Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly
©2022 ISGSR Organizers. *Published by* Research Publishing, Singapore.
doi: 10.3850/978-981-18-5182-7_20-009-cd

Three-Dimensional Marine Geological Modeling for Land Reclamation

Te Xiao[1], Li-Min Zhang[2], Hai-Feng Zou[3], and Yue-Bin Liu[4]

[1]The Hong Kong University of Science and Technology, Hong Kong SAR, China.
E-mail: xiaote@ust.hk
[2]The Hong Kong University of Science and Technology, Hong Kong SAR, China; HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, Shenzhen, China.
E-mail: cezhangl@ust.hk
[3]AECOM Asian Company Limited, Hong Kong SAR, China.
E-mail: Haifeng.Zou@aecom.com
[4]The Hong Kong University of Science and Technology, Hong Kong SAR, China.
E-mail: yliued@connect.ust.hk

Abstract: Understanding marine geological conditions is a primary task for land reclamation, and a three-dimensional (3-D) marine geological model is a straightforward and visualized tool for this purpose. Considering the different advantages of boreholes and geophysical surveys in terms of accuracy and scale, this study proposes a Bayesian framework to integrate the information from both boreholes and geophysical surveys to develop a more reliable 3-D marine geological model. Machine learning is applied to facilitate the prediction of geological units at unknown locations from existing boreholes. A 3-D marine geological model for the East Lantau Metropolis, Hong Kong, is created. Both geological layers of alluvium and marine deposits are thicker than 18 m, and the bedrock surface is more than 50 m below the sea level. The 3-D marine geological model can facilitate the land reclamation design in this area.

Keywords: Land reclamation; geological modeling; marine borehole; Bayesian approach; machine learning.

1 Introduction

Hong Kong is one of the most congested metropolises in the world because of the shortage of land supply. Land reclamation is a cost-effective solution to provide more developable space for coastal cities. More than 74 km^2 of land has been reclaimed in Hong Kong since 1887, about 27% of the built-up area. Towards a more sustainable future, a strategic land reclamation project named East Lantau Metropolis (CEDD, 2022) (Figure 1) is planned to construct artificial islands around Kau Yi Chau and Hei Ling Chau in the Central Waters, which will provide a total of 17 km^2 of land for housing and other purposes. Understanding marine geological conditions in this area becomes a primary task before land reclamation.

Compared with widely-used geological maps that focus on describing the horizontal distribution of geological units, geological models visualize the geological profiles along with depth additionally. It is an effective way to communicate geological conditions with engineers and the general public. In the digital age, the paradigm of the geological model is transforming from two-dimensional (Xiao et al., 2017, 2021; Shi and Wang, 2021) to three-dimensional (3-D) (Wu et al., 2005; Li et al., 2020; Liu et al., 2022), and from geological rule-based (Wu et al., 2005; Calcagno et al., 2008) to data-driven (Wang et al., 2017; Shuku and Phoon, 2021). The development of several 3-D national geological models in the United Kingdom (Mathers et al., 2014) and Singapore (Pan et al., 2019) has significantly improved the understanding of large-scale 3-D geological conditions in these regions. Notably, the majority of existing models are onshore geological models. Studies on offshore or marine geological modeling are relatively limited, owing to the high cost of marine geological investigations.

The Geotechnical Engineering Office of the Hong Kong SAR Government has compiled an extensive ground investigation database including 370,000 boreholes since the 1970s, among which approximately 70,000 are marine boreholes, providing an ideal opportunity to establish a 3-D marine geological model for Hong Kong. Although boreholes are the most direct and accurate information for geological modeling, they only provide stratification at very limited locations. By contrast, the geophysical survey can provide large-scale indirect stratification information that needs manual interpretation and may be associated with high uncertainties. For example, sonar scanning and seismic testing are widely used for bathymetry and seabed stratum identification. Figure 2 presents the seabed level and the base of offshore Quaternary formations interpreted from marine geophysical surveys (Fyfe et al., 2000). A more reliable 3-D marine geological model can be developed if the information from both boreholes and geophysical surveys can be integrated properly.

This study aims to develop a 3-D marine geological model for the land reclamation works of East Lantau Metropolis using machine learning techniques. A Bayesian framework is proposed to integrate the information from both marine boreholes and geophysical surveys. The dominant geological surfaces in this area are identified to facilitate the land reclamation design.

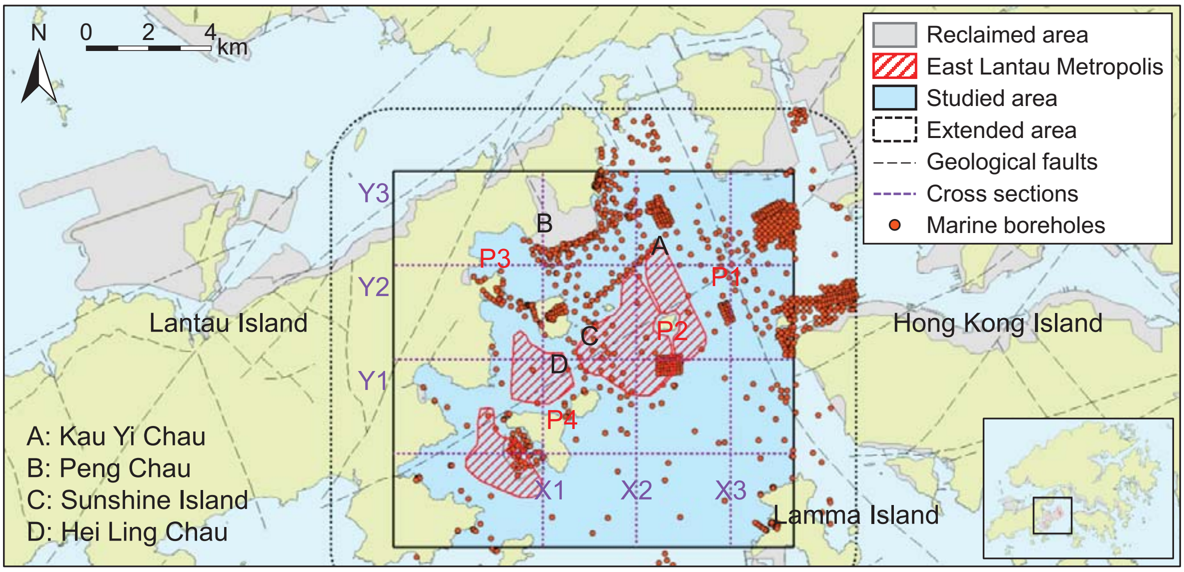

Figure 1. The East Lantau Metropolis.

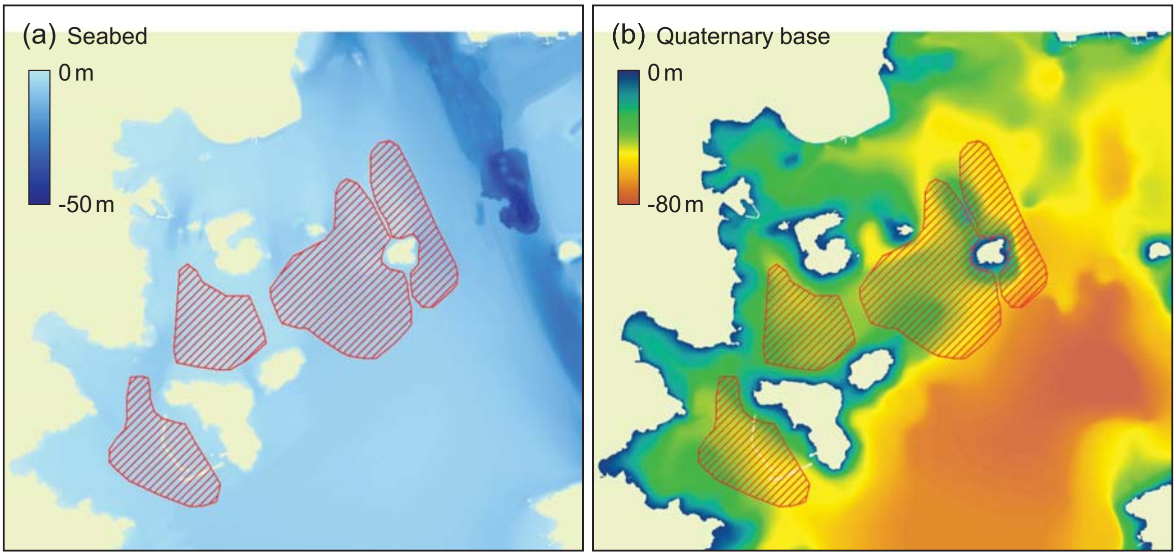

Figure 2. Results of geophysical surveys (adapted from Fyfe et al. (2000)): (a) seabed; (b) Quaternary base.

2 3-D Marine Geological Modeling

2.1 Bayesian framework

Considering the different advantages of boreholes and geophysical surveys in terms of accuracy and scale, a Bayesian framework is proposed to fuse the information from both sources in a probabilistic manner. The Quaternary base interpreted from geophysical surveys (Figure 2(b)) gives the prior distribution of different geological formations (e.g., Quaternary formation like marine deposits and alluvium, and non-Quaternary formation like weathered rock and solid rock), and the borehole information gives the likelihood of observing specific geological units (e.g., marine deposits, alluvium, weathered rock, and solid rock) at the borehole sampling locations. The seabed from geophysical surveys (Figure 2(a)) is treated as a top boundary of the marine geological model due to its relatively high accuracy.

Let $X = [x, y, z]$ be the spatial coordinates of one position in x-, y- and z-directions, and Y be the geological unit at X. In the Bayesian framework, the posterior probability of Y, $P''(Y)$, can be written as:

$$P''(Y) \propto L(Y) \times P'(Y) \tag{1}$$

where $L(Y)$ is the likelihood of Y estimated from the borehole information; and $P'(Y)$ is the prior probability of Y learned from the Quaternary base. They can be calculated, respectively, in the following subsections.

2.2 Likelihood based on boreholes

The probability of observing a specific geological unit, g, i.e., $L(Y=g)$, at an unknown location, X, can be estimated from existing boreholes using machine learning techniques, specifically the weighted k-nearest neighbor (KNN) method applied in this study as follows:

$$L(Y=g)\propto\sum_{i=1}^{k}I_g(n_i)\times w(n_i) \tag{2}$$

where k is the size of neighbors that is the k nearest known locations to X; n_i, i = 1, 2, …, k, is the ith neighbor from existing boreholes; $I_g(\cdot)$ is an indicator function: $I_g(n_i) = 1$ if n_i belongs to geological unit g; else, $I_g(n_i) = 0$; and $w(n_i)$ is the weight of n_i. While the standard KNN algorithm takes equal weight $w(n_i) = 1/k$, a distance-based weight will be taken in this study, i.e., $w(n_i) = 1/[d(X, n_i)]^2$, where $d(X, n_i)$ is the distance between X and n_i. This is similar to the commonly-used inverse distance weighting approach for spatial interpolation. With the distance-based weight, the prediction at a known location will be exactly the same as the observation.

Note that, for 3-D geological modeling, the horizontal separation of boreholes (e.g., a few hundred meters) is far greater than the vertical interval (e.g., a few meters). Such an anisotropy feature naturally selects all neighbors from a single nearest borehole, as the distances to other boreholes are much larger. This will lead to the replication of a borehole in the horizontal plane. To make full use of multiple boreholes, a simple parameter of anisotropy ratio, δ, is introduced to eliminate the anisotropic effect between the horizontal scale and vertical scale. Specifically, a transformed position $X' = [x/\delta, y/\delta, z]$ is used instead of the original X in the KNN modeling. The anisotropy ratio is somewhat similar to the ratio between horizontal and vertical scales of fluctuation in random field modeling and kriging interpolation.

2.3 Prior information from geophysical surveys

The Quaternary base identified from geophysical surveys shares the prior distribution of marine deposits and alluvium but with uncertainties. Assume that the actual Quaternary base follows a normal distribution with a mean, μ_q, at a recorded base level, z_q (see Figure 2(b)), i.e., $\mu_q = z_q$, and a standard deviation, σ_q. The probability of observing $Y = q$ (i.e., Quaternary formation group of marine deposits and alluvium) at X when $z > z_q$ can be derived as:

$$P\left(Y=q\middle|z>z_q\right)=\Phi\left(\frac{z-\mu_q}{c_v\mu_q}\right) \tag{3}$$

where $\Phi(\cdot)$ is the cumulative distribution function of standard normal distribution; and $c_v = \sigma_q/\mu_q$ is the coefficient of variation of the Quaternary base.

Since the specific geological unit is uncertain, the probability of Quaternary formation is equally divided into that of marine deposits and alluvium. Similarly, the probability of non-Quaternary formation is also equally divided into that of weathered rock and solid rock. By this means, the prior probability of Y equal to marine deposits, alluvium, weathered rock and solid rock at every location can be obtained as $[P/2, P/2, (1-P)/2, (1-P)/2]$. If the uncertainty is deemed very large, i.e., a very large c_v, Eq. (3) will give $P = 0.5$ and the probabilities of all possible geological units are equal, which is equivalent to non-informative prior and the posterior probability totally depends on the likelihood from boreholes. If the uncertainty is small, the prior will dominate the identification of Quaternary and non-Quaternary groups, while the borehole information determines the geological unit within each group.

To sum up, the developed geological modeling method involves three parameters, namely the size of neighbors, anisotropy ratio, and coefficient of variation of the Quaternary base. The size of neighbors is found to be insensitive and it will be taken as 10 in this study. The anisotropy ratio should be determined according to the borehole data, as will be introduced later. The coefficient of variation of the Quaternary base reflects the degree of belief regarding the geophysical survey interpretation. When no more information is available, a moderate uncertainty level, i.e., $c_v = 0.2$, is suggested. With the predicted posterior probabilities of all possible geological units, the most probable geological unit at any location can be determined to form a 3-D geological model.

3 Case Study of East Lantau Metropolis

3.1 Studied area and borehole data

This section applies the proposed Bayesian framework to the 3-D marine geological modeling of East Lantau Metropolis. The studied area is shown in Figure 1, with Easting between 817,250 m and 830,000 m and Northing between 809,000 m and 821,000 m in the HK1980 grid system. The elevation concerned ranges from 0 m (i.e., mean sea level) to −100 m. Two major faults are identified from the previous geological investigations.

To minimize extrapolation uncertainty, the studied area is extended for 2 km in horizontal directions to include more related marine boreholes. A total of 1216 marine boreholes are extracted, with ten geological units, including fill, marine deposits, estuarine deposits, beach deposits, debris flow deposits, alluvium, colluvium, residual soil, grade V-IV rocks, and grade III-I rocks. Since some geological units have much fewer records than other dominant units, they are combined into respective dominant units with similar properties, forming five main geological units, namely fill, marine deposits (with estuarine, beach and debris flow deposits), alluvium (with colluvium), weathered rock (with residual soil and grade V-IV rocks), and solid rock (i.e., grade III-I rocks). Figure 3 shows the extracted marine boreholes in 3-D view, whose lengths ranging from 2 to 167 m. The total borehole lengths of fill, marine deposits, alluvium, weathered rock and solid rock are about 200, 9000, 4600, 3500 and 2800 m, respectively. They are discretized into a series of data points with a vertical resolution of 0.5 m. Besides, the map of the Quaternary base (Figure 2(b)) is utilized to describe the prior distribution of marine deposits and alluvium.

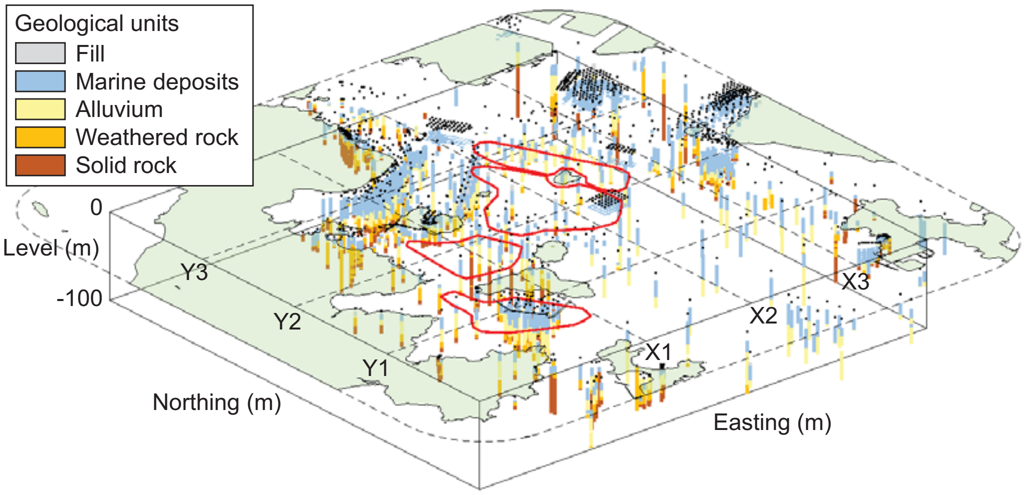

Figure 3. Marine boreholes in the studied area.

3.2 Model training and optimal anisotropy ratio

To determine an optimal anisotropy ratio in the machine learning model, the boreholes are randomly split by a ratio of 70/30 into training and test sets. Note that the partition is performed at the borehole level, rather than the point level, as the purpose is to verify model capacity on horizontal prediction from one borehole to another. If the partition is performed at the point level, only vertical prediction within a single borehole will be verified and the accuracy will easily reach as high as 0.95. In addition, due to the class imbalance issue, the F_1 score (a harmonic mean of the precision and recall) is used to measure the model performance. Macro-averaging is adopted to calculate the mean F_1 score among different classes, highlighting the same weight of infrequent classes (e.g., solid rock) and frequent classes (e.g., marine deposits).

Figure 4 compares the F_1 scores on test sets of KNN models with different anisotropy ratios. The anisotropy ratio affects the performance significantly and the highest F_1 score reaches 0.6 when the anisotropy ratio is taken as 100. This is reasonable as the horizontal separation of boreholes is in a few hundred meters and the vertical interval is in meters. Figure 4 also illustrates the impacts of prior information with different coefficients of variation of the Quaternary base, from very low uncertainty (i.e., c_v = 0.01) to very high uncertainty (i.e., c_v = 5). The F_1 score is less sensitive to the uncertainty of prior information. A lower uncertainty on the Quaternary base slightly decreases the F_1 score since the interpreted Quaternary base from geophysical surveys may conflict with the observed boreholes at some locations. Eventually, the model with an anisotropy ratio of 100 and c_v = 0.2 is used in this study to develop a 3-D marine geological model for the East Lantau Metropolis.

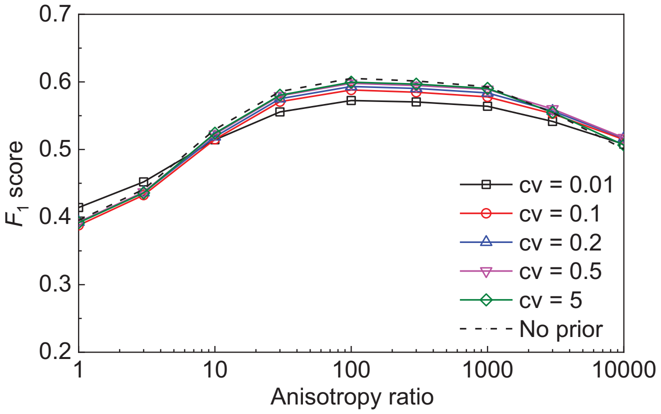

Figure 4. Comparison of model performance with different anisotropy ratios.

Table 1. Statistics on geological layers of East Lantau Metropolis.

Island	Area (km^2)	Mean water depth (m)	Mean thickness of geological layer (m)			Mean bedrock depth (m)
			Marine deposits	Alluvium	Weathered rock	
P1	3.1	8.6	19.2	26.4	11.6	65.8
P2	7.0	7.3	20.0	17.4	6.7	51.4
P3	3.1	4.3	15.9	13.6	11.1	44.9
P4	3.7	6.2	16.0	19.3	8.6	50.1
All	16.9	6.7	18.2	18.8	8.8	52.6

3.3 3-D geological model

With the optimized parameters, a 3-D marine geological model of East Lantau Metropolis (Figure 5) is created by fusing all 1216 marine boreholes and the information of the Quaternary base, with horizontal and vertical resolutions of 50 m and 1 m, respectively. The space above the seabed and within the land is out of modeling. Applying machine learning techniques, such a high-resolution model can be generated efficiently within one minute on a personal desktop computer.

Figure 6 expands the 3-D geological model along the outer profiles of the four artificial islands (i.e., P1 and P2 around Kau Yi Chau, and P3 and P4 around Hei Ling Chau), starting from the northernmost point and in a clockwise direction. The depths of bedrock (i.e., solid rock) below the two islands around Kau Yi Chau are much deeper than that around Hei Ling Chau, particularly at their southeast corners (i.e., about 20 m deeper). The mean thickness of each geological layer is summarized in Table 1. Both alluvium and marine deposits are thicker than 18 m, the weathered rock is less than 9 m thick, and the bedrock surface is more than 50 m below the sea level. The 3-D marine geological model helps design appropriate land reclamation plans in this area. The non-dredged reclamation techniques, such as deep cement mixing, will be a feasible choice in such a geological condition.

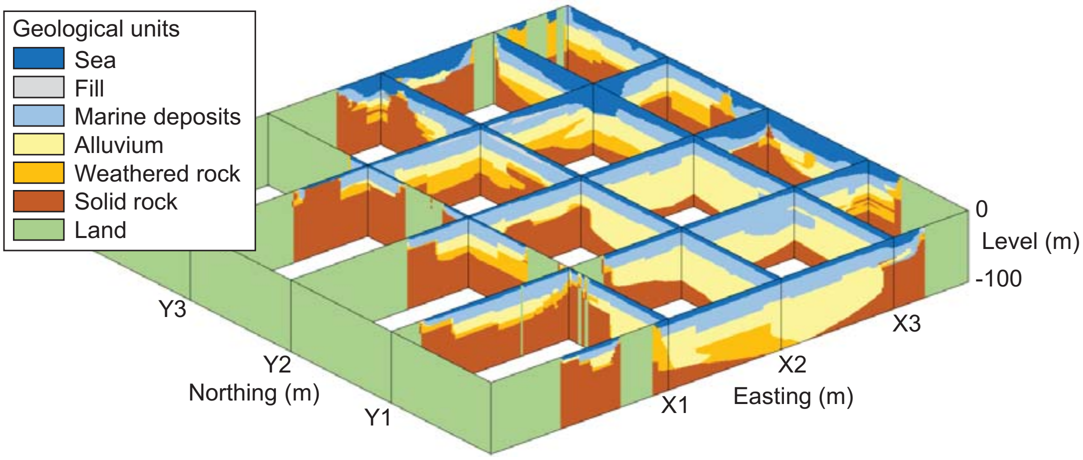

Figure 5. 3-D marine geological model of East Lantau Metropolis.

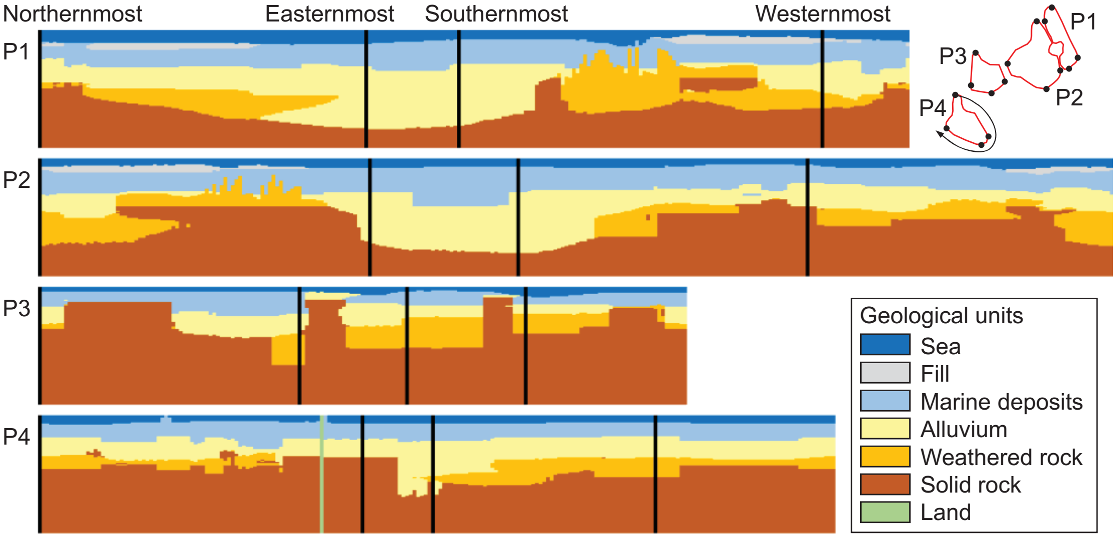

Figure 6. Geological profiles around the artificial islands of East Lantau Metropolis.

Recall that Figure 5 is produced by assuming a moderate uncertainty level (i.e., $c_v = 0.2$) of the Quaternary base obtained from geophysical surveys. By integrating boreholes with geophysical surveys, the overall mean difference between the prior Quaternary base and the updated one is 9.3 m, which is about 23% smaller than that (i.e., 12.1 m) using boreholes only without any prior information. If considering a very small uncertainty (i.e., $c_v = 0.01$) of the Quaternary base, the overall difference can be reduced to 1.9 m. Particularly among the four artificial islands, P4 on the southwest of Hei Ling Chau corresponds to the minimum difference of 3.3 m as many deep marine boreholes were drilled at the Hei Ling Chau Typhoon Shelter. Similarly, the main difference between the prior and borehole-updated Quaternary bases in those areas far away from the land with limited shallow marine boreholes that do not reach the Quaternary base, such as the region between cross sections X2 and X3 (Figure 3). Geophysical surveys are very useful to these regions, and the Bayesian approach provides a powerful tool to fuse the information from both boreholes and geophysical surveys in a sound probabilistic manner.

4 Summary and Conclusions

This study proposes a Bayesian framework to integrate the information from boreholes and geophysical surveys to develop a reliable 3-D marine geological model. The information on the Quaternary base from geophysical surveys provides the prior distribution of different geological units, and the borehole information gives the likelihood of observing specific geological units at sampling locations. Machine learning technique is applied to predict the geological units at unknown locations from existing boreholes.

The land reclamation of East Lantau Metropolis, Hong Kong, is taken as a case study. More than 1200 marine boreholes are used to develop a large-scale 3-D marine geological model. In the area of East Lantau Metropolis, the geological layers of alluvium and marine deposits are thicker than 18 m, the weathered rock is less than 9 m thick, and the bedrock surface is more than 50 m below the sea level. The depths of bedrock below the two artificial islands around Kau Yi Chau are much deeper than that around Hei Ling Chau, particularly at their southeast corners. The 3-D marine geological model helps design appropriate land reclamation plans in this area.

Acknowledgments

This work was supported by the Eunsung O&C Offshore Marine and Construction (EUNSUNG19EG01) and the Hetao Shenzhen-Hong Kong Science and Technology Innovation Cooperation Zone (HZQB-KCZYB-2020083). The authors would like to acknowledge the data support from the Geotechnical Engineering Office of the Civil Engineering and Development Department, the Government of the Hong Kong SAR.

References

Calcagno, P., Chilès, J.P., Courrioux, G., and Guillen, A. (2008). Geological modelling from field data and geological knowledge: Part I. Modelling method coupling 3D potential-field interpolation and geological rules. *Physics of the Earth and Planetary Interiors*, 171(1-4), 147-157.

Civil Engineering and Development Department (CEDD). (2022). Lantau Tomorrow Vision. *https://www.lantau.gov.hk/en/lantau-tomorrow-vision/index.html (accessed on March 5, 2022)*

Fyfe, J.A., Shaw, R., Campbell, S.D.G., Lai, K.W., and Kirk, P.A. (2000). The Quaternary Geology of Hong Kong. *Geotechnical Engineering Office, Hong Kong.*

Li, A., Jafari, N.H., and Tsai, F.T.C. (2020). Modelling and comparing 3-D soil stratigraphy using subsurface borings and cone penetrometer tests in coastal Louisiana, USA. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 14(2), 158-176.

Liu, Y.B., Xiao, T., and Zhang, L.M. (2022). 3-D Hong Kong geological modelling and management system. *Proc. HKIE Geotechnical Division 42nd Annual Seminar, Hong Kong*, p. 67-78.

Mathers, S.J., Terrington, R.L., Waters, C.N., and Leslie, A.G. (2014). GB3D–a framework for the bedrock geology of Great Britain. *Geoscience Data Journal*, 1(1), 30-42.

Pan, X., Chu, J., Aung, Z., Chiam, K., and Wu, D. (2019). 3D geological modelling: a case study for Singapore. *Proc. International Conference on Information Technology in Geo-Engineering, Guimarães, Portugal*, p. 161-167.

Shi, C., and Wang, Y. (2021). Development of subsurface geological cross-section from limited site-specific boreholes and prior geological knowledge using iterative convolution XGBoost. *Journal of Geotechnical and Geoenvironmental Engineering*, 147(9), 04021082.

Shuku, T., and Phoon, K.K. (2021). Three-dimensional subsurface modeling using geotechnical Lasso. *Computers and Geotechnics*, 133, 104068.

Wang, H., Wellmann, J.F., Li, Z., Wang, X., and Liang, R.Y. (2017). A segmentation approach for stochastic geological modeling using hidden Markov random fields. *Mathematical Geosciences*, 49(2), 145-177.

Wu, Q., Xu, H., and Zou, X. (2005). An effective method for 3D geological modeling with multi-source data integration. *Computers & Geosciences*, 31(1), 35-43.

Xiao, T., Zhang, L.M., Li, X.Y., and Li, D.Q. (2017). Probabilistic stratification modeling in geotechnical site characterization. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 3(4), 04017019.

Xiao, T., Zou, H.F., Yin, K.S., Du, Y., and Zhang, L.M. (2021). Machine learning-enhanced soil classification by integrating borehole and CPTU data with noise filtering. *Bulletin of Engineering Geology and the Environment*, 80(12), 9157-9171.