa, M., and 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. doi:10.1023/B:WATE.0000036803.84061.e5.

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. 2015. TensorFlow: large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.

Agapito, L.A., and Bareither, C.A. 2018. Application of a one-dimensional large-strain consolidation model to a full-scale tailings storage facility, Minerals Engineering, **119**: 38–48. doi:10.1016/j.mineng.2018.01.013.

Althnian, A., AlSaeed, D., Al-Baity, H., Samha, A., Bin Dris, A., Alzakari, N., et al. 2021. Impact of dataset size on classification performance: an empirical evaluation in the medical domain. Applied Sciences, **11**(2):796. doi:10.3390/app11020796.

Anand, R., Veni, S., and Aravinth, J. 2021. Robust classification technique for hyperspectral images based on 3D-discrete wavelet transform. Remote Sensing, (13): 7. doi: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. 2021. Standard test methods for laboratory compaction characteristics of soil using standard effort (12.400 ft-lbf/ft3 (600kN-m/m3)) (ASTM D698-12).

ASTM International. 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., and Korwan, D. 2014. Phase angle dependence of sand density observable in hyperspectral reflectance. Remote Sensing of Environment, **150**: 53–65. doi:10.1016/j.rse.2014.03.024.

Bezdek, J.C., Ehrlich, R., and Full, W. 1984. FCM: The fuzzy c-means clustering algorithm. Computers & Geosciences **10**(2–3): 191–203. doi:10.1016/0098-3004(84)90020-7.

Bindner, J., Bareither, C., and Scalia, J. 2025. Geotechnical and hyperspectral dataset for gold tailings (version 1) [data set]. Dryad. doi:10.5061/dryad.v15dv4278.

Blight, G., and Steffen, D. 1979. Geotechnics of gold mine waste disposal. *In* Current geotechnical practice in mine waste disposal. American Society of Civil Engineers, New York. pp. 1–52.

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): 2331–2338. doi:10.1109/TGRS.2002.804721.

Carson, T., Bachmann, C., and Salvaggio, C. 2015. Soil signature simulation of complex mixtures and particle size distributions. Optical Engineering, **54**(9): 094103. doi:10.1117/1.OE.54.9.094103.

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, **60**: 683–692. doi:10.13031/trans.12049.

Chollet, F. 2015. Keras. GitHub. Available from https://github.com/fchollet/keras [accessed 30 June 2023].

Clark, R.N., Swayze, G.A., Gallagher, A.J., King, T., and Calvin, W.M. 1993. The U.S. Geological Survey, digital spectral library: Version 1 (0.2 to 3.0 μm) (Open-File Report 93-592) U.S. Geological Survey. doi:10.3133/ofr93592.

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. doi:10.1080/01431160902893469.

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). doi:10.1139/cgj-2018-0066.

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): 710. doi:10.3390/agriculture11080710.

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.

Entezari, I., McGowan, D., and Glavina, J. 2022. Hyperspectral imaging technology for oil sands tailings characterization: practical aspects. In Proceedings of Tailings and Mine Waste 2022 (1243-1252).

Entezari, I., McGowan, D., Bindner, J., Glavina, J., and DeJong, J. 2024. Enhancing soil and tailings characterization: integration of photonics technology with cone penetration testing. In Proceedings of Tailings and Mine Waste 2024 (2027-2038).

Entezari, I., Rivard, B., Sharp, J., Wells, P.S., Styler, M.A., and McGowan, D. 2018. Real-time prediction of oil sands tailings properties using hyperspectral observations. In Proceedings of 6th International Oil Sands Tailings Conference (102-106).

Entezari, I., Sharp, J., and McGowan, D. 2019. Characterization of oil sands tailings using hyperspectral technology. In Proceedings of Tailings and Mine Waste 2019 (519-528).

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. In Eighth International Conference on Case Histories in Geotechnical Engineering. doi:10.1061/9780784482148.005.

GTR (Global Tailings Review). 2020. Global industry standard on tailings management. Available from https://globaltailingsreview.o

rg/wp-content/uploads/2020/08/global-industry-standard_EN.pdf [accessed January 2024].

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**: 115071. doi:10.1016/j.geoderma.2021.115071.

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.

Hedley, C., Poudier, P., and Maddi, L. 2015. VNIR soil spectroscopy for field soil analysis. Communications in Soil Science and Plant Analysis, **46**(1): 104–121. doi:10.1080/00103624.2014.988582.

ICMM (International Council on Mining & Metals). 2021. Tailings management: good practice guide. ICMM. Available from 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. *In* Proceedings of the 4th International Workshop on Remote Sensing and GIS Applications to Forest Fire Management.

Jefferies, M., and Been, K. 2016. Soil liquefaction–a critical state approach. CRC Press, Taylor & Francis Group, Boca Raton, Florida, USA.

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. doi:10.1016/j.geoderma.2019.114080.

KCB (Klohn Crippen Berger). 2017. Study of tailings management technologies. MEND report 2.50.1. Available from 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. U.S. Bureau of Mines, Washington, DC.

Laukamp, C., Rodger, A., LeGras, M., Lampinen, H., Lau, I., Pejcic, B., et al. 2021. Mineral physiochemistry underlying feature-based extraction of mineral abundance and composition from shortwave, mid and thermal infrared reflectance spectra. Minerals, **11**(4): 347. doi:10.3390/min11040347.

Lee, G., Gommers, R., Waselewski, F., Wohlfahrt, K., and O'Leary, A. 2019. PyWavelets: a python package for wavelet analysis. Journal of Open Source Software, **4**(36): 1237. doi:10.21105/joss.01237.

Liu, L., Ji, M., and Buchroithner, M. 2018. Transfer learning for soil spectroscopy based on convolutional neural networks and its application in soil clay content mapping using hyperspectral imagery. Sensors, **18**(9): 3169. doi:10.3390/s18093169.

Lundberg, S., and Lee, S. 2017. A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems 30. Available from https://proceedings.neurips.cc/paper_files/paper/2017 [accessed 15 November 2023].

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. American Society of Civil Engineers, New York. pp. 103–117.

Malvern Panalytical. 2018. TerraSpec halo user manual (ASD Document 600073). Malvern Panalytical. Available from https://www.malvernpanalytical.com/ [accessed 8 March 2024].

Mayne, P.W., Coop, M., Springman, S., Huang, A.B., and Zornberg, J.G. 2009. Geomaterial behavior and testing. *In* Proceedings of the Seventeenth International Conference on Soil Mechanics and Geotechnical Engineering: State-of-the-art keynote lecture. Alexandria, Egypt. pp. 2777–2872.

METER Group. 2017. PARIO user manual V2017/12. METER Group AG. Available from http://library.metergroup.com [accessed 8 March 2024].

Minasny, B., and McBratney, A.B. 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Computers & Geosciences **32**(9): 1378–1388. doi:10.1016/j.cageo.2005.12.009.

Molle, J. 2005. The accuracy of the interpretation of CPT-based soil classification methods in soft soils. M.Sc. thesis, Section for Engineering Geology, Department of Applied Earth Sciences, Delft University of Technology, Report No. 242, Report AES/IG/05-25.

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. doi: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. Available from https://www.resolutionmineeis.us/documents/fundao-2016 [accessed 6 February 2024].

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. doi: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. doi:10.1016/j.resourpol.2009.12.001.

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**: 112315. doi: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. Global Tailings Review, St Gallen, Switzerland. Available from https://globaltailingsreview.org/ [accessed 5 February 2024].

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. 2011. Scikit-learn: machine learning in Python. Journal of Machine Learning Research, **12**(85): 2825–2830. Available from https://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html [accessed 20 March 2024].

Pei, X., Sudduth, K., Veum, K., and Minzan, L. 2018. Improving in-situ estimation of soil profile properties using a multi-sensor probe. Sensors, **19**(1011). doi:10.3390/s19051011.

Prusa, J., Khoshgoftarr, T., and Seliya, N. 2015. The effect of dataset size on training tweet sentiment classifiers. International Conference on Machine Learning Applications. doi:10.1109/ICMLA.2015.22.

Quinlan, J.R. 1992. Learning with continuous classes. In Proceedings of the 5th Australian Joint Conference on Artificial Intelligence. doi:10.1142/9789814536271.

Quinlan, J.R. 1993. Combining instance-based and model-based learning. In 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, **195–196**: 268–279. doi:10.1016/j.geoderma.2012.12.014.

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). doi:10.5194/isprs-annals-IV-2-W5-615-2019.

Robertson, P. 2016. Cone penetration test (CPT)-based soil behaviour type (SBT) classification system–an update. Canadian Geotechnical Journal, **53**(12): 1910–1927. doi: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. Available from http://www.b1technicalinvestigation.com/ [accessed 6 February 2024].

Rodríguez, R., Muñoz–Moreno, A., Caparrós, A., García–García, C., Brime–Barrios, A., Arranz–González, J., et al. 2021. How to prevent flow failures in tailings dams. Mine Water and the Environment, **40**: 83–112. doi:10.1007/s10230-021-00752-8.

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. doi:10.1016/j.rse.2018.03.028.

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. doi:10.1016/S0003-2670(00)85460-4.

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**: 114208. doi: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). doi:10.3390/w12030737.

United Nations Development Programme (UNDP) and UN Environment. 2018. Managing mining for sustainable development: a sourcebook. United Nations Development Programme, Bangkok.

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. doi:10.1016/j.jag.2003.09.001.

van Rossum, G., and Drake, F.L. 2009. Python 3 reference manual. CreateSpace, Scotts Valley, CA.

Ventola, A., Dolling, R., and Hryciw, R.D. 2020. Image analysis and hardware development for the vision cone penetrometer (VisCPT). GeoCongress 2020. doi:10.1061/9780784482803.068.

Veres, M., Lacey, G., and Taylor, G.W. 2015. Deep learning architectures for soil property prediction. In 12th Conference on Computer and Robot Vision. doi:10.1109/CRV.2015.15.

Vick, S.G. 1990. Planning, design, and analysis of tailings dams [B]. doi:10.14288/1.0394902.

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**: 102007. doi:10.1016/j.gloenvcha.2019.102007.

Williams, D. 2021. Lessons from tailings dam failures–where to go from here?. Minerals, **11**(8): 853. doi:10.3390/min11080853.

Wride, C., Hofmann, D., Sego, D., Plews, H.D., Konrad, J.M., Biggar, K.W., et al. 2000a. Ground sampling program at the CANLEX test sites. Canadian Geotechnical Journal, **37**(3): 530–542. doi:10.1139/t00-045.

Wride, C., Robertson, P., Biggar, K., Campanella, R.G., Hofmann, B.A., Hughes, J.M., et al. 2000b. Interpretation of in situ test results from the CANLEX sites. Canadian Geotechnical Journal, **37**(3): 1505–1529. doi:10.1139/t00-044.

Zhu, X., Vondrick, C., Fowlkes, C., and Remanan, D. 2016. Do we need more training data? International Journal of Computer Vision, **119**: 76–92. doi:10.1007/s11263-015-0812-2.