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/ICSMGE2026-308

GIM-to-FEM: From digital ground information models to probabilistic numerical analysis of underground structures

Ksenija Micić, Novak Joksimović, Miloš Marjanović
Faculty of Civil Engineering, University of Belgrade, Serbia, kmicic@grf.bg.ac.rs

Zehao Ye
School of Engineering, University of Birmingham, United Kingdom

Jelena Ninić
Department of Engineering, University of Durham, United Kingdom

ABSTRACT: Prediction of ground behaviour, design optimization and risk management associated with complex ground-structure interactions requires information rich and accurate subsurface models. However, the diverse nature and scarcity of geological and geotechnical data pose major challenges in modelling, visualization and exchange across digital platforms. This is additionally complicated due to the ongoing demands for seamless interoperability, in terms of spatial variability representation and adaptability to new data. This study presents an advanced modelling framework for simulation of the soil spatial variability and geological uncertainty by joining the ground information modelling (GIM) and probabilistic FEM analysis for the underground construction projects. Voxel-based 3D geological model is generated in GIS environment by novel geoprocessing algorithms. To account for variable ground conditions, Conditional Random Fields (CRFs) are employed to simulate the geotechnical parameter spatial distribution, highlighting the specific behaviour of various soil layers, as well as complex formations like interlayers. Geological voxel model and CRF parameter model are merged into a holistic digital ground information model in Revit, interoperable with the FEM framework. GIM-to-FEM approach performs smooth and efficient subsurface data export to a cell-based computing mesh, thus providing the simulation of realistic conditions and ensuring the enhanced realism in predicting the influencing factors of the ground-structure interaction using BIM framework as main data-exchange platform. Key features of the proposed modelling system deal with complete interoperability and data exchange between digital and numerical models (including model updating with new data), therefore ensuring that the simulations reflect the most up-to-date geological and geotechnical conditions. To demonstrate the efficiency of the proposed modelling framework, a case study of the mechanized tunnelling project is presented. Proposed approach shows advantages over traditional methods in storing and managing subsurface data and allows the comparison between deterministic and probabilistic numerical approaches, offering more accuracy and reliability when modelling heterogeneous soils.

KEYWORDS: Ground information model, conditional random fields, GIS, FEM, BIM.

1 INTRODUCTION

Reliable prediction of ground behaviour and safe design of underground structures depend on accurate and information-rich subsurface models. Geological formations are inherently heterogeneous, and both stratigraphic configuration and geotechnical parameters can vary significantly in space, influencing soil-structure interaction and the outcomes of numerical analysis (Jaksa et al, 2005; Ali et al, 2017; Juang et al, 2019). In practice, site investigation campaigns are often limited to sparse measurements due to cost and logistical constraints (Ching & Phoon, 2020; Phoon et al, 2022), leaving considerable uncertainty in both layer geometry and soil property distributions. These uncertainties propagate directly into numerical simulations, affecting estimates of structural capacity, deformation behaviour, and associated risk (Zhang et al, 2022; Minini et al, 2023; Zhao et al, 2025).

Advances in digital ground modelling offer opportunities to reduce these uncertainties by integrating multi-source datasets into continuous 3D subsurface representations (Shi and Wang, 2022; Khan et al, 2025; Zhao et al, 2025; Zinas et al, 2025). Voxel-based models, which discretize the subsurface into volumetric units enriched with semantic attributes, such as soil type and mechanical properties, provide a scalable and uniform structure for storing, updating, and exchange of this information. Their compatibility with Building Information Modelling (BIM) environments makes them particularly suited for the development of robust digital twins of the underground space (Huang et al, 2021; Khan et al, 2023; Khan et al, 2025).

Despite the high interest and fast pace of improving in this field, there remains a gap between high-fidelity ground information models (GIMs) and their direct, automated use in FEM solvers. Existing workflows often require manual data transfer, risking errors and loss of spatial detail. To address this, we present a fully interoperable GIM-to-FEM framework that: (1) generates a voxel-based stratigraphy model from borehole data using advanced geostatistical tools in GIS; (2) enriches it with layer-specific Conditional Random Fields (CRFs) of key soil parameters to capture spatial variability; and (3) makes use of BIM to seamlessly integrate GIM into a FEM solver for probabilistic soil-structure interaction analysis. This end-to-end approach enables consistent, up-to-date numerical simulations that reflect both geological complexity and parameter uncertainty, supporting more reliable design and risk assessment in underground construction. The proposed workflow is demonstrated on a case study of the mechanized tunnelling project.

2 GROUND INFORMATION MODEL

2.1 *Stratigraphy delineation in ArcGIS*

Stratigraphy modelling requires a workflow capable of handling various-scale, multi-modal datasets, supporting full 3D representation of the subsurface, enabling parametric modelling for iterative refinement, and offering modularity to integrate with downstream geotechnical analyses. As complex and challenging as this may sound, there are modelling systems that are capable of meeting these requirements. ArcGIS is one of the most common systems, offering a robust Geographic Information System (GIS) platform widely used in geosciences for integrating spatial datasets, managing complex workflows, and applying advanced geostatistical methods, all while maintaining computational efficiency. From the wide spectrum of user-friendly algorithms, herein we employ Empirical Bayesian Kriging (EBK) to conduct borehole-based stratigraphy delineation in 3D space. This advanced geostatistical algorithm performs great ability to automatically estimate semivariogram parameters through an iterative Bayesian process, thus reducing the subjectivity in model fitting.

The modelling procedure, illustrated in Figure 1, begins with borehole data preprocessing to form the spatial point feature class, which in this case consists of spatial information on stratigraphic zone boundaries (top and bottom). Next, EBK input parameters such as semivariogram type, subset size, vertical anisotropy, elevation inflation, and search neighbourhood are configured to account for data spread and correlation. The algorithm then performs parameter value prediction at unobserved locations within the Gaussian spatial process framework, by combining the intrinsic random function kriging and linear mixed models (more details given in Krivoruchko and Gribov, 2019; Esri, 2025). The EBK product, stored within geostatistical horizontal transect rasters, is then transformed into multidimensional voxel-based NetCDF format. Finally, the continuous interpolation output is reclassified into discrete soil categories based on the cross-validation results provided by the algorithm.

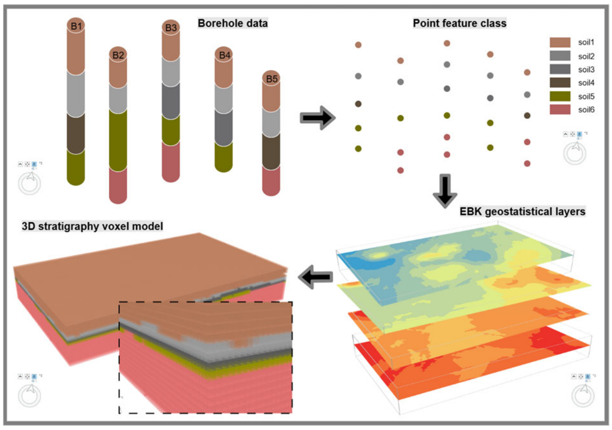

Figure 1. Borehole-based stratigraphy modelling in ArcGIS.

2.2 *Conditional modelling of soil spatial variability*

Realistic and reliable site characterization is greatly dependent on the spatial variability of soil geotechnical parameters. Modelling soil inherent variability and evaluating its influence on soil-structure interactions is mostly being performed within Random Field Theory (Pan and Dias, 2017; Gong et al, 2018; Hu et al, 2025).

To account for the available site 'knowledge', conditional simulation of parameter random fields is being employed. Conditional random field is derived in GSTools (Müller et al, 2022) following a three-step procedure: 1) kriging-based estimation, 2) unconditional Gaussian random field generation ($F_u(x)$), and 3) conditioning via kriging correction ($F_c(x)$):

$$F_u(x) = \sqrt{\frac{\sigma^2}{N}} \sum_{m=1}^{N} [a_m cos(k_m \cdot x) + b_m sin(k_m \cdot x)] \quad (1)$$

$$F_c(x) = F_u(x) + \sum_{j=1}^{n} \lambda_j(x) \left[Z(x_j) - F_u(x_j)\right] \quad (2)$$

where $Z(x_j)$ is the observed data at locations x_j; a_m and b_m are independent standard normal random variables; N is the number of Fourier modes, k_m are independent random wave vectors sampled from the spectral density of the variogram, and $\lambda_j(x)$ are the kriging weights at prediction location x. The unconditional field relies on a covariance model defined by variance, nugget, correlation length, and correlation function, and is simulated using the Randomization method in the Fourier domain (Heße et al. 2014; Kraichnan, 1970). The final CRF is obtained by adding the unconditional field to the difference between kriging estimates from measured data and those from the unconditional simulation (Equation (2)), ensuring the realization meets both the spatial correlation structure and the conditioning data. For more details, the reader is referred to Müller et al. (2022).

2.2.1 *Layer-based conditioning*

Given that different soil types, due to their distinctive geological nature, often exhibit varying characteristics not only compared to other types but also within their own spatial extent, it is recommended to develop the random fields for each stratigraphic layer individually, whenever possible.

Layer-based CRFs are being produced on a structured interpolation grid, inheriting the simulation domain from the voxel-based stratigraphy model. Consequently, this may introduce a challenge: not all soil parameter specimens from boreholes will spatially align with their corresponding soil layer in the ArcGIS EBK stratigraphy model, due to the imprecision of the stratigraphy interpolation. With the possible mismatch of this kind, we risk either the loss of valuable (and most often very scarce) conditioning data or distorting the soil model's spatial integrity.

To address the issue, a carefully designed hybrid domain-expansion strategy is proposed, as illustrated in Figure 2. The inherited original spatial domain is, if necessary, being expanded by forming the voxel buffers around the mismatched conditioning points, to allow for the inclusion of all available measurements in the simulation. The expanded domain retains the original voxel geometrical attributes. After the simulation, only the original EBK-delineated voxels get their CRF values assigned.

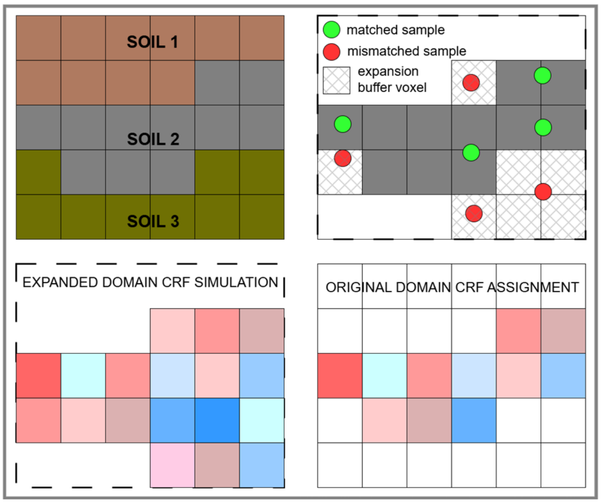

Figure 2. Layer-based CRF expanded domain simulation.

2.3 *Model update*

When developing digital ground information models for various applications in subsurface space utilization and dynamic digital twins, it is crucial to prioritize continuous model updates with new knowledge over time.

In this study, we add new knowledge in the form of additional site investigation data (boreholes and soil specimen laboratory testing). Such update automatically augments existing modelling datasets both for stratigraphy and soil parameter simulations – the stratigraphy interpolation and CRF generation process is repeated, including the new and richer dataset. The update allows for ground model re-evaluation in order to capture additional spatial dependencies, as well as to be able to add completely new instances (e.g. new soil types detected by drilling at new locations).

The new borehole dataset and GIM are considered more reliable, as they are based on higher amount (and usually more recent) geotechnical data. The aim of this concept is not only to confirm the importance of 'certainty', but also to underscore the undeniable need for a greater number of investigation works, as well as their optimal spatial distribution, with the ultimate goal of achieving knowledge-driven and reliable geotechnical modelling.

2.4 *Accuracy evaluation*

To evaluate the accuracy of the GIMs generated using the previously described tools and algorithms, and to demonstrate the improvement as the level of domain knowledge increases, we propose a preliminary simple and efficient metric system, as follows. The predictive performance of stratigraphy models is evaluated using borehole data as the ground truth – predictions being represented by the model itself, and true values being represented by the particular model's dataset. A combination of zone-majority accuracy, balanced accuracy and Matthew's correlation coefficient (MCC) is computed for each model.

Zone-majority accuracy is defined as the proportion of borehole zones for which the predicted majority voxel-level soil types match the observed soil zone interval. This metric reflects the model's ability to capture the primary stratigraphic character of an interval, with higher value illustrating better prediction. It is calculated as the ratio of correctly classified zones over the total number of zones within a borehole record.

Since soil classes are often imbalanced in spatial distribution, the balanced accuracy metric is employed to provide an unbiased evaluation across classes. It represents the macro-average of recall scores per class, treating all classes equally regardless of their abundance and providing a more reliable performance measure in class-imbalanced datasets. Random prediction scores the value of 0, whereas 1 marks the perfect prediction (Mosley, 2013).

The MCC provides a robust single-value summary of classification quality, accounting for all elements of the confusion matrix (true/false positives and negatives) and is suitable for multi-class problems. Originally, it is formulated for the binary case (Matthews, 1975) but is also generalized to fit the multi-class case (Gorodkin, 2004). In its essence, it is a correlation coefficient value that spans between -1 (inverse prediction) and +1 (perfect prediction).

In order to efficiently and directly compare parameter spatial variability across GIMs (represented by per-property per-soil type CRFs), it is convenient to express the predictive capability of each model to a single value. Here we use the Normalized Root Mean Squared Error (NRMSE), an extension of RMSE that allows for comparing different datasets and models with varying scales. The 'prediction vs measurement' RMSE is normalized for the parameter value range for a distinct model, providing a dimensionless and interpretable measure of relative prediction error:

$$NRMSE = \frac{RMSE}{range(y)} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}}{max(y) - min(y)} \tag{3}$$

where y_i are measured values, $\hat{y}_i$ are CRF-predicted values, and n is the number of conditioning points.

3 GIM-TO-FEM FRAMEWORK

The developed framework establishes an automated and repeatable procedure for transforming a voxel-based GIM into a fully generated PLAXIS 2D model of the tunnelling project, ensuring consistency and traceability from geological data to numerical analysis. The process begins with the GIM providing a CSV file that contains the centroid coordinates of each voxel (X, Y, Z), their dimensions (dx, dy, dz), soil type identifiers, and the corresponding geotechnical parameters. We use Revit Dynamo scripting module to: (1) instantiate a custom voxel family and a parametric tunnel family along the project alignment and (2) to pre-process the FEM numerical model input, via application of the active cutting plane within the Revit model, automatically identifying all intersected voxel and tunnel instances, and extracting their geometry and material attributes. These are then compiled into a section-specific CSV file, which contains all necessary data for generating the FE model. This step reconstructs both the geological configuration and the tunnel geometry directly from the GIM dataset, producing a precise BIM representation of the chosen project area.

Then, section-specific CSV file is passed to an automation script via PLAXIS 2D API (in Python) to build the model in its entirety. Since PLAXIS 2D does not support the direct import of the Industry Foundation Class format (IFC), the model preparation must be done stepwise, first by generation of model geometry, followed by assignment of structural properties to geometric entities. By isolating the data preparation (via Dynamo) from the PLAXIS model creation, the framework allows any cross section along the tunnel alignment to be modelled simply by adjusting the cutting plane in Revit. The process eliminates manual geometry recreation, reduces user error, and significantly shortens the turnaround time from ground modelling to FEM analysis. The workflow for transferring data from Revit to PLAXIS, using Dynamo for extraction and Python for automated model generation, is presented in Figure 3.

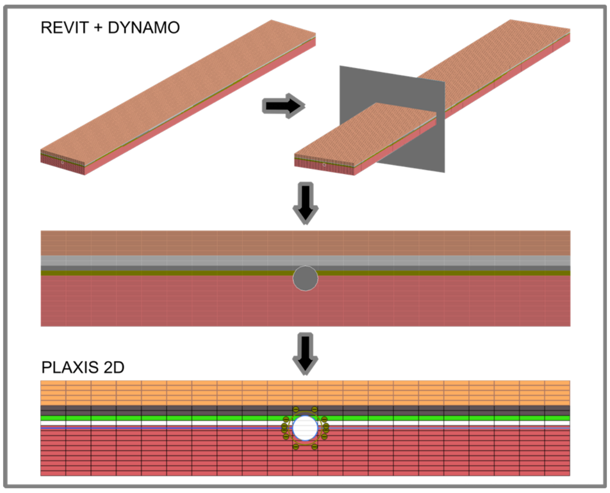

Figure 3. Revit-to-PLAXIS workflow and model visualization.

4 IMPLEMENTATION

4.1 *GIMs generation*

The proposed framework is demonstrated on a real case study covering the area of blocks 61, 62, 63 and 64 in New Belgrade, Serbia. All the available borehole data is first pre-processed, restructured and formatted into a digital borehole database, as described in Micić et al. (2024). The full dataset consists of 94 boreholes carrying information on their spatial positions, existing soil zones, and soil specimens with laboratory testing results.

To account for model updating and evaluate its effects, we generate 5 GIMs, each with an extended modelling dataset, as detailed in Table 1. Modelling datasets are extended in such manner that they augment the previous model's boreholes dataset with additional knowledge (e.g. M_30 consists of model's M_10 boreholes and additional 20 boreholes). Stratigraphy modelling is performed as elaborated in Section 2.1, using the total number of soil zone points presented in Table 1. All 5 models cover the exact same area and consist of the exact same number and size of voxels – 150x93x19 voxels in XYZ, with the size of 10x10x2m in XYZ, respectively. Each model's EBK continuous prediction is reclassified into discrete classes using the corresponding cross-validation metrics.

Table 1. Total number of boreholes, soil points and sample conditioning points used for GIMs generation.

Model	Total boreholes	Total soil zone points	Total conditioning points
M_10	10	82	18
M_30	30	249	50
M_50	50	432	84
M_70	70	600	107
M_94	94	808	148

Conditional realizations of random fields for cohesion, internal friction angle and Young's modulus (as chosen FE model parameters) are derived for each model individually, following the per-layer methodology (see Section 2.2.1), making use of the total number of conditioning points presented in Table 1. Table 2 shows the geotechnical parameter value spans for the full borehole dataset.

Voxel-based GIMs are constructed by incorporating the stratigraphy with soil properties information, such that each voxel centre includes its characteristic soil type and soil property value.

Table 2. Per-soil geotechnical properties value spans for full dataset.

Soil	Full dataset		
	c (kPa)	*φ* (°)	*E (kPa)***
Fill*	5	18	5000
Silty clay	0 - 35	12 - 32.5	585.4 - 11596
Silty sand	0 - 20	21.5 - 40.8	3400 - 11900
Silt	0 - 10	17 - 24	1352 - 3848
Sand	0 - 25	20 - 36	1808.3 - 18516.7
Sand&gravel	0 - 50	18.9 - 38.5	5308.3 - 30833.3

*There is no measured soil specimen data for fill. Adopted constant values are presented. **Es is derived from oedometric modulus E_{oed} and adopted Poisson's ratio ν (Maksimović, 2005).

4.2 *FEM numerical model*

Finite element numerical models are developed in 2D, for a single cross section in each of the described GIMs using the automated data transfer and model generation workflow described in Section 3. The aim of these models is to simulate the excavation of a shallow tunnel with no lining applied, under site-specific ground conditions. The reference tunnel cross section is circular, with an internal diameter of 10 m, and is positioned at a depth where the overburden is relatively small. Given these conditions, the excavation method ensures uniform relaxation of the tunnel face in sequential phases, whereas displacements increase as the tunnel face is fully destressed.

The model domain and geometry are imported directly from the GIM dataset, ensuring that the stratigraphy, soil boundaries and associated geotechnical parameters remain consistent with the ground model. In all analyses, the soil is represented by the elastic–perfectly plastic Mohr–Coulomb constitutive model under drained conditions, a widely accepted choice for preliminary tunnel design and comparative studies. The values of cohesion (c), friction angle (φ), and Young's modulus (E) are read directly from the processed GIM data, thereby avoiding manual data transcription and guaranteeing consistency. This approach ensures that each numerical model provides a realistic representation of soil-structure interaction, allowing for reliable evaluation of tunnel performance under the site-specific geological and operational constraints.

5 RESULTS AND DISCUSSION

5.1 *GIM accuracy*

The results obtained by conducting the predictive performance evaluation of multiple GIMs, derived with updated modelling datasets, are presented in Figures 4 and 5. Generally, it is clearly observed that GIMs with greater scope of available knowledge perform better predictions of lithology and geotechnical parameter spatial distribution.

Figure 4 presents consistent improvement in zone-majority accuracy with an increasing boreholes density, rising from 45.45% in the 10-borehole case to 70.87% in the full dataset model. This indicates that the inclusion of additional boreholes enhances the spatial reliability of stratigraphy prediction by providing denser conditioning constraints. Balanced accuracy increases sharply between 10 and 30-borehole models but also slightly decreases for intermediate datasets before recovering in the full dataset. This suggests that while more boreholes improve the model's performance for dominant soil classes, challenges remain in accurately predicting less represented layers. MCC exhibits a steady increase with borehole count, indicating an overall improvement not only in prediction accuracy but also in the balance and structure of classification errors. Higher MCC values in the denser datasets imply that the model is less prone to systematic misclassification between specific stratigraphy pairs. The observed trends confirm that increasing the number of boreholes substantially improves the model's reliability, although certain soil classes remain challenging to predict due to inherent geological complexity and sampling limitations.

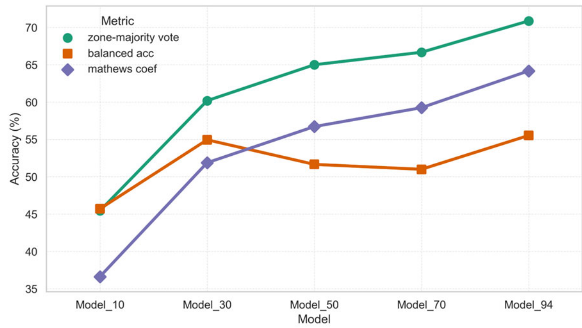

Figure 4. Prediction accuracy evaluation for stratigraphy models trained with an increasing number of conditioning boreholes.

A radar diagram in Figure 5 illustrates the variation of NRMSE for CRF-obtained geotechnical parameters as a function of distinct number of conditioning data used for simulations. From a general perspective, all three parameters exhibit a reduction in NRMSE with increasing borehole data count, which manifests as a progressive inward shift of the plotted graphs from the outer (10-borehole) to the inner (94-borehole) positions. However, the rate and uniformity of improvements differ between parameters. While cohesion and friction angle benefit in a more balanced manner from the increased borehole count, meaning that they show relatively consistent improvements, Young's modulus NRMSE is represented through a more irregular graph, with larger deviations in certain increments, and only slight fluctuations in others. Even though the overall trend is still toward the lower NRMSE, the irregularity suggests that Es is possibly more sensitive to localized data density variations, indicating the potential need for targeted data acquisition or model refinements.

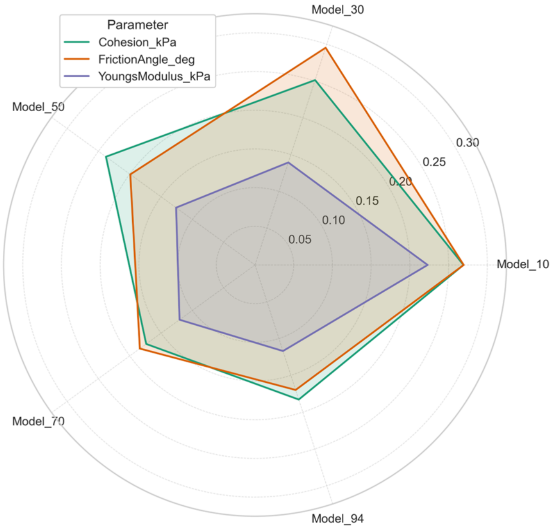

Figure 5. NRMSE values for soil properties across CRF models simulated with an increasing number of conditioning data.

5.2 *FEM analysis*

The results of the FEM analyses indicate a notable variability arising from the inherent heterogeneity of the soil's geomechanical parameters. Such variability has a direct impact on tunnel design, where stability is strongly influenced by the degree of stress relief in the rock mass/soil during excavation. A stronger lining does not necessarily represent a better solution if it is not installed at the optimal time. The convergence diagram (radial displacements) along the tunnel contour (Figure 6) shows a slight asymmetry, reflecting the influence of ground heterogeneity and the resulting variations in geotechnical parameters, which in turn cause non-uniform loading of the tunnel lining. The linearized diagram's starting point is at the left spring-line point, with a counter-clockwise orientation and positive magnitude of inward displacements.

The magnitude of this loading depends on the specific cross section considered. Numerical analyses incorporating CRFs generate non-homogeneous boundary conditions, providing a more realistic representation of ground behaviour compared to conventional analyses with homogeneous materials and constant parameters. Even when parametric studies are performed on homogeneous models, the results may show changes in absolute values, but they fail to reproduce the effects of irregular boundary conditions. This is also seen in the radial displacements along a horizontal cross section, originating from the left and right spring-line points, as shown in Figure 7, which shows unequal displacements in the left and right cross section. The influence of the borehole dataset is evident in the current results; however, its contribution becomes clearer when evaluated across a larger number of 2D sections and would be most effectively visualized within a fully three-dimensional numerical analysis.

Structural response is significantly different when the dataset is increased. GIM with only 10 boreholes seems inadequate for reliable simulation, while with the borehole number increase, the structural responses converge to less variable behaviour.

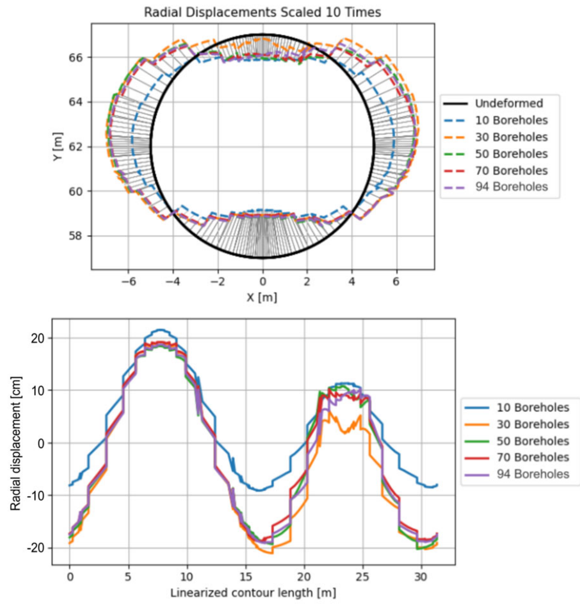

Figure 6. Radial displacements along the tunnel contour.

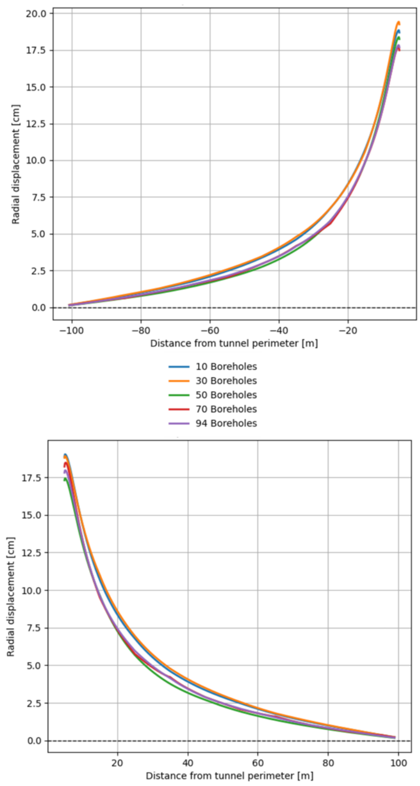

Figure 7. Radial displacements measured from the tunnel perimeter in the left and right horizontal cross sections.

6 CONCLUSIONS

The analysed workflow demonstrates interoperability between different software platforms, enabling a more accurate representation and streamlined transition from archival and investigation data related to geological surveys, to their visualization in a BIM environment, subsequent export from

BIM to numerical modelling software, and potential feedback for result visualization within the BIM environment. By putting the BIM environment (Revit) "in the middle" of the design process, supported by FEM and GIS tools, the current structural design practice can be shifted towards a concept of digital twinning.

The demonstrated GIM-FEM model updating clearly indicates the importance of the ground investigation works scope, especially in large infrastructure projects. Presented results of GIM models accuracy evaluation, as well as the FEM analysis, show convergence only for rich datasets, demonstrating the possibility for smart optimization of ground investigation campaigns.

While this approach offers substantial benefits through the central BIM environment in Revit, careful consideration must be given to the domain size, particularly regarding the data transfer between software. Dynamo, as the primary tool for parameterizing models in Revit, has proven highly effective; however, issues arise with models exceeding 60,000 voxels, likely because it is not optimized for vectorized input but rather for sequential data handling through the Dynamo–Revit interface. In such cases, performance could be improved by batching data input into smaller segments.

Another optimization challenge is the generation of large numbers of voxels in PLAXIS. One possible approach is to classify soil materials to a degree that preserves the CRF variability but avoids assigning a unique material to every voxel. This could be achieved by defining parameter intervals for key geomechanical properties, allowing materials to be grouped. In turn, this would reduce the number of instantiated materials within the model, improving the interface responsiveness. Future development should also address the potential feedback loop for displaying numerical analysis results directly within the Revit BIM model.

7 ACKNOWLEDGEMENTS

This research was supported by the Science Fund of the Republic of Serbia, #345, Digital and Numerical Model Integration for Optimization in Geotechnics – DiNum-GEO. Authors are partially supported by the Serbian Ministry of Science, Technological Development and Innovation via Project 200092.

8 REFERENCES

Ali, A., Lyamin, A. V., Huang, J., Li, J. H., Cassidy, M. J., and Sloan, S. W. 2017. Probabilistic stability assessment using adaptive limit analysis and random fields. *Acta Geotechnica* 12, 937–948.

Ching, J., and Phoon, K. K. 2020. Constructing a site-specific multivariate probability distribution using sparse, incomplete, and spatially variable (MUSIC-X) data. *Journal of Engineering Mechanics* 146(7).

Esri, 2025. What is empirical Bayesian kriging? *ArcGIS Pro documentation*. [Online] Available at: https://pro.arcgis.com/en/proapp/latest/help/analysis/geostatistical-analyst/what-is-empirical-bayesian-kriging-.htm [Accessed 1st June 2025].

Gong, W., Juang, C.H., Martin, J. R., Tang, H., Wang, Q., Huang, H. 2018. Probabilistic analysis of tunnel longitudinal performance based upon conditional random field simulation of soil properties. *Tunnelling and Underground Space Technology*, 73, 1–14.

Gorodkin, J. 2004. Comparing two K-category assignments by a K-category correlation coefficient. *Computational Biology and Chemistry* 28(5–6), 367–374.

Heße, F., Prykhodko, V., Schlüter, S., and Attinger, S. 2014. Generating random fields with a truncated power-law variogram: A comparison of several numerical methods. *Environmental Modelling & Software* 55, 32–48.

Hu, L., Kasama, K., Wang, G., and Takahashi, A. 2025. Assessing the influence of geotechnical uncertainty on existing tunnel settlement caused by new tunnelling underneath. *Tunnelling and Underground Space Technology* 155.

Huang, M. Q., Ninić, J., and Zhang, Q. B. 2021. BIM, machine learning and computer vision techniques in underground construction: Current status and future perspectives. *Tunnelling and Underground Space Technology* 108.

Jaksa, M. B., Goldsworthy, J. S., Fenton, G. A., Kaggwa, W. S., Griffiths, D. V., Kuo, Y. L., and Poulos, H. G. 2005. Towards reliable and effective site investigations. *Géotechnique* 55(2), 109–121.

Juang, C. H., Zhang, J., Shen, M., and Hu, J. 2019. Probabilistic methods for unified treatment of geotechnical and geological uncertainties in a geotechnical analysis. *Engineering Geology* 249, 148–161.

Khan, M. S., Cho, H. S., and Seo, J. 2025. Voxel-based digital twin framework for earthwork construction. *Applied Sciences* 15.

Khan, M. S., Kim, I. S., and Seo, J. W. 2023. A boundary and voxel-based 3D geological data management system leveraging BIM and GIS. *International Journal of Applied Earth Observation and Geoinformation* 118.

Kraichnan, R. 1970. Diffusion by a random velocity field. *Physics of Fluids* 13, 22–31.

Krivoruchko, K., and Gribov, A. 2019. Evaluation of empirical Bayesian kriging. *Spatial Statistics* 32.

Maksimović, M. 2005. Mehanika tla (In Serbian). 3rd Edition. Belgrade: Građevinska knjiga.

Matthews, B. W. 1975. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta-Protein Structure* 405(2), 442–451.

Micić, K., Ye, Z., Marjanović, M., and Ninić, J. 2024. Automated generation of ground model incorporating soil spatial variability. *Proc. 28th European Young Geotechnical Engineers Conference (EYGEC 2024)*, Demir Kapija, North Macedonia, June 25–29.

Minini, J., Zhang, Y., Groslambert, M., and Commend, S. 2023. Finite element-based probabilistic framework including Bayesian inference for predicting displacements due to tunnel excavation. *Computers and Geotechnics* 162.

Mosley, L. 2013. A balanced approach to the multi-class imbalance problem. *PhD Dissertation*. Iowa State University.

Müller, S., Schüler, L., Zech, A., and Heße, F. 2022. GSTools v1.3: A toolbox for geostatistical modelling in Python. *Geoscientific Model Development* 15, 3161–3182.

Pan, Q. J., and Dias, D. 2017. Probabilistic evaluation of tunnel face stability in spatially random soils using sparse polynomial chaos expansion with global sensitivity analysis. *Acta Geotechnica* 12, 1415–1429.

Phoon, K. K., Ching, J. Y., and Shuku, T. 2022. Challenges in data-driven site characterization. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* 16(1), 114–126.

Shi, C., and Wang, Y. 2022. Data-driven construction of three-dimensional subsurface geological models from limited site-specific boreholes and prior geological knowledge for underground digital twin. *Tunnelling and Underground Space Technology* 126.

Zhang, W., Han, L., Gu, X., Wang, L., Chen, F., and Liu, H. 2022. Tunneling and deep excavations in spatially variable soil and rock masses: A short review. *Underground Space* 7(3), 380–407.

Zhao, C., Gong, W., Juang, C. H., Tang, H., Liu, Z., and Hu, X. 2025. Probabilistic performance analysis of longitudinal tunnels based on coupled characterization of stratigraphic and geo-properties uncertainties. *Tunnelling and Underground Space Technology* 161.

Zhao, C., Gong, W., Juang, C. H., Tang, H., Liu, Z., Hu, X. 2025. Probabilistic characterization of subsurface stratigraphic configuration by merging borehole and geophysical data. *Acta Geotechnica* 20, 2421–2439.