Predicting Tailings Properties Using Hyperspectral Sensing and Machine Learning

Joseph Bindner, ConeTec Group, Canada
Joseph Scalia IV, Colorado State University, USA
Christopher Bareither, Colorado State University, USA
Iman Entezari, ConeTec Group, Canada
Dallas McGowan, ConeTec Group, Canada

Abstract

Understanding tailings properties at high spatial resolutions is critical for geotechnical and geochemical analyses. While tailings properties can be characterized using undisturbed samples and in-situ tests, the current methods have limitations. Tailings properties determined from undisturbed samples represent a point measurement and obtaining enough samples to characterize tailings at high spatial resolutions may not be feasible. The characterization of tailings in situ, using methods such as the cone penetration test (CPT), can provide high resolution data, but does not provide information related to all relevant tailings properties. This study explores the use of hyperspectral sensing and convolutional neural networks (CNN) for the prediction of tailings properties, including percent sand, silt, clay, solids content, moisture content, and degree of saturation. Tailings from a gold, silver, lead, and zinc producing mine were used to create samples with diverse properties and hyperspectral scans of each sample were captured. The tailings-hyperspectral dataset was then split into training and testing subsets and a CNN was optimized and trained. Model performance was assessed using the testing data and the results demonstrate strong promise for the use of hyperspectral data and CNNs for the prediction of tailings properties.

Introduction

Tailings properties can inform geotechnical and geochemical analyses and include particle size distribution (PSD), solids content, moisture content, and degree of saturation. The PSD of tailings influences shear strength, hydraulic conductivity, consolidation, dry density, and beach angle, and can inform geochemical and rheological characterization of the material. The moisture content of tailings influences hydraulic conductivity, bearing capacity, compaction behavior, and acid-generation potential. At saturations below approximately 80%, geotechnical stability is improved (Rodriguez et al., 2021) and at saturations above approximately 85%, oxygen diffusion exponentially decreases, indicating reduced acid generation in sulfide tailings (Aachib et al., 2004).

Tailings can be characterized by sampling and laboratory testing. However, the results only represent a point measurement and obtaining enough samples for high resolution characterization of tailings throughout a tailings facility is impractical. In-situ methods can be a cost-effective alternative for high resolution characterization and include the cone penetration test (CPT). Data from the CPT can be used to better understand tailings behavior; however, current methods for CPT interpretation may not apply universally and the CPT does not provide information on all geotechnically relevant tailings properties. Alternatively, hyperspectral sensing has shown promise for the characterization of soil and tailings properties.

Figure 1 illustrates a simplified conception of hyperspectral sensing, which involves illuminating a sample and measuring the intensity of reflected radiation from the material surface. Hyperspectral reflectance can be measured with a spectrometer and varies with tailings properties including water content, PSD, and mineralogy. Hyperspectral sensing is typically used to measure reflectance in the visible, near infrared, and short-wave infrared portions of the electromagnetic spectrum and can be conducted on disturbed or undisturbed samples in either laboratory or field settings.

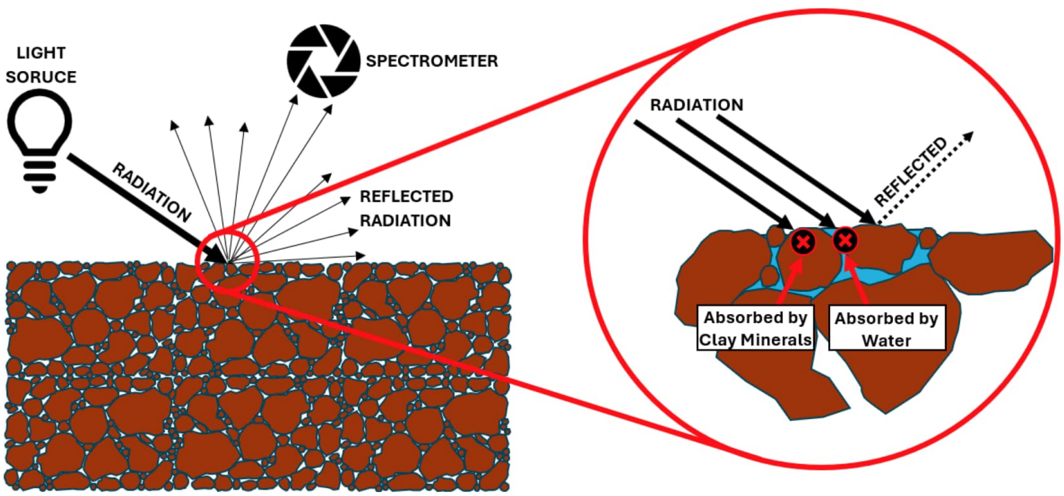

Figure 1: Simplified illustration of hyperspectral sensing for tailings

Machine learning models have been successfully used to predict soil texture from laboratory measured hyperspectral reflectance (Riese and Keller, 2019; Tsakiridis et al., 2020). In agriculture, studies have demonstrated the use of probe-based hyperspectral sensing for the estimation of in-situ soil texture and moisture content for near surface soils (Cho et al., 2017; Pei et al., 2018). Additionally, ex-situ hyperspectral sensing has been used for the prediction of tailings solids content, fines content, and water content for oil-sands tailings (Entezari et al., 2022). While machine learning models, such as convolutional neural networks (CNNs), and hyperspectral reflectance can be used to predict some tailings properties, less is understood regarding the use of hyperspectral sensing for the prediction of tailings percent sand, silt, clay, and degree of saturation.

The objective of this study is to measure hyperspectral reflectance for tailings with diverse properties and assess how hyperspectral reflectance 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 reflectance was captured using a spectrometer and test specimen properties were measured using standard laboratory procedures. The tailings-hyperspectral dataset was separated into training and testing subsets and a CNN was constructed to assess the use of hyperspectral reflectance for the prediction of tailings properties including percent sand, percent silt, percent clay, solids content, gravimetric moisture content, and degree of saturation. The CNN was trained and tested using (1) the full range of measured wavelengths and (2) a limited range of wavelengths, to better understand how short-wave infrared data impacts the model's performance, as a potential cost-reducing measure.

Methodology

Test specimens

Table 1 outlines the engineering characteristics for a bulk tailings sample from a gold, silver, lead, and zinc producing mine (whole tailings). The tailings were processed to artificially prepare tailings specimens with diverse properties including PSD, moisture content, and density (test specimens). The whole tailings properties were determined according to ASTM standards and sensor testing manuals (see Table 1).

The whole tailings underwent a series of sieving and sedimentation procedures to separate out fractions of the whole tailings that were dominated by (1) sand, (2) silt, and (3) clay. Unique mass ratios of each tailings fraction were combined to prepare 100 test specimens with different PSDs (PSD specimens). Figure 2 displays the percentile ranges for the 100 PSD specimens and the typical range for tailings PSD from the literature.

Table 1: Whole tailings properties

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

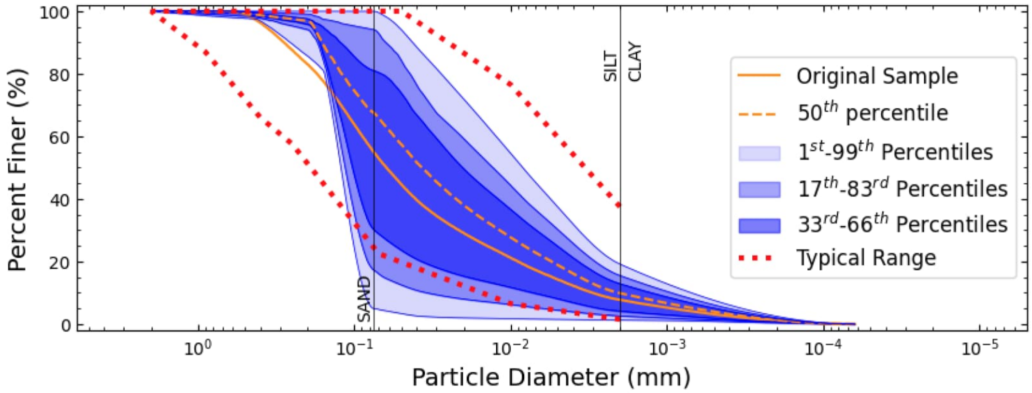

Figure 2: Particle size distribution of the whole tailings and the percentile ranges for all PSD specimens. Solid vertical lines differentiate sand, silt, and clay sized particles. Red dashed lines represent the average lower and upper bounds of PSD for typical mine tailings from the literature (after Gorakhki et al., 2019, adapted from Hamade, 2017)

Each PSD specimen was prepared to three moisture and density conditions, totaling 300 test specimens. Each test specimen was moisture conditioned using de-ionized water, sealed, and allowed to moisture equilibrate for 12 hours. Test specimens were then compacted in a 50 mm inner diameter, 15 mm height, petri dish to target a density, which was randomly determined and between 0.9 and 1.7 Mg/m³. Each specimen's total mass, volume, and gravimetric moisture content was measured and used to calculate solids content and saturation based on phase relationships.

Figure 3 illustrates the frequency distributions for the test specimens' properties. The sand, silt, and clay content ranged between 0–99%, 2–84%, and 1–20%, respectively (Figure 3a-c). Solids content spanned 65–100%, capturing a portion of the tailings continuum including thickened, paste, and filtered tailings (Figure 3d). Gravimetric moisture content ranged from 0–55% and saturation varied from 0–100% (Figure 3e-f).

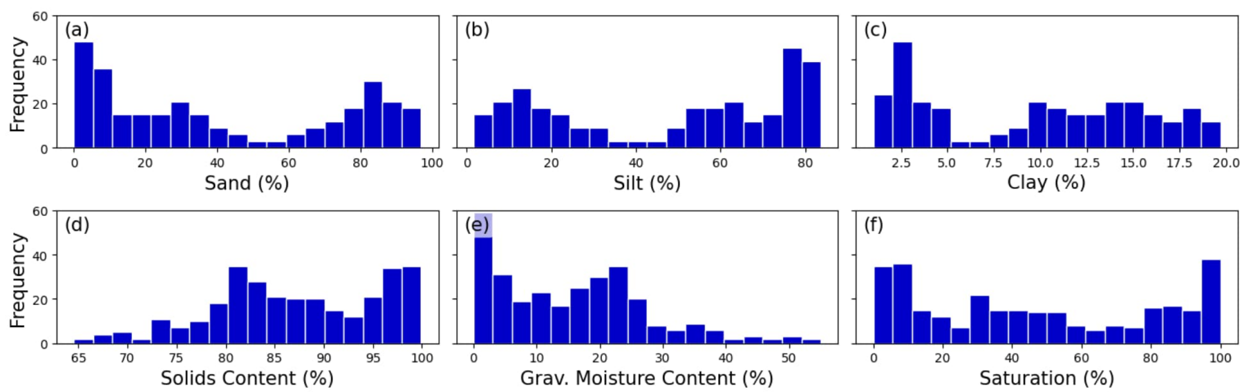

Figure 3: Distribution of test specimen properties including (a) percent sand, (b) percent silt, (c) percent clay, (d) solids content, (e) gravimetric moisture content, and (f) degree of saturation

Hyperspectral sensing

Figure 4 illustrates the preparation of test specimens and the spectrometer used to capture hyperspectral reflectance. Immediately following test specimen preparation, the specimen's reflectance intensity was measured at wavelengths ranging from 350 – 2500 nm. The spectrometer contains: (1) a visible-near-infrared sensor (VNIR) (350–1,000 nm); (2) two short-wave infrared sensors, referred to as SWIR-1 (1001–1,800 nm) and SWIR-2 (1,801–25,00 nm); and (3) an internal halogen light for sample illumination. The spectrometer's viewing window was placed in direct contact with the test specimen and hyperspectral reflectance was measured at two locations for each test specimen. The data from these two locations were averaged to generate a single hyperspectral signal for each test specimen.

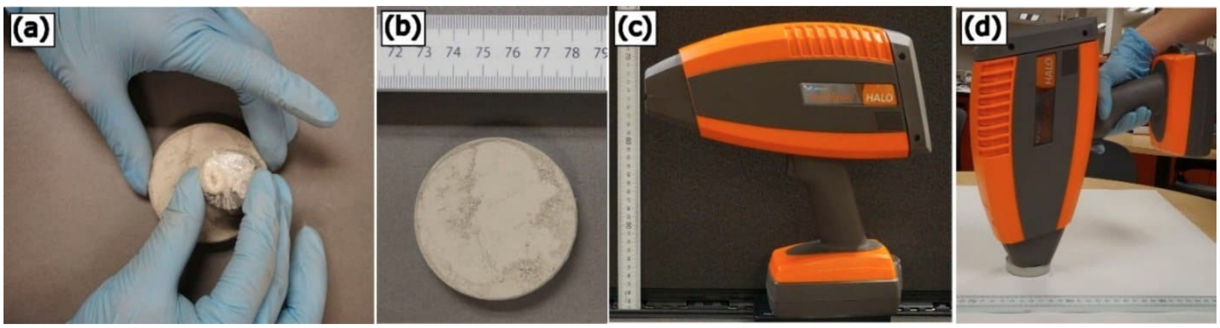

Figure 4: Photographs illustrating (a) test specimen compaction, (b) prepared test specimen, (c) spectrometer, and (d) spectrometer measuring hyperspectral reflectance of a test specimen

Data processing

Various preprocessing procedures were applied to the hyperspectral data including the discrete wavelet transform (Anand et al., 2021; Bruce et al., 2022) and the standard scaler transform (Pedregosa et al., 2011) to reduce data dimensionality and standardize data, which has been associated with improved model performance (Pedregosa et al., 2011; Anand et al., 2021; Bruce et al., 2022). The tailings-hyperspectral dataset was then split into training and testing sub-datasets using conditioned Latin Hypercube Sampling (cLHS) (Minasny and McBratney, 2006). The cLHS algorithm was used to create training and testing datasets that have similar distributions of hyperspectral data, which has been linked to improved model performance (Althnian et al., 2021). The testing data contains 20% of all samples (n=60) and the training data contains 80% (n=240).

Figure 5 illustrates the distribution of properties among test specimens in the training and testing datasets derived from cLHS. Results from cLHS demonstrate that the testing data has a similar distribution and range of properties as that seen in the training data.

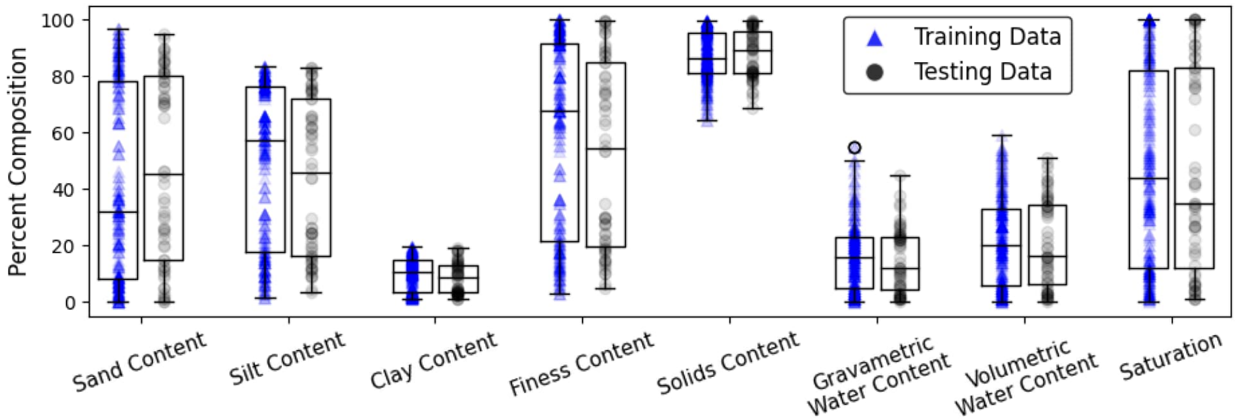

Figure 5: Distribution of tailings properties in the training and testing datasets

Machine learning

Convolutional neural networks have been shown to outperform other common algorithms in soil spectroscopy (Tsakiridis et al., 2020) and were used to predict tailings properties from hyperspectral data in this study. K-fold cross validation helps to generalize models for predictions on unseen data (i.e., the testing data) and was used to select the CNN architecture. The Keras deep learning application program interface was used for machine learning procedures (Chollet, 2015). The model was first trained and tested using the entire hyperspectral dataset, including VNIR, SWIR-1, and SWIR-2 data (V-S2 model). Then, to assess how the SWIR-2 data influences model performance, the model was trained and tested using only VNIR and SWIR-1 data (V-S1 model).

Results

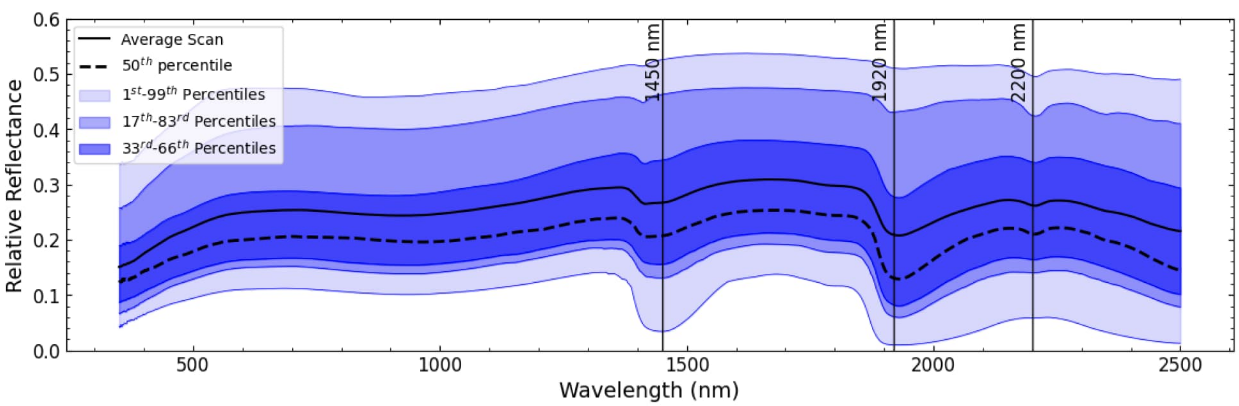

Figure 6: Hyperspectral signal percentile ranges including percentiles 1–99, 17–83, and 33–66. Vertical lines are shown at major absorptive features

Figure 6 illustrates the range of hyperspectral signals for test specimens. Relative reflectance denotes the spectrometer's measurement of radiation reflected from the surface of the test specimen and ranges from 0 to 1. A value of 0 indicates complete absorption, while 1 signifies total reflectance. Distinct absorptive features are observed in the hyperspectral signals at wavelengths near 1,450 nm and 1,920 nm, indicating the presence of water (van der Meer, 2004). Additionally, the absorptive feature at 2,200 nm is attributed to clay minerals containing hydroxyl groups such as illite, chlorite, and kaolinite (Laukamp et al., 2021), which were present in the whole tailings based on X-ray diffraction analysis.

Figure 7 displays the results from the V-S1 and V-S2 models along with the root mean squared error (RMSE), mean bias error (MBE), and coefficient of determination (R^2) of predictions. For predictions of percent sand and silt, the V-S2 model has RMSEs of 5.9% and 5.1% respectively, which are approximately 2.5% lower than the V-S1 model (8.5% and 7.5% respectively). Additionally, the R^2 values are approximately 0.04 higher for the V-S2 model compared to the V-S1 model. Both models have MBEs near 0% for sand and silt predictions. Performance metrics for percent clay and solids content are comparable between both models with RMSEs less than 2%, MBE near 0%, and R^2 greater than 0.94. The V-S2 model performed slightly better than the V-S1 model for prediction of gravimetric moisture content (RMSE of 2.2% and 2.7% respectively). Conversely, for the prediction of degree of saturation, the V-S1 model outperforms the V-S2 model with RMSEs of 7.5% and 8.3% respectively and MBEs of 0.6% and 2.1% respectively.

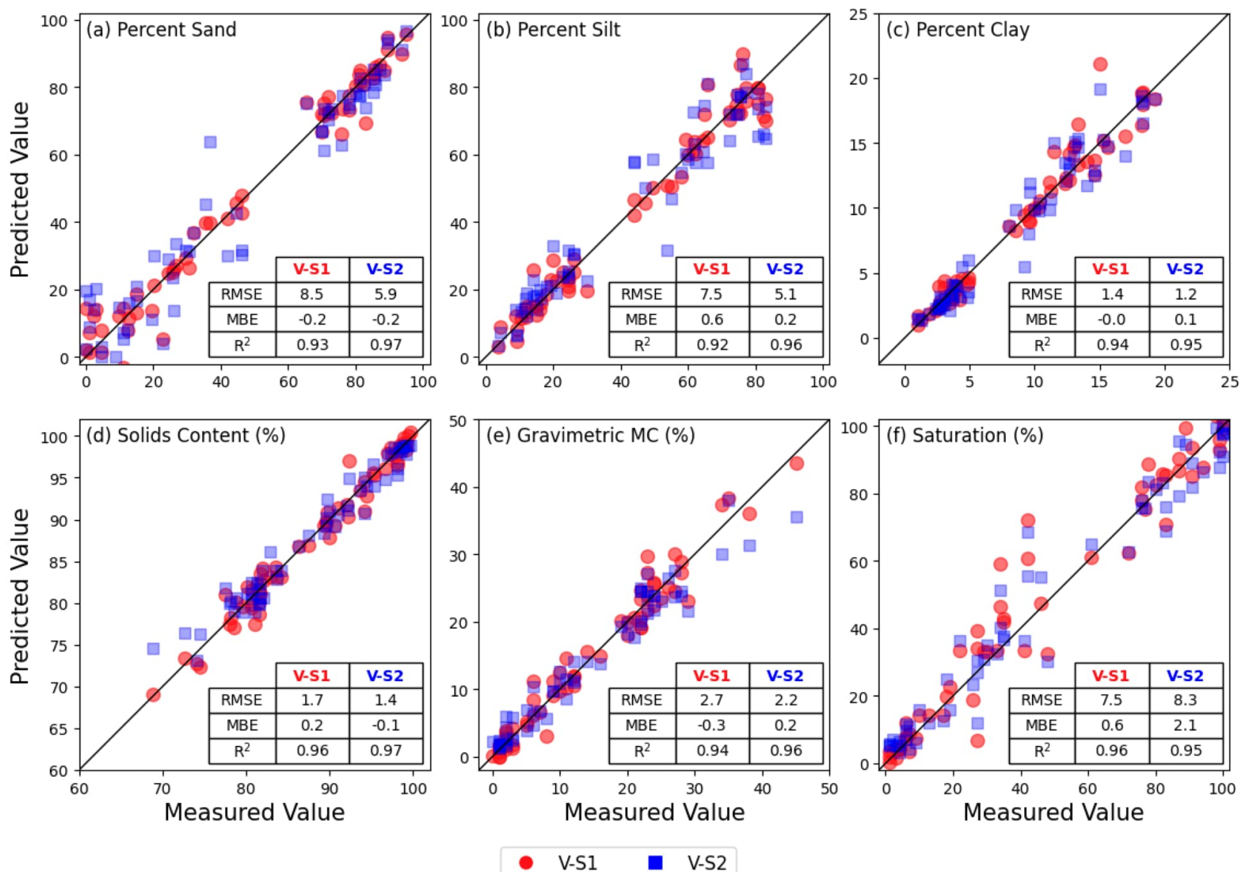

Figure 7: Results from V-S1 and V-S2 models for (a) percent sand, (b) percent silt, (c) percent clay, (d) solids content, (e) gravimetric moisture content, and (f) degree of saturation

Figure 8 shows results from the Shapley Additive Explanations (SHAP) analysis, which uses game theory to explain the importance of various input features in machine learning models (Lundberg and Lee, 2017). The SHAP analysis was used to understand wavelength importance for the V-S2 model, where greater SHAP values correspond to greater influence on the model output for a given wavelength. Results show the predictions of percent sand and silt are strongly influenced by wavelengths from 2,100 – 2,400 nm. For the prediction of percent clay, solids content, gravimetric moisture content, and degree of saturation, most wavelengths appear to have moderate influence on the prediction of these properties.

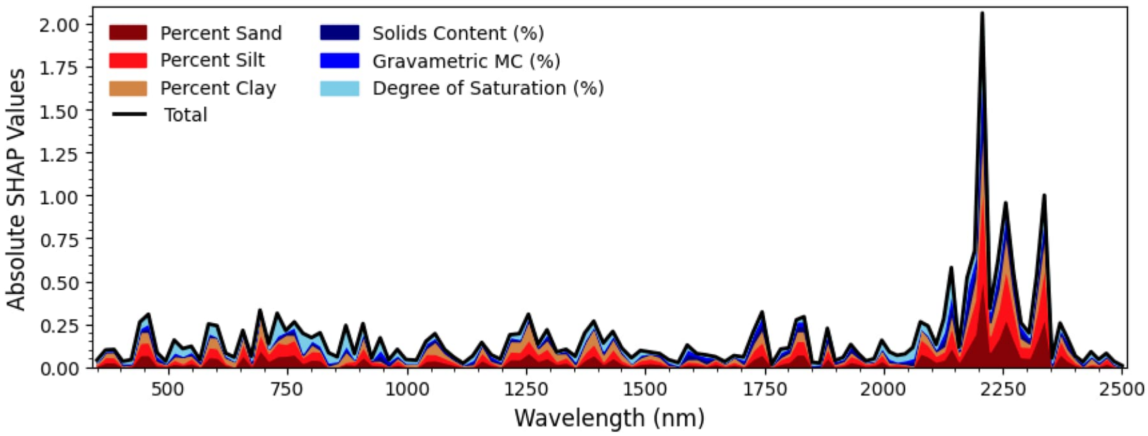

Figure 8: Absolute SHAP values for the V-S2 model where greater values correspond to greater influence of a wavelength on the model outputs. Cumulative absolute SHAP values are shown by the solid back line

Discussion

The model performance demonstrates that hyperspectral sensing can be used to estimate tailings properties including percent sand, silt, clay, solids content, moisture content, and saturation. Tsakiridis et al. (2020) used machine learning and laboratory hyperspectral reflectance to predict percent sand, silt and clay from hyperspectral data on 18,000 natural soil samples and achieved RMSEs of 12.0%, 9.3%, and 4.8% respectively. In comparison, the V-S1 and V-S2 models predictions of tailings sand, silt, and clay have lower RMSEs than Tsakiridis et al. (2020), possibly due to the simplified mineralogical composition of the tailings test specimens, which were derived from a single bulk sample. The coefficients of determination for predicted sand, silt, and clay percentages for the V-S1 and V-S2 models exceed 0.92, highlighting the capacity of hyperspectral reflectance to capture variability in particle size distribution.

For the prediction of tailings moisture content, the model demonstrates high accuracy, which is supported by the strong correlation between soil moisture and hyperspectral reflectance (van der Meer, 2004). Entezari et al. (2022) used hyperspectral sensing and machine learning to predict solids content and gravimetric moisture content of oil-sands tailings, yielding results comparable to those from the V-S1 and V-S2 models. While prior research has not explored the use of hyperspectral sensing for predicting tailings saturation, the model results highlight the potential of hyperspectral sensing for the rapid characterization of tailings saturation.

The strong influence of wavelengths in the SWIR-2 sensor range on the predictions of percent sand and silt support the improved performance of the V-S2 model compared to the V-S1 model for these properties. However, for predictions of percent clay, solids content, moisture content, and degree of saturation, predictive accuracy between the two models did not vary substantially. Omission of the SWIR-2 sensor may result in cost savings for some applications. However, the potential implications of this on predictive accuracy should be considered.

Hyperspectral sensing shows promise for the rapid assessment of tailings properties in field settings. Recent research has used visible light cameras on CPT probes for enhanced soil profile characterization (Ventola et al., 2020). Additionally, hyperspectral sensing has been used in subsurface probes to analyze agricultural soil properties (Cho et al., 2017; Pei et al., 2018). Combining data from multiple sensors, such as hyperspectral sensing, current generation sensors, and the CPT, is anticipated to enhance the accuracy of predictive models for tailings properties (Pei et al., 2018; Riese and Keller, 2019). Future research should investigate the application of probe-based hyperspectral sensing for real-time characterization of in-situ tailings properties.

Conclusion

The objective of this study was to assess the use of hyperspectral sensing for the prediction of various tailings properties. Tailings from a bulk sample were processed to create a diverse tailings-hyperspectral dataset with 300 unique test specimens. Hyperspectral reflectance was captured for each test specimen and geotechnical properties were measured. The dataset was used to construct, train, and test a convolutional neural network for the prediction of percent sand, silt, clay, solids content, gravimetric moisture content, and degree of saturation. The model was evaluated using two scenarios: (1) using only VNIR and SWIR-1 sensor data (V-S1 model), and (2) using VNIR, SWIR-1, and SWIR-2 sensor data (V-S2 model). Based on the findings, the following conclusions were reached:

1. Hyperspectral sensing and machine learning show promise for the prediction of tailings percent sand, silt, and clay. Based on SHAP analysis, the V-S2 model is strongly influenced by SWIR-2 data for the prediction of percent sand and silt, which aligns with the reduced error in the V-S2 model results, compared to the V-S1 model for the properties. Both models have RMSEs less than 9% and R^2 values greater than 0.92 for the prediction of sand, silt, and clay.
2. Predictions of solids content, gravimetric moisture content, and degree of saturation using hyperspectral reflectance shows promise. The V-S1 and V-S2 models performed similarly for the predictions of these properties and have RMSEs less than 9% and R^2 values greater than 0.94.

Hyperspectral sensing can rapidly produce data related to tailings properties and demonstrates the potential for the high-speed characterization of tailings properties for various applications. Including data from hyperspectral sensing and other sensors in machine learning procedures is expected to improve tailings property predictions. Future studies should consider applying similar methodologies in situ for the characterization of tailings and explore how this methodology applies to tailings from other sites.

References

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

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. and 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.

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.

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

Chollet, F. 2015. Keras. GitHub. Accessed at https://github.com/fchollet/keras

Entezari, I., McGowan, D. and Glavina, J. 2022. Hyperspectral imaging technology for oil sands tailings characterization: practical aspects. *Proceedings of Tailings and Mine Waste* 26: 1243–1252.

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. *Proceedings of the Eighth International Conference on Case Histories in Geotechnical Engineering.* Accessed at https://doi.org/10.1061/9780784482148.005.

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.

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.

Lundberg, S., Lee, S. 2017. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems* 30. Accessed at https://proceedings.neurips.cc/paper_files/paper/2017

Minasny, B. and McBratney, A.B. 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computer Geoscience* 32:1 378–1388. https:// doi.org/10.1016/j.cageo.2005.12.009

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. Accessed at 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. *Sensors* 19(1011). http://doi.org/10.3390/s19051011.

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.

Rodríguez, R., Muñoz-Moreno, A., Caparrós, A., García-García, C., Brime-Barrios, A., Arranz-González, J., Rodríguez-Gómez, V., Fernández-Naranjo, F. and 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

Tsakiridis, N.L., 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. https://doi.org/10.1016/j.geoderma.2020.114208.

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

Ventola, A., Dolling and R., Hryciw, R. 2020. Image analysis and hardware development for the vision cone penetrometer (VisCPT). *Proceedings of Geo-Congress 2020*. https://doi.org/10.1061/9780784482803.068.