Proceedings of the 21st ICSMGE, Vienna, Austria, 14 – 19 June 2026. Pistrol, Adam & Schweiger (eds.)
Published by: ÖGG, Austrian Society for Geomechanics, Salzburg, Austria, ISBN 978-3-9503898-4-5
https://doi.org/10.53243/ICSMGE2026-756

Modelling non-stationary subsoil spatial variability using CPT data

Joanna Pieczyńska-Kozłowska
Faculty of Civil Engineering, Wrocław University of Science and Technology, Poland,
joanna.pieczynska-kozlowska@pwr.edu.pl

Giovanna Vessia
Department of Engineering and Geology, University "G. d'Annunzio" of Chieti-Pescara, Italy

ABSTRACT: Engineering Geological Models (EGMs) typically rely on deterministic assumptions, which can lead to overdesign and underestimation of uncertainty. To address these limitations, Bayesian Compressive Sampling (BCS), combined with the Discrete Cosine Transform (DCT), is introduced to model the non-stationary spatial variability of subsoil properties using Cone Penetration Test (CPT) data from the Po River Alluvial Plain, Italy. This novel approach is data-driven, enables explicit uncertainty quantification, adapts to sparse or irregular data, and supports a reliable design. A case study using six CPT profiles – four for modelling and two for validation – demonstrates the method's effectiveness in capturing spatial variability and improving reliability in geotechnical design. Unlike deterministic models that overestimate soil strength and ignore depth-dependent variability, the BCS-DCT framework produces stable, data-consistent random fields (RFs) even with limited measurements, thereby aligning with geological heterogeneity. The study highlights the potential of nonparametric RFs to enhance subsoil characterisation and to integrate uncertainty-aware modelling into engineering practices.

KEYWORDS: Non-Stationary Spatial Variability; CPT; Random Field Theory; EGM; Soil Strength Parameters.

1 INTRODUCTION

The concept of the Engineering Geological Model (EGM), introduced in the 2nd generation of Eurocode 7 (CEN 2024a, CEN 2024b), is defined as a collection of site features of the ground and groundwater, based on results from ground investigations and other available data. In most engineering designs, EGMs are based on and constrained by deterministic data, leading to overdesign and increased costs (Qian and Shi, 2025). Di Curzio et al. (2025) emphasise that EGMs are affected by multiple sources of uncertainty: spatial variability of soil properties, measurement errors, and the model uncertainty concerning the transformation equations used to derive design variables (e.g., undrained shear strength, s_u, internal friction angle, ϕ') from field measurement (i.e. CPTu). As a result, quantifying and propagating these uncertainties through stochastic simulation frameworks – such as sequential Gaussian co-simulation (Vessia et al., 2020) – is essential for building ground models that provide continuous estimates of the target variables and their uncertainties. This probabilistic characterisation aligns with the increasing emphasis on uncertainty-aware modelling in the 2nd generation of Eurocode and international guidelines for geotechnical reliability.

The number of measurements is usually too limited to capture effectively the spatial variability of the subsurface media, however, some achievements in featuring spatial variability through geostatistical tools have been done (Vessia et al., 2020a,; Vessia et al., 2020b, Di Curzio and Vessia 2023. Di Curzio et al., 2025) and the theory of random fields (RFs) (Vanmarke, 1983; Fenton and Griffiths, 2008; Li et al., 2015), where the soil parameters are defined by a probability distribution (PDF) and an autocorrelation function (ACF) with the scale of fluctuation (SOF). Based on dense numerical measurements, such as CPT data, it is possible to estimate spatial variability effectively; however, assessing the horizontal SOF remains challenging due to data sparsity (Ching et al., 2018). Estimating the input parameters required to generate RFs is time-consuming. It implies the selection of a trend function (Vanmarke, 1977; Pieczyńska-Kozłowska, 2015; Ching and Phoon, 2017), which is needed when RFs are used, leveraging the assumption of spatial stationary data, of the autocorrelation function and the estimation of the scale of fluctuation (Pieczyńska-Kozłowska et al., 2017; Camie et al., 2020; Phoon et al., 2022). Stationarity, especially in a weak sense, assumes that the ACF is spatial invariant. While convenient, this is a substantial simplification of the absolute natural heterogeneity of soils and rocks. Studies show that spatial stationarity can bias estimates and underrepresent uncertainty (Fenton and Griffiths, 2008; Ching and Phoon, 2013). Despite the use of locally adaptive models (Zhang et al., 2020), stationary Gaussian fields with fixed autocorrelation structures remain widely used in practice due to their simplicity and data limitations.

New frontiers in modelling the spatial variability structure of geomaterials involve combining RFs with Bayesian inference, such as the Bayesian Compressive Sampling (BCS) approach proposed by Wang and Zhao (2017). The latter allows the use of sparse measurements without requiring the stationarity assumption to generate RFs. The spatial structure varies locally, adapting to changes in subsurface conditions. BSC methods integrate prior knowledge and observed measurements to produce posterior estimates of spatial variability, enhancing reliability in heterogeneous soils (Li et al., 2016; Wang and Zhang, 2022; Hu et al., 2024). Unlike classical RF models, BSC can accommodate irregular data distributions and varying data density, generating nonparametric RFs.

Hereinafter, the BSC is used to generate nonparametric RFs. The case study focuses on the subsurface of the Po River Alluvial Plain (Italy), where extensive geotechnical and geological data were collected. A set of 6 CPT studies was selected; 4 were used for field modelling, and the remaining 2 for validation. Figure 1 shows the location of the CPTs.

2 METHODOLOGY

2.1 *Calculation of the undrained shear strength* s_u

The Po River plain is located in the Emilia-Romagna Region. It is characterised by hundreds of alluvial deposits made up of complex quaternary successions of silty sand, clay, and gravel interbedded at different depths along the meanders of the Po River. Further details on the geological and lithological character of the study area are provided in Di Curzio et al. (2020). Four CPT soundings (CPT1–CPT4) and two deeper reference CPTs (REF1–REF2) were selected within a 3 km² area. An almost regular grid with a horizontal spacing of

1.72 × 1.7 km was created, with a depth of 15 m. Raw CPT data were processed to compute undrained shear strength (s_u) from cone resistance (q_c) using the empirical correlations suggested by Lunne and Kleven (1981):

$$s_u = \frac{q_t - \sigma_z}{N_{kt}} = \frac{q_t - \sigma_z}{15} \tag{1}$$

where q_t is the corrected total cone resistance, σ_z is the total overburden pressure at each depth, N_{kt} is the tip factor, which ranges between 14 and 16 (Robertson and Cabal, 2015). $N_{kt} = 15$ is assumed in the current study.

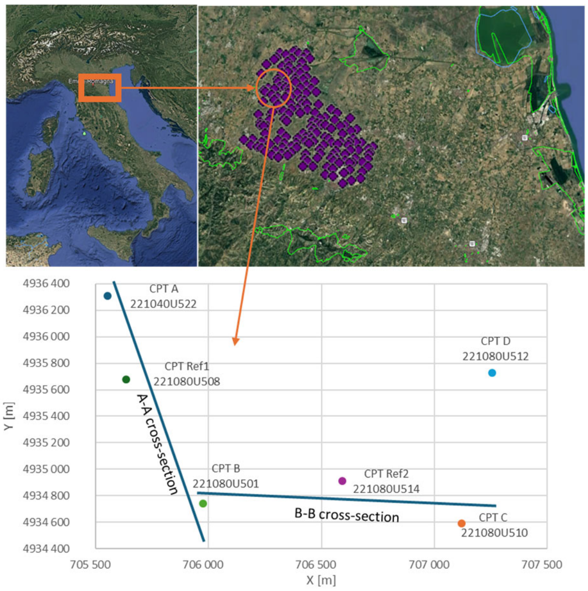


Figure 1. Location of the 6 CPTs adopted in this study.

2.2 *Bayesian Compressive Sampling*

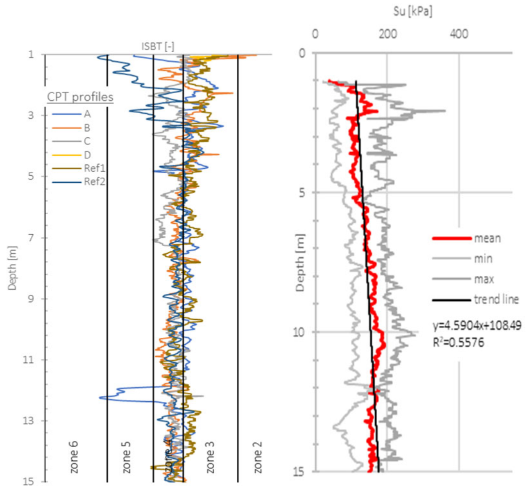
The BCS method provides a probabilistically rigorous alternative to the traditional RFs approach for modelling the spatial correlation of geotechnical parameters. In the present study, a discrete cosine transform (DCT) is used to efficiently represent smooth spatial variations commonly observed in soil profiles. BCS transforms the soil property profile from the spatial domain to the DCT domain, where most coefficients are near zero, yielding a sparse representation. Using a sampling matrix that encodes the measurement locations, the method estimates only the significant coefficients through non-informative priors and type-II maximum likelihood within a Bayesian framework. This process yields both the posterior mean and covariance of the coefficients, enabling explicit uncertainty quantification. The inverse DCT is then applied to reconstruct the high-resolution profile, which involves spatial correlation without a direct variogram function. This approach reduces the dependence on stationarity assumptions and performs well even with sparse CPT measurements (Wang and Zhao, 2017).

2.3 *Random Field Generation*

To generate RFs, the BCS-DCT framework offers a data-driven approach that relies solely on observed measurements, without explicitly calibrated analytical correlation models. The procedure starts with CPT-derived or synthetic measurements generating a measurement equation via a combined sampling (the measurements) and basis matrix (DCT coefficients). Bayesian inference identifies the most significant coefficients, accounting for measurement uncertainty. Once the posterior statistics are obtained, the inverse DCT produces the full-resolution random field, which is inherently consistent with the measurement data and preserves its spatial correlation structure. Unlike conventional approaches based on predefined ACF forms and variogram fitting, this method naturally incorporates uncertainty quantification. It avoids strong stationarity requirements, making it suitable for cases with limited or irregularly spaced data. Hereinafter, the free software ASSD-BCS_v1.2 (Hu et al., 2024) was used to generate RFs.

3 RESULTS AND DISCUSSION

3.1 *Soil Behaviour Type Index*



a) b)

Figure 2. (a) The CPT profiles used in this study and their I_{SBT} classification; (b) the s_u values calculated along zones classified as 3 and 4 according to I_{SBT} values.

According to Robertson's classification (2009), alluvial deposits from the Po River Plain consist of fine soils, such as clay and silty clay. They fall into Zones 3 and 4 of the soil behaviour type index (I_{SBT}) classification. Figure 2a shows the trend of the I_{SBC} along depth. The profiles were considered at depths ranging from 1 m to 15 m below ground level (36 m above sea level).

3.2 *Undrained shear strength*

The value of the s_u is shown up to a depth of 15m for all considered CPT shapes as shown in Figure 2(b). Basic statistics (mean, coefficient of variation, PDF) for each measurement are provided in Table 1.

Table 1. The mean and coefficient of variation of the s_u parameter for each CPT profile with depth.

CPT	s_u	
	Mean (kPa)	COV
A	101.65	0.22
B	168.53	0.23
C	188.43	0.17
D	151.31	0.30
Ref 1	129.21	0.31
Ref 2	131.22	0.29

As shown in Table 1, the traditional deterministic approach significantly overestimates computational parameter values because it does not account for variability, which ranges from 17 to 31%. The literature frequently refers to this range of variability (Cherubini et al., 2010).

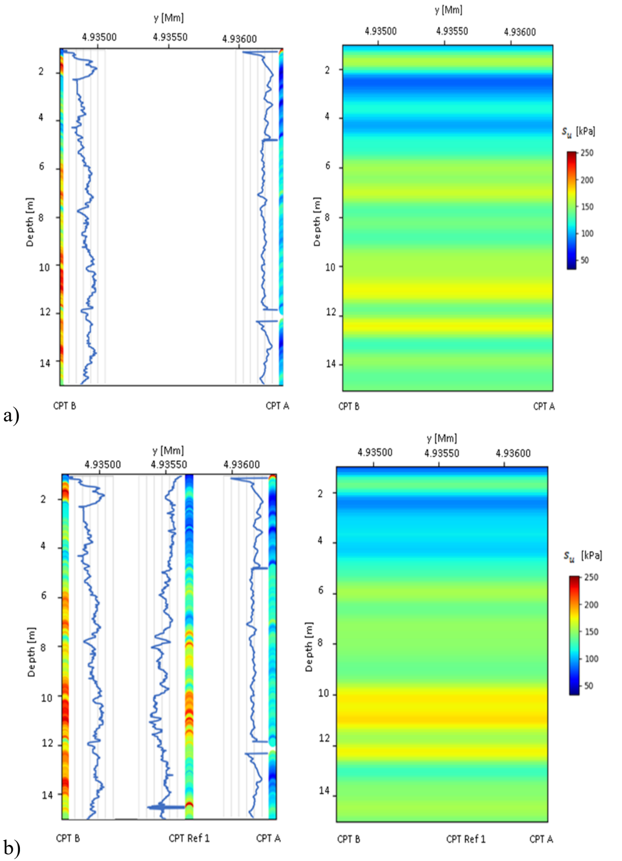


Figure 3. A-A cross-section random fields. (a) random field based on two CPTs profiles (CPT A & CPT B); (b) random field based on three CPTs profiles (CPT A & CPT B and CPT Ref1)

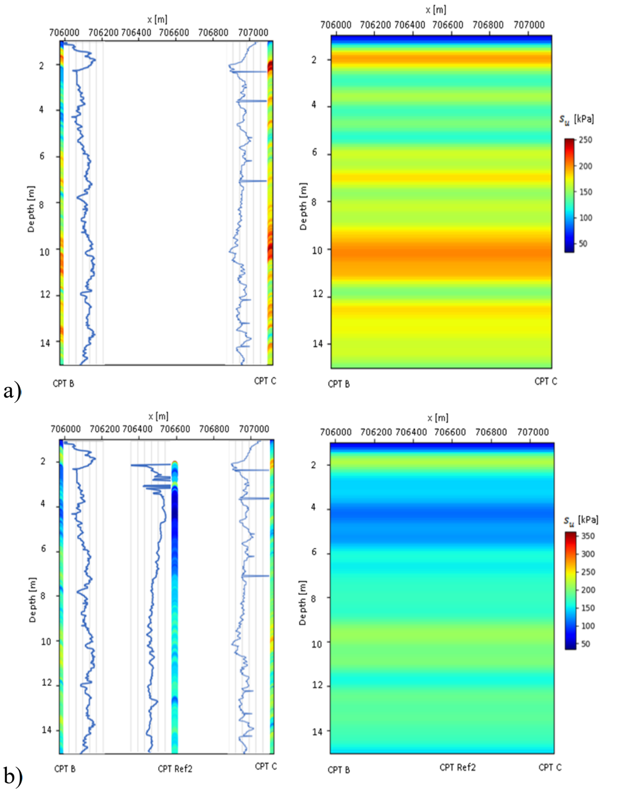


Figure 4. B-B cross-section random fields. (a) random field based on two CPTs profiles (CPT B & CPT C); (b) random field based on three CPTs profiles (CPT B & CPT C and CPT Ref2)

Traditional random fields often assume that a random variable is the residual from a detrended dataset with zero mean and unit standard deviation. The most commonly used ACF is the stationary Gaussian. The new approach suggested in the present work considers the s_u in total, without the entrance detrending procedure. Such an approach led to truly data-driven modelling of the s_u parameter along the depth.

Figure 3a and Figure 4a show the generated 2D random fields of the s_u variable based on two cross sections: (A-A) CPTA - CPTB; (B-B) CPTB - CPTC. The distance between the CPTs is approximately 1.5 km in section A-A (Figure 3a), and in section B-B (Figure 4a) around 1 km. Both random fields exhibit a variable character, with a clear upward trend with depth, ranging from approximately 50 kPa in the near-surface zone to nearly 200 kPa at a depth of 11 m below ground level. Although the s_u value observed in Figure 4a is higher, the zones of extreme values are located at similar depths, which might be caused by the common denominator of both cross-sections – CPT B.

By adding information to profiles A-A and B-B, better unification of both sections is observed. The additional third CPT profile ($CPT_{Ref\,1}$) added to the A-A cross-section reduces the s_u value at the ground surface, moderates the increase in the range of 2-5 m, smoothes the s_u variability between 5-10 m and enhances the s_u value between 10-12 m below the ground level. Figure 4b, with the additional profile $CPT_{Ref\,2}$, only apparently significantly changes the cross-section B-B. As a result of its inclusion, the range of the observed parameter s_u changes. The s_u values at the ground surface is decreseing, a further reduction of values between 3-6 m, and an observed maximum in s_u values at a depth of about 10 m below the ground level.

Research activity in the Po River Plain has been conducted through the classical RF approach, reporting a vertical SOF of s_u ranging between 0.7 m and 2 m (Pieczyńska-Kozłowska and Vessia, 2022). In the random fields obtained in the present study, comparable fluctuation patterns were observed: sequences of similar values formed thinner interlayers and thicker vertical stripes, with the latter not exceeding 2 m. However, the present Bayesian non-stationary framework enhances this picture by dynamically adapting to information from reference CPTs, leading to more consistent identification of fluctuation thicknesses and improved reproduction detection of local heterogeneities, thereby reducing the smoothing effect that characterises classical stationary RF models. As a result, the generated fields not only confirm previously measured fluctuation scales but also refine their interpretation, providing a more reliable description of the subsoil structure.

4 CONCLUSIONS

This study presented the application of the BCS framework as a CPT data-driven method for directly generating random fields. By avoiding the predefinition of correlation structure and stationarity assumptions, the proposed approach provides a more realistic, flexible, and uncertainty-aware alternative to classical random field models.

The CPT profiles used in this study are located at least 1 km apart, resulting in limited variability in the horizontal direction. However, the adopted approach allows the generation of a field with an almost horizontal layer arrangement while reflecting the sparse vertical spatial control. The generated field retained the increase in value, consistent with the data trend.

By adding a third reference CPT, the reconstructed fields adapted dynamically to the new information. However, the field does not change significantly; instead, it adapts (Figure 3b) in the region where CPT Ref1 exhibited higher parameter values. In the B-B cross-section, this was particularly evident, as the inclusion of Ref2 shifted the field toward lower undrained shear strength values, consistent with the measurements (Figure 4b), thereby correcting the overestimation observed in traditional deterministic approaches. The reference soundings thus acted as anchors, improving both the predictive consistency of the fields and their alignment with geological heterogeneity.

The enhancement is twofold: the method not only provides a better description of vertical variability, in line with previously reported fluctuation scales, but also integrates spatially distributed information, thereby refining local field dynamics. This adaptive update demonstrates the BCS-DCT ability to fuse multiple CPT datasets into a coherent, realistic subsurface model, even across large inter-CPT distances.

In conclusion, the study shows that integrating reference CPTs into a Bayesian non-stationary framework enhances the representativeness of random-field simulations and significantly improves subsoil characterisation compared with a stationary RF. This method directly contributes to the development of more reliable ground models, consistent with the principles of the 2nd generation Eurocode 7, and provides a robust foundation for uncertainty-aware geotechnical design.

5 REFERENCES

Cami, B., Javankhoshdel, S., Phoon, K. K., Ching, J. 2020. Scale of fluctuation for spatially varying soils: Estimation methods and values. ASCE-ASME *Journal of Risk and Uncertainty in Engineering Systems*, *Part A: Civil Engineering,* 6(4). https://doi.org/10.1061/AJRUA6.0001083

CEN. 2024a. *Eurocode 7: Geotechnical design – Part 1: General rules*. EN 1997-1:2024, The European Committee for Standardization, Brussels.

CEN. 2024b. *Eurocode 7: Geotechnical design – Part 2: Ground properties*. EN 1997-2:2024, The European Committee for Standardization, Brussels.

Cherubini, C., Giasi, C.I. and Rethati, L. 2020. The coefficients of variation of some geotechnical parameters. In *Probabilistic methods in geotechnical engineering*, pp.179-183. CRC Press.

Ching, J., and Phoon, K.K. 2013. Probability distribution for mobilised shear strengths of spatially variable soils under uniform stress states. *Georisk*, 7(3): 209224. Doi: 10.1080/17499518.2013.801273.

Ching, J., and K. K. Phoon. 2017. Characterising uncertain site-specific trend function by sparse Bayesian learning. *J. Eng. Mech.* 143 (7): 04017028. Doi:10.1061/(ASCE)EM.1943-7889.0001240.

Ching, J., Wu, T.J., Stuedlein, A.W., & Bong, T. 2018. Estimating horizontal scale of fluctuation with limited CPT soundings. *Geoscience Frontiers*, 9 (6), 1597-1608.

Di Curzio D., Vessia G. 2020. Multivariate Geostatistical Analysis of CPT Readings for Reliable 3D subsoil Modeling of Heterogeneous Alluvial Deposits in Padania Plain. International *Journal of Geoengineering Case Histories*, 6(4): 17-34. DOI: 10.4417/IJGCH-06-04-02.

Di Curzio, D., Vessi, G., 2023. Calculating Reliable Engineering Geological Model through Stochastic Co-Simulation Applied to CPTu Data, Geotechnical Special PublicationOpen source preview, 2023-July (GSP 347), pp. 259–267

Di Curzio, D., Castrignano, A. and Vessia, G., 2025. Assessing uncertainty propagation in CPTu-based hydro-mechanical subsoil characterisation using a multivariate stochastic simulation approach. *Engineering Geology*, 352(10806), p.4.

Fenton, GA., Griffiths, DV. 2008. *Risk assessment in geotechnical engineering*. Hoboken, N.J.: John Wiley & Sons

Hu, Y., Wang, Y., Phoon, K.K. and Beer, M., 2024. Similarity quantification of soil spatial variability between two cross-sections using auto-correlation functions. *Engineering Geology*, 331, p.107445.

Li, D.-Q., X.-H. Qi, Z.-J. Cao, X.-S. Tang, W. Zhou, K.-K. Phoon, and C.-B. Zhou. 2015. Reliability analysis of strip footing considering spatially variable undrained shear strength that linearly increases with depth. *Soils Found.* 55 (4): 866–880. https://doi.org/10.1016/j.sandf .2015.06.017.

Li, X. Y., Zhang, L., and Li, J. H. 2016. Using conditioned random field to characterise the variability of geologic profiles. *Journal of Geotechnical and Geoenvironmental Engineering*, 142(4), 04015096. doi:10.1061/(ASCE)GT.19435606.0001428.

Lunne, T., Kleven, A., 1981. Role of CPT in North Sea foundation engineering, In: *Proceedings of Geotechnical Engineering, Division Session*. ASCE National Convention. St. Louis (MO), pp.76–107

Vanmarcke, E. H. 1977. Probabilistic modeling of soil profiles. *J. Geotech. Eng. Div*. 103 (11): 1227–1246.

Vanmarcke EH. *Random fields analysis and synthesis*. Cambridge: MIT Press: 1983 World Scientific.

Pieczyńska-Kozłowska, J., 2015. Comparison between two methods for estimating the vertical scale of fluctuation for modeling random geotechnical problems. *Studia Geotechnica et Mechanica*, 37(4), pp.95-103.

Pieczyńska-Kozłowska J.M., Puła W. and Vessia G. 2017. A Collection of Fluctuation Scale Values and Autocorrelation Functions of Fine Deposits in Emilia Romagna Plain, Italy. Geo-Risk 2017 GSP 284, pp. 290-299, ASCE*, Proc. 6th International Symposium on Geotechnical Safety and Risk (Geo-Risk)*, Denver

Pieczyńska-Kozłowska, J. and Vessia, G., 2022. Spatially variable soils affecting geotechnical strip foundation design. *Journal of Rock Mechanics and Geotechnical Engineering*, 14(3), pp.886-895.

Phoon, K.K., Cao, Z.J., Ji, J., Leung, Y.F., Najjar, S., Shuku, T., Tang, C., Yin, Z.Y., Ikumasa, Y. and Ching, J., 2022. Geotechnical uncertainty, modeling, and decision making. *Soils and foundations*, 62(5), p.101189.

Robertson P K 2009 CPT interpretation – a unified approach. *Canadian Geotech. J*., 46, 1-19. https://doi.org/10.1139/T09-065

Robertson, P. K., Cabal, K. L. 2015. *Guide to Cone Penetration Testing For Geotechnical Engineering* (California: Gregg Drilling & Testing).

Qian, Z., and Shi, C. 2025. Large language model-empowered paradigm for automated geotechnical site planning and geological characterisation. *Automation in Construction*, 173.

Vessia G., D. Di Curzio, Castrignanò A. 2020a. Modeling 3D soil lithotypes variability through geostatistical data fusion of CPT parameters, *Science of The Total Environment*, 6981 134340.

Vessia, G., Di Curzio, D., Chiaudani, A., Rusi, S. 2020b. Regional rainfall threshold maps drawn through multivariate geostatistical techniques for shallow landslide hazard zonation. Science of the Total EnvironmentOpen source preview, 705, 135815

Wang, Y., and Zhao, T. 2017. Statistical interpretation of soil property profile from limited data using Bayesian compressive sampling. *Geotéchnique*, 67(6): 523536. DOI 10.1680/jgeot.16.P.143.

Zhao, T., Wang, Y., 2020. Nonparametric simulation of non-stationary non-gaussian 3D random field samples directly from sparse measurements using signal decomposition and Markov Chain Monte Carlo (MCMC) simulation. *Reliab. Eng. Syst. Saf.* 203, 107087.