Calculating Reliable Engineering Geological Model through Stochastic Co-Simulation Applied to CPTu Data

Di Curzio, Diego; Vessia, Giovanna

DOI
10.1061/9780784484999.027

Publication date
2023

Document Version
Final published version

Published in
Geotechnical Special Publication

Citation (APA)
Di Curzio, D., & Vessia, G. (2023). Calculating Reliable Engineering Geological Model through Stochastic Co-Simulation Applied to CPTu Data. In J. Ching, S. Najjar, & L. Wang (Eds.), *Geotechnical Special Publication* (GSP 347 ed., pp. 259-267). (Geotechnical Special Publication; Vol. 2023-July, No. GSP 347). American Society of Civil Engineers (ASCE). https://doi.org/10.1061/9780784484999.027

Calculating Reliable Engineering Geological Model through Stochastic Co-Simulation Applied to CPTu Data

Diego Di Curzio, Ph.D.[1]; and Giovanna Vessia, Ph.D.[2]

[1]Marie Skłodowska-Curie Postdoctoral Fellow, Dept. of Water Management, Delft Univ. of Technology, Delft, Netherlands. ORCID: https://orcid.org/0000-0002-2465-1014. Email: d.dicurzio@tudelft.nl
[2]Adjunct Professor, Dept. of Engineering and Geology, Univ. "G. d'Annunzio" of Chieti–Pescara (corresponding author). ORCID: https://orcid.org/0000-0003-1733-7112. Email: giovanna.vessia@unich.it

ABSTRACT

To perform a geotechnical reliable design, the spatial variability and the uncertainties related to the adopted engineering geological model (EGM) must be taken into account. However, any conceived EGM is characterized by uncertainties covering (1) the bias in the mathematical expression that transforms the measured parameters into design ones; and (2) the uncertainty associated with the variability of the soil and rock parameters in the prediction equations. Hereinafter, the sequential Gaussian co-simulation method (SGCS) has been applied to propagate the uncertainty in the calculation of the undrained shear resistance s_u from measured CPTu profiles (i.e., q_c, f_s, u_2) through a linear model of co-regionalization. The studied area is located in the Po River alluvial plain (Bologna Province, Italy), where the mixture of silts, sands, and clays gets thicknesses of hundreds of meters. These heterogeneous deposits have been mechanically characterized through a 3D EGM to be used in reliability-based designing.

INTRODUCTION

Any Engineering Geological Model, EGM, used to design according to reliability-based design rules, must take into account the main sources of uncertainties. The need of assessing EGM spatial variability and uncertainties has been confirmed in the current revision process of the Eurocodes (Lesny et al., 2017).

In particular, dealing with uncertainties, the following two sources can be pointed out: (1) the bias in the mathematical expression that transforms the measured parameters into design ones and (2) the uncertainty associated with the variability of the soil and rock parameters in the prediction equations. Hence, to quantify the previous two sources of uncertainties, Lee and Chen (2009) listed five methods to carry out the uncertainty propagation (UP) from the measured to the design variables: 1) the simulation; 2) the local expansion; 3) the most probable point; 4) the functional expansion; 5) the numerical integration.

In this study, the first method has been used through the Sequential Gaussian Co-Simulation method (SGCS). The uncertainty in the calculation of the undrained shear strength s_u from the measured CPTu profiles (i.e., q_c, f_s, u_2) has been assessed by means of the Linear Model of Co-regionalization. This latter was fitted to the matrix of experimental direct and cross-variograms of the input data, and one thousand realizations that honor the experimental data were provided.

These realizations enabled to calculate the distributions of s_u values. Thus, the mean and the standard deviation maps of s_u can be drawn. The site where the present study has been applied is located in the East part of the Bologna district (Po plain, Italy).

THE CASE STUDY OF THE PO RIVER ALLUVIAL PLAIN

The selected study area (Figure 1) is a 900-square-kilometer-wide portion of the eastern Bologna district, located in the southern part of the Po plain (Italy). This latter is a tectonic depression, filled by hundreds-of-meter-thick continental and/or marine-transitional deposits (Amorosi and Farina 1995).

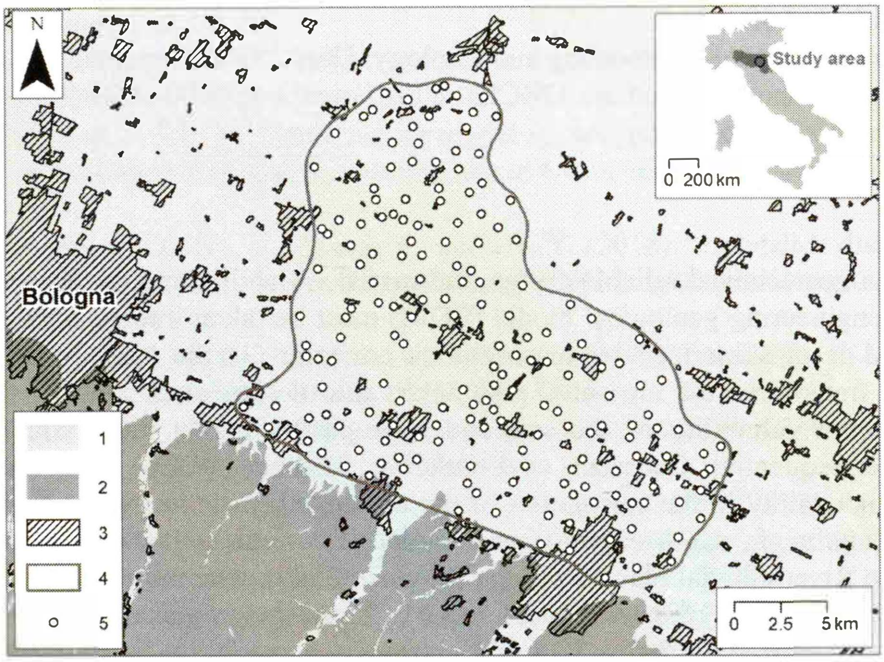

Figure 1. Map showing the location and main geological features of the study area as well as the CPTus' distribution within the selected domain. In legend: 1) alluvial deposits; 2) bedrock; 3) urban areas; 4) geostatistical domain; 5) CPTus' locations.

Within the study area, there are alluvial deposits made up 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 2009a, 2009b). From a lithological point of view, these inclusions consist of sandy, gravelly, and silty-clayey soils.

On one hand, sandy-gravelly alluvial fans are prevalent nearby the Apennine reliefs, in the south; on the other hand, sandy paleo-channels become predominant moving northward. Conversely, fine lacustrine lenses can be found all over the study area. It is worth pointing out that all the types of inclusion have shape, size, and depth that can be predicted through direct investigations, within the whole subsoil volume.

Dataset

The dataset used in this research consists of 182 CPTus performed across a 900-square-kilometer-wide area and 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), subsequently made available by Di Curzio and Vessia (2021).

It is worth pointing out that the distances among CPTu profiles vary between 500 m and 2 km. Although these horizontal distances might not allow recognizing of the typical horizontal scale of fluctuation within a single geological body, this latter is not the objective of the Geostatistical methods because they are data-driven techniques. Hence, the spatial variability structure of the whole subsoil model, instead, is here recognized through the experimental variogram.

THE SEQUENTIAL GAUSSIAN CO-SIMULATION

To get a deeper insight into the propagation of uncertainty while using transformation equations, the Sequential Gaussian Co-Simulation (SGCS) method has been selected (Gooverts, 1997; Webster and Oliver, 2007; Chilès and Delfiner, 2012).

It was derived from the general Sequential Gaussian Simulation approach (SGS). This advanced geostatistical technique is one of the most straightforward and used among Conditional Simulation methods (Delbari et al., 2009; Emery and Peláez, 2011; Nussbaumer et al., 2018), which are the simulation approaches that honor measured data. Unlike kriging methods, stochastic simulation techniques are devoted to assessing spatial uncertainty (Castrignanò and Buttafuoco, 2004).

Furthermore, since these methods can preserve the spatial variability, which is instead smoothed in kriging methods, stochastic simulation approaches can be also used to obtain optimized maps of estimates (ASTM International, 2018).

SGS is based on the multi-gaussian assumption, that is the conditionally simulated values ($\mathrm{z}_{\mathrm{c}}^{*}(\mathbf{x_0})$) at each node of the grid is obtained conditioning results with the Kriging estimator ($\mathrm{z}^{*}(\mathbf{x_0})$), as follows:

$$\mathrm{z}_{\mathrm{c}}^{*}(\mathbf{x_0}) = \mathrm{z}^{*}(\mathbf{x_0}) + [\mathrm{s}^{*}(\mathbf{x_0}) - \mathrm{z}_{\mathrm{s}}^{*}(\mathbf{x_0})] \tag{1}$$

where, $\mathrm{s}^{*}(\mathbf{x_0})$ is the simulated field calculated with the same variogram model as that of experimental data, while $\mathrm{z}_{\mathrm{s}}^{*}(\mathbf{x_0})$ is the kriging estimates obtained by considering the simulated values at the sampling points. In SGS, this process is repeated several times by random seeds, which correspond to different paths through the data. As a result, several equiprobable representations of the spatial distribution of the considered variable can be obtained, namely, realizations, providing a statistical distribution for each node of the grid, instead of an estimated value and the corresponding error (i.e., as in kriging methods).

As in this work we dealt with a multivariate case, the Kriging estimator in Eq. 1 was replaced with the Co-Kriging one. In fact, the method name moves from SGS to SGCS. In this case, the simulation relies on the fitting of a Linear Model of Coregionalization (LMC) of the considered variables (i.e., qc, fs and u2), below represented in matrix notation (Wackernagel, 2003; Castrignanò et al., 2015; Di Curzio et al., 2019, Vessia et al., 2020a):

$$\boldsymbol{\Gamma}(\mathbf{h}) = \sum_{\mathrm{u}=1}^{\mathrm{N_S}} \mathbf{B}^{\mathrm{u}} \mathrm{g}^{\mathrm{u}}(\mathbf{h}) \tag{2}$$

where, $\mathbf{B}^{\mathrm{u}}$ is the Coregionalization matrix of the LMC coefficients, which is symmetric and positive, $\boldsymbol{\Gamma}(\mathbf{h})$ is the n x n matrix with direct variograms (i.e., diagonal elements) and crossvariogram (i.e., non-diagonal elements) modeled as a linear combination of N_S basic variogram functions, $\mathbf{h}$ is the lag, $g^u(\mathbf{h})$ is the spatial structure, and u is the spatial scale.

It is worth pointing out that, since a Gaussian distribution is required, non-Gaussian data transformation is needed. In this study, raw measurements of q_c, f_s, and u_2 have been transformed through the Gaussian Anamorphosis function (Chilès and Delfiner, 2012). This function converts a Gaussian-distributed variable (Y) into a new variable ($Z = \Phi(Y)$), whatever its statistical distribution, by fitting a Hermite polynomial expansion Hi(Y):

$$\Phi(Y) = \sum \Psi_i H_i(Y) \tag{3}$$

where, Ψ_i are the coefficients of the Hermite polynomials. Once the Gaussian Anamorphosis function is defined, the transformation from a non-Gaussian variable into a Gaussian standardized one (i.e., the one required to use SGCS) is performed by inverting the function, as follows:

$$Y = \Phi^{-1}(Z) \tag{4}$$

The selected estimation domain has a cell size equal to 500 m x 500 m x 0.5 m. All the geostatistical analyses have been performed using Isatis 2018, whose results have then been visualized through Isatis.neo (www.geovariances.com/en/software/isatis-neo-geostatistics-software/).

UNDRAINED SHEAR RESISTANCE FROM CPTU PROFILES

The commonest expression that is used to calculate the undrained shear resistance s_u from CPTs follows (Lunne et al., 1997):

$$s_u = \frac{q_t - \sigma_v}{N_{kt}} \tag{5}$$

where q_t is the total cone tip resistance, σ_v the vertical stress at each depth and N_{kt} is the cone factor varying from 10 to 18 with 14 as an average.

It is worth pointing out that Eq. 5 is valid for fine grained materials; thus, uncertainty propagation for s_u has been investigated only where the soil behaviour type index I_{SBTn} values are greater than 2.6 (Figure 2 and 3) (Robertson, 2009).

RESULTS AND DISCUSSION

1000 three-dimensional realizations of raw measurements have been obtained through SGCS considering the directional LMC (anisotropic) described in detail in Table 1, which is a combination of scale-dependent variabilities: one isotropic structure on the horizontal plane and one vertical structure. The large horizontal gap among CPTu profiles does not allow us to infer the ranges at a smaller size than 500 m. In Table 1, the nested variogram consisting of double spherical functions shows that at distances larger than 500 m there is still a spatial correlation among data.

Conversely, in the vertical direction, the accurate profile description through the CPTu measurements enables us to describe accurately the vertical variability structure through four nested variogram functions. The k-Bessel function simulates the vertical trend observed along all the CPTu profiles. It is assumed that the stationarity is reached at a vertical scale larger than the model dimension.

Table 1. Features of the Linear Model of Coregionalization related to the Gaussian transformed variables

Variables	Horizontal LMC structures	Range (m)
gq_c, gf_s, gu_2	Spherical Spherical	1200 12000
Variables	**Vertical LMC structures**	**Range (m)**
gq_c, gf_s, gu_2	Spherical Spherical Spherical k-Bessel	2 6 12 > 100

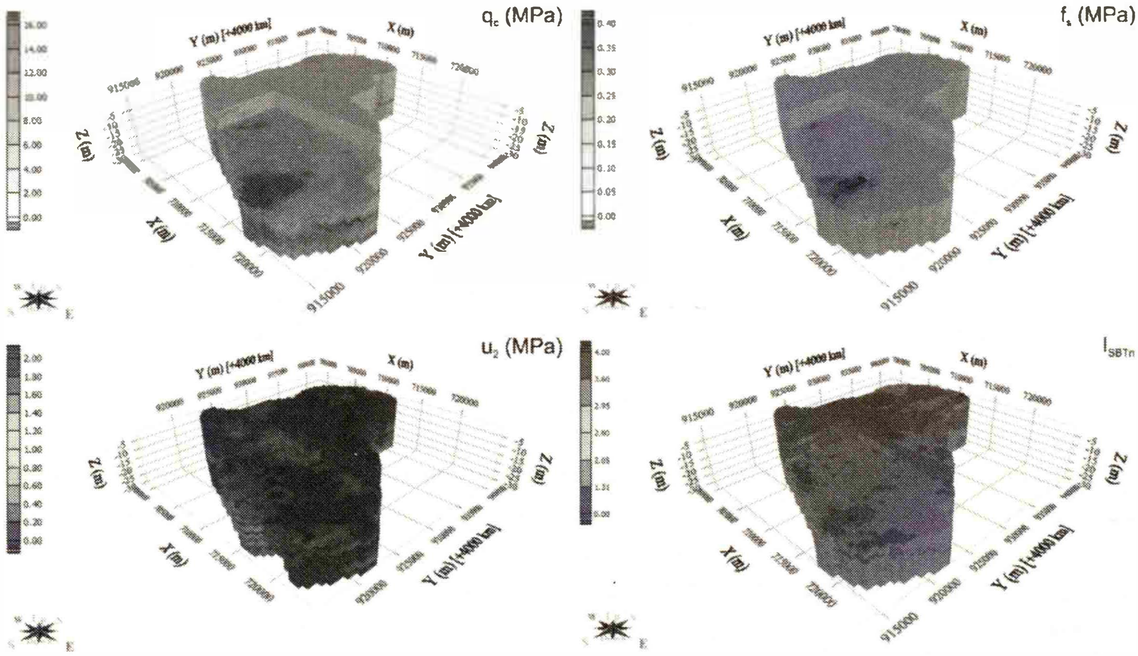

Figure 2. Optimized 3D models of q_c, f_s, u_2, and I_{SBTn}, represented by the mean values of the 1000 realizations.

In turn, these realizations have been used as input variables of Eq. 5, providing the same number of realizations of the selected output variable (i.e., s_u).

Figures 3 and 4 show the optimized 3D models (i.e., mean values of 1000 realizations) of q_c, f_s, and u_2 (including I_{SBTn}, used as a filtering variable in Figure 4), and s_u, respectively. It appears clear that the considered subsoil is mainly characterized by fine-grained alluvial deposits, with coarser sandy gravelly inclusions (i.e., buried alluvial fans, and paleo-channels). As a matter of fact, q_c, f_s, and u_2 values are generally very low for fine lithotypes, while increase considerably in the case of coarse deposits.

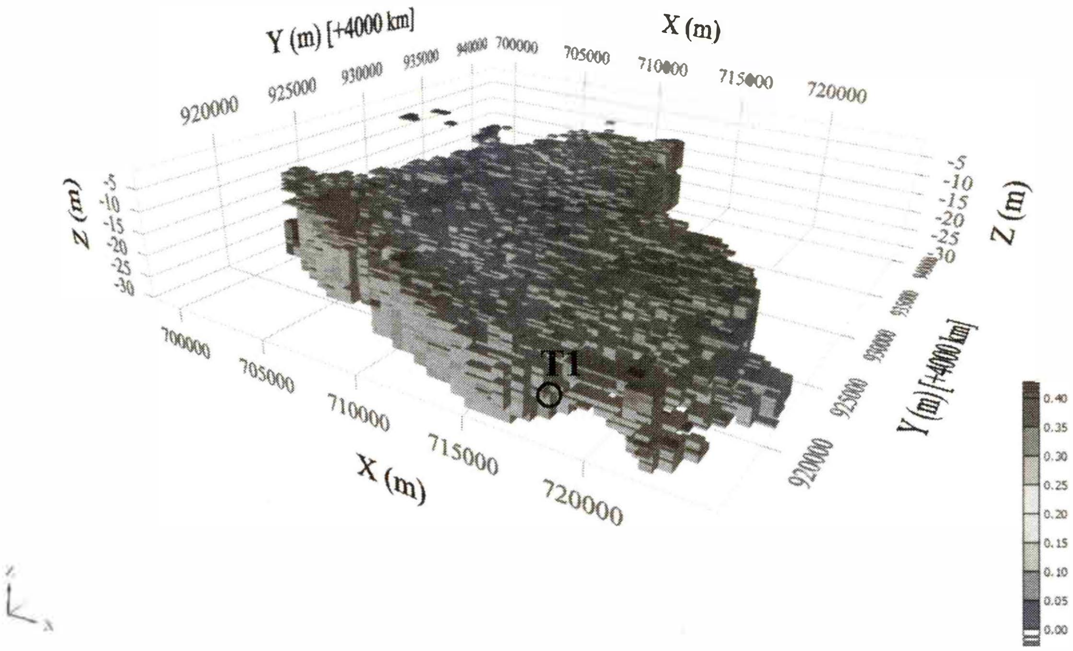

Figure 3. Optimized 3D models of s_u related to I_{SBTn} values greater than 2.60 (i.e., fine deposits). The black circle indicates the test cell taken into account for uncertainty propagation (Figure 4 and Table 2).

Within lithological classes described by I_{SBTn} values larger than 2.60, s_u varies considerably as well, probably because of both the different fine deposits present in the study area and the confining stress.

The histograms in Figure 4 as well as the descriptive statistics in Table 2 depict the propagation of uncertainty when calculating 1000 realizations of s_u from the same number of realizations of input variables (i.e., q_c, f_s, and u_2) obtained through the SGCS. Even though all the statistical distributions are characterized by very large outliers, the values of the majority of the realizations are slightly variable, both for the input and the output variables.

This evidence suggests that the uncertainty quantified for q_c, f_s, and u_2 does not propagate generating a larger variability of s_u, perhaps due to the use of a multivariate variability model (i.e., LMC), which allows reducing the uncertainty in the estimation process.

Table 2. Descriptive statistics of equiprobable values of input variables (i.e., q_c, f_s, and u_2) with the output ones (i.e., s_u), at the test cell T1, shown in Figure 3.

Variable	Mean	Median	Min	Max	I quartile	III quartile	St.D.	Skewness	Kurtosis
q_c	2.433	1.945	0.540	22.490	1.560	2.400	2.286	5.111	30.962
f_s	0.097	0.090	0.010	0.460	0.060	0.120	0.049	1.393	4.032
u_2	0.187	0.120	0.000	0.940	0.060	0.270	0.171	1.438	1.848
s_u	0.155	0.120	0.030	1.580	0.090	0.150	0.162	5.134	31.200

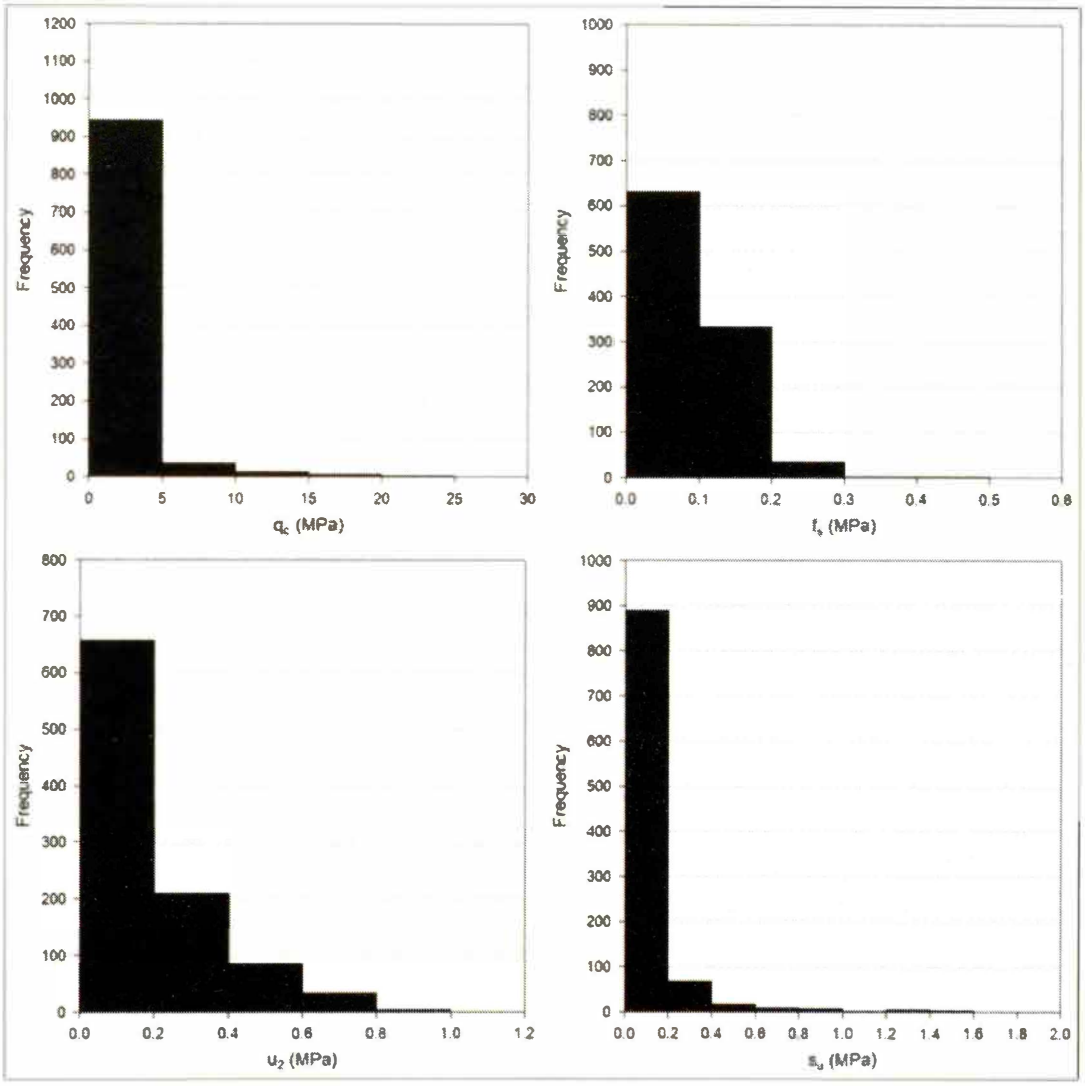

Figure 4. Histograms comparing the statistical distribution of equiprobable values of input variables (i.e., q_c, f_s, and u_2) with the output ones (i.e., s_u), at the test cell, shown in Figure 3.

CONCLUSION

In the present study, the uncertainty propagation has been accomplished starting from CPTu profiles of data (mainly q_c) to derive the design variable s_u by Eq. 5. To this end, the Sequential Gaussian Co-Simulation has been applied by fitting a Linear Model of Coregionalization. In detail, 1000 realizations of q_c, f_s, and u_2, obtained through SGCS, have been used to calculate the same number of realizations of s_u, providing an optimized 3D model of lithotypes' distribution as well as a quantification of the uncertainty associated with the transformation expression.

The same methodological approach can be used to quantitatively assess the uncertainty propagation of other crucial design variables, which are generally obtained using empirical equations and raw measured data. This study shows how to calculate a 3D data model with its uncertainty of a large subsoil volume to be used as a support of the infrastructure designing and urban planning development.

REFERENCES

Amorosi, A., and Farina, M. (1995). "Large scale architecture of a thrust-related alluvial complex from subsurface data: the Quaternary succession of the Po basin in the Bologna area (Northern Italy)." *Giorn. Geol.*, 57(1-2), 3-16.

ASTM. (2018). *Standard Guide for Selection of Simulation Approaches in Geostatistical Site Investigations*, ASTM Standards, D5924-18.

Castrignano, A., and Buttafuoco, G. (2004). "Geostatistical stochastic simulation of soil water content in a forested area of south Italy." *Biosystems Engineering*, 87(2), 257-266.

Castrignanò, A., Landrum, C., and De Benedetto, D. (2015). "Delineation of Management Zones in Precision Agriculture by Integration of Proximal Sensing with Multivariate Geostatistics." *Examples of Sensor Data Fusion Agric.*, 80, 39-45.

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

Delbari, M., Afrasiab, P., and 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., Rusi, S., and Signanini, P. (2019). "Advanced redox zonation of the San Pedro Sula alluvial aquifer (Honduras) using data fusion and multivariate geostatistics." *Science of the Total Environment*, 695.

Di Curzio, D., and Vessia, G. (2021). 'Multivariate Geostatistical Analysis of CPT Readings for Reliable 3D Subsoil Modeling of Heterogeneous Alluvial Deposits in Padania Plain." *ISSMGE International Journal of Geoengineering Case Histories*, 6(4), 17-34.

Emery, X., and Peláez, M. (2011). "Assessing the accuracy of sequential Gaussian simulation and co-simulation." *Computational Geosciences*, 15(4), 673-689.

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

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

ISPRA. (2009b). *Carta Geologica d'Italia (scala 1:50000), Foglio 239 «Faenza», Servizio Geologico d'Italia*, SystemCart s.r.l, Roma.

Lee, S. H., and Chen, W. (2009). "A comparative study of uncertainty propagation methods for black-box-type problems." *Structural Multidisciplinary Optimum*, 37, 239–253.

Lesny, K., Akbas, S. M., Bogusz, W., Burlon, S., Vessia, G., Phoon, K. K., Tang, C., and Zhang, L. (2017). Evaluation and Consideration of Model Uncertainties in Reliability Based Design. Joint TC205/TC304 Working Group on "Discussion of statistical/reliability methods for Eurocodes" – Final Report (Sep 2017), 20–64. London, UK: International Society for Soil Mechanics and Geotechnical Engineering.

Lunne, T., Robertson, P. K., and Powell, J. J. M. (1997). *Cone penetration testing in geotechnical practice*, E & FN Spon Routledge, 352 p, ISBN 0-7514-0393-8.

Nussbaumer, R., Mariethoz, G., Gloaguen, E., and Holliger, K. (2018). "Which path to choose in sequential Gaussian simulation." *Mathematical Geosciences*, 50(1), 97-120.

Phoon, K. K., and Tang, C. (2019). "Characterization of geotechnical model uncertainty." *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*.

Robertson, P. K., and Wride, C. E. (1998). "Evaluating cyclic liquefaction potential using the Cone Penetration Test." *Canadian Geotechnical Journal*, 35 (3), 442–459.

Robertson, P. K. (2009). "CPT interpretation – a unified approach." *Canadian Geotechnical Journal*, 46, 1-19.

Vessia, G., Di Curzio, D., Chiaudani, A., and Rusi, S. (2020a). "Regional rainfall threshold maps drawn through multivariate geostatistical techniques for shallow landslide hazard zonation." *Science of the Total Environment*, 70525, 135815.
Vessia, G., Di Curzio, D., and Castrignanò, A. (2020b). "Modeling 3D soil lithotypes variability through geostatistical data fusion of CPT parameters." *Science of the Total Environment*, 698.
Wackernagel, H. (2003). *Multivariate Geostatistics: An Introduction with Applications*, Springer-Verlag, Berlin.
Webster, R., and Oliver, M. A. (2007). *Geostatistics for Environmental Scientists*, John Wiley & Sons, New York.