A case study on optimizing the geotechnical site investigation using Kriging method

- **Kieu Le Thuy Chung**

- **Phan Thi San Ha**

  University of Technology, VNU-HCM

- **Le Minh Son**

  flatGEO Consulting Co., Ltd., Vietnam

ABSTRACT:

One of the major tasks in geotechnical investigation is the stratigraphy distribution and the physico-mechanical properties of strata encountered in the investigation area. In order to reduce the project risks associated with uncertainties in predicting the distribution of strata (in area and in depth), a geotechnical investigation plan is usually designed with as many as possible of the boreholes. And this, in turn, increases the investigation costs. On the contrary, the owner of the project is expected to gather as much information about the subsurface soils as possible at the lowest cost. To deal with this contradiction, geotechnical engineers not only mobilize their knowledge on the investigation area and their experiences in the field of geotechnical investigation but also should be supported by geostatistical tools, especially the interpolation method of Kriging. Based on the real data taken from a geotechnical investigation project in Saigon (Vietnam), this paper will introduce two geotechnical investigation plans: a) an actual investigation plan; b) an optimized investigation plan designed with the support of Kriging method. From these two plans, the ability of Kriging in optimization of geotechnical investigation will be evaluated.

Keywords: *Kriging, Optimization, Geotechnical investigation plan, Validation, TOST.*

1. INTRODUCTION

To understand the behavior and characteristics of the ground, geotechnical engineers build up a geotechnical investigation plan in which several methods such as excavation, drilling, penetration… are applied to gather as much geological data as possible.

However, the contradiction arises during that planning. The more amount of data is collected, the higher reliability of data is and therefore it is very costly and time-consuming. On the contrary, if less amount of data is gathered then this sampling strategy will be cost-effective and time-saving but the data reliability is lower.

To overcome the contradiction *"less data – more information"*, geotechnical engineers should be supported by geostatistical tools, especially Kriging method to unveil the spatial characteristics and make use of intrinsic information in available geological data.

Kriging method investigates the spatial relationship of geological data by building up a semi-variogram (see Fig 1) which presents the relationship between semi-variances of data pairs $\gamma(h)$ with their distances lag(h). From this chart, some important parameters should be taken into account: theoretical semi-variogram model, nugget, sill and range. Readers who are interested in Kriging method could get more information from [1]-[3].
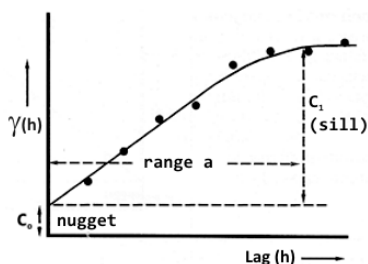
Figure 1. A typical semi-variogram

The value of data at one unsampling point is estimated by the formula:

$$\hat{z}_p = \sum_{i=1}^{n} w_i z_i, \text{ with } \sum_{i=1}^{n} w_i = 1 \qquad (1)$$

where

$\hat{z}_p$ the predicted value at point p

w_i weight factor at point i

z_i the measured value at point i

One of the advantages of Kriging method is that it estimates not only the data value but also the magnitude of estimation error. The error variance of the interpolation is calculated by the following formula in which the values of w_i and λ are derived from the semi-variogram of data points:

$$\sigma_p^2 = \sum_{i=1}^{n} \left[w_i \times \gamma(h_{pi}) \right] + \lambda \qquad (2)$$

where

σ_p^2 the error variance for the interpolation at point p

h_{pi} distance between point p and point i

$\gamma(h_{pi})$ the value of the semi-variogram model for the distance h_{pi}

w_i weight factor at point i

λ Lagrange multiplier used to minimize possible estimation error

With the support of the map of error variance, geotechnical engineers could design an optimized sampling network in which the selection of the borehole locations for the next stage is based on the error map from the current investigation sampling (see [2], [3]).

This paper is a pilot study to demonstrate the optimization of geotechnical investigation plan by using Kriging method for the top layer in the study area.

The study area is located in the south of Saigon, Vietnam (see Fig. 2). In an actual geotechnical investigation plan, the project owner had conducted 41 boreholes and 20 CPTu tests in order to understand the geological conditions of the study area.

According to boring logs of 41 boreholes, the top layer is a soft soil layer whose bottom elevation values vary from -32.2m to -16.5m. Meanwhile, the elevation of borehole collars vary from +0.3m to +1.6m.

Figure 2. Location of the study area

DATA SET

Available data of the bottom elevation of the top layer is divided into two sets: borehole data is used for prediction by Kriging method and CPTu data is for the validation. Fig. 3 shows the location of boreholes and CPTu.
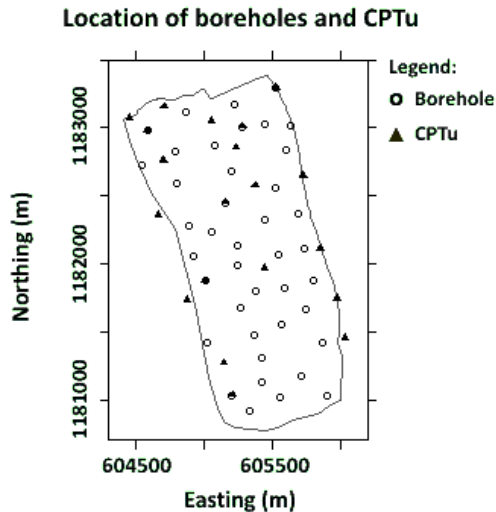

Figure 3. Location of boreholes and CPTu

2. METHODOLOGY

To optimize the sampling network, the geotechnical investigation is undertaken in multi stages. The location of new boreholes in the current stage is decided from the error map of the previous stage (see Fig. 4).

The multistage investigation process will be terminated when the error variance at the current stage is not higher than the expected error value or the difference between two prediction maps from two consecutive stages is not statistically significant. A statistical method TOST (Two One-Sided t-Test) is applied to test the equivalence of two prediction maps from two stages. The null hypothesis in TOST equivalence test is that the two prediction maps are totally different. And vice versa, the alternative hypothesis is that the two prediction maps are similar. More information about the TOST equivalence method should be referred in [4], [5].

The TOST method compares two group means and their two one-sided α-level confidence intervals by comparing them to a predefined equivalence limit ($\pm\varepsilon$). An indifference region of $\varepsilon = \pm25\%$ of the standard deviation is commonly used. If the indifference region completely encompasses the confidence interval then the two populations are deemed significantly similar (the null hypothesis is rejected). If not, then the null hypothesis is not rejected. In this case, there is a lack of sufficient evidence to conclude that the two prediction maps are similar. At each interpolation stage, the semi-variogram models will be validated with the independent data set of 20 CPTu holes.
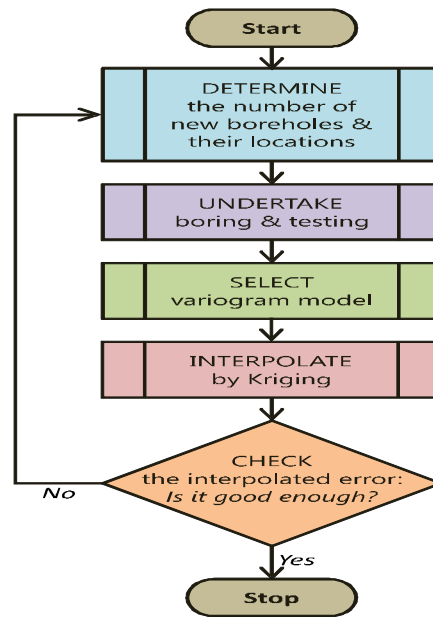

Figure 4. Flowchart of the multistage geotechnical investigation

3. RESULTS

Interpolation of stage 1

At the initial stage, 10 boreholes are selected randomly to cover the whole area (see Fig. 5). Their coordinates and the corresponding bottom elevations of the top layer are presented in Table 4.

The semi-variogram of stage 1 and its parameters are displayed in Fig. 6. Apparently, the quality of the semi-variogram is not good enough and results in the high error variance in the error map (see Fig. 9).

Interpolation of stage 2

Based on the error map of stage 1, 10 new boreholes which locate at the positions of high error on the error map are added for the interpolation of stage 2 (see Fig. 5). The semi-variogram of 20 boreholes used in

stage 2 is presented in Fig. 6. The prediction map and error map of stage 2 are presented in Fig. 8 and Fig. 9, respectively.

The variation of estimated error values of stage 2 is narrower than that of stage 1. TOST test for the difference between two prediction maps of stage 1 and stage 2 proves that there is no strong evidence to conclude that two prediction maps are similar (the null hypothesis could not be rejected).

Table 1. Equivalence test for maps of stage 1 and 2

Mean of difference	α-level	TOST result	Statistical conclusion
-0.4183	0.05	not rejected	Two maps are different

Interpolation of stage 3

Twelve new boreholes added in stage 3 are based on the result of the error map of stage 2 (see Fig. 5). The result of the interpolation and error variance are displayed in Fig. 8 and Fig. 9, respectively.

The semi-variogram and its parameters of 32 boreholes are in Fig. 7. The theoretical semi-variogram is fitted well with the experimental semi-variogram up to the distance of 1200m. Beyond this range, the difference between theoretical and experimental semi-variograms is gradually increased. This behavior could be caused by a spatial trend imposed on the bottom elevation data of the top layer.

TOST test for the difference between stage 2 and stage 3 proves that there is no strong evidence to conclude that two prediction maps are similar.

Table 2. Equivalence test for maps of stage 2 and 3

Mean of difference	α-level	TOST result	Statistical conclusion
-0.3085	0.05	not rejected	Two maps are different

Interpolation of stage 4

All boreholes are used in stage 4 which is similar to the actual investigation plan carried out by the project owner of this project (see Fig. 5). The result of interpolation and error variance are displayed in Fig. 8 and Fig. 9, respectively. The semi-variogram and its parameters are presented in Fig. 7.

Table 3. Equivalence test for maps of stage 3 and 4

Mean of difference	α-level	TOST result	Statistical conclusion
-0.1230	0.05	rejected	Two maps are similar

TOST test for the difference between stage 3 and stage 4 confirms that the difference between two prediction maps is not statistically significant. Therefore, it is reasonable to stop the investigation at stage 3 with 32 boreholes instead of at stage 4 with 41 boreholes.

4. VALIDATION OF THE SEMI-VARIOGRAM MODELS

The reliability of the semi-variogram models will be validated by using CPTu data as the independent data set. The validation results presented in Table 5 show that the quality of the interpolation gradually increase from stage 1 to stage 3. The interpolated values at CPTu holes using the semi-variogram models of stage 3 and stage 4 are not significantly different.

5. CONCLUSIONS

The approach of multistage investigation plan together with TOST equivalence test shows that the geotechnical investigation plan with 32 boreholes could reveal the same information as the actual investigation plan with 41 boreholes conducted by the project owner of the project.

Geostatistics tools, especially Kriging method are really helpful to unveil the spatial characteristics of geological data and to optimize the sampling network.

Using geostatistics tools or Kriging method will reduce the uncertainty and increase the reliability of information derived from available geological data.

The location of measurements should be based on the predicted error map – an added value of the Kriging interpolation method. The location of sampling in the next stage should be placed at the high error regions from the error map. The investigation process could be ended whenever the estimated error is lower than the expected error or two prediction maps of two consecutive stages are not different.

Table 4. Bottom elevation of the top layer from 41 boreholes in the study area

No.	Stage	ID	Easting (m)	Northing (m)	Ground elevation (m)	Bottom elevation (m)	No.	Stage	ID	Easting (m)	Northing (m)	Ground elevation (m)	Bottom elevation (m)
1	1	BH01	604583	1182980	1.20	-19.60	21	3	BH04	604786	1182820	1.20	-20.20
2	1	BH09	605205	1182680	0.40	-23.00	22	3	BH05	605073	1182870	0.80	-20.20
3	1	BH12	604887	1182280	0.85	-22.75	23	3	BH11	605154	1182450	1.30	-19.20
4	1	BH25	605527	1183290	1.25	-17.35	24	3	BH17	605005	1181870	1.62	-24.78
5	1	BH27	605686	1182370	1.00	-19.30	25	3	BH20	605522	1182560	0.76	-19.94
6	1	BH30	605383	1181800	0.75	-26.85	26	3	BH21	605448	1182320	1.54	-19.86
7	1	BH42	605873	1181420	1.23	-24.87	27	3	BH24	605224	1183170	0.60	-17.00
8	1	BH43	605027	1181420	0.80	-26.50	28	3	BH26	605635	1183020	1.30	-16.50
9	1	BH48	605339	1180920	0.53	-25.47	29	3	BH28	605549	1182070	1.25	-25.25
10	1	BH63	605902	1181030	1.10	-32.20	30	3	BH31	605594	1181830	1.20	-22.10
11	2	BH02	604869	1183120	1.50	-19.80	31	3	BH40	605267	1181680	1.20	-25.10
12	2	BH03	604545	1182720	1.46	-22.04	32	3	BH61	605568	1181550	1.28	-23.72
13	2	BH06	604800	1182600	1.00	-21.80	33	4	BH13	605054	1182230	0.30	-22.20
14	2	BH16	605244	1182130	1.42	-23.08	34	4	BH15	604919	1182060	1.40	-22.30
15	2	BH18	605283	1183010	1.25	-17.35	35	4	BH22	605245	1181980	0.70	-22.60
16	2	BH32	605800	1181880	1.10	-21.20	36	4	BH29	605732	1182110	1.00	-20.50
17	2	BH37	605601	1182830	1.22	-17.58	37	4	BH34	605745	1181670	0.60	-19.70
18	2	BH44	605419	1181310	0.67	-27.33	38	4	BH36	605442	1183020	1.47	-17.03
19	2	BH45	605200	1181030	1.17	-28.13	39	4	BH41	605369	1181480	1.28	-27.52
20	2	BH47	605709	1181170	0.70	-25.30	40	4	BH46	605421	1181130	0.78	-27.52
							41	4	BH49	605562	1181020	1.21	-26.69

Table 5. Results of the validation with CPT data (independent data set)

No.	ID	Easting	Northing	Measured value	Predicted value			
					Stage 1	Stage 2	Stage 3	Stage 4
1	CPTu01	604453	1183072	-18.44	-19.58	-19.82	-19.87	-19.88
2	CPTu02	604583	1182977	-18.75	-19.60	-19.60	-19.60	-19.60
3	CPTu03	604704	1183155	-19.10	-19.42	-19.51	-19.51	-19.51
4	CPTu04	605527	1183292	-17.45	-17.35	-17.35	-17.35	-17.35
5	CPTu05	604701	1182757	-22.08	-20.88	-21.35	-21.02	-21.03
6	CPTu06	605055	1183052	-19.62	-20.00	-19.17	-18.85	-18.85
7	CPTu07	605283	1183009	-17.30	-19.85	-17.35	-17.35	-17.35
8	CPTu08	605231	1182855	-19.28	-21.08	-19.99	-19.98	-19.97
9	CPTu09	604663	1182357	-22.02	-22.25	-22.83	-22.90	-22.72
10	CPTu10	605154	1182454	-20.45	-22.77	-22.61	-19.26	-19.26
11	CPTu11	605374	1182573	-20.00	-21.61	-21.22	-20.44	-20.47
12	CPTu12	605729	1182648	-18.54	-18.94	-18.24	-18.31	-18.27
13	CPTu13	604879	1181732	-25.18	-25.40	-25.49	-25.35	-25.10
14	CPTu14	605005	1181874	-23.98	-25.25	-25.18	-24.78	-24.78
15	CPTu15	605447	1181965	-23.30	-24.83	-24.15	-25.10	-24.76
16	CPTu16	605853	1182113	-20.40	-21.56	-20.28	-20.86	-20.05
17	CPTu17	605977	1181742	-24.04	-23.81	-22.50	-22.45	-21.45
18	CPTu18	606035	1181452	-25.18	-25.52	-25.21	-25.34	-25.00
19	CPTu19	605147	1181270	-27.44	-26.45	-27.35	-27.29	-27.60
20	CPTu20	605206	1181035	-27.83	-25.74	-28.05	-28.05	-28.07

No.	ID	Easting	Northing	Measured value	Predicted value			
					Stage 1	Stage 2	Stage 3	Stage 4

+   *Boreholes used at the previous stage*
o   *New boreholes added for the current stage*
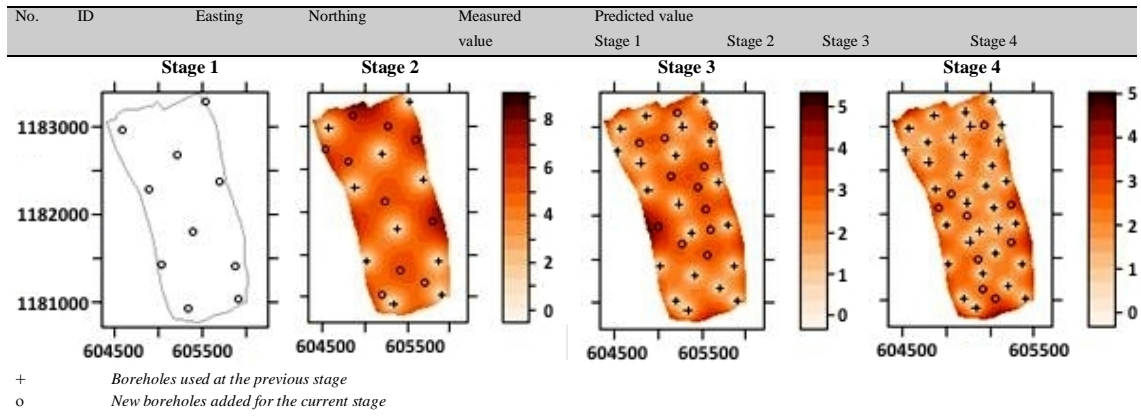
Figure 5. Borehole locations at four stages

Model	Circular	Circular
Nugget	0	0
Sill	25	35
Range	2000	3600
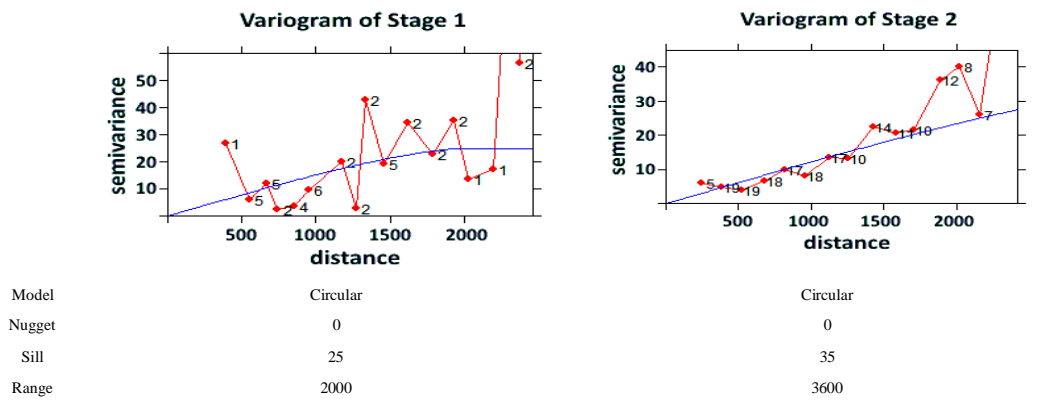
Figure 6. Semi-variogram of stage 1 (left) and stage 2 (right)

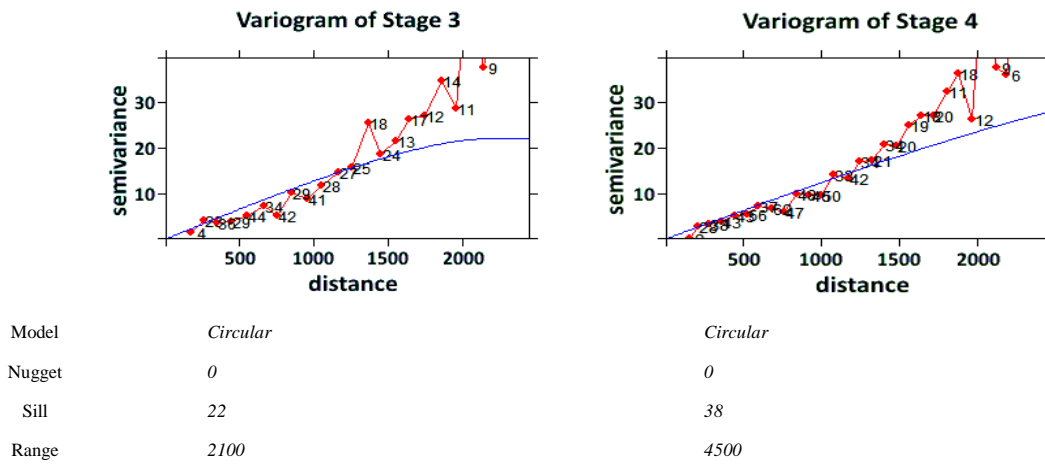| Model	*Circular*	*Circular*
Nugget	*0*	*0*
Sill	*22*	*38*
Range	*2100*	*4500*

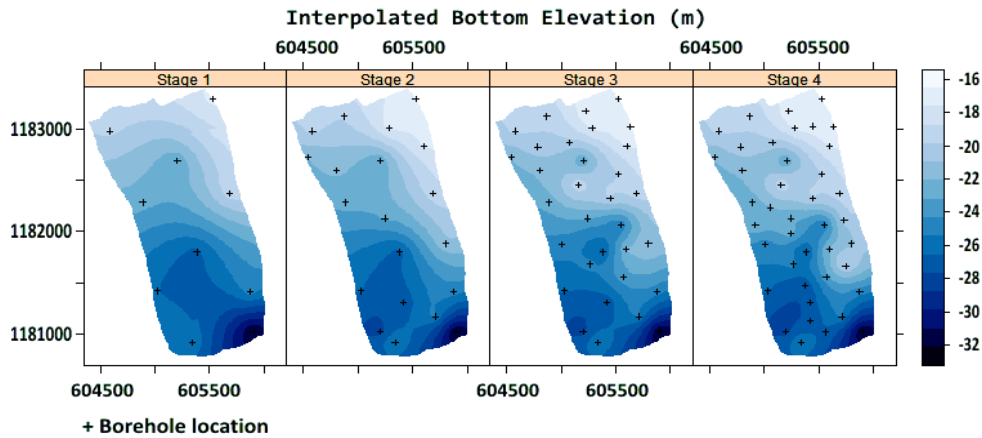**Figure 7.** Semi-variogram of stage 3 (left) and stage 4 (right)

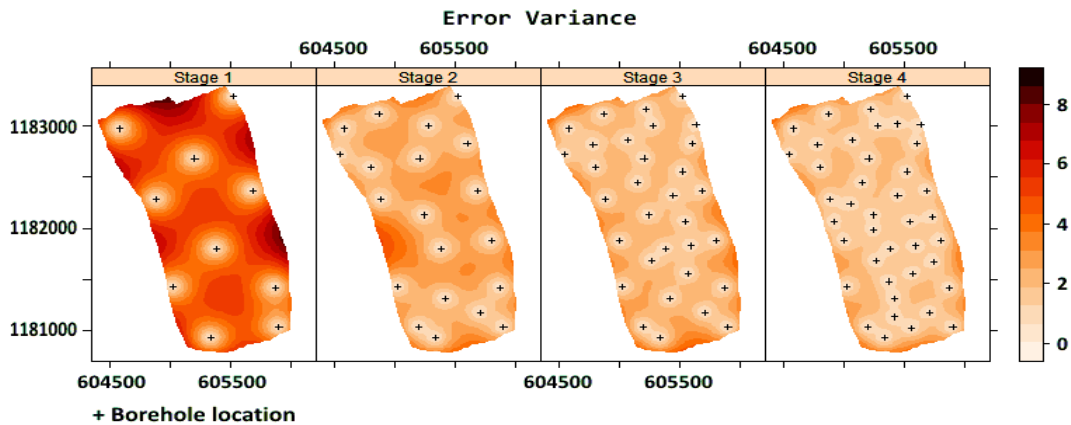**Figure 8.** Interpolated map of four stages

Figure 9. Error map of four stages

Ví dụ tối ưu hóa phương án khảo sát địa Kỹ thuật bằng phương pháp nội suy Kriging

- **Kiều Lê Thủy Chung**

- **Phan Thị San Hà**

  Trường Đại học Bách Khoa, ĐHQG-HCM

- **Lê Minh Sơn**

  Công ty TNHH Tư Vấn Địa Chất Phẳng, Việt Nam

TÓM TẮT:

Một trong những nhiệm vụ chính của công tác khảo sát địa kỹ thuật là xác định các lớp đất có mặt trong khu vực khảo sát cũng như đặc tính cơ lý của các lớp đất đó. Để giảm thiểu sai số trong *việc dự đoán sự phân bố các lớp đất (theo diện và theo độ sâu), một phương án khảo sát địa kỹ thuật thường được thiết kế với rất nhiều hố khoan, dẫn đến việc tăng chi phí khảo sát. Ngược lại, chủ đầu*

tư luôn muốn có được càng nhiều thông tin về khu vực khảo sát càng tốt nhưng với chi phí thấp nhất. Để giải quyết nghịch lý này, kỹ sư địa kỹ thuật không chỉ vận dụng toàn bộ hiểu biết về khu vực khảo sát cũng như kinh nghiệm khảo sát địa kỹ thuật ngoài hiện trường, mà còn cần phải được hỗ trợ bởi các công cụ địa thống kê, đặc biệt là phương pháp nội suy Kriging. Dựa trên tập dữ liệu thực tế của một dự án khảo sát địa kỹ thuật ở Sài Gòn (Việt Nam), bài báo sẽ giới thiệu hai phương án khảo sát địa kỹ thuật: a) phương án khảo sát đã được sử dụng trong thực tế; b) phương án khảo sát được tối ưu bằng phương pháp nội suy Kriging. Từ đó, bài báo đánh giá khả năng sử dụng phương pháp nội suy Kriging nhằm tối ưu hóa phương án khảo sát địa kỹ thuật.

Từ khóa: Kriging, tối ưu hóa, phương án khảo sát địa kỹ thuật, kiểm định, TOST.

TÀI LIỆU THAM KHẢO

[1]. Burrough P.A., McDonnell R.A., "Principles of Geographical Information System", Oxford University Press, 1998.

[2]. Houlding S.W., "Practical Geostatistics: Modelling and Special Analysis", Springer, 2000.

[3]. Kitanidis P.K., "Introduction to Geostatistics: Applications to Hydrogeology", Cambridge University Press, 1999.

[4]. Robinson A.P., Froese R.E., "Model Validation Using Equivalence Tests", Ecological Modelling, Vol. 176, No. 3-4, pp. 349-358, 2004.

[5]. Wellek S., "Testing Statistical Hypotheses of Equivalence", Chapman & Hall/CRC, 2010.