Assessing uncertainty propagation in CPTu-based hydro-mechanical subsoil characterization using a multivariate stochastic simulation approach

Di Curzio, Diego; Castrignanò, Annamaria; Vessia, Giovanna

DOI
10.1016/j.enggeo.2025.108064

Publication date
2025

Document Version
Final published version

Published in
Engineering Geology

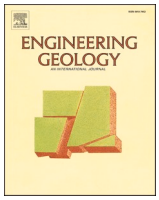
Citation (APA)
Di Curzio, D., Castrignanò, A., & Vessia, G. (2025). Assessing uncertainty propagation in CPTu-based hydro-mechanical subsoil characterization using a multivariate stochastic simulation approach. *Engineering Geology*, *352*, Article 108064. https://doi.org/10.1016/j.enggeo.2025.108064

Engineering Geology 352 (2025) 108064

Contents lists available at ScienceDirect

Engineering Geology

journal homepage: www.elsevier.com/locate/enggeo



Assessing uncertainty propagation in CPTu-based hydro-mechanical subsoil characterization using a multivariate stochastic simulation approach

Diego Di Curzio [a,*], Annamaria Castrignanò [b], Giovanna Vessia [b]

[a] Department of Water Management, Delft University of Technology, Stevinweg 1, 2628 CN Delft, Netherlands
[b] Department of Engineering and Geology (InGeo), University "G. d'Annunzio" of Chieti-Pescara, Via dei Vestini 31, 66013 Chieti, Italy

ARTICLE INFO

Keywords:
3D subsoil models
Hydro-mechanical properties
Propagated uncertainty
Alluvial deposits
Sequential Gaussian co-simulation
CPTu profiles

ABSTRACT

Estimating the spatial distribution of hydromechanical properties in the investigated subsoil by defining an Engineering Geological Model (EGM) is crucial in urban planning, geotechnical designing and mining activities. The EGM is always affected by (i) the spatial variability of the measured properties of soils and rocks, (ii) the uncertainties related to measurement and spatial estimation, as well as (iii) the propagated uncertainty related to the analytical formulation of the transformation equation. The latter is highly impactful on the overall uncertainty when design/target variables cannot be measured directly (e.g., in the case of piezocone Cone Penetration Test–CPTu measurements). This paper focuses on assessing the Propagated Uncertainty (PU) when defining 3D EGMs of three CPTu-derived design/target variables: the undrained shear resistance (s_u), the friction angle (φ'), and the hydraulic conductivity (k). We applied the Sequential Gaussian Co-Simulation method (SGCS) to the measured profiles of tip (q_c) and shaft resistance (f_s), and the pore pressure (u_2), measured through CPTus in a portion of Bologna district (Italy). First, we calculated 1000 realizations of the measured variables using SGCS; then, we used the available transformation equations to obtain the same number of realizations of s_u, φ', and k. The results showed that PU is larger when the transformation equation used to obtain the design/target variable is very complex and dependent on more than one input variable, such as in the case of k. Instead, linear (i.e., for s_u) or logarithmic (i.e., for φ') transformation functions do not contribute to the overall uncertainty of results considerably.

1. Introduction

The reliability assessment of the subsoil geological model for mining and engineering purposes is a pressing issue in designing duty (Phoon and Tang, 2019) and public reports by mining companies (McManus et al., 2021). Due to the risks related to building (infra)structures, mining activities, as well as their environmental impact, the quantification of the uncertainty affecting the modeling of the continuous subsurface volume in the Engineering Geological Model (EGM) from sparse spatial datasets of different nature is an urgent need (Baynes and Parry, 2022). EGMs mostly consist of heterogeneous geological bodies, whose hydro-mechanical properties are described by qualitative and quantitative variables. These variables are assessed or measured through various geological, geophysical, and geotechnical investigations. These include both point-based methods (e.g., borehole logging, penetration tests, in-hole geophysical tests) and 2D/3D methods (e.g., indirect geophysical methods, geological survey).

EGM of estimated design/target variables are always more uncertain than measured ones because they suffer from both measurement error of input data (Cherubini et al., 2007) and model uncertainty (Zhao et al., 2022), whose combination can be referred to as "prediction uncertainty". Phoon and Tang (2019) highlighted the need to quantify the degree of confidence in EGM by additionally considering its spatial variability.

Most of the quantitative design/target are calculated through transformation equations, which convert field measurements into hydro-mechanical subsoil properties (Robertson and Cabal, 2014). The uncertainty introduced by empirically derived deterministic equations from the literature is often unknown, especially because in most cases they have not been validated. Hence, the prediction uncertainty, together with the natural spatial variability of subsoil hydro-mechanical properties, propagates through the transformation equations. Assessing this "propagated uncertainty" (PU) is crucial in geo-engineering and engineering geology (Erharter et al., 2024), as it is a fundamental

* Corresponding author.
E-mail address: DiCurzio@tudelft.nl (D. Di Curzio).

https://doi.org/10.1016/j.enggeo.2025.108064
Received 13 July 2024; Received in revised form 25 March 2025; Accepted 3 April 2025
Available online 5 April 2025

requirement for the upcoming revised version of Eurocodes (Lesny et al., 2017) and the Joint Ore Reserves Committee guidelines (JORC, 2012).

Several methods exist to assess the uncertainty in hydro-mechanical characterization quantitatively (McManus et al., 2021; Zhang et al., 2023). Among them, one of the most used is Monte Carlo Simulation (Pownuk, 2018), which assumes the independence of the observations. Another widely used approach is the Bayesian approximation (Savoy et al., 2017; Wang et al., 2016, 2022; Zhan et al., 2023), which is often used in combination with the Monte Carlo Markov Chain and Gibbs sampling (Rahman et al., 2016; Hsu et al., 2022). However, these methods are sensitive to (i) the number of variables, (ii) their non-Gaussian distribution, and (iii) the high spatial correlation (Lee and Chen, 2009; McManus et al., 2021; Shao et al., 2023). In geo-engineering and environmental analysis (Phoon, 2020; Di Curzio et al., 2021; Kokkala and Marinos, 2022), machine learning techniques, such as the Bayesian ones (Wei and Wang, 2022; Phoon, 2023; Zhao et al., 2023), represent an important class of data-driven models able to quantify the uncertainty of an inference. Nevertheless, Heuvelink and Webster (2022) remarked that machine learning approaches play an increasingly relevant role due to the availability of large datasets and data fusion, but they are in essence non-spatial. They operate on a "regression matrix" that stores paired observations of the dependent (i.e., target) and independent variables, regardless of their spatial coordinates. This implies that there are many situations in which machine learning is less effective, such as (i) in the case of relatively small sets of observations for reliable calibration, (ii) when available environmental variables are only weakly related to design/target variables, and (iii) when data are spatially correlated.

In most of these approaches, the spatial correlation of data represents a clear limitation, whereas geostatistics have always represented a suite of methods able to explicitly and effectively handle the spatial correlation, since its early time (Matheron, 1962). Several methods have been developed to deal with non-Gaussian distributions, auxiliary covariates, geographic trends, and external drift, and specifically used in estimating the natural variability of subsoil volumes (Gao et al., 2018; Vessia et al., 2020a, 2020b; Heuvelink and Webster, 2022; Lachérade et al., 2023; Liang et al., 2024). Among these, stochastic simulation approaches have been increasingly used to investigate properly the uncertainty associated with the prediction rather than just the estimation itself, because these methods can reproduce the actual spatial variability (Castrignanò and Buttafuoco, 2004; Castrignanò et al., 2008; ASTM International, 2018, Heuvelink and Webster, 2023). Simulation methods are not weakened by large uncertainty, enabling the prediction of every type of data, both stationary and non-stationary, and it is truly data-driven (Lilla Manzione et al., 2021). According to Delbari et al. (2009), Emery and Peláez (2011), and Nussbaumer et al. (2018), the Sequential Gaussian Simulation technique is one of the most widely used conditional parametric simulation methods and it performs even better in multivariate cases, such as the Sequential Gaussian Co-Simulation (Emery and Peláez, 2011).

In the previous studies, stochastic simulation has always been used to quantify local estimation uncertainty; however, it has never been applied to quantitatively assess how prediction uncertainty (i.e., input uncertainty) propagates through the transformation equations used to estimate design/target hydro-mechanical variables. Therefore, this paper proposes for the first time a rationale to quantitatively assess propagated (PU) in subsoil characterization. We applied the Conditional Sequential Gaussian Co-Simulation method (SGCS) to the measured profiles of tip (q_c) and shaft resistance (f_s), and the pore pressure at shoulder position (u_2) from piezocone Cone Penetration Tests (CPTu) and used specific equations to estimate three key hydro-mechanical variables (i.e., the undrained shear resistance s_u, the friction angle φ', and the hydraulic conductivity k) in a 3D EGM. The set of equiprobable realizations obtained for the input variables with SGCS was used to calculate the same number of realizations of the design/target variables, allowing the direct investigation of the propagated uncertainty.

2. Material and methods

2.1. Study area and dataset

The considered study area is located in Emilia Romagna Region (Italy), in the Po plain eastwards of Bologna, where hundreds-of-meter of alluvial deposits are present. They consist of undifferentiated fine silty-sandy deposits (i.e. flooding plain), characterized by coarser (i.e. alluvial fans and paleo-channels) and finer (i.e. lacustrine lenses) geological bodies (ISPRA, 2009). From a lithological point of view, these are sandy, gravelly, and silty-clayey soils (Fig. 1). More details about the geological features of this area can be found in Vessia et al. (2020b), and Di Curzio and Vessia (2021).

The dataset used in this research consists of 182 CPTus performed across an area with an extension of 900 km^2 and investigating a volume of 92 km^3 (with a depth of 30 m). These CPTu profiles were collected in a comprehensive database by the Regional Office for Territorial Protection and Development of the Emilia-Romagna region (http://geoportale.regione.emilia-romagna.it/it), and subsequently made available by Di Curzio and Vessia (2021).

2.2. Sequential Gaussian co-simulation

To assess PU in the process of obtaining design/target hydro-mechanical variables from input measurements by using transformation equations, the Sequential Gaussian Co-Simulation (SGCS) method was adopted (Goovaerts, 1997; Webster and Oliver, 2007; Chilès and Delfiner, 2012). This advanced geostatistical technique represents a multivariate version of Sequential Gaussian Simulation approach (SGS), which is one of the most straightforward and used among Conditional Simulation methods (Delbari et al., 2009; Emery and Peláez, 2011; Nussbaumer et al., 2018).

Since SGCS assumes a multi-Gaussian behavior, the variables that do not follow the Gaussian distribution, such as the input variables considered in this work (i.e., q_c, f_s and u_2), were transformed into Gaussian ones through the Gaussian Anamorphosis (Chilès and Delfiner, 2012). It consists in estimating a function converting a standardized Gaussian variable (Y) into a non-Gaussian one ($Z = \Phi(Y)$), through a Hermite polynomial expansion ($H_i(Y)$) truncated at a finite number of terms:

$$\Phi(Y) = \sum \Psi_i H_i(Y) \quad (1)$$

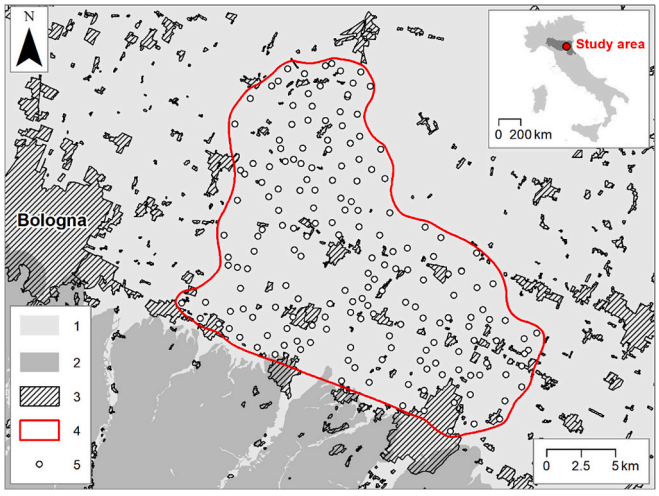


Fig. 1. Map showing the location and main geological features of the study area as well as the CPTu distribution within the selected domain. In legend: 1) alluvial deposits; 2) bedrock; 3) urban areas; 4) geostatistical domain; 5) CPTu locations.

where, Ψ_i are the coefficients of the Hermite polynomials to be estimated. Once defined the Gaussian Anamorphosis function, the transformation from a raw variable into the standardized gaussian one was performed by inverting the function, as follows:

$$Y = \Phi^{-1}(Z) \tag{2}$$

As in other multivariate geostatistical approaches (e.g., Co-Kriging), SGCS relies on the fitting of a Linear Model of Coregionalization (LMC) to the matrix of direct (i.e., diagonal) and cross- (i.e., non-diagonal) experimental variograms of the considered input variables (Wackernagel, 2003; Castrignanò et al., 2015; Di Curzio et al., 2019, Vessia et al., 2020a):

$$\boldsymbol{\Gamma}(\mathbf{h}) = \sum_{u=1}^{N_S} \mathbf{B}^u g^u(\mathbf{h}) \tag{3}$$

where, $g^u(\mathbf{h})$ is the spatial structure standardized to unit sill, and u is the spatial scale, $\mathbf{B}^u$ is the Coregionalization matrix of the LMC partial sills corresponding to the scale u, which is symmetric and semi definite positive, $\boldsymbol{\Gamma}(\mathbf{h})$ is the n x n matrix with direct variograms (i.e., diagonal elements) and cross-variograms (i.e., non-diagonal elements) modeled as a linear combination of N_S basic variogram functions, being N_S total number of spatial scales, and $\mathbf{h}$ is the lag.

The conditional sequential Gaussian simulation algorithm (Deutsch and Journel, 1998) was used to generate the simulations. This algorithm entails defining a randomly spaced grid covering the region of interest and a path through the grid that visits each node only once. Ordinary kriging estimator is used at each unsampled location to estimate the expected value of the multi-Gaussian probability distribution function based on data values that have been approximately normal score converted. Next, a random value is selected randomly from the cumulative probability distribution function, and the variable's corresponding value is taken to be one plausible simulated value at that particular location in space. The procedure is repeated at the simulated value and added to the conditioning data set. This process is repeated many times by different random seeds and paths crossing all nodes of the grid only once. As a result, several equiprobable representations of the spatial distribution of the considered variable are obtained, namely realizations. These realizations provide a statistical distribution at each node of the grid, instead of one estimated value and the corresponding error (i.e., as in kriging methods). Furthermore, their recurrence offers a quantitative and visual indicator of spatial uncertainty (Goovaerts, 1996).

In this paper, we calculated 1000 realizations, considering a simulation domain with a cell size equal to 500x500x0.5 m, which allowed describing the spatial variability of hydro-mechanical variables at a regional scale effectively. Even though the optimal number of realizations is equal or more than 500 (Szatmári et al., 2021; Angelini et al., 2023), we used the Error Variance (i.e., average squared differences between the measured and predicted values; Cressie, 2015) to select the best set of realization among several options (i.e., 200, 400, 600, 800, and 1000 realizations, respectively). By using the LMC described above in the SGCS method, it was possible to obtain 1000 three-dimensional realizations of the Gaussian transformed input variables, which were then back-transformed into the original units by the Gaussian Anamorphosis function. Subsequently, we applied the transformation equations described in Section 2.3 to the 1000 realizations of the input variables (i.e., q_c, f_s, and u_2) to calculate 1000 realizations of each of the design/target variables considered in this study (i.e., s_u, ϕ', and k). Hereinafter, we define the expected values "optimized" values when they are calculated through the stochastic simulation according to ASTM International (2018).

All the geostatistical analyses were performed using Isatis 2018, whose results have then been visualized through Isatis.neo (Geovariances, 2021).

2.3. Transformation equations

Starting from the set of 1000 equiprobable realizations of q_c, f_s, and u_2, we calculated the considered design/target variables: the undrained shear resistance (s_u), the friction angle (φ'), and the hydraulic conductivity (k). We selected the most appropriate transformation equations available in the literature to calculate the selected design/target variables, based on authors' prescriptions and according to the qualitative features of the investigated soils.

Since both s_u and φ' are valid only for certain lithotypes, it was first necessary to calculate the normalized Soil Behavior Type Index (I_{SBTn}; Robertson, 2009). This index allows identifying the litho-technical behavior of the investigated soils:

$$I_{SBTn} = \left[(3.47 - \log(Q_{tn}))^2 + (\log F_R + 1.22)^2\right]^{0.5} \tag{4}$$

where, $Q_{tn} = \left(\frac{q_t - \sigma_{v0}}{P_a}\right) \bullet \left(\frac{P_a}{\sigma'_{v0}}\right)^n$ is the normalized tip resistance, $F_R = \frac{f_s}{q_t - \sigma_{v0}} \bullet 100\%$ is the friction ratio, $q_t = q_c - u_2(1 - a)$, coefficient a equal to 0.8 (i.e., average value), and $n = 0.381 \bullet I_c + 0.05 \bullet \left(\frac{\sigma'_{v0}}{P_a}\right) - 0.15$, I_c is the non-normalized Soil Behavior Type Index (Robertson, 1990):

$$I_c = \left[\left(3.47 - \log\left(\frac{q_c}{p_a}\right)\right)^2 + \left(\log\left(\frac{f_s}{q_c} \bullet 100\right) + 1.22\right)^2\right]^{0.5} \tag{5}$$

whereas σ'_{v0} and σ_{v0} are the effective and total lithostatic stresses at each depth, respectively. The non-normalized Soil Behavior Type Index (Eq. (5)) was only used to estimate the coefficient n, needed for the tip resistance normalization.

Table 1 shows the relation between I_{SBTn} ranges of values and the classes of lithotypes, which are characterized by specific prevalent soil fractions.

This classification was used to obtain 1000 equiprobable configurations of lithotypes starting from the sets of realizations for q_c, f_s, and u_2. Then, it was possible to extract the optimized model of I_{SBTn} providing the reference lithotype distribution used to classify the simulation domain cell by cell.

For finer lithotypes (i.e., SBT classes 4, 5, and 6), the undrained shear strength (s_u, in MPa) was derived through the formula (Robertson and Cabal, 2014):

$$s_u = \frac{q_t - \sigma_v}{N_{kt}} \tag{6}$$

Typically, the cone factor, N_{kt}, varies from 10 to 18; in this study, we decided to use the average value, which is equal to 14 as recommended by Robertson and Cabal (2014) when the cone diameter is not specified.

For sand-like soils (i.e., SBT classes 3, and 4), Robertson and Campanella (1983) suggested Eq. (6) to estimate the peak friction angle (φ', in °) for uncemented, unaged, moderately compressible, predominately quartz sands, based on calibration chamber test results:

$$\tan(\varphi') = \frac{1}{2.68}\left[\log\left(\frac{q_c}{\sigma'_{vo}}\right) + 0.29\right] \tag{7}$$

Table 1
Soil behavior type classes as defined by Robertson (2009), with the corresponding I_{SBTn} values.

Soil Behavior Type	I_{SBTn}	Class
Gravelly sand to dense sand	< 1.31	SBT2
Sands – clean sand to silty sand	1.31–2.05	SBT3
Sand mixtures – silty sand to sandy silt	2.05–2.60	SBT4
Silt mixtures – clayey silt to silty clay	2.60–2.95	SBT5
Clays – silty clay to clay	2.96–3.60	SBT6
Organic soils – clay	> 3.60	SBT7

The lithotype defined by the SBT5 class (Table 1) was considered to have an intermediate mechanical behavior; thus, it was considered in the estimation of both s_u and φ'.

In addition to providing the reference lithotype model used to decide where to calculate either s_u or φ', the 1000 configurations of I_{SBTn} were used to obtain the hydraulic conductivity (k, in m/s) for all the lithotypes in Table 1, as follows (Robertson and Cabal, 2014):

$$k = \begin{cases} 10^{(0.952-3.04\bullet I_c)} \text{ when } 1.00 < I_{SBTn} \ 3.27 \\ 10^{(-4.52-1.37\bullet I_c)} \text{ when } 3.27 < I_{SBTn} < 4.00 \end{cases} \quad (8)$$

2.4. Uncertainty evaluation indicators

To quantify the propagated uncertainty, we defined the limits of the 68 % (that is the expected value plus or minus one standard deviation) confidence intervals by extracting from the 1000 realizations the values corresponding to their lower (LL; 16th quantile) and upper (UL; 84th quantile) limits by post-processing the simulation (i.e., q_c, f_s, and u_2) and transformation (i.e., s_u, φ', and k) results.

From all the LL and UL values, the Underestimation (UE%) and Overestimation (OE%) percentages were calculated (Vessia et al., 2020a, 2020b), according to the following equations:

$$UE(\mathbf{x})\% = \frac{|\bar{z}(\mathbf{x}) - LL(\mathbf{x})|}{\bar{z}(\mathbf{x})}\% \quad (9)$$

$$OE(\mathbf{x})\% = \frac{|\bar{z}(\mathbf{x}) - UL(\mathbf{x})|}{\bar{z}(\mathbf{x})}\% \quad (10)$$

These indicators measure the distance between the expected value of the predictions ($\bar{z}(\mathbf{x})$) and the limits of the corresponding confidence interval at each location, enabling the quantification and mapping of the local uncertainty of prediction, both in underestimation and overestimation.

To investigate the lithotype-dependent contribution to the PU, we proposed the Coefficient of Variation (CV) to compare the mean variability and the total variability of the 1000 realizations of both input and design/target variables, for each SBT class identified in the considered within the 3D subsoil model. The first CV is hereafter named "CV mean" whereas the second is named "CV realizations" (Lachérade et al., 2023). In addition, the Coefficient of Quartile Variation–CQV (Bonnet, 2006) was even calculated for each SBT class. This latter measure of relative dispersion is preferred when the variables are nonnormal distributed:

$$CQV = \frac{Q_3 - Q_1}{Q_3 + Q_1} \quad (11)$$

3. Results and discussion

3.1. Dataset description

Table 2 shows the basic statistics related to the raw CPTu measurements used in this work. In this table Q1, Q2, and Q3 are the first quantile (15 %), the second (50 %), and the third quantile (75 %), respectively. As typically observed for soils similar to the investigated ones (Lunne et al., 2002), q_c is generally two to three orders of magnitude larger than f_s (i.e., means equal to 2.54 and 0.093 MPa, respectively), while u_2 mean is equal to 0.33 MPa.

All the measured variables exhibited a non-Gaussian distribution, as suggested by the significant differences between mean and median values, the skewness and kurtosis values (i.e., consistently, > 0 and > 3, respectively), as well as the very large differences between Q3 and maximum values. Similar positively skewed, leptokurtic statistical distributions suggest the measured variables may refer to soils behaving hydro-mechanically in different ways, in terms of, e.g., cohesion, soil structure, grain size, and saturation. In a geological context as the one investigated in this paper, the presence of both finer and coarser inclusions within the predominant lithotype may explain these skewed distributions.

The use of the SGCS method then required the preliminary transformation of such non-Gaussian distributed variables through the Gaussian Anamorphosis function, during the pre-processing stage.

3.2. Modeling local uncertainty and its propagation

The experimental direct and cross-variograms calculated from the previously Gaussian transformed variables (i.e., gq_c, gf_s, and gu_2) were fitted with the Linear Model of Coregionalization, whose features are shown in Table 3 and Fig. 2.

This LMC used in the SGCS to obtain 1000 realizations of each of the input variables was modeled in a previous paper as a zonal anisotropy (Di Curzio and Vessia, 2021). The horizontal variability (i.e., spatial structures for the X and Y directions were considered the same according to a superficial isotropic assumption) was modeled through two spherical structures with ranges equal to 1200 and 12,000 m, respectively; the anisotropic vertical variability was modeled by three spherical structures with ranges equal to 2, 6, and 12 m, likely attributable to three different depositional structures. To reproduce the vertical trend observed in all three variables (Fig. 2) and be able to use a stationary geostatistical method like SGCS, we adopted a long-range (i.e., 100,000 m beyond the investigated depth) K-Bessel structure. The presence of this local trend at the extent of study depth is caused by the increasing lithostatic stress that influences the measured mechanical parameters along depth.

The LMC performances were tested through a cross-validation (Cressie, 2015), and the results are reported in Di Curzio et al. (2021).

Fig. 3 shows the 3D models (i.e., expected values of the 1000 realizations per variable) of both the input and design/target variables. To help refer the volumetric distributions of all the variables to the geo-lithological features of the investigated subsoil, we calculated the model of the I_{SBTn} index, describing the spatial distribution of lithotype classes. The predicted subsoil geometries were consistent with the literature (ISPRA, 2009).

Table 3
Features of the Linear Model of Coregionalization (LMC) related to the Gaussian transformed variables.

Variables	Horizontal LMC structures	Range (m)
gq_c, gf_s, gu_2	Spherical	1200
	Spherical	12,000

Variables	Vertical LMC structures	Range (m)
gq_c, gf_s, gu_2	Spherical	2
	Spherical	6
	Spherical	12
	K-Bessel	100,000

Table 2
Descriptive statistics of the three variables measured through CPTus and used as inputs in SGCS.

Variable	Unit	Count	Mean	Median	St. Dev.	Q1	Q3	Min	Max	Skewness	Kurtosis
q_c	MPa	187,705	2.54	1.79	3.27	1.35	2.4	0.02	68.22	5.944	52.53
f_s	MPa	187,634	0.093	0.086	0.057	0.053	0.119	0.007	0.917	1.726	9.323
u_2	MPa	166,057	0.33	0.26	0.28	0.09	0.52	0.01	2.73	1.018	4.075

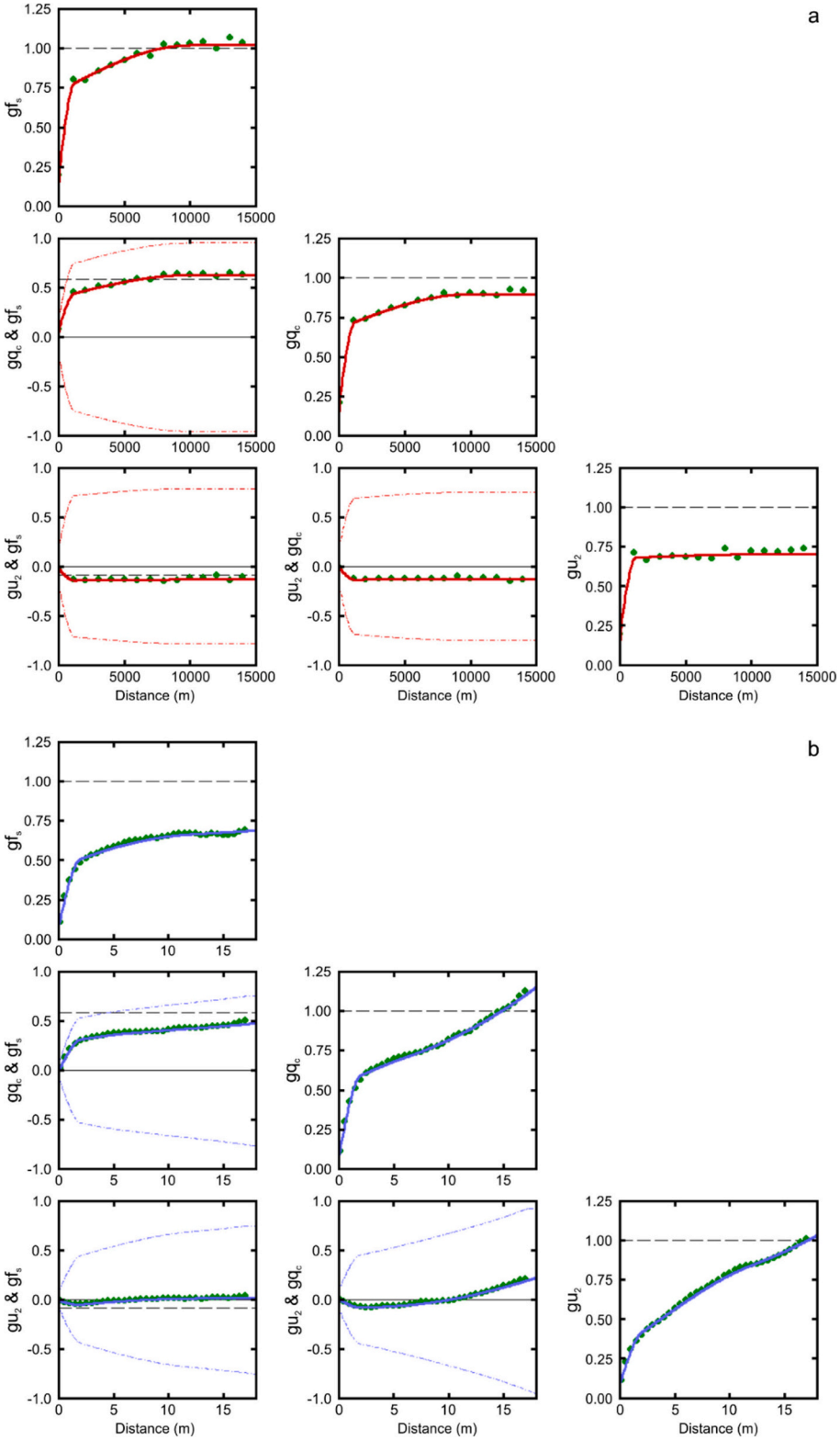


Fig. 2. LineModel of Coregionalization used in SGCS and referring to a) the horizontal (i.e., X and Y; red solid lines) and b) vertical (blue solid lines) directions, fitted to the experimental (cross-)variograms (green circles) calculated from the Gaussian transformed CPTu measurements (i.e., gq_c, gf_s, and gu_2). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

As a result, the highest estimated values of both the tip (q_c) and shaft resistance (f_s), correspond to the I_{SBTn} ranges referring to sand and sand mixtures (i.e., SBT3, and SBT4; Table 1). In this area these lithotypes are present as inclusions (i.e., alluvial fans, and paleochannels, as in the southeastern portion of the study area) within the finer deposits (i.e., silt mixtures of the SBT5 class), from about 10 m depth downward (Fig. 3). On the contrary, the shallowest coarse deposits, which likely describe the topsoil, do not show the same inverse correlation between both q_c and f_s and the I_{SBTn} values. The fact that deeper sediments typically have a higher density because of the increasing lithostatic stress with depth is

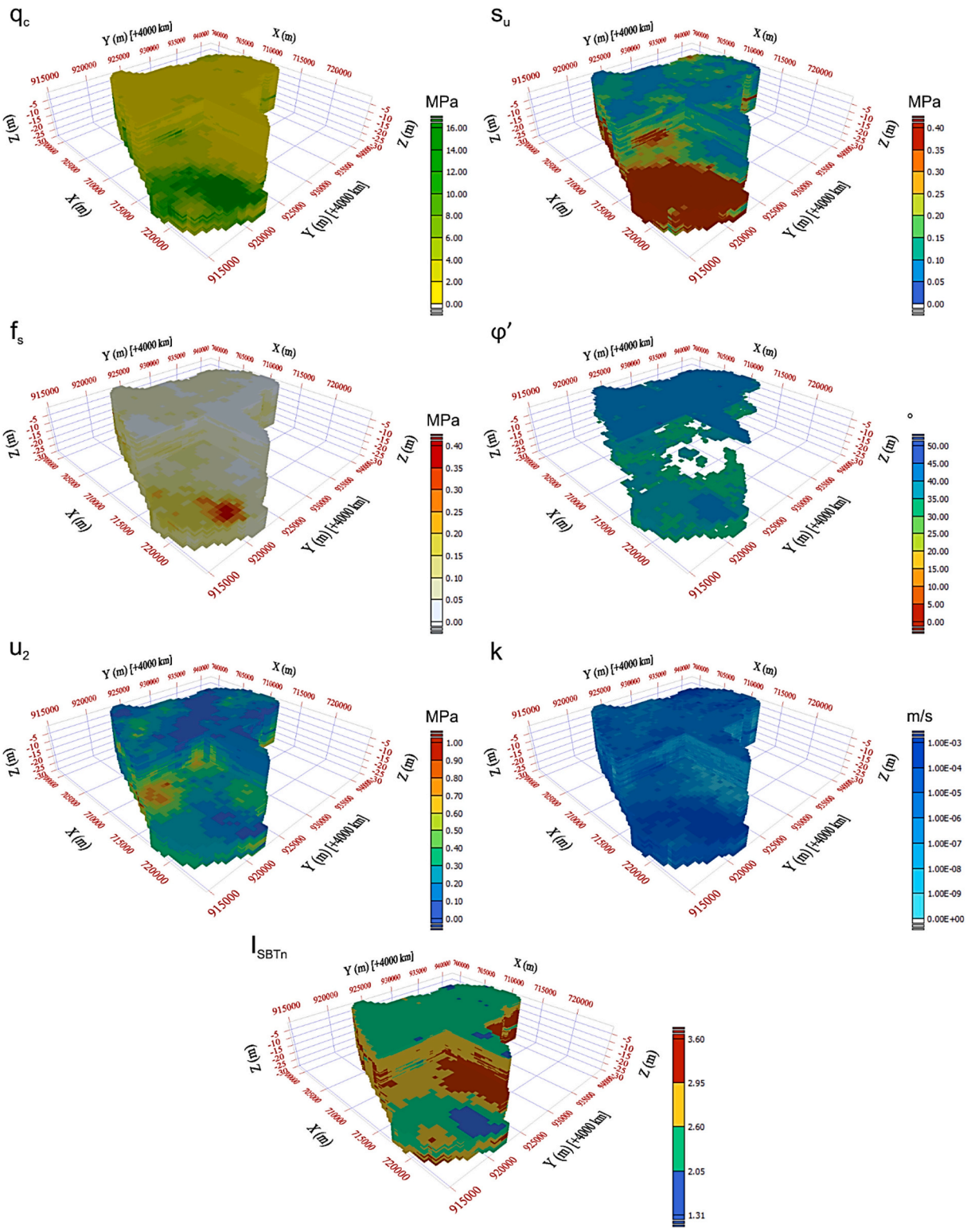


Fig. 3. 3D models of the input (i.e., q_c, f_s, and u_2) and design/target variables (i.e., s_u, φ′, and k), compared to the spatial distribution lithotypes, as defined by the I_{SBTn} in Table 1.

well known.

Comparing the 3D spatial variability of the pore pressure (u_2) and the 3D lithotype model, it appears clear that u_2 and I_{SBTn} are generally positively correlated, meaning that the highest u_2 values are typical of finer lithotypes (i.e., SBT5, and SBT6). In this area, the alluvial deposits host aquifers (Zuccarini et al., 2024), and the groundwater flow is usually more active in sands and sand mixtures than in silts and clays. This effect is mostly evident for the lithotype described by the SBT3 class (I_{SBT} value varying from 1.31 to 2.05). On the other hand, the typically high volume of water stored in fine deposits with a high clayey fraction is not free water but adsorbed.

As can be seen from Fig. 3, both the undrained shear resistance (s_u) and the friction angle (φ′) were estimated only in some portions of the investigated subsoil. s_u values were calculated only in those areas where

fine deposits (i.e., SBT5, and SBT6) or lithotypes likely containing a significant fraction of fine soils (i.e., SBT4) are present. Conversely, φ' was estimated only for granular lithotypes, described by SBT classes 3 and 4. Consequently, the number of cells in both s_u and φ' model is less than that of the variables estimated across the entire model grid. Moreover these design/target variables appear linked to the combined effect of mechanical properties of the considered lithotypes and local conditions.s_u is inversely proportional to I_{SBTn}, meaning that the finer the grain size the lower is the resistance of the considered deposits. Interestingly, however, the shallowest granular sediments containing a finer fraction (i.e., SBT4) – and likely describing the topsoil – show a higher variability and values similar to classes SBT5 and SBT6. A possible explanation can be either the higher content of organic matter and fine fraction or a lower density due to a very low lithostatic stress. φ', instead, is coherently higher where the granular fraction is predominant (i.e., with peak values above 40°), as in the sandy deposits described by the class SBT3, whereas it ranges between 25° and 40° elsewhere.

Regarding the hydraulic conductivity (k), its values are in the range of 10^{-4}–10^{-3} m/s where lithotypes with a prevalent granular fraction (i.e., SBT3, and SBT4) are predominant, typical of sandy aquifers. Differently, k decreases below 10^{-8} m/s in the case of clayey deposits (i.e., SBT6), identifying impervious geological bodies, which act as aquiclude in these kinds of porous aquifer. The SBT5 class describing the silty mixtures is characterized by intermediate values of k (i.e., in the order of 10^{-7}–10^{-5} m/s), which describe a typical hydrogeological behavior as 'aquitard'.

Figs. 4 and 5 show the spatial distribution of the local uncertainty for the input (i.e., q_c, f_s, and u_2) and design/target (i.e., s_u, φ', and k) variables, expressed in terms of underestimation (UE%) and overestimation (OE%) percentages, considering the 68 % confidence interval. The choice of displaying the uncertainty similarly comes from the necessity to take into account the experimental statistical distribution of the measured data, which was non-Gaussian (Table 2). Thus, this asymmetry needed to be included in the uncertainty analysis, and the UE% and OE% indexes offer such an opportunity.

As can be seen in Fig. 4, the distribution of the uncertainty associated with the input variables are significantly different. q_c shows UE% and OE% ranging between 12.5 and 37.5 % in those subsoil portions characterized by the presence of fine deposits (i.e., SBT5, and SBT6). However, the uncertainty increases significantly in the presence of coarse

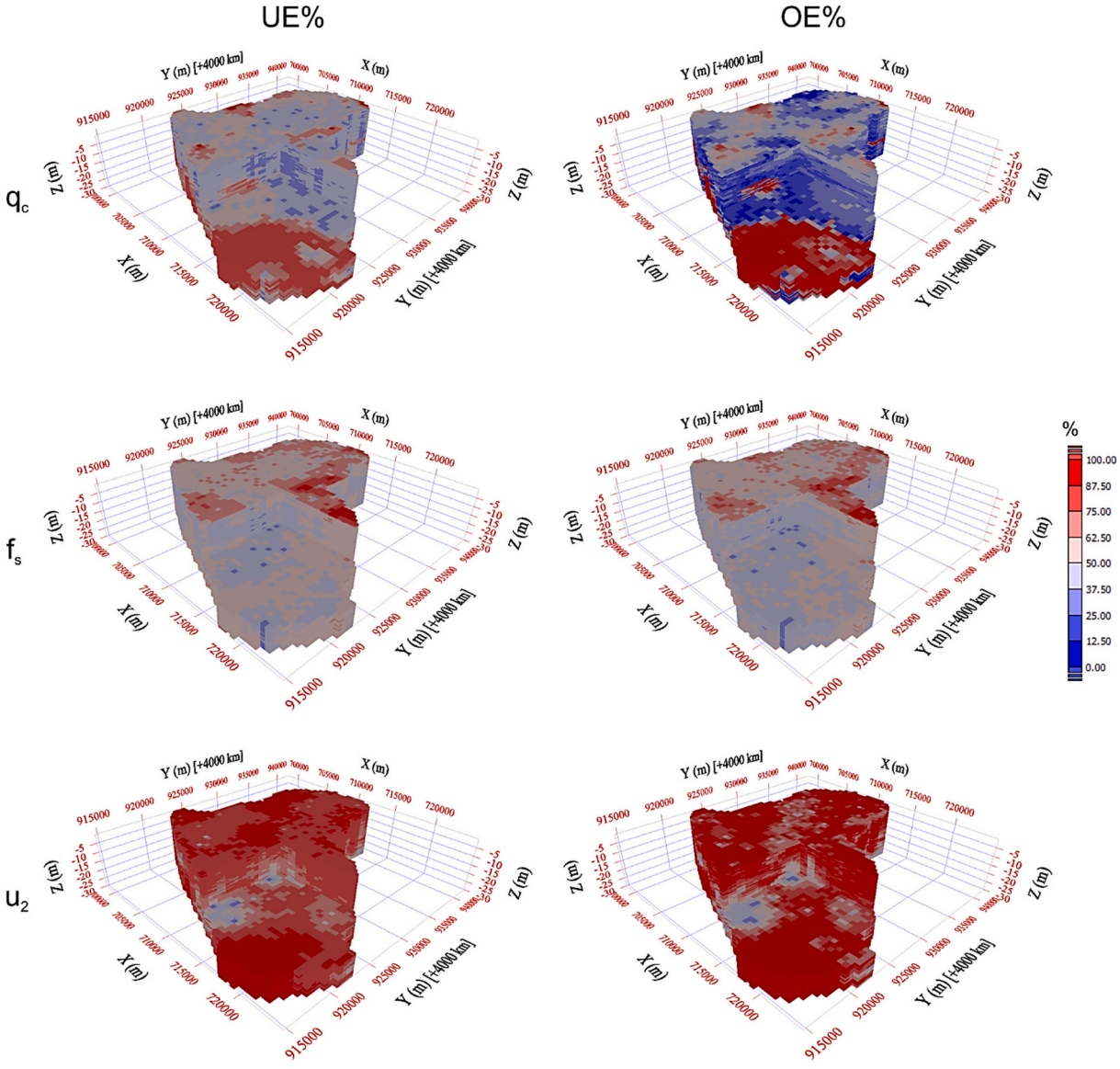


Fig. 4. Spatial distribution of the local uncertainty of the 1000 realizations of the input variables (i.e., q_c, f_s, and u_2), described by the Underestimation (UE%) and Overestimation (OE%) percentages and related to the 68 % confidence interval.

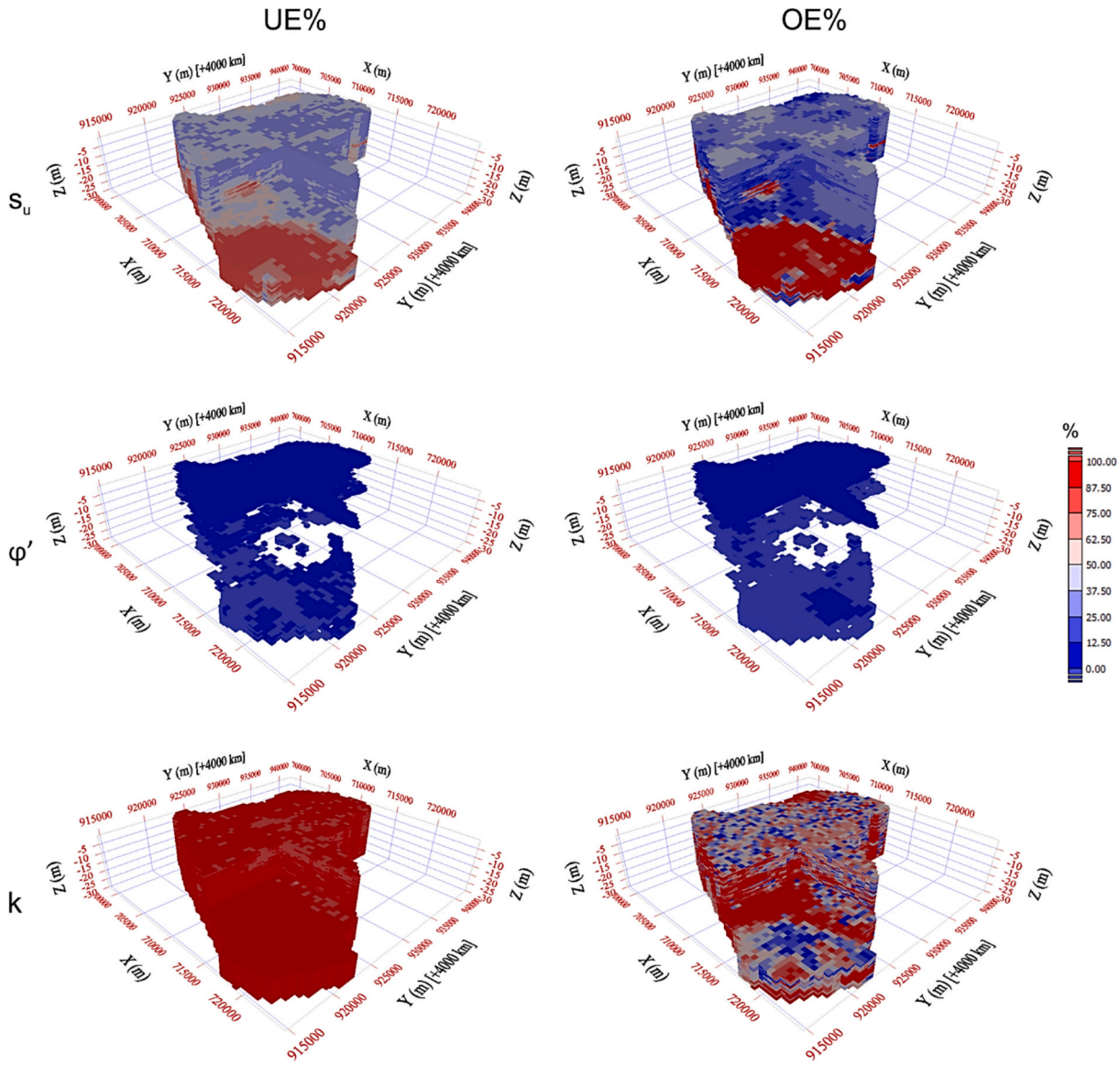


Fig. 5. Spatial distribution of the local uncertainty of the expected values of the design/target variables (i.e., s_u, φ', and k), described by the Underestimation (UE%) and Overestimation (OE%) percentages and related to the 68 % confidence interval.

deposits, displaying values above 62.5 % for both UE% and OE%. For f_s, UE% and OE% values are less variable and consistently in the range 37.5–62.5 %, regardless of the lithotype. This discrepancy in the spatial patterns of the uncertainty between q_c and f_s can be attributed to the fact that mixtures of coarse alluvial deposits (i.e. lithotypes SBT3 and SBT4) provide more variable cone tip resistance readings because of the local presence of larger grains of gravel or even boulders. This effect is less relevant for f_s because the readings are acquired through the lateral shaft, which provides more average measurements and is less sensitive to occasional pebbles. Instead, u_2 resulted the most uncertain i.e., UE% and OE% mostly above 75 %) because of the temporal variability of hydraulic condition on this input variable.

The accuracy associated with the predictions of the design/target variables (Fig. 5) results from the combination of the spatial variability of the input variables, and on the reliability of the considered transformation equations. It should be emphasised that in our approach we can only estimate the uncertainty of the design/target variables derived from the uncertainty of the input variables transformed by transformation equations. The linear dependence between s_u and q_t (i.e., a corrected version of q_c) introduced by Eq. (5) resulted in similar patterns of both UE% and OE% in lithotypes SBT4, 5 and 6, with values in the same range as q_c. Although q_t depends on both $q_c$and u_2, the variability of s_u is much more influenced by q_c due to the lower values of u_2. In the case of φ', values of both UE% and OE%, calculated for coarse deposits (i.e., SBT3 and SBT4), are less variable and always below 25 %. The friction angle is linked to q_c by a logarithmic transformation function (Eq. 6), which keeps the level of uncertainty very low. Instead, the hydraulic conductivity k shows a very high uncertainty, both in underestimation and overestimation. UE% values are higher than 75 % throughout the entire domain, while and OE% has relatively low values in the case of coarse deposits. In addition to the considerably high values of both UE% and OE%, the overall spatial distribution of the uncertainty of k is rather scattered. This behavior is attributable to the combination of functions used to obtain hydraulic conductivity values. First, the 1000 realizations of q_c, f_s, and u_2 are combined in Eq. (4) to calculate I_c realizations; then, they are used in the exponential equation Eq. (7) to obtain the 1000 realizations of k. In addition, the fact that OE% values are higher for fine deposits can be interpreted as an effect of the limited reliability of the function to estimate k for I_{SBTn} values larger than 3.27. However, even considering PU values relating to the 68 % confidence interval (Fig. 5),

the estimated s_u, φ', and k fall into the observed hydro-mechanical value ranges from the literature (Vessia and Di Curzio, 2018). This evidence is also confirmed for the very uncertain hydraulic conductivity values. In this case, UE% and OE% larger than 100 % will still keep the estimated values somewhat within the same order of magnitude, which is the highest precision possible with the traditional methods for field measurements (Elhakim, 2016).

3.3. Lithotype-dependent propagation uncertainty

As already pointed out in the previous section, PU can vary depending on the type of deposits. For this reason, we divided the simulation results according to the SBT class and calculated the Coefficient of Variation (CV) for both the specific lithotype and the 1000 realizations. In the first case, the CV values represent the inherent spatial variability within a specific lithotype; in the second case, CV also incorporates the uncertainty over the whole investigated volume. Associated with the estimates, which is the spatial variability for the input variables and a comparison of spatial variability and the propagated uncertainty for the design/target variables.

Fig. 6 and Table 4 compare the two CV values. Due to the propagated uncertainty, CV values for q_c, f_s, and u_2 increase by a percentage in the range 35–240 %, which depends on the volumetric size and/or inherent spatial variability of each lithotype in the considered portion of subsoil, and the actual number of measurements.

Since s_u and φ' are both mostly dependent on q_c, the CV values increased by the same percentages, even though the absolute CV values diverge by about one order of magnitude. In the case of SBT4, the increase of the realizations' CV of φ' was less relevant that the one observed for q_c. this evidence confirmed that the transformation equations used to calculate both s_u and φ' do not add PU. On the contrary, Eq. (7) tends to reduce the effect of the model uncertainty because of its logarithmic form, as observed for sandy mixtures. This means that the use of Eq. (6) and (7) provides reliable results for geotechnical designing purposes, as they do not contribute to the overall uncertainty with significant PU. Instead, the hydraulic conductivity k, whose calculation is achieved by a two-fold exponential transformation combining the uncertainty of all the three input variables, shows the largest difference between the two types of CV values. The CV increase is in the order of 700–800 % for coarse deposits and even higher for fine deposits (i.e., it increases by more than one order of magnitude). This difference between coarse and fine deposits is in line with what was already observed previously, confirming that the reliability of Eq. (8) is even lower when it comes to SBT5 and SBT6. However, k spatial variability is always very difficult to grasp, as the measurements are usually performed through pumping or slug tests. This means that measurements are either too averaged values, or are too local, or both (Elhakim, 2016) to provide a reliable spatial distribution of the subsoil hydraulic conductivity. Thus, although the use of Eq. (8) exposes to high level of PU, the vertical density of CPTu measurements can be beneficial to catch the

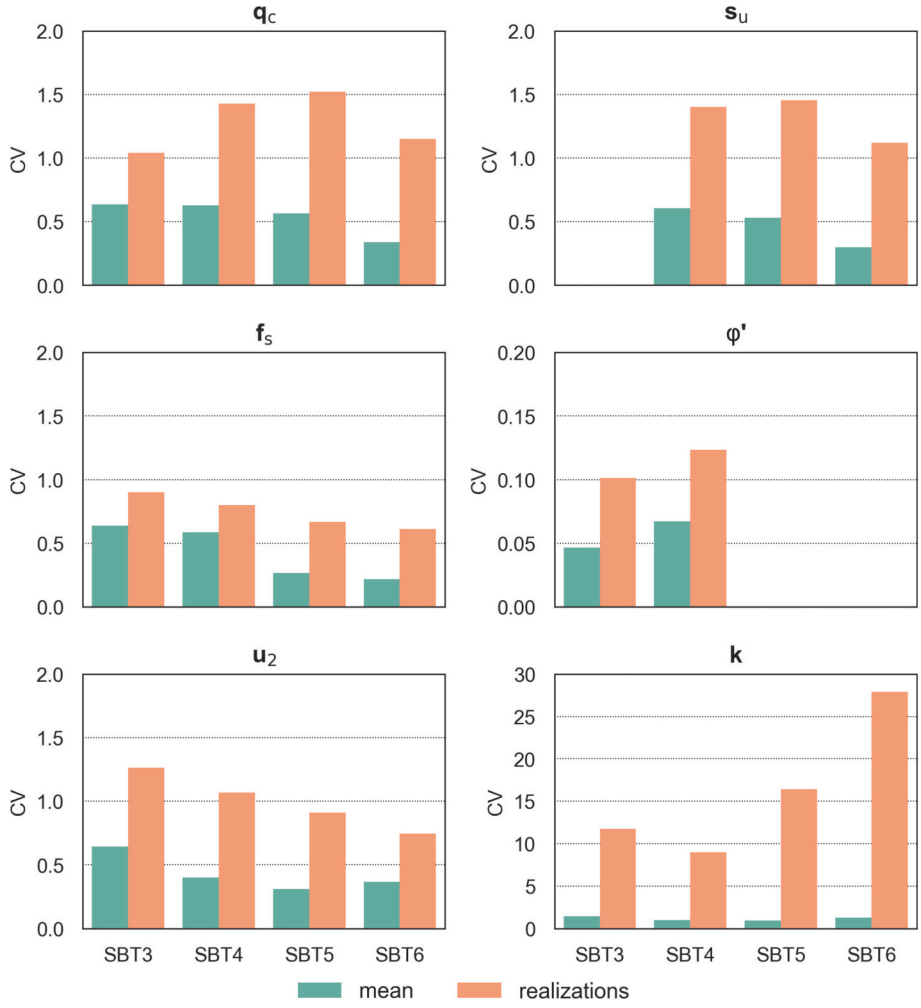


Fig. 6. Comparison between the inherent variability of the optimized 3D models (i.e., mean model) of both input and output variables and the total uncertainty (i.e., inherent variability, plus model and transformation uncertainty) related to the 1000 realizations, using the Coefficient of Variation (CV) as the selected reference statistical indicator.

Table 4
CV values describing the inherent variability of the optimized 3D models of both input and output variables and the total uncertainty related to the 1000 realizations.

	Mean						*Realizations*					
SBT class	q_c	f_s	u_2	s_u	ϕ'	k	q_c	f_s	u_2	s_u	ϕ'	k
3	0.64	0.64	0.65		0.05	1.43	1.04	0.90	1.26		0.10	11.73
4	0.63	0.59	0.40	0.61	0.07	1.01	1.43	0.80	1.07	1.40	0.12	8.99
5	0.56	0.27	0.31	0.53		0.94	1.52	0.67	0.91	1.46		16.39
6	0.34	0.22	0.37	0.30		1.30	1.15	0.61	0.75	1.12		27.89

heterogeneity of k spatial distribution, which for instance is known to be a critical factor in the reactive transport of contaminants in groundwater (Thouement and Van Breukelen, 2020; Wu et al., 2024) and the corresponding risk assessment (Høyer et al., 2019; Locatelli et al., 2019).

Finally, Figs. 7 and 8 show the comparison between the values of the indicators CV and CQV, which take into account just the inherent variability of the variables (i.e., CV mean and CQV mean) and the PU related to the 1000 realizations (i.e., CV realizations and CQV realizations), respectively. It is worth noticing that the mean values of the indicators do not greatly differ but for k values. However, PU increases CV values of realizations (Fig. 8), which are significantly larger than the CQV ones, especially for k variable. This means that the distributions of the derived variables are far from being symmetrical. Furthermore, CV values of realizations are greatly affected by outliers and tails of the distributions compared to the CQV ones. The clearest example is the hydraulic conductivity k, for which these indicators differ by more than one order of magnitude. This implies that dispersion indicators that are able to consider the effect of asymmetrical distribution and filter out the effect of outliers, such as CQV, should be preferred when investigating PU for design purposes or risk assessment. Moreover, the variability of PU among the SBT classes points out that equations can affect estimation reliability differently. Thus, it is recommended to assess PU with a lithotype/based approach.

4. Conclusions

The application of the Sequential Gaussian Co-Simulation method to CPTu data through fitting a Linear Model of Coregionalization enabled assessing the uncertainty propagation through empirical transformation equations, used to derive design/target variables (i.e., s_u, φ', and k) from the measurements q_c, f_s, and u_2.

The obtained results showed that the propagation uncertainty

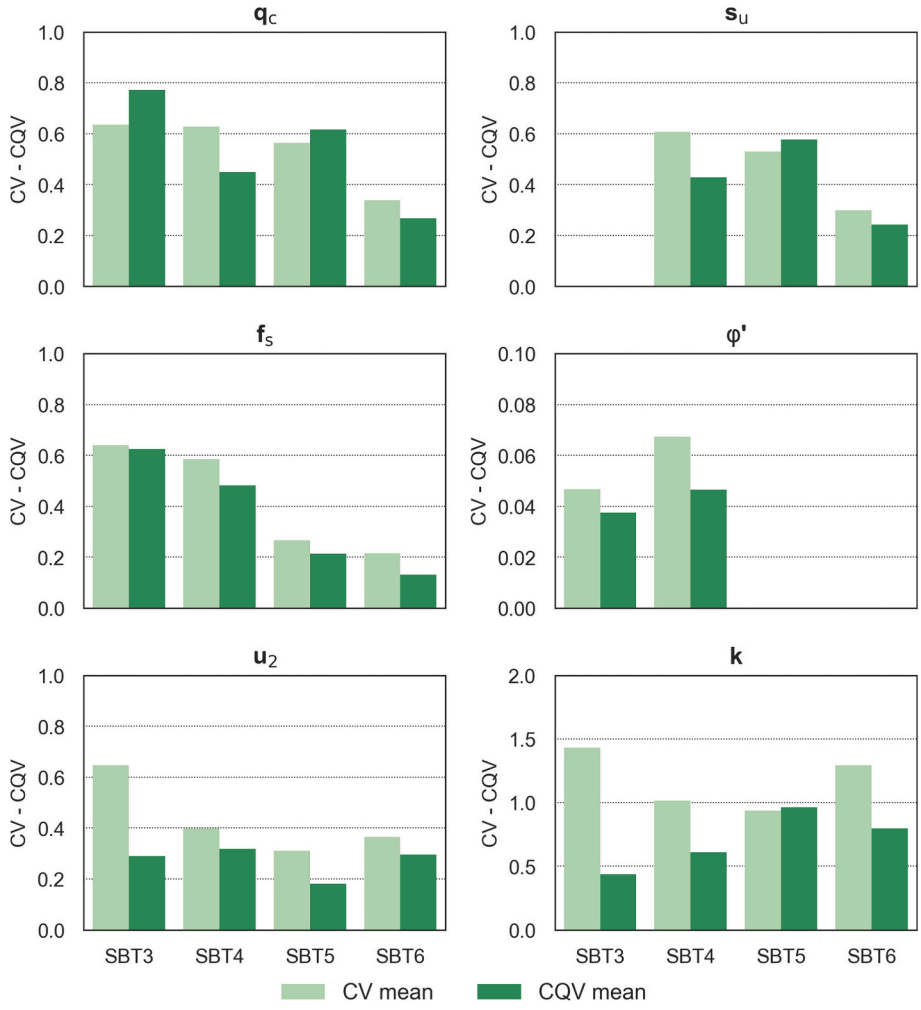


Fig. 7. Comparison between the inherent variability of the optimized 3D models (i.e., mean model) of both input and output variables using the Coefficient of Variation (CV) and the Coefficient of Quantile Variation (CQV) indicators.

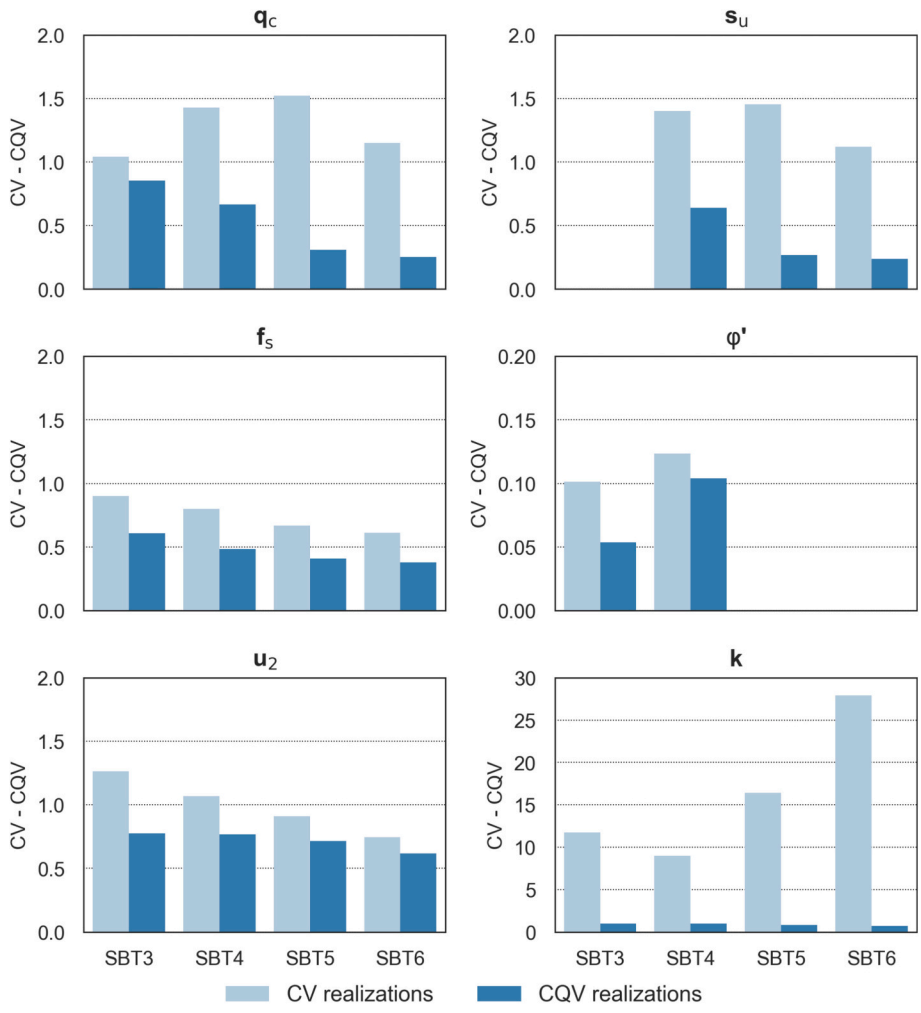


Fig. 8. Comparison between the total uncertainty (i.e., inherent variability, plus model and transformation uncertainty) related to the 1000 realizations related to the input and output variables, using the Coefficient of Variation (CV) and the Coefficient of Quantile Variation (CQV) indicators.

strongly depends on the spatial variability of the measured variables. Additionally, depending on the design/target variable of interest, some lithotypes are more prone to be affected by larger propagation uncertainty. Thus, the propagation uncertainty is affected by the heterogeneous spatial distribution of hydro-mechanical soil properties and might be taken into account through 3D spatial structure models.

Despite the significant level of uncertainty associated with s_u, φ', and k 3D models, the Sequential Gaussian Co-Simulation certainly proved to be a method for quantitatively assessing the reliability of Engineering Geological Models. It effectively described the spatial variability of hydro-mechanical subsoil properties, predicting subsoil geometries and values consistent with the literature. Thus, this model serves as a credible starting point for more detailed, smaller-scale investigations within the selected study area. Evidently, in such cases, denser measurement locations are needed to improve model accuracy.

As a result, the same geostatistical approach can be used to quantitatively assess the uncertainty propagation for other critical design/target variables and provide three-dimensional models of subsoil hydro-mechanical properties (i.e., Engineering Geological Models), which are generally obtained using empirical transformation equations applied to raw measured data. Therefore, further studies should investigate how the uncertainty of the measured variables used in transformation equations translates into the design/target variables' uncertainty, as this can significantly affect the reliability of outcomes.

Overall, this study underlines that the use of transformation equations to calculate relevant design/target variables might affect how uncertainty propagates and offers a reliable rationale to quantitatively assess this propagated uncertainty to enhance the reliability of risk assessment, hazard mapping, geotechnical design, and mining.

CRediT authorship contribution statement

Diego Di Curzio: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Annamaria Castrignanò:** Writing – review & editing, Methodology, Conceptualization. **Giovanna Vessia:** Writing – original draft, Supervision, Resources, Project administration, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research received financial support from the PRIN project "SLIDE – Stochastic Modeling of Compound Events" CUP: D53D23018920001 - P2022KZJTZ (PRIN2022PNRR_IP), granted by the Italian Ministry of University and Research to the University "G. d'Annunzio" of Chieti-Pescara, Italy. The authors extend their sincere

gratitude to the editor and the anonymous reviewers for their insightful comments and constructive feedback, which greatly contributed to improving the quality of this work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.enggeo.2025.108064.

Data availability

Data will be made available on request.

References

Angelini, M.E., Heuvelink, G.B.M., Lagacherie, P., 2023. A multivariate approach for mapping a soil quality index and its uncertainty in southern France. Eur. J. Soil Sci. 74 (2), e13345.

ASTM International, 2018. Standard guide for selection of simulation approaches in geostatistical site investigations. ASTM Standards. D5924-18.

Baynes, F.J., Parry, S., 2022. Guidelines for the development and application of engineering geological models on projects. International Association for Engineering Geology and the Environment (IAEG) Commission 25 Publication No. 1, 129 pp. https://www.iaeg.info/c25egmguidelines/.

Bonnet, D.G., 2006. Confidence interval for a coefficient of quartile variation. Comput. Stat. Data Anal. 50, 2953–2957.

Castrignanò, A., Buttafuoco, G., 2004. Geostatistical stochastic simulation of soil water content in a forested area of South Italy. Biosyst. Eng. 87 (2), 257–266.

Castrignanò, A., Buttafuoco, G., Canu, A., Zucca, C., Madrau, S., 2008. Modelling spatial uncertainty of soil erodibility factor using joint stochastic simulation. Land Degradation & Development 19 (2), 198–213.

Castrignanò, A., Landrum, C., De Benedetto, D., 2015. Delineation of Management zones in Precision Agriculture by Integration of Proximal Sensing with Multivariate Geostatistics. Exam. Sens. Data Fus. Agric. 80, 39–45.

Cherubini, C., Vessia, G., Pula, W., 2007. Statistical soil characterization of Italian sites for reliability analyses. In: Characterisation and Engineering Properties of Natural Soils. Taylor and Francis group, pp. 2681–2706.

Chilès, J.-P., Delfiner, P., 2012. Geostatistics: Modeling Spatial Uncertainty, 2nd ed. Wiley, Hoboken, NJ, USA.

Cressie, N., 2015. Statistics for Spatial Data. John Wiley & Sons.

Delbari, M., Afrasiab, P., Loiskandl, W., 2009. Using sequential Gaussian simulation to assess the field-scale spatial uncertainty of soil water content. Catena 79 (2), 163–169.

Di Curzio, D., Vessia, G., 2021. Multivariate geostatistical analysis of CPT readings for reliable 3D subsoil modeling of Heterogeneous Alluvial Deposits in Padania Plain. ISSMGE Int. J. Geoeng. Case Hist. 6 (4), 17–34.

Di Curzio, D., Rusi, S., Signanini, P., 2019. Advanced redox zonation of the San Pedro Sula alluvial aquifer (Honduras) using data fusion and multivariate geostatistics. Sci. Total Environ. 695.

Deutsch, C.V., Journel, A.G., 1998. GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press.

Di Curzio, D., Castrignanò, A., Fountas, S., Romić, M., Viscarra Rossel, R.A., 2021. Multi-source data fusion of big spatial-temporal data in soil, geo-engineering and environmental studies. Sci. Total Environ. 788, 147842.

Elhakim, A.F., 2016. Estimation of soil permeability. Alex. Eng. J. 55 (3), 2631–2638.

Emery, X., Peláez, M., 2011. Assessing the accuracy of sequential Gaussian simulation and co-simulation. Comput. Geosci. 15 (4), 673–689.

Erharter, G.H., Lacasse, S., Tschuchnigg, F., Tentschert, E., Becker, D., Phoon, K.K., 2024. A consistent terminology to communicate ground-related uncertainty. Eng. Geol. 342, 107744.

Gao, X., Yan, E.C., Yeh, T.C.J., Cai, J.S., Liang, Y., Wang, M., 2018. A geostatistical inverse approach to characterize the spatial distribution of deformability and shear strength of rock mass around an unlined rock cavern. Eng. Geol. 245, 106–119.

Geovariances, 2021. Isatis.neo 2021.07.1. www.geovariances.com/en/software/isatis-neo-geostatistics-software.

Goovaerts, P., 1996. Stochastic simulation of categorical variables using a classification algorithm and simulated annealing. Math. Geol. 28 (7), 909–921.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press.

Heuvelink, G.B.M., Webster, R., 2022. Spatial statistics and soil mapping: a blossoming partnership under pressure. Spat. Stat. 50, 100639.

Heuvelink, G.B.M., Webster, R., 2023. Uncertainty assessment of spatial soil information. In: Encyclopedia of Soils in the Environment, Second edition vols. 1-5, pp. V4-671–V4-683. Open Access Volume 4. 1 January 2023.

Høyer, A.S., Klint, K.E.S., Fiandaca, G., Maurya, P.K., Christiansen, A.V., Balbarini, N., Møller, I., 2019. Development of a high-resolution 3D geological model for landfill leachate risk assessment. Eng. Geol. 249, 45–59.

Hsu, Y.-H., Lu, Y.-C., Khoshnevisan, S., Juang, H., Hwang, J.-H., 2022. Influence of geological uncertainty on the design of OWTF monopiles. Eng. Geol. 303, 106621.

ISPRA, 2009. Carta Geologica d'Italia (scala 1:50000), Foglio 221 «Bologna», Servizio Geologico d'Italia. SystemCart s.r.l, Roma.

JORC, 2012. The JORC Code 2012 Edition, Joint Ore Reserves Committee of the Australasian Institute of Mining and Metallurgy. Australian Institute of Geoscientists and Minerals Council of Australia (JORC).

Kokkala, A., Marinos, V., 2022. An engineering geological database for managing, planning and protecting intelligent cities: the case of Thessaloniki city in Northern Greece. Eng. Geol. 301, 106617.

Lachérade, L., Marache, A., Denis, A., Halfon, I., Rohmer, J., Closset, L., Quesnel, F., 2023. A sectorisation-based method for geostatistical modeling of pressuremeter test data: Application to the Grand Paris Express project (France). Eng. Geol. 324, 107270.

Lee, S.H., Chen, W., 2009. A comparative study of uncertainty propagation methods for black-box-type problems. Struct. Multidiscip. Optim. 37, 239–253.

Lesny, K., Akbas, S.M., Bogusz, W., Burlon, S., Vessia, G., Phoon, K.K., Tang, C., Zhang, L., 2017. Evaluation and consideration of model uncertainties in reliability based design. In: Joint TC205/TC304 Working Group on "Discussion of Statistical/Reliability Methods for Eurocodes" – Final Report (Sep 2017), 20–64. International Society for Soil Mechanics and Geotechnical Engineering, London, UK.

Liang, Y., Xia, R., Yeh, T.C.J., Sun, Z., Zhang, H., Xu, B., 2024. Characterizing preferential infiltration of loess using geostatistical electrical resistivity tomography. Eng. Geol. 340, 107692.

Lilla Manzione, R., de Oliveira Ferreira Silva, C., Castrignanò, A., 2021. A combined geostatistical approach of data fusion and stochastic simulation for probabilistic assessment of shallow water table depth risk. Sci. Total Environ. 765, 142743.

Locatelli, L., Binning, P.J., Sanchez-Vila, X., Søndergaard, G.L., Rosenberg, L., Bjerg, P.L., 2019. A simple contaminant fate and transport modelling tool for management and risk assessment of groundwater pollution from contaminated sites. J. Contam. Hydrol. 221, 35–49.

Lunne, T., Powell, J.J., Robertson, P.K., 2002. Cone penetration testing in geotechnical practice. CRC press.

Matheron, G., 1962. Traité de géostatistique appliquée. Tome 1. Editions Technip, Paris, 334 pp.

McManus, S., Rahman, A., Coombes, J., Horta, A., 2021. Uncertainty assessment of spatial domain models in early stage mining projects – A review. Ore Geol. Rev. 133, 104098.

Nussbaumer, R., Mariethoz, G., Gloaguen, E., Holliger, K., 2018. Which path to choose in sequential Gaussian simulation. Math. Geosci. 50 (1), 97–120.

Phoon, K.K., 2020. The Goldilocks Dilemma—too little or too much data? GeoStrata Mag. Arch. 24 (1), 14–16.

Phoon, K.K., 2023. What geotechnical engineers want to know about reliability. ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A 9 (2), 03123001.

Phoon, K.K., Tang, C., 2019. Characterization of geotechnical model uncertainty. Georisk. https://doi.org/10.1080/17499518.2019.1585545.

Pownuk, A., 2018. Combining Interval, Probabilistic, and Other Types of Uncertainty in Engineering Applications. Springer International Publishing, Cham.

Rahman, A., Gao, J., D'Este, C., Ahmed, S.E., 2016. An assessment of the effects of prior distributions on the Bayesian predictive inference. Int. J. Stat. Prob. 5, 31.

Robertson, P.K., 1990. Soil classification using the cone penetration test. Can. Geotech. J. 27, 151–158.

Robertson, P.K., 2009. CPT interpretation – a unified approach. Can. Geotech. J. 46, 1–19.

Robertson, P.K., Cabal, K.L., 2014. Guide to Cone Penetration Testing for Geotechnical Engineering. Gregg Drilling & Testing, Inc, Signal Hill, California.

Robertson, P.K., Campanella, R.G., 1983. Interpretation of Cone penetration tests part I (sand) part II (clay). Can. Geotech. J. 20 (4), 719–745.

Savoy, H., Heße, F., Rubin, Y., 2017. anchoredDistr: A package for the Bayesian inversion of geostatistical parameters with multi-type and multi-scale data. R J. 9, 6–17.

Shao, W., Chen, S., Li, M., Su, Y., Ni, J., Dong, J., Zhang, Y., Yang, Z., 2023. Reducing uncertainties in hydromechanical modeling with a recently developed Rosetta 3 podeotransfer function. Eng. Geol. 324, 107250.

Szatmári, G., Pásztor, L., Heuvelink, G.B., 2021. Estimating soil organic carbon stock change at multiple scales using machine learning and multivariate geostatistics. Geoderma 403, 115356.

Thouement, H.A., Van Breukelen, B.M., 2020. Virtual experiments to assess opportunities and pitfalls of CSIA in physical-chemical heterogeneous aquifers. J. Contam. Hydrol. 231, 103638.

Vessia, G., Di Curzio, D., 2018. Lacustrine deposits. In: Bobrowsky, P.T., Marker, B. (Eds.), Encyclopedia of Engineering Geology (Encyclopedia of Earth Sciences Series), vol. PartF1. Springer.

Vessia, G., Di Curzio, D., Chiaudani, A., Rusi, S., 2020a. Regional rainfall threshold maps drawn through multivariate geostatistical techniques for shallow landslide hazard zonation. Sci. Total Environ. 70525, 135815.

Vessia, G., Di Curzio, D., Castrignanò, A., 2020b. Modeling 3D soil lithotypes variability through geostatistical data fusion of CPT parameters. Sci. Total Environ. 698.

Wackernagel, H., 2003. Multivariate Geostatistics: An Introduction with Applications. Springer-Verlag, Berlin.

Wang, Y., Cao, Z., Li, D., 2016. Bayesian perspective on geotechnical variability and site characterization. Eng. Geol. 203, 117–125.

Wang, Y., Hu, Y., Phoon, K.K., 2022. Non-parametric modelling and simulation of spatiotemporally varying geo-data. Georisk 16 (1), 77–97.

Webster, R., Oliver, M.A., 2007. Geostatistics for Environmental Scientists. John Wiley & Sons, New York.

Wei, X., Wang, H., 2022. Stochastic stratigraphic modeling using Bayesian machine learning. Eng. Geol. 307, 106789.

Wu, C.L.R., Wagterveld, R.M., van Breukelen, B.M., 2024. Reactive transport modeling for exploring the potential of water quality sensors to estimate hydrocarbon levels in groundwater. Water Resour. Res. 60 (4), e2023WR036644.

Zhan, W., Baise, L.G., Moaveni, B., 2023. An uncertainty quantification framework for logistic regression based geospatial natural hazard modeling. Eng. Geol. 324, 107271.
Zhang, Z., Wang, G., Carranza, E.J.M., Liu, C., Li, L., Fu, C., Liu, X., Chen, C., Fan, J., Dong, Y., 2023. An integrated machine learning framework with uncertainty quantification for three-dimensional lithological modeling from multi-source geophysical data and drilling data. Eng. Geol. 324, 107255.
Zhao, Z., Duan, W., Cai, G., Wu, M., Liu, S., 2022. CPT-based fully probabilistic seismic liquefaction potential assessment to reduce uncertainty: Integrating XGBoost algorithm with Bayesian theorem. Comput. Geotech. 149, 104868.
Zhao, C., Gong, W., Juang, C.H., Tang, H., Hu, X., Wang, L., 2023. Optimization of site exploration program based on coupled characterization of stratigraphic and geo-properties uncertainties. Eng. Geol. 317, 107081.
Zuccarini, A., Giacomelli, S., Severi, P., Berti, M., 2024. Long-term spatiotemporal evolution of land subsidence in the urban area of Bologna, Italy. Bull. Eng. Geol. Environ. 83 (1), 35.