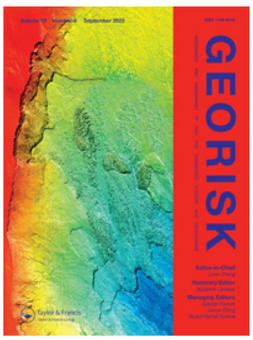

Taylor & Francis
Taylor & Francis Group

Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ngrk20

Inherent variability assessment from sparse property data of overburden soils and intermediate geomaterials using random field approaches

Opeyemi E. Oluwatuyi, Rebecca Holt, Rasika Rajapakshage, Shaun S. Wulff & Kam Ng

To cite this article: Opeyemi E. Oluwatuyi, Rebecca Holt, Rasika Rajapakshage, Shaun S. Wulff & Kam Ng (2022): Inherent variability assessment from sparse property data of overburden soils and intermediate geomaterials using random field approaches, Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, DOI: 10.1080/17499518.2022.2046783

To link to this article: https://doi.org/10.1080/17499518.2022.2046783

GEORISK
https://doi.org/10.1080/17499518.2022.2046783

Taylor & Francis
Taylor & Francis Group

Inherent variability assessment from sparse property data of overburden soils and intermediate geomaterials using random field approaches

Opeyemi E. Oluwatuyi [a], Rebecca Holt[b], Rasika Rajapakshage[b], Shaun S. Wulff [b] and Kam Ng [a]

[a]Civil and Architectural Engineering Department, University of Wyoming, Laramie, WY, USA; [b]Mathematics and Statistics Department, University of Wyoming, Laramie, WY, USA

ABSTRACT

This study assesses the inherent variability in the geomaterial parameter by quantifying the parameter uncertainty and develops a site investigation plan with a low degree of uncertainty. A key research motivation was using sparse borehole data to predict a site geomaterial configuration in order to determine the design of a site investigation plan. This study develops a systematic methodology for carrying out a study of inherent variability in light of the limitations posed by borehole data. The data in this study was provided by the Iowa Department of Transportation which consisted of eight boreholes from which 92 associated SPT N-values was considered as the geomaterial parameter of interest. The systematic methodology then involved the following steps. A general linear model was employed to fit and compare various spatial covariance models with and without a nugget. These spatial covariance models were also evaluated with variograms. Predicted SPT N-values were generated using universal kriging. Simulations were performed conditionally and unconditionally to identify optimal site investigation plans. The results identified site investigation plans with reduced parameter uncertainty. The proposed approach can produce site investigation plans that target any or all geomaterial layers to reduce uncertainty with respect to any geomaterial parameter of interest.

ARTICLE HISTORY
Received 27 January 2021
Accepted 8 February 2022

KEYWORDS
Geomaterial layer; geostatistics; kriging; spatial statistics; SPT; variogram

1. Introduction

Geotechnical engineering must address geomaterial heterogeneity that comes from inherent variability and geological uncertainty as it affects structure stability. Engineering judgment, experience, factor of safety, rules of the thumb, and subjective reasoning have been used previously to deal with geomaterial heterogeneity (Elkateb, Chalaturnyk, and Robertson 2003; Simpson 2003; Baecher and Christian 2005). However, this has led to potentially dangerous under design or expensive over design. Therefore, there is a need to optimally quantify and integrate inherent variability into geotechnical design. Inherent variability arises mainly from spatial variations of geomaterial properties/parameter (Deng et al. 2017). It is regarded as the most influential type of uncertainty in geotechnical engineering (Christian, Ladd, and Baecher 1994; Phoon and Kulhawy 1999; Lloret-Cabot, Fenton, and Hicks 2014).

Inherent variability has often been assessed with random field methods (Vanmarcke 2010; Christakos 2012; Stuedlein et al. 2012; Wang, Au, and Cao 2010; Dasaka and Zhang 2012; Li, Zhang, and Li 2016). The random field (RF) approach generates a random field based on a geomaterial parameter from which to simulate the subsurface region. Several studies have used RF approaches, including those by Wang, Zhao, and Phoon (2018), where they developed a RF generator based on Bayesian compressive sampling (BCS) and Karhunen–Loève (KL) expansion to generate samples from limited data while incorporating the associated statistical uncertainty. Conditional RF has been applied to account for the random behaviour of ground properties to examine the performance of a tunnel liner via finite-difference modelling within a Monte Carlo simulation framework (Yu et al. 2019). A probability-based geotechnical analysis using Bayesian RF models has been implemented to incorporate the inherent variability of soil friction angle (Tian et al. 2016). In this approach, a three-dimensional (3D) conditioned RF with a low prediction error was used for soil resistance estimation at unsampled locations based on limited cone penetration tests (Cai et al. 2019). A two-dimensional (2D) RF model based on the Karhunen–Loève expansion was generated to account for hydraulic conductivity uncertainties and spatial variation in a layered soil profile (Cho 2012). This stationary RF was modified

CONTACT Kam Ng kng1@uwyo.edu Civil and Architectural Engineering Department, University of Wyoming, 1000 E. University Ave, EN3050, Laramie, WY 82071-2000, USA

to a more accurate non-stationary RF to account for the mean and variance of the soil properties (Jamshidi Chenari and Kamyab 2015).

The study of inherent variability is also important for constructing a site investigation plan. This is because inherent variability assessed through parameter uncertainty can be reduced through the determination of an adequate number of well-placed boreholes (BHs). Heatmaps were generated with the random finite element method to determine the optimal location for a single borehole in a single-layered soil (Crisp et al. 2020). In another study by Crisp, Jaksa, and Kuo (2020a), several factors like soil parameters, number of piles, stiffness ratios between soil layers, and layer boundaries were examined to determine an optimal investigation plan. The effect of these factors on the optimal site investigation were studied individually rather than collectively. The authors concluded that the optimum number of boreholes were affected by the number of piles, stiffness ratio, and degree of undulation in soil layers. In another related study, the optimal site investigation plan was examined for the design of an undrained slope (Yang et al. 2019). These authors determined the number of boreholes required to characterise the slope with a low degree of inherent variability as measured through the standard deviation. The standard deviation is a commonly used statistical measure of uncertainty, which is used in this study to assess the inherent variability of the geomaterial parameter. A low degree of inherent variability implies a small standard deviation value for the geomaterial parameter. Quantification of variability in different rock types had shown low variability in bulk density, specific gravity, and unit weight, while rock properties with high variability include strength, porosity, and water content (Aladejare and Wang 2017). An approach to detect the minimal number of rock specimens needed for laboratory uniaxial compressive strength testing using variograms and directional covariance maps were developed by Boyd, Trainor-Guitton, and Walton (2018).

Though these studies served as background to this present study, the previous analyses were limited to layered soil (Zhao et al. 2013) or rock (Chen et al. 2019). However, especially in the Rocky Mountain region, deep foundations for bridges are driven into shallow bedrock stratigraphy known as intermediate geomaterials (IGMs), which are an acclaimed variable material with properties between soil and rock (Adhikari et al. 2020). Apart from the challenges with their geotechnical and geological characterisation, presently little is known on the inherent variability of IGMs. The contributions of this study were, therefore, aimed at (a) predicting geomaterial parameter of interest for the study area using sparse borehole data, (b) assessing inherent variability through the quantification of parameter uncertainty of multilayered geomaterials including the IGM layer, and (c) obtaining a site investigation plan with a low degree of parameter uncertainty. Thus, the RF estimation approach in this study needed to be modified accordingly over previous approaches. For example, ordinary kriging, which treats all the geomaterial layers with a constant mean, was used in previous studies for the estimation of the RF (Lloret-Cabot, Hicks, and van Den Eijnden 2012; Al-Bittar, Soubra, and Thajeel 2018; Soubra et al. 2019). This study uses universal kriging so that all geomaterial layers are analyzed with different means with respect to the geomaterial parameter. By combining the layers this way, it is possible to gain additional and important information on the spatial structure of the whole site while also improving geomaterial parameter prediction at unsampled location of the random field. The standard deviation (SD) reduction method was suggested for the selection of site investigation plan in a single layer soil (Crisp, Jaksa, and Kuo 2020b). However, in this study, SD reduction method is performed in an optimal manner as follows: First, to reduce uncertainty with respect to a geomaterial parameter in pile foundation design, boreholes are recommended near or on a foundation (Arsyad et al. 2010; Goldsworthy et al. 2007). Subsequent boreholes are thereafter selected to reduce the standard deviation of the predictions with respect to the IGM layer. RF modelling in this study was conducted with both 2D and 3D representations. The 2D representation provided a simpler analysis while the 3D provided specific x and y coordinates for borehole drilling. Because of these analyses contribution, site-specific recommendations were made to determine a site investigation plan. The plan involved selection of an optimal numbers of boreholes, the locations of these boreholes, and the sampling depths within the random field (study site). This information is particularly beneficial for engineering projects with relatively similar dimensions and subsurface profiles to this present case study, where there are no test data, or the data is very sparse. The paper presents the techniques and methods used in the inherent variability study (Section 2), the dataset used as a case study (Section 3), the results and discussions from the analyses (Section 4), and conclusions and recommendations (Section 5).

2. Methodology

This section details the methods used to characterise inherent variability in this study. The R program

was chosen due to its accessibility and versatility for conducting the required analyses. Specific R packages and functions will be listed in the discussions as appropriate. The proposed systematic process of assessing inherent variability involved specifying the general linear model and variogram fit after the geomaterial property had been assigned to the corresponding grid of a random field. Universal Kriging is thereafter used in predicting the geomaterial property of the discretised region. Virtual boreholes, which form a proposed sampling plan, were then drilled on geostatistically simulated realisations of the site. Inherent variability was then assessed from the quantified standard deviation obtained from the statistical analysis on the drilled virtual boreholes. A flowchart showing the process, and the specific R package functions, is displayed in Figure 1. The methodology subsections elucidate more on the proposed systematic process of assessing inherent variability.

2.1. General linear model

The random response of interest ($\underline{Y}$) is a geomaterial based quality characteristic, which could vary within the study area both spatially and in the mean. The general linear model (GLM) was used with a specific spatial covariance structure to model $\underline{Y}$ as given by:

$$\underline{Y} = X\underline{\beta} + \underline{E} \tag{1}$$

where it is assumed that the model errors satisfy $\underline{E} \sim \text{Normal}_n(0, \Sigma)$ so that the expected value is $\varepsilon(\underline{Y}) = X\underline{\beta}$ and the covariance matrix $\mathbb{C}(\underline{Y}) = \Sigma$.

The covariance matrix Σ is important to account for the spatial structure. The spatial correlation function was specified via $H(\rho)$ as:

$$\Sigma = \sigma^2 H(\rho) \tag{2}$$

where ρ represents the range parameter and σ^2 represents the partial sill. A nugget (τ^2) was used to account for the measurement error or additional spatial variation that would be observed at the origin of the variogram (Diggle and Ribeiro Jr 2002). With the addition of the nugget, the covariance matrix becomes:

$$\Sigma = \sigma^2 H(\rho) + \tau^2 I\,. \tag{3}$$

Four spatial correlation functions considered for $H(\rho)$ which were Exponential, Gaussian, Linear, and Spherical functions given in Equation (4) through Equation (7), respectively (Pinheiro and Bates 2006):

$$H(\rho) = \exp(-h/\rho) \tag{4}$$

$$H(\rho) = \exp(-(h/\rho)^2) \tag{5}$$

$$H(\rho) = \left[1 - \left(\frac{h}{\rho}\right)\right] I(h < \rho) \tag{6}$$

$$H(\rho) = \left[1 - 1.5\left(\frac{h}{\rho}\right) + 0.5\left(\frac{h}{\rho}\right)^3\right] I(h < \rho) \tag{7}$$

The above equations give correlation values $H(\rho)$ as a function of the Euclidean distance (h) between the spatial locations. The null model has a correlation function $H(\rho)$ equal to 1.

In this study, the geomaterial layers were assumed to have different true means that affect the expected value in Equation (1). Thus, the unknown regression coefficient vector ($\underline{\beta}$) was used to estimate these means. The corresponding design matrix was then given by

$$X = I_p^{(n_l)} \otimes \underline{1}_{n_l} = \text{diag}(\underline{1}_{n_1}, \underline{1}_{n_2}, \ldots, \underline{1}_{n_m}) \tag{8}$$

which is the Kronecker product ($\otimes$) of the identity matrix (I) and the vector with values 1 ($\underline{1}$) where p denotes the number of geomaterial layers and n_l denotes the number of observations per layer. Generalised least squares (GLS) was used to estimate the unknown regression coefficients ($\underline{\beta}$). The GLS estimate ($\widehat{\underline{\beta}}$) is given by Equation (9) and the standard errors given by Equation (10) as

$$\widehat{\underline{\beta}} = (X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}\underline{y} \tag{9}$$

$$\widehat{se}(\widehat{\underline{\beta}}) = (\text{diag}[(X'\hat{\Sigma}^{-1}X)^{-1}])^{1/2} \tag{10}$$

The GLM was fit with restricted maximum likelihood (REML) to obtain the estimate of the covariance matrix ($\hat{\Sigma}$). This study illustrates how to evaluate and choose covariance structures using the GLM. Given these estimates $\widehat{\underline{\beta}}$ and $\hat{\Sigma}$, the model can be evaluated using the $-2 \times$ REML log-likelihood (n2llik) (Pinheiro and Bates 2006). The value of n2llik is obtained from gls{nlme} in R for the non-nugget GLM in Equation (2), and from lme{nlme} in R for the nugget GLM in Equation (3). From Harville (1977), the $-2 \times$ REML log-likelihood (n2llik) is given by

$$\begin{aligned} n2llik = {} & (n-p)\ln(2\pi) + \ln(\hat{\Sigma}) + \ln(X'\hat{\Sigma}^{-1}X) \\ & + (\underline{y} - X\widehat{\underline{\beta}})'\hat{\Sigma}^{-1}(\underline{y} - X\widehat{\underline{\beta}}) \end{aligned} \tag{11}$$

Information criteria were obtained from Equation (11). In particular, the Akaike information criterion (AIC), Akaike information criterion with correction for small sample sizes (AICc), and Bayesian information criterion (BIC), were used to compare different spatial covariance structures. Table 1 describes the penalties that define these information criteria, which include

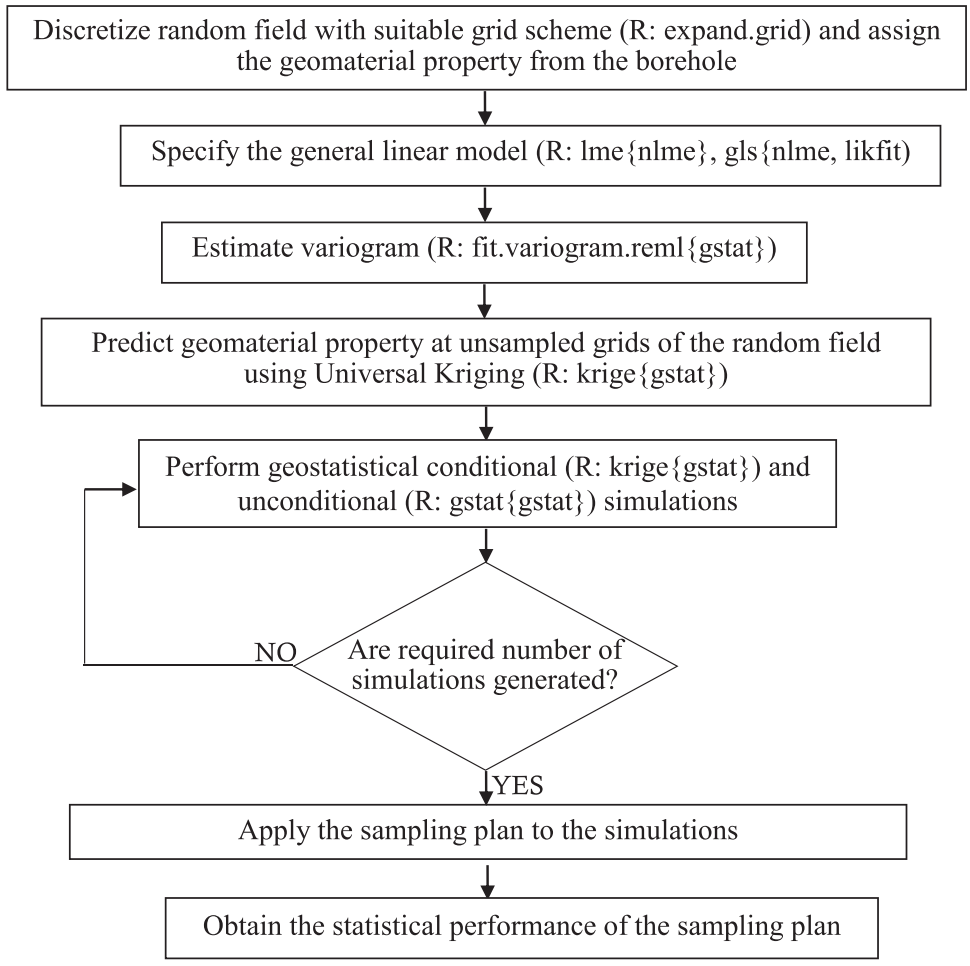

Figure 1. Flowchart for inherent variability and sampling assessment (*R* functions and packages are listed in parentheses).

the number of observations n, p is the number of columns of X, and the number of covariance parameters (c). The assumption of isotropy is also assessed by comparing the information criteria between two models where one model was fit with isotropy and one model was fit with anisotropy using the chosen specification for Σ (R: likfit{geoR}).

This study also shows how to use Cholesky residuals to assess the assumption of multivariate normality and to identify outliers for the GLM. Given the estimated value $\hat{\underline{\beta}}$, these residuals ($\widehat{\underline{r}}$) are given by

$$\widehat{\underline{r}} = (\hat{L}^{-1})'(\underline{y} - X\hat{\underline{\beta}}) = (\hat{L}^{-1})'\widehat{\underline{e}} \tag{12}$$

Table 1. Information criteria used for model selection.

Information criteria	Penalty
AIC	n2llik + $2(p+c)$
AICc	n2llik + $\frac{2(p+c)\times n}{(n-(p+c)-1)}$
BIC	n2llik + $2(p+c)\ln(n-p)$

where $\hat{L}$ is the Cholesky decomposition of $\hat{\Sigma}$ with $(\hat{L}^{-1})(\hat{L}^{-1})' = \hat{\Sigma}^{-1}$ and $(\hat{L}^{-1})'\hat{\Sigma}(\hat{L}^{-1}) = I$ (Houseman, Ryan, and Coull 2004).

2.2. Variogram

The errors ($\underline{E}$) for the GLM in Equation (1), along with the spatial correlation functions in Equations (4), (5), (6) and (7), assume stationarity and isotropy. Under these assumptions, the spatial correlation structure can be represented by the semi-variogram (γ) as a function of lag distance (h):

$$\gamma(h) = \frac{1}{2}\varepsilon[(E_{\underline{s}} - E_{\underline{v}})^2] \tag{13}$$

where $h = ||\underline{s} - \underline{v}||$ represents the Euclidean distance between coordinate $\underline{s}$ and coordinate $\underline{v}$ (Pinheiro and Bates 2006).

The semi-variogram in Equation (13) consists of the scaled expected squared differences as a function of h. An estimate of the semi-variogram, $\hat{\gamma}(h)$, can be obtained by replacing the errors in Equation (13) with

the mean of the residuals $\widehat{\underline{e}}$ in Equation (12) that have a distance h (Pinheiro and Bates 2006, 231). The residuals are standardised by the variance of the errors. Ordinary least squares estimation is used to obtain $\widehat{\underline{\beta}}$, which corresponds to Equation (9) with $\hat{\Sigma} = I$. The sample semivariogram is then created to visualise the averages from the semi-variogram cloud over the standardised residuals at a distance h.

The variogram outlines the geomaterial property variation within a region, and it is needed for kriging prediction (Webster and Oliver 1992). Variogram modelling is often helpful when the variogram is to be used for spatial prediction or simulation. A parametric variogram model is specified and fit to the sample variogram. The variogram models are chosen from among the spatial correlation functions by Equations (4), (5), (6), and (7). REML estimates are obtained for the partial sill, range, and nugget parameters (R: fit.variogram.reml{gstat}).

2.3. Kriging

Kriging is used in spatial statistics to obtain predictions for unsampled locations. Kriging utilises best linear unbiased prediction (BLUP) at an unobserved location based upon the observed data. This study describes and performs universal kriging due to the presence of covariates (R: krige{gstat}) (Bivand et al. 2008). The BLUP for universal kriging is given by:

$$\hat{y}(\underline{s}_0) = \underline{x}(\underline{s}_0)'\hat{\underline{\beta}} + \hat{\underline{\gamma}}'\hat{\Sigma}^{-1}(y(\underline{s}) - X\hat{\underline{\beta}}), \qquad (14)$$

where $\underline{s}$ is the observed spatial location, $\underline{s}_0$ is the unobserved spatial location to be predicted, $\hat{\underline{\beta}}$ is the generalised least squares estimate from Equation (9), $\hat{\Sigma}$ is the REML estimate of Σ, $\underline{x}(\underline{s}_0)$ is value of the design matrix at $\underline{s}_0$, $y(\underline{s})$ is the value of the geomaterial parameter at $\underline{s}$, and $\hat{\underline{\gamma}}$ is a vector of covariances with entries $\widehat{\text{cov}}(Y(\underline{s}_0), Y(\underline{s}_i))$ for $i = 1, \ldots, n$ (Pebesma 2004; Bivand et al. 2008). If there is no nugget, then the predictions at observed locations will match the observed values. However, if a nugget is present, then the predictions at observed locations may not match the observed values (Banerjee, Carlin, and Gelfand 2014).

2.4. Geostatistical simulation

Geostatistical simulation (GS) refers to the simulation of possible realisations of a random field given the mean structure, residual spatial structure, and possibly observed data (Bivand et al. 2008). These simulations are typically Gaussian GS (GGS) based upon the original data or transformed data. GGS is needed to account for the uncertainties associated with the plug-in estimates used to perform kriging using Equation (14). These uncertainties can then be built into the next stage of an analysis. The general idea is to simulate a vector of standard normal values and transform these values using the multivariate standard normal transformation. These algorithms were developed to make the simulation process less computationally expensive (Pebesma 2004). It is possible to include covariates by simulating a value at simulation (s) of the fixed effect parameters ($\hat{\underline{\beta}}^{(s)}$) from the multivariate normal distribution with mean $\hat{\underline{\beta}}$ obtained from Equation (9) and covariance matrix $(X'\hat{\Sigma}^{-1}X)^{-1}$ from Equation(10).

In this study, simulations were conducted using conditional and unconditional approaches. Conditional simulation produces realisations that match the data at the observed locations by working with the conditional distribution given the data (R: krige{gstat}) (Bivand et al. 2008). Unconditional simulation is based upon the marginal normal distribution to reproduce the means and spatial covariance structure (R: gstat{gstat}). In this way, the uncertainties in both the parameter estimates and the responses are maintained and can be simultaneously propagated into additional analyses.

For each simulation, the sample mean of the simulated geomaterial based property was obtained for each geomaterial type across the sampled locations. In this study, these sample means were then summarised across the simulations in terms of the mean, standard deviation, and 95% equal tails percentile interval. By using this sampling-based approach, it was not necessary to assume normality or independence of the sample means across the simulations to obtain the summaries.

The standard deviation, or the length of the percentile interval, was then used to assess the inherent variability of a geomaterial based quality characteristic for a given site investigation plan. The standard deviation reduction method was used whereby the simulated standard deviation of the mean of the simulated parameter is reduced until further reduction was no longer cost effective. The standard deviation formula used in this study is given by

$$SD_l = \frac{1}{n_s - 1}\sum_{i=1}^{n_s}(\hat{y}_{i,l} - \bar{\hat{y}}_l)^2 \qquad (15)$$

where l indexes the layer, i indexes the simulation, n_s is the total number of simulations, $\hat{y}_{i,l}$ is the predicted log SPT N for layer l and simulation i, and $\bar{\hat{y}}_l$ is the sample mean of the predicted log SPT N for layer l and simulation i. Virtual boreholes are initially located near or on the foundation to reduce uncertainty during the

pile foundation design. Subsequent boreholes are thereafter located so as to minimise SD_l in Equation 15 for any of the layers or combination of layers. Thus, optimal boreholes are added to the sampling plan if they suitably reduce the difference between the predicted log mean SPT N and the mean of the predicted log mean SPT N for layer l across the simulations. This method is nicely visualised with scree plots involving the SD_l in Equation 15 versus the number of boreholes. Different site investigation sampling plans were compared and evaluated using these standard deviations.

3. Data

The data for this study was taken from a project site to replace an existing US 63 bridge and culverts over a drainage ditch about 0.3 miles south of Ottumwa, Wapello County, Iowa. A Standard Penetration Test (SPT) was chosen as the in-situ test method because it was the current practice of the Iowa Department of Transportation (IADOT) in investigating the subsurface with IGM layer. The SPT N-value is the number of hammer blows within a borehole to drive a 2-inch split-spoon sampler 12 inches into a geomaterial, and SPT N-values were used as the geomaterial parameter in this study. The advantages of using a borehole with SPT included (i) its suitability for dense geomaterials like IGMs, (ii) sampling of disturbed and undisturbed geomaterials for laboratory testing, and (iii) allowing for in-situ determination of geomaterial parameters. Figure 2 shows the eight boreholes represented with black–white circles along with SPT tests conducted at the borehole locations with respect to the four bridge foundation locations. The foundations were north abutment (N.ABUT), centre pier 1 (C.PIER 1), centre pier 2 (C.PIER2), and south abutment (S.ABUT). The eight boreholes were in line with AASHTO recommendations for bridges (superstructure) with over 100 ft. in width.

The subsurface was characterised in four geomaterial layers (Figure 3) with layer 1 consisting of dark brown-gray stiff clay, layer 2 consisting of brown-gray clayey/fine sand to coarse sand, layer 3 consisting of gray firm to stiff silt with brown firm sand clay with brown-gray firm clay and gravel, and layer 4 consisting of gray slightly to moderately weathered shale, which was considered as the IGM layer. The SPT hammer blows obtained in most samples in layer 4 could not achieve a full 12-inch penetration, and hence the hammer blows were linearly extrapolated using Equation (16) to obtain an equivalent N-value for a consistent

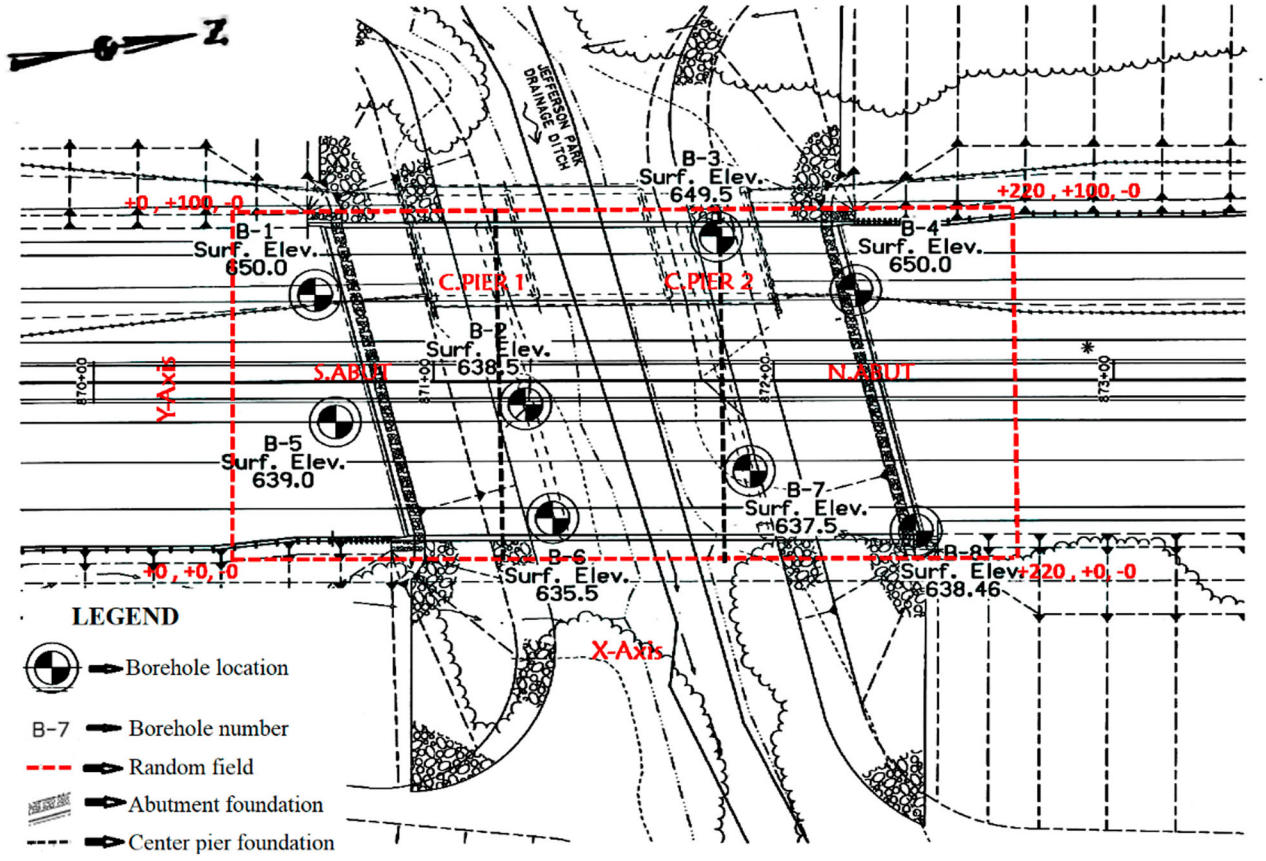

Figure 2. Eight boreholes at the bridge site in Wapello County, Iowa.

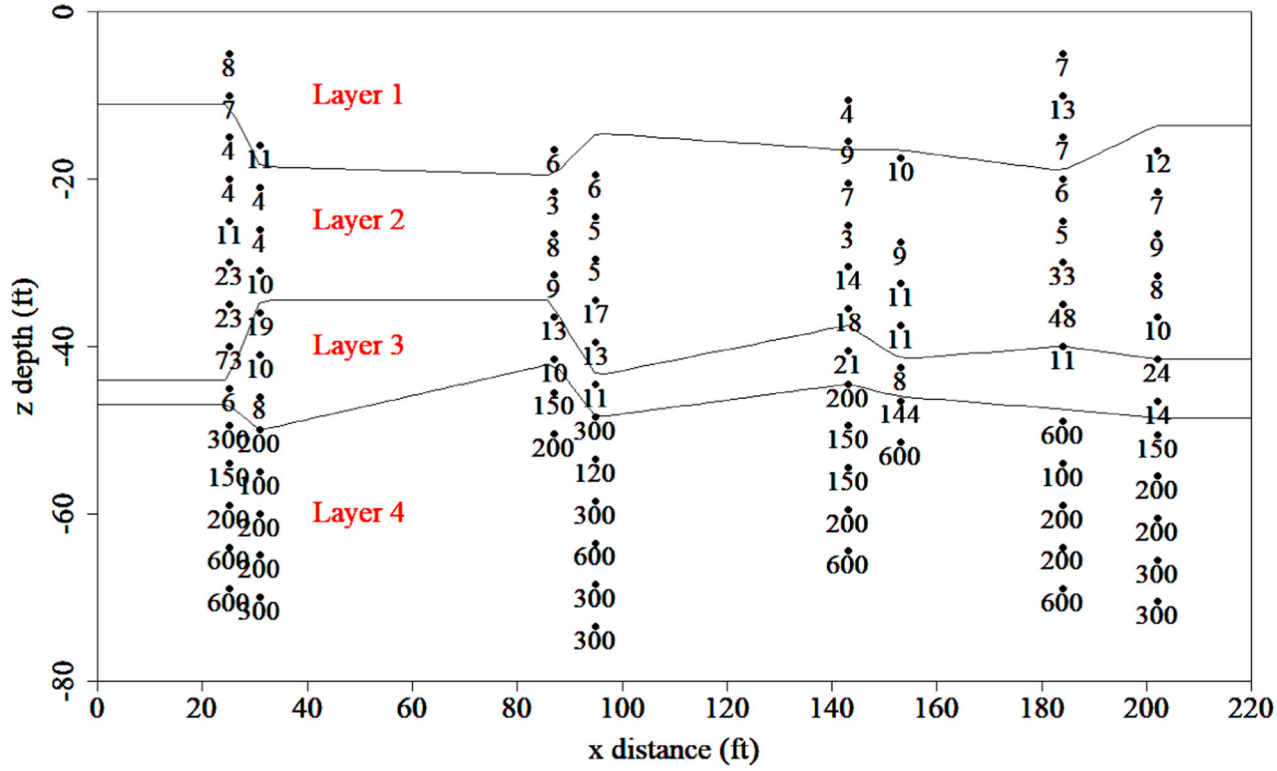

Figure 3. Observed 90 SPT N-values from eight boreholes for the 2D analysis.

and comparable analysis. The study simultaneously considered all four geomaterial layers, so it was necessary to have the SPT N-values on a comparable scale in the analysis. By simultaneously combining the layers this way, there is an important gain in information on the spatial structure of the whole study site. This approach also allows for improved estimation and prediction of the SPT N-values.

$$\text{Extrapolated SPT N value} = \frac{\text{Measured SPT hammer blows}}{\text{Penetration depth before refusal (inches)}} \times 12\ \textit{inches} \quad (16)$$

The data were analyzed both in two and three dimensions. The 2D analysis had all boreholes projected to a straight line at the centre of the y-axis. The 2D analysis allowed for easy and efficient computations. The SPT N-value, layers, x-axis (South–North), and z-axis (Depth) were of interest as shown in Figure 3. The 3D analysis was conducted to reflect the coordinates consistent at the site. The 3D random field of SPT N-values is shown in Figure 4. The dimension of the random field is +220 ft. by +100 ft. by −80 ft. for the x, y, and z axes, respectively, starting at the origin (Figure 2). The negative sign in the z axis is due to the fact that depth is described below the ground surface.

There were a total of 92 SPT N-values obtained from the eight boreholes drilled in the study area. Initial analyses of the Cholesky residuals in Equation (12) from fits of the general linear model in Equation (1) revealed that a lognormal transformation was needed for the SPT N-values to satisfy normality assumptions. The lognormal distribution was archetypal of geomaterial parameters like SPT N-values (Grasmick et al. 2020). With this logarithm transformation, two outliers were identified which had Cholesky residual values with magnitude larger than 3. The outliers corresponded to borehole 4 (x = +184 ft., z = −44.0 ft., SPT N-value = 120, layer = 3) and borehole 7 (x = +153 ft., z = −22.5 ft., SPT N-value = 1, layer = 2). After the removal of outliers, all analyses were conducted on the dataset containing 90 observations.

4. Results and discussion

Various spatial covariance models were examined assuming different mean log SPT N-values for each of the four layers. The system of analysis of inherent variability proposed in this study included identifying an adequate spatial covariance structure using the GLM, comparing fits with the variogram, generating prediction surfaces, and obtaining simulations for assessing site investigation plans. Each of these aspects are addressed in the following subsections.

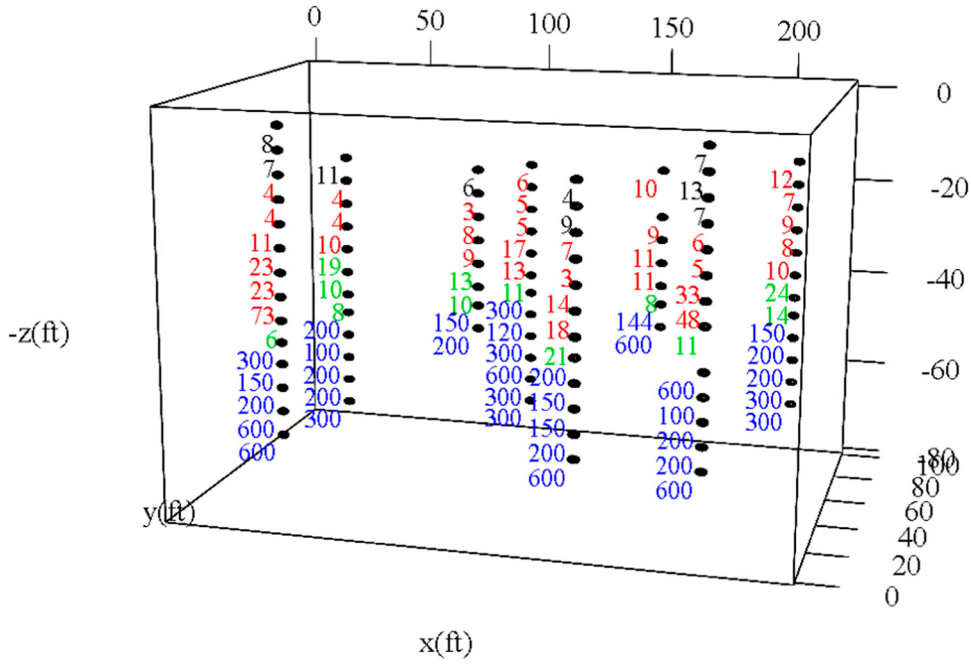

Figure 4. Observed 90 SPT N-values from eight boreholes for the 3D analysis.

4.1. General linear model

The GLM in Equation (1) was fit to log SPT N-values for the spatial covariance structures in Equations (4), (5), (6), and (7) for both the 2D and 3D analyses. For the 2D analysis, the Euclidean distances (h) were calculated based upon pairs of x and z coordinates, whereas for the 3D analysis, the Euclidean distances (h) were calculated based upon pairs of the x, y, and z coordinates.

For all spatial structures, the nugget variance τ^2 in Equation (3) was estimated to be 0, and so it was not considered for inclusion in the model. The information criteria for the GLM given in Table 2 were calculated for both 2D and 3D based distances, where smaller is better for these values. It is evident that the Null model with no spatial correlation is inferior to the other four models that incorporate spatial correlation. The information criteria values for the spatial correlation structures were nearly identical for the 2D analysis and very similar for the 3D analysis. The Spherical structure had slightly smaller information criteria values for the 3D analysis. Nevertheless, the Exponential covariance structure was selected since it was close to the smallest values for both the 2D and 3D analyses, it is a simple structure, and it has been applied in other studies (Li, Genton, and Sherman 2008; Littell, Pendergast, and Natarajan 2000; Hicks 1996).

The assumption of isotropy was checked in this study by fitting an anisotropic model for log SPT N-values with the Exponential spatial structure. This assumption was checked based upon 2D analysis of the (x,z) and (y, z) coordinate systems. The values for n2llik, AIC, AICc, and BIC using the formulas in Table 2 are shown in Table 3 for the Exponential spatial structure. The information criteria favor the more parsimonious isotropic model.

The null hypothesis that the model errors were all normally distributed was also assessed using the Cholesky residuals in Equation (12) for the estimated Exponential covariance function in the 2D and 3D analyses. The Shapiro–Wilk (W) normality test of the Cholesky residuals had little evidence against the claim that the log SPT N responses were normally distributed for the 2D analysis (W test statistic = 0.9783, p-value = 0.1382) and for the 3D analysis (W test statistic = 0.9792, p-value = 0.1591) as these p-values exceed 0.1.

Table 2. Information criteria (smaller is better) of the GLM for the 2D and 3D analyses.

	2D				3D			
GLM	n2llik	AIC	AICc	BIC	n2llik	AIC	AICc	BIC
Null	169.299	179.299	180.013	191.571	169.299	179.299	180.013	191.571
Exponential	163.995	175.995	177.007	190.721	165.160	177.160	178.172	191.886
Gaussian	163.438	175.438	176.449	190.164	164.075	176.075	177.087	190.801
Linear	163.794	175.794	176.806	190.520	164.071	176.071	177.083	190.797
Spherical	163.458	175.458	176.470	190.184	163.988	175.988	177.000	190.714

Table 3. Information criteria (smaller-is-better) of the GLM with Exponential spatial correlation for the 2D analyses with the (*x*, *z*) and (*y*, *z*) coordinate systems.

	(*x*, *z*)				(*y*, *z*)			
GLM	n2llik	AIC	AICc	BIC	n2llik	AIC	AICc	BIC
Isotropic	163.995	175.995	177.007	190.721	163.811	175.811	176.824	190.538
Anisotropic	162.283	178.283	180.061	197.918	159.935	175.935	177.712	195.569

4.2. Variogram

The variogram of the logarithm SPT N-values was fit using the spatial correlation functions in Equations (4), (5), (6), and (7). These variograms for the 2D and 3D analyses are displayed in Figure 5(a,b), respectively, to visualise the partial sill, where the variogram reaches its plateau, and the range, which corresponds to the distance *h* at which the sill is attained (Bivand et al. 2008). It is evident that the Null model does not provide a good fit. However, the Exponential (Exp), Gaussian (Gaus), Linear (Lin), and Spherical (Spher) spatial correlation functions all fit similarly, which confirms the previous information criteria comparisons for the GLM. In fact, the Gaussian and Spherical fits are nearly identical in both the 2D and 3D analyses. The most visible differences in these fits are in the range, which differ the most between the Exponential and Linear spatial correlations. Nevertheless, the spatial correlation fits are quite similar as suggested by the information criteria.

4.3. Kriging

Kriging was used to obtain predicted 2D and 3D surfaces of SPT N values on the log and original scales for the 2D analysis (Figure 6) and the 3D analysis (Figure 7). These predictions were based upon modelling log SPT N-values with the Exponential spatial correlation function, and back-transformation of the log SPT N-values was used to obtain predictions on the original scale. The predicted SPT N-values are the same as the data at observed locations since the covariance model does not include a nugget. For example, a SPT N-value of 8 in Figure 3 has the same value at the exact position ($x = 25$ft., $z = -5$ft.) on the kriging plot in Figure 6(b).

4.4. Geostatistical simulation and optimal borehole location determination

Gaussian Geostatistical Simulation (GGS) was used to evaluate site investigation plans in terms of the number and placement of boreholes. One hundred simulations were obtained both conditionally and unconditionally from the predicted surfaces of log SPT N-values. The number of simulations was adequate as additional simulations resulted in less than ±1% change in variation of the log SPT N-values (Pyrcz and Deutsch 2014).

To reduce uncertainty during pile foundation design, boreholes are better located near or on a foundation (Arsyad et al. 2010; Goldsworthy et al. 2007). Thus, to

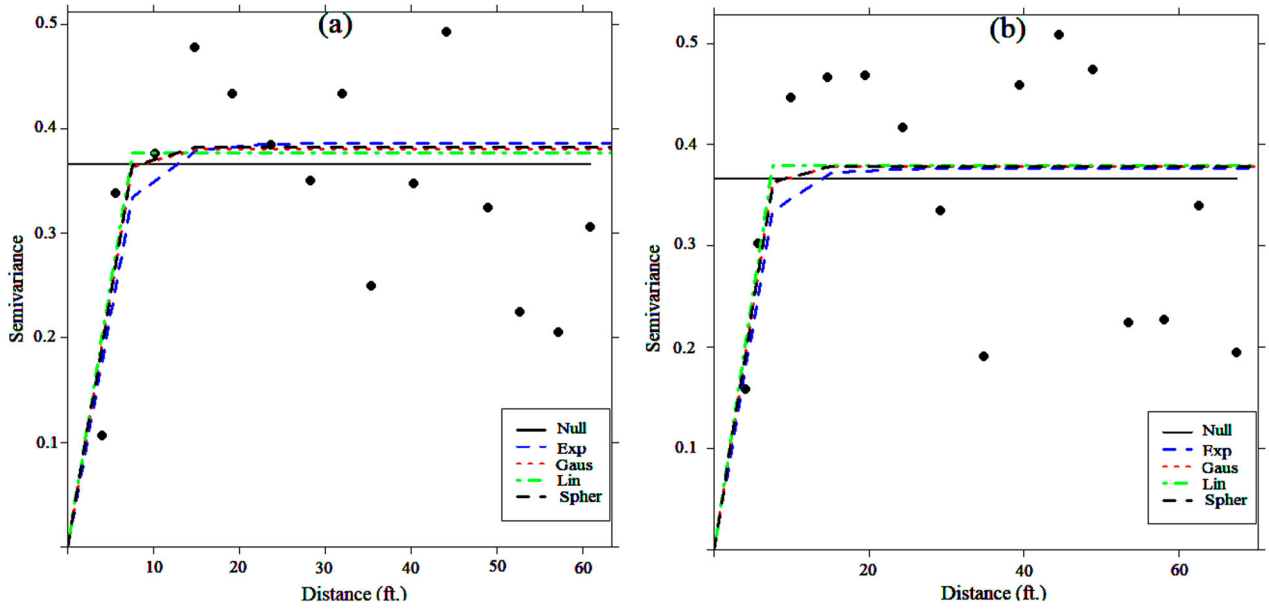

Figure 5. Variograms for the (a) 2D analysis and (b) 3D analysis.

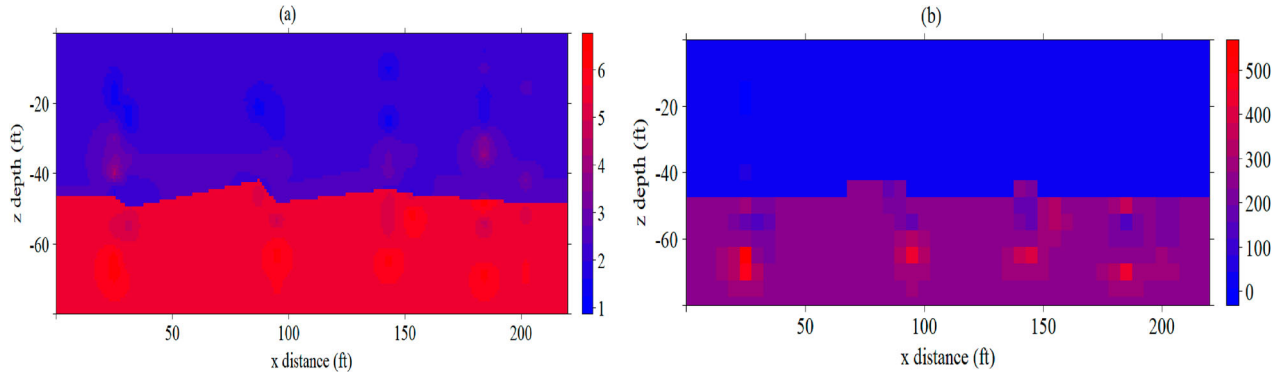

Figure 6. 2D plot of (a) predicted log SPT N and (b) back-transformed SPT N values using universal kriging.

obtain a site investigation plan with the lowest standard deviation (uncertainty), four virtual boreholes, marked as 1st to 4th BHs, were each placed at the four foundation locations of the 100 simulations. Subsequent borehole locations were determined in the 2D analysis by placing virtual boreholes at every 5 ft. position throughout the *x*-axis (over the 220 ft. in the *x*-axis), except for where virtual boreholes were already located. The idea was to select those boreholes that yielded the lowest standard deviation for the highly variable IGM layer (layer 4, the target layer). This was done by selecting boreholes which minimised SD_l in Equation 15 with respect to layer 4. Once optimal boreholes were identified, they were then added to the bulk of previously selected boreholes and analyzed to determine the uncertainty for the site. Another aspect of the site investigation plan is the sampling frequency. The two sampling frequencies considered in this study were (i) taking SPT N samples at every 5 ft. from ground to the terminating depth of 80 ft. (all samples per layer) and (ii) taking one sample per layer. Thus, sampling frequency approach (i) is dependent upon the layer thickness whereas sampling frequency approach (ii) is not.

The subsequent borehole locations for the 2D analysis by taking all samples per layer for the conditional and unconditional simulation is shown in Figures 8 and 9, respectively. The black circular markers in these figures suggest the two virtual borehole locations on the *x*-axis with the lowest standard deviation values in the IGM layer. An important observation from Figures 8 and 9 is that the line plot flattens out especially in layer 4 (IGM layer) after the addition of the 6th borehole. This suggests that six (6) boreholes are optimally sufficient for the reduction of parameter uncertainty of the site in the IGM layer.

In this study, the 3D analysis was conducted to determine the corresponding *y*-axis for the already obtained *x*-axis locations in the 2D analysis described in the previous paragraph. Virtual boreholes were then placed at every 5 ft. along the breadth of the *y*-axis corresponding to the already obtained *x*-axis locations. The position on the *y*-axis with the lowest standard deviation after the analysis was chosen as the corresponding *y* coordinate as shown in Figure 10. The sampling locations for 12 boreholes on the *x*- and *y*-axes as obtained from the conditional simulation (Figure 8) and the unconditional simulation (Figure 9) are displayed in Table 4.

The previously described methodology was also used to identify borehole locations for taking one sample per layer. The 12 borehole locations on the *x* and *y* axes

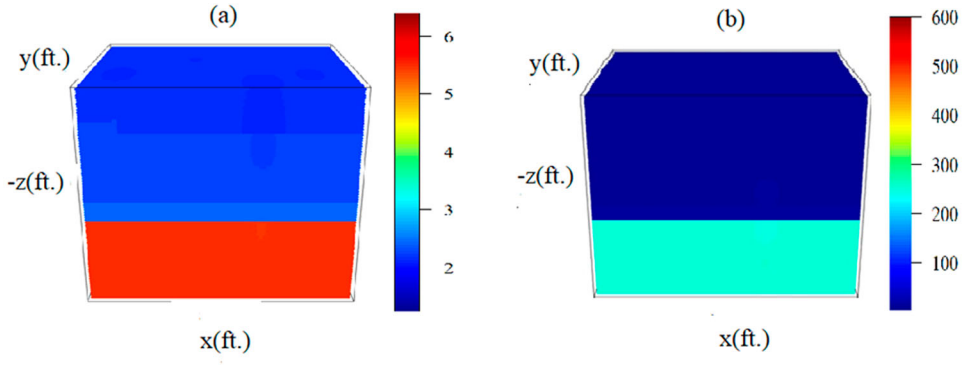

Figure 7. 3D plot of (a) predicted log SPT N and (b) back-transformed SPT N values using universal kriging.

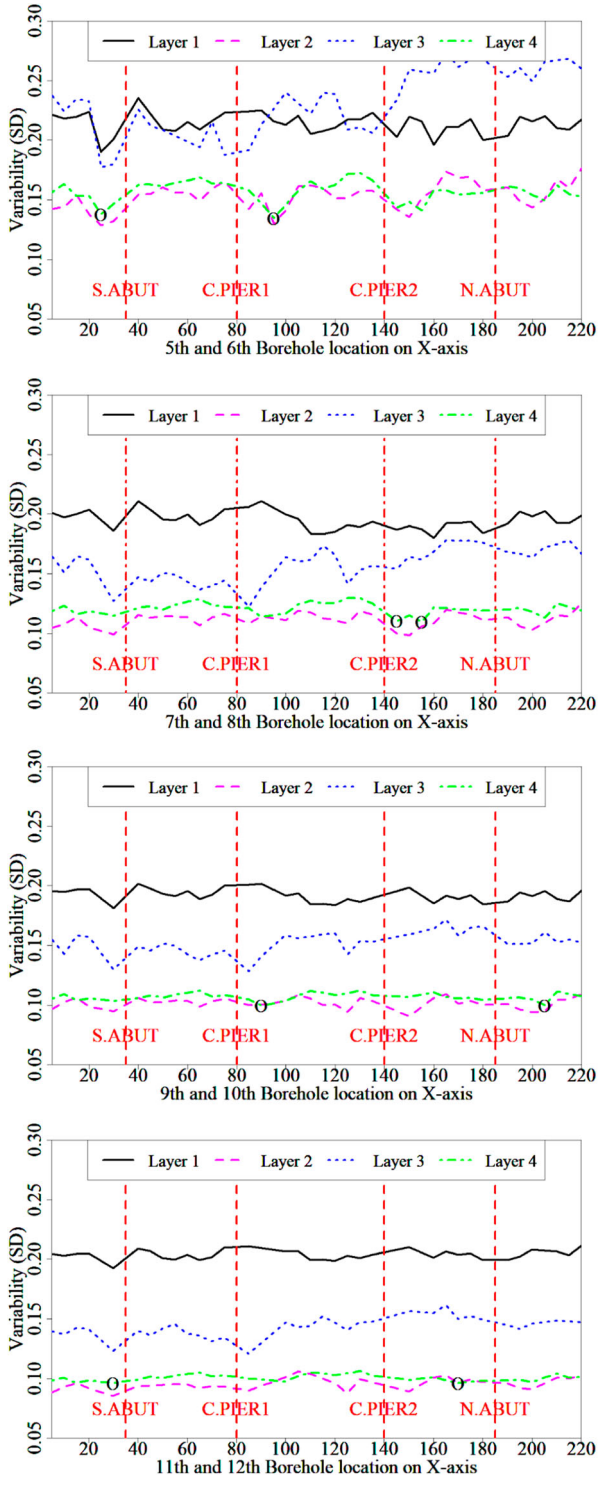

Figure 8. Subsequent borehole locations from conditional simulation for sampling all depths per layer.

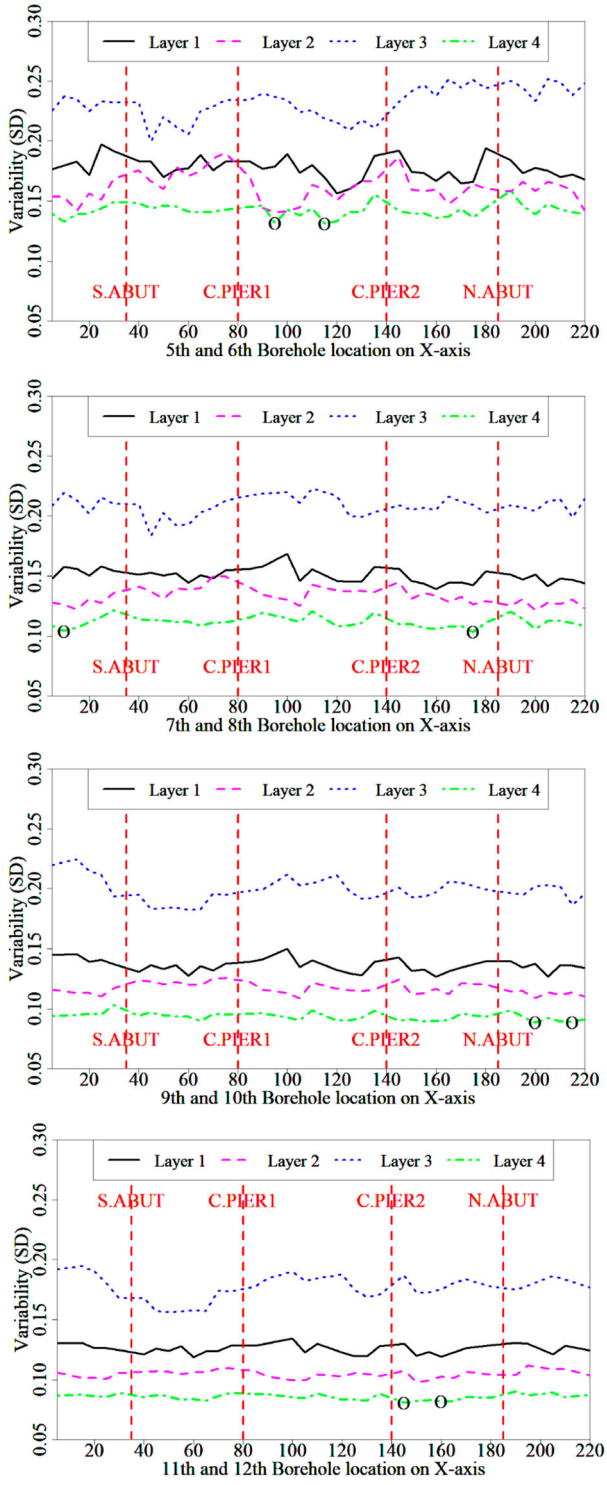

Figure 9. Subsequent borehole locations from unconditional simulation for sampling all depths per layer.

obtained from the conditional and unconditional simulation are displayed in Table 5. These sampling locations had a range of ±3 ft. Sampling from these range resulted in no change or less than 1% increase in the standard deviation.

4.5. Discussion

The inherent variability in the IGM layer, which was assessed by quantifying the standard deviation of the unconditional simulation for the different site investigation plans and different sampling frequencies, is

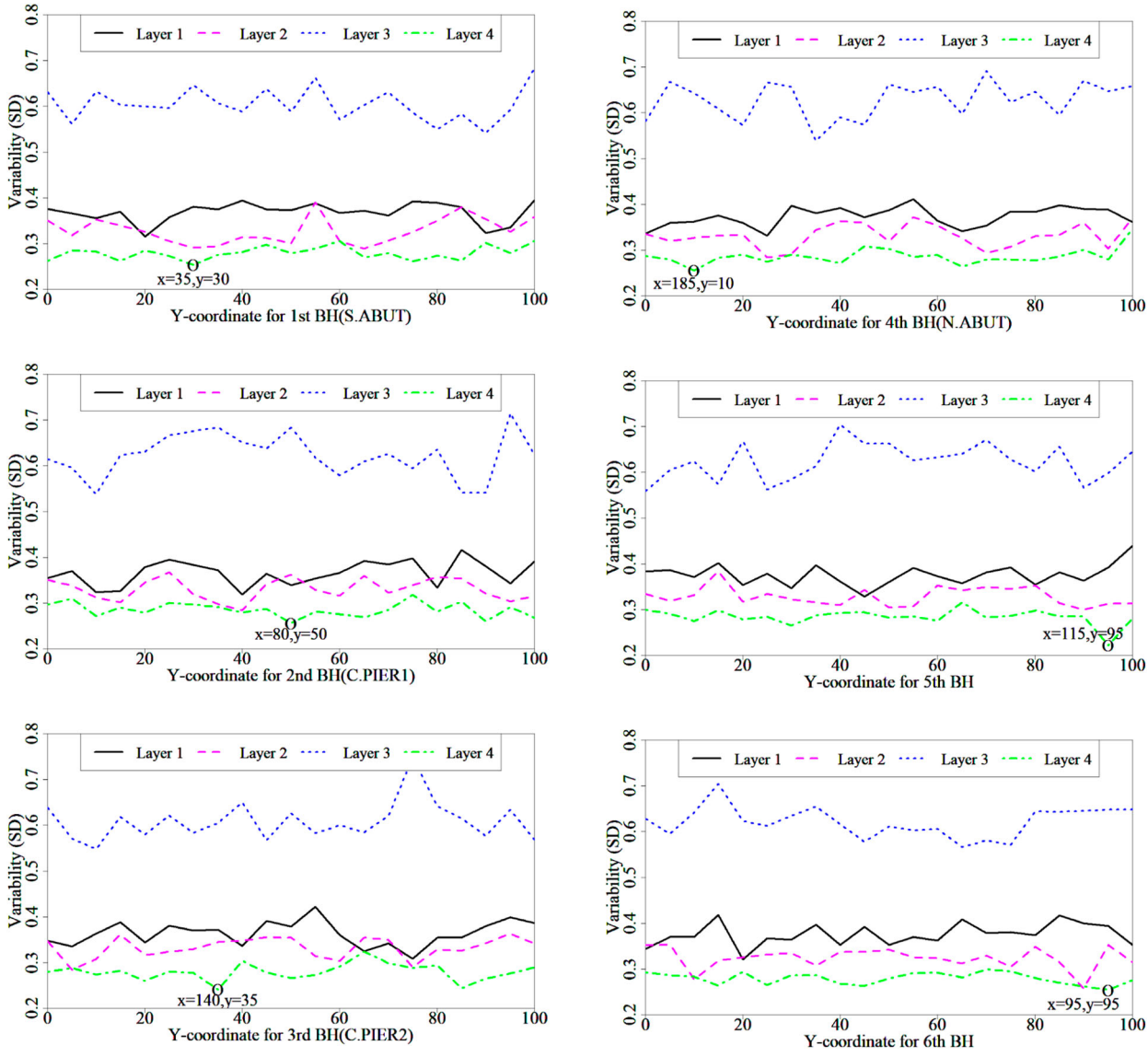

Figure 10. The *y* coordinates from unconditional simulation of sampling all depths per layer for the first six boreholes.

Table 4. Conditional and unconditional borehole position coordinates for sampling at all depths per layer.

		Location of Borehole Number (ft)											
Simulation type	Coordinate	1	2	3	4	5	6	7	8	9	10	11	12
Conditional	*X*-axis	35	80	140	185	95	25	155	145	90	205	30	170
	Y-axis	0	5	15	30	95	30	20	10	25	5	70	65
Unconditional	*X*-axis	35	80	140	185	115	95	175	10	200	215	145	160
	Y-axis	30	50	35	10	95	95	15	45	85	90	35	90

Table 5. Conditional and unconditional borehole position coordinates for sampling at one depth per layer.

		Location of Borehole Number (ft)											
Simulation type	Coordinate	1	2	3	4	5	6	7	8	9	10	11	12
Conditional	*X*-axis	35	80	140	185	215	170	95	150	125	65	90	45
	Y-axis	75	15	100	30	25	60	30	75	90	0	95	40
Unconditional	*X*-axis	35	80	140	185	170	5	150	145	10	60	100	115
	Y-axis	15	30	30	90	50	35	75	0	65	45	5	100

Table 6. Standard deviation (SD) in the IGM layer from the unconditional simulation of the different site investigation plans and sampling frequencies.

		SD based on number of boreholes							
Sampling frequency	Sampling plan	1	2	3	4	5	6	7	8
All samples per layer	Unconditional	64.90	46.11	35.76	31.05	26.49	24.46	22.54	21.60
	Conditional	65.52	51.68	42.50	37.50	30.47	27.67	25.89	25.61
	IADOT	69.89	46.30	40.68	34.61	31.33	29.87	28.18	25.68
One sample per layer	Unconditional	133.22	100.02	82.44	72.50	68.05	58.84	55.80	48.51
	Conditional	143.29	110.22	83.35	77.98	71.34	67.60	61.90	53.58
	IADOT	178.93	104.26	91.87	77.01	67.76	63.99	57.91	55.02

presented in Table 4. The unconditional simulation does not match the data at the observed locations, because it is completely independent and only shares spatial moments with data (Pebesma and Wesseling 1998). Thus, it provided a basis to compare the standard deviation in the geomaterial parameter of different sampling plans at sites with a similar geomaterial profile. The standard deviation values were a bit high as shown in Table 6 due to the high mean SPT N-value (≈ 290) in layer 4. Table 6 shows that the unconditionally obtained sampling plan had the lowest standard deviation for both sampling frequencies. However, sampling at all depths per layer (equivalent of every 5 ft.) had more than 50% reduction in the SPT N-value standard deviation compared to sampling at one depth per layer. A comparison of eight boreholes from the proposed site investigation plans with the actual site investigation plan conducted by IADOT is shown in Figure 11. According to Figure 11, the first six boreholes on all three sampling plans had almost the same longitudinal coordinates, except for the 5th borehole in IADOT and unconditional sampling plan and the 6th borehole in the conditional sampling plan. This was also in agreement with the results in Figures 8 and 9, as the line plot flattens out especially in the IGM layer 4 after the 6th borehole. The first six boreholes on unconditional and IADOT sampling plans had almost the same transverse coordinates, except for the 5th BH in IADOT sampling plan. Table 6 shows that for all samples per layer, the difference in the standard deviation after adding the 7th borehole became less significant compared to the addition of the previous boreholes. For example, the addition of the 5th, 6th, and 7th boreholes in the unconditional sampling plan for all samples per layer (Table 6) resulted in approximately 17.2%, 8.4%, 8.5% decrease in the standard deviation, respectively. The addition of the 7th borehole resulted in little or no difference in standard deviation reduction which may be responsible for the curve flattening out in Figures 8 and 9 after 6 boreholes. One could conclude that the optimal number of boreholes for the present case study was six (6). Another vivid observation from the study was the fact that location matters for reduced standard deviation when drilling borehole(s). For example, according to all samples per layer in Table 6, six boreholes from the sampling plan obtained from unconditional simulation

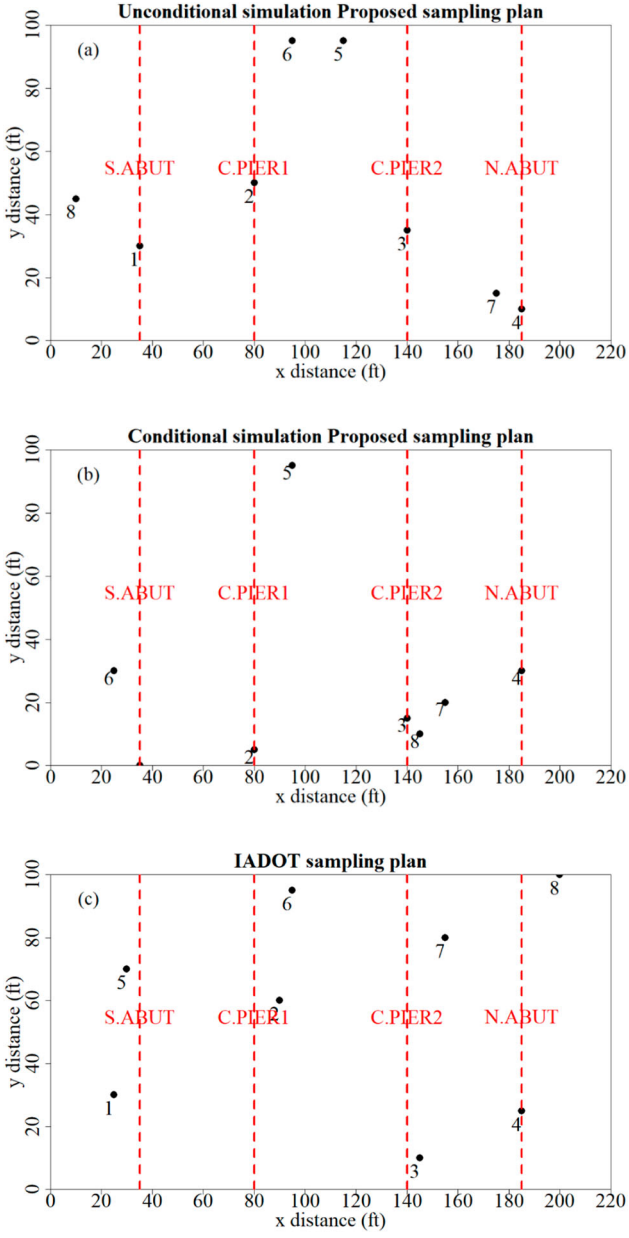

Figure 11. Comparison of different site investigation plans.

of present data had a lower standard deviation (24.46) compared to that of eight (8) boreholes in the plans from the conditional simulation (25.61) and IADOT (25.68). The lower standard deviation of unconditional simulation compared to the conditional simulation could be due to the assumption of the normal distribution as the conditional simulation makes no such assumptions, and so is nonparametric.

5. Conclusions

This study presented an effective and systematic method to quantify inherent variability with respect to a geomaterial parameter, and to reduce this variability through proposed geotechnical investigation plans. Various spatial covariance structures were fit with and without a nugget. The fits of the spatial correlation functions were visualised with the variogram and compared via information criteria associated with fits of the Generalised Linear Model. The chosen covariance model in this study had no nugget, an Exponential correlation function, and assumed isotropy. Predictions were obtained with respect to the selected covariance model in 2D and 3D analyses using universal kriging. Conditional and unconditional simulations were conducted based upon the predictions at different sampling frequencies with depth. The site investigation plan was obtained, which best reduced the standard deviation while also minimising the required number of boreholes. It was found, as expected, that taking samples at all depths (an equivalent of every 5 ft.) resulted in approximately 50% lower variability compared to taking one sample per layer. However, it must be remembered that the amount of variability depends on layer thickness when taking samples at all depths. The unconditional simulation tended to produce an approximately 18.6% lower variability than the conditional simulation, which is likely due to the assumption of normality, which appeared reasonable for the log SPT N-values at this site. The proposed plan fared well in standard deviation reduction compared to the actual plan carried out by IADOT at this site. The proposed unconditional sampling plan can also be applied at another site with a similar subsurface profile. The proposed method is applicable to addressing inherent variability and site investigation plans in any multilayered geomaterial. This is because layers are analyzed collectively with different means and with respect to the geomaterial parameter of interest. For future analyses with limited data, a Bayesian approach could be considered with carefully constructed informative priors on the covariance parameters.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Wyoming Department of Transportation as the lead agency, Colorado Department of Transportation, Iowa Department of Transportation, Kansas Department of Transportation, North Dakota Department of Transportation, Idaho Transportation Department and Montana Department of Transportation [grant number RS05219].

Data availability statement

The data that support the findings of this study are available from the corresponding author, Kam Ng, upon reasonable request.

ORCID

Opeyemi E. Oluwatuyi http://orcid.org/0000-0002-0499-6998
Shaun S. Wulff http://orcid.org/0000-0002-5695-4925
Kam Ng http://orcid.org/0000-0001-5099-5454

References

Adhikari, P., K. W. Ng, Y. Z. Gebreslasie, and S. S. Wulff. 2020. "Static and Economic Analyses of Driven Steel H-Piles in IGM Using the WyoPile Database." *Journal of Bridge Engineering* 25 (5): 040200161–0402001615. doi:10.1061/(ASCE)BE.1943-5592.0001546.

Al-Bittar, Tamara, Abdul Hamid Soubra, and Jawad Thajeel. 2018. "Kriging-Based Reliability Analysis of Strip Footings Resting on Spatially Varying Soils." *Journal of Geotechnical and Geoenvironmental Engineering* 144: 10. doi:10.1061/(ASCE)GT.1943-5606.0001958.

Aladejare, Adeyemi Emman, and Yu Wang. 2017. "Evaluation of Rock Property Variability." *Georisk* 11 (1): 22–41. doi:10.1080/17499518.2016.1207784.

Arsyad, A., M. Jaksa, W. Kaggwa, and Y. Mitani. 2010. "Effect of Radial Distance of a Single CPT Sounding on the Probability of Over-and Under-Design of Pile Foundation." In *Proceedings of the First Makassar International Conference on Civil Engineering (MICCE2010)*, 1–5. Makassar, Indonesia.

Baecher, G. B., and J. T. Christian. 2005. *Reliability and Statistics in Geotechnical Engineering*. Chichester: John Wiley & Sons.

Banerjee, S., B. P. Carlin, and A. E. Gelfand. 2014. *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton, FL: CRC Press.

Bivand, R. S., E. J. Pebesma, V. Gómez-Rubio, and E. J. Pebesma. 2008. *Applied Spatial Data Analysis with R*. New York: Springer.

Boyd, D. Lane, Whitney Trainor-Guitton, and Gabriel Walton. 2018. "Assessment of Rock Unit Variability

Through Use of Spatial Variograms." *Engineering Geology* 233: 200–212. doi:10.1016/j.enggeo.2017.12.012.

Cai, Yongmin, Jinhui Li, Xueyou Li, Dianqing Li, and Limin Zhang. 2019. "Estimating Soil Resistance at Unsampled Locations Based on Limited CPT Data." *Bulletin of Engineering Geology and the Environment* 78 (5): 3637–3648. doi:10.1007/s10064-018-1318-2.

Chen, Dongfang, Dingping Xu, Gaofeng Ren, Quan Jiang, Guofeng Liu, Liangpeng Wan, and Ning Li. 2019. "Simulation of Cross-Correlated Non-Gaussian Random Fields for Layered Rock Mass Mechanical Parameters." *Computers and Geotechnics* 112: 104–119. doi:10.1016/j.compgeo.2019.04.012.

Cho, Sung Eun. 2012. "Probabilistic Analysis of Seepage That Considers the Spatial Variability of Permeability for an Embankment on Soil Foundation." *Engineering Geology* 133–134: 30–39. doi:10.1016/j.enggeo.2012.02.013.

Christakos, G. 2012. *Random Field Models in Earth Sciences*. Chelmsford: Courier Corporation.

Christian, J. T., C. C. Ladd, and G. B. Baecher. 1994. "Reliability Applied to Slope Stability Analysis." *Journal of Geotechnical Engineering* 120 (12): 2180–2207. doi:10.1061/(ASCE)0733-9410(1994)120:12(2180).

Crisp, M. P., M. B. Jaksa, and Y. L. Kuo. 2020a. "Effect of Borehole Location on Pile Performance." *Georisk*, 1–16. doi:10.1080/17499518.2020.1757721.

Crisp, M. P., M. B. Jaksa, and Y. L. Kuo. 2020b. "Toward a Generalized Guideline to Inform Optimal Site Investigations for Pile Design." *Canadian Geotechnical Journal* 57 (8): 1119–1129. doi:10.1139/cgj-2019-0111.

Crisp, Michael, Mark Jaksa, Yien Kuo, Gordon A. Fenton, and Vaughan Griffiths. 2020. "Characterizing Site Investigation Performance in a Two Layer Soil Profile." *Canadian Journal of Civil Engineering*, 1–9. doi:10.1139/cjce-2019-0416.

Dasaka, S. M., and L. M. Zhang. 2012. "Spatial Variability of in Situ Weathered Soil." *Geotechnique* 62 (5): 375–384. doi:10.1680/geot.8.P.151.3786.

Deng, Zhi Ping, Dian Qing Li, Xiao Hui Qi, Zi Jun Cao, and Kok Kwang Phoon. 2017. "Reliability Evaluation of Slope Considering Geological Uncertainty and Inherent Variability of Soil Parameters." *Computers and Geotechnics* 92: 121–131. doi:10.1016/j.compgeo.2017.07.020.

Diggle, P. J., and P. J. Ribeiro Jr. 2002. "Bayesian Inference in Gaussian Model-Based Geostatistics." *Geographical and Environmental Modelling* 6 (2): 129–146. doi:10.1080/1361593022000029467.

Elkateb, T., R. Chalaturnyk, and P. K. Robertson. 2003. "An Overview of Soil Heterogeneity: Quantification and Implications on Geotechnical Field Problems." *Canadian Geotechnical Journal* 40 (1): 1–15. doi:10.1139/t02-090.

Goldsworthy, J. S., M. B. Jaksa, G. A. Fenton, W. S. Kaggwa, V. Griffiths, and H. G. Poulos. 2007. "Effect of Sample Location on the Reliability Based Design of Pad Foundations." *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* 1 (3): 155–166. doi:10.1080/17499510701697377.

Grasmick, Jacob G., Michael A. Mooney, Whitney J. Trainor-Guitton, and Gabriel Walton. 2020. "Global Versus Local Simulation of Geotechnical Parameters for Tunneling Projects." *Journal of Geotechnical and Geoenvironmental Engineering* 146 (7): 04020048. doi:10.1061/(asce)gt.1943-5606.0002262.

Harville, D. A. 1977. "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems." *Journal of the American Statistical Association* 72 (358): 320–338. doi:10.2307/2286796.

Hicks, Paul J. 1996. "Unconditional Sequential Gaussian Simulation for 3-D Flow in a Heterogeneous Core." *Journal of Petroleum Science and Engineering* 16 (4): 209–219. doi:10.1016/S0920-4105(96)00041-1.

Houseman, E. A., L. M. Ryan, and B. A. Coull. 2004. "Cholesky Residuals for Assessing Normal Errors in a Linear Model with Correlated Outcomes." *Journal of the Americal Statistical Association* 99 (466): 383–394. doi:10.1198/016214504000000403.

Jamshidi Chenari, Reza, and Hassan Kamyab Farahbakhsh. 2015. "Generating Non-Stationary Random Fields of Auto-Correlated, Normally Distributed CPT Profile by Matrix Decomposition Method." *Georisk* 9 (2): 96–108. doi:10.1080/17499518.2015.1033429.

Li, B., M. G. Genton, and M. Sherman. 2008. "Testing the Covariance Structure of Multivariate Random Fields." *Biometrika* 95 (4): 813–829. doi:10.1093/biomet/asn053.

Li, X. Y., L. Zhang, and J. H. Li. 2016. "Using Conditioned Random Field to Characterize the Variability of Geologic Profiles." *Journal of Geotechnical and Geoenvironmental Engineering* 142 (4): 1–11. doi:10.1061/(ASCE)GT.1943-5606.0001428.

Littell, R. C., J. Pendergast, and R. Natarajan. 2000. "Modelling Covariance Structure in the Analysis of Repeated Measures Data." *Statistics in Medicine* 19 (13): 1793–1819. doi:10.1002/1097-0258(20000715)19:13<1793::AID-SIM482>3.0.CO;2-Q.

Lloret-Cabot, M. F. G. A., G. A. Fenton, and M. A. Hicks. 2014. "On the Estimation of Scale of Fluctuation in Geostatistics." *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* 8 (2): 129–140. doi:10.1080/17499518.2013.871189.

Lloret-Cabot, M., M. A. Hicks, and A. P. van Den Eijnden. 2012. "Investigation of the Reduction in Uncertainty Due to Soil Variability When Conditioning a Random Field Using Kriging." *Geotechnique Letters* 2 (7–9): 123–127. doi:10.1680/geolett.12.00022.

Pebesma, E. J. 2004. "Multivariable Geostatistics in S: The Gstat Package." *Computers & Geosciences* 30 (7): 683–691. doi:10.1016/j.cageo.2004.03.012.

Pebesma, E. J., and C. G. Wesseling. 1998. "Gstat: A Program for Geostatistical Modelling, Prediction and Simulation." *Computers and Geosciences* 24 (1): 17–31. doi:10.1016/S0098-3004(97)00082-4.

Phoon, K. K., and F. H. Kulhawy. 1999. "Characterization of Geotechnical Variability." *Canadian Geotechnical Journal* 36 (4): 612–624. doi:10.1139/t99-038.

Pinheiro, J., and D. Bates. 2006. *Mixed-Effects Models in S and S-PLUS*. New York, NY: Springer Science & Business Media.

Pyrcz, M. J., and C. V. Deutsch. 2014. *Geostatistical Reservoir Modeling*. New York: Oxford University Press.

Simpson, B. 2003. *Eurocode 7. In Limit State Design in Geotechnical Engineering Practice: (With CD-ROM)*. Singapore: World Scientific.

Soubra, Abdul Hamid, Tamara Al-Bittar, J. Thajeel, and Ashraf Ahmed. 2019. "Probabilistic Analysis of Strip Footings

Resting on Spatially Varying Soils Using Kriging Metamodeling and Importance Sampling." *Computers and Geotechnics* 114: 103107. doi:10.1016/j.compgeo.2019.103107.

Stuedlein, A. W., S. L. Kramer, P. Arduino, and R. D. Holtz. 2012. "Geotechnical Characterization and Random Field Modeling of Desiccated Clay." *Journal of Geotechnical and Geoenvironmental Engineering* 138 (11): 1301–1313. doi:10.1061/(ASCE)GT.1943-5606.0000723.

Tian, Mi, Dian Qing Li, Zi Jun Cao, Kok Kwang Phoon, and Yu Wang. 2016. "Bayesian Identification of Random Field Model Using Indirect Test Data." *Engineering Geology* 210: 197–211. doi:10.1016/j.enggeo.2016.05.013.

Vanmarcke, E. 2010. *Random Fields: Analysis and Synthesis*. Singapore: World Scientific.

Wang, Y., S. K. Au, and Z. Cao. 2010. "Bayesian Approach for Probabilistic Characterization of Sand Friction Angles." *Engineering Geology* 114 (3–4): 354–363. doi:10.1016/j.enggeo.2010.05.013.

Wang, Yu, Tengyuan Zhao, and Kok Kwang Phoon. 2018. "Direct Simulation of Random Field Samples from Sparsely Measured Geotechnical Data with Consideration of Uncertainty in Interpretation." *Canadian Geotechnical Journal* 55 (6): 862–880. doi:10.1139/cgj-2017-0254.

Webster, R., and M. A. Oliver. 1992. "Sample Adequately to Estimate Variograms of Soil Properties." *Journal of Soil Science* 43 (1): 177–192. doi:10.1111/j.1365-2389.1992.tb00128.x.

Yang, Rui, Jinsong Huang, D. V. Griffiths, Jingjing Meng, and Gordon A. Fenton. 2019. "Optimal Geotechnical Site Investigations for Slope Design." *Computers and Geotechnics* 114. doi:10.1016/j.compgeo.2019.103111.

Yu, Xin, Jinguo Cheng, Chunhui Cao, Erqiang Li, and Jili Feng. 2019. "Probabilistic Analysis of Tunnel Liner Performance Using Random Field Theory." *Advances in Civil Engineering*, doi:10.1155/2019/1348767.

Zhao, H. F., Limin Zhang, Y. Xu, and D. S. Chang. 2013. "Variability of Geotechnical Properties of a Fresh Landslide Soil Deposit." *Engineering Geology* 166: 1–10. doi:10.1016/j.enggeo.2013.08.006.