Discussion about correlation distance of loess parameter

Yang Yong¹², Zhang Haifeng¹, Zhao Zhihai¹, Fan Wen², Jiang Guichun¹

¹ College of Geology Engineering and Geometrics in Chang an university 710054;
² Northwest Research Institute of Engineering Investigation and Design 710003

Abstract. The correlation distance of soil parameter is a basic indicator which use to describe the spatial variability and autocorrelation of soil. It is of great significance to study the rule of its change with different calculate methods and test samples. Based on the Xi'an loess, the problem of data collection which used to set up the soil random field model was discussed in this paper. Using borehole CPT data as samples, correlation distance of different loess layers were calculated. The differences of correlation distance like between with different method and different indexes was discussed, And questions some the change law of this difference with the space H was discussed.

1. Introduction

In geotechnical engineering design and calculation, The classical statistics methods think of geotechnical test or test parameters as mutually independent random variables, the variance estimators of soil parameters obtained by using the method of statistical parameter, it ignores the autocorrelation properties of rock and soil. In 1977, Vanmarcke proposed a random field model of soil profile, the fluctuation range (scale of fluctuation) this concept to describe the spatial variability of soil parameters and the correlation of the correlation distance, namely commonly referred to as soil. In theory, as a basic attribute of soil spatial variability and correlation of expression, it should be a constant. Therefore, the variance reduction function should be a fixed value in a certain spatial range. That is to say, for the soil of the same site, under a certain selected computational space H, when the different calculation methods and physical mechanical indexes are taken as samples, and the calculated correlation distances should be consistent.

Then, for the Xi'an loess, what's the difference in the calculated soil correlation distance under the same calculation method? I had to test data of loess in Xi'an as a sample, the one-dimensional Xi'an loess soil parameters of MRF model established, to test the applicability of the model, it discusses the statistical distribution characteristics of the calculation of correlation distance in different landforms of Xi'an loess.

This paper takes Xi'an loess drilling CPT data of QC resistance and friction resistance of FS as samples (sample spacing is z0=0.1m), using recursive method and spatial correlation function method (used to calculate spatial range average h=3.5m) quaternary early Pleistocene aeolian loess paleosol layer, residual update and quaternary soil related from the Aeolian Loess Layer were calculated. At the same time, it discusses the data collected to establish soil distance field model, we use the different calculation methods and different test index as samples to analyze so that we can calculate the correlation distance of soil property differences, and this difference with the space average is used to calculate the changing trend of the H problem.

2. Discussion on the interval of sample data acquisition

The establishment of soil random field model is scientific and reasonable to test data as the basis, the
kind of soil index which solve the test method, test sample capacity and test density or spacing issues such as how much is required to consider the first original data acquisition. Zhu Dengfeng, Gao Dazhao, and other research think: the main factors affecting the spatial average variance calculation result is the sample size of N, and the sample size depends on the N used to calculate the average thickness (i.e. the space above the concept of H) and the sampling distance. According to the definition of the correlation distance, if the sampling interval is Z0, then the space range h is required to be infinite. On the other hand, if the space range h is certain, the delta Z0 must be guaranteed to be infinitely small. It is pointed out that the literature, the smaller the distance between the sampling data, more fully, so as to determine the correlation distance is closer to the true value. However, the Vanmarcke, has pointed out that "in order to avoid the related properties between different samples (so that the information collected to become wasteful), choose the sampling interval is greater than the correlation distance". Wilson.H. Tang, also said: if the sample is very close, the correlation distance of soil in between the experimental observations is the assumption of statistical independence will no longer set up. Huang Chuanzhi, Sun Wanhe, and other people have been raising doubts about this view, they think this view confuses the independence between the sample data (i.e. independent random test) the correlation between random variables and random fields with certain distance of (or independent) concept. The former refers to the probability of random phenomena, the probability of each test result is not affected by other tests, it is related to the test process, also known as the independence of the test. The latter is a property of the thing itself, that is the correlation between the soil properties, and we use the correlation function to express it. Li Xiaoyong, Cheng qiang and other persons think: when we calculate the correlation distance Delta, the sample data collection interval is approximately equal to the correlation distance of soil when the reliable results can be obtained.

The author believes that the independent nature of the random test (i.e., each in situ test or sampling or other geotechnical tests) is adequately guaranteed in the project and is not the focus of our study. On the contrary, the correlation and independence between soil indexes are the focus of our research. These two concepts should be distinguished. Theoretically, it should be that the smaller the test spacing, the more likely to reflect the correlation between soil properties, the more conducive to the study of the correlation between soil parameters. Therefore, the smaller the sampling interval, the better, but in actual engineering, our geotechnical test data is always discrete, and it is impossible to have completely continuous test datas. Dense sampling frequency is good, but if the sampling interval is too dense, it will inevitably lead to unnecessary waste, and do not meet the economic principles. But if the sampling interval is too large, it can not fully reflect the correlation between soil properties. How large is the sampling interval? The author thinks that the sampling distance is a little smaller than the correlation distance. In this way, the correlation between the samples is guaranteed and the economic rationality is guaranteed.

Considering the small test distance of CPT test data, the general test and data acquisition interval are 0.1m. According to the results of the study show it: for Xi’an loess, this data is generally less than the correlation distance. Therefore, it is reasonable and feasible to analyze the soil profile random field model by using CPT data with 0.1M test spacing. Based on this data, previous researchers have done a lot of research on the soil random field model, which is the source of the sample.

3. The influence of different calculation methods on calculation results

How to calculate the fluctuation range of soil parameters accurately is very important, and a lot of work has been done by domestic and foreign scholars. Vanmarcke (1977) puts forward three basic methods for calculating soil correlation distance, that is, space recurrence averaging method, correlation function method and average zero span method. The domestic scholars have extensively discussed the advantages and disadvantages, applicability and differences among them, and put forward some new methods.

Yan Shuwang, Jia Xiaoli and other scholars, respectively using correlation function method, recursive method, zero span space average method and statistical simulation method, employ 30 CPT borehole data as the sample to calculate the correlation distance of soil layers. The results show that the results
calculated by the statistical simulation method are too large and the variability is great, while the results of other three methods have little difference, and they are relatively stable. Yan Shuwang and Zhu Hongxia come to a conclusion through theoretical analysis: the correlation distance calculated by recursive space method and that calculated by correlation function method should be consistent. Besides, they employ the pier rear land static penetration test value (QC) of Tanggu petrochemicalCone resistancein the south of Tianjin as a sample, using recursive method and spatial correlation function method to calculate the reduction function of variance and correlation distance (sampling distance of 0.05 meters, the sample size of 121). The results show that the variance reduction function calculated by the two methods is basically consistent with the correlation distance. But some scholars think that the correlation distance calculated correlation function method is bigger than that calculated by recursive method and the ratio is around 2.

This paper mainly discusses the differences between the two widely accepted methods, namely the spatial recurrence method and the correlation function method when they are used to calculate the correlation distance of Xi’an loess soil. Using the CPT data of Xi’an loess drilling end and side friction resistance QC FS as samples (sampling spacing is \(z_0=0.1m\)), I respectively use recursive method and spatial correlation function method (used to calculate spatial range average \(h=3.5m\)) to calculate the correlation distance of it. The correlation distances of statistical calculation are listed in table 1.

Table 1  Statistical table of correlation distance

| Calculation method | Calculate the index | Soil layer | Number of samples | maximum | Minimum |
|-------------------|---------------------|------------|-------------------|---------|---------|
| A Recursive space method | End resistance \(q_e\) | \(Q_3\)Loess | 176 | 0.723 | 0.118 |
|                     | \(Q_3\)Paleosol      | 176        | 0.686             | 0.089  |
|                     | \(Q_3\)Loess         | 128        | 0.702             | 0.126  |
|                     | \(Q_2\)Loess         | 176        | 0.702             | 0.126  |
|                     | \(Q_3\)Paleosol      | 176        | 0.686             | 0.089  |
|                     | \(Q_3\)Loess         | 176        | 0.702             | 0.126  |
|                     | \(Q_2\)Loess         | 176        | 0.702             | 0.126  |
| B Correlation function method | End resistance \(q_e\) | \(Q_3\)Loess | 176 | 0.665 | 0.100 |
|                     | \(Q_3\)Paleosol      | 176        | 0.686             | 0.095  |
|                     | \(Q_3\)Loess         | 176        | 0.702             | 0.136  |
|                     | \(Q_3\)Paleosol      | 176        | 0.665             | 0.100  |
|                     | \(Q_3\)Loess         | 176        | 0.702             | 0.136  |
|                     | \(Q_3\)Paleosol      | 176        | 0.665             | 0.100  |
|                     | \(Q_3\)Loess         | 176        | 0.702             | 0.136  |
| mean               | Standard deviation  | Coefficient of variation | Mean ratio \(q_e / f_e\) | Mean ratio \(B / A\) |
| 0.343              | 0.134              | 0.390         | 0.855          | 1.271  |
| 0.265              | 0.116              | 0.438         | 0.813          | 1.091  |
| 0.314              | 0.129              | 0.411         | 0.863          | 1.213  |
| 0.401              | 0.170              | 0.424         | 0.826          | 1.195  |
| 0.326              | 0.131              | 0.402         | 0.863          | 1.113  |
| 0.364              | 0.163              | 0.446         | 0.910          | 1.266  |
| 0.436              | 0.203              | 0.466         | 0.910          | 1.266  |
| 0.289              | 0.128              | 0.443         | 0.796          | 1.133  |
| 0.381              | 0.184              | 0.482         | 0.826          | 1.091  |
| 0.479              | 0.201              | 0.419         | 0.826          | 1.091  |
| 0.363              | 0.154              | 0.423         | 0.826          | 1.091  |
| 0.461              | 0.236              | 0.512         | 0.826          | 1.091  |

According to table 1, it can be seen that the calculated correlation distance between the Xi’an loess layers (using the average data including spatial range of \(h=3.5m\) and the experimental spacing \(z_0=0.1m\)) is generally between 0.1m and 1.0m. The average value calculated by recursive method is generally between 0.3m and 0.4m and the average value calculated by correlation function method is generally between 0.3m to 0.5m. The result of correlation function method is generally larger than that
of recursive space method, and the ratio between the two is about 1.1~1.3.

The difference can be explained as follows: When we calculate the correlation distance using the correlation function method, we need to determine the coefficient of correlation function by curve fitting, and any form of function will actually have soil fitting error with actual relevance of soil properties. At the same time, when the correlation function is fitted, there is a large amount of the back part discarded, making the calculation distorted. The spatial recursive averaging method is using the direct test data. From the point of the physical meaning of variance reduction, the standard deviation of the spatial mean is gradually deduced by gradually increasing the spatial mean range. The method is clear and reasonable. It avoids the error caused by fitting error and data rejection, and the calculation result is relatively true and accurate.

In the following study, this difference is used to calculate the average spatial extent of H changes. Taking 10 sets of CPT data in Xi’an Loess loess Q3 layer end resistance QC and side resistance FS as samples (sampling spacing is z0=0.1m) and standardizing data. On this basis, we firstly need prove that the soil profile is in accord with the second airport parameters with the test of the stationary rows and ergodic properties. And then the recursive space method and the correlation function method are used to calculate the correlation distance of H in different spatial range. Compare the difference of correlation distances in different spatial ranges of H calculated by the two methods, as shown in figure 1-3.

![Fig. 1 Comparison of correlation distance calculated](image1)

![Fig. 2 Comparison of correlation distance calculated by two methods with $f_s$ index by two methods with $q_c$ index](image2)
As shown in figure 1-3. In the following study, this difference is used to calculate the average spatial extent of H changes. With 10 sets of CPT data in Xi'an Loess loess Q3 layer end resistance QC and side resistance FS samples (sampling spacing is $z_0=0.1m$), on standardized data, firstly the stationarity and ergodicity of the test proved that the soil profile is in accord with the second airport parameters with stationary random field conditions. Then the recursive space method and the correlation function method are used to calculate the correlation distance of H in different spatial range. Compare the difference of correlation distances calculated by the two methods in different spatial ranges of H.

4. The influence of different calculation indexes on results

There have been a lot of discussions about different experimental indexes, namely, the influence of sample data types of soil random field model on soil properties. Gao Dazhao who in the recursive space method of correlation distance of soil in Shanghai area were calculated after that: for the same soil, the correlation distance difference calculated by different soil parameters in general: in situ test results are slightly larger than the indoor test, and that this difference is related to the undisturbed soil and that disturbance generated damage caused by. Li Xiaoyong, in Taiyuan silty clay and silt and silty clay from Hangzhou and clay layer after the similar results were obtained by calculation.

For the Xi'an loess, what is the difference between the calculated distances of the soil random field model and different test indexes. According to table 1 of this paper, it can be seen that the calculated distance of Xi'an loess in each layer (used to calculate the average spatial range of $h=3.5m$, the experimental spacing $z_0=0.1m$) is generally between 0.1M and 1.0m. Two methods of using the calculation results of different indicators of the difference is not big, but there are general rules: using QC as an index of the sample results are FS as a kind of index calculation results, the ratio of the two in general is about 0.8~0.9.

This difference can be explained as follows: first, as the experimental object of loess, which contains a lot of calcareous nodules, its strength is hight, the mechanical index is ularger than the loess. Therefore, in the process of collecting CPT data, the distribution points of these calcium nodules become some data points with greater variability, which increases the variability of statistical data. This variability is different for the side resistance FS and the terminal blocking QC for CPT data acquisition, It is shown in the cone penetration probe that the sensitivity of soil resistance, lateral resistance, FS and end resistance QC are not the same, although the two can consistently respond to changes in soil properties in a general trend. Because of this difference in sensitivity, the two reactions differ in their response to variability and relevance of soil properties. Second, the data used in this paper for the unit end resistance QC value is "MPa", and the lateral resistance of FS value of the unit is "KPa", the difference between the two count units also lead to the difference of soil was sensitive to different degrees.

In the following study, this difference is used to calculate the average spatial extent of H changes. In 10 groups of CPT data in Xi'an Loess loess Q3 layer end resistance QC and side resistance FS samples (sampling spacing is $z_0=0.1m$), on standardized data, firstly the stationarity and ergodicity of the test
proved that the soil profile is in accord with the second airport parameters with stationary random field conditions. Then the recursive space method and the correlation function method are used to calculate the correlation distance of H in different spatial range. Compare the difference of correlation distances calculated by different test indexes in different spatial range h, as shown in Figure 4, figure 6.

Fig. 4  Comparison of correlation distance calculated by Recursive Space Method with $f_s$ and $q_c$

Fig. 5  Comparison of correlation distance calculated by Correlation function Method with $f_s$ and $q_c$

Fig. 6  The ratio of the results of two index
As can be seen from figure 4-, Figure 6, the difference in the correlation distance calculated by using the two indices as samples is different from that used to calculate the average spatial extent of H. Generally speaking, the correlation distances calculated by the two methods are increasing with the increase of H in the spatial range, but the rate of increase of the two is not the same as that of the two methods. When the H is less than a certain value, the correlation distance computed by FS is larger than the calculated value of QC, but later it is the opposite. When the recursive space method is used, the ratio of the two increases gradually with the change of the spatial extent of H, then decreases and tends to be stable, and the ratio is roughly between 1.1 and 1.2. When the correlation function method is used, the ratio of the two increases with the change of the spatial extent of H first, then decreases.

5. Conclusion
1) for Xi'an loess, it is reasonable and feasible to use the data acquisition distance of 0.1M, and the driling CPT data as the test sample of soil correlation distance calculation.
2) for Xi'an loess, the correlation distance calculated by correlation function method is generally larger than the recursive space method, and the ratio between the two is generally about 1.1~1.3. At the same time, this different changes with the calculation of spatial range H.
3) for Xi'an loess, two kinds of calculation methods using the calculation results of different indicators of the difference is not big, but there are general rules: the calculation results of QC index calculation are the result of FS, the ratio of the two in general about 0.8~0.9. At the same time, this different changes with the calculation of spatial range H.

References
[1] Yang Yong, Jiang Gui Chun, et al. Discussion and application of the statistical method of geotechnical parameters and its simplified calculation formula. Coalfield geology and prospecting, 2013, 41(2):60~63.
[2] Vanmarcke E.H. Probability modeling of soil Profile [J]. Journal of Geotechnical Engineering Division, ASCE, 1977, 103:1227-1245.
[3] Vanmarcke E.H. Random fields: analysis and synthesis [M]. MIT Press, Cambridge, Mass.1983.
[4] Yang Yong, Study on the random filed model of xi’an loess parameter. Xi’an : chang an university,2013.
[5] Yang Yong, FAN Wen, et al. Applicability of random field model about Xi’an loess property parameter. Journal of engineering geology, 2013, 21(3):357-362.
[6] Yang Yong, ZHANG Hai Feng, et al. Statistical distribution characteristics of correlation distance of Xi’an loess. Journal of engineering geology, 2014, 22(1): 160-165.
[7] Zhu Dengfeng, Gao Dazhao. Space statical characteristics analysis of smooth stochastic field of soil characteristics. Rock and Mechanics, 2003, 24(3):455-462
[8] Rowe R K. Chairman report Testing methods[A].Proc. of the Int. Symposium on Earth Reinforcement[C]. Rotterdam: Balkema A.A, 1996:1055-1065.
[9] Wilson.H.Tang. Probabilistic characteristics of soil. The principle of reliability analysis on Geotechnical Engineering (1). Geotechnical engineering reliability analysis of feasibility of PR group..1991.9. Yangtz River Scientific Research Institute
[10] Huang Chuan Zhi, Sun Wan He. Several problems in Probability analysis of Soil parameters. Port & Waterway Engineering, 2003, 24(3):455-462
[11] Li Xiao Yong, Xie Kang He, Yu Yan. Research of the characteristics of correlation distance on soil properties indexes. China civil engineering journal, 2003, 36(8): 91-95.
[12] Cheng Qiang, Luo Shu Xue. Correlation of scale of fluctuation with soil property parameters and its calculation method. Journal of southwest jiaotong university,2000,35(5):496-500.
[13] Yan Shu Wang, Jia Xiao Li, Guo Huai Zhi. Examination of stationarity and ergodicity on soil profile. Chinese journal of geotechnical engineering, 1995; 17(3) :1-9.
[14] Yan Shu Wang, Zhu Hong Xia, et al. Study on application of random field theory to reliability analysis. Chinese journal of geotechnical engineering, 2006,28(12): 2053-2059.
[15] Xu Bin, Wang DaTong, Gao DaZhao. Discussion on determining relative distance of static cone penetration curves by relative function method. Rock and Mechanics, 1998, 19(1): 55-59.
[16] Tan XiaoHui, Wang JianGuo, et al. Study on the calculation methods of the correlation distance of soil property parameters. Journal of Hefei university of technology, 2004(11):1420-1424.
[17] Gao DaZhao. Safety index of geotechnical engineering and its application. Geotechnical investigation & surveying, 1996, (1):1-6.