The stratigraphic reconstruction of longitudinal tunnel based on improved coupled Markov chains

Qihao Jiang[1], *Jinzhang* Zhang[1], and *Dongming* Zhang[1#]

[1]*Dept. of Geotechnical Engineering, Tongji Univ, Shanghai China*

[#]*Corresponding author: 09zhang@tongji.edu.cn*

ABSTRACT

The distribution of natural strata is uncertain due to tectonic movements and sedimentation. Capturing geological uncertainty is a challenge for traditional deterministic models. In this study, an improved three-dimensional coupled Markov chains method for probabilistic stratigraphic reconstruction was developed. This method considers the correlation between the field borehole data. On this basis, an inversion analysis method for horizontal transition probability matrix estimation is proposed. This method makes the predictions more suitable for possible stratigraphic distributions. The accuracy of the method was further verified by different borehole schemes from the Mawan Tunnel in Shenzhen. The results show that the proposed method can still have high accuracy when the number of boreholes is sparse. This method can reflect the asymmetry, continuity and anisotropy of three-dimensional strata.

Keywords: Geological uncertainty; Coupled Markov chains; Tunnel; Three-dimensional.

1. Introduction

The properties of natural geological formations exhibit significant uncertainty due to complex geological, environmental, and physicochemical processes, rarely being homogeneous, a widely accepted notion (Elkateb et al., 2003; Phoon & Kulhawy, 1999). Geological variability significantly impacts the subsurface geological structures (Yeh et al., 2021; Zhang et al., 2020). Geological uncertainty manifests primarily in two forms (Elkateb et al., 2003): one being the geological variability within heterogeneous formations, often presenting as irregular embeddings of one geological material within another (Elfeki, 2006); the other uncertainty arises from spatial variability of the same geological parameters within a single formation, manifesting as spatial variations of geological parameter properties from one point to another (Griffiths et al., 2002; Phoon & Kulhawy, 1999). Due to insufficient borehole data in practical engineering, research on subsurface variability remains relatively scarce. One choice for modeling subsurface variability is based on geostatistical variogram simulation methods, such as the Kriging method (Chiles & Delfiner, 1999), Gaussian threshold models (Matheron et al., 1987), and multipoint geostatistics methods (Strebelle, 2002). However, these methods have limitations, such as the inability of the Kriging method to fully utilize the cross-correlation information contained in the data (Deutsch & Journel, 1998). Gaussian threshold models and multipoint geostatistics methods heavily rely on the quality of site-specific borehole data and require a sufficient number of borehole data to infer the spatial correlation structure of geological heterogeneity, which is often inadequate in practical engineering, making it difficult to meet the requirements of these methods (Carle, 2000; Zhang & Dasaka, 2010). From the aforementioned research findings, it is evident that subsurface variability is objectively present, and the coupled Markov chain model is a computationally efficient simulation method. However, due to the typically limited amount of field borehole data, estimation of the horizontal transition probability matrix in the coupled Markov chain model based on finite field data is a major research focus.

2. Methodology

2.1. Improved three-dimensional coupled Markov chain model

In geotechnical engineering, the uncertainty simulation of subsurface involves different soil types distributed across various spatial locations within the strata, corresponding to different states in the Markov chain. By dividing the strata into units and considering the distance between soil units as the difference in order between states in the Markov chain, Markov chains can be used to describe the transition of soil states. A three-dimensional coupled Markov chain (ICMC3D) provides a more realistic simulation of subsurface uncertainty compared to one-dimensional Markov chains limited to two directions. As depicted in Figure 1, the soil within the simulation area is divided into Nx × Ny × Nz units of equal size, with each unit corresponding to its own soil state. Additionally, it is necessary to determine the soil types exposed at the ground surface and the information from the nearest boreholes. The specific

calculation formula is shown in Equation (1):

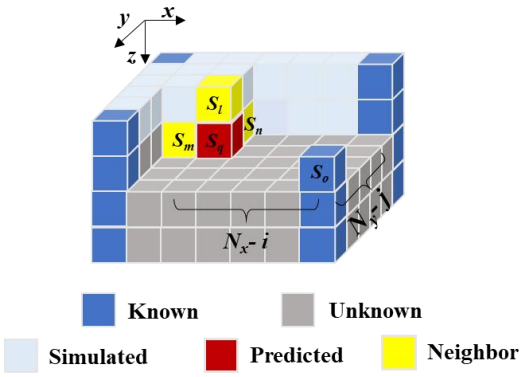

Figure 1. Schematic diagram of three dimensional modified coupled Markov chain.

$$\begin{aligned}&\Pr\{Z_{i,j,k}=S_q \mid Z_{i-1,j,k}=S_m, Z_{i,j-1,k}=S_n, Z_{i,j,k-1}=S_l, Z_{N_x,N_y,k}=S_o\}\\&=C' P^x_{mq} P^y_{nq} P^z_{lq} \Pr\{Z_{i,j,k}=S_q \mid Z_{N_x,N_y,k}=S_o\}\\&=\frac{P^x_{mq} P^y_{nq} P^z_{lq} (P^x)^{(N_x-i)}_{qo} (P^y)^{(N_y-j)}_{qo}}{\sum_f P^x_{mf} P^y_{nf} P^z_{lf} (P^x)^{(N_x-i)}_{fo} (P^y)^{(N_y-j)}_{fo}}\end{aligned} \quad (1)$$

where C' represents the normalized parameters. P^x_{mq} represents the transition probability between the state of S_q and S_m in the x-direction. The other elements P^y_{nq} and P^z_{lq} have the same meaning. N_x-i is the relative distance between the state of S_q and S_o along the x-direction and N_y-j is the relative distance between them along the y-direction. $(P^x_{mq})^{(Nx-i)}$represents the transition probability between the state of S_q and S_o in the x-direction. The other elements $(P^y_{nq})^{(Ny-j)}$ have the same meaning.

2.2. Estimate transition probability matrix

The vertical transition probability matrix can be estimated based on borehole data. Firstly, boreholes are discretized in the vertical direction to determine the rock-soil type for each cell. By counting the transitions between various soil types, the vertical transition count matrix is derived. Secondly, the vertical transition probability matrix can be calculated using Eq.(2) as follows:

$$p^v_{lk}=\frac{T^v_{lk}}{\sum_{q=1}^{u} T^v_{lq}} \quad (2)$$

where T^v_{lk} represents the number of transitions from state S_l to S_k in the vertical direction; p^v_{lk} defines the ratio of transition count matrix T^v_{lk} to the sum of the transition counts from state S_l to all other soil states. All elements of the vertical transition probability matrix can be sequentially computed.

This section illustrates the computation of the transition probability matrix using three types of geotechnical materials as an example for clarity. A method akin to the vertical transition matrix is utilized to calculate initial transition count matrices for the x-direction and y-direction. They are denoted as $T^{x'}$, $T^{y'}$ and $T^{z'}$, as shown in the Eq.(3):

$$T^{x'}=\begin{pmatrix}T^x_{11} & T^x_{12} & T^x_{13}\\ T^x_{21} & T^x_{22} & T^x_{23}\\ T^x_{31} & T^x_{32} & T^x_{33}\end{pmatrix}$$
$$T^{y'}=\begin{pmatrix}T^y_{11} & T^y_{12} & T^y_{13}\\ T^y_{21} & T^y_{22} & T^y_{23}\\ T^y_{31} & T^y_{32} & T^y_{33}\end{pmatrix} \quad (3)$$
$$T^{z'}=\begin{pmatrix}T^z_{11} & T^z_{12} & T^z_{13}\\ T^z_{21} & T^z_{22} & T^z_{23}\\ T^z_{31} & T^z_{32} & T^z_{33}\end{pmatrix}$$

Based on the Walter Principle, diagonal elements of the transition count matrix are transformed to estimate the horizontal transition probability matrix. The term txyz represents the maximum value among cell positions in $T^{x'}_{ij}$, $T^{y'}_{ij}$ and $T^{z'}_{ij}$, as illustrated below:

$$T^x=\begin{pmatrix}\frac{l_z\cot a}{l_x}T^{xyz}_{11} & T^{xyz}_{12} & T^{xyz}_{13}\\ T^{xyz}_{21} & \frac{l_z\cot a}{l_x}T^{xyz}_{22} & T^{xyz}_{23}\\ T^{xyz}_{31} & T^{xyz}_{32} & \frac{l_z\cot a}{l_x}T^{xyz}_{33}\end{pmatrix}$$
$$T^y=\begin{pmatrix}\frac{l_z\cot \beta}{l_y}T^{xyz}_{11} & T^{xyz}_{12} & T^{xyz}_{13}\\ T^{xyz}_{21} & \frac{l_z\cot \beta}{l_y}T^{xyz}_{22} & T^{xyz}_{23}\\ T^{xyz}_{31} & T^{xyz}_{32} & \frac{l_z\cot \beta}{l_y}T^{xyz}_{33}\end{pmatrix} \quad (4)$$

where *cota* is the cotangent value of the dip angle of strata in the x-direction; *cotb* is the cotangent value of the dip angle of strata in the y-direction. l_z is the simulation cell step length in the vertical z-direction; l_x is the simulation cell step length along the x-direction; And l_y is the simulation cell step length along the y-direction.

It is worth noting that the initial transition count matrix serves as an intermediate matrix offering additional information on the x-directional and y-directional transitions. The final transition probability matrices for the x-direction and y-direction will be derived through further computations. The specific calculation equation is as follows:

$$T^x_{cal}=\begin{pmatrix}\frac{n_z l_z}{n_x l_x}T^{x'}_{11} & T^{x'}_{12} & T^{x'}_{13}\\ T^{x'}_{21} & \frac{n_z l_z}{n_x l_x}T^{x'}_{22} & T^{x'}_{23}\\ T^{x'}_{31} & T^{x'}_{32} & \frac{n_z l_z}{n_x l_x}T^{x'}_{33}\end{pmatrix}$$

$$T_{cal}^{y} = \begin{pmatrix} \frac{n_z l_z}{n_y l_y} T_{11}^{y'} & T_{12}^{y'} & T_{13}^{y'} \\ T_{21}^{y'} & \frac{n_z l_z}{n_y l_y} T_{22}^{y'} & T_{23}^{y'} \\ T_{31}^{y'} & T_{32}^{y'} & \frac{n_z l_z}{n_y l_y} T_{33}^{y'} \end{pmatrix} \quad (5)$$

where $T_{ij}^{x'}$ represents the transition count from lithology i to lithology j. $T_{ij}^{y'}$ and $T_{ij}^{z'}$ have a similar meaning of representation. n_x, n_y and n_z are the number of simulation cells in the x-direction, y-direction and z-direction, respectively.

With N proposed dip angles set for both x-direction and y-direction, a total of N^2 combinations of angles are considered. Error matrices of input and output transition probabilities in x-direction and y-direction are calculated for each combination. Each angle combination needs *nsim* simulations to achieve consistent output results. The mean squared error (MSE) for the stratigraphic dip angles is calculated, with the following equation:

$$MSE = \frac{1}{nsim} Tr[(Tp^{x} - Tp_{cal}^{x})^2 + (Tp^{y} - Tp_{cal}^{y})^2] \quad (6)$$

where Tp^x is the input transition probability matrix and $T_{cal}^{x'}$ is the output transition probability matrix in x-direction. Tp^y is the input transition probability matrix and $T_{cal}^{y'}$ is the output transition probability matrix in y-direction. *diag* means the diagonal elements of a matrix; *Tr* means a trace of a matrix. The dip angle combination corresponding to the minimal MSE among the computed N^2 combinations is selected. This combination is deemed the optimal stratigraphic dip angle set for simulating the site in question. Finally, The most probable transition probability matrix can be calculated using Eq.(2).

3. Case Study

This paper applies the aforementioned method to the Ma Wan Cross-Sea Tunnel project in Shenzhen, China. Based on the geological survey data before tunnel construction, a corresponding geological model is generated, and the input borehole data are shown in Figure 2.

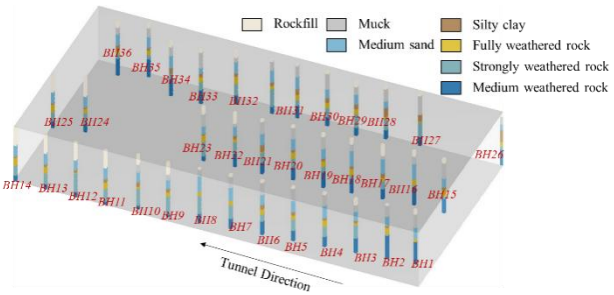

Figure 2. Borehole location and exposed strata diagram.

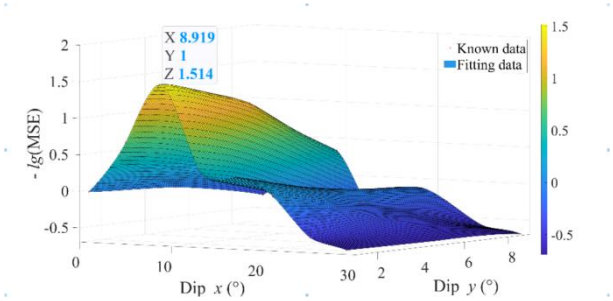

Figure 3. The optimum dip angles combined value optimizes the surface diagram.

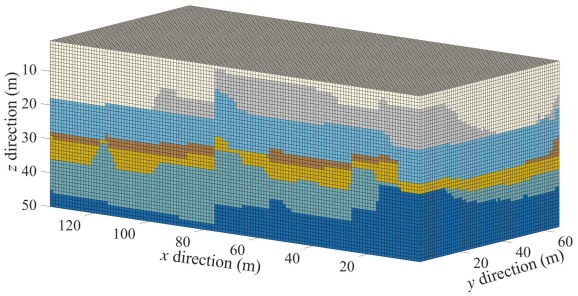

Figure 4. Most likely 3D geological map based on ICMC3D.

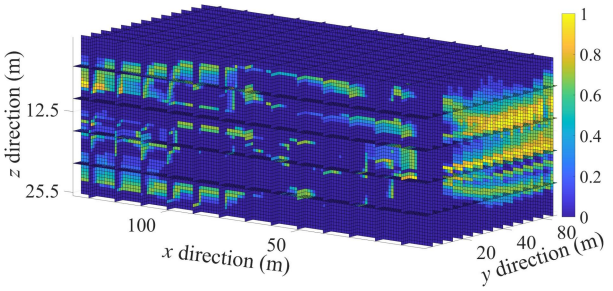

Figure 5. 3D simulation information entropy diagram.

An inverse analysis method for the formation dip angle mentioned in Section 2 is employed. The ICMC3D method is used for stochastic geological modeling. Finally, the optimal combination of dip angles is obtained through calculation, as illustrated in Figure 3. The values of the transition probability matrix for the corresponding ICMC3D model are provided in the appendix. Figure 4 shows the most probable three-dimensional stratum distribution obtained after 400 calculations based on the optimal dip angle combination. Similarly, the spatial distribution morphology of uncertainty in subsurface simulation, commonly assessed using the entropy index in probability geological modeling, is obtained through computation, as shown in Figure 5. The average entropy of the simulated space is calculated to be 0.1383.

In order to further validate the accuracy of the model and the extensibility of geological simulation, this study reduces the number of boreholes inputted into the ICMC3D model by one-third, specifically removing boreholes 1 to 14 (BH1-BH14). The spatial distribution of these specific boreholes is illustrated in Figure 2. Boreholes 15 to 36 are then inputted into the model for three-dimensional geological simulation. Boreholes 1 to 14 are utilized as a training set to evaluate the predictive accuracy of the model. Specifically, the simulated stratigraphic sequence is compared with the

corresponding borehole-exposed stratigraphic sequence. The comparative results are shown in Figure 6. The results indicate that, except for a few individual boreholes, the geological simulation accuracy can reach above 70%, with borehole 14 achieving an accuracy of 98%. This demonstrates the good accuracy of the ICMC3D model in three-dimensional geological modeling.

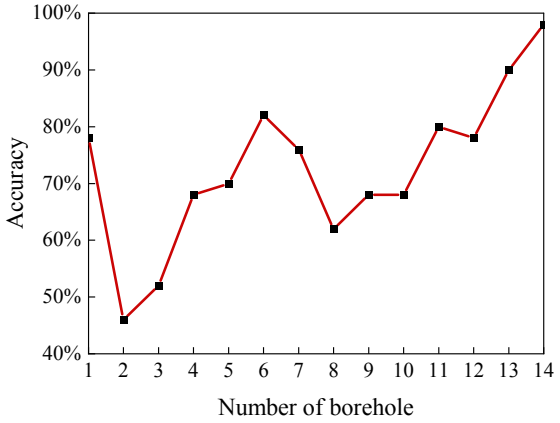

Figure 6. Different borehole accuracy curves.

4. Conclusion

This study applied the methodology to the Ma Wan Cross-Sea Tunnel project in Shenzhen, China, generating a geological model based on pre-construction surveys, with borehole data inputted as previously described. Utilizing the inverse analysis method for formation dip angles and ICMC3D for stochastic geological modeling, optimal dip angle combinations were computed. The most probable three-dimensional stratum distribution was obtained after 400 calculations. The spatial distribution morphology of uncertainty in subsurface simulation was analyzed using the entropy index, resulting in an average entropy of 0.1383 for the simulated space. To validate the model's accuracy, one-third of the borehole data were removed, and the remaining data were used for simulation. Comparative analysis showed high accuracy, with one borehole reaching 98%. This underscores the robustness of the ICMC3D model in three-dimensional geological modeling.

Acknowledgements

The authors are grateful for the financial support provided by funding agency the Young Scientists Program (2021YFF0502200), the National Natural Science Foundation of China (No. 52022070), and China National Postdoctoral Program for Innovative Talents (BX20220234).

Appendix

Table 1. The calculated x-direction transition probability matrix

x-direction transition probability matrix							
T	***1***	***2***	***3***	***4***	***5***	***6***	***7***
ype							
1	***0.997***	***0.002***	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*
2	***0.004***	***0.993***	***0.003***	*0.000*	*0.000*	*0.000*	*0.000*
3	***0.001***	***0.001***	***0.997***	***0.001***	***0.001***	*0.000*	*0.000*
4	*0.000*	*0.000*	*0.000*	***1.000***	*0.000*	*0.000*	*0.000*
5	*0.000*	*0.000*	*0.000*	*0.000*	***1.000***	*0.000*	*0.000*
6	*0.000*	*0.000*	*0.000*	*0.000*	***0.002***	***0.996***	***0.001***
7	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	***0.003***	***0.997***

Table 2. The calculated y-direction transition probability matrix

y-direction transition probability matrix							
T ype	***1***	***2***	***3***	***4***	***5***	***6***	***7***
1	***0.977***	***0.021***	***0.002***	*0.000*	*0.000*	*0.000*	*0.000*
2	***0.034***	***0.943***	***0.023***	*0.000*	*0.000*	*0.000*	*0.000*
3	***0.005***	***0.010***	***0.974***	***0.006***	***0.006***	*0.000*	*0.000*
4	*0.000*	*0.000*	***0.007***	***0.985***	***0.007***	***0.001***	*0.000*
5	*0.000*	*0.000*	***0.002***	***0.004***	***0.989***	***0.005***	*0.000*
6	*0.000*	*0.000*	*0.000*	***0.004***	***0.019***	***0.965***	***0.012***
7	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	***0.025***	***0.975***

Note: Types 1-7 indicate rockfill, muck, medium sand, silty clay, fully weathered rock, strongly weathered rock and medium weathered rock.

References

Elkateb, T., Chalaturnyk, R., & Robertson, P. K. "An overview of soil heterogeneity: quantification and implications on geotechnical field problems." Can. Geotech. J., 40(1), 1-15, 2003.

Phoon, K. K., & Kulhawy, F. H. "Characterization of geotechnical variability." Can. Geotech. J., 36, 612-624, 1999.

Yeh, C. H., Dong, J. J., Khonevisan, S., et al. "The role of geological uncertainty in a geotechnical design – A retrospective view of Freeway No. 3 Landslide in Northern Taiwan." Eng. Geol., 106233, 2021.

Zhang, W. G., Han, L., Gu, X., et al. "Tunneling and deep excavations in spatially variable soil and rock masses: A short review." Underground Space., 2020.

Elfeki, A. M. M. "Prediction of contaminant plumes (shapes, spatial moments and macrodispersion) in aquifers with insufficient geological information." J. Hydraul. Res., 44(6), 841-856, 2006.

Griffiths, D. V., Fenton, G. A., & Manoharan, N. "Bearing capacity of rough rigid strip footing on cohesive soil:

probabilistic study." J. Geotech. Geoenviron. Eng., 128(9), 743-755, 2002.

Chiles, J. P., & Delfiner, P. "Geostatistics: modeling spatial uncertainty." New York: Wiley, 1999.

Matheron, G., Beucher, H., De Fouquet, C., et al. "Conditional simulation of the geometry of fluvio-deltaic reservoirs." SPE Ann. Tech. Conf. Exhib.: Soc. Pet. Eng., 1987.

Strebelle, S. "Conditional simulation of complex geological structures using multiple-point statistics." Math. Geol., 34(1), 1-21, 2002.

Deutsch, C. V., & Journel, A. G. "GSLIB, geostatistical software library and user's guide." 2nd ed. New York: Oxford University Press, 1998.