CPTu-Based Spatial Variability Assessment of Thickened and Conventional Mine Tailings

Jorge Macedo, M.ASCE[1]; Luis Vergaray[2]; Chenying Liu[3]; James Sharp[4]; Kimberly Finke Morrison[5]; and Brett Byler[6]

Abstract: The Global Industry Standard on Tailings Management (GISTM) promotes performance-based approaches in geotechnical assessments. Hence, characterizing the spatial variability of deposited tailings is expected to be a key input for some tailings storage facilities (TSFs); however, it has seldom been investigated. In this study, we assess the spatial variability of thickened and conventional tailings, which have been deposited into the same TSF, providing a unique opportunity to investigate two tailings technologies. A dense array of 15 cone penetration tests (CPTus) with an average offset of 1.5 m has been conducted to collect data. In addition to evaluating the spatial variability, the collected information is also used to assess the potential of machine learning (ML) for detrending when deriving random fields. Using a new proposed stationarity score, we find that an ML-based detrending outperforms traditional procedures for most scenarios. In terms of correlation lengths, we find similar ranges for thickened and conventional tailings (vertical: $\delta_{wv} = 0.2$–0.6 m, horizontal $\delta_{wh} = 1.5$–4.5 m) and similar distributions, likely influenced by the depositional processes. In contrast, the variance in the conventional tailings is higher, which we attribute to its segregating nature. Finally, by inspecting previous studies on natural soils, we find that the variability of mine tailings ($\delta_{wh}/\delta_{wv} = 2$–$21$) resembles that observed in alluvial deposits, which we attribute to the parallels in the depositional processes. **DOI: 10.1061/JGGEFK.GTENG-11969.** © *2024 American Society of Civil Engineers.*

Author keywords: Spatial variability; Mine tailings; Random fields; Machine learning; Cone penetration testing.

Introduction

Mine tailings are deposited within tailings storage facilities (TSFs) to prevent their release into the environment. Historically, different deposition methods associated with different technologies have been used. Depending on the adopted technology, mine tailings can often be classified as conventional, thickened, paste, and filtered (Vick 1990). Conventional tailings are associated with relatively low solids content (25%–40%), thickened tailings are produced in a thickener plant and have solids content in the range from 40% to 65%, paste tailings have even larger solids content (65%–80%), and filtered tailings are often associated with solids contents larger than 80% (Cacciuttolo and Marinovic 2022). Additional details on these technologies can be found in Morrison (2022). The conventional deposition is associated with spigot systems that are used to discharge tailings into a TSF; hence, the mine tailings are deposited as a slurry. This technology often allows for lower costs and higher production rates than other technologies, making it common, especially for mines with high production rates (Watson et al. 2010). Thickened tailings are produced in thickening plants and pumped into the TSF by displacement pumps; this technology was first used in 1973 (Robinsky 1979) and has been typically limited to low to moderate production rate operations. Due to the contrasting solid contents, conventional and high-density thickened tailings are expected to have different properties. For instance, conventional tailings form large ponds and are expected to develop more segregation (Watson et al. 2010). Moreover, due to the processes involved in the deposition of conventional and thickened tailings within a TSF, the spatial variability of geotechnical properties is expected to be a relevant factor; however, it is not commonly addressed. In this context, this study is focused on investigating the spatial variability of conventional and high-density thickened tailings, hereinafter referred to simply as thickened tailings, using cone penetration testing (CPTu), taking advantage of data collected from a TSF where the two types of tailings are deposited.

Previous efforts on characterizing the spatial variability of soil deposits have been mainly focused on natural soils, often using the random field theory (RFT) (e.g., Jaksa et al. 1999, 2005; Breysse et al. 2005; Stuedlein et al. 2012; Fenton and Griffiths 2005; Griffiths et al. 2006; Hu and Ching 2015; Bong and Stuedlein 2017; Stuedlein and Bong 2017; Cai et al. 2017; Ching et al. 2018; Cary 2021). In terms of man-made waste materials, we are only aware of the Bagińska et al. (2016) study, which characterized the spatial variability of a mine waste dump in a coal mine using CPTu. However, they focused only on the vertical variability due to the large spacing between the CPTus. Regarding mine tailings, to the best of the author's knowledge, there have been no previous efforts to characterize the extent of their spatial variability at a local scale

[1]Associate Professor, Frederick Olmsted Early Career Professorship, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332 (corresponding author). ORCID: https://orcid.org/0000-0002-0457-4824. Email: jorge.macedo@gatech.edu

[2]Postdoctoral Research Fellow, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332. ORCID: https://orcid.org/0000-0003-4230-0768. Email: luis.vergaray@gatech.edu

[3]Graduate Student, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332. ORCID: https://orcid.org/0000-0001-5183-4265. Email: cliu662@gatech.edu

[4]CEO, ConeTec Group, 201 - 8327 Eastlake Dr., Burnaby, BC, Canada V5A 4W2. Email: jsharp@conetec.com

[5]Senior Director, Global Tailings Management, Newmont, 6900 E. Layton Ave., Suite 700, Denver, CO 80237. Email: kim.morrison@newmont.com

[6]Director, Tailings and Dams, Newmont, 6900 E. Layton Ave., Suite 700, Denver, CO 80237. Email: brett.byler@newmont.com

Note. This manuscript was submitted on May 27, 2023; approved on February 20, 2024; published online on July 27, 2024. Discussion period open until December 27, 2024; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Geotechnical and Geoenvironmental Engineering***, © ASCE, ISSN 1090-0241.

considering vertical and horizontal directions, which is one of the key contributions of our study. Advancing the spatial variability characterization of mine tailings is a step forward to enable better treatment of uncertainties in performance-based assessments, a framework recommended in the Global Industry Standard on Tailings Management (International Council on Mining and Metals 2020). In addition, gaining insights into the spatial variability of mine tailings provides valuable information to improve the planning of site investigation programs for TSFs. The spatial variability assessments of this study rely on data collected on conventional and thickened tailings using a dense array of CPTus. In addition to the traditional use of polynomial functions in the detrending process to assess the extent of spatial variability, we discuss the potential of taking advantage of modern machine learning-based procedures for trend removal, which is another contribution from our study. In contrast to traditional polynomial functions with limited flexibility due to their fixed forms, machine learning-based methods have greater flexibility in capturing trends in CPTu data.

This study is organized as follows: After a brief introduction, details of the characterized site and conducted CPTus are provided. Then, we discuss the methods considered to assess the spatial variability of mine tailings, including the role of machine learning on data detrending. The results considering the different approaches are then presented, including vertical and horizontal correlation lengths and other statistical assessments. We then discuss the results obtained for the conventional and thickened tailings in the context of previous studies for natural soils and the potential role of deposition processes. Finally, the paper closes with conclusions and recommendations for future work.

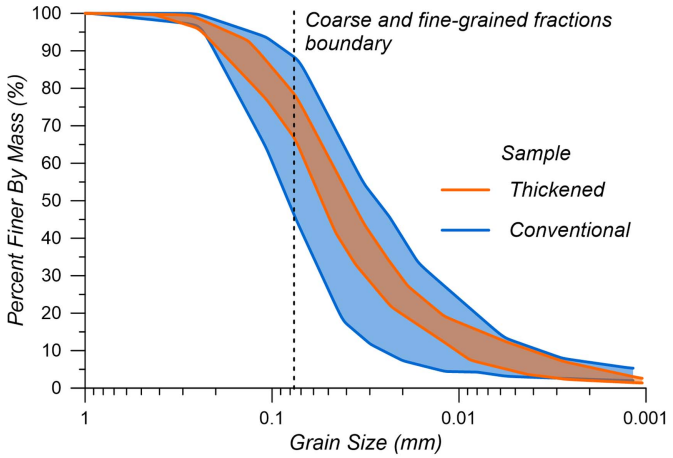

Fig. 1. Representative particle size distributions of conventional and thickened tailings.

Site Characterization

The investigated tailings started to be deposited back in the 1990s using conventional deposition until 2010, when the deposition method was changed to high-density thickened tailings, providing a unique opportunity to investigate both conventional and thickened tailings in the same TSF. The solids contents for the conventional and high-density thickened tailings are about 50% and 62%–70% (the variation is due to seasonal change from winter to summer), respectively. Fig. 1 illustrates the range of particle size distributions of both tailings. The field characterization was planned to collect data on both the conventional and thickened tailings and included 15 CPTu conducted with a compression-type piezocone with a diameter of $d = 4.37$ cm (i.e., cone area $a = 15$ cm^2). Conventional strain-gauged load cells record the tip (q_t) and friction resistance (f_s), whereas a strain-gauged transducer records dynamic and static pore pressure in the u_2 position. In addition, vertical and horizontal geophone sensors are included in the body of the cone to assess shear and compression wave velocities. The CPTu array was oriented parallel to the discharge points in the dam, as illustrated by Fig. 2(a). The CPTus were conducted at a standard penetration rate of 2 cm/s and were spaced at a distance of 1.5 m, as shown in Fig. 2(b). The separation was selected following the recommendations of Cary et al. (2022) to minimize any potential disturbance between nearby CPTus. The central CPTu [$X = 0.0$ m in Fig. 2(b)] was pushed to a depth of 24 m, and complemented with pore pressure dissipation tests every meter. The dissipation tests confirmed a phreatic surface approximately 6.1 m below the surface. The remaining CPTus were pushed to a depth of 16 m.

Figs. 3(a–c) show q_t, f_s, and u_2 profiles, soil layering, and the phreatic surface inferred from the CPTu tests. The inferred soil profile can be divided into three layers. The layer from the surface to 2 m is a pad access built to access the site, which is not considered in the interpretations. The second layer, extending from 2.0 to about

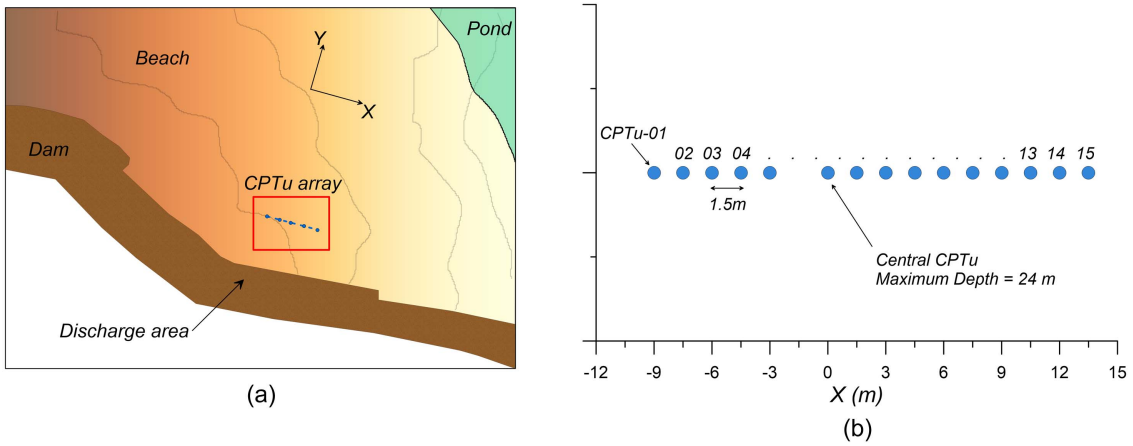

Fig. 2. (a) CPTu array location relative to the discharge area; and (b) CPTu array zoom-in and labels.

J. Geotech. Geoenviron. Eng., 2024, 150(10): 04024091

8.5 m, corresponds to the thickened tailings showing relatively uniform q_t, f_s, and a linear trend in u_2. The conventional tailings underlay the thickened tailings and extend to the bottom of the profile, showing a more erratic variation of q_t, f_s, and u_2 values. The separation of the conventional and thickened tailings at about 8.5 m is consistent with the tailing's deposition history. The thickened tailings present relatively low q_t (i.e., 1 to 5 MPa) and f_s (i.e., 5 to 40 kPa). The conventional tailings exhibit q_t values between 1 and 8 MPa and f_s values between 5 and 70 kPa. The u_2 values are negligible above 4 m, positive in most CPTus for depths between 6.1 and 8.5 m within the thickened tailings, and tend to be hydrostatic in the conventional tailings. The observed vertical patterns for the different CPTus suggest some degree of homogeneity in the horizontal direction for a given depth (Fig. 3). However, these qualitative interpretations may be subject to judgment, and quantifying the variability based on statistical procedures should be preferred, as demonstrated in later sections.

Fig. 4(a) uses the central CPTu to illustrate the typical variation of the soil behavior type (I_c), proposed by Robertson (2016), which is estimated as $I_c = \sqrt{[(3.47 - \log Q_{tn})^2 + (\log F_r + 1.22)^2]}$, where $Q_{tn} = [q_t - \sigma_v/P_a](P_a/\sigma'_v)^n$, $F_r = [f_s/(q_t - \sigma_v)] \times 100\%$, σ_v, σ'_v, and P_a are the total in situ stress, effective vertical stress, and the atmospheric pressure, respectively, and $n = 0.381I_c + 0.05(\sigma'_v/P_a) - 0.15$. Based on the expected suction levels from representative soil water characteristic curves for the examined tailings up to 50 kPa with a contribution to the effective stresses up to 25 kPa (considering χ, the effective stress parameter, as 0.5), the effect of matric suction on I_c was estimated not to be significant; hence, I_c in Fig. 4(a) is estimated without considering suction above the phreatic surface.

Fig. 4(b) shows I_c histograms to aid in visualizing the I_c variation in the two tailings types. The thickened tailings exhibit I_c values from $I_c = 1.8$ to 2.6, suggesting they behave as sandy silt and silty sand mixtures based on Robertson (2016). Consistent with their more segregating nature, the conventional tailings exhibit higher I_c values [up to 3.1, as noted in Figs. 4(a and c)]. According to Robertson (2016), the conventional tailings are expected to behave mainly in the range of sands to silty sands and silty clays. Fig. 4(c) shows the CPTus plotted in the Robertson (2016) classification chart. Using Robertson's (2016) terminology, the thickened tailings show a sand-like and transitional contractive behavior for the bulk of the data. The conventional tailings exhibit a broader range of behaviors, with a significant fraction in the zone of transitional materials. In addition, the conventional tailings exhibit a more balanced partition between the contractive/dilative regions than the thickened tailings; however, they are predominantly contractive. The more contractive nature of the thickened tailings is likely associated with their lower depositional energy compared to the conventional tailings. The conventional and thickened tailings are deposited using subaerial deposition; however, the amount of water in the conventional tailings is higher, likely creating a higher deposition energy. As highlighted by Reid and Jefferies (2018), depositional energy is a key factor in dictating the contractive nature of deposited mine tailings. The observations just discussed in the context of different depositional environments are consistent with Reid and Jefferies's (2017) study for conventional tailings and Reid and Jefferies's (2018) study for thickened tailings. Another factor that might have contributed to the more contractive nature of the thickened tailings is their more recent deposition (thickened tailings are still being deposited), as consolidation may not have been completed due to large drainage paths promoted by their non-segregating nature; however, this likely applies only to the first few meters.

Spatial Variability Assessment

Random Field Theory and Spatial Soil Variability

The inherent spatial variability of soils can be modeled with the RFT theory, which has been used in several previous studies mainly focused on natural soils (e.g., Vanmarcke 1977; DeGroot and Baecher 1993; Lacasse and Nadim 1996; Phoon and Kulhawy 1999; Fenton and Griffiths 2005; Jaksa et al. 2005; Griffiths et al. 2006; Breysse et al. 2005; Hu and Ching 2015; Stuedlein and Bong 2017; Cary 2021). In the following, we provide a brief overview of the RFT framework and its use in assessing spatial soil variability. Interested readers can refer to Jaksa et al. (1997), Stuedlein (2008), and Bong and Stuedlein (2017) for additional details. Under the RFT framework, a measurement of some spatially varying soil property of interest $g(z)$ at a given depth z can be expressed as the sum of a deterministic trend $[t(z)]$ and a fluctuating component $[w(z)]$ as expressed by Eq. (1) and schematically presented in Fig. 5. The spatially varying parameter of interest is then characterized by the correlation length (δ), and the coefficient of variation (COV) of $w(z)$. δ is a measure of the distance within which a particular measurement is correlated, and COV is the standard deviation (σ) of $w(z)$ normalized by the mean of $t(z)$

$$g(z) = t(z) + w(z) \tag{1}$$

δ can be calculated by different geostatistical procedures (e.g., Jaksa 1995; Vanmarcke 1977; DeGroot and Baecher 1993). In particular, the use of autocorrelation models to fit the autocorrelation in the observed data is common in geotechnical applications. In this procedure, the sample autocorrelation function $\rho(\tau_j)$ – i.e., based on the observed data, which defines the spatial correlation for a given lag distance (τ_j) on the measurement of interest, is evaluated according to Eq. (2)

$$\rho(\tau_j) = \frac{\sum_{i=1}^{n_d - j} w_i \times w_{i+j}}{\sum_{i=1}^{n_d - j} w_i^2} \tag{2}$$

As suggested by Box and Jenkins (1970), different lag distances τ_j can be considered by varying the sampling space (Δz) for $j = 1, 2, 3 \ldots n_d/4$, where n_d is the sample size. Once $\rho(\tau_j)$ is assessed, δ can be estimated by fitting different autocorrelation models (ACMs) to $\rho(\tau_j)$. The ACMs considered in this study (Table 1) are the single exponential (SNX), cosine exponential (CSX), second-order Markov (SMK), and squared exponential (SQX) models. These ACMs have been used in geotechnical applications (e.g., Phoon et al. 2003; Uzielli et al. 2005).

In using the RFT framework, it is highly beneficial to have stationarity in w, i.e., the covariance structure only depends on the distance between observations (τ). Thus, a key step in assessing spatial soil variability is detrending the data [i.e., the removal of $t(z)$], ideally leading to a stationary $w(z)$ random field. The selection of the functional form $t(z)$ is key to detrending, as discussed in more detail in the next section. Once the detrending is conducted, there are several methods for assessing stationarity conditions [e.g., Spearman's rank coefficient–(Spearman 1904); Kendall's tau test–(Kendall 1938); Bartlett's test–(Bartlett 1937); Phoon et al. 2003)]. In particular, the Kendall and Bartlett tests have been used in recent studies for natural soils (e.g., Bong and Stuedlein 2017; Cary 2021). Uzielli et al. (2005) and Stuedlein (2011) compared the Kendall and Bartlett tests, concluding that the Bartlett test is more stringent and discriminative in assessing stationarity. In this context, in this study, we use the Bartlett test as the basis to define a stationarity score, as discussed in subsequent sections.

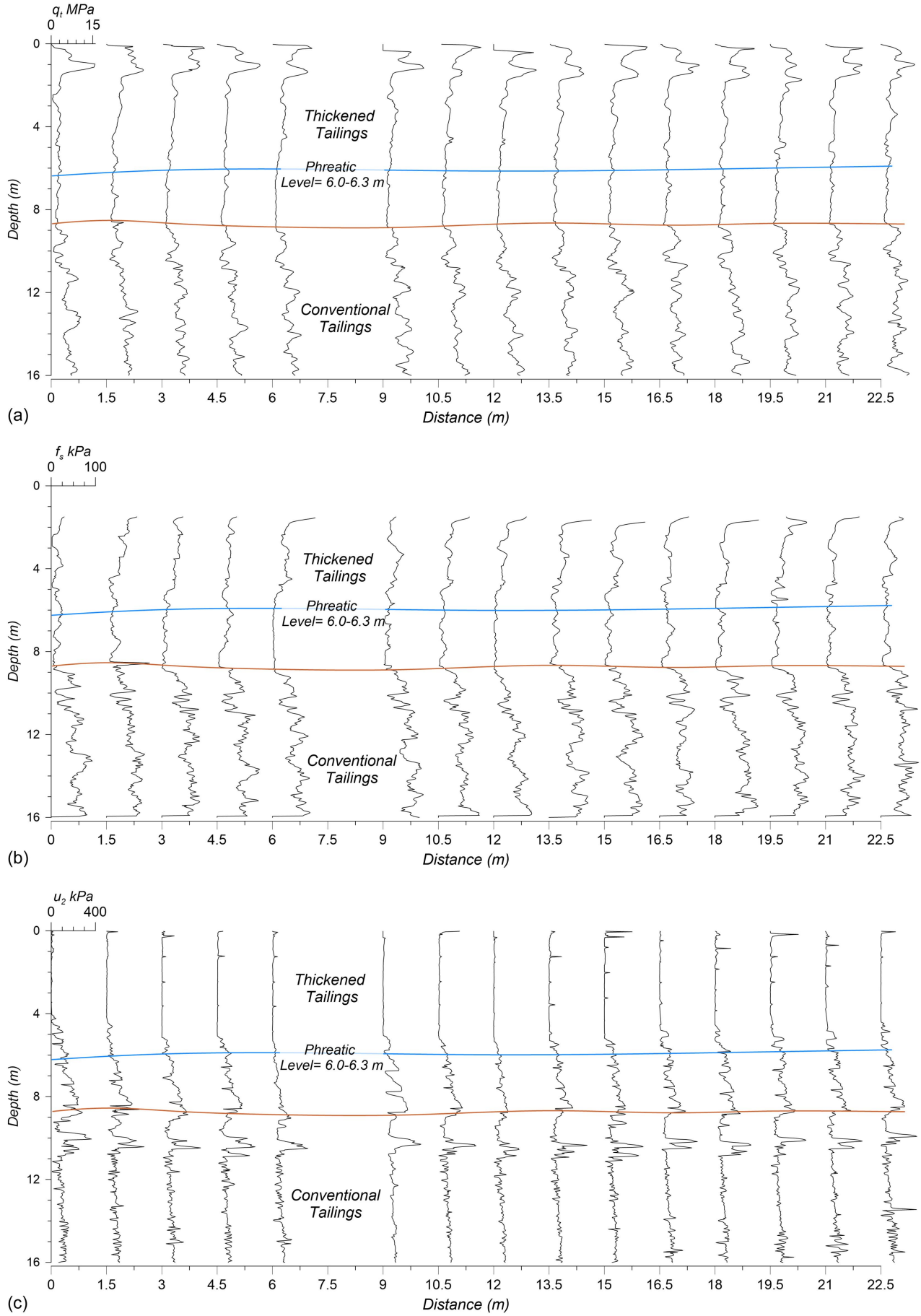

Fig. 3. (a) q_t; (b) f_s; and (c) u_2 CPTu profiles.

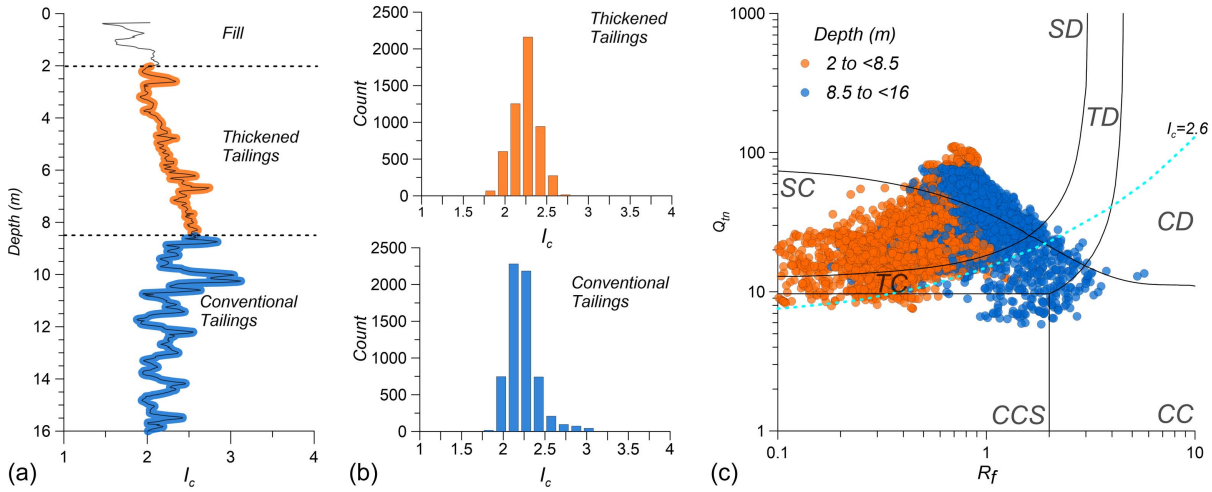

Fig. 4. (a) Representative I_c profile varying with the depth; (b) histograms of I_c; and (c) soil behavior type (SBTn) chart of thickened and conventional tailings obtained from all CPTus. The SBTn zones are defined as SC = sand-like contractive; SD = sand-like dilative; TC = transitional-like contractive; TD = transitional-like dilative; CC = clay-like contractive; CD = clay-like dilative; and CCS = clay-like contractive sensitive.

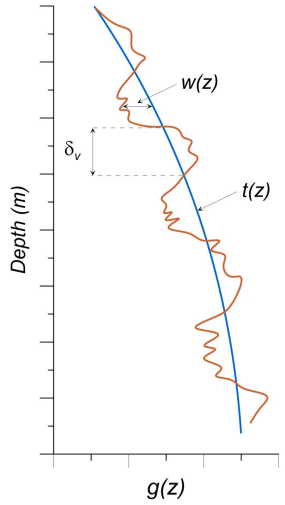

Fig. 5. Sketch of spatially varying soil property $g(z)$, the trend $t(z)$, the fluctuating component $w(z)$, and the correlation length (δ). (Adapted from Phoon et al. 2003.)

Table 1. Autocorrelation models and corresponding correlation length

Autocorrelation model	Functional form	Correlation length
SNX	$\rho(\tau) = e^{-k\|\tau\|}$	$\delta = 2/k$
CSX	$\rho(\tau) = e^{-k\|\tau\|}\cos k\tau$	$\delta = 1/k$
SMK	$\rho(\tau) = (1 + k\|\tau\|)e^{-k\|\tau\|}$	$\delta = 4/k$
SQX	$\rho(\tau) = e^{-(kt)^2}$	$\delta = \sqrt{\pi}/k$

Source: Adapted from Uzielli et al. (2005).

Machine Learning-Based Methods for Data Detrending and Vertical Correlation Length Estimation

The current practice commonly considers polynomial functions in the detrending of CPTu data. For instance, a 1st order polynomial function is typically used to remove trends, and then statistical tests are performed to check the stationarity of the residuals. If the residuals fail the stationarity check, one must increase the order of the polynomial function and repeat the calculation until the residuals are stationary. Often, polynomials up to degree three have been used (Cary 2021). A limitation of using polynomial-based trend functions is their restricted flexibility. A polynomial function can only represent trends that belong to the polynomial family and may not be efficient in capturing trends in CPTu measurements, potentially producing nonstationarity. In this context, modern machine learning-based methods can be advantageous as they offer more flexibility (e.g., nonparametric-based procedures) than polynomials with fixed functional forms. In addition, the potential use of various ML algorithms can also enable better treatment of epistemic uncertainties in the estimation of random field parameters, e.g., by providing alternative correlation lengths. In this study, we have selected three nonparametric ML methods, namely, random forest (Breiman 2001), gradient boosting decision tree (GBDT) (Friedman 2001), and the k-nearest neighbor (KNN) (Peterson 2009) for conducting the CPTu data detrending and assessing the potential of machine learning on informing the assessment of spatial variability. In the following, we briefly describe these methods and illustrate how they have been used in this study. Interested readers are referred to Bishop (2006), Breiman (2001), Friedman (2001), and Peterson (2009) for more details on these methods. The parameters that control the behavior of an ML method are often referred to as hyperparameters and need to be tuned during the training process. In the case of random forest, the considered hyperparameters are the tree depth (how deep the tree can grow), the minimum samples (required to split a tree node), the minimum samples required to form a leaf (often referred to as the subspace), and the number of trees. In the case of GBDT, the considered hyperparameters are the tree depth (how deep the tree can grow), learning rate (the optimization speed of parameters in the model), row subsample ratio (the subsampling ratio of the data to train the tree), number of trees, and the minimum child weight (minimum number of data to form a leaf or subspace). Finally, in the KNN method, k, defining the number of the nearest neighbors, is considered to be the hyperparameter. Standard machine learning literature provides detailed information on these hyperparameters; interested readers can refer to Bishop (2006). The hyperparameters were optimized during the training process using a grid search strategy (Bishop 2006); the considered values for training are listed in the Supplemental Materials.

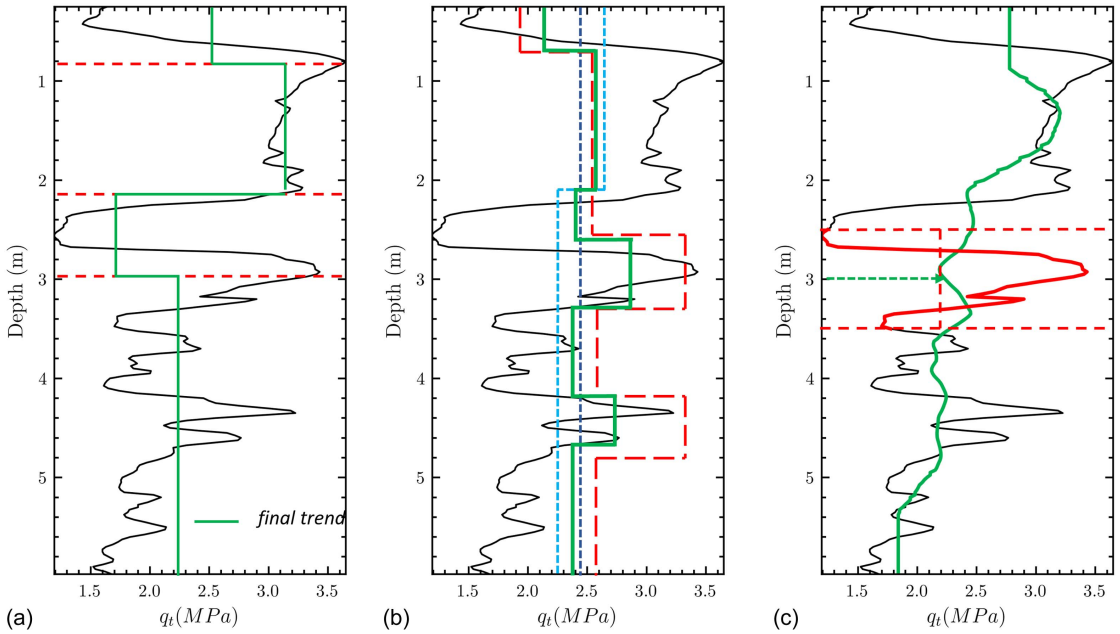

Fig. 6. Schematic illustration of the (a) random forest; (b) gradient boosting decision tree (dark blue line: constant trend, light blue line: trend based on the residuals of the constant trend, red line: trend based on further updated residuals, green line: final trend function that combines previous fittings); and (c) k-nearest neighbor (green curve: trend function, horizontal red lines: interval considered for a prediction at 3 m) models used to remove the trends in a CPTu profile.

Random forest is an ensemble learning method that combines the prediction of multiple estimators known as decision trees (Breiman 2001). Random forest uses a modified bootstrap aggregating algorithm to generate multiple tree learners. For each tree learner, a fraction of the CPTu measurement is randomly sampled with replacement (bootstrapping) for training. Then, the prediction results are averaged over all tree learners (aggregating). For one-dimensional data (e.g., q_t or f_s), the main hyperparameter that controls the complexity of the trend function derived from random forest is the maximum tree depth, which controls how deep each tree can grow. Trees with larger depths will lead to more flexible trend functions but also have more potential to overfit the data (Breiman 2001), i.e., the derived model could perform well for the training data but not for other datasets. Fig. 6(a) shows a schematic illustration of building one tree in the random forest algorithm using CPTu 14 and q_t. The q_t data is first automatically split by the algorithm into four subregions with splitting boundaries at 0.8, 2.2, and 3.0 m. These regions are subspaces created by the algorithm and do not imply any stratification. Then, the average value of q_t in each subregion is used for prediction. This process is repeated multiple times—each repetition represents a *tree*—and the predictions across all the trees are averaged to generate the final prediction, which is represented by the piecewise linear trend.

GBDT is another ensemble learning method. The major difference with respect to random forest is the training procedure. GBDT combines multiple weak tree learners into a single strong learner in an iterative manner. A weak tree learner in GBDT refers to a simple decision tree model with limited predictive capability when considered alone. The GBDT model starts with a weak tree learner, and at each iteration, a new tree learner is added using the residual between the previous learner and the observed response variable. The algorithm terminates when reaching a prescribed maximum number of iterations or when there is no improvement relative to a previous iteration, which in GBDT is implemented through a loss function (Chen and Guestrin 2016). Finally, the predictions from all tree learners are summed up. Similar to random forest, the maximum tree depth is key in controlling the complexity of the generated trend function. The GBDT use on the CPTu data is illustrated in Fig. 6(b) using the same CPTu discussed for random forest. Initially, a simple constant line (i.e., the mean value of q_t) is considered. Based on the residuals between q_t and the constant line, another tree is fitted to the residuals. Then, the new residuals are calculated as the difference between the light blue line and the previous residuals, and another tree is fitted to the new residuals. Finally, all the trees (represented by the previous fittings) are combined to produce the final trend function.

The KNN algorithm builds a prediction model based on the distances between the predictor and its nearest neighboring data points. Specifically, a predicted value is based on the average value of its k-nearest neighbors. The value of k serves as a hyperparameter and controls the complexity of the model. For instance, for $k = 1$, the generated trend function is complex and attempts to include every CPTu measurement. In contrast, when k is equal to the total number of data points, the generated trend function passes through the mean value of the CPTu measurements. Any k value in between will produce continuous trend functions with different complexities controlled by k. Fig. 6(c) schematically illustrates the use of the KNN method. The trend function is generated by averaging the q_t values for the 20 nearest neighboring points at a given depth. For example, the prediction at 3 m is computed by averaging the q_t values in a window from 2.5 to 3.5 m [i.e., as illustrated in Fig. 6(c)]. In this study, we start with a large k value, gradually decreasing it to generate more complex functions until its performance is optimized.

Quantifying the Performance of Trend Removal (Stationarity Score)

In this section, we propose a stationarity score for assessing the performance of different models that describe spatial variability

based on CPTu data. In estimating random field parameters, two factors are of particular relevance: the stationarity of the random field and the fitting of autocorrelation models. The fitting of the autocorrelation models is usually represented by the coefficient of determination (R^2) and considered adequate when R^2 is higher than 0.9 (Phoon et al. 2003). The fitting is conducted for the regions where the empirical autocorrelation function [Eq. (2)] shows values higher than the Bartlett's limit (Phoon et al. 2003) defined as $r_B = 1.96/(n_d)^{0.5}$, where n_d is the number of observations. The stationarity of a random field is evaluated using the modified Bartlett's test. In the test, the Bartlett's statistic is calculated at each depth using the data after detrending, and the maximum value is compared with a Bartlett's critical value (B_{crit}) defined in Phoon et al. (2003). If the maximum value of the Bartlett's statistic is below the critical value, then the random field is considered stationary. However, the Bartlett's test does not quantify the extent of stationarity (or nonstationarity). Thus, we propose a combined stationarity score (S) to quantify the extent of stationarity and the goodness of fit of autocorrelation models, defined in Eq. (3) as:

$$S = \begin{cases} \dfrac{\int_z B_{\text{crit}} - B_{\text{stat}} dz}{B_{\text{crit}} \times z} \times \beta, & \max(B_{\text{stat}}) < B_{\text{crit}} \\ \dfrac{\int_z H(B_{\text{stat}} > \text{B}_{\text{crit}})(B_{\text{stat}} - B_{\text{crit}}) dz}{\max(B_{\text{stat}}) \times z} \times (1-\beta), & \max(B_{\text{stat}}) \geq B_{\text{crit}} \end{cases} \tag{3}$$

$$\beta = \frac{1}{1 + \exp -15(R^2 - 0.5)} \tag{4}$$

where $H(B_{\text{stat}} > B_{\text{crit}})$ = Heaviside function with a value of 1 only if $B_{stat} > B_{crit}$ and 0 otherwise. B_{stat} = calculated Bartlett's statistic at each depth; B_{crit} is the critical value for Bartlett's statistics; and z = depth at which Bartlett's statistic is calculated. β = weighting parameter [Eq. (4)], which is a sigmoid function of R^2 with bounded values from 0 to 1. When R^2 is large, β approaches 1, and when R^2 is small, β approaches 0. Hence, S is a normalized scalar quantity (bounded from −1 to 1), where negative values indicate nonstationary and positive values indicate stationary. S combines the contributions from both R^2 and Bartlett's statistic and can be used as a single metric to evaluate the performance of different trend functions. More negative S values represent more nonstationarity, while more positive values represent more stationarity. S is evaluated using detrended models. In the case of ML-based models, the detrending is conducted by fitting them to a variable of interest, optimizing the ML model hyperparameters using a grid search strategy (Bishop 2006). Then, the residuals between CPTu measurement and the trend functions are used to further calculate S for a given ML model and an autocorrelation model.

Fig. 7 illustrates the process of using S to assess the performance of ML-based models by considering GBDT in detrending q_t data on the thickened tailings collected in CPT04. Maximum tree depths from 1 to 3 are considered in addition to the hyperparameters listed in the Supplemental Materials. In addition, the SMK model is used to fit the autocorrelation in the data. Since GBDT uses decision trees, it generates piecewise functions, as observed in Fig 7(a). Notice how the increase in the tree depth generates more flexible trends. Once the trend functions are generated, the residuals are estimated (i.e., by subtracting the trends from the original data) as illustrated in Fig 7(b) and used to fit the SMK model [Fig. 7(c)]. Figs. 7(d–e) show the estimation of Bartlett's statistics for increasing tree depths and the corresponding R^2 and S estimates. In Fig. 7(d), some nonstationarity is observed for a tree depth of one; then, as the tree depth increases, S and R^2 also increase suggesting a benefit in increasing the GBDT flexibility. Of note, as the tree depth was increased, other hyperparameters were also optimized. Once the tree depth is 3, the S score is close to 1, and the Bartlett's statistics are consistently lower than the critical value. At this stage, one can stop increasing the complexity of the trend functions and select the trend function that leads to the highest S (i.e., the GBDT model with a tree depth of three). In general, the process of finding the optimal ML models used in this study considers varying the hyperparameters and tracking the S values identifying the maximum values that represent more stationarity. Figs. S1 and S2 provide similar examples for random forest and KNN. The same process can also be applied to polynomial models.

Assessment of ML-Based Procedures and Polynomial-Based Models

In this section, we compare the performance of ML and polynomial-based functions in generating stationary data considering both the thickened and conventional tailings. For each CPTu measurement, four different autocorrelation models previously discussed (Table 1) are considered to assess the vertical correlation lengths (δ_{wv}) for both the thickened and conventional tailings. We increase the order of the polynomial functions (up to the fourth order as commonly used in practice) and tune the flexibility of the ML-based trend functions, as illustrated in the previous section. For each autocorrelation model and CPTu, we select the trend functions (i.e., the polynomial-based or ML-based model) that lead to the highest S. First, we discuss the process using a representative CPTu, and then present the results considering all CPTus.

Fig. 8 illustrates the process by showing the results obtained for the q_t profile of CPT-14 for the conventional tailings using the CSX autocorrelation model and polynomial, KNN, random forest, and GBDT detrending methods. The hyperparameters that produce the highest S for each method are noted in the figure titles. In this case, a polynomial function of order two was selected because it gives the highest S. However, S is negative, implying nonstationarity at some depths. In contrast, all ML-based methods produce stationarity (Fig. 8, right column). Hence, the CSX model would be discarded if one only considers traditional polynomials. Using ML-based methods, the extent of potentially useful autocorrelation models is increased, which favors the treatment of uncertainties (i.e., more alternative models to treat epistemic uncertainties can be included). For instance, in this case, the use of the CSX model derives δ_{wv} values from 0.25 to 0.28 m.

Fig. 9 shows similar results, but now considering the q_t profile of CPT-14 for the thickened tailings considering the SMK autocorrelation model. In this case, all the trend functions satisfy the stationarity requirement given by the modified Bartlett's test. However, the polynomial trend function (second order) just passes the stationarity check at an S of zero, whereas the ML-based methods produce random fields that are more stationary (larger S), highlighting a better performance in removing trends in this case.

Figs. 10 and 11 show a summary of the performance of different models (polynomial and ML-based) in terms of their S values for all the CPTus and autocorrelation models previously discussed, considering q_t data. In presenting the results of a model family (e.g., GBDT), only the results for the highest S are considered. Fig. 10 shows the results for the thickened tailings. Out of the four autocorrelation models, SMK and SQX produce more stationarity when combined with different trend functions, as reflected by the larger positive bars in Fig. 12. Regardless of the autocorrelation model, the ML-based trend functions show, in general, a higher S for most CPTus. Specifically, for the SNX and CSX models, the polynomial functions fail stationarity checks for several CPTus

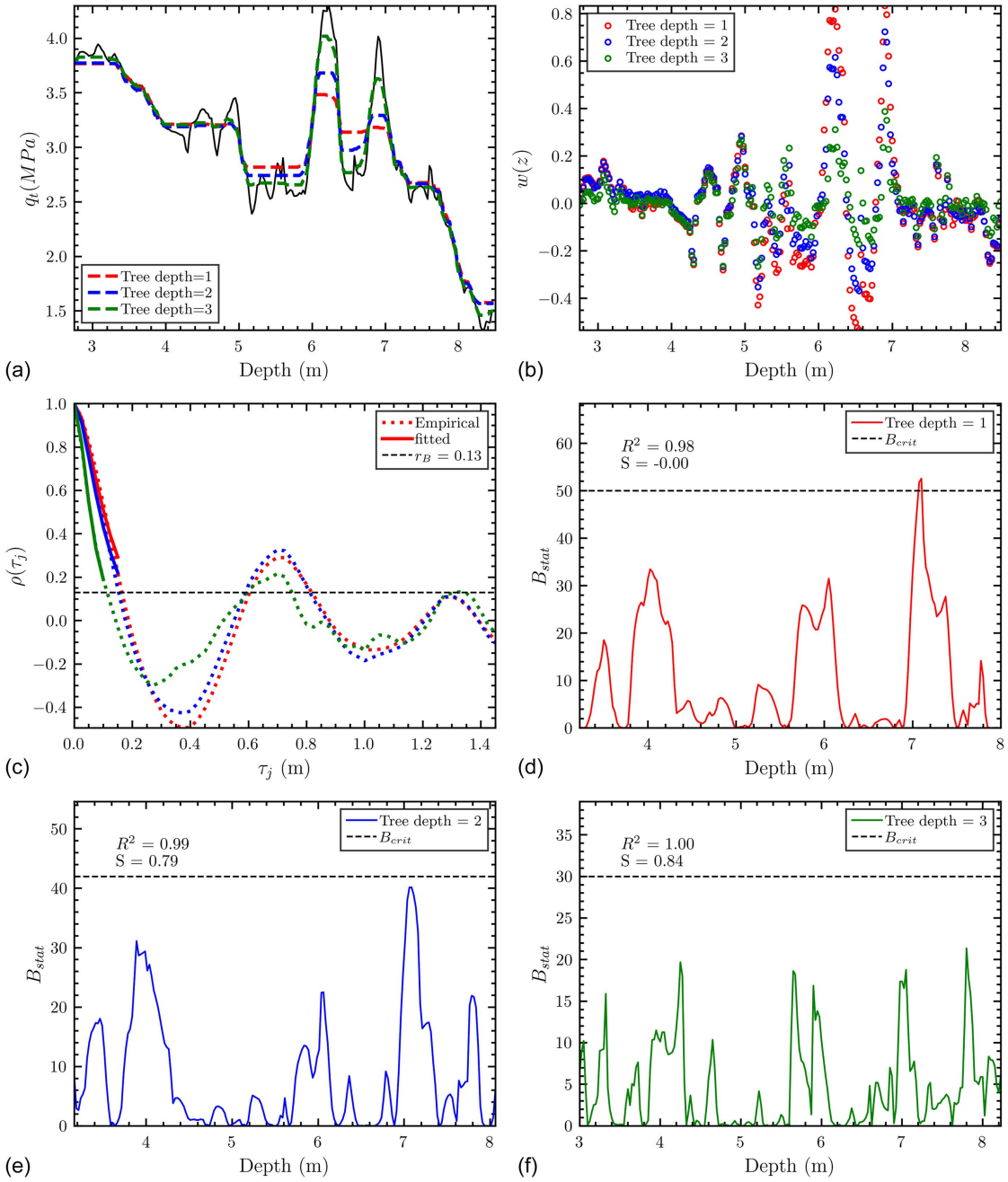

Fig. 7. Example of results obtained using GBDT with variable tree depths considering the test CPT04.

(i.e., negative S). In the case of the SMK model, the traditional and ML-based trend functions have stationary random fields for most CPTus, but the ML-based models show, on average, larger S. In the case of the SQX model, both ML-based and polynomial trend functions show similar performance. It is worth highlighting that there are also a few combinations of CPTus and ACMs where polynomial trend functions perform slightly better than ML-based methods, such as for CPT-12 using the SNX model. In comparing the performance of different trend functions across the four ACMs, the KNN model performs better in the case of thickened tailings, as reflected by the more positive S in Fig. 10.

Fig. 11 shows similar results, but now considering the conventional tailings. In this case, the SQX model shows the best performance, with different trend functions being comparable for this ACM. ML-based models generally provide higher S values compared to polynomials for most combinations. Of note, polynomial models produce negative S values for some cases in which at least one ML-based model provides a positive S (e.g., SNX and CPT-06, SMK and CPT-12), enabling the use of more autocorrelation models as previously discussed. It is also interesting to note that in the case of the CSX model and CPT-06, none of the trend functions derive in full stationarity, suggesting that this model is unsuitable for modeling the CPTu profile's random field. Similar to the thickened tailings, there are also a few combinations where polynomial models produce higher S values than ML-based models (e.g., SQX model and CPT-15). Lastly, in this case, the different ML-based models show a similar performance when combined with different autocorrelation models. Figs. S3 and S4 show the

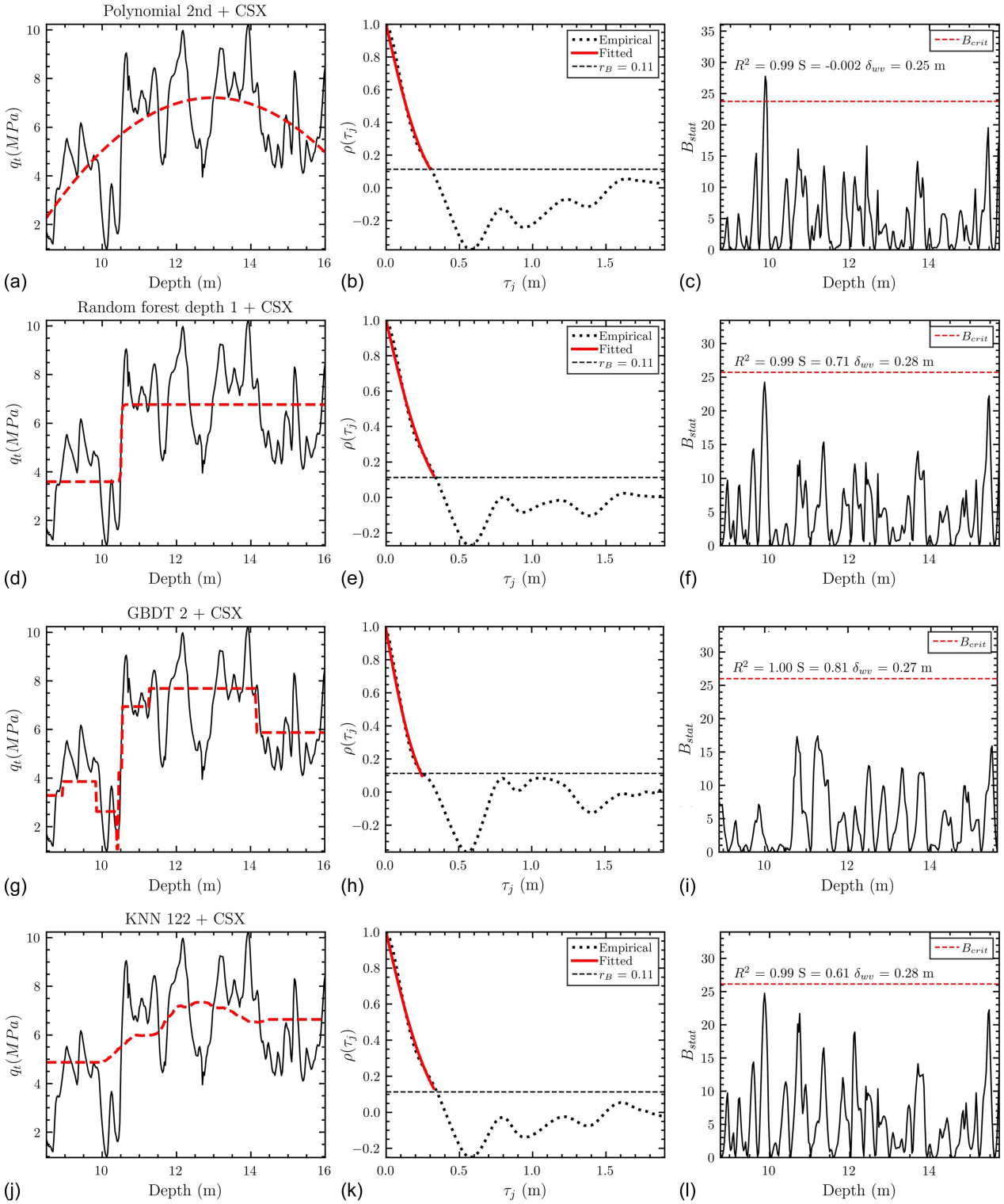

Fig. 8. Results using polynomial and ML-based trend functions for the q_t profile of CPT-14, considering the conventional tailings.

results considering f_s for all CPTus, autocorrelation models, and trend functions (polynomial and ML-based). In both the conventional and thickened tailings, the SQX model performs better. Moreover, the performance of KNN and random forest algorithms is comparable and overall better than that of other trend functions, as they can generate stationary random fields for several CPTus that are regarded as nonstationary by other trend functions. This is the case, in particular, for conventional tailings.

Spatial Variability of Mine Tailings

Vertical Correlation Length

As previously described, δ_{wv} was calculated for the conventional and thickened tailings by fitting four different autocorrelation models (Table 1) to the sample autocorrelation functions, as illustrated in Figs. 8 and 9. Fig. 12 presents the cumulative distribution

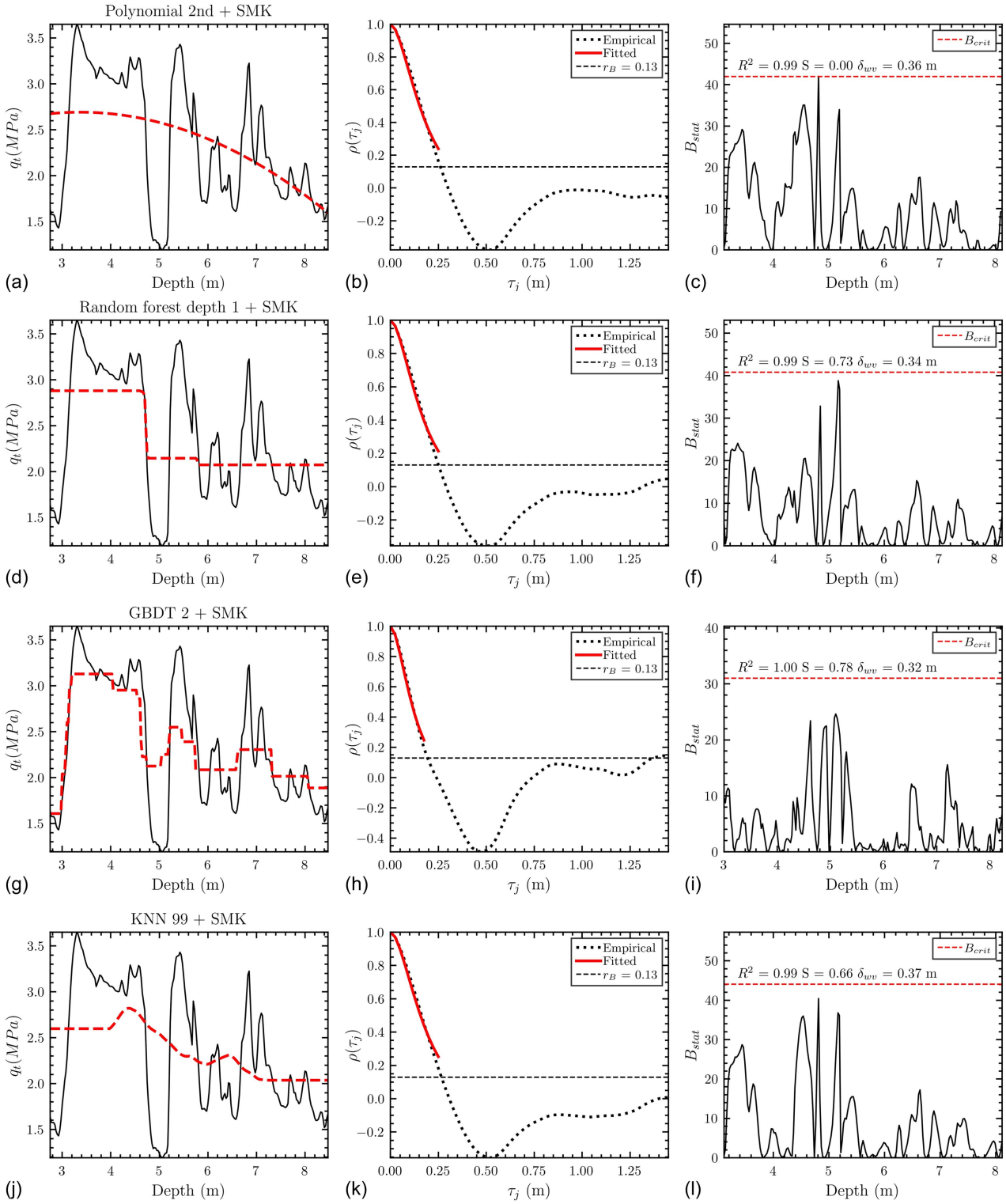

Fig. 9. Results using polynomial and ML-based trend functions for the q_t profile of CPT-14, considering the thickened tailings.

functions (CDFs) of δ_{wv} and COV_{wv} for q_t [Figs. 12(a and b)] and f_s [Figs. 12(c and d)], respectively, considering different methods (i.e., ML-based and polynomials with the highest S) and the CSX autocorrelation model for the thickened and conventional tailings (additional plots for other ACMs are included in Figs. S5–S7). The CDFs calculated with different methods for a given CPTu represent the epistemic uncertainty. Interestingly, the estimated δ_{wv} and COV_{wv} values are in a relatively narrow range. In the case of thickened tailings, the mean δ_{wv} estimated from q_t is 0.33 m, which is smaller, but still comparable with the δ_{wv} of 0.49 m based on f_s. In the conventional tailings, the mean δ_{wv} estimated from q_t is 0.35 m, quite close to the mean δ_{wv} of 0.31 m considering f_s. The q_t-based δ_{wv} CDFs for both tailings are also consistent. The f_s-based δ_{wv} CDFs for the thickened tailings are shifted to the right with respect to the CDFs for conventional tailings, but the δ_{wv} are still comparable.

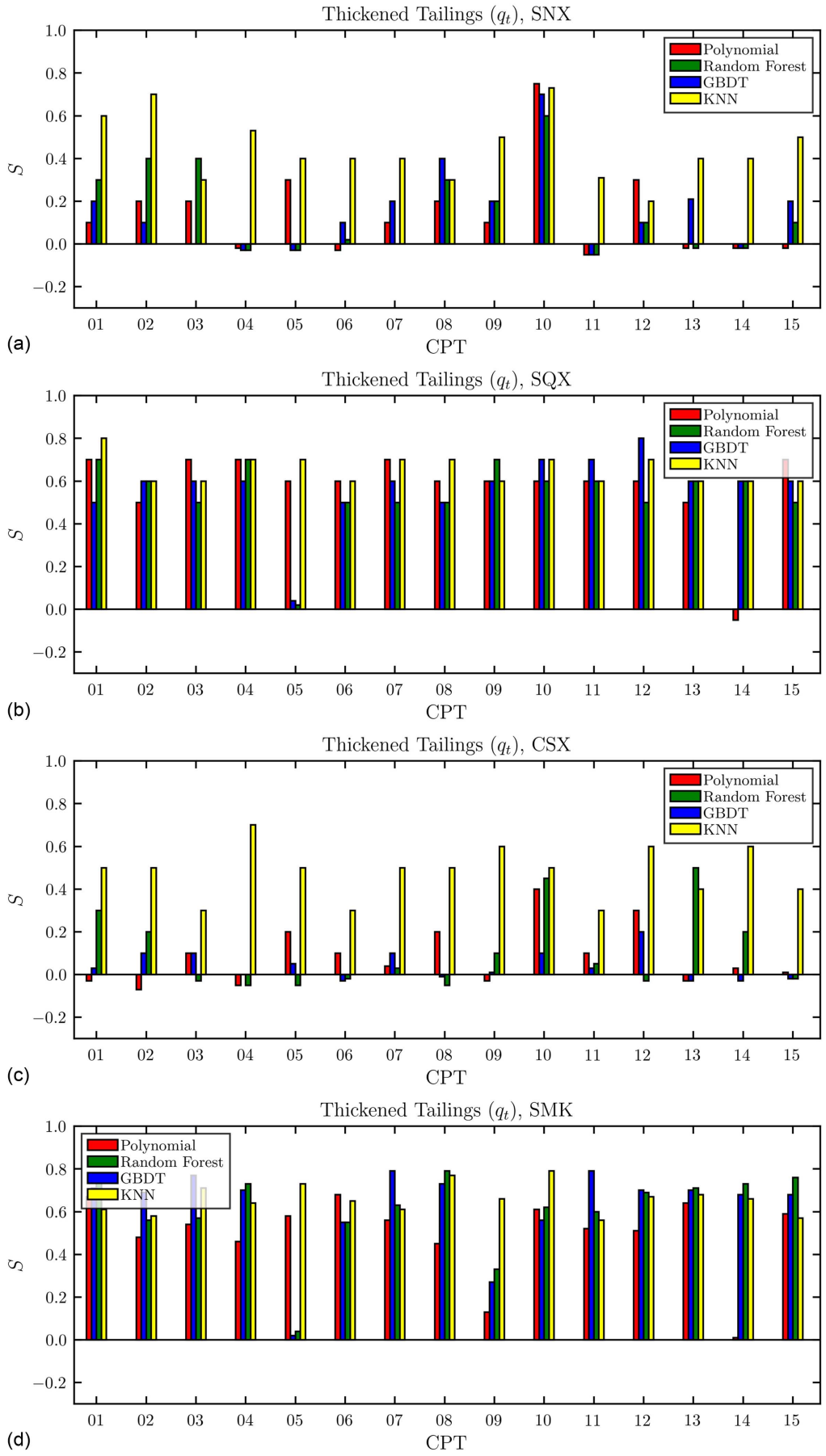

Fig. 10. Comparison of different trend functions (polynomials and ML-based) considering four different ACMs and all CPTus. The results are for q_t in the thickened tailings.

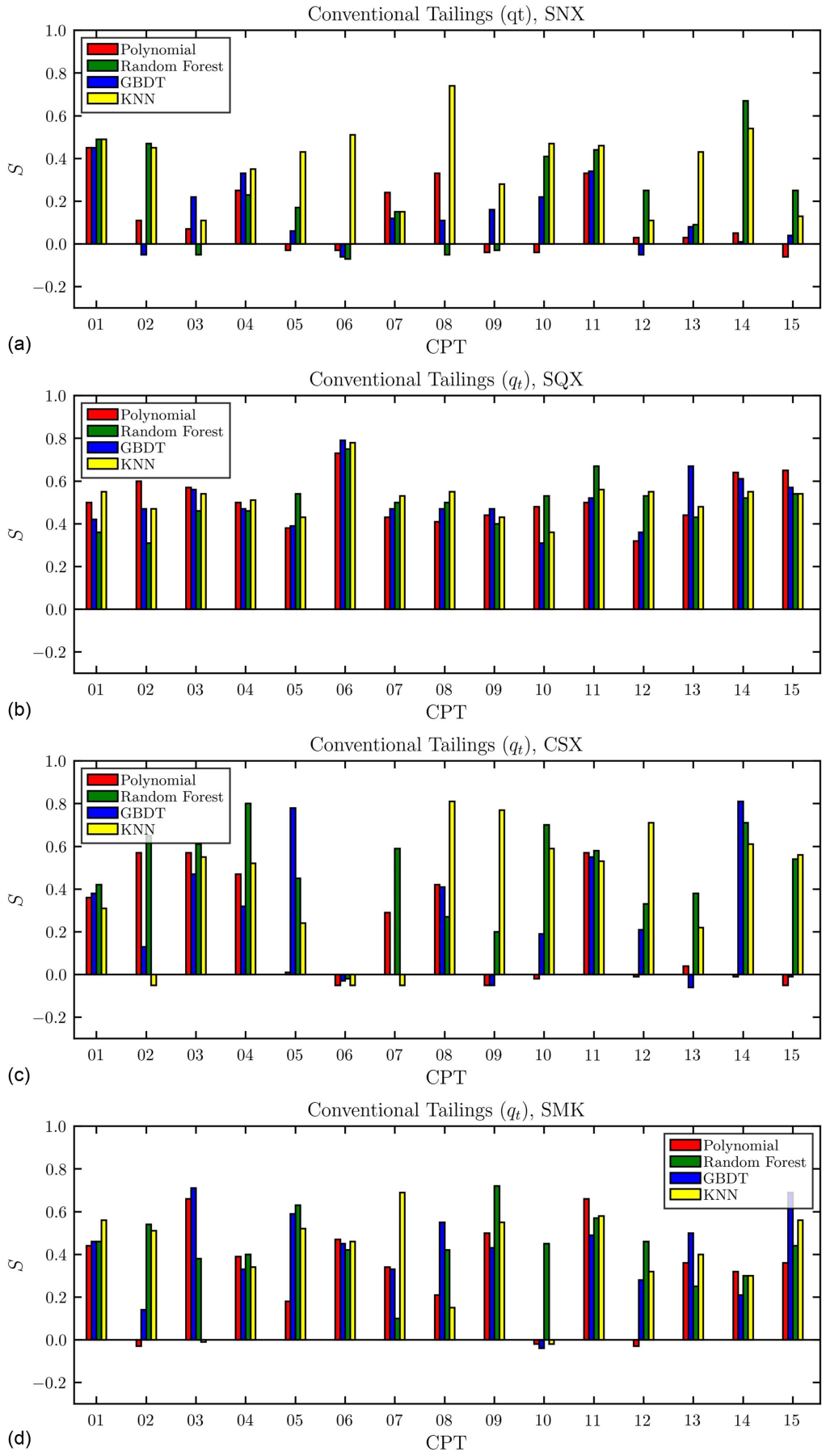

Fig. 11. Comparison of different trend functions (polynomials and ML-based) considering four different ACMs and all CPTus. The results are for q_t in the conventional tailings.

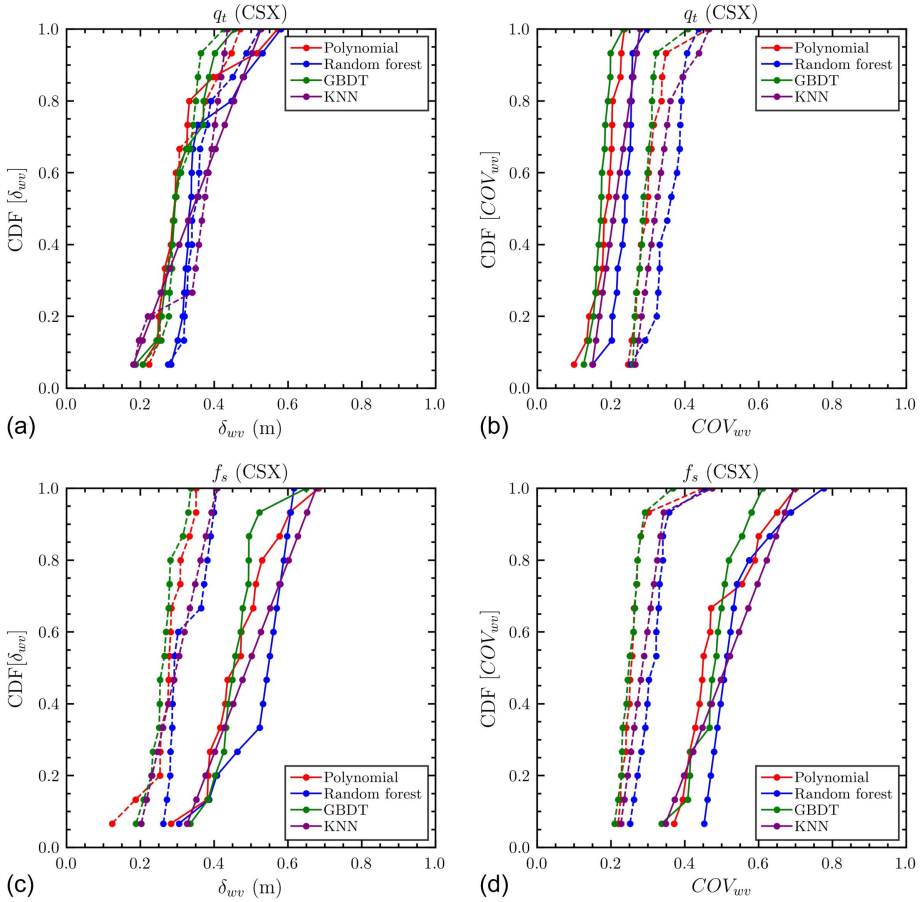

Fig. 12. (a) δ_{wv} CDF estimated from q_t data; (b) COV_{wv} CDF estimated from q_t data; (c) δ_{wv} CDF estimated from f_s data; and (d) COV_{wv} CDF estimated from f_s data. Results for thickened tailings are presented with solid curves, and for conventional tailings, with dashed curves.

The similar δ_{wv} values can be attributed to the combination of two aspects (1) the deposition scheme for the conventional and thickened tailings, and (2) averaging effects in CPTu measurements. In terms of the deposition process, similar phenomena are involved in both the deposition of the conventional and thickened tailings. These phenomena include pauses in the tailings discharge (the deposition is conducted on layers), drying-wetting cycles, induced consolidation when new layers are deposited, and induced flow gradients. Moreover, even though the thickened tailings have a less segregating nature, it is still reasonable to expect some segregation under field conditions during the deposition process. In terms of the averaging during CPTu penetration, it is well known that q_t is influenced by the response of the units that are at a distance of 1 to 3 cone diameters for soft soils and 20 to 30 cone diameters for stiffer soils (Lunne et al. 1997; Ahmadi and Robertson 2005). Based on the Ahmadi and Robertson (2005) findings and considering the relatively low tip resistance in the thickened and conventional tailings, the influence penetration zones are expected to be closer to the lower range from Lunne et al. (1997), with values on the order of 5 cone diameters for the thickened tailings and 10 cone diameters for the conventional tailings. Similarly, f_s also represents an averaged measurement. Thus, the relatively increased heterogeneity on the conventional tailings may have also been averaged out to some extent by the nature of the CPTu penetration process.

In terms of the COV_{wv} CDFs, it is important to put in context that COV_{wv} is a measurement of the fluctuation (i.e., σ) of $w(z)$ with respect to its order of magnitude [i.e., $t(z)$]. The relatively larger COV_{wv} for q_t in the conventional tailings is a reflection of the more significant fluctuation of q_t, potentially influenced by the more segregating nature of the conventional tailings. In contrast, the larger COV_{wv} for f_s in the thickened tailings is a reflection of the significantly lower f_s in the thickened tailings compared to the conventional tailings (i.e., a factor of about 3), which counteracts the σ contribution.

Horizontal Correlation Length

Due to the relatively large sampling interval in the horizontal direction compared to the vertical, deriving on significantly fewer points at a given depth, the horizontal correlation length (δ_{wh}) is estimated following the expeditive method, which was originally proposed by Vanmarcke (1977) and has been used in previous studies for natural soils (e.g., Stuedlein et al. 2012; Bong and Stuedlein 2017; Cary 2021). The expeditive method estimates δ_{wh} as $\sqrt{(2/\pi)}\,\overline{d}$, where $\overline{d}$ = average length of the segments obtained at the crossings of $g(z)$ and its linear trend function $t(z)$, where $g(z)$ = metric of interest (i.e., q_t or f_s). Following the recommendations of Stuedlein et al. (2012) and Bong and Stuedlein (2017), in applying the expeditive method, we considered a minimum of five crossings to get stable COV_{wh} and δ_{wh} estimates. Fig. 13(a) illustrates the q_t-based δ_{wh} estimation in the thickened and conventional tailings at depths of 4 and 12 m, and Fig. 13(b) illustrates the f_s-based δ_{wh} estimation at the same depths. The same procedure illustrated in these figures was applied to all depths.

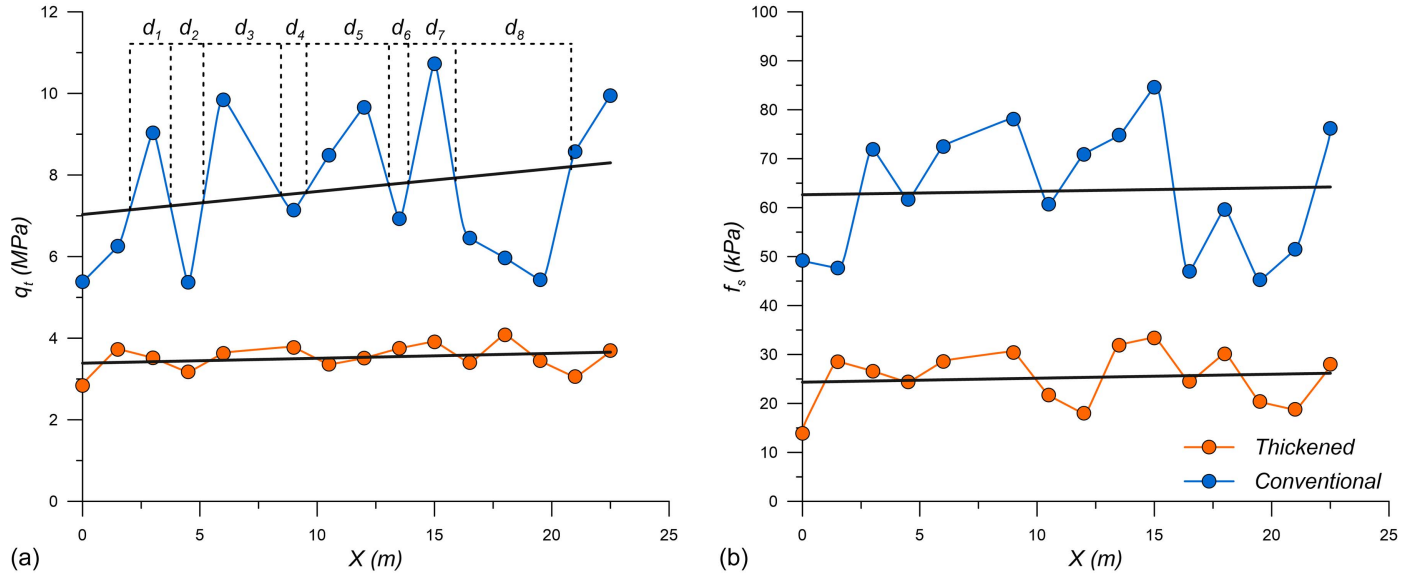

Fig. 13. Application of Vanmarcke's expeditive method for estimating the horizontal correlation length based on (a) q_t; and (b) f_s, considering the thickened and conventional tailings at selected depths (4 m for the thickened tailings and 12 m for the conventional tailings).

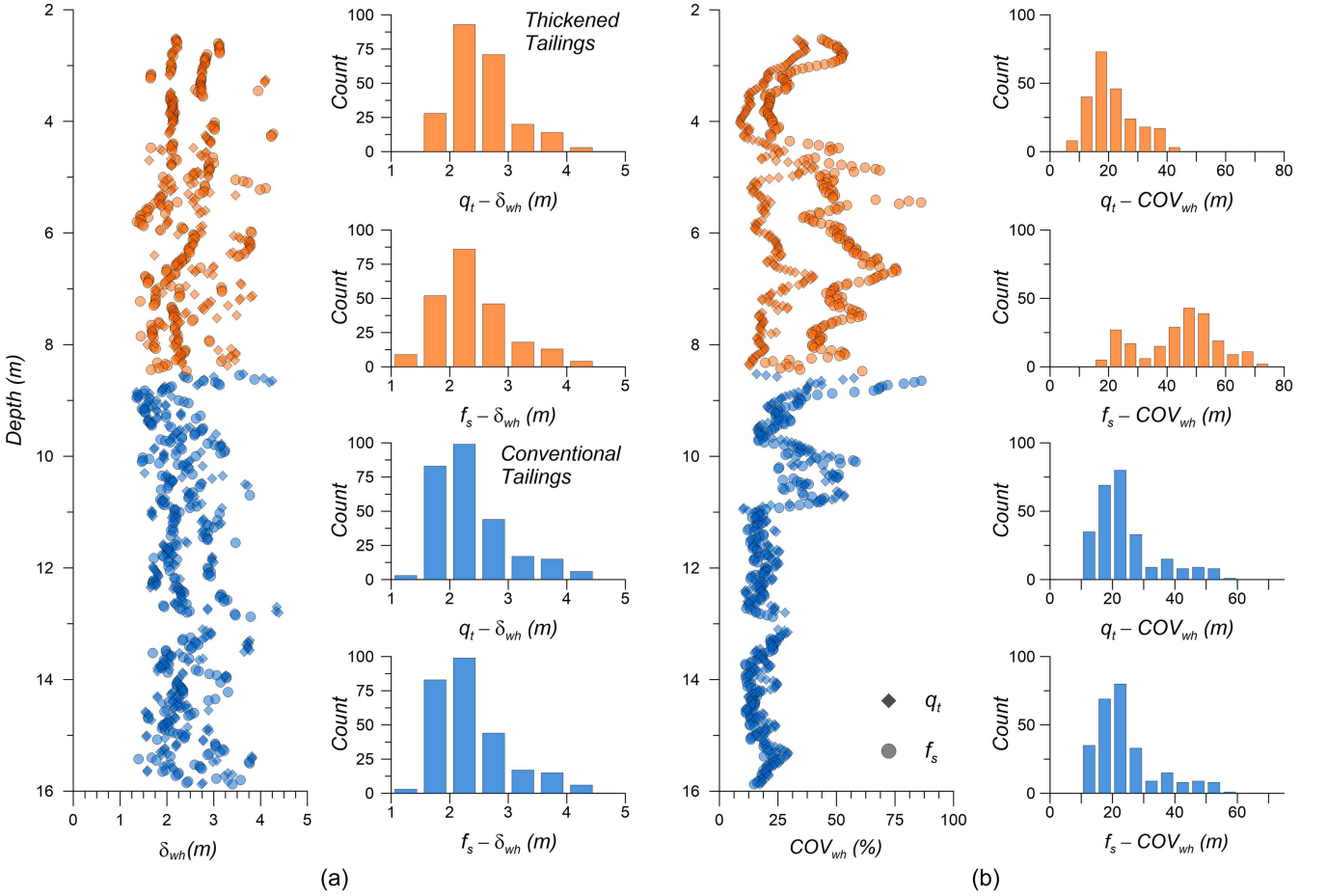

Fig. 14. (a) δ_{wh}; and (b) COV_{wh} profiles based on q_t and f_s, considering the thickened and conventional tailings.

Fig. 14(a) shows the estimated δ_{wh} considering q_t and f_s for the entire deposit, and Fig. 14(b) presents the associated horizontal coefficient of variations (COV_{wh}). It can be observed that the thickened and conventional tailings exhibit a similar δ_{wh}. For instance, the 16–84 percentile for the q_t-based δ_{wh} is in the range of 2 to 3.5 m for the thickened tailings and in the range of 1.8 to 3.0 m for the conventional tailings. In the case of f_s-based estimations, the 16–84 percentile δ_h is in the range of 1.7 to 3.2 m for the

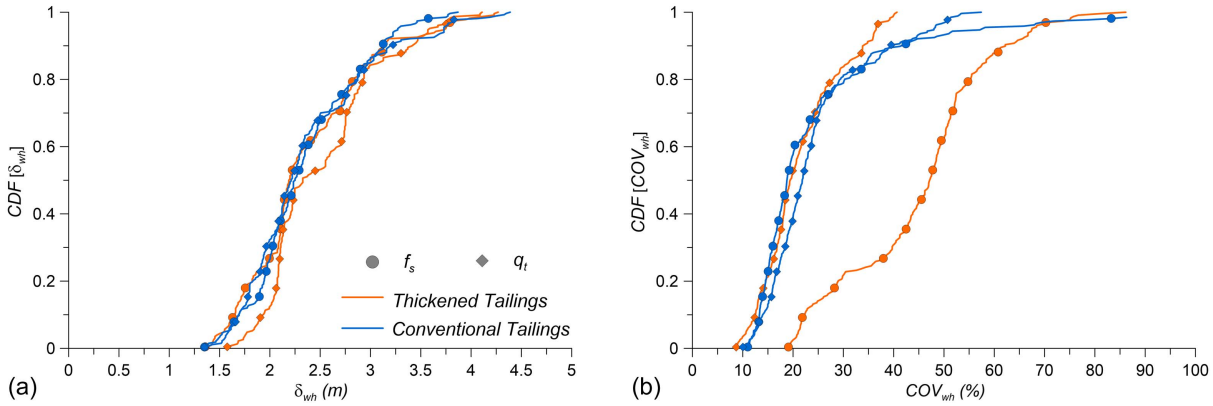

Fig. 15. CDFs of q_t and f_s based (a) δ_{wh}; and (b) COV_{wh} in the thickened and conventional tailings.

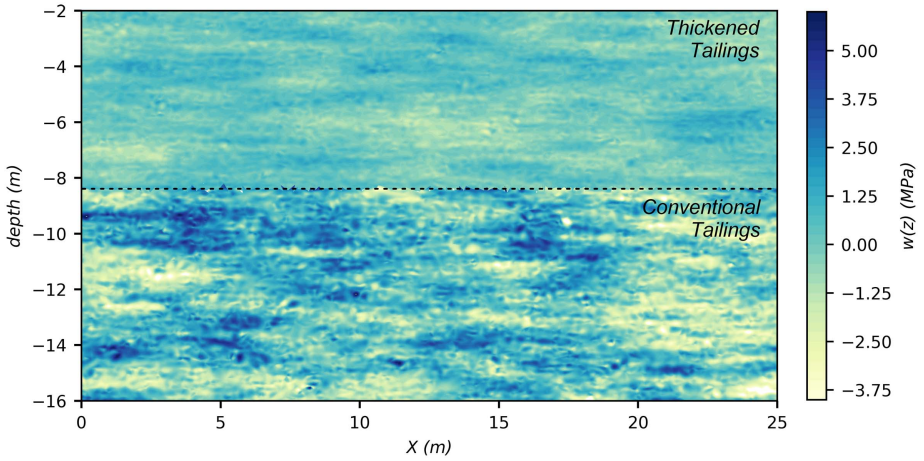

Fig. 16. Random field realization of a q_t-based $w(z)$ on the thickened (2 to 8.5 m) and conventional (8.5 to 16 m) tailings.

thickened tailings and 1.9 to 2.9 m for the conventional tailings. The similar δ_{wh} in both tailings can be attributed to the deposition process and associated phenomena (e.g., pauses in the discharge, flow gradient, drying-wetting cycles, freezing, and consolidation) as discussed before for δ_{wv}. Moreover, it is reasonable to expect segregation to have less effect in the horizontal direction compared to the vertical, which, again, is reflected in a similar δ_{wh}. The q_t-based COV_{wh} values tend to be slightly larger for the conventional tailings as the existing variance controls them, whereas the f_s-based COV_{wh} values are larger for the thickened tailings as they are controlled by the lower f_s values in the thickened tailings as previously discussed.

Fig. 15 presents q_t- and f_s-based CDFs for δ_{wh} and COV_{wh}; it can be observed that the δ_{wh} CDFs for q_t and f_s are consistent for both the thickened and conventional tailings. In addition, similar q_t-based COV_{wh} CDFs for the thickened and conventional tailings are also observed, being also consistent with the f_s-based COV_{wh} for the conventional tailings, whereas the f_s-based COV_{wh} CDF for the thickened tailings is shifted to the right, showing larger values for reasons already discussed.

The generation of random fields depends on δ and COV (or variance, σ^2); hence, having similar δ, which is the case for the thickened and conventional tailings, does not necessarily imply similar random fields. This can be illustrated by generating random fields with the parameters evaluated for the thickened and conventional tailings using the GSTools package (Muller and Schuler 2019). Fig. 16 shows one of these realizations considering the SNX autocorrelation model and parameters estimated using the GBDT model for detrending. In this case, the q_t-based estimated parameters for the thickened tailings are $\sigma^2 = 0.3$ MPa, $\delta_{wv} = 0.37$ m, and $\delta_{wh} = 2.45$ m, whereas the parameters for the conventional tailings are $\sigma^2 = 2.6$ MPa, $\delta_{wv} = 0.41$ m, and $\delta_{wh} = 2.29$ m. As can be inferred from these parameters, the main difference is in σ^2. It is interesting to note that the extent and shape of zones with similar $w(z)$ values are comparable for the two types of tailings due to the similar δ. However, the random field for the thickened tailings is significantly more stable (i.e., it presents less variability) due to the lower σ^2. Thus, despite the similar δ for both tailings, the simulated random field for the thickened tailings shows a more homogeneous structure compared to the conventional tailings. Lastly, Fig. 16 also shows the impact of different correlation lengths (i.e., δ_{wv}, and δ_{wh}), which derives in an anisotropic random field with properties correlated over more considerable (and comparable) distances in the horizontal direction for both the thickened and conventional tailings as δ_{wh} is larger than δ_{wv}.

Discussion

It is relevant to put in context that even though the generation of mine tailings entails complex processes to extract the valuable ore, the tailings gradation is regulated by the processing operations,

J. Geotech. Geoenviron. Eng., 2024, 150(10): 04024091

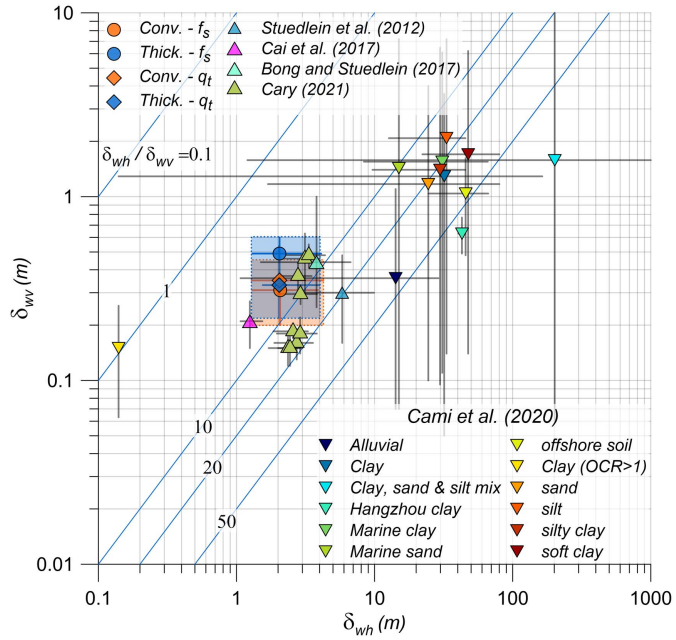

Fig. 17. δ_{wh} versus δ_{wv} for the mine tailings examined in this study in the context of different natural soils. Cary (2021): alluvial sands, Bong and Stuedlein (2017): estuarine sand deposit, Cai et al. (2017): alluvial plain, and Stuedlein et al. (2012): alluvial floodplain clay. δ_{wh} and δ_{wv} ranges from Cami et al. (2020) for different natural soils are also included. The shaded regions represent interpreted ranges for the conventional and thickened tailings examined in this study.

with the grain size and rheological properties being dependent on the ore source. As a result, the deposited material has been *prefabricated* to some extent, which contrasts with natural soils. Hence, we found it instructive to examine the estimated correlation lengths in this study for the thickened and conventional tailings in the context of natural soils. In Fig. 17, we compiled correlation length ranges using information reported by the recent state-of-the-art review of the International Society of Soil Mechanics and Geotechnical Engineering (ISSMGE) TC304 committee (2021) for natural soils. We considered the study from Cami et al. (2020), who compiled values for different soil types, including sands, silts, and clays, and the studies conducted by Stuedlein et al. (2012), Bong and Stuedlein (2017), and Cary (2021) for floodplain clays, estuarine, and alluvial deposits. The values reported in these efforts are compared with the values estimated in this study for tailings, considering δ_{wh}/δ_{wv}, δ_{wv}, and δ_{wh}. The best estimates are presented with different markers in Fig. 17; the horizontal/vertical bars represent the range of estimates. It can be observed that the estimated δ_{wh}/δ_{wv} values for the thickened tailings (2 to 18) and conventional tailings (4 to 21) are on the lower bound of values reported for most of the considered natural soils (i.e., 10 to 50).

The estimated δ_{wv} for mine tailings is generally within the ranges reported by Cami et al. (2020) for natural soils, whereas δ_{wh} is generally more consistent with the lower range reported by the same authors (Fig. 17). The estimated δ_{wv} is also consistent with studies on deltaic soils (Wickremesinghe and Campanella 1993) and alluvial plains (Cai et al. 2017). Considering studies that estimated vertical and horizontal correlation lengths, the δ_{wv}, δ_{wh}, and δ_{wh}/δ_{wv} estimated values in this study for mine tailings are comparable with values estimated in low-energy environments such as the Beaumont alluvial floodplain clay deposit examined by Stuedlein et al. (2012), the estuarine sand deposit investigated by Bong and Stuedlein (2017), and the alluvial sand deposits examined by Cary (2021), as illustrated by Fig. 17. These deposits are mainly alluvial in low-energy environments, corresponding to sediments or soils transported and deposited by running water, such as rivers or streams. For instance, deltas are low-lying plains formed by sediments deposited by rivers when they meet seas (i.e., estuaries) or stagnant water bodies. On the other hand, a floodplain is an area adjacent to a river stream characterized by unconsolidated sedimentary materials and periodically inundated by that stream. In the context of mine tailings, spigots are typically used to deposit them; tailings flow into multiple streams (associated with changes in spigot locations) toward the pond until they sediment, which can be associated to some extent with braided river channels in alluvial deposits. The processes just discussed resemble the formation of the alluvial deposits discussed in this section (see Fig. 18 for an example), hence potentially explaining the similarities in the spatial variability parameters (i.e., δ_{wh}/δ_{wv}) of alluvial deposits and mine tailings. As the energy environment during the deposition impacts the correlation length (Jones et al. 2002), it is also reasonable to expect that lacustrine formations deposited under low-energy environments may have comparable correlation lengths to the tailings examined in this study. However, we were unable to find recent studies quantifying δ_{wh} in lacustrine deposits with a dense array of CPTus, as considered in this study. The similarities/differences in correlation lengths between mine tailings and natural deposits

(a) (b)

Fig. 18. Aerial views: (a) Example of a thickened tailings impoundment (image © Google, © 2024 Airbus); and (b) Kachemak Bay estuary in Alaska [image courtesy of ShoreZone, under Creative Commons-BY-3.0 license (https://creativecommons.org/licenses/by/3.0/)].

should be further investigated once more data on spatial variability for mine tailings and low-energy environments is collected as part of future efforts.

It is also relevant to highlight that the estimated horizontal correlation length for the examined tailings varies between 1.5 and 4 m; thus, the lower end of the range is similar to the average CPTu spacing. In this context, the estimated horizontal correlation length lower bound is relatively uncertain. Several previous studies providing correlation lengths for natural soils (e.g., Jaksa et al. 1999; Stuedlein et al. 2012; Lloret-Cabot et al. 2014; Cai et al. 2017; Bong and Stuedlein 2017; Cary 2021) also share this potential limitation. As the correlation length is unknown a priori, in this study, consistent with Cary et al. (2022), the average horizontal scale of observation was 1.5 m, and it was set to minimize the CPTu spacing while also minimizing the potential disturbance of a CPTu push to neighboring CPTus. Some previous efforts also suggest that the scale of observation can play a role in determining the horizontal correlation length. For example, Jaksa (1995) and Jaksa et al. (1999) compared correlation lengths obtained from vertically and horizontally pushed CPTus (data collected every 1.0 and 0.2 m, respectively) on Keswick clay and postulated the idea of a nested correlation structure with correlation lengths that can be influenced by the observation scale. Conducting horizontal CPTus in a TSF is currently impractical. In this context, future efforts could consider pushing small-diameter CPTus, allowing closer spacing while minimizing potential disturbance effects on neighboring CPTus. For example, Meisina et al. (2021) recently used a 2 cm^2 CPTu to investigate natural silts; similar CPTu probes could be used for investigating mine tailings and assessing correlation lengths. One challenge for this is that data quality from a 2 cm^2 CPTu may be inferior compared to a standard CPTu. Last, there are also ongoing research efforts to develop bioinspired, robotic-based technologies that would potentially allow penetration devices to investigate subsurface properties in multiple directions (e.g., Chen et al. 2021; Borela et al. 2021). Such next-generation penetrometers would likely allow further investigation of the challenging problem of estimating horizontal correlation lengths and the postulate of nested correlation structures.

Conclusions

As the application of performance-based approaches to tailings management advances (Morgenstern 2018), the quantification of variabilities (spatial and aleatory) is expected to become crucial, particularly for TSFs where the geotechnical characterization of mine tailings is key in assessing the overall response (e.g., the overall physical stability in upstream TSFs). In this context, we have collected data using a dense CPTu array to assess the spatial variability of two different types of tailings deposited in the same TSF. We also evaluated the potential of modern machine learning methods for informing the estimation of random field parameters; the examined methods provided advantages, which we attribute to the interpolation nature of CPTu detrending (caution should be exerted for extrapolation problems where machine learning can act as a black box). This study's findings may prove valuable in geotechnical analyses that require adopting δ values to account for spatial variability in mine tailings similar to those examined in this study, specifically thickened and conventional tailings produced from gold ore. Given the lack of spatial correlation information on mine tailings, we expect that it would also be helpful as a benchmark for future efforts. Of note, case studies exist documenting the importance of spatial variability parameters for water dams (Sanchez-Lizarraga and Lai 2014; Boulanger and Montgomery 2016; Guo et al. 2019). Thus, it is reasonable to expect similar spatial variability studies focused on TSFs to translate into engineering practice in the future.

In terms of the assessment of random fields, our results show that ML-based detrending provides higher stationarity than traditional polynomials in most cases, which was assessed using a proposed new stationarity score. It is worth highlighting that utilizing ML models in addition to traditional polynomial methods improves the treatment of epistemic uncertainties by providing more options to create stationary random fields. The results show that the vertical correlation length (δ_{wv}) and the horizontal correlation length (δ_{wh}) for the thickened tailings (considering the 16–84 percentile) are in the range of 0.2–0.45 m and 2.0–3.2 m. On the other hand, δ_{wv} and δ_{wh} for the conventional tailings (again considering the 16–84 percentile) are in the range of 0.2–0.4 m and 1.75–2.85 m. Thus, the δ_{wv} and δ_{wh} and their distributions are comparable. As discussed in the manuscript, we attribute this to similar processes in the deposition of the examined tailings (i.e., pauses in the discharge, flow gradient, drying-wetting cycles, freezing, and consolidation). Potential averaging effects on the CPTu measurements may also contribute to some extent. In contrast, the standard deviation (σ) for the conventional tailings is about ten times higher than that estimated for the thickened tailings, which we attribute to the segregating nature of the conventional tailings. Ultimately, the higher variance in the conventional tailings results in a more variable random field, which is consistent with what would be expected in field conditions.

Finally, the estimated autocorrelation lengths and the associated anisotropy ($\delta_{wh}/\delta_{wv} = 2$–$21$) for the investigated tailings are put into the context of information available for natural soils. We found that the estimated δ_{wv}, δ_{wh}, δ_{wh}/δ_{wv} values are comparable with reported values in delta and floodplain alluvial deposits, which we attribute to similarities in the deposition processes, but this needs to be further investigated once more data is collected. In closing, it is also relevant to highlight that even though this study provided the unique opportunity of characterizing spatial variability properties for two different types of tailings, future additional studies on other TSFs and tailings technologies with varying sources of ore are encouraged. In particular, one limitation of our study is that we only considered a line of CPTus parallel to the discharge line (Fig. 2); future studies should consider multiple directions relative to the discharge line.

Data Availability Statement

Some or all data, models, or code generated or used during this study are available from the corresponding author by request.

Acknowledgments

This material is based upon work supported by the National Science Foundation (NSF) under Grant No. CMMI 2145092. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF. The PRONABEC program of the Peruvian government also provided complementary support. In addition, we would like to thank ConeTec, Newmont, and WSP-Golder for supporting the site characterization efforts. Finally, we thank Prof. Armin Stuedlein for sharing VBA codes we used to validate our implementations of calculations based on polynomial fittings and Prof. Jason Dejong for discussions when planning the CPTu campaign for the spatial variability characterization.

Supplemental Materials

Table S1 and Figs. S1–S7 are available online in the ASCE Library (www.ascelibrary.org).

References

Ahmadi, M. M., and P. K. Robertson. 2005. "Thin-layer effects on the CPT qc measurement." *Can. Geotech. J.* 42 (5): 1302–1317. https://doi.org/10.1139/t05-036.

Bagińska, I., M. Kawa, and W. Janecki. 2016. "Estimation of spatial variability of lignite mine dumping ground soil properties using CPTu results." *Stud. Geotech. Mech.* 38 (1): 3–13. https://doi.org/10.1515/sgem-2016-0001.

Bartlett, M. S. 1937. "Properties of sufficiency and statistical tests." *Proc. R. Soc. London* 160 (901): 268–282. https://doi.org/10.1098/rspa.1937.0109.

Bishop, C. 2006. *Pattern recognition and machine learning*. New York: Springer.

Bong, T., and A. W. Stuedlein. 2017. "Spatial variability of CPT Parameters and silty fines in liquefiable beach sands." *J. Geotech. Geoenviron. Eng.* 143 (12): 04017093. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001789.

Borela, R., J. D. Frost, G. Viggiani, and F. Anselmucci. 2021. "Earthworm-inspired robotic locomotion in sand: An experimental study using x-ray tomography." *Géotech. Lett.* 11 (1): 1–22. https://doi.org/10.1680/jgele.20.00085.

Boulanger, R. W., and J. Montgomery. 2016. "Nonlinear deformation analyses of an embankment dam on a spatially variable liquefiable deposit." *Soil Dyn. Earthquake Eng.* 91 (Mar): 222–233. https://doi.org/10.1016/j.soildyn.2016.07.027.

Box, G., and G. Jenkins. 1970. *Time series analysis: Forecasting and control*. San Francisco: Holden-Day.

Breiman, L. 2001. "Random forests." *Mach. Learn.* 45 (1): 5–32. https://doi.org/10.1023/A:1010933404324.

Breysse, D., H. Niandou, S. Elachachi, and L. Houy. 2005. "A generic approach to soil–structure interaction considering the effects of soil heterogeneity." *Géotechnique* 55 (2): 143–150. https://doi.org/10.1680/geot.2005.55.2.143.

Cacciuttolo, C., and A. Marinovic. 2022. "Sustainable management of thickened tailings in Chile and Peru: A review of practical experience and socio-environmental acceptance." *Sustainability* 14 (17): 10901. https://doi.org/10.3390/su141710901.

Cai, G., J. Lin, S. Liu, and A. J. Puppala. 2017. "Characterization of spatial variability of CPTU data in a liquefaction site improved by vibro-compaction method." *KSCE J. Civ. Eng.* 21 (1): 209–219. https://doi.org/10.1007/s12205-016-0631-1.

Cami, B., S. Javankhoshdel, K.-K. Phoon, and J. Ching. 2020. "Scale of fluctuation for spatially varying soils: Estimation methods and values." *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.* 6 (4): 03120002. https://doi.org/10.1061/ajrua6.0001083.

Cary, J. 2021. "An investigation into the role of spatial variability on liquefaction consequence severity." Master's thesis, School of Civil and Construction Engineering, Oregon State Univ.

Cary, J. R., A. W. Stuedlein, C. R. McGann, B. A. Bradley, and B. W. Maurer. 2022. "Effect of refinements to CPT-based liquefaction triggering analysis on liquefaction severity indices at the Avondale playground site, Christchurch, NZ." In *Proc., 4th Int. Conf. on Performance Based Design in Earthquake Geotechnical Engineering*, 1454–1466. Berlin: Springer.

Chen, T., and C. Guestrin. 2016. "XGBoost: A scalable tree boosting system." In *Proc., 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD '16)*, 785–794. New York: Association for Computing Machinery.

Chen, Y., A. Khosravi, A. Martinez, and J. DeJong. 2021. "Modeling the self-penetration process of a bio-inspired probe in granular soils." *Bioinspiration Biomimetics* 16 (4): 046012. https://doi.org/10.1088/1748-3190/abf46e.

Ching, J., T.-J. Wu, A. W. Stuedlein, and T. Bong. 2018. "Estimating horizontal scale of fluctuation with limited CPT soundings." *Geosci. Front.* 9 (6): 1597–1608. https://doi.org/10.1016/j.gsf.2017.11.008.

DeGroot, D. J., and G. B. Baecher. 1993. "Estimating autocovariance of in situ soil properties." *J. Geotech. Geoenviron. Eng.* 119 (1): 147–166. https://doi.org/10.1061/(ASCE)0733-9410(1993)119:1(147).

Fenton, G. A., and D. V. Griffiths. 2005. "Three-dimensional probabilistic foundation settlement." *J. Geotech. Geoenviron. Eng.* 131 (2): 232–239. https://doi.org/10.1061/(ASCE)1090-0241(2005)131:2(232).

Friedman, J. 2001. "Greedy function approximation: A gradient boosting machine." *Ann. Stat.* 29 (5): 1189–1232. https://doi.org/10.1214/aos/1013203451.

Griffiths, D. V., G. A. Fenton, and N. Manoharan. 2006. "Undrained bearing capacity of two-strip footings on spatially random soil." *Int. J. Geomech.* 6 (6): 421–427. https://doi.org/10.1061/(ASCE)1532-3641(2006)6:6(421).

Guo, X., D. Dias, and Q. Pan. 2019. "Probabilistic stability analysis of an embankment dam considering soil spatial variability." *Comput. Geotech.* 113 (Mar): 103093. https://doi.org/10.1016/j.compgeo.2019.103093.

Hu, Y.-G., and J. Ching. 2015. "Impact of spatial variability in undrained shear strength on active lateral force in clay." *Struct. Saf.* 52 (Dec): 121–131. https://doi.org/10.1016/j.strusafe.2014.09.004.

International Council on Mining and Metals. 2020. *Global industry standard on tailings management*.

ISSMGE (International Society of Soil Mechanics and Geotechnical Engineering). 2021. *State-of-the-art review of inherent variability and uncertainty in geotechnical properties and models*. ISSMGE-TC304. London: ISSMGE.

Jaksa, M. B. 1995. "The influence of spatial variability on the geotechnical design properties of a stiff, over consolidated clay." Ph.D. thesis, Faculty of Engineering, Univ. of Adelaide.

Jaksa, M. B., P. I. Brooker, and W. S. Kaggwa. 1997. "Inaccuracies associated with estimating random measurement errors." *J. Geotech. Geoenviron. Eng.* 123 (5): 393–401. https://doi.org/10.1061/(ASCE)1090-0241(1997)123:5(393).

Jaksa, M. B., J. S. Goldsworthy, G. A. Fenton, W. S. Kaggwa, D. V. Griffiths, Y. L. Kuo, and H. G. Poulos. 2005. "Towards reliable and effective site investigations." *Géotechnique* 55 (2): 109–121. https://doi.org/10.1680/geot.2005.55.2.109.

Jaksa, M. B., W. S. Kaggwa, and P. I. Brooker. 1999. "Experimental evaluation of the scale of fluctuation of a stiff clay." In *Proc., 8th Int. Conf. on Application of Statistics and Probability*, 415–422. Rotterdam, Netherlands: A.A. Balkema.

Jones, A. L., S. L. Kramer, and P. Arduino. 2002. *Estimation of Uncertainty in Geotechnical Properties for Performance-Based Earthquake Engineering*. PEER Rep. 2002-16. Berkeley, CA: Pacific Earthquake Engineering Research Center, Univ. of California.

Kendall, M. G. 1938. "A new measure of rank correlation." *Biometrika* 30 (1–2): 81–93. https://doi.org/10.1093/biomet/30.1-2.81.

Lacasse, S., and F. Nadim. 1996. "Uncertainties in characterising soil properties." In *Uncertainty in the geologic environment: From theory to practice*, Geotechnical Special Publication No. 58, 49–75. Reston, VA: ASCE.

Lizarraga, H. S., and C. G. Lai. 2014. "Effects of spatial variability of soil properties on the seismic response of an embankment dam." *Soil Dyn. Earthquake Eng.* 64 (Mar): 113–128. https://doi.org/10.1016/j.soildyn.2014.03.016.

Lloret-Cabot, M., G. A. Fenton, and M. A. Hicks. 2014. "On the estimation of scale of fluctuation in geostatistics." *Georisk* 8 (2): 129–140. https://doi.org/10.1080/17499518.2013.871189.

Lunne, T., P. K. Robertson, and J. M. Powell. 1997. *Cone penetration testing in geotechnical practice*. London: Blackie Academic & Professional.

Meisina, C., P. S. Öztürk Kardoğan, R. Bonì, S. Stacul, D. Castaldini, D. Fontana, S. Lugli, M. Bordoni, and D. Lo Presti. 2021. "Development and use of a minicone for liquefaction risk evaluation in layered soil deposits." *J. Geotech. Geoenviron. Eng.* 147 (2): 04020169. https://doi.org/10.1061/(ASCE)GT.1943-5606.0002457.

Morgenstern, N. 2018. "Geotechnical risk, regulation, and public policy." *Soil Rocks* 41 (2): 107–129. https://doi.org/10.28927/SR.412107.
Morrison, K. 2022. *Tailings management handbook: A life-cycle approach*. Englewood, CO: Society for Mining, Metallurgy and Exploration (SME).
Muller, S., and L. Schuler. 2019. *GeoStat-framework/GSTools: Bouncy blue (v1.0.1)*.
Peterson, L. 2009. "K-nearest neighbor." *Scholarpedia* 4 (2): 1883. https://doi.org/10.4249/scholarpedia.1883.
Phoon, K. K., and F. H. Kulhawy. 1999. "Characterization of geotechnical variability." *Can. Geotech. J.* 36 (4): 612–624. https://doi.org/10.1139/t99-038.
Phoon, K. K., S. T. Quek, and P. An. 2003. "Identification of statistically homogenous soil layers using modified Bartlett statistics." *J. Geotech. Geoenviron. Eng.* 129 (7): 649–659. https://doi.org/10.1061/(ASCE)1090-0241(2003)129:7(649).
Reid, D., and M. Jefferies. 2017. "State parameter as a geological principle in tailings." In *Proc., of Tailings and Mine Waste*, 305–314. Edmonton, AL, Canada: Univ. of Alberta.
Reid, D., and M. Jefferies. 2018. "A geological principle for the density of thickened tailings." In *Paste 2018: Proc., 21st Int. Seminar on Paste and Thickened Tailings*. Perth, Australia: Australian Centre for Geomechanics.
Robertson, P. K. 2016. "Cone penetration test (CPT)-based soil behaviour type (SBT) classification system–an update." *Can. Geotech. J.* 53 (12): 1910–1927. https://doi.org/10.1139/cgj-2016-0044.
Robinsky, E. I. 1979. "Tailing disposal by the thickened discharge method for improved economy and environmental control." In Vol. 2 of *Proc., 2nd Int. Tailings Symp.*, 75–95. San Francisco: Miller Freeman Publications.
Spearman, C. 1904. "The proof and measurement of association between two things." *Am. J. Psychol.* 15 (1): 72. https://doi.org/10.2307/1412159.
Stuedlein, A. W. 2008. "Bearing capacity and displacement of spread footings on aggregate pier reinforced clay." Ph.D. thesis, Dept. of Civil and Environmental Engineering, Univ. of Washington.
Stuedlein, A. W. 2011. "Random field model parameters for Columbia River silt." In *Georisk 2011: Risk Assessment and Management*, Geotechnical Special Publication 224, 169–177. Reston, VA: ASCE.
Stuedlein, A. W., and T. Bong. 2017. "Effect of spatial variability on static and liquefaction-induced differential settlements." In *Geo-Risk 2017*. Reston, VA: ASCE.
Stuedlein, A. W., S. L. Kramer, P. Arduino, and R. D. Holtz. 2012. "Geotechnical characterization and random field modeling of desiccated clay." *J. Geotech. Geoenviron. Eng.* 138 (11): 1301–1313. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000723.
Uzielli, M., G. Vannucchi, and K. Phoon. 2005. "Random field characterization of stress-normalised cone penetration testing parameters." *Géotechnique* 55 (1): 3–20. https://doi.org/10.1680/geot.2005.55.1.3.
Vanmarcke, E. H. 1977. "Probabilistic modeling of soil profiles." *J. Geotech. Eng. Div.* 103 (11): 1227–1246. https://doi.org/10.1061/AJGEB6.0000517.
Vick, S. 1990. *Planning, design, and analysis of tailings dams*. Richmond, BC: BiTech Publishers.
Watson, A. H., P. G. Corser, E. E. Garces Pardo, T. E. Lopez Christian, and J. Vandekeybus. 2010. "A comparison of alternative tailings disposal methods–The promises and realities." In *Mine Waste 2010: Proc., 1st Int. Seminar on the Reduction of Risk in the Management of Tailings and Mine Waste*, edited by R. Jewell and A. B. Fourie, 499–514. Perth, Australia: Australian Centre for Geomechanics.
Wickremesinghe, D., and R. G. Campanella. 1993. "Scale of fluctuation as a descriptor of soil variability." In *Proc., Conf. on Probabilistic Methods in Geotechnical Engineering*, 233–239. Rotterdam, Netherlands: A.A. Balkema.