Optimization of CPT Soundings to Reduce Uncertainty in Interpretation of Subsurface Stratigraphy

Jason T. DeJong
University of California Davis, Davis, CA, USA, jdejong@ucdavis.edu
Christopher P. Krage
GEI Consultants, Inc., Rancho Cordova, CA, USA, ckrage@geiconsultants.com
Ross W. Boulanger
University of California Davis, Davis, CA, USA, rwboulanger@ucdavis.edu
Don J. DeGroot
University of Massachusetts Amherst, Amherst, MA, USA, degroot@umass.edu

ABSTRACT: The site characterization objectives for a project are typically focused on determining the types of soil present, the engineering properties of these soils, and the stratigraphic structure of the subsurface. Often, there is a greater emphasis and more systematic evaluation of the soil properties, while the subsurface stratigraphic units, including their boundaries, vertical and lateral continuity, etc., are assessed in a more ad-hoc manner with greater reliance on engineering judgement. The engineering judgement employed often assumes the subsurface to be more uniform that it actually is. This paper presents a framework using geostatistical methods to more quantitatively design and adapt site investigation programs to better estimate the subsurface geologic conditions while still evaluating soil types and engineering properties. This is exemplified first using an idealized synthetic 3D site and second through a 2D cross-section of a real project site. For the former, an idealized 3D site is used to evaluate the effect of sounding number and spatial distribution on calibration of a geostatistical model and the subsequent calibrated realization of subsurface conditions. Grid, nested, and combined grid-nested sampling patterns are used to calibrate and condition 3D transition probability geostatistical models of subsurface variability. The ability to identify spatial correlation from various number and patterns of exploration locations is presented and discussed. Analysis of a 2D cross-section at a real project site is then used to apply the observations made, delineating how assumptions in the geostatistical model and the number of CPT soundings can be used to refine a subsurface geologic model. This set of analyses lead to recommendation of a sounding spacing to correlation length ratio that is necessary to accurately quantify the spatial variability conditioned on discrete sampling locations.

Keywords: Geostatistics; Spatial Variability; Site Investigation; CPT

1. Introduction

Site characterization is a multi-step process with the primary purpose of evaluating a project site for design and construction of a geosystem (e.g., building or bridge foundations, dam and levee construction, etc.). The process consists of multiple stages, with the effort required for each dependent on project complexity. As detailed further by DeJong et al. (2016), the process typically consists of hypothesizing the performance mechanisms and integrating them with a geologic model, performing preliminary analyses to verify or eliminate possible scenarios, performing site investigation using in situ and laboratory tools, distilling and idealizing the site conditions through re-evaluation of the hypothesized geologic model against the data collected followed by sub-division into representative zones to which properties are assigned, performing analysis and design, and finally performing the observational method through construction and the project service life.

Site investigation is the stage within the site characterization process where specific subsurface information is obtained for the project site. This stage typically consists of a planning phase, where information about the proposed project extent and anticipated performance mechanisms are used to design a site investigation layout, ideally consisting of multiple investigation tools. A combination of in-situ penetration tests (e.g., cone penetration test, CPT, standard penetration test, SPT), non-destructive tests (e.g., shear wave velocity, SASW, etc.), and drilling and sampling, combined with subsequent laboratory testing, are used to evaluate subsurface conditions and obtain engineering parameters for analysis and design.

Information obtained from the site investigation is used to idealize site conditions by assessing the spatial variability in terms of measurement variability, property variability, and variability in the existence and location of individual soil types and layers (Phoon and Kulhawy 1999). Site idealization for design often requires simplifying assumptions for the subsurface properties and soil types. Due to budget constraints that tend to limit the extent of a site investigation, engineers must rely on measurements from a small subset of the project site and extrapolate observations to non-sampled locations for analysis and design. This latter step typically relies heavily on engineering judgement, in part because the executed site investigation is often more focused on obtaining engineering properties and achieving site coverage. Engineering judgement is then heavily leveraged to generate a subsurface cross-section, which often portrays layers as being more uniform and continuous than reality due to the geotechnical engineer's common propensity to overlook the complexity and discontinuity of typical depositional processes.

The purpose of this paper is to investigate and demonstrate how geostatistical concepts of spatial correlation and conditional realizations using transition probability geostatistics can provide more firm guidance in how site investigation programs can be performed to obtain information useful for evaluating subsurface variability. To this end, background on the use of geostatistical methods applied in geotechnical engineering is presented, followed by an overview of the transition probability geostatistical approach used in this study. A theoretical sampling exercise of an idealized synthetically generated 3D site is used to provide insight into the accuracy and error produced when the number and spatial distribution of the CPT soundings is varied. Subsequently, a 2D cross-section from a linear infrastructure project site along which a large number of CPT soundings was performed is studied. The effect of the number of CPT soundings as well as the incorporation of 'soft information' (e.g. geologic depositional process informed estimates) into the geostatistical modeling is investigated. Finally, a workflow for considering integrated site characterization using adaptive sampling layout planning is presented.

2. Background

Consideration of spatial variability in geotechnical engineering has historically focused on the selection of properties and parameters for engineering design (e.g., Phoon and Kulhawy 1999) or on the use of random field modeling for performing engineering analysis (Fenton and Vanmarcke 1990). These approaches for reliability based design provided early advances in the field, with a focus on capturing how properties and subsurface conditions can vary and affect system performance. However, they have not fully captured the stratigraphic complexity inherent to the depositional processes, nor on how a geotstatical model can be accurately conditioned to site data. As a result, the standard of practice largely assumes uniformity of site conditions (soil types, properties, layers, etc.) in absence of information definitively proving otherwise, which can be either a conservative or unconservative assumption depending on project-specific loading conditions and failure mechanisms.

Methods to robustly analyze spatial variability have been proposed in the geotechnical literature. Developing spatially variable random field models for evaluating soil stratigraphy and continuity began largely with the work of Vanmarcke (1977). In addition, researchers have studied the variability of properties captured during site investigation (e.g., Phoon and Kulhawy 1999, Baecher and Christian 2005) and the ability to generate random fields for numerical simulations (e.g., Fenton and Vanmarcke 1990, Griffiths et al. 2002, Jaksa et al. 2005, Boulanger and Montgomery 2016, Montgomery and Boulanger 2017). The results presented in this prior research often separate the modelling of property variability within a single stratigraphic unit from modeling the presence of distinct stratigraphic units.

Variability is often captured using spherical or exponential covariance functions that model the correlation of properties in space or variogram functions that model the dissimilarity of properties in space (e.g., Elkateb et al 2003). It is often represented in spatial models using a term to describe the extent of spatial correlation, such as range, correlation length, or scale of fluctuation (Elkateb et al. 2003).

Approaches to capture correlation extent and calibrate spatial correlation models to measured field data depend on the types and quality of data and the sampling patterns used to obtain measurements for calibration. Borings and in-situ tests provide vertical profiles that capture subsurface information at varying frequencies. CPT testing has high vertical resolution for soil profiling, with data reported every 2 to 10 cm. Sonic core samples provide continuous sampling beneficial for interpretation, although recovered samples are not suitable for obtaining soil properties due to significant disturbance during sampling. More discrete sampling and testing methods (e.g., SPT, VST, PMT, DMT, tube sampling, etc.) do not have the resolution to finely resolve depositional stratigraphy but do provide measurements of engineering behavior. Information from borings or in-situ tests can assist interpretation of laboratory tests, and vice versa. In addition, geophysical surveys can greatly assist in identifying stratigraphic features and providing information that infills between widely spaced borings and soundings (Cotterill et al. 2017). Collectively, this information can be used to obtain measurements of engineering properties for analysis and design or be used to define depositional boundaries and contacts between units.

The ability to evaluate correlation structure depends on the physical layout of the sampling plan. DeGroot and Baecher (1993) studied the ability of different sampling plans to capture spatial correlation structure. Using sampling approaches of clustered, nested, stratified (or grid), and random sampling, they found that the nested approach provided the least biased estimate of the autocovariance distance. They also found that separation distances less than the estimated spatial correlation provide the best estimates of the autocovariance distance. This is reasonable since closely spaced data is most helpful in defining the initial decay slope of a correlation function or exponential variogram. When the spatial correlation structure is known the best sampling plan is one with systematic distribution of sampling locations, as these provide even conditioning for the site (Olea 1984). However, the ability to capture spatial correlation structure with a given sampling plan also informs the design of the sampling program, such that a balance between grid and nested approaches may be warranted (DeGroot and Baecher 1993).

Generating simulations of subsurface stratigraphy using spatially correlated properties requires sampling, models for spatial correlation, and site interpolation. Models of spatial correlation can be applied to conditioned or unconditioned random fields, where conditioning a correlation model on the measured data constrains the spatial correlation and honors data measured at sampling locations (Lloret-Cabot et al. 2012). One such technique for conditioning random fields is kriging, which is a best linear unbiased estimator technique for interpolation that honors conditioning data and is the cornerstone for geostatistical methods (Gooverts 1997).

Geostatistical simulations increase in complexity with added spatial dimensions. 1-D implementation examples include modeling the spatial correlation of tip resistance and stratigraphic breaks from vertical CPTs (e.g., Fenton 1999, Cao and Wang 2013, Li, X.Y. et al. 2015, Bong and Stuedlein 2017). 2-D applications include estimates of lateral spatial correlation (e.g., Lloret-Cabot et al. 2012, 2014) and estimates of spatially varying foundation settlement (e.g., Griffiths et al. 2002, Al-Bittar and Soubra 2014). 3-D applications have investigated the settlement behavior of a spread footing foundation on spatially variable soil (Jaksa et al. 2005), estimated the depth to bedrock for deep foundations (Li, J. et al. 2015) and used neural network models to train and predict CPT tip resistance in sands (Juang et al. 2001).

In addition to information obtained via site investigation, the incorporation of 'soft information' is often necessary and can aid in development of spatial correlation models. Examples of soft information include inferences about subsurface conditions from maps, reports, trenching, etc. Consider, for example, a fluvial depositional environment (e.g. braided river, meandering river) where the number, extent, and continuity of layers is dependent on flow rate, sediment load, geometry, changes to river course, avulsion rates, subsidence rates, etc. (Nichols 2009). While details of each of these contributing factors over depositional time may remain unknown, a depositional process hypothesis developed by a geologist could inform a site model that includes reasonable range estimates for the number, extent, sequence, and connecitvity of layers across a site. This type of 'soft information' can be used to supplement 'hard information' (e.g. CPT sounding data) in geostatistical model development. Importantly, this geologic model can also help consider the presence of other geologic layers, features, etc. that the 'hard information' may not encouter or effectively reveal.

The extent of considered subsurface variability depends on the length scales relevant to the performance of the geosystem. As exemplified by Paull et al. (2020), the length scale of spatial variability is often proportional to the length scale of the failure mechanism. When the mechanism length scale is small relative to the stratigraphic variability then the system performance will likely depend on the lower values within the weaker stratigraphic units. On the other hand, when the mechanism length scale is large compared to the stratigraphic variability the average values across all the stratigraphic units engaged in the failure mechnaism may be controlling. Exploring the relationship between the depositional environment and spatial extent of the failure mechanism is necessary to understand the length scales of interest for geotechnical analysis and design as the depositional processes influences depositional length scales.

3. Transition Probability Approach to Modeling Spatial Correlation

The transition probability approach to modeling spatial correlation was developed in the field of groundwater hydrology for categorizing spatial continuity of flood plain deposits for contaminant transport (Carle 1996, Carle and Fogg 1996). This approach uses the probability of transitioning between categories to model spatial correlation structure (Carle and Fogg 1996). A Markov chain transition probability model is used to model cross-correlation relationships for geostatistical modeling, which assumes that spatial occurrences depend entirely on the nearest data.

Calibration of the transition probability model requires knowledge of the dataset sill, the background category, mean length(s), and cross-correlation relationships. The sill is the volumetric proportion of each category and is calibrated independently of model direction. The background category is typically assigned to what is expected to be the most prevalent category in the random field. For example, in floodplain depositional environments, overbank deposits may be set as the background category. Selecting a background category constrains the mathematical formulation of the Markov chain transition probability model. Defining a background category means the number of user defined matrix entries increases by the square of the number of categories (let i=number of categories; # of entries = i^2), such that the number of background entries increases (# background entries = 2i-1), and the number of matrix entries that must be calibrated increases (# of user calibrated entries = i^2-2i+1). In a two-category system, the transiogram is fully defined with one autocorrelation input (i.e., the mean length of the foreground category), while for a three-category system four inputs are required (two autocorrelation and two cross-correlation estimates), and so on. In a system consisting of three or more categories, additional insight into categorical transitions is required to determine the transition probability model. If enough data is available this information can be obtained by calibrating to the sampled data, otherwise these relationships must be inferred.

Transition rates are selected after specifying the sill and background category. The autocorrelation transition rates are equal to the negative inverse of the mean length, such that either transition rates or mean lengths can be specified for the transition probability correlation model. Herein the term mean length is used to describe the spatial correlation model, which is specified for each non-background category.

Transition probabilities are measured by sorting measured data into lag (or distance) bins, calculating the transition probability for each lag bin, and finally calibrating the transition probability models against these bins for subsequent geostatistical simulation. Lag bins can be set equal to the typical separation distance between observed data points, with the lag distance over which to calculate transition probabilities typically set to a value of half the data range in the given direction.

Using the developed spatial correlation model, the subsurface is then idealized using the kriging technique described above. This technique minimizes error between the observed data and the objective function (i.e., the calibrated model) which is accomplished by determining the kriging weights necessary to minimize function error. As a result, kriging produces a singular realization that has the minimum error between observed data and interpreted data (Gooverts 1997).

The specific transition probability approach applied herein was implemented in TPROGS (Carle 1999) using a categorical approach to geostatistical modeling, where categories can be assigned using CPT correlations (e.g., soil behavior type index, I_c, normalized penetration resistance, q_{c1N}, etc.), classifications from soil samples (e.g. sand, clay, SP, CH), or in terms of performance mechanisms (e.g., expected liquefaction, permeability, etc.). Subjective development of category mean lengths is another advantage of this approach, where direct measurement of category extents from geologic samples or inferences based on the depositional environment can be used as 'soft information' to augment observations from site investigation or the acquisition of sparse data. Ordering of facies (e.g., stratigraphic units) relationships used to describe depositional relationships is also captured using this categorical approach (Carle and Fogg 1996) and can be useful in cases with sequential depositional structure (e.g., fining upward sequence).

3.1. Data Gathering and Reduction

Data gathering and reduction occurs by first gathering site data (e.g., site investigation) followed by identifying of zones of interest (e.g., geologic units) to reduce data and establish a stationary dataset. Random field properties of the zoned and categorized data (e.g., mean, variance) are evaluated to assess ergodicity and stationarity. Verifying stationarity assumptions of constant mean and variance in space are necessary to rigorously assess the applicability of spatial correlation models for estimating the properties of the given random field (e.g., identifying geologic boundaries that may be present at the project site).

Cumulative distributions can be useful in initial assessment of soil variability. The mean and standard deviation of each individual profile are easily distinguished from these distributions by considering the 50th percentile and the shape of the curve. Compilation of cumulative distributions from all CPT profiles from a project site or zone are useful in identifying the distribution of properties across the site and can be useful for selecting representative properties for analysis (e.g., DeJong et al. 2016, Boulanger and Montgomery 2016, Montgomery and Boulanger 2017).

3.2. Development of a Transition Probability Model

Transition probabilities are measured by aggregating categorical observations to lag bins for each spatial, orthogonal direction. Calibration of the spatial correlation model is performed in two stages: (1) determining the sill of the random field and (2) determining the transition rates for each component of the transition probability matrix.

The sill is determined from the global proportion of a given category or calibrated from observations in a single direction. Despite how the sill is selected, the sill is a property of the dataset and maintains the same value for each direction. Herein the sill is selected as the average number occurrences for each category and is most easily determined with category observations in the elevation (vertical) direction.

For two and three category systems, the transition rates are (1) qualitatively calibrated using visual fitting of observed transition probabilities vs. lag bin data, (2) quantitatively calibrated using a best fit least squares regression of observed transition probabilities, or (3) specified using inferred mean lengths (i.e., soft information). Sensitivity analysis with a range of transition probability calibrations are useful for evaluating the robustness for certain project sites.

3.3. Geostatistical Simulation

Geostatistical simulation using transition probability models in TPROGS consists of an objective function (e.g. the transition probability model), an algorithm to minimize the error of the objective function, and a quenching-based algorithm to further reduce error in the objective function. In TPROGS, the sequential indicator simulation (SIS) and simulated quenching algorithms are used in tandem to reduce simulation error.

Sequential indicator simulation (SIS) incorporated in TPROGS uses an indicator kriging approach to develop geostatistical simulations of spatial variability (Carle 1999). In its implementation each grid node is treated incrementally to determine the category value. Once defined, the node is set and serves as an additional conditioning location for undefined nodes. In this method, the first node is specified, and the simulation results in unique randomly generated realizations. This allows for Monte Carlo analysis of the subsurface using unique subsurface realizations that have the same spatial properties.

Simulated annealing (or quenching) iteratively improves the simulation by eliminating 'impurities' that disrupt the objective function. This procedure cycles through each grid node and evaluates whether a change in the node category will reduce the objective function. The process continues until the objective function is minimized below a specified threshold or the maximum number of iterations are reached.

Successful implementation of SIS and simulated annealing results in a statistically robust simulation that honors conditioning data and the geostatistical model. Uncertainty in simulation realizations can be quantified using both internal and external validation methods. TPROGS contains an internal check in the quenching algorithm based on a user set conversion tolerance.

4. Case I: Analysis of Simulated 3D Site

A simulated site investigation was performed with an idealized, synthetically generated 3D subsurface using the transition probability technique described above. This example provides insight into the ability of transition probability geostatistics to capture spatial correlation structure and produce realistic simulations using a range of sampling patterns and model calibrations. The idealized synthetic subsurface was developed as an analog for an alluvial depositional environment, defined such that categories at each X, Y, and Z location were fully known. The idealized subsurface was then 'investigated' with CPTs to obtain profiles of the subsurface conditions, which were then used to develop spatial correlation models and produce conditional simulations using the transition probability approach. Three site investigation patterns were considered; equally spaced grid sampling, nested sampling, and a combined nested-grid sampling approach.

Using a fully known idealized subsurface allows for determination of what constitutes a sufficiently rigorous site investigation, which is not practically possible in natural alluvial soil deposits. Since the synthetic subsurface is fully defined, uncertainty is restricted to only model and simulation uncertainty. This section is organized by first presenting the idealized site, then presenting the simulated sampling methodology, followed by the results and discussion of the simulated investigations.

4.1. Idealized Site

An idealized site was generated to represent a simplified alluvial depositional environment comprised of channel and overbank soil layers or more simply, 'sand-like' and 'clay-like' soils. This simulation represents a defined 'reality' that was used to quantify uncertainty derived from simulations conditioned on a sampled subset of the 'reality'. The site was developed using TPROGS and modeled with specified correlation lengths in each direction without conditioning constraint. The volumetric proportion of channel deposits is defined as 30% (e.g., the sill) and mean lengths of 90, 30, and 3 m were specified for the downstream (X), cross-channel (Y), and elevation (Z) directions, respectively. The extent of the simulated site dimensions are 540, 180, and 18 m in the X, Y, and Z directions, respectively, representing a domain that is six times the respective mean length in each direction. The idealized site shown in Figure 1 is the realization of the simulation calibrated to these parameters and dimensions that was defined as 'reality'.

4.2. Sampling Methodology

Simulated CPT investigation programs of the idealized subsurface were designed and performed in a manner consistent with typical site investigation programs. Idealized sampling was performed using between 9 to 2,025 synthetic CPTs that penetrated the full depth of the idealized site and reported the measured categories at each intersected simulation node. Three investigation patterns were considered for synthetic site investigation: even grid sampling, nested sampling, and a combined nested-grid approach (e.g. Figure 2). For each simulated CPT the extracted categories were considered as measured site information. An example sampling summary is shown for a 36 CPT grid pattern in Figure 3. Note that the large number of simulated CPTs used in some of the investigation patterns is not a practical suggestion but is instead used to capture and demonstrate the modeled spatial structure of a densely sampled subsurface.

Grid sampling locations were evenly subdivided in each direction based on the number of simulated CPTs selected for each direction (e.g., Figure 2a, Figure 2c). The minimum spacing between soundings was equal to approximately 1/8th of the mean length for the given direction and the maximum spacing between soundings was twice the mean length for the given direction. The spacing for each direction was determined by:

$$\theta_i = \frac{\theta_{max}}{N_\theta}(i - 0.5)\ for\ i = 1{:}\,N_\theta \quad (1)$$

where θ represents each orthogonal direction (either X or Y), θ_{max} represents the maximum dimension of the simulation domain of the given direction, and N_θ represents the number of investigation locations in the given direction. The spacing locations were rounded to the nearest

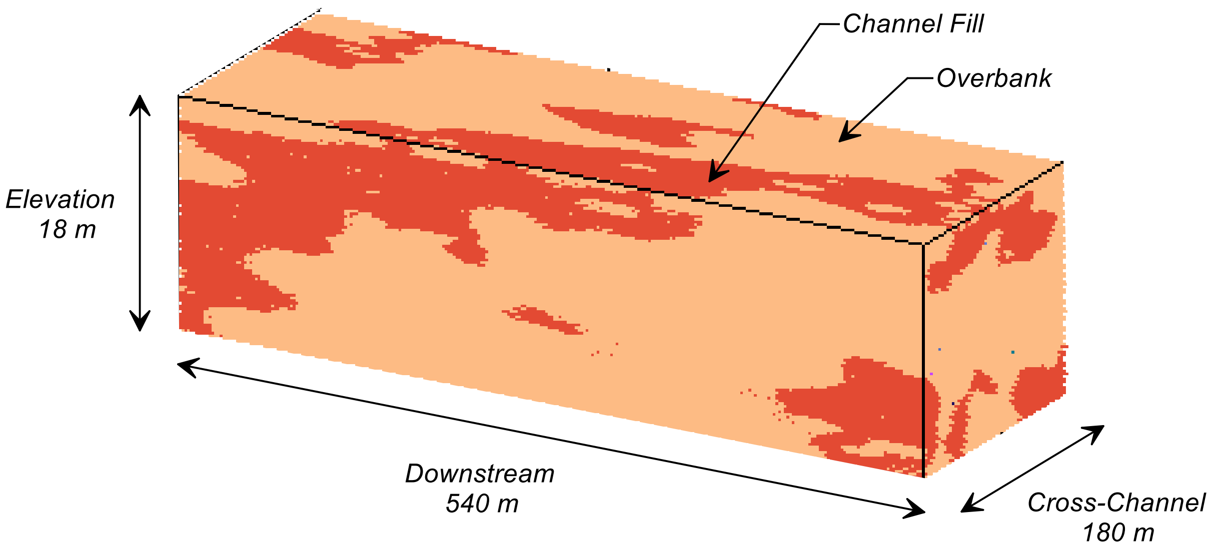

Figure 1. Idealized simulation used as an analog for a simplified alluvial depositional environment. Correlation lengths of 90, 30, and 3 m in the downstream, cross-channel, and elevation directions, respectively.

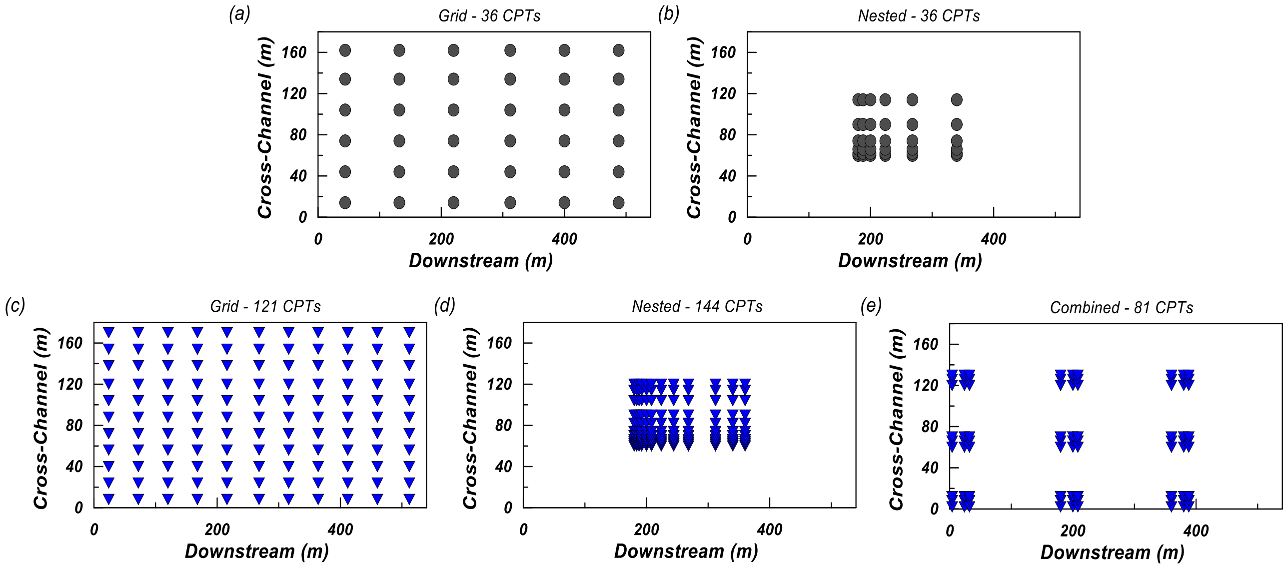

Figure 2. Some of the sampling patterns used in the idealized sampling example: (a) grid sampling with 36 CPTs, (b) nested sampling with 36 CPTs, (c) grid sampling with 121 CPTs, (d) nested sampling with 144 CPTs, and (e) combined nested grid with 81 CPTs.. Note that the combined nested grid pattern in (e) is a 9 CPT nested pattern replicated 9 times. Other combined nested grid patterns follow this same approach, though the number of CPTs used in the original nest varies.

grid node integer (e.g., divisible by a value of 2 in the Y direction and 4 in the X direction). Grid sampling used a total of 9, 36, 121, 484, and 2025 CPTs.

Nested sampling locations were selected to provide a variety of sampling spacings, from as small as 1/16th to 2 times the mean length for the given direction. Nested sampling was performed with the same number of CPTs in each direction with up to 12 soundings located along a given sampling transect. The location of the nested pattern occupies the middle third of the project site in both horizontal directions, resulting in a grouping of CPTs in the middle of the project site (e.g., Figure 2b). Nested patterns use a total of 9, 36, and 144 CPTs.

The combined nested-grid approach takes each nested pattern and repeats it across the remainder of the site, resulting in nine combined nested sampling subsets (e.g., Figure 2e). The combined approach resulted in a total of 81, 324, and 1296 CPTs.

Transition probability models were determined by either calibrating the spatial correlation model to simulated sampled observations (CPT sounding data) using a best fit approach (as shown in Figure 4) or by using geologic inferences (i.e., 'soft information') of spatial correlation lengths for each orthogonal direction. For this example, the geologically "inferred" spatial correlation lengths were set equal to the specified mean length of the idealized subsurface in each direction (i.e., "perfect" soft information). The sill was determined from the category proportions of the synthetically sampled observations. Uncertainty in transition probability model calibration was quantified in terms of true calibration error, which is defined as the percent difference between the calibrated model and the actual mean lengths used to simulate the idealized site.

The calibrated model from the given spatial sampling pattern was used to generate simulations conditioned on the sampled data to estimate the site's spatial structure.

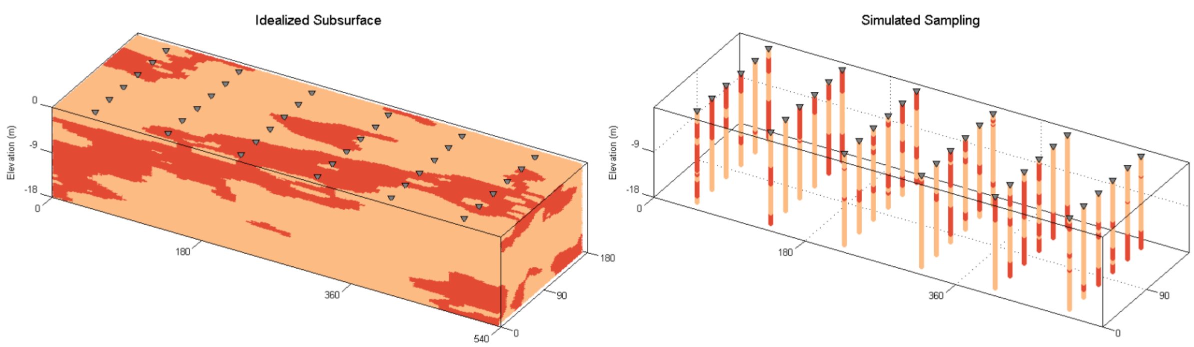

Figure 3. Example shown for 36 CPT site investigation in a grid pattern where (a) shows the sampling locations of the idealized reality superimposed on the idealized subsurface and (b) summarizes the resulting simulated sampling investigation to be used in assessing spatial structure and simulation error.

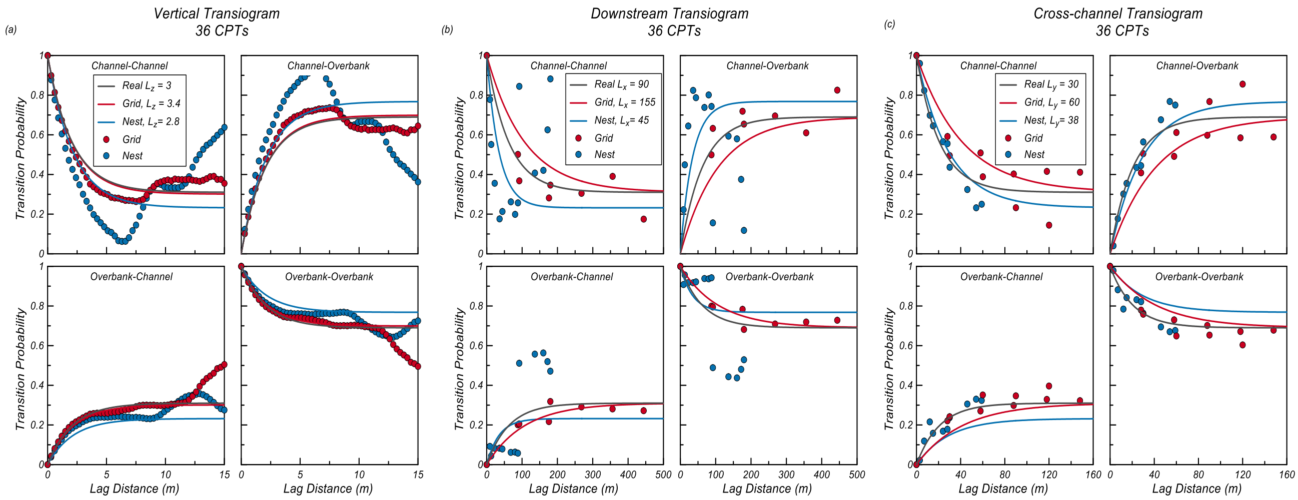

Figure 4. Summary of transition probability measurements and calibrations for 36 CPTs where (a) is the vertical transiogram, (b) is the downstream transiogram, and (c) is the cross-channel transiogram. Note both grid and nested patterns are represented herein.

Results from one realization of each simulation performed in TPROGS using the same grid dimensions as the idealized site were used to determine simulation error. Simulation error was defined as the difference between simulation realizations and the idealized site and was determined by identifying the locations of channel deposits in the idealized reality and comparing against simulated categories for the corresponding grid locations from the simulated realization. The error is calculated as:

$$Sim.Error\ (\%) = \frac{1}{N}\sum_X \sum_Y \sum_Z \left(Val_{X,Y,Z} - Val_{X,Y,Z,Ideal}\right) \in Val_{X,Y,Z\ Ideal} = Channel \qquad (2)$$

where N is the total number of simulation grid nodes and X, Y, and Z are the grid nodes belonging to each direction. For idealized simulation grid nodes that are channel deposits (e.g., $Val_{X,Y,Z,Ideal} = 1$), $Val_{X,Y,Z}$ is the value of the sampling simulation for the corresponding grid nodes. Grid nodes that indicate the same value (i.e., a value of 1) do not contribute to the simulation error. Increased similarity between the idealized reality and the given simulation results in reductions in the measured simulation error. Simulation error is affected by errors in the estimated sill, correlation length for each direction, and limited conditioning locations.

4.3. Results and Discussion

Three sampling patterns consisting of between 9 and 2025 simulated CPTs were used to characterize an idealized site for the purpose of evaluating the sampling pattern's ability to capture spatial structure and model subsurface stratigraphy. Results are presented in Figure 5 that compare isometric views of the idealized geometry

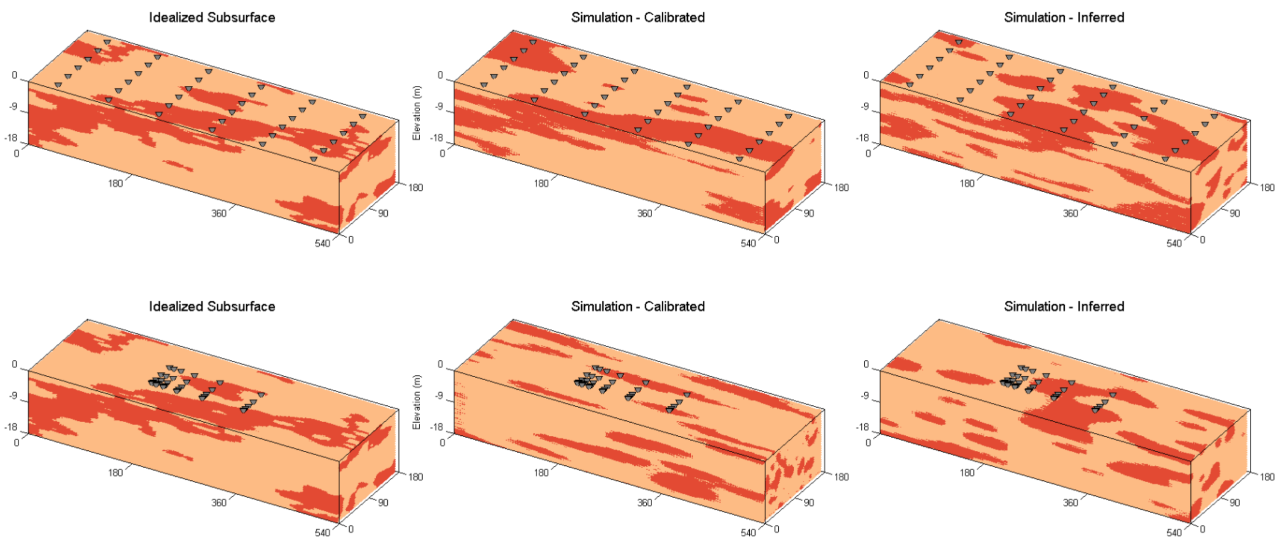

Figure 5. Simulation realizations for different sampling patterns using 36 CPTs. The left column shows the base subsurface stratigraphy with the proposed sampling pattern displayed. The middle column shows the conditional simulation when mean lengths are calibrated to observed data. The right column shows the conditional simulation when mean lengths are inferred from site knowledge. The top row is a grid sampling approach while the bottom row is a nested sampling pattern, both with 36 total CPTs.

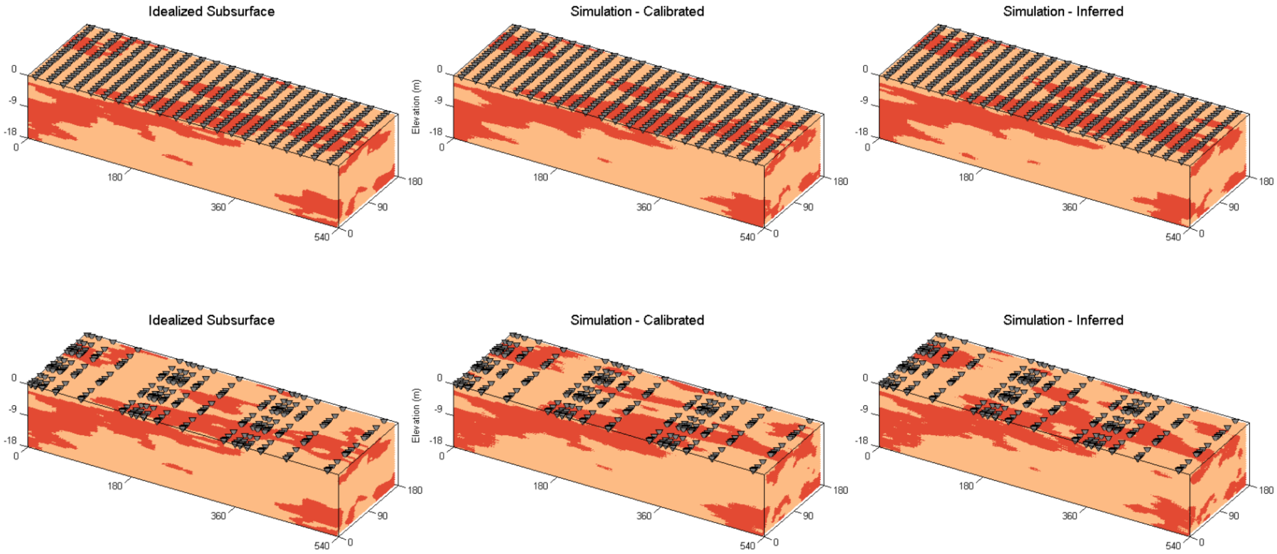

Figure 6. Simulation realizations for different sampling patterns using a large number of CPTs. The left column shows the base subsurface stratigraphy with the proposed sampling pattern displayed. The middle column shows the conditional simulation when mean lengths are calibrated to observed data. The right column shows the conditional simulation when mean lengths are inferred from site knowledge. The top row is a grid sampling approach with 484 CPTs while the bottom row is a combined nested grid sampling pattern with 324 CPTs.

and resulting subsurface realizations interpreted from simulations based on grid and nested simulations with 36 CPTs. The subsurface realizations were based on either calibrated or inferred correlation lengths, and are conditioned on the CPT locations for each sampling pattern. Figure 6 presents similar results as Figure 5 but with a larger number of CPTs (>324).

The sill estimated from each simulation pattern and number of CPTs is shown in Figure 7. Results indicate that the sampled sill approaches the global defined sill as the number of CPTs increases. The grid pattern provides reasonable estimates of the sill for the range of CPTs used, while the nested pattern is less effective site the site is not fully covered.

The calibrated transition probability mean lengths of channel fill deposits in the vertical, cross-channel, and downstream directions are shown in Figure 8. It is noted that the inferred data shown must, by definition, produce the correct value since it was defined as such, and is therefore included for context. For simulations in which the mean length is calibrated to measured transition probabilities, the nested approach typically underestimates mean lengths while the grid approach overestimates mean lengths when fewer CPTs are used. When investigations included more CPTs the various sampling approaches converge to the defined mean length in all three orthogonal directions.

The error in estimating the transition probability mean lengths of channel fill deposits in the vertical, cross-channel, and downstream directions is shown in Figure 9. For simulations in which the mean length is calibrated to measured transition probabilities, the nested approach results in smaller error when fewer CPTs were used. When more CPTs were used the three sampling approaches result in smaller mean length calibration error. The spacing to mean length ratio (S/L) reported in Figure 9 indicates

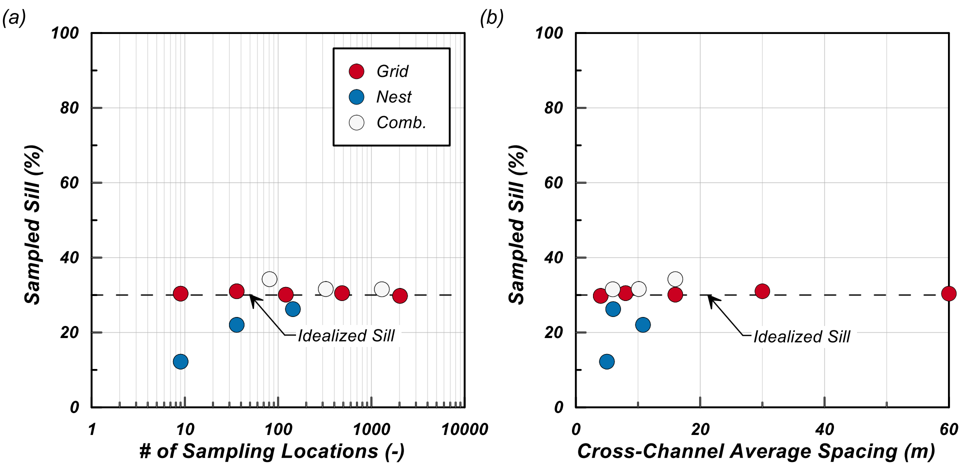

Figure 7. Estimates of global sill obtained from site investigation patterns.

decreasing calibration error with decreasing S/L for grid sampling patterns in the horizontal directions (e.g., Figures 9d, 9f).

Simulation error decreases with an increasing number of CPTs for all sampling patterns and calibration methods. Measured simulation error is shown in Figure 10 for all sampling patterns with both calibrated and inferred mean lengths. For grid sampling patterns with a small number of CPTs the inferred simulations result in larger simulation error than the calibrated approach. This is due to the observations evident in Figures 8 and 9 where the calibrated mean lengths are much greater than actual mean lengths for cases with fewer CPTs. The increase in calibrated mean length implies continuity, therefore removing more isolated pockets of channel deposits. The simulation error is largest for nested patterns with calibrated mean lengths, even with inclusion of additional CPTs. Differences between simulation error for calibrated and inferred mean lengths tend to decrease with increasing number of CPTs. Simulation error as a function of S/L plotted in Figure 10 shows a decrease in simulation error with decreasing S/L.

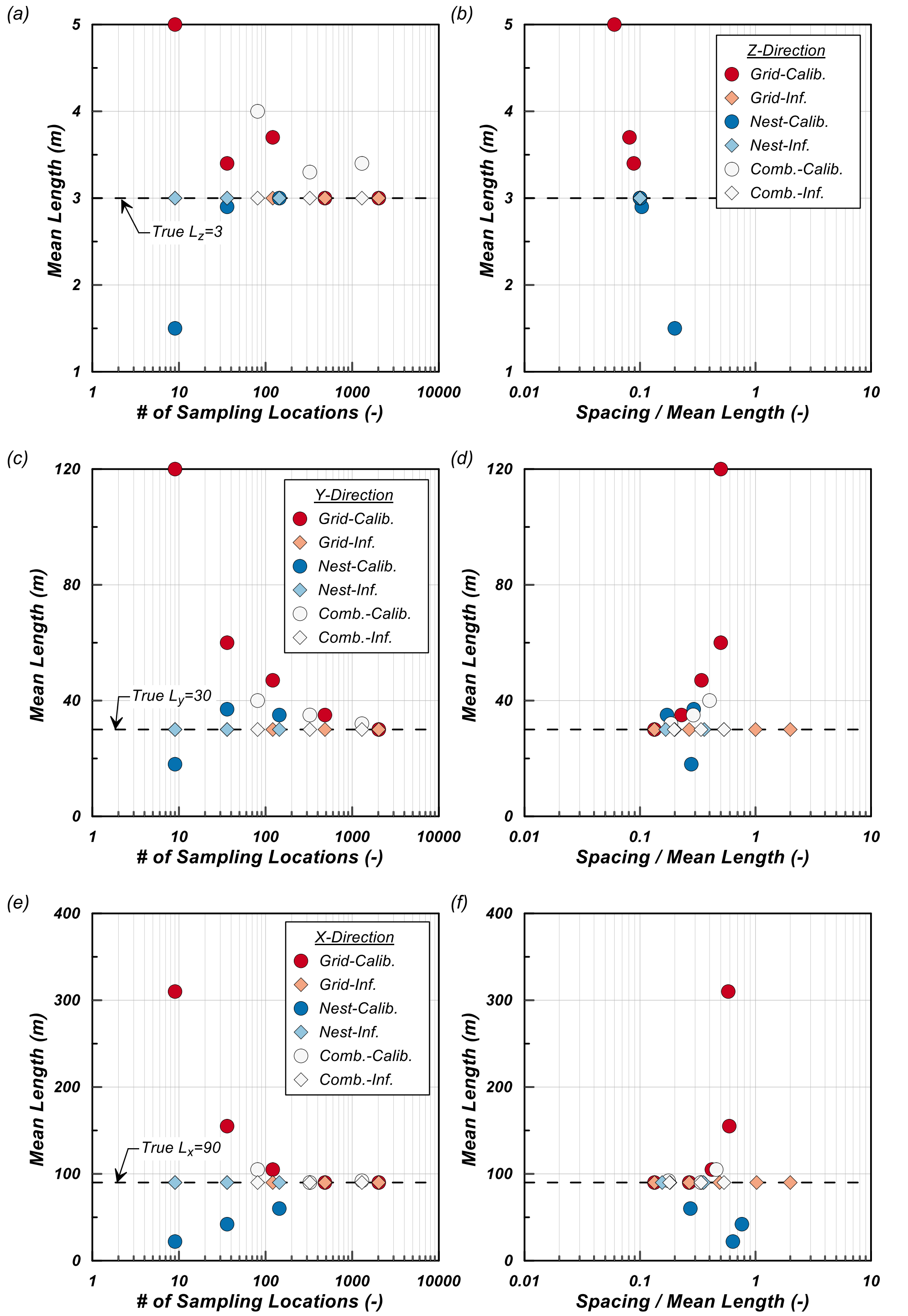

Figure 8. Calibrated transitiion probability mean lengths for each direction as a function of the number of sampling locations and the spacing to mean length value, S/L for the three sampling patterns investigated: grid, nested, and combined nested grid.

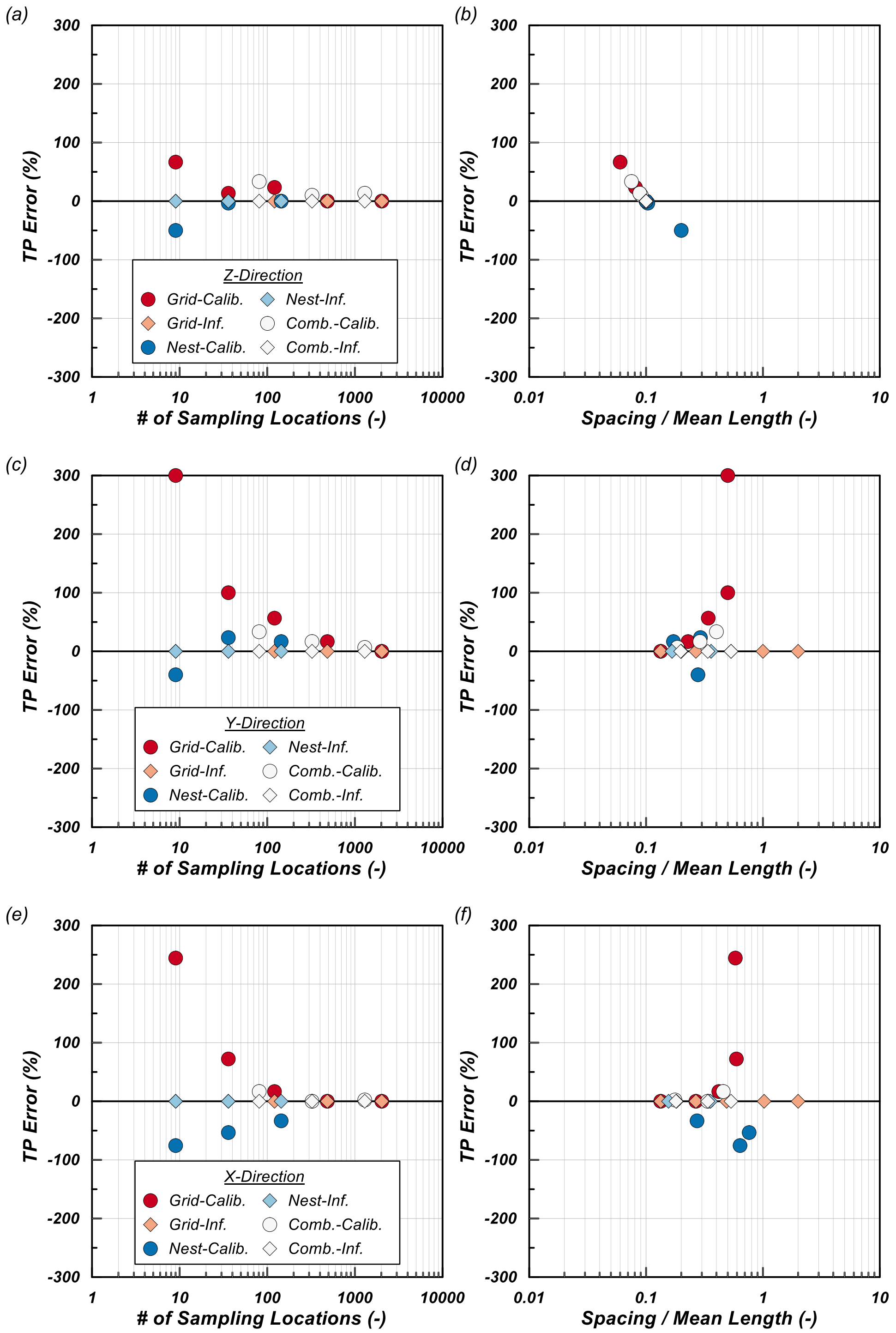

Figure 9. Error in estimating transitiion probability mean lengths for each direction as a function of the number of sampling locations and the spacing to mean length value, S/L.

These results indicate that the global sill is best estimated using the grid approach, where greater site coverage occurs with a small number of CPTs. The nested approach has a smaller spacing to mean length ratio and is less distributed for a similar number of CPTs, resulting in localized error for sill estimation. The combined nested grid produces stable sill estimates, in part due to the balance between approaches and the relatively large number of CPTs used in the smallest combined pattern (e.g., 81 CPTs).

The ability to accurately capture mean length depends more on S/L than on the total number of CPTs used, as evident in the comparison between grid and nested patterns with 9 or 36 CPTs in Figure 9. The transition probability framework uses an exponential autocorrelation model, where the mean length represents the initial decay of the exponential model. If the value of S/L is large, the initial decay of the exponential function and the mean length is difficult to determine and typically overestimated with calibration errors greater than 100% for certain sampling patterns and S/L values greater than 0.5.

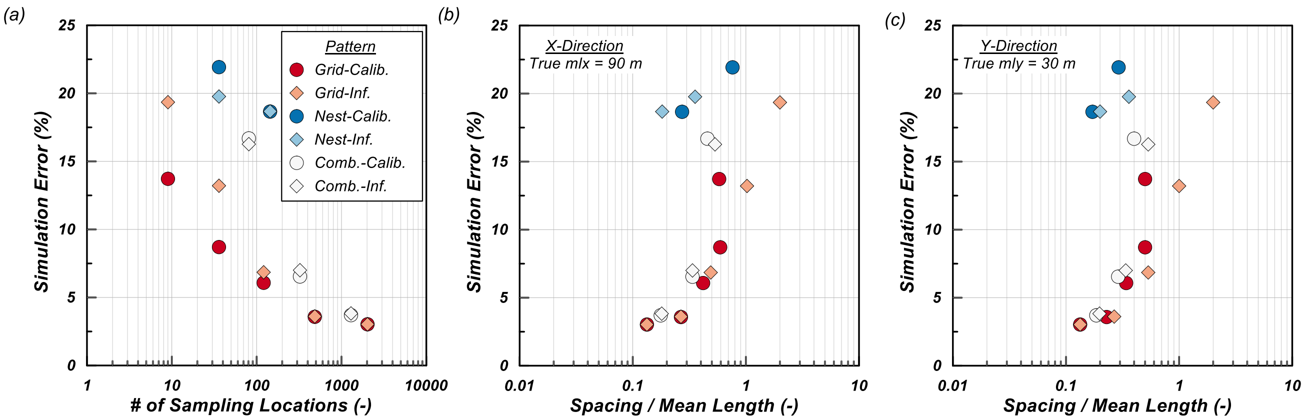

Figure 10. Simulation error for the idealized subsurface as a function of number of sampling locations and the spacing to mean length ratio, S/L. Note that the simulation error is the aggregated errors in the X, Y, and Z directions.

Accurate estimates of mean length are observed for decreasing value of S/L in both X and Y directions for each sampling pattern. Since the CPT spacing in Z is fixed by the vertical sampling frequency, the error and value of S/L is small for both sampling patterns explored and depends mostly on the aggregated length of vertical penetration for all considered CPTs.

For simulations where mean lengths were calibrated to sampled data and few CPTs are used, the nested sampling patterns typically underestimated mean lengths while grid patterns typically overestimated mean lengths. As the number of grid pattern CPTs increases, the spacing decreases, resulting in more lag bins, smaller mean lengths, and reduced S/L values. Because the spacing between CPTs is variable in the nested pattern, additional nested CPTs do not add additional lag bins but instead add more data points to the considered lag bins and refining the measured transition probabilities, thereby improving the mean length estimate. As the number of CPTs increases, the simulation error of the calibrated sampling patterns converge.

Visual comparison between the idealized subsurface and resulting simulations developed using simulated CPT site investigation indicate that, for a small number of CPTs, simulations are highly variable and are not well constrained. As expected, visual evaluation of the more thoroughly sampled site in Figure 6 indicates greater consistency between the idealized subsurface and the resulting simulation.

The simulation error for the sampling patterns, calibration types, and number of sounding locations used as a function of either number of CPTs or the value of S/L for each calibration direction shown in Figure 10 reveal interdependencies. Differences in simulation error for calibrated versus inferred grid sampling patterns range between 3 to 20% and depend primarily on the number of CPTs used in the simulations. Overprediction of correlation length (e.g., assuming greater continuity) may result in global reductions in error, however localized error may increase. For large length scales of interest, the result may be acceptable, however for smaller length scales of interest, the lack of adequate conditioning on the smaller length scale may result in conservative or unconservative predictions of performance.

Nested sampling patterns result in the largest simulation error for calibrated conditional simulations which is attributed to the lack of conditioning data across the site (Figure 10). While mean length estimates slightly underpredict the actual values, the nested sampling pattern only covers at most 11% of the project site, resulting in a poor distribution of conditioning data. Performing nested explorations with increased density of conditioning locations within a given location further constrains the simulations near conditioning locations, but the regions without conditioning locations remain unconstrained with greater simulation error in these areas.

For grid and combined nested-grid approaches, widely spaced sampling locations do not constrain simulations when sampling spacing is large relative to mean length estimates (e.g., larger value of S/L). In the case of a 36 CPT grid pattern, the estimated cross-channel mean length is 60 m versus the actual value of 30 m. Given the cross-channel spacing between CPTs of 30 m, the S/L is 0.5 and 1.0 for the estimated calibration and actual value spatial correlation, respectively. The S/L artificially defines a smaller value, due to the overestimation of mean length in the calibration. Overestimation of mean length results in more simulation continuity and the artificially smaller S/L value increases the influence of conditioning locations during simulation.

Improvements to spatial correlation estimates have little effect in reducing simulation error in regions where the simulation is unconstrained (i.e., large values of S/L). The uncertainty in site simulations with larger values of S/L requires a combination of more accurate spatial correlation models and closely spaced conditioning locations to reduce simulation uncertainty. Site investigation layouts that provide this balance will improve estimations of subsurface stratigraphy.

For this idealized simulation, nested patterns appear to better capture spatial correlation structure with fewer investigation locations due to the increased variety of lag spacing, while grid sampling patterns are better at anchoring the conditional simulations due to the more distributed CPT pattern. The use of the combined nested-grid approach is an attempt to balance correlation calibration with added conditioning constraint. Grid patterns result in lower simulation error when the spatial structure is known or when sufficient site investigation is performed such that S/L is small. Ideally, additional sampling locations would increase the presence and distribution of conditioning data, which would reduce simulation uncertainty and limit the impact of a poorly defined spatial correlation structure. The ratio of sampling separation distance to the calibrated mean length (i.e., S/L) is an indicator of site spatial variability in a given direction. An inferred spatial correlation structure from knowledge of the geologic depositional environment reduces simulation uncertainty by improving assessment of the spatial correlation structure. These simulations show the need to balance site coverage (i.e., number and location of conditioning data) and assessment of spatial correlation structure (i.e., the measured or inferred correlation lengths). Simulations with a balanced approach result in reduced simulation uncertainty.

5. Case II: Analysis of Project Site 2D Cross-Section

An industry project case history where 31 CPT soundings were performed along the axis of a proposed earthen embankment dam provided the opportunity to use the conclusions drawn from the above idealized synthetically generated scenario and to apply it to an industry project. The site, located east of the Sierra Nevada mountain range in California, consists of an incised fluvial channel bounded by late Pleistocene and early Holocene fan deposits to the west, and early to mid-Holocene fan deposits to the east. A substantial site exploration program consisting of geotechnical borings, sonic borings, CPTs, SPTs, iBPT soundings, and trenching investigations was performed to investigate the subsurface stratigraphy and evaluate liquefaction susceptibility of these soils.

Geostatistical simulations of subsurface variability were performed using the methods presented above. The selection of the decision variable, calibrations of the spatial correlation model, and resulting conditional simulations are presented. Sensitivity of the correlation model calibration and conditional simulation to the number and location of CPT soundings used is also investigated and discussed. Additionally, sensitivity to correlation length calibrations is considered using correlation lengths equal to 0.5 to 2.0 times the calibrated value.

5.1. Site Investigation

The project owner commissioned several exploratory studies beginning in 2002 to compile a large dataset of subsurface explorations consisting of borings, in-situ tests, and geologic studies. Many CPTs were performed at the project site over multiple investigation campaigns, with a total of 31 unique CPTs performed along the proposed dam alignment, resulting in an average horizontal CPT sampling spacing of 16.1 m. Along the proposed dam axis, the closest CPTs were spaced 7.9 m apart and the furthest spaced CPTs were 21.9 m.

Geologic units of lower channel infill and upper channel infill deposits were delineated in the site investigation. The incised channel deposits are bounded by Pleistocene alluvial fan deposits to the east and west, with Holocene channel infill deposits overlying Pleistocene alluvial channel infill deposits. Characterization of the upper alluvium, denoted as Zone 1 alluvium, was examined in this study.

5.2. Geostatistical Simulations

Geostatistical simulations were developed for the Zone 1 alluvium at the project site using the 31 CPT soundings obtained along the proposed dam alignment.

5.2.1. Decision Variable Specification

The overburden corrected, normalized penetration resistance (q_{c1N}) was selected as the decision variable for the geostatistical simulations. q_{c1N} is a unit independent, stress normalized measurement that removes non-stationary trends from the data, such that variations in measured q_{c1N} primarily reflect the variation in soil type, density, and strength. The use of the soil behavior type index, I_c, to delineate CPT measurements into bins of 'sand-like' and 'clay-like' behavior is not an effective decision variable in this analysis since the Zone 1 deposits are 98% "sand-like" using the I_c threshold of 2.6. Additionally, given adequate seismic hazard, this suggests the soil is capable of earthquake-induced excess pore pressure generation and cyclic liquefaction.

The threshold for q_{c1N} was selected based on the potential for liquefaction. A design earthquake with a moment magnitude of 7.75 and 84th percentile peak ground acceleration (PGA) of 0.85g results in cyclic stress ratios (CSR) of approximately 0.4 to 0.6 in the Zone 1 alluvium. A simplified two-category threshold q_{c1N} of 170 was selected to delineate liquefiable and non-liquefiable deposits, where the liquefiable materials exceed the 15% probability of liquefaction threshold indicated by the Boulanger and Idriss (2015) empirical CPT triggering curves.

The CPT traces are summarized in Figure 11 in terms of elevation, cumulative percentile, mean and standard deviation. Approximately 50% of the CPT traces in Zone 1 indicate expected liquefaction (e.g., $q_{c1N} < 170$), however the patterns of these CPT traces is less easily discernable (Figure 11c).

5.2.2. Model Calibration

Transition probabilities were calculated in TPROGS and fit with Markov chain transition probability models. A two-category simulation with a background category of non-liquefiable (NL) (e.g., $q_{c1N} > 170$) and a foreground cateory of liquefiable (L) (e.g., $q_{c1N} < 170$) was selected for the series of simulations, with model calibration sill and mean lengths determined based on the data obtained from all 31 CPTs. The calibration shown in Figure 12 results in an expected liquefaction categorical sill of 49% and mean lengths of 1.05 and 61 m were determined via regression for the vertical and horizontal directions, respectively. This resulted in a horizontal to vertical correlation length anisotropy (L_x/L_v) of 58 and a horizontal spacing to mean length ratio (S_x/L_x) of 0.27.

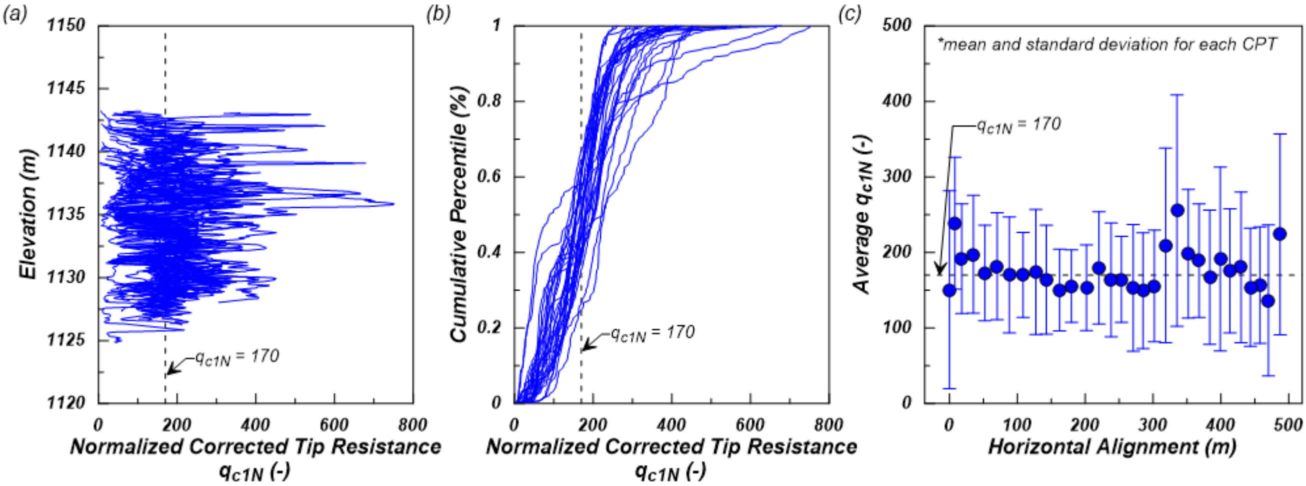

Figure 11. Summary of normalized corrected tip resistance for the 31 CPTs performed along the proposed dam allignment. A q_{c1N} value of 170 is chosen as the categorical cutoff between expected liquefaction and non-liquefaction based on triggering correlations from Boulanger and Idriss (2015).

5.3. Results and Discussion

Simulation realizations were performed as described above using available conditioning data for simulation. A single conditional realization for the 31 CPT simulation for the project site is shown in Figure 13. The realization indicates continuity of liquefiable material at the upper east of the site and an approximately 1.5 m thick continuous liquefiable lens extending 305 m across the western to central region of the site at elevation 1135 m. This particular layer may be a sheet flow deposit from the slope west of the site. The remainder of the deposition is alluvial channel deposits.

Multiple realizations of the same calibrated model and conditional simulation provide insight into the sensitivity of the simulations to considered modeling parameters. Figure 14 shows a heat map of the average grid node category for 10 realizations generated from the same simulation parameters. Bold colors (i.e., black or cream) indicate that the category across the realizations was consistently produced in the different simulations. The transitional gray colors indicate the relative frequency that either category 1 or category 2 values were produced at the given grid node, with a lighter gray indicating that category 1 was produced more frequently and a darker gray indicating that category 2 was produced more frequently. These gray regions indicate regions where the calibrated model is unconstrained such that the category value depends primarily on the sill (i.e., 49% in this case).

Simulation sensitivity to calibrated horizontal mean length for the 31 CPT cross-section is evaluated using

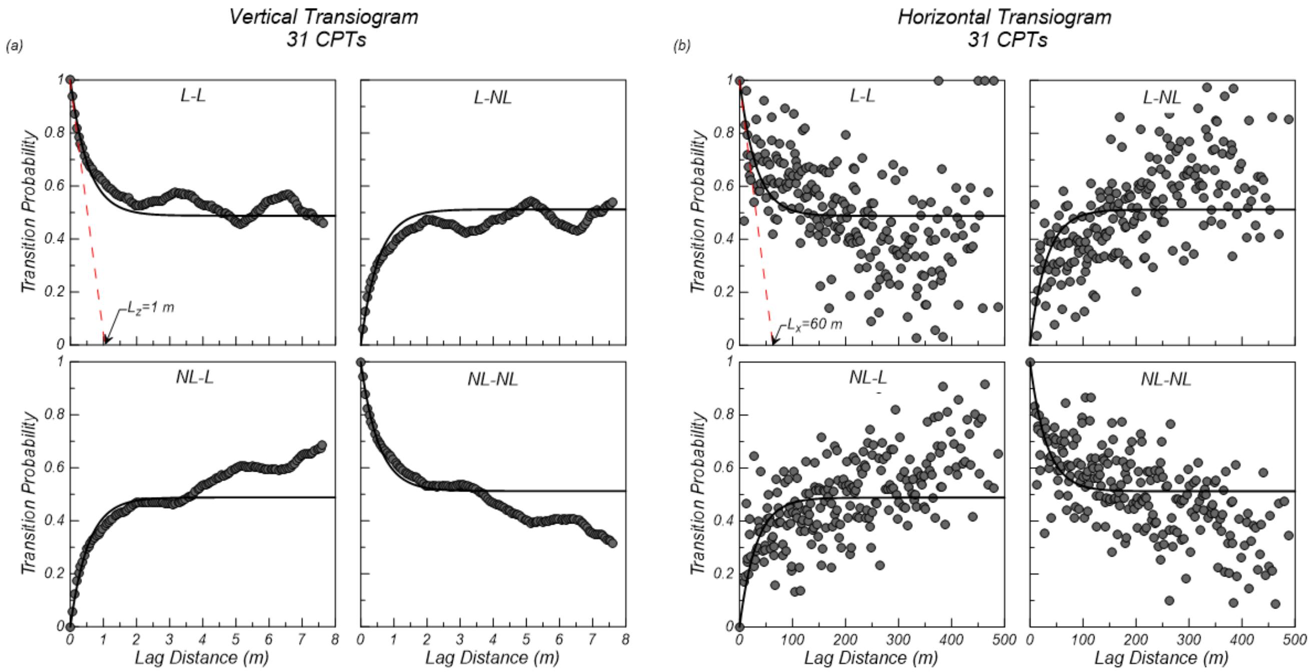

Figure 12. Transition probability calibrations using 31 CPTs for (a) vertical and (b) horizontal (cross-channel) measurements. The dashed red line represnets the mean length calibration for the category of interest.

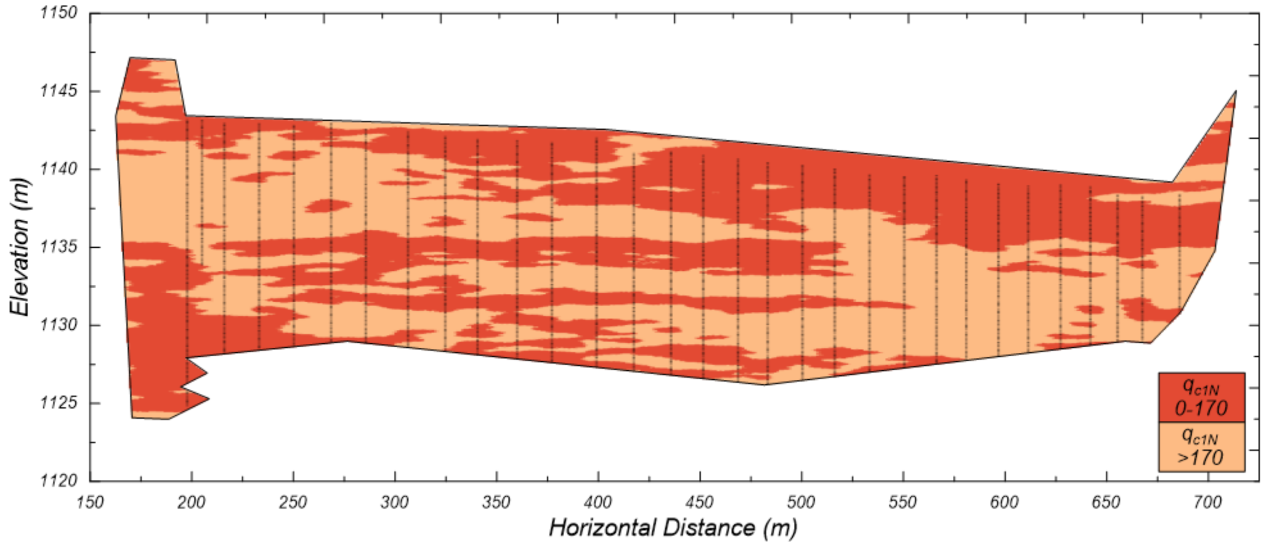

Figure 13. A single simulation realization for the 31 CPT cross-section. Dashed black lines represent available CPT data.

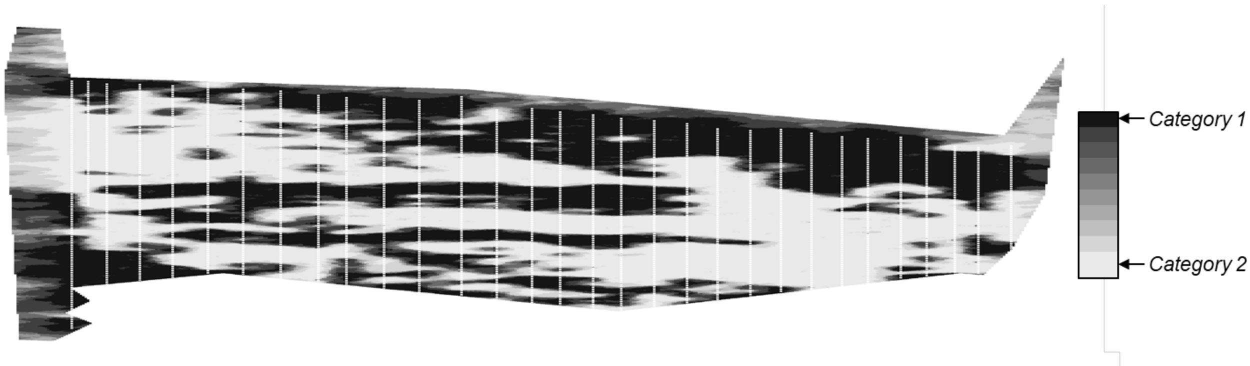

Figure 14. The average of 10 simulation realizations described using a heat map. The colorscale represents the individual categories where black corresponds to category 1 (expected liquefaction) and cream corresponds to category 2 (expected nonliquefaction). Varying shades of gray indicate preference towards either category. Dashed white lines represent available CPT data.

horizontal mean lengths of 30.5, 61.0 and 122 m, representing values of 0.5, 1.0 and 2.0 times the determined horizontal correlation length. Figure 15 shows the resulting average of 10 conditional simulations for these variations (while the vertical mean length was held constant at 1.05 m). The average simulation error does not vary significantly between the simulations, with the values being 18.5, 12.8, and 14.8% for simulations with horizontal mean lengths of 30.5, 61.0 and 122 m, respectively.

Calibration sensitivity is influenced by the number of CPTs performed as varying the number and selection of considered CPTs affects the sampled sill, calibrated mean lengths, quantity and location of conditioning data, and separation distance between conditioning locations. The number of CPTs used in each sensitivity simulation were 3, 5, 9, 16, and 31. In addition, the CPTs were selected to maintain a similar spacing between sounding locations that were distributed across the site.

The additional calibrations for increasing the number of CPTs used in site investigation is shown in Figure 16, which indicates the category 1 measured transition probabilities in the horizontal cross-channel direction and the calibrated correlation lengths calibrated to each limited data set. Widely distributed CPTs produce large lag distances and correspondingly large horizontal mean lengths. With increased number of CPTs, the separation distance between CPTs decreases and the resulting calibrated mean lengths correspondingly decrease.

The average of 10 conditional simulations using the sill and correlation lengths individually calibrated to 3, 5, 9, 16, and 31 CPTs are shown in Figure 17. As expected, each increase in CPTs improves definition of the estimated subsurface structure. The calibrations and conditional simulations presented in Figures 16 and 17 show dependency of modeled subsurface conditions on both correlation length calibrations and conditioning locations.

The full CPT data set results in a well-defined vertical mean length of 1.05 m and a horizontal mean length of 61.0 m, respectively. The sensitivity of geostatistical simulations to the number and location of conditioning CPTs is illustrated using the correlation length calibrations from 31 CPTs. Figure 18 summarizes the simulation calibrations in terms of category 1 horizontal mean length. In this application, additional CPTs serve solely as conditioning constraints. When the CPTs have separation distances greater than the horizontal mean length (61 m) the estimated subsurface conditions in the region between the soundings is highly variable. Adding CPTs increases the conditioning locations, thereby reducing simulation variance. Maintaining the correlation structure from the 31 CPT cross-section but reducing the number

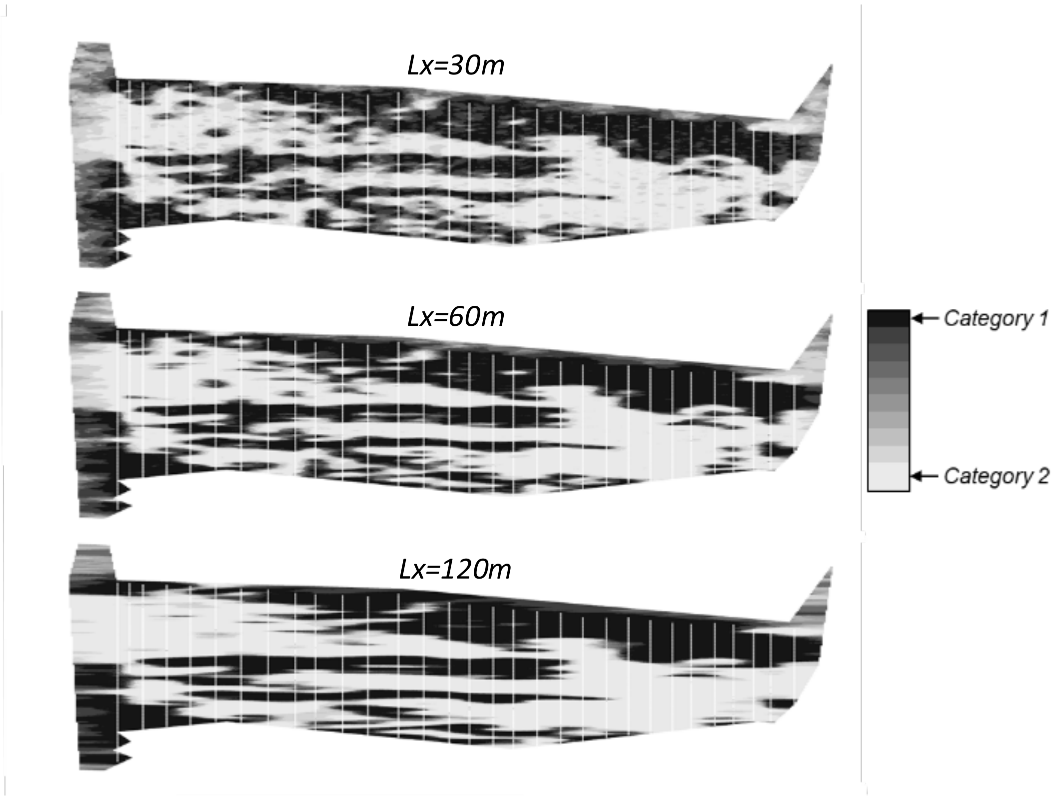

Figure 15. Comparison of 0.5, 1.0 and 2.0 times the calibrated mean length of 60 m for the 31 CPTs cross-section. Dashed white lines represent available CPT data.

of conditioning locations used in simulation shows the value of conditioning locations when the correlation structure is known.

The summary of calibration and simulation results from varying CPTs used in developing conditional simulations are shown in Figure 19, which reports the number of CPTs, individually calibrated sill, mean lengths, anisotropy, sampling spacings, and simulation error for each simulation. The sill is a relatively stable value since it is based on the vertical CPT information. Additional CPTs further refine the sill estimate from an initial value of 56% to a final value of 49%. The calibrated correlation lengths decrease with increased number of CPTs for both vertical and horizontal directions. The vertical calibration decreases from an initial estimate of 1.8 m to a final value of 1.05 m while the horizontal calibration decreases from an initial estimate of 436 m to a final value of 61 m. The mean length anisotropy is reported as L_x/L_z and the initial structural anisotropy of 243 at an average spacing of 244 m decreases to a final value of 58 at an average spacing of 16.2 m. In terms of S/L, the 3 CPT cross-section results in a horizontal S/L value of 0.56 while the final 31 CPT cross-section has a S/L value of 0.27 when using correlation lengths calibrated to the respective number of CPTs. Assigning the correlation length for 31 CPTs calibration to each cross-section decreases the denominator in the S/L ratio while not affecting the spacing, resulting in estimates of S/L as large as 4.0. The simulation error indicates decreasing error with increasing number of CPTs or decreasing value of S/L when simulation error is referenced to a single realization of the 31 CPT simulation.

If simulations are developed solely on the available CPTs, increasing the number of CPTs used in geostatistical simulations will affect both model calibration and conditional simulations. This provides insight to the stability of geostatistical simulations using incremental site investigation to consider the evolution of model calibration and conditional simulations with increasing site investigation information.

An important observation is that the ability to rigorously evaluate subsurface structure requires the ability to quantify correlation length and simulation uncertainty. Correlation uncertainty using the S/L framework suggests that values less than 0.3 are well constrained and values above 0.5 are either minimally constrained or unconstrained. Quantifying simulation uncertainty is more challenging because it requires establishing an objective metric with which to compare simulations. It is impossible to establish a rigorous simulation of the subsurface conditions for unconstrained sites with larger S/L values, since there is no ground truth to compare against. Unconstrained simulations can occur with both small and large numbers of CPTs performed or with large or small spacings between nearby CPTs, since subsurface features can have both large or small correlation lengths. Normalizing spacing with correlation length provides an indicator for when to use stratigraphic simulations and when to perform alternate analyses using equivalent uniform analyses (e.g., Boulanger and Montgomery 2016).

6. Incorporating Geostatistics into Integrated Site Characterization Framework

The incorporation of geostatistics into an integrated site characterization framework is conceptualized in the diagram shown in Figure 20. In this approach, a geologic

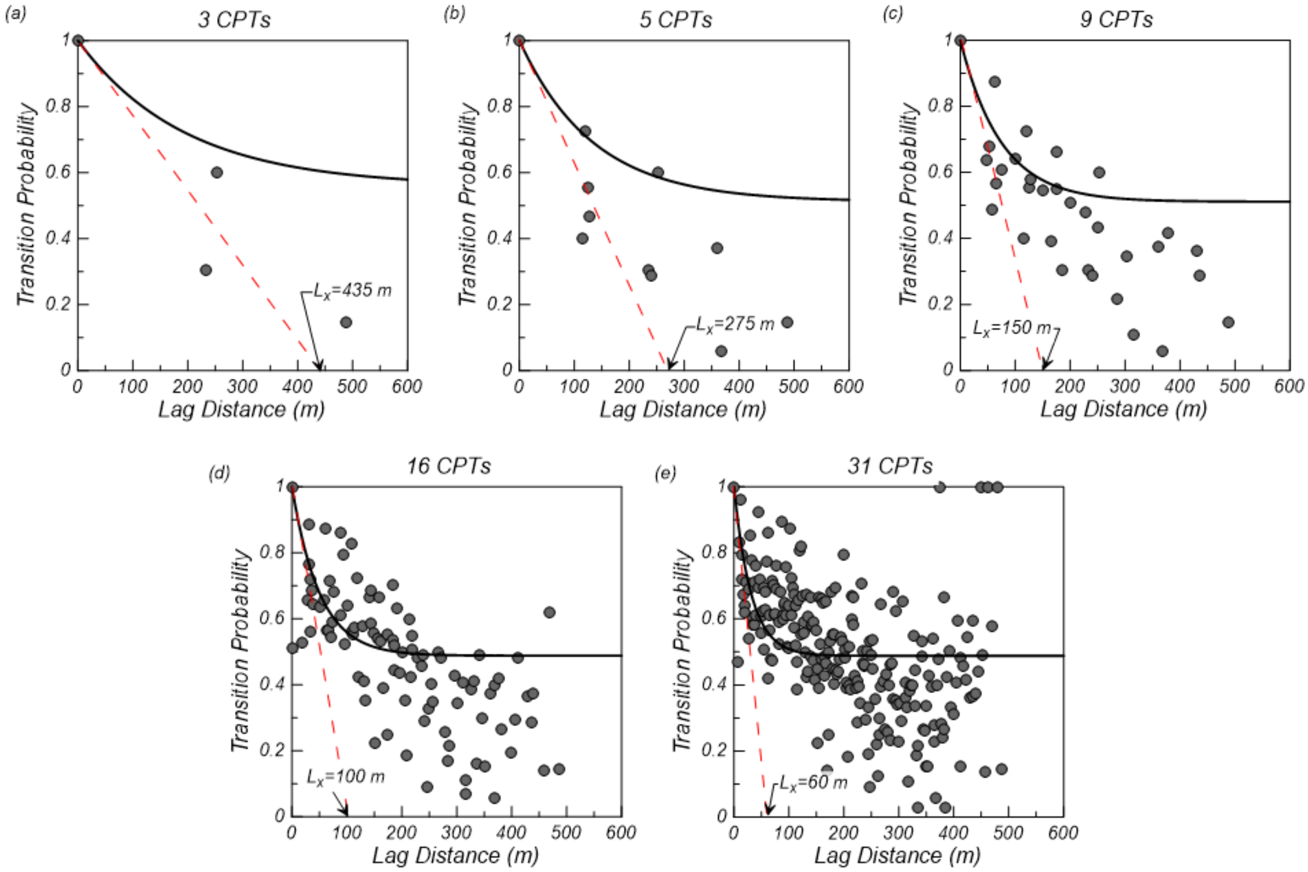

Figure 16. Subset of transition probability calibrations for the autocorrelation for zones of expected liquefaction in the horizontal (cross-channel) direction. Calibrations from left (a) to right (f) represent increasing subsurface knowledge through increased number of CPT soundings and resulting lag distance pairs.

model hypothesis of the subsurface is developed along with the relevent performance mechanisms. This information can be used to design an initial site investigation plan, followed by performing a portion (first phase) of the site investigation. Subsurface information can be gathered from the site investigation and statistically evaluated to determine representative properties, cumulative distributions, etc. Following the statistical evaluation of subsurface information, the subsurface may be analyzed to estimate correlation lengths of the subsurface spatial structure. The outcomes of this can be used to assess, verify, and refine the initial geologic modeling hypotheses. Depending on what stage the site characterizaton process is in, this information can be used to update and refine the investigation plan for the next phase or work.

The iterative loop can occur at different layers of complexity and across different time scales. One approach is to perform simplified versions of this analysis as each piece of information is gathered during site investigation, in near real-time. The information obtained is analyzed and added to a working database to assess subsurface properties and spatial structure. Each new sounding provides additional information that can confirm existing hypotheses or change the current understanding of subsurface conditions. Another approach is to perform site investigation in incremental campaigns, where units of subsurface information are gathered, the geostatistical model is updated, the main sources of uncertainty identified, the next phase of the site characterization program is planned, and so on.

7. Conclusions

This paper has examined an idealized 3D site and a 2D cross-section of an industry project using a transition probability geostatistics method to provide more firm guidance in how site investigation programs can be performed to obtain information useful for evaluating subsurface variability in addition to identifying soil types and properties. This has led to the following observations:

- The category sill value (absolute proportion of material types across the site) tends to stabilize with relatively few CPT soundings due to the high sampling frequency in the vertical direction in CPT soundings. Once the category sill is defined, the focus can shift toward defining the correlation structure (i.e. mean lengths in the horizontal directions).
- The mean length is typically overpredicted when the CPT spacing is large, a condition common for many site investigations when the number of soundings performed is relatively small (resulting in large spacing) and the mean lengths are also relatively small.

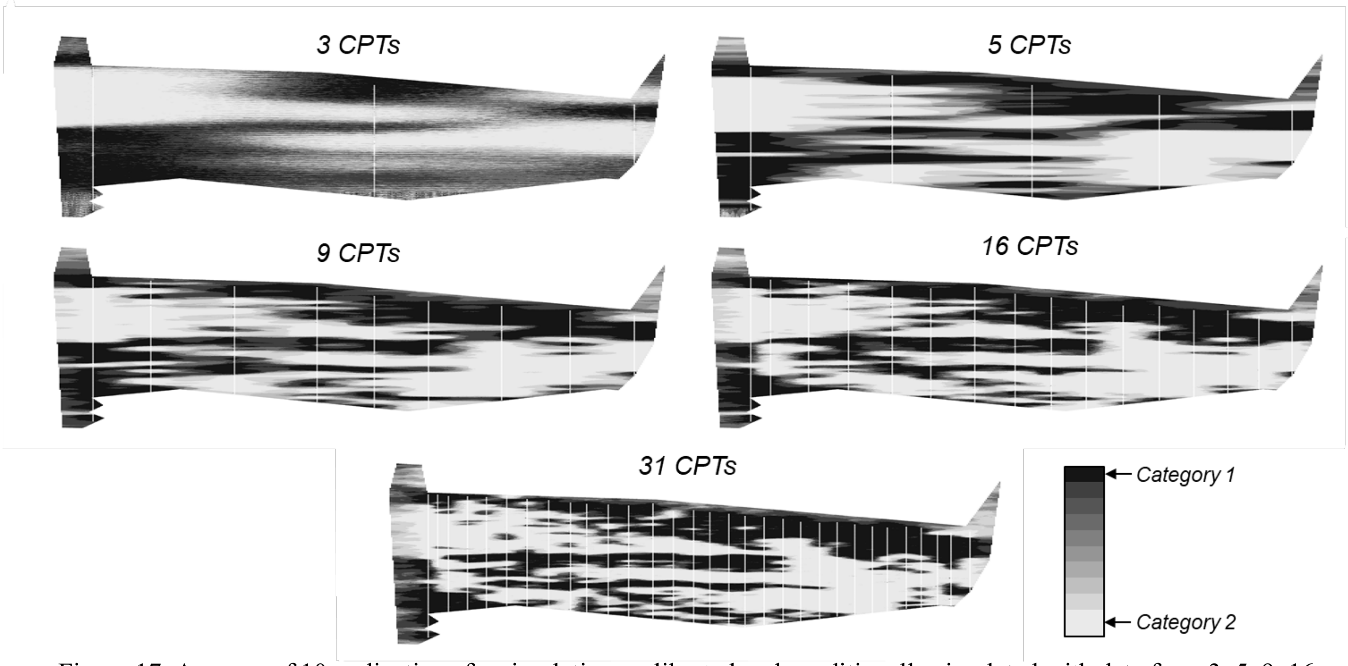

Figure 17. Average of 10 realizations for simulations calibrated and conditionally simulated with data from 3, 5, 9, 16, and 31 CPTs with color shades representing the average category value at each simulation node. Dashed white lines represent available CPT data.

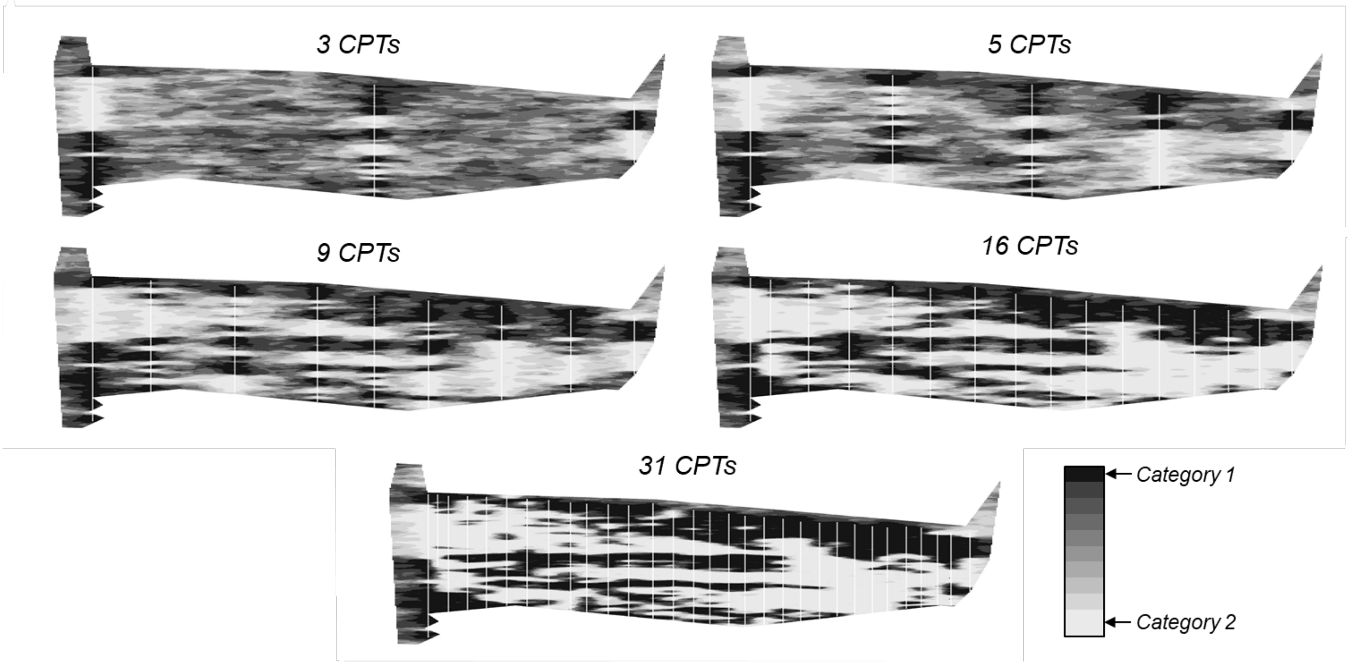

Figure 18. Average of 10 realizations for simulations calibrated using 31 CPTs with correlation lengths of 1.05 and 61.0 m in the vertical and horizontal directions, respectively. The 31 CPT correpation model is conditionally simulated with data from 3, 5, 9, 16, and 31 CPTs, with color shades representing the average category value at each simulation node. Dashed white lines represent available CPT data.

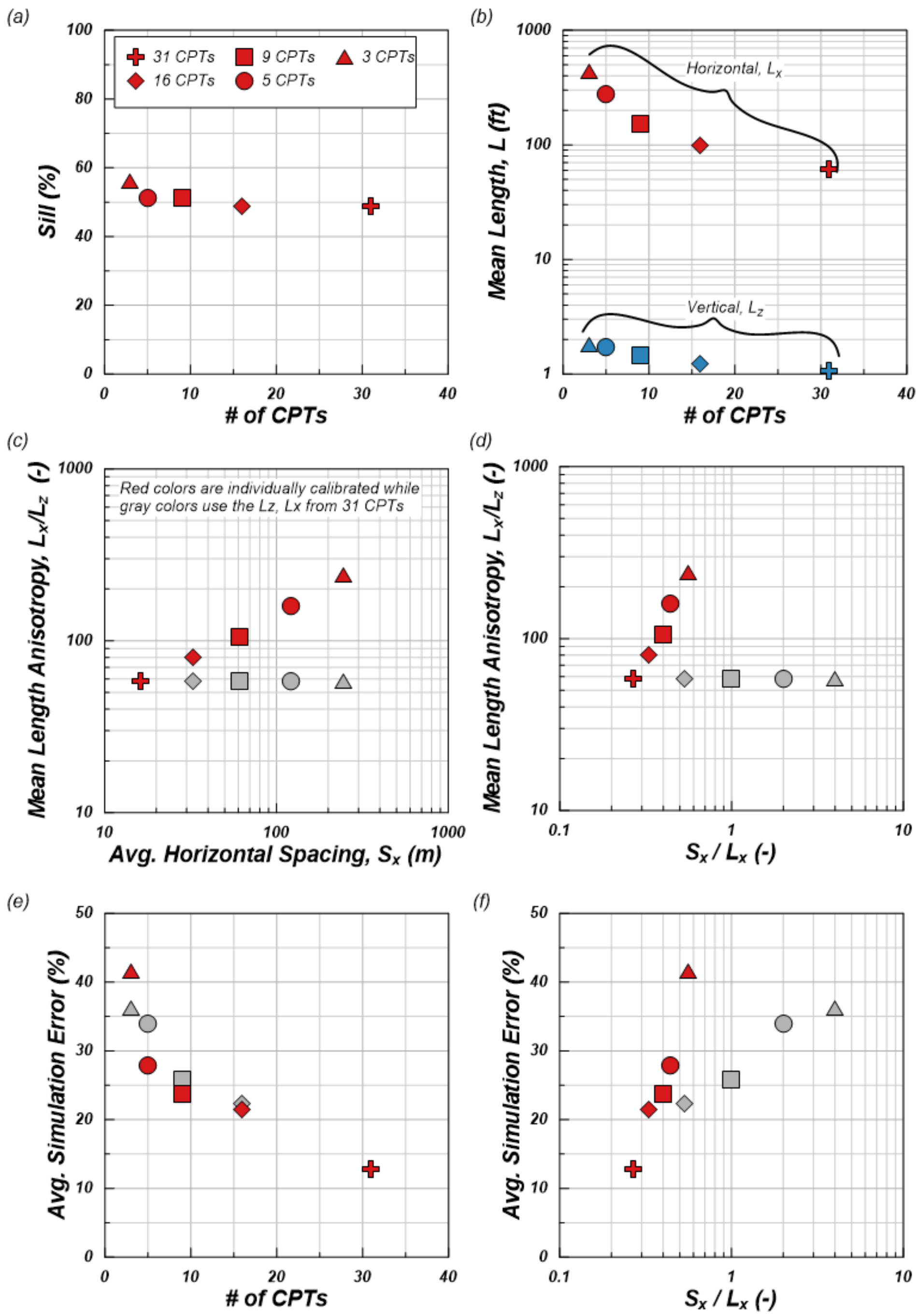

Figure 19. Summary of sensitivity to multiple correlation calibrations and conditional simulations for the 2-D cross-section. Simulation error is compared to a single realization of the 31 CPTs cross-section.

- An equally spaced grid and a nested layout for CPT soundings each have advantages and disadvantages. The grid layout is effective at distributing conditioning locations for simulation across the entire site but less effective at estimating the correlation structure. In contrast, the nested layout is more effective at estimating correlation structure when fewer CPTs are used, but is less effective at distributing conditioning locations for simulation. In practice, a combined approached may be most appropriate, where grid spacing provides sufficient site coverage while the nested layout improves mean length calibrations.
- The transition probability categorical approach allows category bins to be defined in terms of direct measurement values, soil types, or performance metrics. The simulated 3D example demonstrated the effectiveness of site investigation at capturing spatial correlation structure and conditional simulations of subsurface conditions. The latter 2D industry example demonstrated the effectiveness of defining the category bins based on a performance threshold and the demonstrated the evolution of geostatistical calibration parameters and resulting simulations using additional site investigation information.
- The robustness of spatial correlation calibrations using a transition probability geostatistical approach depends on the spacing to mean length ratio (S/L). A value of S/L that is less than 0.3 will likely produce

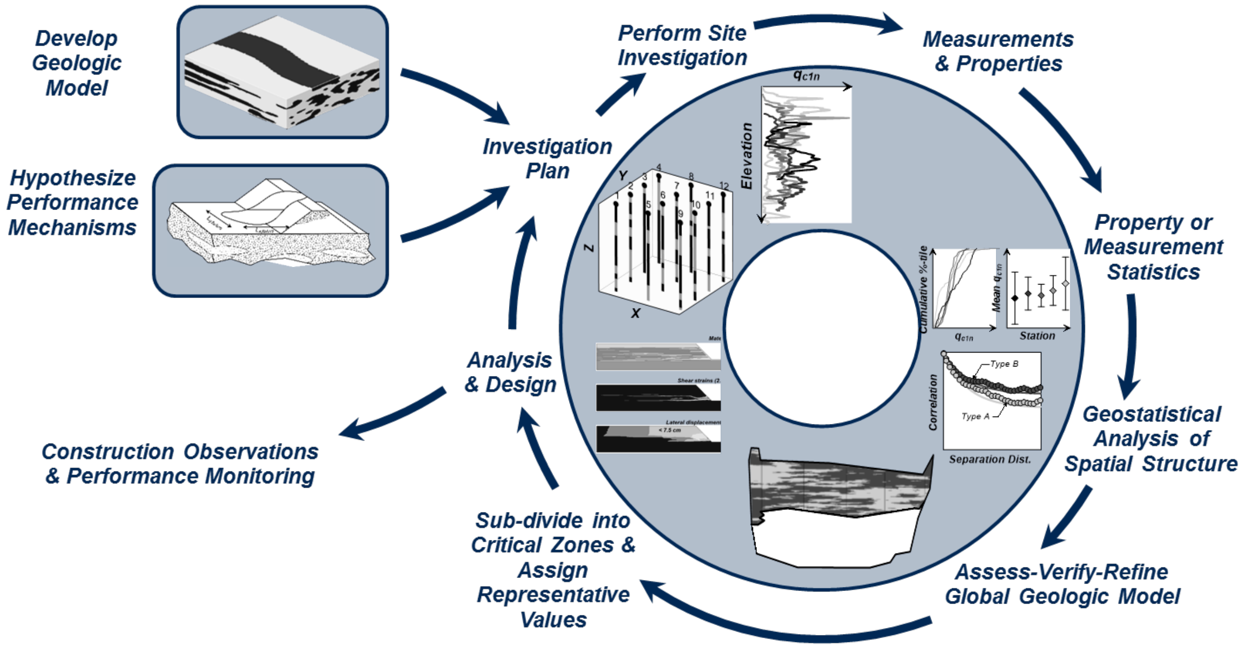

Figure 20. Framework for implementing geostatistical simulations into site characterization.

a (reasonably) well constrained correlation model calibration while a value that is 0.5 or greater will likely produce an unconstrained correlation calibration. It is important to note that estimate of mean length (L) generally decreases as the sounding spacing decreases; therefore, in practical application it may be prudent to examine this ratio with respect to both the calibrated mean length and a range of possible L values defined based on the calibrated mean length and supplemental soft information (e.g. depositional processes, anticipated length scales of possible features, etc.)

- Implementation of the approach presented herein will likely constitute an iterative process, as depicted in Figure 20. This approach is begins with development of a geologic model, which is then iteratively refined during the site investigation program or between phases of the site investigation program. In all likelihood the desired amount of information will not be collected due to time and financial project constraints, and it will be necessary to employ judgement when deciding when a sufficient amount of data has been obtained. The process as presented is intended to bring a systematic structure to integrating soft and hard information as it is collected, and to determine the potential value of pursuing additional data.

Acknowledgements

This study was primarily supported by the US National Science Foundation (NSF) under Grant Nos. CMMI-1436793 and CMMI-1436617. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the NSF.

References

[1] Al-Bittar, T. and Soubra, A.H. (2014). "Probabilistic analysis of strip footings resting on spatially varying soils and subjected to vertical or inclined loads." ASCE JGGE 140(4), https://doi.org/10.1061/(ASCE)GT.1943-5606.0001046.

[2] Baecher, G.B. and Christian, J.T. (2005). *Reliability and statistics in geotechnical engineering*. John Wiley & Sons.

[3] Bong, T., and Stuedlein, A.W. (2017). "Spatial variability of CPT parameters and silty fines in liquefiable beach sands." ASCE JGGE, 143 (12), https://doi.org/10.1061/(ASCE)GT.1943-5606.0001789.

[4] Boulanger, R. W., and Idriss, I. M. (2015). "CPT-based liquefaction triggering procedure." ASCE JGGE, 142(2), https://doi.org/10.1061/(ASCE)GT.1943-5606.0001388.

[5] Boulanger, R. W., and Montgomery, J. (2016). "Nonlinear deformation analyses of an embankment dam on a spatially variable liquefiable deposit." Soil Dynamics and Earthquake Engineering, 91(2016), 222-233, https://doi.org/10.1016/j.soildyn.2016.07.027.

[6] Cao, Z., and Wang, Y. (2013). "Bayesian approach for probabilistic site characterization using cone penetration tests." ASCE JGGE, 139(2), 267-276, https://doi.org/10.1061/(ASCE)GT.1943-5606.0000765.

[7] Carle, S.F. (1996). "A transition probability-based approach to geostatistical characterization of hydrostratigraphic architecture." PhD Dissertation, University of California, Davis, Davis, CA.

[8] Carle, S.F. and Fogg, G.E. (1996). "Transition probability-based indicator geostatistics." Mathematical Geology, 28 (4).

[9] Cotterill, C., Phillips, E., James, L., Forsbery, C., Tjelta, T.I. (2017). "How understanding past landscapes might inform present-day site investigations: a case study from Dogger Bank, southern central North Sea." Near Surface Geophysics, 15(4):403-414, https://doi.org/10.3997/1873-0604.2017032

[10] DeGroot, D.J., and Baecher, G.B. (1993). "Estimating the autocovariance of in-situ soil properties." ASCE JGGE 119(1): 147-166.

[11] DeJong, J.T., Sturm, A., Ghafghazi, M. "Characterization of Gravelly Alluvium", Soil Dynamics and Earthquake Engineering, 2016, 91: 104–115, https://doi.org/10.1016/j.soildyn.2016.09.032.

[12] Elkateb, T., Chalaturnyk, R., and Robertson, P. K. (2003). "An overview of soil heterogeneity: quantification and implications on geotechnical field problems." Can. Geotech. J., 40(1), 1-15, https://doi.org/10.1139/t02-090.

[13] Fenton, G. A. (1999). "Random field modeling of CPT data." ASCE JGGE, 125(6), 486-498, https://doi.org/10.1061/(ASCE)1090-0241(1999)125:6(486).
[14] Fenton, G. A., and Vanmarcke, E. H. (1990). "Simulation of random fields via local average subdivision." ASCE JGGE, 116(8), 1733-1749, https://doi.org/10.1061/(ASCE)0733-9399(1990)116:8(1733).
[15] Gooverts, P. (1997). *Geostatistics for environmental applications*. Oxford University Press.
[16] Griffiths, D.V., Fenton, G.A., Manoharan, N. (2002). "Bearing capacity of rough rigid strip footing on cohesive soil: a probabilistic study." ASCE JGGE, 128(9), https://doi.org/10.1061/(ASCE)1090-0241(2002)128:9(743).
[17] Jaksa, M. B., Goldsworthy, J. S., Fenton, G. A., Kaggwa, W. S., Griffiths, D. V., Kuo, Y. L., and Poulos, H. G. (2005). "Towards reliable and effective site investigations." Géotechnique, 55(2), 109-121, https://doi.org/10.1680/geot.2005.55.2.109.
[18] Juang, C. H., Jiang, T., and Christopher, R. A. (2001). "Three-dimensional site characterization: neural network approach." Géotechnique, 51(9), 799-809, https://doi.org/10.1680/geot.2001.51.9.799.
[19] Li, J., Tian, Y., and Cassidy, M. J. (2015). "Failure Mechanism and Bearing Capacity of Footings Buried at Various Depths in Spatially Random Soil." ASCE JGGE, 141(2), 04014099, https://doi.org/10.1061/(ASCE)GT.1943-5606.0001219.
[20] Li, X. Y., Zhang, L. M., and Li, J. H. (2015). "Using Conditioned Random Field to Characterize the Variability of Geologic Profiles." ASCE JGGE, 142(4), https://doi.org/10.1061/(ASCE)GT.1943-5606.0001428.
[21] Lloret-Cabot, M., Hicks, M. A., and van den Eijnden, A. P. (2012). "Investigation of the reduction in uncertainty due to soil variability when conditioning a random field using Kriging." Géotechnique Letters, 2(3), 123-127, https://doi.org/10.1680/geolett.12.00022.
[22] Lloret-Cabot, M., Fenton, G. A., and Hicks, M. A. (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, https://doi.org/10.1080/17499518.2013.871189.
[23] Montgomery, J. and Boulanger, R.W. (2017). "Effects of Spatial Variability on Liquefaction-Induced Settlement and Lateral Spreading". ASCE JGGE, 143(1):04016086, https://doi.org/10.1061/(ASCE)GT.1943-5606.0001584.
[24] Nichols, G. (2009). Sedimentology and stratigraphy. John Wiley & Sons.
[25] Olea, R.A. (1984). "Systematic sampling of spatial functions." Series of spatial analysis No. 7, Kansas Geological Survey, Lawrence, Kans.
[26] Paull, N.A., Boulanger, R.W., and DeJong J.T. (2020). "Accounting for Spatial Variability in Nonlinear Dynamic Analyses of Embankment Dams on Liquefiable Deposits", J. Geotech. Geoenviron. Eng., 146(11), https://doi.org/10.1061/(ASCE)GT.1943-5606.0002372.
[27] Phoon, K. K., and Kulhawy, F. H. (1999). "Characterization of geotechnical variability." Can. Geotech. J., 36:612-624, https://doi.org/10.1139/t99-038.
[28] Vanmarcke, E. H. (1977). "Probabilistic modeling of soil profiles." ASCE JGGE, 103(11), 1227-1246.