

Effective Method to Diagnose Health of Earth-Fill Dams  
based on Evaluation of Spatial Variability  
of Soil Properties

January, 2019

Kazunari IMAIDE

Graduate School of Environmental and Life Science  
(Doctor's Course)  
OKAYAMA UNIVERSITY

## Acknowledgements

This study presents the results of research conducted during the years from 2014 to 2019 by the author who has been studying as a doctoral student at the Graduate School of Environmental and Life Science at Okayama University.

First and foremost, the author would like to express his sincere gratitude to Professor Shin-ichi Nishimura of Okayama University not only for his supervision, but also for his enthusiastic guidance and encouragement. Professor Nishimura allowed the author an opportunity to study and suggested the evaluation of the spatial variability of soil properties using the cone penetration test. It would not have been possible to complete this thesis without his invaluable academic guidance and continuous support.

The author also deeply appreciates Associate Professor Toshifumi Shibata for his consistent encouragement and constructive suggestions. The author would also like to offer his thanks to Associate Professor Takayuki Shuku for his valuable and instructive discussions, which led the author to have a better understanding of the random field theory. Professor Akira Murakami and Associate Professor Kazunori Fujisawa have also given precious advice. The author is grateful for their very detailed comments regarding amendments and improvements to the draft of this paper, which comprise essential parts of this thesis.

The author is indebted to the members of his Dissertation Committee, Professor Yuji Takeshita and Associate Professor Minoru Komatsu for their discussions and constructive suggestions in reviewing the manuscript.

The author has obtained the kind encouragement of Mr. Minoru Kaneshige and Mr. Toshiyuki Tamoto of the Graduate School of Okayama University. Thanks from the author are also extended to all the student members of his laboratory, especially Mr. Tatsuya Ueta and Ms. Yoshiko Narasaki of Okayama University. The author was very fortunate to have shared a pleasant time with these persons while at Okayama University. Special thanks are also due to Ms. Heather Griswold for proofreading almost all of the author's paper.

Finally, the author would like to express his sincere gratitude to his parents and his brother for their inexhaustible support and encouragement.

# Notations and abbreviations

## Notations

$\alpha$	sampling point
$\gamma_x$	approximation function of semi-variogram in $x$ direction
$\hat{\gamma}_x$	estimated value of semi-variogram from measurements in $x$ direction
$\gamma_z$	approximation function of semi-variogram in $z$ direction
$\hat{\gamma}_z$	estimated value of semi-variogram from measurements in $z$ direction
$\Delta x$	separation distances of data in $x$ direction
$\delta x$	difference value of distance between $x$ and $x'$
$\Delta z$	separation distances of data in $z$ direction
$\varepsilon_r$	random variable which follows standard normal distribution $N(0,1)$ : error term
$\varepsilon_j$	random variable which follows standard normal distribution $N(0,1)$ : error term
$\theta$	parameter to control density function
$\kappa$	Lagrange multiplier
$\lambda_\alpha, \lambda_\beta$	weights of linear combinations
$\rho(x, x'; y, y'; z, z')$	indicates autocorrelation function
$\sigma$	standard deviation
$\sigma_b^2$	priori variance of random function $B$
$\sigma_k^2(x, z)$	squared residual between true function value $b$ and estimated value $b^*$
$\sigma_{v0}$	total overburden stress
$\sigma'_{v0}$	effective overburden stress
$\Phi$	cumulative standard normal distribution function
$a$	net area ratio
$a_{\max}$	peak ground acceleration of dam
$a_0, a_1, a_2, a_3, a_4,$ and $a_5$	approximation coefficients of mean function
$A_c$	cross-sectional area of base of cone
$A_n$	cross-sectional area of load cell or shaft
$(B_1, B_2, \dots, B_M)$	measured data: random variables
$b(x, z)$	realization of random variable $B$ which expresses soil properties
$b^*$	estimated value by kriging
$b^{(l)}(x, z)$	$l$ th realization of random function $B$ without conditioning
$b^{*(l)}(x, z)$	kriged estimation using the simulated values of $b^{(l)}(x, z)$ at sampling points

$b_\alpha$	sample value of random function $B$ at point $\alpha$
$b_c^{(l)}(x, z)$	realization of conditional simulation
$C(\ )$	covariance function
$\mathbf{C}$	$M \times M$ covariance matrix
$[C_{ij}]$	$i$ - $j$ component of covariance matrix
$C_{0x}$	parameter used for nugget effect in $x$ direction
$C_{0z}$	parameter used for the nugget effect in $z$ direction
$C_{1x}$	parameter used to express shape of semi-variogram function in $x$ direction
$C_{1z}$	parameter used to express shape of semi-variogram function in $z$ direction
$D$	difference value defined as $D =  R_M - T $
$D_H$	difference value calculated from upper threshold value defined as $D_H(x, z) = T_H - R_M(x, z)$
$D_L$	difference value calculated from lower threshold value defined as $D_L(x, z) = R_M(x, z) - T_L$
$D_x$	number of combinations of $U(x_k) - U(x_k +  x_i - x_j )$ in which $ x_i - x_j $ mean separation distances in $x$ direction
$D_z$	number of combinations of $U(z_k) - U(z_k +  z_i - z_j )$ in which $ z_i - z_j $ mean separation distances in $z$ direction
$D_{50}$	mass median diameter
$E(x, z)$	mean value calculated from realizations of random field of $\mathbf{R}_G$
$f$	normalized variable of $S$ as $f = (S - m) / \sigma$ to remove trend
$f'$	realization of conditional simulation of normalized variable
$f(\ )$	probability density function
$\mathbf{f}_H$	vector of conditional simulation realizations of normalized variable of high group
$\mathbf{f}_L$	vector of conditional simulation realizations of normalized variable of low group
$\mathbf{f}_M$	vector of conditional simulation realizations of normalized variable of middle group
$f_s$	measured sleeve friction
$F()$	cumulative distribution function of standardized variable $f$
$F_c$	finer content
$F_{cic}$	finer content obtained from $I_c$

$F_L$	liquefaction resistance factor
$F_R$	normalized friction ratio
$g()$	function between $P_f$ and $F_L$
$H$	seismic hazard curve over next 50 years at top of dam
$I_c$	soil behavior type index
$K$	number of unknown parameters included in equation
$l$	correlation distance
$l_x$	correlation lengths for $x$ direction
$l_z$	correlation lengths for $z$ direction
$L$	dynamic load
$L( )$	likelihood function
$L_x$	horizontal length of site investigation at studied site
$m(x, y, z)$	mean function
$\mathbf{m}^t=(m_1, m_2, \dots, m_M)$	indicates mean vector of random function $\mathbf{S}^t=(S_1, S_2, \dots, S_M)$
$M$	number of test points
$n$	number of sample points used for interpolation
$N_a$	corrected $N$ -value including effects of particle size distribution
$N_c$	$N$ -values calculated from results of cone penetration test (CPT)
$N_{\delta x}$	number of combinations of $[b(x_i + dx) - b(x_i)]$
$N_e$	nugget effect parameter
$N_{HOB}$	number of measured data included in high group
$N_{HSIM}$	number of locations selected from simulation results of high group for re-composition
$N_{LOB}$	number of measured data classified in low group
$N_{LSIM}$	number of locations selected from simulation results of low group for re-composition
$N_{MSIM}$	number of locations selected from simulation results of middle group for re-composition
$N_{OB}$	total number of measured data
$N_r$	number of prediction targets
$N_{SIM}$	number of evaluated locations in re-composition
$N_{SPT}$	$N$ -value obtained from standard penetration test (SPT): $N$ -value
$N_{SWS}$	$N$ -value obtained from Swedish weight sounding (SWS)
$N_T$	total number of evaluated points in conditional simulation
$N_1$	converted $N$ -value as effective overburden stress equivalent to 100 kPa

$P_f$	liquefaction probability
$P_{fE}$	expected value for liquefaction probability
$\bar{P}_{fE}(a_{\max})$	spatial average of $P_{fE}$
$P_{fE50}$	liquefaction probability over next 50 years
$\bar{P}_{fE50}$	liquefaction probability of whole dam over next 50 years
$P_H$	assembly ratio of high group for re-composition
$P_L$	assembly ratio of low group for re-composition
$q_c$	measured cone resistance
$q_n$	normalized cone penetration resistance
$q_t$	corrected cone resistance
$Q_t$	normalized CPT penetration resistance
$r_d$	reduction factor
$R$	liquefaction resistance
$\mathbf{R}_G$	vector of simulated results for re-composition of three groups
$\mathbf{R}_H$	vector of $N_c$ obtained from conditional simulation of high group
$\mathbf{R}_L$	vector of $N_c$ obtained from conditional simulation of low group
$\mathbf{R}_M$	vector of $N_c$ obtained from conditional simulation of middle group
$\mathbf{s}' = (s_1, s_2, \dots, s_M)$	realization of random vector $\mathbf{S}' = (S_1, S_2, \dots, S_M)$
$s_u$	undrained shear strength
$SD(x, z)$	standard deviation calculated from realizations of random field of $\mathbf{R}_G$
$\mathbf{S}' = (S_1, S_2, \dots, S_M)$	random vector of random variable $S$
$t$	iteration number of simulation
$T$	threshold value
$T_H$	threshold value between high group and middle group
$T_L$	threshold value between middle group and low group
$u$	pore water pressure
$u_2$	pore water pressure measured at cylindrical extension part of cone: $u=u_2$
$v(x, y, z)$	variance function
$w$	standardized residual
$x$	real number in horizontal coordinate
$\mathbf{X} = (x, y, z)$	function of special location
$y$	other horizontal coordinate: $y$ is perpendicular to embankment axis
$Y$	standardized variable of normalized variable $f$
$Y'$	realization of conditional simulation of standardized variable
$\mathbf{Y}_H$	vector of standardized variable, $Y$ , of high group

$\mathbf{Y}'_H$	vector of conditional simulation realizations of standardized variable of high group
$\mathbf{Y}_L$	vector of standardized variable, $Y$ , of low group
$\mathbf{Y}'_L$	vector of conditional simulation realizations of standardized variable of low group
$\mathbf{Y}_M$	vector of standardized variable, $Y$ , of middle group
$\mathbf{Y}'_M$	vector of conditional simulation realizations of standardized variable of middle group
$z$	vertical coordinate

## **Abbreviations**

AIC	Akaike Information Criterion
CPT	Cone Penetration Test
CPTU	Piezometer Cone Penetration Test
FORM	First Order Reliability Method
ISO	International Organization for Standardization
J-SHIS	Japan Seismic Hazard Information Station
MLE	Maximum Likelihood Estimation
RMSE	Root Mean Square Error
SPT	Standard Penetration Test
SWS	Swedish Weight Sounding
RBD	Reliability-Based Design
VST	Vane Shear Test

## List of tables

Table 4.1 Statistical models of $\log N_c$ estimated by MAICE .....	35
Table 4.2 Constants of covariance functions of $\log N_c$ determined by semi-variograms .....	38
Table 4.3 Statistical models for $\log N_c$ introduced into simulation .....	42
Table 5.1 Statistical models determined by MAICE .....	49
Table 5.2 Influence of removal of outliers on goodness to fit for model. ....	52
Table 5.3 Statistical models of $Y_H$ , $Y_L$ , $Y_M$ .....	53
Table 5.4 Variables for re-composition of simulation results .....	60
Table 6.1 Statistical models of $\log N_c$ and $\log F_{clc}$ estimated by MAICE .....	73
Table 6.2 Constants of covariance functions of $\log N_c$ and $\log F_{clc}$ determined by semi-variograms .....	75
Table 6.3 Statistical models for $\log N_c$ and $\log F_{clc}$ introduced into simulation.....	77

## List of figures

Fig. 1.1 Concept of sequence of uncertainties incorporated into the RBD of geotechnical structures (partially modified figure of Honjo, 2011) .....	1
Fig. 2.1 Cross section of an example of a cone penetrometer .....	10
Fig. 2.2 Concept of net area ratio .....	11
Fig. 2.3 Cone penetration machine and attached cone .....	11
Fig. 2.4 Comparison of $N_c$ and $N_{SPT}$ .....	13
Fig. 2.5 Comparison of $\log F_{c1c}$ and $\log F_c$ . This figure was derived from data included in Suzuki et al. (2003) .....	13
Fig. 3.1 Examples of random processes of different correlation distances .....	16
Fig. 3.2 Several types of autocorrelation functions .....	17
Fig. 3.3 Example of autocorrelation function .....	17
Fig. 4.1 Plan views .....	31
Fig. 4.2 Cross section of each dam and legend .....	32
Fig. 4.3 Histograms of $N_c$ .....	33
Fig. 4.4 Values measured by CPTUs and mean functions of $\log N_c$ .....	36
Fig. 4.5 Semi-variograms and approximation function of $Y$ .....	37
Fig. 4.6 Spatial distribution of statistics of $N$ -value .....	41
Fig. 4.7 Values measured by SPT and statistical values of interpolated values at $x=125m$ .....	42
Fig. 5.1 Distribution of tip resistance in cone penetration test (CPT) .....	45
Fig. 5.2 Flowchart to model data of soil strength .....	46
Fig. 5.3 Plan views (these figures already shown in Fig. 4.1) .....	48
Fig. 5.4 Cross section of each dam and legend (these figures already shown in Fig. 4.2) .....	48
Fig. 5.5 Spatial distributions of $\log N_c$ .....	49
Fig. 5.6 Values measured by CPTUs and mean function of $\log N_c$ .....	50
Fig. 5.7 Relationship between removal ratio of outliers and $RMSE$ .....	52
Fig. 5.8 Classification of in-situ data in case of (10, 90) at $x = 2$ m .....	53
Fig. 5.9 Relationship between semi-variogram and removal of outliers at C dam .....	55

Fig. 5.10 Relationship between semi-variogram and removal of outliers at D dam .....	55
Fig. 5.11 Concept of re-composition of simulated results of three groups.....	58
Fig. 5.12 Flowchart of re-composition of simulation results .....	59
Fig. 5.13 Definitions of $D_H$ and $D_L$ .....	59
Fig. 5.14 Spatial distribution of statistics by proposed method.....	63
Fig. 5.15 Comparison of soil strength distribution between expected values and in-situ data .....	65
Fig. 5.16 Comparison of probability density function between simulated values and in-situ data .....	65
Fig. 5.17 Probability density function of standardized residuals, $w$ .....	66
Fig. 6.1 Plan view of C dam and testing points of CPTs and SPT .....	71
Fig. 6.2 Geological cross section of C dam.....	71
Fig. 6.3 Distribution of $N$ -values from SPT in depth direction .....	72
Fig. 6.4 Geological columnar section at boring point .....	72
Fig. 6.5 Values measured by CPTUs and mean function with intervals of 4 m .....	74
Fig. 6.6 Semi-variograms and approximation function of $\log N_c$ .....	75
Fig. 6.7 Semi-variograms and approximation function of $\log F_{clc}$ .....	75
Fig. 6.8 Spatial distribution of statistical values of $F_c$ .....	78
Fig. 6.9 Spatial distribution of statistical values of $N_{SPT}$ .....	78
Fig. 6.10 Time histories of bedrock acceleration at studied site assuming Nankai Trough earthquake .....	80
Fig. 6.11 Seismic hazard curve of Nankai Trough earthquake for studied dam.....	80
Fig. 6.12 Relationship between liquefaction probability $P_f$ and liquefaction resistance factor FL (Relationship was derived from data included in Iwasaki et al., 1984).....	81
Fig. 6.13 Seismic hazard curve and fragility curve of dam.....	84
Fig. 6.14 Cumulative distribution function of spatial average of liquefaction probability of dam over next 50 years $\bar{P}_{JE50}$ .....	84
Fig. 6.15 Spatial distribution of expected values of liquefaction probability when $a_{max}$ is equivalent to 140 gal.....	85
Fig. 6.16 Spatial distribution of expected values of liquefaction probability over next 50 years .....	85

# Contents

Acknowledgements .....	i
Notations and abbreviations .....	ii
List of tables .....	viii
List of figures .....	ix
Contents .....	xi
Chapter 1 Introduction	
1.1 Background and objectives.....	1
1.2 Summary of previous research .....	6
1.2.1 Spatial variability of soil properties .....	6
1.2.2 Liquefaction of earth-fill dams .....	7
1.3 Composition of thesis .....	7
Chapter 2 Cone penetration test	
2.1 Introduction .....	9
2.2 Outline of cone penetration test.....	9
2.3 Applications of results of cone penetration test.....	11
2.4 Conclusion .....	14
Chapter 3 Fundamental theories of statistical modeling	
3.1 Introduction .....	15
3.2 Random field theory .....	15
3.3 Akaike Information Criterion (AIC).....	19
3.4 Geostatistical method.....	21
3.4.1 Semi-variograms .....	22
3.4.2 Kriging .....	23
3.4.3 Conditional simulation.....	24
3.5 Statistical modeling of ground.....	25

Chapter 4	Estimation of correlation lengths for several earth-fill dams	
4.1	Introduction .....	29
4.2	Site investigation .....	30
4.3	Statistical modeling of CPT results .....	34
4.4	Evaluation of spatial distribution of $N$ -value using conditional simulation	39
4.5	Conclusion .....	43
Chapter 5	Evaluation of spatial variability of cone penetration resistance inside earth-fill dams composed of materials with different particle size distributions	
5.1	Introduction .....	45
5.2	Site investigation .....	47
5.3	Statistical modeling of soil strength .....	49
5.4	Interpolation of soil strength using conditional simulation .....	56
5.5	Re-composition of simulated values of three groups .....	58
5.6	Validation of simulated values using proposed method .....	62
5.6.1	Estimated statistical model for C dam .....	62
5.6.2	Evaluation of spatial distribution of soil strength .....	62
5.6.3	Validation of interpolated results by proposed method.....	64
5.7	Conclusion .....	66
Chapter 6	Evaluation of liquefaction probability of earth-fill dam over next 50 years using geostatistical method based on CPT	
6.1	Introduction .....	68
6.2	Site investigation .....	70
6.3	Statistical model identification of studied dam .....	72
6.4	Interpolation of measured values of CPTs.....	76
6.5	Liquefaction resistance factor $F_L$ .....	79
6.6	Calculation of liquefaction probability .....	81
6.7	Evaluation of liquefaction probability over next 50 years.....	83
6.8	Conclusion .....	86

Chapter 7	Concluding remarks and future work	
7.1	Summary of each chapter .....	88
7.2	Summary of research .....	90
7.3	Future works .....	90
References	.....	92

## Chapter 1

### Introduction

Chapter 1 consists of three parts. First, the background and objectives of this thesis are described. Second, the past literary works related to the topics of this thesis are summarized. Third, the composition of this thesis is presented.

#### 1.1 Background and objectives

In the design of geotechnical structures, the uncertainties included in the soil properties have not been explicitly introduced into the design as numerical values; and thus, the deterministic values for the design of these structures have been decided based on the judgment of engineers. Nowadays, however, the reliability-based design (RBD) method is applied to obtain a realistic response using analytical models and to discuss economical designs. The RBD method quantitatively addresses heterogeneity and the uncertainties of the material properties and external forces. In the International Organization for Standardization (ISO) and Specifications for Highway Bridges, the uncertainties of the load and the material properties are beginning to be taken into account, as presented by Phoon and Retief (2016) and the Japan Road Association (2017). Similar discussions are also being conducted in the fields of irrigation, drainage and rural engineering (Murakami et al., 2009a; Murakami et al., 2009b).

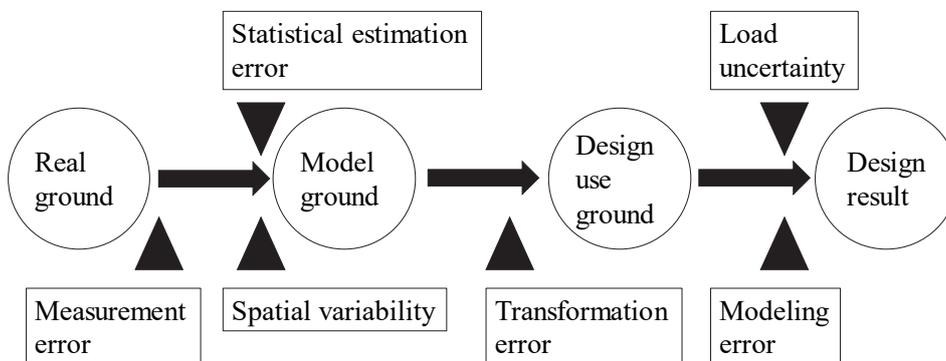


Fig. 1.1 Concept of sequence of uncertainties incorporated into the RBD of geotechnical structures (partially modified figure of Honjo, 2011)

In recent years, foundation disasters have occurred frequently. For example, a river dike collapsed due to heavy rain when a piping phenomenon arose inside the dike. Also, a geotechnical structure failed due to liquefaction inside the structure. It is necessary, therefore, to perform a numerical analysis, in advance, considering complex factors, such as rainfall and earthquakes, the soil properties, etc., in order to prepare for heavy rain events, which have been increasing due to global warming, and huge earthquakes, which are both of great concern. In addition, in order to evaluate the points where these disasters are likely to occur, an analysis of the complex factors is required. A model to evaluate the ground in detail would play an important role in making sure that this type of sophisticated analysis will function effectively. However, the design of geotechnical structures generally includes many kinds of unavoidable uncertainties of the soil, and a modeling method to properly treat these uncertainties is necessary for obtaining results that reflect the actual situations.

Honjo and Otake (2012) indicated that the four main uncertainties which should be considered in the reliability-based design of geotechnical structures are (1) spatial variability of the soil properties, (2) statistical errors within estimations caused by determining the values of the soil properties from limited investigations, (3) transformation errors caused by converting the measured data into the soil properties in question for a design, and (4) modeling errors accompanying the calculation model. Honjo (2011) presented an example of the sequence of uncertainties incorporated into the RBD of geotechnical structures, as shown in Fig. 1.1. In addition, Baecher and Christian (2003) stated that the uncertainties in geotechnical engineering can be categorized into two types, namely, aleatory and epistemic. Aleatory means random uncertainty, like throwing dice, while epistemic implies uncertainty due to a limitation of knowledge, for example, playing card games like poker. Among the four main uncertainties in geotechnical engineering, (1) spatial variability of the soil properties inside grounds is classified as aleatory, and the others are classified as epistemic. Outlines of the four main uncertainties and of the past studies related to them are given below.

First, the spatial variability of the soil properties primarily results from the natural geologic processes that are produced and continually modify the in-situ soil masses. This uncertainty is modeled based on the random field theory in the RBD of geotechnical engineering. A random variable, which is a function of one variable, for instance, time, is called a random process. A variable which is a function of several variables, for example, spatial coordinates, is called a random field. The soil properties are determined by themselves and already exist at each location. However, due to

epistemic uncertainties, they are modeled using the random field theory for convenience. This theory simplifies and idealizes the problems. There is a vast amount of literature on this topic. Lumb (1966, 1974) conducted one of the first studies; it introduced the formal random process to geotechnical engineering. Vanmarcke (1977, 1983) introduced the systematic random field theory.

Second, due to the limited amount of measured samples obtained from on-site investigations, the sample statistics, for instance, the sample mean and the sample variance, which are used to estimate the population parameters of the ground, include statistical estimation errors. The errors decrease along with the increase in the number of samples. In evaluating statistical estimation errors, Honjo and Setiawan (2007) and Honjo (2008) pointed out that it is important to distinguish “general estimation” and “local estimation”. In a general estimation, the relative positions of the location under investigation and the location of the structure to be built are not taken into account in the estimation of the soil parameters. In a local estimation, however, the relative positions of these two locations are taken into account in the estimation. Honjo and Setiawan (2007) presented a formulation for these two cases for a particular situation, and compared the results of their proposed method with those of the traditional statistical theory. Honjo (2008) discussed this problem with an illustrative example, namely, the determination of the characteristic values of the soil parameters in design codes.

Third, errors associated with the conversion of the measured soil properties by a soil investigation to the geotechnical parameters to be used in the design calculation are termed transformation errors. For example, the  $N$ -value can be utilized to estimate the friction angle and Young's modulus. Kulhawy and Mayne (1990) gave a comprehensive reference for this problem including a considerable amount of quantitative information. Moreover, Otake and Honjo (2014) comprehensively summarized the errors in the conversion formulas used in the design codes in Japan.

Lastly, design calculation model errors are caused by the prediction capabilities of simplified and idealized design calculation models against the real phenomena. In geotechnical engineering, tests and experiments on as-large-as-real-structure scales (e.g., pile load tests, plate loading tests, etc.) are commonly performed, and several cases of failure are available for reference especially on earth structures, such as embankments, cut slopes, and excavations. These facts make it easier to quantitatively evaluate the model errors in the design of geotechnical structures. For example, Matsuo and Asaoka (1976) analyzed failed embankments on soft grounds by adding the error term to evaluate the factor of safety, and the modeling error was quantitatively evaluated.

From the above sources, it is the physical uncertainty that will mainly be discussed here. Since the stability of geotechnical structures is related to the spatial variability of the soil properties, evaluations of the spatial variability of the soil properties are important. In the design of geotechnical structures, addressing the uncertainties especially of the spatial variability of the soil properties, in order to evaluate the stability of structures, could be useful for decision-making.

In this thesis, particular focus is placed on earth-fill dams. Based on an evaluation of the spatial variability of the soil properties inside earth-fill dams, the safety of the dams is quantitatively evaluated. There are many earth-fill dams in Japan, and most of them have aged and deteriorated. In addition, over the next 30 years, the probability of an earthquake occurring in the Nankai Trough is predicted to be about 70%. And the magnitude of this earthquake is predicted to be as large as 9.0. Such an earthquake would affect a wide area from western to central Japan and could cause the further decay and/or collapse of these earth-fill dams. To mitigate the disasters caused by this type of huge earthquake, a quantitative evaluation of the stability of the dams and the subsequent sufficient reinforcement of them should be performed as soon as possible. In general, however, the intervals between past investigations of these earth-fill dams have been too long to evaluate the spatial variability of the soil properties. To appropriately evaluate the safety of earth-fill dams under realistic conditions, dealing with the spatial variability of the soil properties, the testing of earth-fill dams should be conducted at short intervals.

In the modeling of the spatial variability of the soil properties, the random field theory has generally been assumed in past literature. The degree of the spatial correlation of the random field is represented by the correlation distance. Since the correlation distance is an essential parameter for modeling the spatial variability of the soil properties using the random field, special focus is placed on the correlation distance when examining earth-fill dams.

There are two methods for evaluating the soil properties of the ground. One employs laboratory tests and the other employs in-situ tests. Since laboratory tests use relatively small specimens, which are commonly obtained from boring cores, it is difficult to extensively apply the results to the soil properties of the whole ground. It is also difficult to maintain the natural stress conditions during the test procedure. In addition, the interpretation of the laboratory test results becomes complicated when the ground in question has complex geological layers or when the ground consists of materials with different particle size distributions. From this perspective, laboratory testing using specimens obtained by boring has shortcomings, and in-situ tests, namely,

sounding tests, which can directly measure the soil properties in their natural stress conditions, are advantageous. Moreover, to evaluate the spatial variability of the soil properties, a great deal of data from the studied site are required. In general, neither the standard penetration test (SPT), which is one kind of sounding test, nor boring can provide appropriate estimations for the spatial model of the soil properties. This is because it is difficult to collect a sufficient amount of information on the soil due to the economic factor. On the other hand, the cone penetration test (CPT), which is another kind of sounding test, can be conducted economically and speedily. Its strong point is that it can be used to detect the detailed locations of the weak areas inside earth-fill dams. The results obtained from CPTs were employed here to model the spatial variability of the soil properties.

The objective of this study is to propose a method for the effective diagnosis of earth-fill dams based on an evaluation of the spatial variability of the soil parameters. For this purpose, the following three topics are presented. First, the database of the spatial structure of the soil properties has not been sufficiently examined. In particular, it is common for the intervals of boring tests in the horizontal direction to be several hundred meters. Thus, the information on the correlation distance in the horizontal direction is quite limited. In the present study, therefore, the correlation distance of the soil properties has been summarized for five earth-fill dams based on the results obtained from CPTs performed at short intervals. Second, in the modeling of the soil properties inside dams, the outliers included in the measured data affect the estimation of the spatial structures. However, several materials, which have different particle size distributions, have been intentionally added in some cases to reinforce the dams. Thus, a method is proposed here to appropriately evaluate the spatial distribution of the soil strength inside earth-fill dams composed of materials with different particle size distributions. Third, as an example of the application of the estimated spatial distribution of the soil properties, the liquefaction probability of earth-fill dams is discussed. In addition, to quantify the risk for large earthquakes, incorporating the time factor, a method is presented to evaluate the liquefaction probability of earth-fill dams over the next 50 years.

## 1.2 Summary of previous research

### 1.2.1 Spatial variability of soil properties

Among the four uncertainties mentioned in section 1.1, particular focus is placed in this study on (1) the spatial variability of the soil properties. Since the spatial variability of the soil properties affects the stability of geotechnical structures, the modeling of the variability and the evaluation of the stability considering the spatial variability of the soil properties have been vastly studied.

The heterogeneity of the ground was addressed as an important issue back in the 1970s. Matsuo and Kuroda (1974) calculated the failure probability of an embankment by taking into account the variability of the soil properties. As a study on the modeling of the two-dimensional heterogeneity of the ground, for example, Griffiths et al. (2009) compared two stability analyses of a slope. One applied the stochastic finite element method and the other applied the first order reliability method (FORM). By comparing these analyses, they pointed out the usefulness of the former method. In addition, as an approximate evaluation method to easily design a ground structure considering the spatial variation, Honjo and Otake (2012) proposed a theory for obtaining the estimation of the local average of the soil properties especially focusing on the neighborhood of a structure. Furthermore, Otake and Honjo (2012) compared two calculation methods. One was the local average theory, proposed by Honjo and Otake (2012), and the other was a shallow foundation settlement problem, studied using the stochastic finite element method presented by Fenton and Griffiths (2008). As a result, it was confirmed that the method proposed by Honjo and Otake (2012) could easily take into account the spatial variability of the ground in shallow foundation settlement problems. Furthermore, Kasama and Zen (2010) examined the influence of the spatial variability of the soil strength on the probability of the collapse of an improved ground in the design of a shallow foundation.

As an example of a study using the geostatistical method, Nishimura and Shimizu (2008) evaluated the spatial distribution of the liquefaction probability in a three-dimensional space based on the autocorrelation of the soil properties of an embankment. They also took into account the correlation between two soil properties by using the co-kriging method. Moreover, Nishimura and Shimizu (2011) evaluated the spatial distribution of the liquefaction probability of an embankment considering the spatial distribution of the soil properties characterized by the correlation distance. Furthermore, the optimal design for soil improvement was discussed based on the

liquefaction probability by minimizing the expected costs.

In addition, as an example of a study based on the CPT, Vivek and Raychowdhury (2014) and Chen et al. (2015) created a hazard map for the liquefaction risk in an area using the geostatistical method. Vivek and Raychowdhury (2014) showed a case where the liquefaction probability was underestimated when the spatial variability of the soil strength was not taken into consideration.

As indicated by the above studies, in the stability analysis of geotechnical structures, the heterogeneity of the soil properties has been considered for problems of various scales, such as areas, slopes, embankments, etc.

### **1.2.2 Liquefaction of earth-fill dams**

There are many earth-fill dams for agriculture in Japan. Most of them were constructed 150-500 years ago and have become old and weak. In addition, as stated previously, the probability of an earthquake occurring in the Nankai Trough over the next 30 years is predicted to be about 70%. And the magnitude of this earthquake is predicted to be as large as 9.0. Such an earthquake would affect a wide area from western to central Japan, and would most likely cause further damage to these earth-fill dams. In the 2011 off the Pacific coast of Tohoku Earthquake, the failure of the Fujinuma Dam was found to have been caused by liquefaction inside the dam (e.g., Tatsuoka et al., 2017; Ono et al., 2011). Since then, the design guidelines for earth-fill dams for irrigation have required an evaluation of the probable liquefaction damage (Ministry of Agriculture, Forestry, and Fisheries of Japan, 2015). For the above-mentioned reasons, the efficient improvement of the dams must be conducted within a limited time and with a limited budget. Evaluating and comparing the seismic risk of various dams enables a quantitative prioritization of the dams such that a decision can be made on which dams among many are to be improved. In general, the seismic risk is expressed by the multiplication of the failure probability by the cost of failure. Therefore, a new procedure is proposed for evaluating the liquefaction probability of a dam against a potential Nankai Trough earthquake.

## **1.3 Composition of thesis**

This thesis is composed of seven chapters. Chapter 2 describes the outline of the cone penetration test. Chapter 3 presents several methods to model the spatial variability of

the soil parameters. Chapter 4 summarizes the correlation distance of the soil strength at five earth-fill dams. Chapter 5 presents a method for evaluating the spatial distribution of the soil strength inside earth-fill dams composed of different particle size distributions. Chapter 6 shows a method for evaluating the liquefaction probability of earth-fill dams over the next 50 years considering the spatial variability of the soil properties and the risk of an earthquake. The conclusion and summary of this thesis are given in Chapter 7.

## Chapter 2

# Cone penetration test

### 2.1 Introduction

The cone penetration test is an on-site investigation method with which the soil properties can be evaluated continuously and economically. It is widely used for ground investigations and designs throughout Europe. The cone penetration test is explained in this chapter and employed to model the geotechnical engineering properties of soils.

The outline of this chapter is as follows. In section 2.2, an outline of the cone penetration test is presented. Section 2.3 shows the conversion formulas for the cone penetration test used to estimate several soil properties. Finally, a conclusion is given in section 2.4.

### 2.2 Outline of cone penetration test

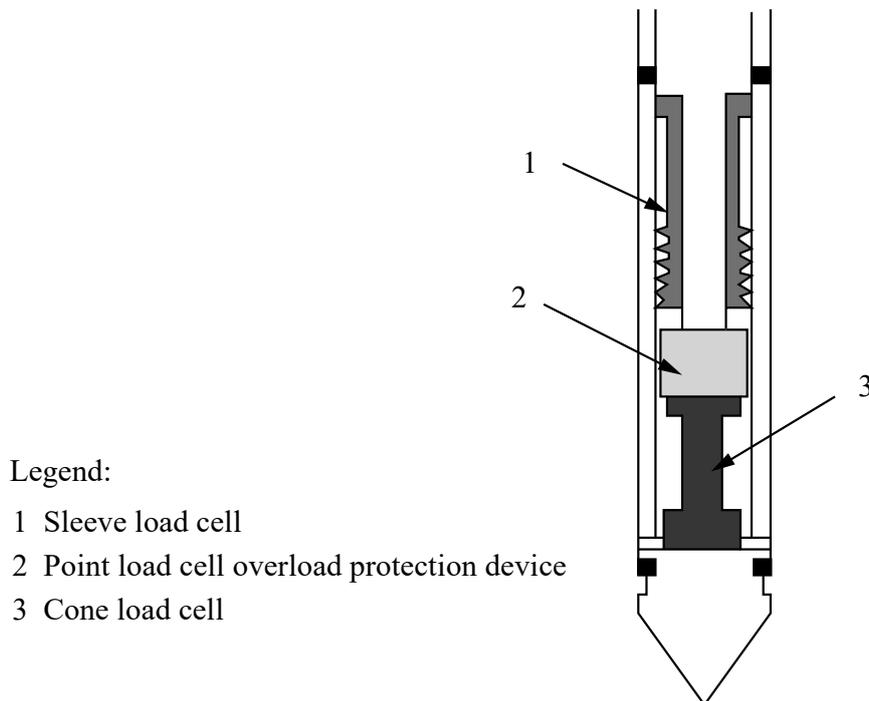
The electric cone penetration test with pore pressure measurements (piezometer cone penetration test, CPTU) is one type of sounding test (Japan Geotechnical Society, 2016). The test results can be used for estimating the ground composition and the mechanical properties of the soils. The CPTU measures the cone penetration resistance, the skin friction resistance, and the pore water pressure. In the test, a cone penetrometer is fitted to the end of a rod and pushed at a constant rate of penetration into the ground, and the force is measured electrically. The pore water pressure around the cone is also measured.

In Fig. 2.1, the cross section of an example of a cone penetrometer is presented. The following values are measured in the electric cone penetration test:

Measured cone resistance  $q_c$  (MPa)

Measured sleeve friction  $f_s$  (MPa)

Pore water pressure  $u$  (MPa)



- Legend:
- 1 Sleeve load cell
  - 2 Point load cell overload protection device
  - 3 Cone load cell

Fig. 2.1 Cross section of an example of a cone penetrometer

The pore water pressure  $u=u_2$  is measured at the cylindrical extension part of the cone. The measured cone resistance and measured sleeve friction are affected by the surrounding pore water pressure. The measured cone resistance is corrected using Eq. (2.1) only when the pore water pressure is measured at the filter of the cylindrical extension part of the cone ( $u_2$ ).

$$q_t = q_c + u_2 \times (1 - a) \quad (2.1)$$

where  $q_t$  is the corrected cone resistance (MPa),  $q_c$  is the measured cone resistance (MPa),  $u_2$  is the pore water pressure at the cylindrical extension part of the cone (MPa), and  $a$  is the net area ratio, namely,  $a=(A_n/A_c)$ , where  $A_c$  is the cross-sectional area of the base of the cone and  $A_n$  is the cross-sectional area of the load cell or shaft. The concept of the net area ratio,  $a$ , is shown in Fig. 2.2.

The CPT has the advantage of often providing more detailed and precise data at a higher speed and a lower cost. In addition, the results obtained from the CPT and those obtained from the CPTU can be converted into many soil properties for use in soil engineering design problems.

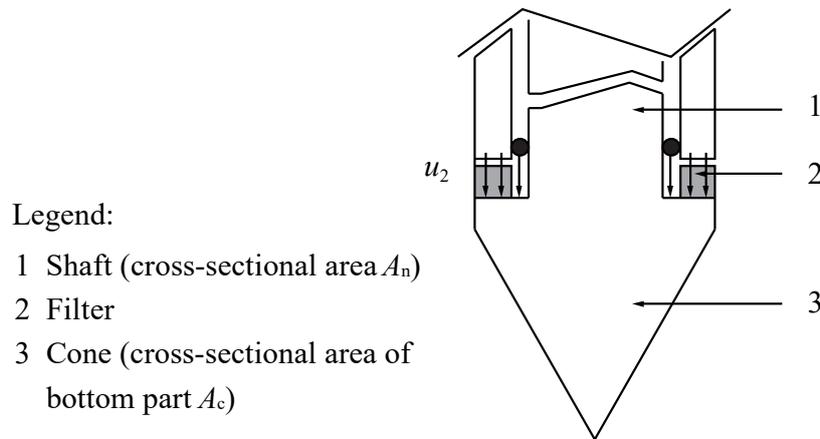


Fig. 2.2 Concept of net area ratio



Fig. 2.3 Cone penetration machine and attached cone

The penetration machine and the cone utilized in the site investigation of this thesis are shown in Fig. 2.3. The weight of the machine is about 3t; the weight works as a counter force to penetrate the cone into the ground. One of the disadvantages of the CPT is the low flexibility in its transportation to the studied sites. This is because the penetration machine requires a sufficient amount of counter force as the weight, and the size of the machine generally becomes large.

### 2.3 Applications of results of cone penetration test

Based on the results obtained from the CPT, several soil properties are calculated through the use of conversion formulas. The CPT is disadvantageous in that soil

samples for visual/lab inspections cannot be obtained with it. However, through the use of the conversion formulas, the results obtained from the CPT allow for estimates of the type of soil at the measured location, the  $N$ -values calculated from the CPT,  $N_c$ , and the fines content,  $F_c$ .

One of the primary applications of the CPT is for stratigraphic profiling. Considerable experience exists concerning the identification and classification of soil types from CPT data. A soil classification chart (Robertson, 1990) and soil behavior type index  $I_c$  (Robertson and Fear, 1995) were proposed for the CPT and the CPTU.

To conduct the soil classification based on the CPT, soil behavior type index  $I_c$  was proposed, as seen in Eq. (2.2) (Robertson and Fear, 1995).  $I_c$  includes two normalized parameters,  $Q_t$  and  $F_R$ , given by Eqs. (2.3) and (2.4) (Robertson, 1990), respectively.

$$I_c = \left\{ (3.47 - \log Q_t)^2 + (1.22 + \log F_R)^2 \right\}^{0.5} \quad (2.2)$$

$$Q_t = (q_t - \sigma_{v0}) / \sigma'_{v0} \quad (2.3)$$

$$F_R = (f_s / (q_t - \sigma_{v0})) \times 100\% \quad (2.4)$$

where  $Q_t$  is the normalized CPT penetration resistance,  $F_R$  is the normalized friction ratio (%),  $q_t$  is the corrected cone resistance (MPa),  $f_s$  is the measured sleeve friction (MPa), and  $\sigma_{v0}$  and  $\sigma'_{v0}$  are the total and the effective overburden stresses (MPa), respectively. Corrected cone resistance  $q_t$  and measured sleeve friction  $f_s$  are directly measured values in the CPT.

Based on  $I_c$  and  $q_t$ , obtained from the CPT, the CPT  $N$ -value,  $N_c$ , and the fines content,  $F_{clc}$ , obtained from  $I_c$ , can be estimated by the following equations. Eq. (2.5) was proposed by Suzuki et al. (2003), while Eq. (2.6) was derived from the data included in Suzuki et al. (2003).

$$\begin{aligned} N_c &= 0.341 I_c^{1.94} (q_t - 0.2)^{(1.34 - 0.0927 I_c)} & (q_t > 0.2 \text{ MPa}) \\ N_c &= 0 & (q_t \leq 0.2 \text{ MPa}) \end{aligned} \quad (2.5)$$

$$F_{clc} = 1.0 I_c^{3.2293} \cdot 10^{0.3024} \quad (\%) \quad (2.6)$$

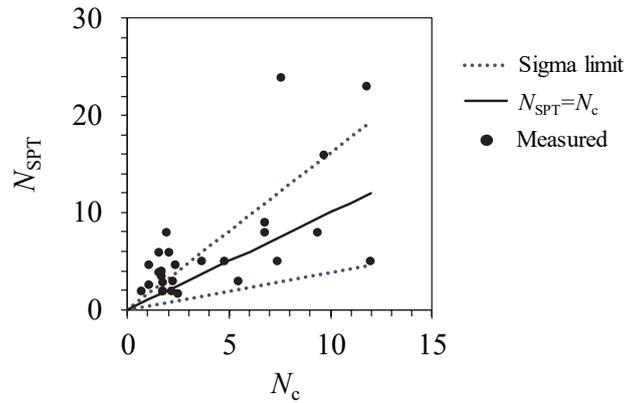


Fig. 2.4 Comparison of  $N_c$  and  $N_{SPT}$

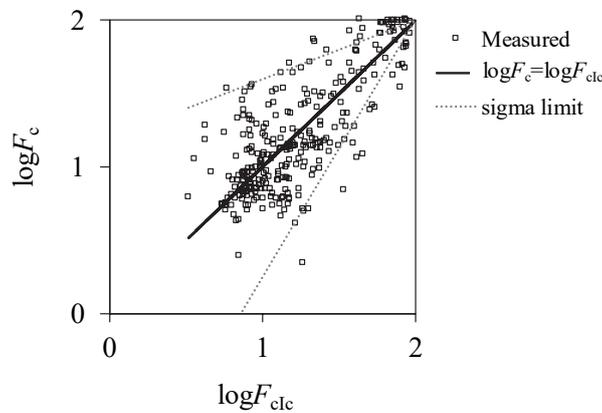


Fig. 2.5 Comparison of  $\log F_{clc}$  and  $\log F_c$ . This figure was derived from data included in Suzuki et al. (2003).

Since the above-mentioned conversion formulas are based on experimental data, conversion errors could be included. The accuracy of the conversion formulas should be taken into account in order to properly design improvements for dams. However, in Eqs. (2.5) and (2.6), the conversion errors are not explicitly considered. Thus, the conversion errors are quantified based on the measured data.

The relationship between  $N_c$  and the standard penetration test (SPT)  $N$ -values,  $N_{SPT}$ , is shown in Fig. 2.4, and  $N_c$  is converted into  $N_{SPT}$  considering the conversion error by the following equation:

$$N_{SPT} = N_c(1 + 0.62\varepsilon_r) \quad (2.7)$$

where error term  $\varepsilon_r$  is assumed to follow standard normal distribution  $N(0,1)$ . Based on the proposal of Suzuki et al. (2003),  $N_c = N_{SPT}$  is assumed in Eq. (2.7), and the quantity of the conversion error is given by the coefficient of variation calculated from the values of the ratio of  $N_{SPT}$  to  $N_c$ . In addition, the conversion error is proportional to  $N_c$  and corresponds to 0 when  $N_c = 0$ .

The relationship between fines content  $\log F_{cfc}$ , obtained from  $I_c$ , and the proper fines content,  $\log F_c$ , is given in Fig. 2.5.  $\log F_{cfc}$  is converted into  $\log F_c$  considering the conversion error by the following equation:

$$\log F_c = 2 - (2 - \log F_{cfc})(1 + 0.598\varepsilon_j) \quad (2.8)$$

where error term  $\varepsilon_j$  is assumed to follow standard normal distribution  $N(0,1)$ .  $\log F_{cfc}$  is assumed to be equal to  $\log F_c$ , and the quantity of the conversion error is given by the coefficient of variation calculated from the values for the ratio of  $\log F_{cfc}$  to  $\log F_c$ . In addition, the conversion error is in inverse proportion to  $\log F_c$  and corresponds to 0 when  $\log F_c = 2$ .

## 2.4 Conclusion

The cone penetration test (CPT) is suitable for investigations deep inside earth-fill dams. Since CPTs can provide quicker and more economical testing than boring tests, the testing intervals of CPTs in the horizontal direction can be significantly shorter than those of boring tests. Moreover, the considerable advantage of the CPT compared to other sounding tests is its ability to continuously and precisely collect data. Therefore, the CPT is capable of detecting the spatial variability of soil properties in detail. In addition, since several conversion formulas have been proposed to estimate soil properties using the results of CPTs, the many aspects of the soil properties can be evaluated with CPTs. For these reasons, CPTs are employed here to evaluate earth-fill dams.

## Chapter 3

# Fundamental theories of statistical modeling

### 3.1 Introduction

The statistical model is a method for modeling the measured values based on a probabilistic distribution. In this chapter, the fundamental theories of the statistical modeling of the soil properties are explained. The spatial variability of the soil properties is modeled using the random field theory. To choose the most appropriate model from the many candidates, the Akaike Information Criterion (AIC) is employed. The estimated statistical model of the random field can be introduced into the geostatistical method to visualize the spatial distribution of the soil properties considering the measured results.

The outline of this chapter is as follows. In section 3.2, the basic random field theory is presented. In section 3.3, the Akaike Information Criterion (AIC) is shown. In section 3.4, several techniques for the geostatistical method are given. Finally, the modeling procedure for the soil properties is explained in section 3.5.

### 3.2 Random field theory

When  $x$  is a real number as a form of scalar and  $B(x)$  behaves as a random variable against any fixed  $x$ ,  $B(x)$  is called the random process. In addition, in a similar manner, the soil properties in the same soil layer, which can continuously change, are defined as the random variables. They depend on a function of the spatial locations as the random field. In such cases, the random variable of the soil property in question is expressed by a function of spatial locations as  $\mathbf{X}=(x,y,z)$ . To describe the variability of the soil properties using the random field, the second order of the statistical values is normally employed, namely, mean, variance, and the autocorrelation function, as follows:

$$\text{Mean: } m(x,y,z) \tag{3.1}$$

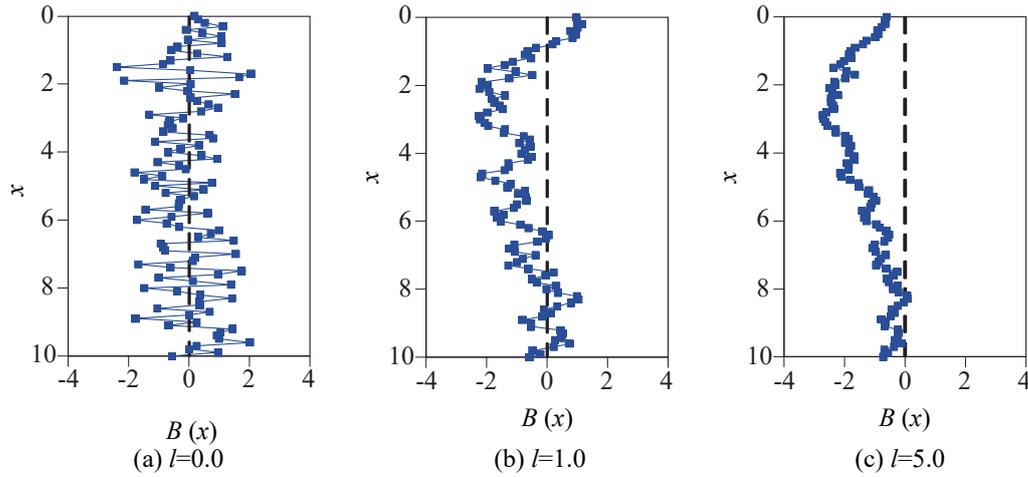


Fig. 3.1 Examples of random processes of different correlation distances

$$\text{Variance: } v(x, y, z) \tag{3.2}$$

$$\text{Autocorrelation function: } \rho(x, x'; y, y'; z, z') \tag{3.3}$$

The autocorrelation function expresses the coefficient of correlation between two points. The random field, called weakly stationary, has the following assumptions. The mean and the variance are assumed to be constant at any location, and the autocorrelation function does not depend on the locations of the two points, but only the distance between them. The theory of the random process (i.e., one-dimensional random field) and that of the random field are essentially the same. Thus, in order to give a concise description of the outline of the random field, the theory of the random field is explained one-dimensionally in the following.

The random field is significantly affected by the autocorrelation function. When the autocorrelation function is considered in one dimension, it depends on the difference value,  $\delta x$ , which is calculated by the distance between  $x$  and  $x'=x+\delta x$ . The value of the autocorrelation will decrease when the absolute value of  $\delta x$ ,  $|\delta x|$  increases. In the modeling of the soil properties using the random field, the practical form of the autocorrelation function consists of three types, namely, the delta function, the exponential function, and the Gaussian function. The equations for these functions are given as follows:

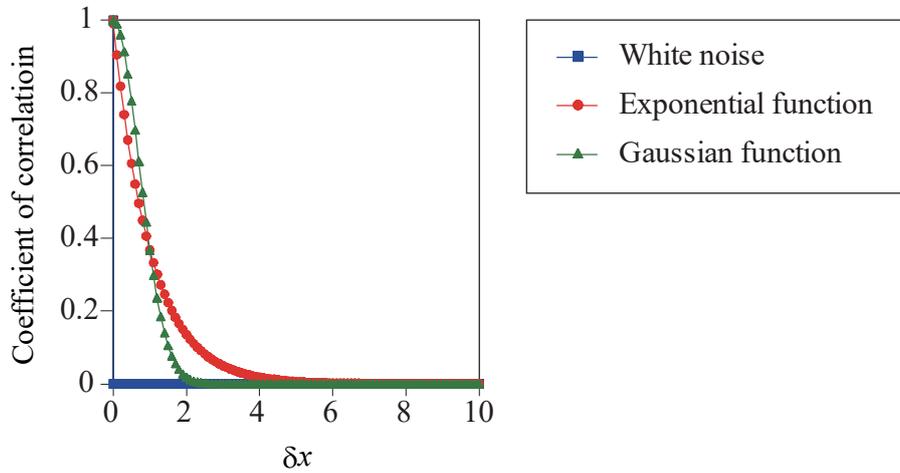
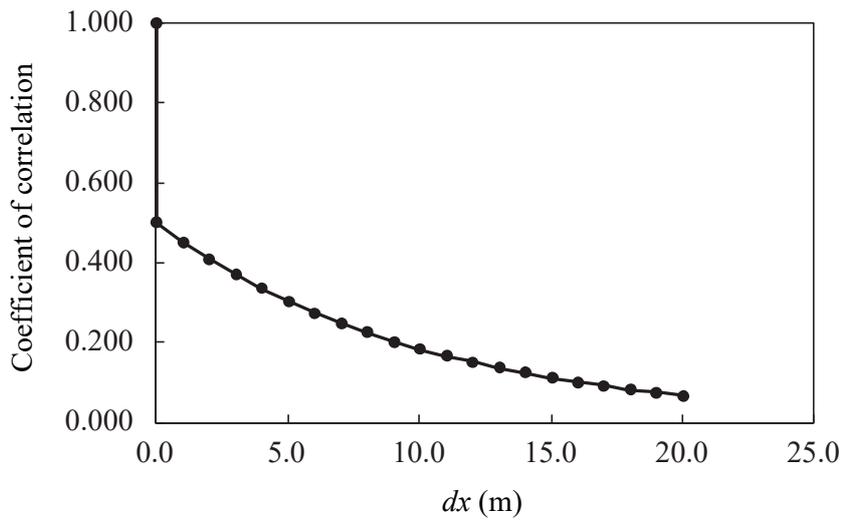


Fig. 3.2 Several types of autocorrelation functions



Exponential function with nugget effect ( $l=10$  m)

Fig. 3.3 Example of autocorrelation function

$$\text{delta function: } \rho(\delta x) = \begin{cases} 1 & \delta x = 0 \\ 0 & \text{otherwise.} \end{cases} \quad (3.4)$$

$$\text{exponential function: } \rho(\delta x) = \exp\left(-\frac{\delta x}{l}\right) \quad (3.5)$$

$$\text{Gaussian function: } \rho(\delta x) = \exp\left[-\left(\frac{\delta x}{l}\right)^2\right] \quad (3.6)$$

where  $\delta x$  is the distance between the two points and  $l$  is the correlation distance. For example, the soil properties between two points have a strong correlation in cases where the distance between the two points is shorter than the correlation distance. On the other hand, the correlation becomes small in cases where the distance between the two points is longer than the correlation distance. In other words, when the distance between the two points corresponds to  $l$ , the coefficient of correlation becomes 0.38. When the distance between the two points becomes even longer, the correlation eventually disappears. In Fig. 3.1, the simulation results of the normal random process are shown. The results are assumed to be  $m=0$  and  $v=1$ , and the autocorrelation functions are the delta function and the exponential function. The exponential function considers two types of correlation distances, namely,  $l=1$  or  $l=5$ . According to Fig. 3.1, the behaviors of the simulated values are different despite the fact that the value for the mean and the variance are the same. In addition, Fig. 3.2 shows the shapes of the autocorrelation functions of the white noise, the exponential function, and the Gaussian function, respectively. In this figure, the correlation distance of the exponential function and that of the Gaussian function are assumed to be 1 (unit length). When the distance between the two points is shorter than the correlation distance, the Gaussian function shows a higher correlation than the exponential function. On the other hand, this relationship between the two functions is reversed when the distance is longer than the correlation distance.

The discontinuity at the origin of the autocorrelation function is modeled as the nugget effect. The nugget effect is caused when the random field consists of a component that has a relatively long correlation distance and quite a short frequency component, like the white noise. In Fig. 3.3, the simulation results for the autocorrelation function, considering the nugget effect, are shown, and the equation is given as follows:

$$\rho(\delta x) = \begin{cases} 0.5 \exp\left(-\frac{\delta x}{l}\right) & (\delta x > 0.0) \\ 1.0 & (\delta x = 0.0) \end{cases} \quad (3.7)$$

This kind of correlation structure is confirmed in soil properties, for instance, the  $N$ -value. A long frequency component indicates the physical correlation structure of soil properties, while a short frequency component is assumed to represent the inherent randomness of the soil properties.

The past studies related to the evaluation of the spatial variability of soil properties assuming the random field are shown below. The physical uncertainty of the soil is modeled as a random field, which can be described concisely by the coefficient of variation (COV) and the correlation distance. Phoon and Kulhawy (1999) presented a summary of the COV of the physical uncertainty for various test measurements. The general soil type and the approximate range of mean values, for which the COV is applicable, were also presented. In addition, the typical correlation distances for a variety of common geotechnical parameters were summarized based on extensive literature reviews. However, it is apparent that the amount of information on the correlation distance was relatively limited in comparison to the amount of information on the COV of the physical uncertainty of soil. DeGroot and Baecher (1993) examined the effect of the intervals of the testing points to estimate the correlation distance. They concluded that it was important to provide information on the range in correlation distances in order to obtain good estimates.

### **3.3 Akaike Information Criterion (AIC)**

As a criterion for selecting a model from the several models obtained by the maximum likelihood method, the Akaike Information Criterion is used. The definition of the Akaike Information Criterion is explained in this section.

Phenomena that occur by accident can be treated as realized values of a random variable following a probability distribution. The model to approximate the distribution is also expressed in the form of a probability distribution. Therefore, the goodness of a model against the true probability distribution, which create data, can be evaluated by the goodness of fit of the probability distribution controlled by the model. In addition, applying data to a model can be regarded as the estimation of the true probability distribution of the data. In this manner, the evaluation of a model or the estimation of a model is performed by assuming that a true distribution and a model can be expressed as a probability distribution, respectively.

The Kullback-Leibler divergence was introduced by Kullback and Leibler (1951) as the directed divergence between two distributions. In the modeling of a true

distribution, the divergence can provide useful information for comparing the goodness of fit between the model and the true distribution. Although, in general, the true distribution is unknown when estimating the model, the goodness of fit can be approximately evaluated according to the large value of log-likelihood, which indicates that the model fits the true distribution well. In this manner, since a log-likelihood can be treated as the estimated value of the Kullback-Leibler divergence, the goodness of fit of the models based on different types can be compared. The joint probability density function of random variables  $(B_1, B_2, \dots, B_M)$  can be given as  $f(b_1, \dots, b_M | \theta)$ , where  $b$  is the observed outcome of an experiment and  $\theta$  is the parameter used to control the density function. In other words, when  $f(b_1, \dots, b_M | \theta)$  is viewed as a function of  $b$  with  $\theta$  fixed, it is a probability density function, and when it is viewed as a function of  $\theta$  with  $b$  fixed, it is a likelihood function, which is given by the following equation:

$$L(\theta) = f(b_1, \dots, b_M | \theta) \quad (3.8)$$

In particular, when the random variables are independent, the density function of  $(B_1, B_2, \dots, B_M)$  equals the product of these density functions  $B_i (i = 1, \dots, M)$ , and the equation is given as follows:

$$L(\theta) = f(b_1 | \theta) f(b_2 | \theta) \cdots f(b_M | \theta) \quad (3.9)$$

Taking the log of both sides, the following equation is derived:

$$l(\theta) = \sum_{i=1}^M \log f(b_i | \theta) \quad (3.10)$$

The log-likelihood can be used in place of the likelihood in maximum likelihood estimation. Finding the maximum of a function often involves taking the derivative of a function and solving for the parameter being maximized. This is often easier when the function being maximized is a log-likelihood function rather than an original likelihood function, because the probability of the conjunction of several independent variables is the product of the probabilities of the variables, and solving an additive equation is usually easier than solving a multiplicative one.

However, the maximum value of log-likelihood has the bias that the value

tends to increase depending on the number of parameters included in the estimated model. The definition of Akaike's Information Criterion, AIC, and the interpretation of AIC are explained below. To cope with the problem of the bias of maximum log-likelihood, AIC is defined by Eq. (3.11).

$$\begin{aligned} \text{AIC} = & -2 \times (\text{maximum log-likelihood}) \\ & + 2 \times (\text{number of parameters in model}) \end{aligned} \quad (3.11)$$

The AIC can be the criterion used to select a model. The model to minimize the AIC is called the Minimum AIC Estimator, MAICE (Akaike, 1974); it is regarded as an optimum model.

### 3.4 Geostatistical method

In the field of mining engineering, geostatistics was proposed in Matheron (1963) to evaluate the spatial distribution of the ore grade using the random field theory. Geostatistics can probabilistically describe the spatial distribution of estimation errors as the uncertainty of the soil properties. When the spatial distribution of the soil properties is estimated using the random field, information on the location of the in-situ testing data is not treated as a deterministic value. On the other hand, the interpolation method called "kriging" is one of the geostatistical methods. The estimated value by kriging at the sampled point coincides with the sample value. Therefore, the spatial distribution of the soil properties is treated as the sample field in the calculation process of kriging, and the spatial distribution of the soil properties can be estimated based on the in-situ data.

To model and visualize the spatial distribution of soil properties, three methods of geostatistics are utilized. First, the semi-variogram is used to evaluate the spatial correlation structure of the soil properties. Second, kriging is employed to estimate the spatial distribution of the soil properties considering the location of the data. Third, a conditional simulation is utilized. The simulation indicates the procedure for creating a sample field based on several statistical parameters, namely, mean, variance, and autocorrelation, respectively. In the simulation, the simulated value at the sample point coincides with the sample value.

The outline of this section is as follows. In sub-section 3.4.1, semi-variograms

are explained. Sub-section 3.4.2 presents the basic theory of kriging. Finally, the procedure for the conditional simulation is shown in sub-section 3.4.3.

### 3.4.1 Semi-variograms

In the characterization of the random field of weak-sense stationarity, autocorrelation function  $\rho(\delta x)$  or covariance function  $C(\delta x) = \sigma^2 \rho(\delta x)$ , where  $\sigma$  is a standard deviation, has a crucial role. In the geostatistical method, a semi-variogram is a functional equivalent to an autocorrelation function, and the semi-variogram can be used to identify the structure of the autocorrelation (Journel and Huijbregts, 1978). In a practical problem, the autocorrelation function and the semi-variogram are estimated from measured data using the following equations, respectively:

$$\hat{\rho}(\delta x) = \frac{\frac{1}{N_{\delta x}} \sum_{i=1}^{N_{\delta x}} [b(x_i + \delta x) - m][b(x_i) - m]}{\sum_{i=1}^{N_{\delta x}} [b(x_i) - m]^2} \quad (3.12)$$

$$\hat{\gamma}(\delta x) = \frac{1}{2N_{\delta x}} \sum_{i=1}^{N_{\delta x}} [b(x_i + \delta x) - b(x_i)]^2 \quad (3.13)$$

where  $N_{\delta x}$  denotes the number of combinations of  $[b(x_i + \delta x) - b(x_i)]$ . When  $B(x)$  is the random field of weak-sense stationarity, the relationship among the semi-variogram, the covariance function, and the autocorrelation function is given as follows:

$$\begin{aligned} \gamma(\delta x) &= \frac{1}{2} E \left\{ [B(x_i + \delta x) - B(x_i)]^2 \right\} \\ &= C(0) - C(\delta x) \\ &= \sigma^2 [1.0 - \rho(\delta x)] \end{aligned} \quad (3.14)$$

Fitting an appropriate model to the semi-variogram, the model provides a mathematical relationship that can be used to describe the difference in variance among the neighboring sample values with distance.

### 3.4.2 Kriging

To spatially interpolate the soil parameters obtained as point information, kriging is often employed. Kriging is one of the geostatistical methods. A 2-D statistical model is introduced here in this method with horizontal coordinate  $x$  and vertical coordinate  $z$ . Estimated parameter  $b^*$  is defined at any location as Eq. (3.15) in kriging.

$$b^* = \sum_{\alpha=1}^n \lambda_{\alpha} b_{\alpha} \quad (3.15)$$

in which  $\lambda_{\alpha}$  is an interpolation coefficient, which is a function of coordinates  $x$  and  $z$  at point  $\alpha$ , and  $b_{\alpha}$  is a sample value of a parameter at point  $\alpha$ . Symbol  $n$  signifies the number of sample points that are used for an interpolation. The kriging method is characterized by parameter  $\lambda_{\alpha}$  determined from the spatial correlation of the parameter values. In order to use Eq. (3.15), it is necessary to determine an appropriate  $\lambda_{\alpha}$ . The way to do this is indicated simply as seen below.

Although the expansion of the middle of the formula is omitted here (Journel and Huijbregts, 1978), the expected value for the square residual of a true random process  $b(x, z)$  and the interpolated value  $b^*(x, z)$  are given in the following equation.

The squared residual  $\sigma_k^2(x, z)$  between the true parameter value  $b$  and the estimated value  $b^*$  is defined as:

$$\begin{aligned} \sigma_k^2(x, z) &= E \left[ \left| b(x, z) - b^*(x, z) \right|^2 \right] \\ &= \sigma_b^2 - 2 \sum_{\alpha=1}^n \lambda_{\alpha} C_b(x, z; x_{\alpha}, z_{\alpha}) \\ &\quad + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} C_b(x_{\alpha}, z_{\alpha}; x_{\beta}, z_{\beta}) \end{aligned} \quad (3.16)$$

where variable  $\lambda_{\alpha}$  is determined by minimizing  $\sigma_k^2(x, z)$  in a restrained condition under which  $\sum_{\alpha=1}^n \lambda_{\alpha} = 1$ . Parameter  $\sigma_b^2$  is the priori variance of random variable  $B$ .

$$\frac{\partial}{\partial \lambda_\alpha} \left\{ E \left[ |b(x, z) - b^*(x, z)|^2 \right] - 2\kappa \left( \sum_{\beta=1}^n \lambda_\beta - 1 \right) \right\} = 0 \quad (3.17)$$

From Eq. (3.17), the following simultaneous linear equations are given, and coefficient  $\lambda_\alpha$  is obtained by solving Eq. (3.18).

$$\begin{aligned} & \sum_{\beta=1}^n \lambda_\beta C_b(x_\alpha, z_\alpha; x_\beta, z_\beta) - \kappa \\ &= C_b(x_\alpha, z_\alpha; x, z) \\ & \forall \alpha = 1, 2, \dots, n \\ & \sum_{\beta=1}^n \lambda_\beta = 1 \end{aligned} \quad (3.18)$$

in which  $\kappa$  is a Lagrange multiplier. Coefficient  $\lambda_\alpha$  is determined as a function of the spatial coordinates  $(x, z)$ .

### 3.4.3 Conditional simulation

Although the average value is estimated by Eq. (3.15) in kriging, the variance is not zero, except at the sampling point,  $\alpha$ , and then the interpolated value for the parameter has variability. Therefore, to consider the variability, many realizations are required and the simulation is repeated many times. In the present study, the sequential Gaussian simulation method is conducted; the analysis code SGSIM (Deutsch and Journel, 1992) is employed for the task. Realization  $b_c^{(l)}$ , by SGSIM, is given by Eq. (3.19); it is conditioned at the sample points so that the realized values coincide with the sample values.

$$b_c^{(l)}(x, z) = b^*(x, z) + b^{(l)}(x, z) - b^{*(l)}(x, z) \quad (3.19)$$

in which  $b_c^{(l)}(x, z)$  means the  $l$ th realization,  $b^*(x, z)$  is an estimated value by kriging, and  $b^{(l)}(x, z)$  is the  $l$ th realization without conditioning.  $b^{*(l)}(x, z)$  is the kriged estimation using the values of  $b^{(l)}(x, z)$  at the sampling points. On the right side of Eq. (3.19), the first item is an estimated value of kriging as an average. The remainder of the second and third items shows an estimation error of kriging.

### 3.5 Statistical modeling of ground

Based on the methods presented in previous sections, the procedure to statistically model the ground is explained below. In the modeling of the ground in question, the ground is assumed to be geologically homogeneous. To model the ground using the random field, the soil properties are assumed for the sum of the trend component and the random component. The trend component is given as the function of a certain coordinate point, and the random component is expressed as the random field, which is called weakly stationary. The equation to model the soil properties is generally given as follows.

A representative variable for the soil properties,  $b$ , is defined by Eq. (3.20) as a function of special location  $\mathbf{X}=(x, y, z)$ . Variable  $b$  is assumed to be expressed as the sum of mean function  $m$  and random variable  $U$ .

$$B(\mathbf{X}) = m(\mathbf{X}) + U(\mathbf{X}) \quad (3.20)$$

where the random variable function,  $B(\mathbf{X})$ , is discretized spatially into a random vector  $\mathbf{B}' = (B_1, B_2, \dots, B_M)$ , in which  $B_k$  is a point estimation value at location  $\mathbf{X}=(x_k, y_k, z_k)$ . Symbol  $M$  signifies the number of test points. The soil parameters, which are obtained from the tests, are defined here as  $\mathbf{b}' = (b_1, b_2, \dots, b_M)$ . Vector  $\mathbf{b}'$  is considered as a realization of the random vector  $\mathbf{B}' = (B_1, B_2, \dots, B_M)$ . If variables  $b_1, b_2, \dots, b_M$  constitute the  $M$  - variate normal distribution, the probability density function of random variable  $B$  can then be given by the following equation:

$$f(\mathbf{B}) = (2\pi)^{-M/2} |\mathbf{C}|^{-1/2} \times \exp \left\{ -\frac{1}{2} (\mathbf{B} - \mathbf{m})' \mathbf{C}^{-1} (\mathbf{B} - \mathbf{m}) \right\} \quad (3.21)$$

where  $\mathbf{m}' = (m_1, m_2, \dots, m_M)$  is the mean vector of random function  $\mathbf{B}' = (B_1, B_2, \dots, B_M)$ ; it is assumed to be the following approximation function. In this research, a 2-D statistical model is introduced to model the spatial distribution of the soil properties inside an earth-fill dam; it consists of horizontal coordinate  $x$ , which is parallel to the embankment axis, and vertical coordinate  $z$ . The other horizontal coordinate,  $y$ , which is perpendicular to the embankment axis, is disregarded.

$$m_k = a_0 + a_1 x_k + a_2 z_k + a_3 x_k^2 + a_4 z_k^2 + a_5 x_k z_k \quad (3.22)$$

in which  $(x_k, z_k)$  means the coordinate corresponding to the position of parameter  $B_k$ , and  $a_0, a_1, a_2, a_3, a_4$ , and  $a_5$  are the approximation coefficients.

$\mathbf{C}$  is the  $M \times M$  covariance matrix, which is selected from the following five types in this study:

$$\begin{aligned}
 \mathbf{C} &= [C_{ij}] = \\
 &\left\{ \begin{array}{l} \sigma^2 \exp(-|x_i - x_j|/l_x - |z_i - z_j|/l_z) \quad \text{(a)} \\ \sigma^2 \exp\left[-(x_i - x_j)^2/l_x^2 - (z_i - z_j)^2/l_z^2\right] \quad \text{(b)} \\ \sigma^2 \exp\left[-\sqrt{(x_i - x_j)^2/l_x^2 + (z_i - z_j)^2/l_z^2}\right] \quad \text{(c)} \\ N_e \sigma^2 \exp(-|x_i - x_j|/l_x - |z_i - z_j|/l_z) \quad \text{(d)} \\ N_e \sigma^2 \exp\left[-\sqrt{(x_i - x_j)^2/l_x^2 + (z_i - z_j)^2/l_z^2}\right] \quad \text{(e)} \end{array} \right. \quad (3.23) \\
 &i, j = 1, 2, \dots, M \\
 &\left\{ \begin{array}{l} N_e = 1 \quad (i = j) \\ N_e \neq 1 \quad (i \neq j) \end{array} \right.
 \end{aligned}$$

in which the symbol  $[C_{ij}]$  signifies an  $i$ - $j$  component of the covariance matrix,  $\sigma$  is the standard deviation, and  $l_x$  and  $l_z$  are the correlation lengths for the  $x$  and  $z$  directions, respectively. Parameter  $N_e$  is related to the nugget effect. Akaike's Information Criterion, AIC, is defined by Eq. (3.24) considering the log-likelihood of  $\mathbf{b}$ , which is modeled as the  $M$ -variate normal distribution.

$$\begin{aligned}
 \text{AIC} &= -2 \cdot \max[\ln f(\mathbf{b})] + 2K \\
 &= M \ln 2\pi + \min\left[\ln |\mathbf{C}| + (\mathbf{b} - \mathbf{m})^t \mathbf{C}^{-1} (\mathbf{b} - \mathbf{m})\right] + 2K \quad (3.24)
 \end{aligned}$$

in which  $K$  is the number of unknown parameters included in Eq. (3.21). According to the procedure to identify the Minimum AIC Estimator (MAICE), the approximation coefficients of the mean function, the number of approximation coefficients, the standard deviation,  $\sigma$ , a type of the covariance function, the nugget effect parameter,  $N_e$ , and the correlation lengths,  $l_x$  and  $l_z$ , are determined.

To execute the procedure of the MAICE, the correlation structure of the

multi-points distributed multi-dimensionally are evaluated simultaneously; and thus, this procedure sometimes creates difficulty. On the other hand, since calculating the semi-variogram (Journel and Huijbregts, 1978) is the method for identifying the correlation distance one-dimensionally, it is easier to use the semi-variogram than the procedure with the MAICE for finding the correlation distance. In the calculation of the semi-variograms, the measured values are assumed to comprise standard normal distribution  $N(0, 1)$ . For example,  $b$  is normalized as  $f = (b - m)/\sigma$  to remove the trend where  $m$  is the mean value and  $\sigma$  is the standard deviation; they are obtained from the procedure of the MAICE. In addition, in order to set the  $f$  values to be  $N(0, 1)$  with certainty, the following equation is used here:

$$Y = \Phi^{-1} [F(f)] \quad (3.25)$$

where  $F$  is the cumulative distribution function of  $f$  and  $\Phi$  is the standard normal distribution function.

The calculated values for the semi-variograms in the horizontal direction,  $\hat{\gamma}_x$ , and the depth direction,  $\hat{\gamma}_z$ , are defined by the following equations, respectively:

$$\begin{aligned} \hat{\gamma}_x(|x_i - x_j|) &= \frac{1}{2D_x} \sum_{k=1}^{D_x} [Y(x_k) - Y(x_k + |x_i - x_j|)]^2 \\ \hat{\gamma}_z(|z_i - z_j|) &= \frac{1}{2D_z} \sum_{k=1}^{D_z} [Y(z_k) - Y(z_k + |z_i - z_j|)]^2 \end{aligned} \quad (3.26)$$

where  $D_x$  and  $D_z$  denote the number of combinations of  $Y(x_k) - Y(x_k + |x_i - x_j|)$  and  $Y(z_k) - Y(z_k + |z_i - z_j|)$ , respectively, in which  $|x_i - x_j|$  and  $|z_i - z_j|$  mean the distance between the two points in the horizontal and vertical directions.  $Y(x_k)$  is  $|x_i - x_j|$  away from  $Y(x_k + |x_i - x_j|)$  in the same depth, and similarly,  $Y(z_k)$  is  $|z_i - z_j|$  away from  $Y(z_k + |z_i - z_j|)$  in the same horizontal coordinate. To identify the geostatistical parameters of a standardized value  $Y$ , the approximation curve of the semi-variograms in the horizontal and depth directions,  $\gamma_x$  and  $\gamma_z$  are simply modeled by exponential functions as the following equations, respectively:

$$\begin{aligned}
 \gamma_x(|x_i - x_j|) &= C_{0x} + C_{1x} \left[ 1 - \exp\left(-|x_i - x_j|/l_x\right) \right] \quad (i \neq j) \\
 \gamma_z(|z_i - z_j|) &= C_{0z} + C_{1z} \left[ 1 - \exp\left(-|z_i - z_j|/l_z\right) \right] \quad (i \neq j) \\
 \gamma_x(0) &= \gamma_z(0) = 0
 \end{aligned} \tag{3.27}$$

where  $C_{0x}$  and  $C_{0z}$  are the parameters used for the nugget effect in the  $x$  and  $z$  directions, respectively, while  $C_{1x}$  and  $C_{1z}$  are the parameters used to express the shapes of the semi-variogram functions.

Based on the semi-variograms in the  $x$  and  $z$  directions, covariance matrix  $\mathbf{C}$ , composed of the standardized values, is assumed to have the following form:

$$\begin{aligned}
 \mathbf{C} = [C_{ij}] &= N_e \sigma^2 \exp\left(-\frac{|x_i - x_j|}{l_x} - \frac{|z_i - z_j|}{l_z}\right) \quad (i \neq j) \\
 N_e &= C_{1x} C_{1z} \quad (|x_i - x_j| \neq 0, |z_i - z_j| \neq 0) \\
 N_e &= C_{1x} \quad (|x_i - x_j| \neq 0, |z_i - z_j| = 0) \\
 N_e &= C_{1z} \quad (|x_i - x_j| = 0, |z_i - z_j| \neq 0) \\
 [C_{ij}] &= \sigma^2 \quad (i = j)
 \end{aligned} \tag{3.28}$$

In Eq. (3.28), in order to faithfully express the correlation structure of the measured data, the anisotropy of the nugget effect is considered (e.g., Zimmerman, 1993; Banerjee et al., 2014). In general, the approximation curves against the semi-variograms are derived by the least squares method, and finally, parameters  $C_{0x}$ ,  $C_{1x}$ ,  $C_{0z}$ ,  $C_{1z}$ ,  $l_x$ , and  $l_z$  are determined.

## Chapter 4

### **Estimation of correlation lengths for several earth-fill dams**

#### **4.1 Introduction**

The spatial variability of the ground properties has been modeled as a random field, and the correlation distance is used as a parameter to represent the degree of the spatial correlation of the soil properties. In addition, the correlation distance is an essential parameter for describing how much the soil properties vary spatially in a random field. Therefore, in this study, focus is placed on the correlation distance of the soil strength inside earth-fill dams. However, DeGroot and Baecher (1993) pointed out that a testing program, which is shorter than the correlation distance, is necessary in order to properly estimate the correlation distance. Since standard penetration tests (SPTs) are generally conducted only at several places a few hundred meters into a river dike and at one to a few places in an earth-fill dam, the estimation of the correlation distance is difficult due to the limited results.

The modeling of the soil properties based on the results of CPTs has been carried out in many past works. For example, Bong and Stuedlein (2017) conducted CPTs on silty sand, which may liquefy, and examined the correlation distances of the soil strength in the vertical and horizontal directions, respectively. In addition, the authors studied a method to evaluate the correlation distance of the soil strength inside earth-fill dams through the use of the maximum likelihood method and the geostatistical method based on the results of CPTs (Imaide et al., 2018; Ueta et al., 2018).

The objectives of this study are to collect the correlation distances of the soil strength at five earth-fill dams in accordance with the proposed method by the authors and to provide the basic data necessary to appropriately evaluate the spatial variability of the soil strength. To achieve these goals, an outline of a site investigation using CPTUs at the studied sites is given in section 4.2. In section 4.3, the correlation distance of the soil strength is calculated for each result obtained from the CPTUs at five earth-fill dams, and the distances are utilized to determine the statistical model of the soil strength. The spatial distribution of the soil strength can be visualized in detail by applying the statistical model estimated in section 4.3 to the geostatistical method. As an example of the application of the correlation distance of the soil strength, in section 4.4, the spatial distribution of the soil strength at an earth-fill dam is evaluated using the

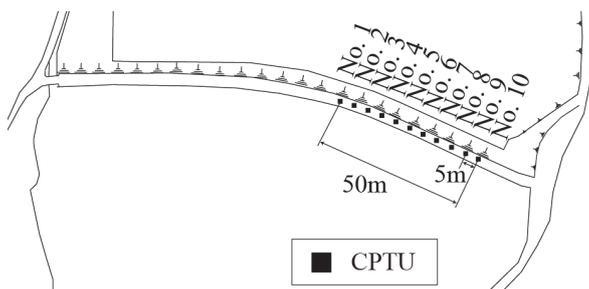
geostatistical method. Furthermore, the interpolated value by the geostatistical method and the measured value are compared at the same location. Finally, the conclusions of this chapter are given in section 4.5.

## 4.2 Site investigation

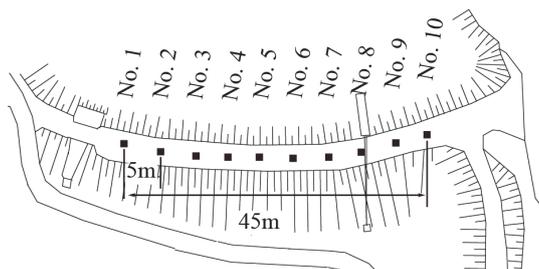
First, for all five earth-fill dams to be surveyed, the plan views of the testing site and the geological cross sections are shown in Figs. 4.1 and 4.2, respectively. Since the  $N$ -value is generally used in Japan as an index for the soil strength, the soil properties obtained from the CPTUs are converted into  $N$ -values based on the conversion formulas presented in section 2.3.

Figs. 4.1 (a) - (e) show the plan views of the testing points with A pond to E pond, respectively. Moreover, Fig. 4.2 (a) - (e) show the geological cross section of each of the surveyed five earth-fill dams with A pond to E pond. The Bs layer of each dam is an intermediate soil containing a large amount of clayey soil. In these investigations, CPTUs were carried out at 2-m, 5-m, and 50-m intervals at the top of the dam. From A pond to D pond, CPTUs were carried out at regular intervals of 2 m or 5 m. On the other hand, two kinds of test spans were set for E pond in order to examine the relatively long embankment. One was a high-density test span, where CPTUs were performed at 5-m intervals, as part of the whole investigation, and the other was a low-density test span, where CPTUs were performed at 50-m intervals.

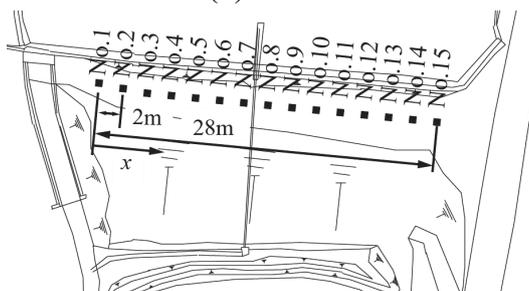
In this thesis, the soil strength of the earth-fill dams was evaluated by calculating the  $N$ -value from the soil properties obtained from the CPTUs. This is because, in Japan, the  $N$ -value obtained by SPTs is generally used to evaluate the soil strength for the design. As explained in Chapter 2, the  $N_c$  is calculated with Eq. (2.5), which is a conversion formula proposed by Suzuki et al. (2003).



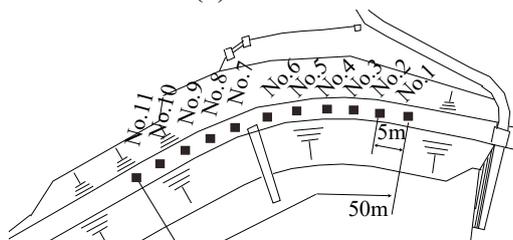
(a) A dam



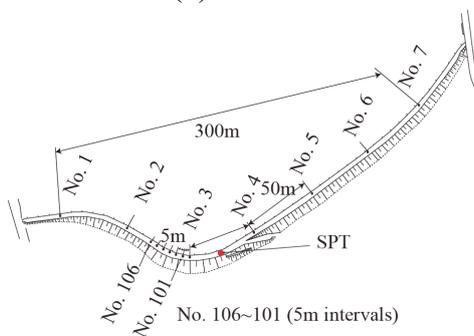
(b) B dam



(c) C dam

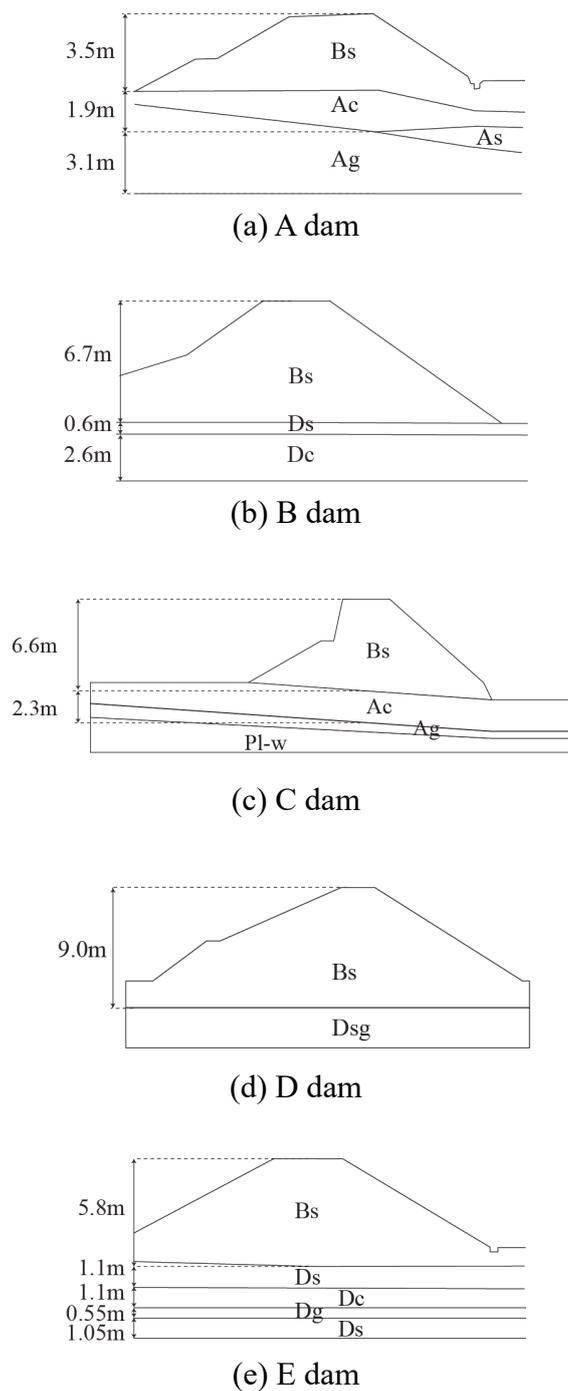


(d) D dam



(e) E dam

Fig. 4.1 Plan views



Bs : Back fill            Ac : Alluvial clay  
 As : Alluvial sand      Ag : Alluvial gravel  
 Dc : Diluvial clay      Ds : Diluvial sand  
 Dg : Diluvial gravel    Dsg : Diluvial sandy gravel  
 Pl-w : Weathered slate

Fig. 4.2 Cross section of each dam and legend

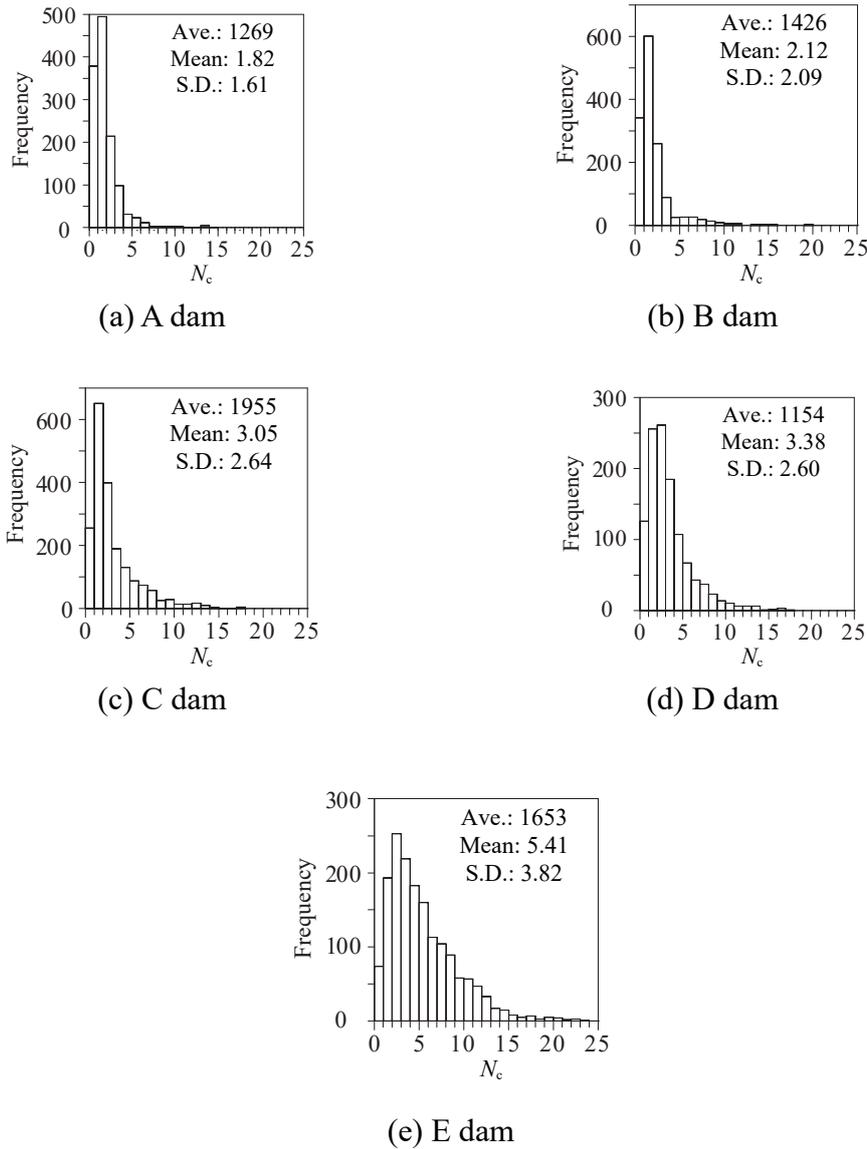


Fig. 4.3 Histograms of  $N_c$

Fig. 4.3 (a) - (e) show the histograms of  $N_c$  for each of the five dams. It is confirmed that the shape of the histograms of  $N_c$  for all the dams has a distribution close to the lognormal distribution. From the viewpoint of the testing procedure for SPTs, the  $N$ -value is considered to be the average measured value of the ground, inherently 30 cm. However, in order to satisfy, at least to some extent, both the high applicability of the  $N$ -value and the rigor in terms of geostatistics, as an eclectic concept, the results of the CPTUs translate into  $N$ -values, but make use of the spatial information of  $q_t$ . In other words, when considering the application of the spatial structure to the geostatistical simulation and the like, it is possible to mathematically handle it more rigorously by evaluating the  $N$ -value at intervals as close to the point estimation as possible.

### 4.3 Statistical modeling of CPT results

The spatial distribution of the soil strength can be estimated by a statistical model. As the data obtained by sounding tests, such as CPTs, are the values for the point estimation, it is necessary to spatially interpolate the data using the statistical model to evaluate the spatial variability of the soil strength. The statistical model is a modeling method based on the random field theory; it is possible to estimate the value at a point where data are not obtained, as a realized value given by the random field.

In section 4.3, a procedure is briefly described to statistically model  $\log N_c$ , which is taken as the logarithm of the  $N$ -value converted from CPTUs, and finally, the estimated statistical models are presented. Here,  $\log N_c$  is employed for the model so that  $N_c$  is not less than 0, while the model is assumed as a two-dimensional model, which is limited to the direction of the longitudinal section of the dam. (i.e., the direction of embankment axis is shown as  $x$  and the depth direction is given as  $z$ , respectively.)

First, statistical modeling will be described. Here, an optimum model is determined by the MAICE described in Chapter 2. With the MAICE, the correlation structure of the multi-points distributed multi-dimensionally are evaluated simultaneously, and this procedure sometimes creates difficulty. Therefore, in this study, when the correlation distance cannot be properly identified by the MAICE, the semi-variogram (Journel and Huijbregts, 1978), which is one of the geostatistical methods, is used to identify the spatial structure one-dimensionally.

Lark (2000) pointed out that semi-variograms are susceptible to outliers. It can be difficult to properly estimate the correlation distances when the semi-variogram is contaminated by outliers. Therefore, the covariance function is determined based on the semi-variogram, which secured the stationarity, by excluding outliers here. In order to deal with this problem, random variable  $b$ , which represents the soil properties, is transformed as seen below. First, the mean function,  $m$ , and the standard deviation,  $\sigma$ , are determined by the MAICE. Next, a normalized random variable,  $f = (b(x, z) - m(x, z)) / \sigma$ , is defined by using the mean function,  $m$ , and the standard deviation,  $\sigma$ . With reference to the results of the influence of outliers on the semi-variogram confirmed by Imaide et al. (2018), it was defined that the outliers are less than or equal to the 5 percentile value of the  $f$  and are more than or equal to the 95 percentile value of the  $f$ , respectively. Based on this criterion, the outliers were removed from the normalized random variable,  $f$ . Although, the calculation of the semi-variogram assumes that  $f$  follows  $N(0, 1)$ , to estimate a correlation distance, the variance of  $f$  decreases by the removal of the outliers. Therefore,  $f$  is surely transformed into the

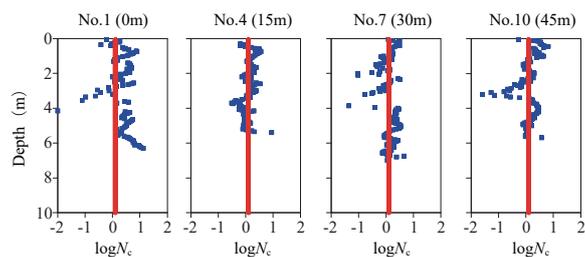
Table 4.1 Statistical models of  $\log N_c$  estimated by MAICE

Mean function	Covariance function ( $i, j = 1, 2, \dots, M$ )
A dam $m = 1.028$	$C = [C_{ij}] = 0.682^2 \exp\left(-\sqrt{ x_i - x_j ^2 / 10.00^2 +  z_i - z_j ^2 / 0.51^2}\right) \quad (i \neq j)$ $C = 0.682^2 \quad (i = j)$
B dam $m = 0.259$	$C = [C_{ij}] = 0.461^2 \exp\left(- x_i - x_j ^2 / 8.11^2 -  z_i - z_j ^2 / 0.52^2\right) \quad (i \neq j)$ $C = 0.461^2 \quad (i = j)$
C dam $m = 3.207$	$C = [C_{ij}] = 0.371^2 \exp\left(- x_i - x_j  / 0.01 -  z_i - z_j  / 0.21\right) \quad (i \neq j)$ $C = 0.371^2 \quad (i = j)$
D dam $m = 0.903 - 0.014x - 0.273z$ $+ 0.0002x^2 + 0.026z^2 + 0.002xz$	$C = [C_{ij}] = 0.365^2 \exp\left(- x_i - x_j  / 0.01 -  z_i - z_j  / 0.20\right) \quad (i \neq j)$ $C = 0.365^2 \quad (i = j)$
E dam $m = 0.409 + 0.051z$	$C = [C_{ij}] = 0.471^2 \exp\left(-\sqrt{ x_i - x_j ^2 / 4.88^2 +  z_i - z_j ^2 / 0.50^2}\right) \quad (i \neq j)$ $C = 0.471^2 \quad (i = j)$

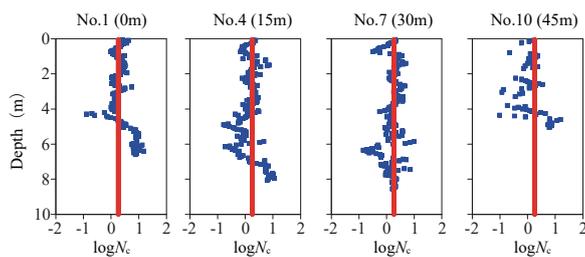
$M$ : Number of measured points,  $x$ : Horizontal coordinate (m),  $z$ : Depth(m).

variable of standard normal distribution  $Y$  by using the normal transformation by the Eq. (3.26). The procedure for the normal transformation is presented in Chapter 2.

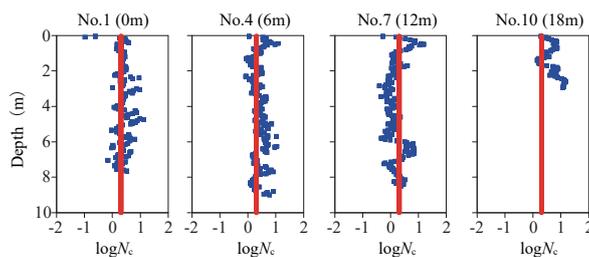
By applying the MAICE and the semi-variogram, the statistical model for  $\log N_c$  inside the dams in question were estimated as below. First, Table 4.1 shows the statistical model determined by the MAICE, and Fig. 4.4 indicates the mean function identified by the MAICE and the spatial distribution in the depth direction of the measured values, respectively. In addition, Fig. 4.5 presents the semi-variogram for every dam; each was calculated for the standardized variable,  $Y$ , and the approximated function of the semi-variogram. Finally, Table 4.2 shows the parameters of the covariance function, which were obtained by the fitting of the approximation function to the semi-variogram.



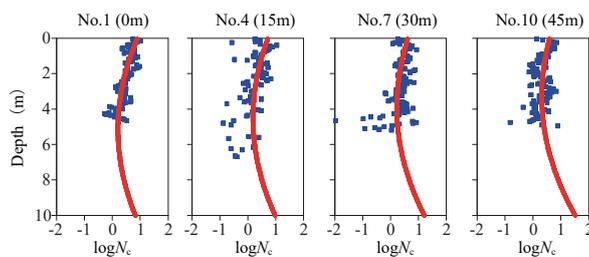
(a) A dam



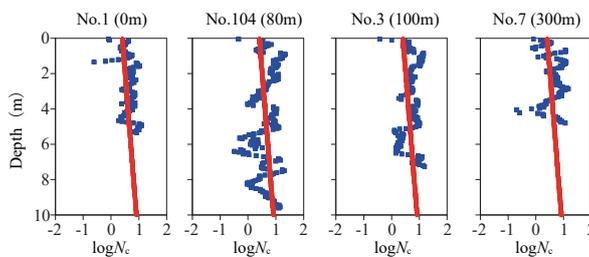
(b) B dam



(c) C dam

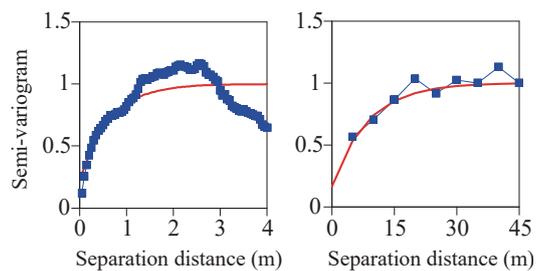


(d) D dam

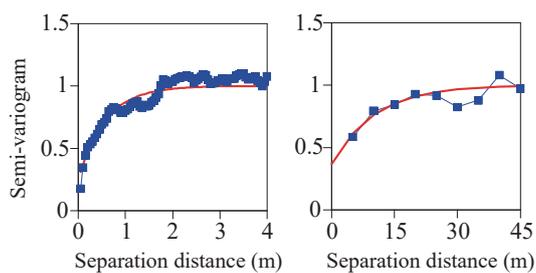


(e) E dam

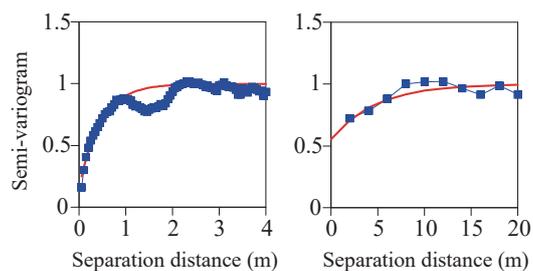
Fig. 4.4 Values measured by CPTUs and mean functions of  $\log N_c$



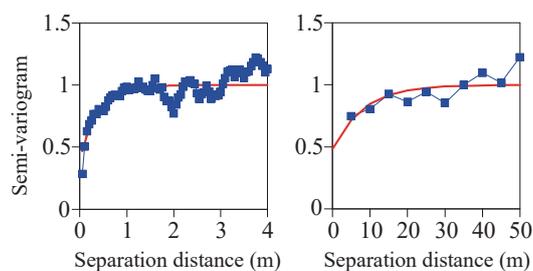
(a) A dam



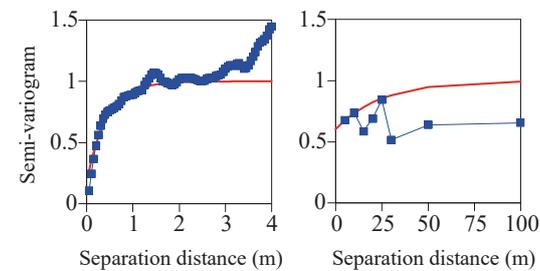
(b) B dam



(c) C dam



(d) D dam



(e) E dam

Fig. 4.5 Semi-variograms and approximation function of  $Y$

Table 4.2 Constants of covariance functions of  $\log N_c$  determined by semi-variograms

	$C_0$	$C_1$	Correlation distance (m)
A dam	$C_{0x}=0.17, C_{0z}=0.22$	$C_{1x}=0.83, C_{1z}=0.78$	$l_x=8.58, l_z=0.63$
B dam	$C_{0x}=0.37, C_{0z}=0.25$	$C_{1x}=0.63, C_{1z}=0.75$	$l_x=10.25, l_z=0.57$
C dam	$C_{0x}=0.55, C_{0z}=0.16$	$C_{1x}=0.45, C_{1z}=0.84$	$l_x=4.72, l_z=0.46$
D dam	$C_{0x}=0.48, C_{0z}=0.39$	$C_{1x}=0.52, C_{1z}=0.61$	$l_x=8.29, l_z=0.39$
E dam	$C_{0x}=0.61, C_{0z}=0.18$	$C_{1x}=0.39, C_{1z}=0.82$	$l_x=25.00, l_z=0.44$

By using the semi-variogram, the correlation distance of the soil strength could be appropriately evaluated, for the following two reasons. First, a reliable semi-variogram was used to identify the correlation distance. Second, it was confirmed that the estimated value of the correlation distance was consistent with the value given in past literature.

Uzielli et al. (2005) pointed out that the reliability of the semi-variogram decreases as the distance between two data increases. Also, the distance between two data is defined as the separation distance in this paper. For that reason, by using the least squares method, the approximation function of a semi-variogram in the depth direction was determined considering the data, which had a separation distance up to 0.9 m in the vertical direction, in the calculations of all the dams. The separation distance of 0.9 m corresponded to about 1.5 to 2 times the correlation distance in the depth direction, and the correlation between the two points almost disappeared. On the other hand, in the horizontal direction, the data which had a separation distance of 15 m for the A, B, and D ponds, of 6 m for the C pond, and of 10 m for the E pond, were needed. In other words, the separation distance of the data used to obtain the approximate function was different among the dams. The reason is because the data interval between the two points was assumed to be one span, and since the first three spans were considered to have relatively large numbers of data, three spans were used to determine the approximate functions. In the E pond, since the correlation tended to decrease up to two spans in proportion to the distance between the data, two spans were used to determine the approximate function. By handling the data in such a way, the correlation distance obtained from the semi-variogram tended to be longer than that from the MAICE.

In addition, it was confirmed that the correlation distance obtained here roughly corresponds to the results of past findings. At first, Nishimura et al. (2016) and Nishimura et al. (2011) can be cited as studies that address the soil strength of earth-fill

dams. Nishimura et al. (2016) conducted Swedish weight sounding (SWS) tests to calculate the  $N$ -value obtained from SWS,  $N_{\text{sws}}$ . The correlation distances of  $N_{\text{sws}}$  in the vertical direction were estimated to be 2.06 - 2.66 m, and those in the horizontal direction were estimated to be 9.88 - 27.1 m. Moreover, Nishimura et al. (2011) organized the correlation distances of  $N_{\text{sws}}$  with respect to three earth-fill dams. The correlation distances in the vertical direction were 0.46 - 4.76 m, and those in the horizontal direction were 9.3 - 62.5 m. Bombasaro and Kasper (2016), Stuedlein et al. (2012), Uzielli et al. (2005), etc. can be cited as studies that target viscous soils, such as those used for the materials of earth-fill dams. For example, Bombasaro and Kasper (2016) examined the correlation distance of normalized cone penetration resistance  $q_n$  to discuss the soil strength for marine clay. The correlation distance of  $q_n$  in the vertical direction was 0.08 - 0.78 m, and that in the horizontal direction was 12.15 - 15.67 m, respectively. As described above, it was confirmed that the correlation distances identified here are reasonable, because of the agreement with the correlation distance of the findings of past researches on earth-fill dams and viscous soil. In addition, the correlation distance in the horizontal direction is larger by one order or more than the vertical one; this also coincides with past findings.

#### **4.4 Evaluation of spatial distribution of $N$ -value using conditional simulation**

The geostatistical method can provide spatially interpolated values based on point estimates of the soil properties. Assuming that the spatial distribution of the soil properties is given as a random field, the population that satisfies the probability characteristics given by the mean, the variance, and the covariance function of the estimated statistical model, is the object of analysis. However, it is reasonable to think that the random field to be analyzed here is not a population, but a sample process (one sample from a population), in which some soil properties are known by in-situ investigation. Therefore, a geostatistical method is introduced as a solution to treat the obtained sample data as a fixed value of the random field and to estimate the spatial distribution of the soil properties. The geostatistical method has been used in numerous studies, such as to evaluate the spatial distribution of the soil properties and that of the liquefaction probability (for example, Suzuki and Ishii, 1988; Otake et al., 2014; Chen et al., 2015).

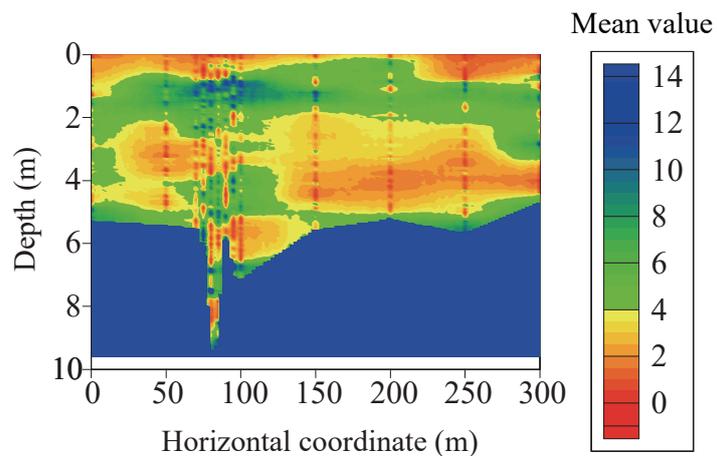
In this section, first, the geostatistics method used to evaluate the spatial

distribution of the soil strength is briefly described. Next, the spatial distribution of the soil strength obtained by applying the statistical model for E pond to the geostatistical method is shown. Furthermore, the interpolated value is compared with the value measured in the standard penetration test (SPT) at the same location.

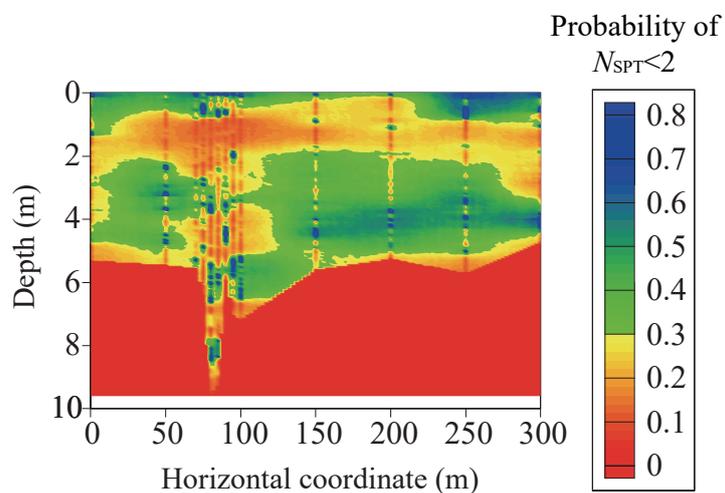
The spatial distribution of the soil properties can be thought of as one sample field of a random field. A simulation is done to create another sample field having the same probability characteristics. A conditional simulation treats the measured data as deterministic values, and assumes that the uncertainty of the soil properties is a problem only when estimating values other than the sampled points. As the code to perform the conditional simulation, SGSIM (Deutsch and Journel, 1992) is used.

Since the calculation procedure for preparing a sample field using the conditional simulation was described in Chapter 3, the procedure will not be repeated here. Only the simulation conditions are explained. As an example, the E pond with a relatively long length was examined; the direction of the longitudinal section of the E pond (i.e., the direction of the dam axis is shown as  $x$  and the depth direction is shown as  $z$ ), was analyzed. In the analysis section, the output points of the simulated results were set to a lattice shape with an interval of 2.5 m in the dam axis direction and an interval of 0.05 m in the depth direction. The output interval set here is determined based on the interval for data acquisition in the CPTUs. The simulation was repeated 2000 times using the Monte Carlo method, and the spatial distribution of the statistics of the soil strength was evaluated.

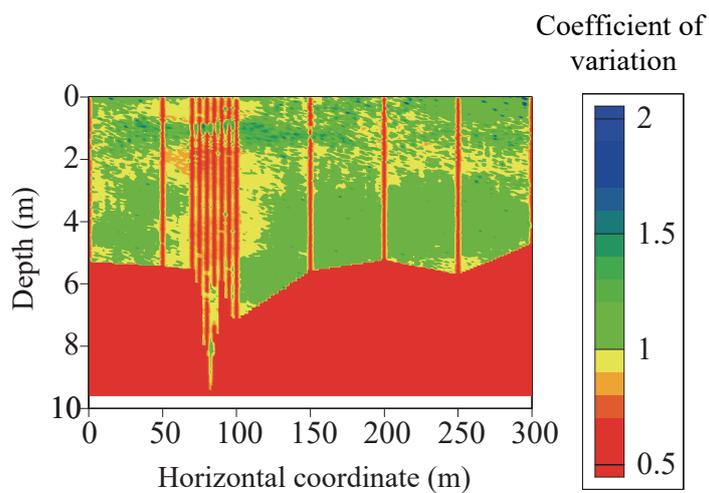
Table 4.3 shows the statistical model of the E pond used in the simulation, and Fig. 4.6 (a) - (c) show the spatial distribution of the statistics of the  $N$ -value obtained by applying the statistical model to the conditional simulation, respectively. In Table 4.3, the mean function and the standard deviation were determined by the MAICE, and the covariance function was determined by the semi-variogram. Fig. 4.6 (a) shows the spatial distribution of the mean values for the  $N$ -value. Fig. 4.6 (b) presents the spatial distribution of the probability that  $N_{\text{SPT}} < 2$ , and Fig. 4.6 (c) indicates the spatial distribution of the coefficient of variation. According to Fig. 4.6 (a), the  $N$ -value inside the dam is mostly less than 8, and the strength is not high as a whole. Fig. 4.6 (b) presents two areas where the probability that  $N_{\text{SPT}} < 2$  is high. The first area is located within  $x=70-100$  m and  $z=5.5 - 6.5$  m, and the other is located within  $x=150-300$  m and  $z=3.5 - 5.0$  m. In Fig. 4.6 (c), the coefficient of variation is about 0.5 at the points where the measured values obtained from the CPTUs exist, while it was confirmed that the coefficient of variation is relatively large at the interpolated points.



(a) Expected value



(b) Probability of  $N_{SPT} < 2$



(c) Coefficient of variation

Fig. 4.6 Spatial distribution of statistics of  $N$ -value

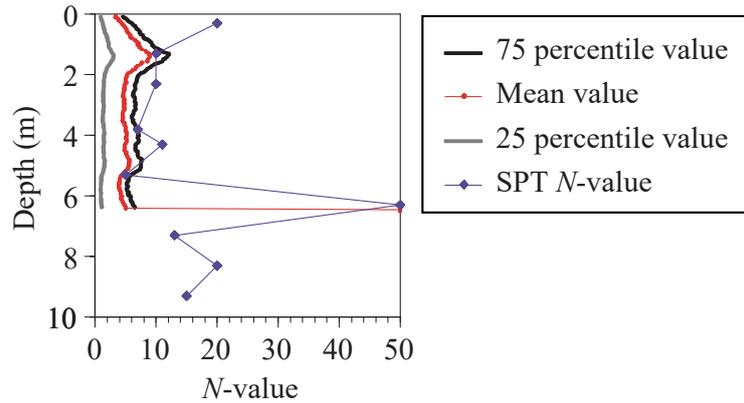


Fig. 4.7 Values measured by SPT and statistical values of interpolated values at  $x=125\text{m}$

Table 4.3 Statistical models for  $\log N_e$  introduced into simulation

Mean function	S.D.
$m = 0.409 + 0.051z$	$\sigma = 0.471$
Covariance function ( $i, j = 1, 2, \dots, M$ )	
$C = [C_{ij}] = \sigma^2 \cdot N_e \exp(- x_i - x_j /25.00 -  z_i - z_j /0.44) \quad (i \neq j)$	
$N_e = 0.322 \quad ( x_i - x_j  \neq 0,  z_i - z_j  \neq 0)$	
$N_e = 0.394 \quad ( x_i - x_j  \neq 0,  z_i - z_j  = 0)$	
$N_e = 0.816 \quad ( x_i - x_j  = 0,  z_i - z_j  \neq 0)$	
$C = \sigma^2 \quad (i = j)$	

S.D.: Standard deviation,  $x$  : Horizontal coordinate (m),  $z$  : Depth (m),

$M$  : Number of measured points,  $N_e$  : Nugget parameter.

In addition, the  $N$ -value measured in the standard penetration test is compared with the  $N$ -value calculated in the CPTUs at the same location. Fig. 4.6 shows the depth distribution of the  $N$ -values of the interpolated values and measured values at  $x = 125$  m. From this figure, it was confirmed that the tendencies of the interpolated values and the measured values are different at the firm part near the ground surface. This is considered to be due to the fact that the measured value of the SPT at  $x=125$  m is greatly different from the measured values of the CPTUs at  $x=100$  m and  $x=150$  m. At the depth of  $z = 1.3$  m or more in Fig. 4.6, the  $N$ -value obtained by the conditional simulation shows a slightly smaller value, on the whole, than the SPT  $N$ -value. However, since it was assumed that the CPTUs could not penetrate  $z = 6.4$  m or deeper at  $x=125$  m, the

$N$ -value calculated from the CPTUs was given as  $N=50$ . Moreover, on the whole, a large variation is seen in the interpolated value. This is because the CPTU results do not exist around  $x=125$  m and the error to convert  $N_c$  into the  $N$ -value is taken into account by Eq. (2.7). In Fig. 4.6, the depth distributions of the 25 percentile values and the 75 percentile values are presented. It can be seen that the  $N$ -value calculated from the CPTUs varies from 0 to about 10 at any depth. In order to reduce the variations in the interpolated values, additional investigations around  $x = 125$  m would be effective.

## 4.5 Conclusion

In this chapter, special focus was placed on earth-fill dams and the correlation distance of the soil strength was examined. For that purpose, the CPTUs were carried out at short intervals at five earth-fill dams, and a statistical model of the soil strength was estimated for each of the dams. Moreover, as an application of the statistical model, the spatial distribution of the soil strength was visualized in detail using a conditional simulation, which is one type of geostatistical method. Finally, the interpolated value obtained by the conditional simulation and the measured value obtained from the SPT, taken at the same location, were compared. The concluding remarks are given below.

1) A method to calculate the  $N$ -value, considering the quantity of the conversion error from the results obtained by the CPTUs, was presented. The method enables the determination of the statistical model of the  $N$ -value based on sufficient information.

2) The correlation distances in the vertical direction and the horizontal direction were estimated for each earth-fill dam using the semi-variogram that ensured the stationarity of the data by removing the outliers. It was confirmed that the obtained correlation distances here are consistent with those in past findings for earth-fill dams and viscous soil.

3) Based on the detailed evaluation of the correlation distances of the dams, the generalized values of the correlation distances were confirmed, as seen below. Except for the E dam, the range in the correlation distances in the depth direction was about 0.4 m to 0.6 m. Similarly, the range in the horizontal direction was about 5 m to 10 m. Moreover, the range of the ratio between  $l_x$  and  $l_z$  was 10 to 20 times.

The correlation distances of the soil strength at the five earth-fill dams were summarized here. Although the database was found to be insufficient in the present situation, a contribution has been made, namely, a database for the correlation distances that can be referred to when evaluating the spatial variation in the soil strength. It would

actually be very difficult to evaluate the spatial variability of the soil strength for each and every earth-fill dam in Japan based on site investigations. However, even for an earth-fill dam where a site investigation is not possible or is insufficient, it is desirable to take into account the spatial variability of the soil strength by using the general value of the correlation distance. Therefore, the database containing the correlation distances of the soil properties should be continually expanded and updated.

## Chapter 5

# Evaluation of spatial variability of cone penetration resistance inside earth-fill dams composed of materials with different particle size distributions

### 5.1 Introduction

In this study, the spatial variability of the cone penetration resistance inside earth-fill dams, which are composed of materials with different particle size distributions, was evaluated with use of the CPTU. To address this problem, the probability theory is employed here to model the spatial variability of the ground. Many of the methods using a stochastic model assume the weak-sense stationarity (i.e., second-order stationarity). Therefore, a stochastic model, such as a random field, is characterized by the first- and second-order moments with respect to the difference value between two arbitrary points, and the spatial structure parameter for expressing the autocorrelation of the soil properties becomes important. Fig. 5.1 shows the concept of the distribution of the tip resistance obtained from the CPT. The earth-fill dams in Japan, for instance, sometimes partially contain gravel, as gravel is commonly used to reinforce soil structures. Due to the outliers caused by gravel, it can become difficult to evaluate the spatial variability of the soil strength based on the random field theory of weak-sense stationarity. Lark (2000) examined the influence of the inclusion of outliers on the estimation of the robust types of semi-variograms using the data, which intentionally includes outliers. As a result, Lark pointed out that, due to the influence of the outliers, it can become difficult to properly estimate the parameters, which are important for

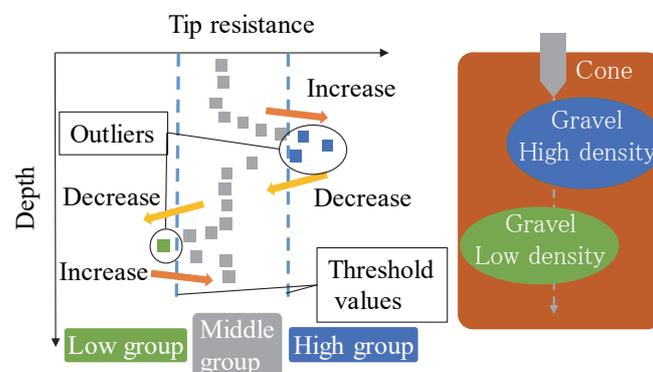


Fig. 5.1 Distribution of tip resistance in cone penetration test (CPT)

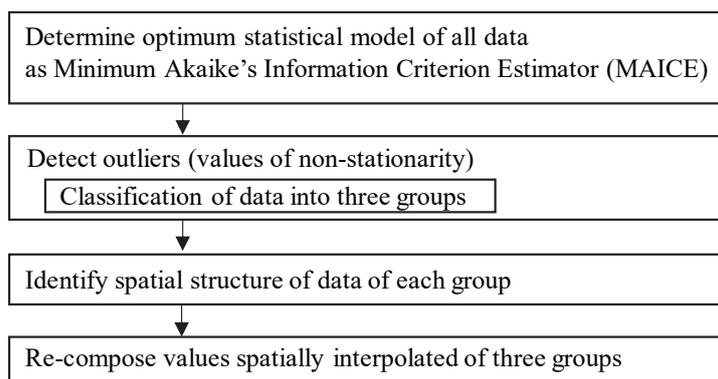


Fig. 5.2 Flowchart to model data of soil strength

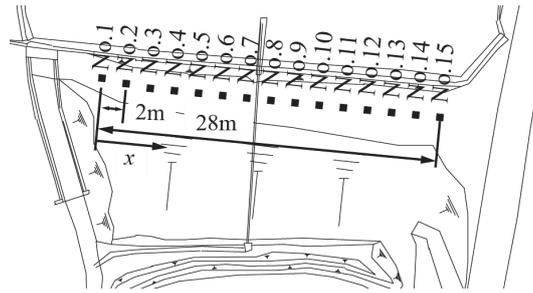
evaluating the heterogeneity of a ground. Therefore, as shown in Fig. 5.1, the outliers are defined in this study as a part of the spike-like distributed data observed by the extent to which the gravels are included at the measured points.

As a method to model the data that include outliers, a theory of robust estimation can be given. In the robust estimation, the data including outliers are weighted, and the entire data are modeled as one probability distribution. On the other hand, in this study, as shown in Fig. 5.1, the threshold values to classify the data are divided into three groups, i.e., a high group, a middle group, and a low group, respectively. In addition, a method is proposed to re-compose the probability distribution of each of the three groups. The purpose of the re-composition is twofold. The first objective is to secure the stationarity of the random field, with the exception of the outliers. By classifying the outliers, the data in the middle group can secure the stationarity, and the parameters of the spatial structure of the data can be appropriately evaluated using the semi-variogram. In other words, the semi-variograms with less disturbance of the data due to outliers can be obtained, and the proper correlation distance of the soil properties can be estimated. The second objective is to equally and separately model the data which are classified as outliers, because the outliers play an important role in the evaluation of the structural stability. In the existing method, the weight of the information on the outliers is generally reduced for estimating a statistical model. However, in the proposed method, the data outside the threshold values, namely, the data included in the high group and the low group in Fig. 5.1 are equally modeled. Generally, since the data classified as outliers have few observed numbers, a smooth probability density function cannot be obtained. However, the proposed method faithfully reproduces the frequency distribution of the obtained data at least to try to use the information obtained from the measured values for the evaluation of the stability analysis.

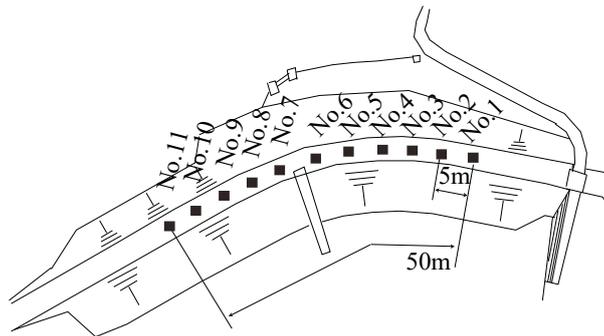
Fig. 5.2 shows the flowchart of this study. First, the threshold values divide the data into the three groups, namely, the high group, the middle group, and the low group, and the statistical model for each group is determined. Furthermore, based on the statistical model for each group, the geostatistical method is used to estimate the spatial distribution of the soil strength for each group. The unique feature of this study is in the re-composition of the spatial distributions of the soil strengths of the three groups, and the aim of the procedure is to equally consider the outliers. The locations where the outliers appear are evaluated as follows. As presented in Fig. 5.1, the spike-like distribution of the tip resistance is shown to be affected by the outliers. Using this kind of tendency of the distribution of the obtained data, the points where the outliers are likely to exist can be evaluated. In order to validate the applicability of the proposed method, the soil strength estimated by the proposed method was compared with the measured data taken at the same location. The estimation error for each estimated value and the frequency distribution of the estimated values were also investigated, and the accuracy of the estimated value calculated by the proposed method was confirmed.

## 5.2 Site investigation

In this study, the C dam and the D dam were examined. The C dam is 6.6 m in height and the back fill layer has a fines content of 48.5%; it is classified as sand of a fine fraction nature with gravel. The D dam is 9.0 m in height and the back fill layer has a fines content of 47.5%; it is also classified as sand of a fine fraction nature with gravel. Figs. 5.3 (a) and (b) show (again) the plan views of the surveys at the C dam and the D dam, respectively. CPTUs were carried out at the top of each dam, namely, CPTUs were performed at 15 points with 2-m intervals at the C dam and at 11 points with 5-m intervals at the D dam. In addition, Figs. 5.4 (a) and (b) show (again) the geological cross sections, estimated from the boring tests, of the C dam and the D dam, respectively. Although corrected cone resistance  $q_t$  is considered to reflect the soil strength with high accuracy, the  $N$ -value is generally used to design soil structures in Japan. Therefore, the conversion formulas for the  $N$ -value obtained from the CPT,  $N_c$ , proposed by Suzuki et al. (2003), was employed here. The procedure to calculate  $N_c$  is shown in Chapter 2. A  $\log N_c$  is taken as the logarithm of  $N_c$ , while Figs. 5.5 (a) and (b) show the depth distribution of the  $\log N_c$  at the C dam and the D dam, respectively.

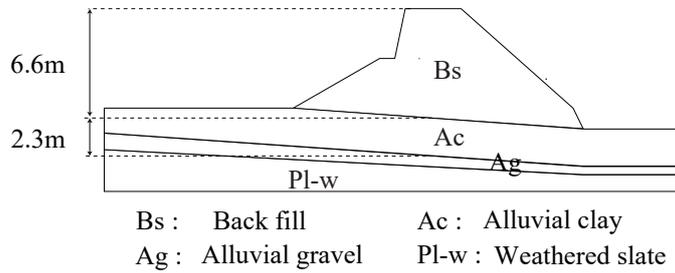


(a) C dam.



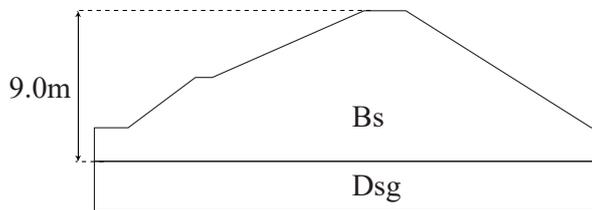
(b) D dam.

Fig. 5.3 Plan views (these figures already shown in Fig. 4.1)



Bs : Back fill      Ac : Alluvial clay  
 Ag : Alluvial gravel      Pl-w : Weathered slate

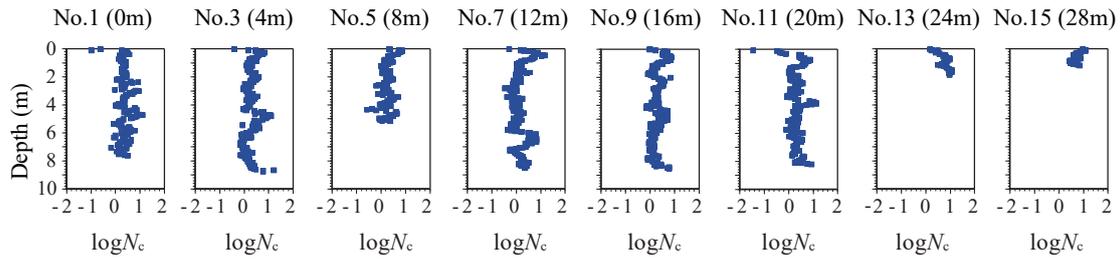
(a) C dam.



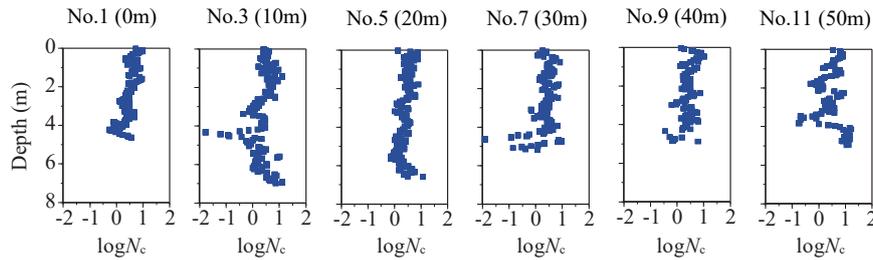
Bs : Back fill      Dsg : Diluvial sandy gravel

(b) D dam.

Fig. 5.4 Cross section of each dam and legend (these figures already shown in Fig. 4.2)



(a) C dam.



(b) D dam.

Fig. 5.5 Spatial distributions of  $\log N_c$

Table 5.1 Statistical models determined by MAICE

	Mean function	S.D.
C dam	$m = 0.321$	$\sigma = 0.371$
D dam	$m = 0.531 - 0.0490z$	$\sigma = 0.374$
Covariance function ( $i, j = 1, 2, \dots, M$ )		
C dam	$C = \sigma^2 \exp\left(- x_i - x_j  / 0.01 -  z_i - z_j  / 0.21\right)$	
D dam	$C = \sigma^2 \exp\left(- x_i - x_j  / 0.01 -  z_i - z_j  / 0.22\right)$	

$M$ : Number of measured points,  $x$ : Horizontal coordinate (m),  
 $z$ : Depth(m), S.D.: Standard deviation

### 5.3 Statistical modeling of soil strength

The optimum model of  $\log N_c$  is determined by minimizing the Akaike Information Criterion, and the model called MAICE (Minimum AIC Estimator). The procedure for the MAICE is presented in Chapter 3. Using the procedure for MAICE, the mean function,  $m$ , the standard deviation,  $\sigma$ , and the covariance function were determined for the C dam and the D dam, as shown in Table 5.1. According to this table, the covariance

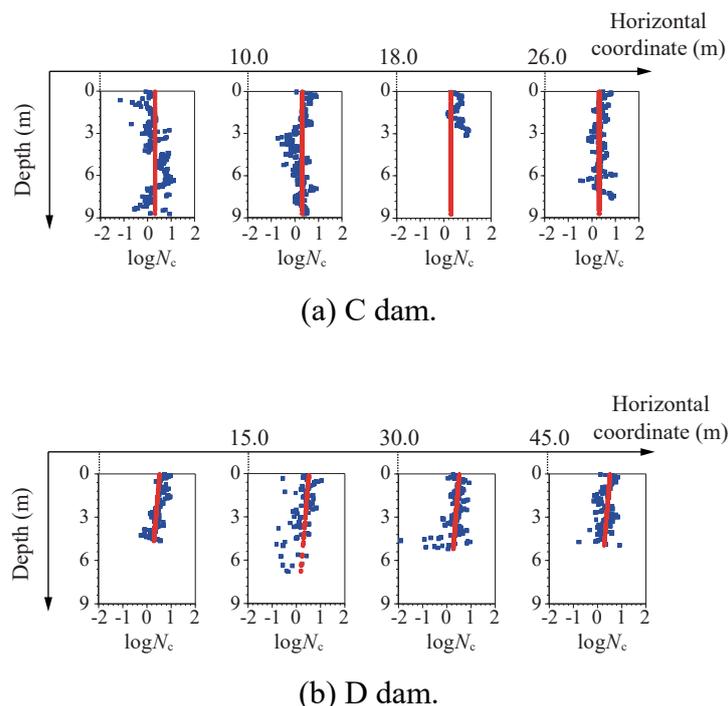


Fig. 5.6 Values measured by CPTUs and mean function of  $\log N_c$

functions for the C dam and the D dam showed that the correlation distance in the horizontal direction,  $l_x$ , equals 0.01 m (lower boundary value in the analysis), which means it is uncorrelated. Figs. 5.6 (a) and (b) show the mean function obtained for each of the depth distributions of  $\log N_c$  at the C dam and the D dam, respectively. In the C dam, a constant value was chosen for the mean function, while in the D dam, a model with a linear gradient in the depth direction was selected.

Since the correlation distance of  $\log N_c$  in the horizontal direction in the C dam and the D dam was determined to be uncorrelated, the procedure for the MAICE could not estimate an appropriate spatial structure. Therefore, since the semi-variogram can easily identify the spatial structure one-dimensionally, the spatial structure was evaluated by the semi-variogram. The procedure for this method is shown in Chapter 3.

As a feature of the CPT, although the continuous data in the depth direction can be obtained with high accuracy, the measured data have large variability in many cases. Therefore, due to the influence of the outliers included in the data, the random field cannot satisfy the assumption of the stationarity, and this makes it difficult to obtain the appropriate semi-variogram. In order to evaluate the influence caused by the outliers, part of the spike-like distribution of the data, shown in Fig. 5.1, was defined here as the outliers that cause the disturbance to the data. Based on the definition, it was assumed that the outliers were close to the maximum value and the minimum values of the  $\log N_c$ .

Moreover, the effect of eliminating these data on the value of the root mean square error (*RMSE*), calculated from the semi-variogram, and its approximate function will be discussed.

As a procedure for detecting the outliers, it is proposed that the threshold values be set as shown in Fig. 5.1. The data are classified into three groups, namely, the high group, the middle group, and the low group, respectively. However, when the data are classified into each of the groups, they lose their normality. Therefore, before computing the semi-variogram, the normal transformation, given by Eq. (3.25), is applied to the normalized random variable,  $f$ , of each group. The standardized random variable,  $Y$ , is defined by using the mean function,  $m$ , and the standard deviation,  $\sigma$ . Letting  $F$  represent the cumulative distribution function in an arbitrary section, with respect to the data classified into each group,  $f$  is converted into normal random variable  $Y$  by Eq. (3.25).

The semi-variogram of normal random variable  $Y$  is modeled by a simple exponential function given by Eq. (3.27). The goodness of fit of the approximate curve to the semi-variogram was evaluated by the *RMSE* expressed by the following equation:

$$RMSE = \sqrt{\frac{1}{N_r} \sum_{i=1}^{N_r} (\gamma_i - \hat{\gamma}_i)^2} \quad (5.1)$$

where  $N_r$  is the number of prediction targets,  $\hat{\gamma}_i$  is the semi-variogram obtained from the measured value, and  $\gamma_i$  is the predicted value of the semi-variogram based on the approximation function.

As the calculation procedure for the *RMSE*, first, an  $a$  percentile and a  $b$  percentile of the standard normal distribution with respect to the normalized random variable,  $f$ , are set as the threshold value between the high group and the middle group and that between the middle group and the low group, respectively. Here, the case of excluding  $f$  less than or equal to the  $a$  percentile and the case of excluding  $f$  greater than or equal to the  $b$  percentile is denoted as  $(a, b)$ . In the C dam, the comparison was made for a total of six cases, including (5, 95), (10, 90), (15, 85), (20, 80), and (25, 75), and all of the data. In addition, in the D dam, the comparison was made for a total of six cases, including (0.5, 99.5), (1, 99), (2.5, 97.5), (5, 95), and (10, 90), and all of the data. The results of the comparisons are summarized in Table 5.2. The reason why the thresholds of the C dam and the D dam are different is that the ranges in the minimum values were different in the process of examining the minimum value of the *RMSE*.

Table 5.2 Influence of removal of outliers on goodness to fit for model

(a) C dam.

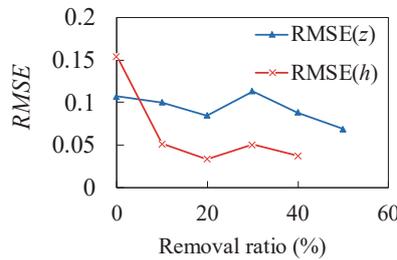
Removal ratio (%)	Case	$C_{0z}$	$C_{0x}$	$l_z$	$l_x$	$RMSE(z)$	$RMSE(h)$
0	All data	0.21	0.62	0.70	5.68	0.107	0.154
10	(5, 95)	0.16	0.55	0.46	4.72	0.100	0.052
20	(10, 90)	0.21	0.62	0.41	4.45	0.085	0.033
30	(15, 85)	0.05	0.80	0.26	4.44	0.113	0.050
40	(20, 80)	0.42	0.80	0.36	4.44	0.088	0.037
50	(25, 75)	0.53	*	0.38	*	0.069	*

\* : Can not be computed, Red color : Minimum value.

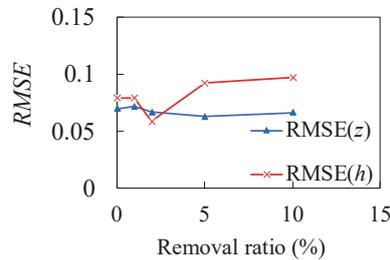
(b) D dam.

Removal ratio (%)	Case	$C_{0z}$	$C_{0x}$	$l_z$	$l_x$	$RMSE(z)$	$RMSE(h)$
0	All data	0.30	0.55	0.48	9.93	0.070	0.079
1	(0.5, 99.5)	0.33	0.53	0.47	9.59	0.072	0.079
2	(1, 99)	0.36	0.57	0.50	10.13	0.067	0.059
5	(2.5, 97.5)	0.40	0.44	0.39	6.57	0.063	0.092
10	(5, 95)	0.43	0.51	0.37	6.59	0.066	0.097
20	(10, 90)	*	*	*	*	*	*

\* : Can not be computed, Red color : Minimum value.



(a) C dam.



(b) D dam.

Fig. 5.7 Relationship between removal ratio of outliers and  $RMSE$

Assuming that the ratio of the data outside the thresholds is defined as the removal ratio of the outliers, three tendencies were confirmed as the range in the removal ratio

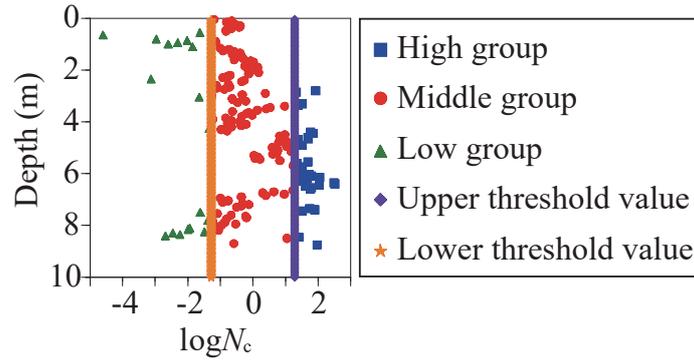


Fig. 5.8 Classification of in-situ data in case of (10, 90) at  $x = 2$  m

Table 5.3 Statistical models of  $\mathbf{Y}_H$ ,  $\mathbf{Y}_L$ ,  $\mathbf{Y}_M$

Covariance function ( $i, j=1, 2, \dots, M$ )	
$\mathbf{Y}_M$	$\mathbf{C} = N_e \exp(- x_i - x_j /4.45 -  z_i - z_j /0.41)$ ( $i \neq j$ )
	$N_e = 0.298$ ( $ x_i - x_j  \neq 0,  z_i - z_j  \neq 0$ )
	$N_e = 0.377$ ( $ x_i - x_j  \neq 0,  z_i - z_j  = 0$ )
	$N_e = 0.790$ ( $ x_i - x_j  = 0,  z_i - z_j  \neq 0$ )
$\mathbf{Y}_L, \mathbf{Y}_H$	$\mathbf{C} = 1^2$ ( $i = j$ )
	$\mathbf{C} = 0$ ( $i \neq j$ )
	$\mathbf{C} = 1^2$ ( $i = j$ )

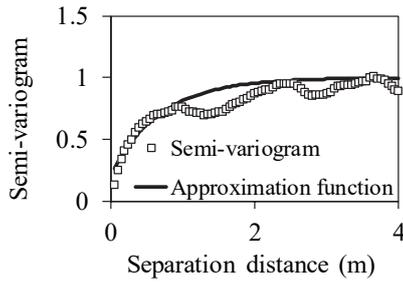
$M$ : Number of measured points,  $N_e$ : Nugget parameter.

became wider. First, there is a tendency for the nugget effect to become larger. Second, there is a tendency for the correlation distance to become shorter. Finally, there is a tendency for the *RMSE* to become smaller. However, the optimum case of the removal ratio is seen in the horizontal direction according to the *RMSE*.

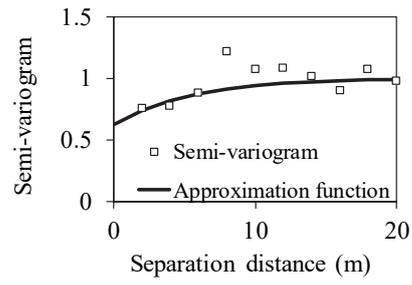
Figs. 5.7 (a) and (b) show the changes in the *RMSE* in the six cases for the C dam and the D dam, respectively. As a result, the case where the *RMSE* in the horizontal direction becomes the minimum, was the case of (10, 90) in the C dam and the case of (1, 99) in the D dam. Since the amount of information is abundant in the depth direction, the change in the semi-variogram in the depth direction is less sensitive than the one in the horizontal direction, to the removal ratio. Therefore, the *RMSE* in the horizontal direction is regarded as top priority, and the case where the *RMSE* becomes the minimum, was considered as the optimum removal ratio. Fig. 5.8 shows the figure in which the measured data at  $x = 2$  m of the C dam is classified into the three groups in the case of (10, 90).

Figs. 5.9 and 5.10 show the changes in the goodness of fit between the

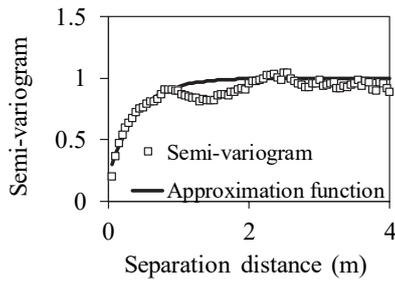
semi-variogram and the approximate function by excluding the outliers. From these results, it is seen that the fitness of the approximate function to the semi-variogram was improved visually. First, Fig. 5.9 shows the semi-variogram in the case of using all the data for the C dam and the case of a removal ratio of 20%, namely, (10, 90). In particular, when comparing Figs. 5.9 (b) and (d), it can be confirmed that the goodness of fit between the semi-variogram and the approximate function was greatly improved. Table 5.3 shows the statistical models determined by the semi-variogram for the high group,  $Y_H$ , the middle group,  $Y_M$ , and the low group,  $Y_L$ , when the removal ratio of the data is (10, 90). In the statistical model obtained for  $Y_M$ , the correlation distance in the horizontal direction was estimated to be about 10 times the correlation distance in the vertical direction. These results were confirmed to show the same trend as the results seen by Nishimura (2007) and Phoon and Kulhawy (1999). On the other hand, because the density of the data included in  $Y_H$  and the  $Y_L$  was insufficient, the spatial structure of the data was estimated to be uncorrelated. Here, in order to determine the approximate functions, the data in which the separation distance of the data are less than or equal to 0.9 m in the vertical direction for both the C dam and the D dam, were used. This procedure is the same as that stated in section 4.3. On the other hand, in the horizontal direction, the least squares method was applied to the data, which have the separation distance of the data up to 6 m in the C dam and the data up to 15 m in the D dam. In other words, it was assumed that the interval between the two sets of data was one span, and the first three spans were used for determining the approximate function. This is because the amount of data in these spans is sufficient and reliable. This assumption is also the same as that stated in section 4.3. By treating the data in this manner, the correlation distance obtained by the semi-variogram tends to be longer than that obtained by the MAICE. In addition, the range to calculate the *RMSE* in the depth direction was up to 2 m for both the C dam and the D dam. Moreover, the range to calculate the *RMSE* in the horizontal direction was up to 10 m for the C dam and up to 20 m for the D dam, respectively. This is because, since the accuracy of the semi-variogram decreases as the distance between the data increases, the *RMSE* was examined in the range of about 40% of the length of each dam.



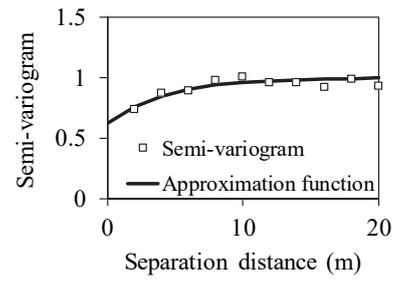
(a) Depth direction (all data).



(b) Horizontal direction (all data).

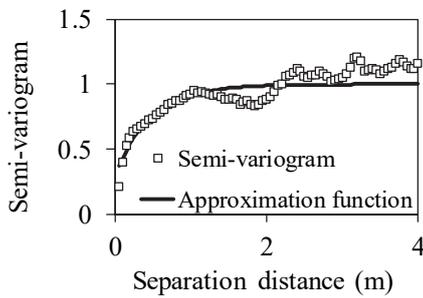


(c) Depth direction (10, 90).

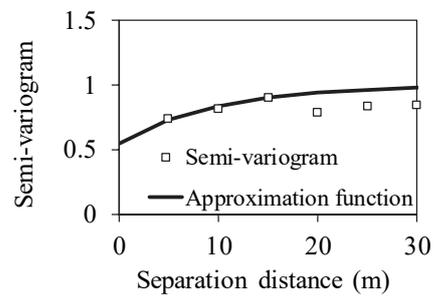


(d) Horizontal direction (10, 90).

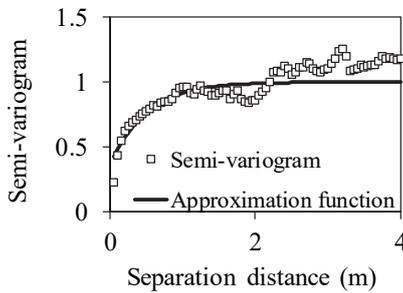
Fig. 5.9 Relationship between semi-variogram and removal of outliers at C dam



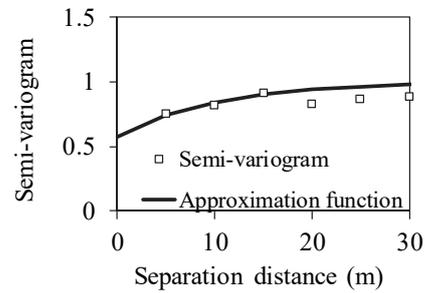
(a) Depth direction (all data).



(b) Horizontal direction (all data).



(c) Depth direction (1, 99).



(d) Horizontal direction (1, 99).

Fig. 5.10 Relationship between semi-variogram and removal of outliers at D dam

## 5.4 Interpolation of soil strength using conditional simulation

In section 5.3, it was confirmed that the removal of the outliers was effective for appropriately obtaining the parameters of the random field. On the other hand, since it is also important to consider the data classified as the outliers when determining the stability of an earth-fill dam, it is necessary to incorporate the outliers into the evaluation of the spatial distribution of the soil strength. Therefore, in this section and the next section, in order to evaluate the spatial distribution of the soil strength considering the outliers, a method of re-composing the results of the conditional simulation obtained for each of the three groups is proposed. However, since the physical phenomena of the high group and the low group are clearly non-stationarity, the random field theory is not applied to these groups. On the other hand, by excluding the outliers, a method is proposed to secure the stationarity of the middle group. In the following sections 5.4, 5.5, and 5.6, the C dam will be analyzed in particular; it is thought to be greatly influenced by the outliers.

Kriging, which is one kind of geostatistical method, is often used to spatially interpolate the soil properties obtained as the point estimation. In kriging, the variance of the coefficient of interpolation is not 0, excluding the sampled points, and the interpolated values of the soil properties have variability. Therefore, in order to take into account the variability, a method for generating a large number of realizations (simulated values) and executing a lot of simulations is called a conditional simulation. In this study, the conditional simulation is employed through the use of the analysis code SGSIM (Deutsch and Journel, 1992).

In the conditional simulation, it is assumed that the input data follow a standard normal distribution. Therefore, the normal transformed variables, namely, the high group,  $Y_H$ , the middle group,  $Y_M$ , and the low group,  $Y_L$  is applied to the conditional simulation. As a result, a group of simulated results of the normal transformation,  $Y'_H$ ,  $Y'_L$ , and  $Y'_M$ , was obtained by the conditional simulation. Also, the normal inverse transformation, shown in Eq. (5.2), was performed for them, and a group of simulated results for the normalized random variable,  $f'_H$ ,  $f'_L$ , and  $f'_M$ , was obtained.

$$f' = F^{-1}(\Phi(Y')) \quad (5.2)$$

Here,  $f'$  is a representative value of the simulated results of the normalized random variable,  $F$  is the cumulative distribution function in an arbitrary section for each of the

three groups, and  $Y'$  is a representative value of the simulated results of the normal transformation. Furthermore,  $\mathbf{f}_H$ ,  $\mathbf{f}_L$ , and  $\mathbf{f}_M$  were converted into the  $N$ -value calculated from the CPTUs,  $N_c$ , by using Eq. (5.3), and the simulated results for the  $N_c$  of the conditional simulation were obtained for each group, namely, the high group,  $\mathbf{R}_H$ , the low group,  $\mathbf{R}_L$ , and the middle group,  $\mathbf{R}_M$ .

$$N_c(x, z) = 10^{(m + f'(x, z) \cdot \sigma)} \quad (5.3)$$

where  $m = 0.321$  and  $\sigma = 0.371$  are taken from Table 5.1.

The output points of all the simulation results are set in a lattice shape, and the lattice consists of 5249 points in total. In the lattice, 29 points are set to have 1-m intervals in the horizontal direction and 181 points are set to have 0.05-m intervals in the depth direction. Since 418 of these points are near the bottom surface and are assumed to be the base layer, they are treated as constant values. Thus, the total number of evaluated points in the conditional simulation,  $N_T$ , is 4831.

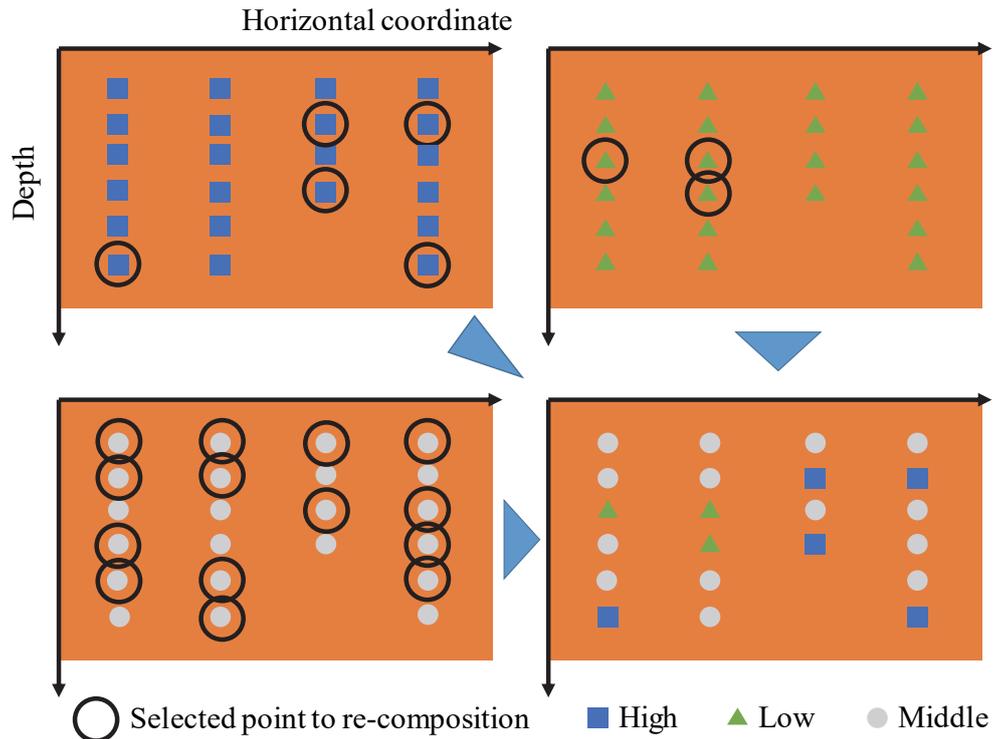


Fig. 5.11 Concept of re-composition of simulated results of three groups

### 5.5 Re-composition of simulated values of three groups

In order to evaluate the strength distribution of a ground with large spatial variability, a method is proposed to re-compose the simulated results of the three groups. Fig. 5.11 shows the concept of the re-composition of the results of the conditional simulation for each of the groups. In addition, Fig. 5.12 shows the flowchart for conducting the re-composition; it shows the procedure for integrating the values from each point of the three results of the conditional simulation. The unique feature of this study is that focus was placed on the spike-like distribution of the data of the cone tip resistance, which was continuously measured by CPTUs in the depth direction. And the locations where outliers are likely to appear were evaluated. To investigate the locations of the outliers, the difference value between the simulated value,  $R_M$ , calculated from the data in the middle group, and the threshold value,  $T$ , used for the classification of the data, was defined in the following equation:  $D = |R_M - T|$ . The reason for using difference value  $D$  and a detailed definition of  $D$  are given below.

First, as shown in Fig. 5.1, when there is a relatively high strength (or low

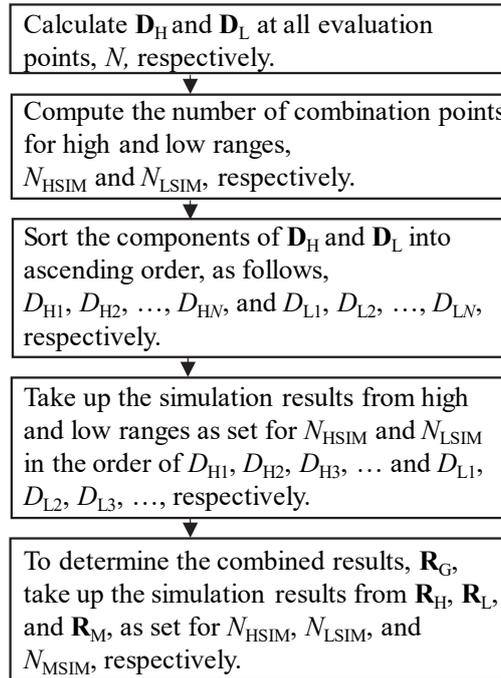


Fig. 5.12 Flowchart of re-composition of simulation results

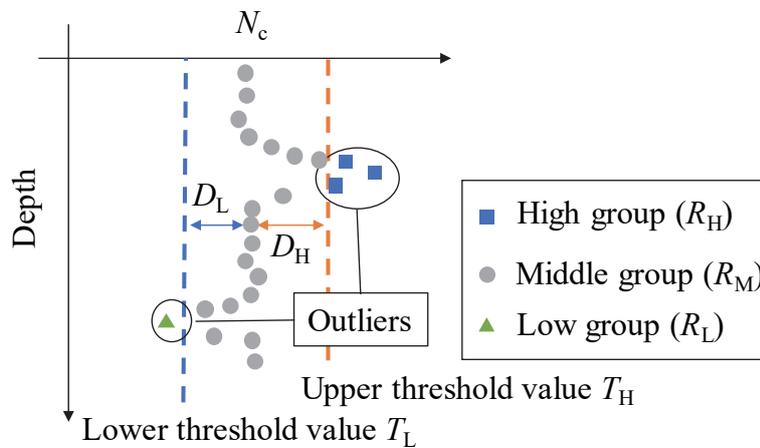


Fig. 5.13 Definitions of  $D_H$  and  $D_L$

strength) part inside a ground, the tip resistance sharply increases (or decreases) with the depth of the penetration, and the value at the spike point becomes the maximum value (or the minimum value). In addition, as the penetration depth becomes even deeper, the value sharply decreases (or increases) again. For example, in Fig. 5.5, spike-like distributions of the  $\log N_c$  were detected in the measured data of the C dam and the D dam. Also, in Fig. 5.8, spike-like distributions were confirmed in the classification diagram of the data for the C dam at  $x = 2$  m. Therefore, it is obvious that the value of a

Table 5.4 Variables for re-composition of simulation results

$N_T$	4831	(Number of all evaluation points)
$N_{OB}$	1955	(Number of all in-situ data) $N_{OB} = N_{HOB} + N_{MOB} + N_{LOB}$
$N_{SIM}$	2876	$N_{SIM} = N - N_{OB}$ $= N_{HSIM} + N_{MSIM} + N_{LSIM}$
$N_{HOB}$	224	(In-situ data of high group)
$N_{MOB}$	1619	(In-situ data of middle group)
$N_{LOB}$	112	(In-situ data of low group)
$P_H$	11.5%	$P_H = N_{HOB}/N_{OB}$
$P_L$	5.7%	$P_L = N_{LOB}/N_{OB}$
$N_{HSIM}$	329	$N_{HSIM} = P_H N - N_{HOB}$
$N_{MSIM}$	2382	$N_{MSIM} = N_{SIM} - (N_{HSIM} + N_{LSIM})$
$N_{LSIM}$	165	$N_{LSIM} = P_L N - N_{LOB}$

simulated result in the middle group around an outlier is close to the threshold value due to the tendency shown in Fig. 5.1. In this manner, when the difference value,  $D$ , calculated from the simulated value of the middle group and the threshold value, is small, the locations with a high possibility for the appearance of outliers, can be judged.

Based on the assumptions stated above, the locations with a high possibility for outliers are evaluated from the simulated results of the middle group,  $\mathbf{R}_M$ , and the threshold values. In Fig. 5.13, the definitions of difference values  $D_H$  and  $D_L$  are given as follows:  $D_H(x, z) = T_H - R_M(x, z)$  and  $D_L(x, z) = R_M(x, z) - T_L$ , respectively. In the calculation, the simulated value,  $\mathbf{R}_M$ , calculated from the conditional simulation using the data in the middle group, and the threshold values used to classify the data into three groups are employed.  $T_H$  is the threshold value between the high group and the middle group, and  $T_L$  is the threshold value between the middle group and the low group. Small values for  $D_H$  and  $D_L$  indicate the locations where outliers of high strength or low strength are likely to appear.

Next, Table 5.4 shows the number of measured data classified in each of the groups and the ratio of the number of classified data against the number of all the data. The ratio is defined as the classification ratio here. The number of evaluated points in the re-composition is defined as  $N_{SIM}$ . Also, in the re-composition of the simulated results of the three groups, the number of points which are selected from each group, correspond to the classification ratio. In other words, the assembly ratio of the high group,  $P_H$ , and that of the low group,  $P_L$ , are defined as  $P_H = N_{HOB}/N_{OB}$  and

$P_L = N_{LOB}/N_{OB}$ . Based on the above assumptions, the re-composition from the three groups is carried out by selecting the simulated results using the assembly ratio of  $P_H$  and  $P_L$  from the simulated results of each group. The process to determine a point to be re-composed from  $\mathbf{R}_H$  and  $\mathbf{R}_L$  will be described in detail in accordance with the flowchart shown in Fig. 5.12, given below.

First, the total number of evaluated points in the conditional simulation,  $N_T$ , where the soil properties are interpolated in the conditional simulation, is defined as  $N_T = N_{OB} + N_{SIM}$ , where  $N_{OB}$  is the total number of measured data. Also, as shown in Table 5.4, since the re-composed results consist of the simulated results of the three groups,  $N_{HSIM}$ ,  $N_{MSIM}$ , and  $N_{LSIM}$  mean the selected number of points from each group, respectively, for the re-composition. The definitions of  $N_{HSIM}$ ,  $N_{MSIM}$ , and  $N_{LSIM}$  are as follows:

$$N_{HSIM} = P_H \cdot N_T - N_{HOB} \quad (5.4)$$

$$N_{LSIM} = P_L \cdot N_T - N_{LOB} \quad (5.5)$$

$$N_{MSIM} = N_{SIM} - (N_{HSIM} + N_{LSIM}) \quad (5.6)$$

where  $N_{HOB}$  is the number of measured data included in the high group and  $N_{LOB}$  is the number of measured data included in the low group.

Next, in order to evaluate the locations where the outliers of high strength or low strength are likely to appear, the components of  $D_H$  and  $D_L$  are sorted into ascending order, as follows,  $D_{H1}, D_{H2}, \dots, D_{HN}$ , and  $D_{L1}, D_{L2}, \dots, D_{LN}$ , respectively. Then, the simulated results from the high group and the low group, as sets for  $N_{HSIM}$  and  $N_{LSIM}$ , are taken in the order of  $D_{H1}, D_{H2}, D_{H3}, \dots$  and  $D_{L1}, D_{L2}, D_{L3}, \dots$ , respectively. The values of  $N_{MSIM}$  are selected from the simulated results of the middle group. As a result, the simulation values of  $N_{SIM} = N_{HSIM} + N_{MSIM} + N_{LSIM} = 2876$  points are determined for the entire region of the analysis, and the simulated results of the re-composition of the three groups,  $\mathbf{R}_G$ , are obtained. Furthermore, in order to obtain a large quantity of realized values from the random field, the procedure given in Fig. 5.12 is repeated as many times as necessary for the conditional simulation.

## 5.6 Validation of simulated values using proposed method

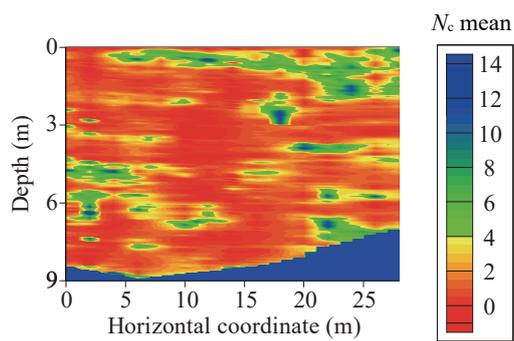
### 5.6.1 Estimated statistical model for C dam

The spatial distribution of the soil strength inside the C dam is modeled using the logarithm of the  $N$ -value calculated from the CPTUs,  $\log N_c$ . First, the mean function,  $m$ , and the standard deviation,  $\sigma$ , were determined by following the procedure of MAICE as presented in Chapter 3. As a result, the estimated models were obtained as shown in Table 5.1.

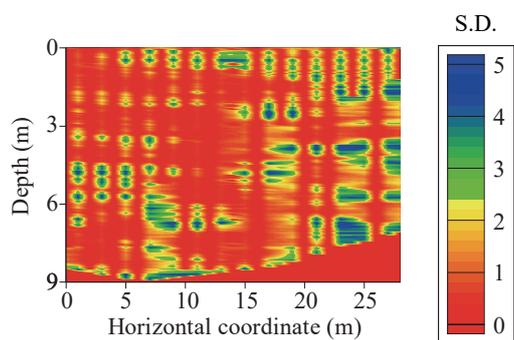
As shown in section 5.3, in the C dam, when data is removed by setting the values of the 10 percentile and the 90 percentile of the standard normal distribution against the normalized random variable,  $f$ , as the threshold values, the  $RMSE$  becomes minimum. Based on the results, the value of the 90 percentile of the  $f$  is set as the upper threshold value and the data of the 90 percentile or more are classified as belonging to the high group. Similarly, the value of the 10 percentile is set as the lower threshold value and the data of 10 percentile or less are classified as belonging to the low group. The other data are classified into the middle group. Table 5.3 shows the statistical model determined by the semi-variogram for all the data classified into the three groups. The spatial distributions of the soil strength inside the C dam using these statistical models are shown in sub-section 5.6.2.

### 5.6.2 Evaluation of spatial distribution of soil strength

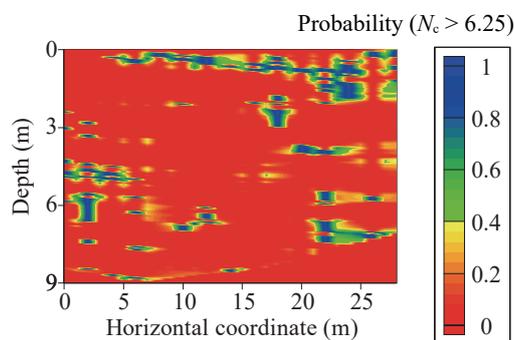
By the method proposed in section 5.5, the simulated results of the re-composition of the three groups,  $\mathbf{R}_G$ , is obtained. However, in order to converge the statistics of the random field obtained from the random field of  $\mathbf{R}_G$ , the conditional simulation was repeated 2000 times. As a result, the spatial distribution of the statistics related to the soil strength were given as shown in Figs. 5.14 (a) to (d).



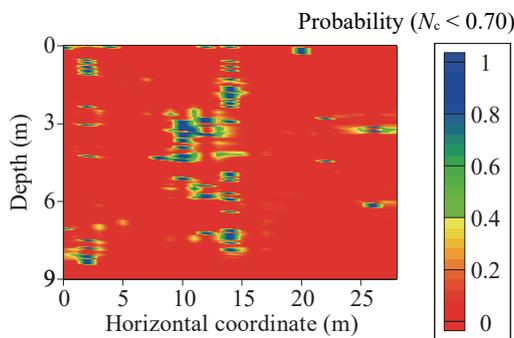
(a) Mean values for  $N_c$ .



(b) Standard deviation.



(c) Probability of  $N_c > 6.25$ .



(d) Probability of  $N_c < 0.70$

Fig. 5.14 Spatial distribution of statistics by proposed method

Fig. 5.14 (a) shows the spatial distribution of the mean value of the  $N_c$ -value calculated from the CPTUs,  $N_c$ . According to this figure, the mean value for the  $N_c$  in the dominant space is  $N_c < 4$ , while there are particularly weak areas of  $N_c$  equal to or less than 2 around  $x = 10 - 15$  m and  $z = 3 - 9$  m. In addition, Fig. 5.14 (b) shows the spatial distribution of the standard deviation. It can be confirmed that the standard deviation is almost 0 at any point where measured data exist or where  $N_c$  is small. Figs. 5.14 (c) and (d) show the spatial distribution of the probability that  $N_c$  is greater than the upper threshold value of 6.25 and the spatial distribution of the probability that  $N_c$  is lower than the lower threshold value of 0.70, respectively. Fig. 5.14 (c) visualizes the degree of the spread of the relatively high strength areas inside the dam as the probability of  $N_c > 6.25$ . The probability that the high strength parts continuously spread around  $x = 5 - 25$  m and  $z = 1 - 2$  m in the C dam were confirmed to be high. Moreover, Fig. 5.14 (d) shows that the probability that relatively weak parts exist around  $x = 10 - 15$  m and  $z = 3 - 4.5$  m are high.

### 5.6.3 Validation of interpolated results by proposed method

In this study, a method is proposed to model the spatial distribution of the soil strength inside earth-fill dams composed of materials with different particle size distributions by appropriately re-composing the three groups of results from a conditional simulation. In order to validate the random field of the strength distribution inside the C dam, estimated by the proposed method, some of the data were intentionally removed from the measured data, and the remaining data were applied to the simulation. In addition, the statistics of the random field estimated by the proposed method and the removed data were compared at the same location to verify their correspondence. The following two cases were verified. One is a case in which the measured data at  $x = 6$  m were intentionally removed, and the other is a case in which the measured data at  $x = 20$  m were intentionally removed.

In order to evaluate the strength distribution inside the C dam, the simulation was repeated 2000 times, and the mean value and the standard deviation of the random field at each of the points in the analysis region were obtained. To evaluate the accuracy of the statistics of the random field, the standardized residuals,  $w$ , with mean = 0 and standard deviation = 1, are defined by the following equation:

$$w(x, z) = (b(x, z) - E(x, z)) / SD(x, z) \quad (5.7)$$

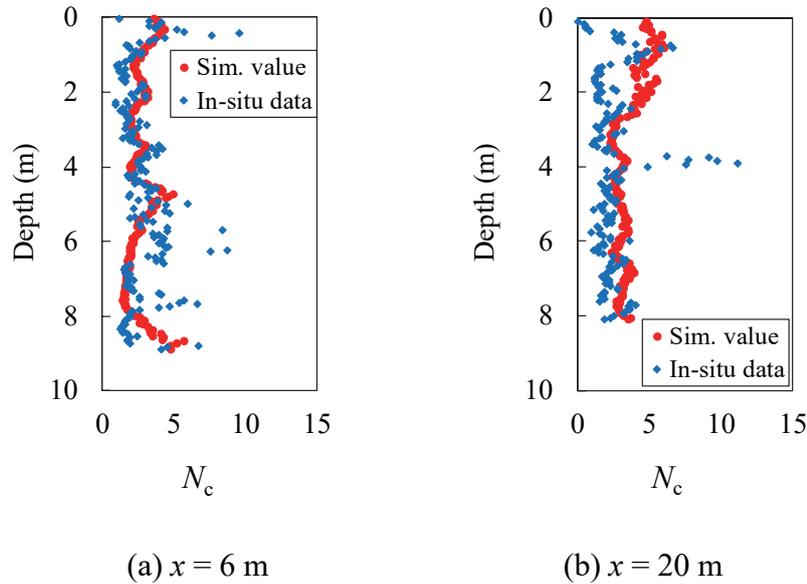


Fig. 5.15 Comparison of soil strength distribution between expected values and in-situ data

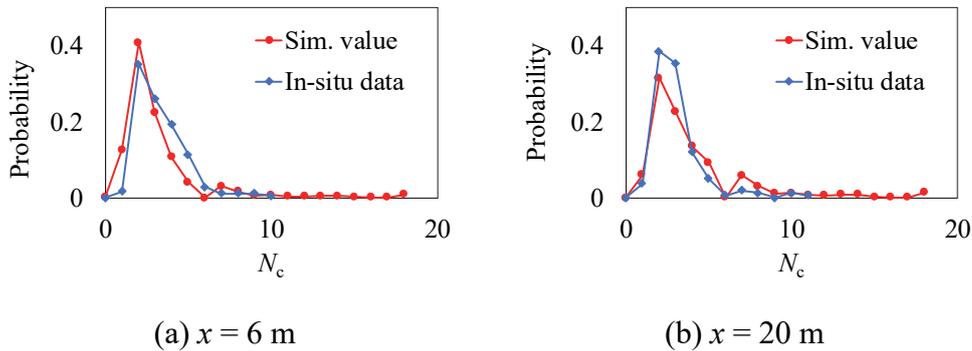


Fig. 5.16 Comparison of probability density function between simulated values and in-situ data

where  $b(x, z)$  is measured data,  $E(x, z)$  is the mean value calculated from the realizations of the random field of  $\mathbf{R}_G$ , and  $SD(x, z)$  is the standard deviation calculated from the realizations of the random field of  $\mathbf{R}_G$ .

Fig. 5.15 (a) and (b) show the mean value of the random field and the distribution of the measured data at  $x = 6$  m and at  $x = 20$  m, respectively. In Fig. 5.15 (a), the mean values roughly follow the measured values, except for around  $z = 6, 8$  m. In addition, Fig. 5.15 (b) shows that the mean values roughly follow the measured values, except for around  $z = 2, 4$  m. It seems that the difference is caused by the great

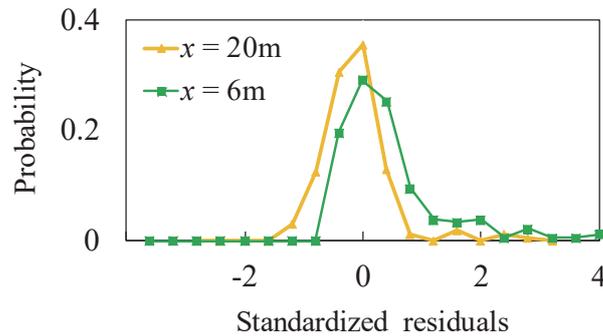


Fig. 5.17 Probability density function of standardized residuals,  $w$

variability in the soil strength of the measured values. The mean values simulate the trend of the measured values well, but the mean values occasionally yield mis-estimations of the measured values.

Fig. 5.16 (a) and (b) show the probability density function of the  $N$ -value calculated from the CPTUs,  $N_c$ , at  $x = 6$  m and at  $x = 20$  m, respectively, and that of the measured data, and the shape of the probability density function of the measured value and the simulated value was compared. As shown in these figures, the shape of the probability density function of the measured value and that of the simulation value corresponded roughly for both  $x = 6$  m and  $x = 20$  m. In addition, the density of these distributions were concentrated in the low strength. Therefore, it was confirmed that the proposed method can reasonably simulate the removed data.

Fig. 5.17 shows the probability density function of the standardized residuals,  $w$ , calculated at  $x = 6$  m and  $x = 20$  m. Since the standardized residuals are concentrated in the part close to 0, the simulated values by the proposed method are considered to be reasonable.

## 5.7 Conclusion

In this chapter, the spatial distribution of the soil strength of earth-fill dams composed of materials with different particle size distributions, was evaluated by applying the geostatistical method to the results obtained from CPTUs. The concluding remarks are given below.

- 1) Since CPTs can measure the soil strength precisely and with high spatial resolution, CPTUs were conducted at a dams composed of materials with different particle size

distributions. As a result, the measured data obtained from the CPTUs contained the local areas of high strength as well as those of low strength.

- 2) The measured values were classified into three groups, namely, high, middle, and low, by the threshold values between the high and low groups. The threshold values were determined so that the *RMSE* would be minimized. Since the influence of outliers on the random field modeling of the  $Y_M$  group was reduced by their removal, the correlation distances could be properly estimated. As a result, the horizontal value was about 10 times that of the vertical one.
- 3) It is seen in the CPTUs that the spatial distributions of the weak areas and the strong areas originated from the amount of gravel mixed into the soil and affect the soil strength. In other words, the weak areas contain a smaller amount of gravel, while the strong areas contain a larger amount of gravel. The novelty of the proposed method is in the re-composition of the simulation values of the three groups, namely, high,  $Y_H$ , middle,  $Y_M$ , and low,  $Y_L$ . The  $Y_H$  and  $Y_L$  groups model the outliers of high strength and low strength, respectively. In the proposed method, the rate of outliers is determined from the measured values. The simulated values for each of the three groups are re-composed so as to follow the determined rate, and the locations of the outliers are determined based on the simulated values of the middle group,  $Y_M$ .
- 4) As a result of a comparison between the simulated values and the measured values, the distribution of these values at the same place roughly corresponded, and the shapes of the probability density functions were also similar. Thus, it has been verified that the proposed method can be used to reasonably simulate the spatial variability of the soil strength considering the outliers.
- 5) One of the problems for the practical application of the proposed method is that it requires a large amount of data in the horizontal direction in order to identify the parameters of the spatial structure. As a practical option, it is necessary to conduct site investigations which include a section for short-interval testing as part of the whole survey line. It is also important to develop a database of the correlation distances of the soil properties.

## Chapter 6

# **Evaluation of liquefaction probability of earth-fill dam over next 50 years using geostatistical method based on CPT**

### **6.1 Introduction**

Since the CPT is a highly efficient method for obtaining data at short intervals, the test data from CPTs enable the mapping of the spatial distribution of the liquefaction probability considering the spatial variability of the soil parameters. Many previous works have dealt with the liquefaction probability through evaluations using CPTs (e.g., Dawson and Baise, 2005; Vivek and Raychowdhury, 2014). Although the seismic hazard should be considered when properly evaluating the liquefaction risk, the above-mentioned studies did not deal with the seismic hazard, and the magnitude of an earthquake was assumed to be the deterministic parameter in the above articles. The present study, on the other hand, considers the frequency of earthquakes by using the seismic hazard. Since the seismic hazard of many earth-fill dams in Japan varies, depending on the location, the regional characteristics of the seismic hazard should be introduced when evaluating the liquefaction probability for each dam. For these reasons, the CPT and the seismic hazard are both utilized to evaluate the liquefaction probability of earth-fill dams.

Although the CPT is often preferable for identifying the liquefaction probability, the CPT is still not widely used in Japan. Therefore, focus will be placed here on the SPT  $N$ -value and the fines content,  $F_c$ , because the majority of the accumulated site characterization data in Japan consists of SPT  $N$ -values and  $F_c$ . In the design guidelines for earth-fill dams for irrigation (Ministry of Agriculture, Forestry, and Fisheries of Japan, 2015), the liquefaction resistance factor  $F_L$  method, which was originally based on the study of Iwasaki et al. (1984), is used. When applying the method properly, the liquefaction resistance of soils is evaluated based on the SPT  $N$ -value and  $F_c$  obtained from laboratory tests. In this paper, however, the liquefaction resistance is calculated through the modification of Iwasaki's method based on the CPT. Therefore, the measured values from the CPT are converted into SPT  $N$ -values and  $F_c$  using the conversion formulas presented by Suzuki et al. (2003). However, the formulas do not take conversion errors into account, and this may have a significant influence on the liquefaction probability. Conversion formulas, including the quantity of the

conversion errors, are proposed to estimate the  $N$ -value and  $F_c$  from the CPT later in this paper.

The soil parameters inside soil structures, such as river dikes and earth-fill dams, generally have spatial variability, which has a great influence on the probability of failure. Since the values taken from CPTs are point estimation values, the measured values need to be interpolated in order to continuously determine the spatial variability of the soil parameters. There are many works which deal with the spatial variability of soil parameters interpolated by geostatistical methods (e.g., Lenz and Baise, 2007; Chen et al., 2015). Since the soil parameters are spatially correlated, the correlation must be evaluated in order to present the spatial variability. The degree of the spatial correlation of the soil parameters is modeled by the correlation distance. The values interpolated through the use of geostatistical methods are largely influenced by the correlation distance of the soil parameters. Thus, the distance of several soil parameters was examined in horizontal and vertical directions, respectively, in past literature (e.g., Phoon and Kulhawy, 1999; Lloret-Cabot et al., 2014; Nishimura et al., 2016). Phoon and Kulhawy (1999) showed the distance of corrected CPT tip resistance  $q_t$  and undrained shear strength  $s_u$  from vane shear tests (VSTs). Lloret-Cabot et al. (2014) examined the distance of CPT tip resistance  $q_c$  utilizing the correlation function and geostatistical methods. Nishimura et al. (2016) presented the distance of the  $N$ -value from SWS by using the maximum likelihood method and geostatistical methods. Although fines content  $F_c$  was taken into account for the calculation of the liquefaction resistance, which was based on SPTs (e.g., Iwasaki et al. 1984; Cetin et al., 2004), the correlation distance of  $F_c$  was not sufficiently examined. Thus, the correlation distance of  $F_c$  inside an earth-fill dam is examined accurately in the horizontal and depth directions, respectively, based on the  $F_c$  data derived from CPTs.

A unique feature of the proposed work is its potential capability to take into account the accumulated site characterization data, such as the SPT  $N$ -value and  $F_c$ , adjacent to the studied sites. An illustrative example assessing the liquefaction probability of an earth-fill dam in Japan is presented to demonstrate the capability of the proposed method. This chapter is composed from eight sections. Section 6.2 provides information on the studied dam. Section 6.3 presents the model identification method of the soil parameters. Section 6.4 shows the spatially interpolated results of the soil parameters by a geostatistical method. Section 6.5 presents the liquefaction resistance factor  $F_L$  method. Section 6.6 describes the seismic hazard analysis of the dam. Section 6.7 shows the results of the seismic hazard analysis. The conclusion and a summary are given in Section 6.8.

## 6.2 Site investigation

An earth-fill dam in Hiroshima Prefecture, Japan was analyzed. The C dam is one of the dams for which there is a fear of liquefaction damage should a Nankai Trough earthquake occur. In Japan, to design countermeasures against the liquefaction of earth-fill dams, the  $N$ -value obtained from SPTs is generally used. In the Japanese design code for earth-fill dams for irrigation, the liquefaction resistance factor  $F_L$  method, originally based on the study of Iwasaki et al. (1984), is used. In Iwasaki et al.'s method, the  $N$ -value and fines content  $F_c$  are utilized to calculate  $F_L$ . Therefore, the measured values from the CPTs on the dam are firstly converted into the soil behavior type of index,  $I_c$ , proposed by Robertson and Fear (1995). Then,  $I_c$  is used to obtain the  $N$ -value and  $F_c$  by utilizing the conversion formulas which were originally proposed by Suzuki et al. (2003). Since the procedure for obtaining the  $N$ -value and  $F_c$  from the results of CPTs was presented in chapter 2, no explanation of the conversion formulas will be given here.

The CPT is employed to obtain several soil parameters, such as corrected cone resistance  $q_t$ , measured sleeve friction  $f_s$ , and pore water pressure  $u$  at short intervals. CPTs were conducted at 15 points at the top of the dam at intervals of 2 m along the crest, as shown in Fig. 6.1. The geological cross section at the studied site is given in Fig. 6.2. The height of the dam is 6.6 m, and the soil profile is classified into four layers, namely, backfill sand (Bs), alluvial clay (Ac), alluvial gravel (Ag), and weathered slate (Pl-w). The strength of the soil is weak around the depth of 9 m from the top of the dam in the Bs and Ac layers.

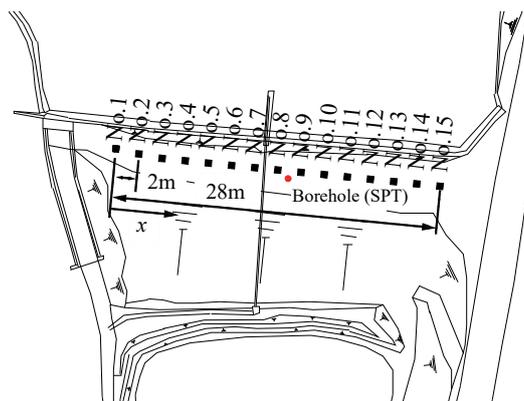
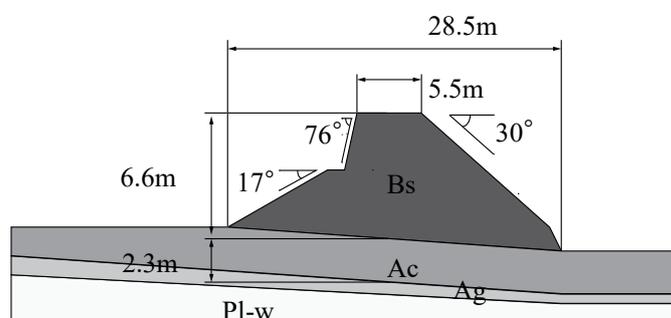


Fig. 6.1 Plan view of C dam and testing points of CPTs and SPT



Bs : Backfill sand      Ag : Alluvial gravel  
 Ac : Alluvial clay      Pl-w : Weathered slate

Fig. 6.2 Geological cross section of C dam

In addition, an SPT was performed at a boring point near the location of the CPTs, as shown Fig. 6.1. In Fig. 6.3, the distribution of the  $N$ -values obtained from the SPT is given. The  $N$ -values were examined at 1-m intervals in the depth direction. Fig. 6.4 shows the geological columnar section derived from laboratory tests at the same boring point as that of the SPT. Based on the laboratory tests, the fines content  $F_c$  of the backfill sand (Bs) layer is estimated as 48%.

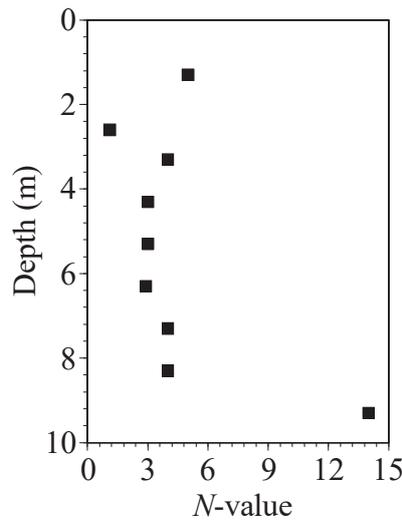


Fig. 6.3 Distribution of  $N$ -values from SPT in depth direction

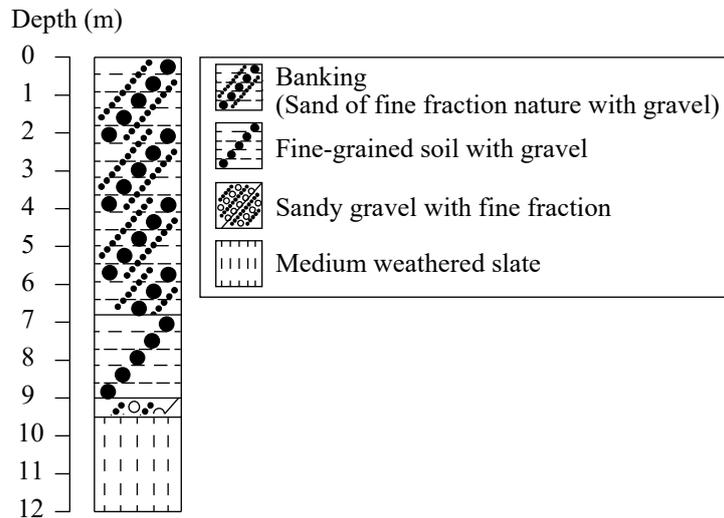


Fig. 6.4 Geological columnar section at boring point

### 6.3 Statistical model identification of studied dam

Statistical models for soil parameters are useful for visualizing the spatial variability of the soil parameters. This is because the soil parameters for the points where test results do not exist can be estimated using the statistical models. The statistical models of the CPT results are determined based on a geostatistical method (Journel and Huijbregts, 1978) and the random field theory. Although the soil parameters utilized to calculate the liquefaction probability have spatial variability, the variability can be introduced into the

Table 6.1 Statistical models of  $\log N_c$  and  $\log F_{clc}$  estimated by MAICE

	Mean function	Covariance function ( $i, j = 1, 2, \dots, M$ )
$\log N_c$	$m = 0.370 - 0.012x - 0.030z$ $+ 0.001x^2 + 0.005z^2 - 0.001xz$	$C = [C_{ij}] = 0.246^2 \exp(- x_i - x_j /0.01 -  z_i - z_j /0.28)$ ( $i \neq j$ ) $C = 0.246^2$ ( $i = j$ )
$\log F_{clc}$	$m = 0.874 + 0.021x + 0.217z$ $- 0.001x^2 - 0.019z^2 - 0.0001xz$	$C = [C_{ij}] = 0.123^2 \cdot 0.688 \cdot \exp(- x_i - x_j /2.28 -  z_i - z_j /1.10)$ ( $i \neq j$ ) $C = 0.123^2$ ( $i = j$ )

$m$ : Mean function,  $x$ : Horizontal coordinate(m),  $z$ : Depth(m).

statistical models explicitly. By using CPTs and the conversion formulas for CPTs, since a sufficient amount of data with high spatial resolution inside the dam is converted into  $N$ -values and fines content  $F_c$ , a large amount of  $N$ -values and  $F_c$  data can be utilized for the statistical model identification.

The representative variable for soil parameter  $b$  is defined by Eq. (3.20). The hypothesis is commonly used to model soil parameters (e.g., Vanmarcke, 1983; Phoon and Kulhawy, 1999). Herein, based on the method of maximum likelihood estimation (MLE), the statistical models are determined for CPT  $N$ -value  $\log N_c$  and fines content  $\log F_{clc}$ . The modeling method employed here is similar to that in Matsuo's work (Matsuo and Asaoka, 1977), and the detailed procedure for the statistical model identification is described in Nishimura et al. (2016a) and also in Chapter 3. In the model identification based on MLE, the Akaike Information Criterion (AIC) is used. The mean functions, the standard deviations, and the covariance functions of  $\log N_c$  and  $\log F_{clc}$  are determined based on the procedure of the Minimum Akaike Information Criterion Estimator (MAICE) (Akaike, 1974). The covariance function is assumed to be of an exponential type and composed of several parameters, such as the standard deviation, the correlation distance in the horizontal and depth directions, and the nugget effect parameter.

Table 6.1 presents the mean function and the covariance function of  $\log N_c$  and  $\log F_{clc}$  estimated by the MAICE. Fig. 6.5 (a) and (b) show the distributions of  $\log N_c$  and  $\log F_{clc}$ , respectively, with the mean functions at the site. As the mean functions, quadratic equations are selected in the horizontal and vertical directions.

As shown in Table 6.1, lateral correlation distance  $l_x$  in the model of  $\log N_c$ , obtained by the MAICE, is not reasonable, because the value is identified as being the lower boundary value of 0.01 m. With the MAICE, the correlation structure of the multi-points distributed multi-dimensionally is evaluated simultaneously, and this procedure sometimes creates difficulty. On the other hand, since calculating the semi-variogram (Journel and Huijbregts, 1978) is the method used to identify the

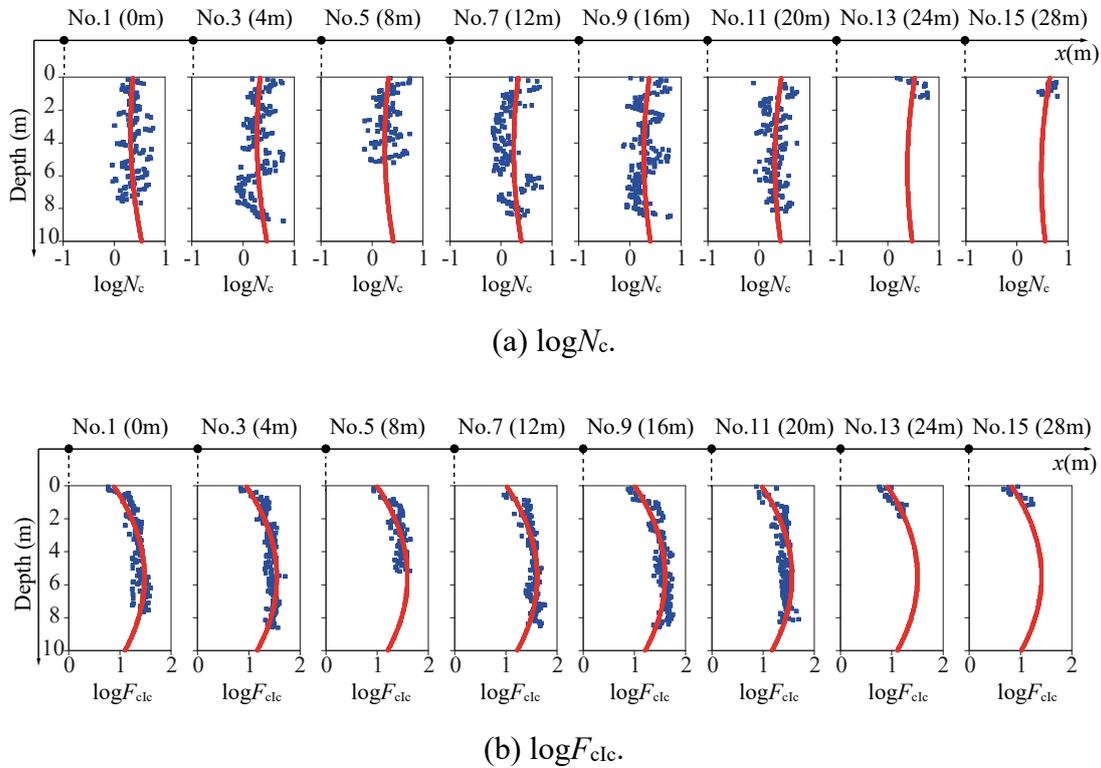
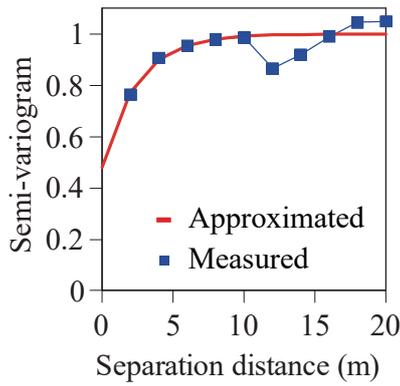


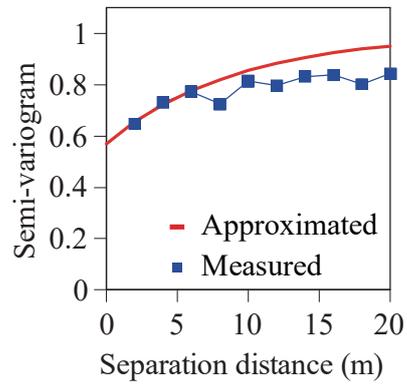
Fig. 6.5 Values measured by CPTUs and mean function with intervals of 4 m

correlation distance one-dimensionally, it is easier to use the semi-variogram than the MAICE for finding the correlation distance. The semi-variograms of horizontal direction  $\hat{\gamma}_x$  and depth direction  $\hat{\gamma}_z$  are defined by Eq. (3.26). Fig. 6.6 and Fig. 6.7 show the semi-variograms of  $\log N_c$  and  $\log F_{clc}$  for the horizontal and depth directions, respectively, of the dam. In the calculation of semi-variograms, the measured values are assumed to be standard normal distributions  $N(0, 1)$ . For example,  $\log N_c$  is normalized as  $f(x, z) = (b(x, z) - m(x, z)) / \sigma$  to remove the trend, where  $b$  is the measured value (i.e.,  $\log N_c$  here),  $m$  is the mean value, and  $\sigma$  is the standard deviation; the  $m$  and  $\sigma$  are obtained with the MAICE. In addition, in order to set the  $f$  values to be  $N(0, 1)$  surely, the  $f$  is converted into a standardized value,  $Y$ , using Eq. (3.25). To identify the geostatistical parameters of a standardized value  $Y$ , the approximation curve is simply modeled here by an exponential function as Eq. (3.27), and the covariance function is assumed as Eq. (3.28) here.

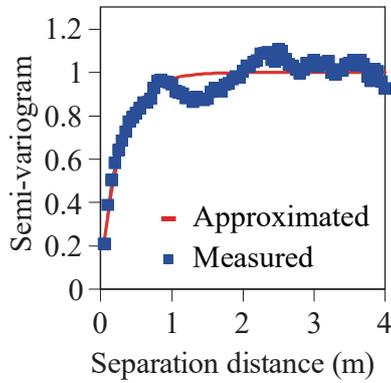
As shown in Fig. 6.6 and Fig. 6.7, the approximation curves roughly correspond to the observed values. These approximation curves are derived by the least



(a) Horizontal direction.

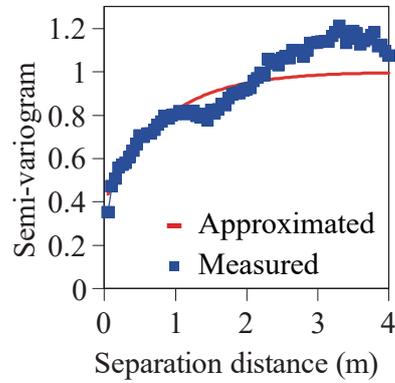


(a) Horizontal direction.



(b) Depth direction.

Fig. 6.6 Semi-variograms and approximation function of  $\log N_c$



(b) Depth direction.

Fig. 6.7 Semi-variograms and approximation function of  $\log F_{clc}$

Table 6.2 Constants of covariance functions of  $\log N_c$  and  $\log F_{clc}$  determined by semi-variograms

	$C_0$	$C_1$	C.L.(m)
$\log N_c$	$C_{0x}=0.48, C_{0z}=0.08$	$C_{1x}=0.52, C_{1z}=0.92$	$l_x=2.42, l_z=0.29$
$\log F_{clc}$	$C_{0x}=0.57, C_{0z}=0.40$	$C_{1x}=0.43, C_{1z}=0.60$	$l_x=9.17, l_z=0.82$

C.L. : Correlation length

squares method, and finally, the parameters of Eq. (3.27), namely,  $C_{0x}$ ,  $C_{1x}$ ,  $C_{0z}$ ,  $C_{1z}$ ,  $l_x$ , and  $l_z$ , are determined. The separation distances of the data,  $\Delta x = |x_i - x_j|$  and  $\Delta z = |z_i - z_j|$ , are less than 6 m in the horizontal direction and less than 0.9 m in the depth direction. They are applied to identify the approximate functions, since the accuracy of the semi-variograms is high within the range of small values for  $\Delta x$  and  $\Delta z$ . Table 6.2 shows the constants determined by the semi-variograms. The lateral correlation distance obtained from the semi-variograms is about 10 times that of the vertical one. The

identified correlation distances are supposed to be reasonable, because this relationship agrees with the published values (Nishimura, 2007; Phoon and Kulhawy, 1999).

#### 6.4 Interpolation of measured values of CPTs

The spatial variability of the soil parameters inside an earth-fill dam could influence the liquefaction probability. The novelty of the proposed work comes from the statistical modeling of the SPT  $N$ -value and  $F_c$  based on the CPT, including the conversion of the measured value of the CPT into the SPT  $N$ -value and  $F_c$  for evaluating the liquefaction probability.

Moreover, since the data measured by the CPTs are point estimation values, the interpolation of these values is required for visualizing the spatial variability of the soil parameters. Thus, the geostatistical software library GSLIB (Deutsch and Journel, 1992) is used as a conditional simulation tool. In the conditional simulation, the realizations of the simulation are based on the statistical models of the soil parameters which are characterized by their first two moments and their covariance function. The realizations generated by the simulation are conditioned by the measured data, and the values of the realizations coincide with the measured values at the measuring points.

Utilizing the simulation tool, the cross section of the studied dam is analyzed along the embankment axis. The cross section is meshed into several grids at intervals of 1 m in the horizontal direction and 0.05 m in the depth direction. The grid spacing is selected such that it corresponds to the intervals of the CPT data. In the simulation, the Monte Carlo method is repeated 2000 times to evaluate the spatial distribution of the expected values for  $N_{\text{SPT}}$  and  $F_c$  at the grid points. The spatial structures of  $\log N_{\text{SPT}}$  and  $\log F_c$  are assumed to be the same as those of  $\log N_c$  and  $\log F_{c,c}$ , respectively.

The statistical models for  $\log N_c$  and  $\log F_{c,c}$ , introduced into the simulation, are summarized in Table 6.3. In this table, the mean functions and the standard deviations are derived from the MAICE, and the covariance functions are determined by the semi-variograms. The covariance function is assumed to be an exponential type of function.

Fig. 6.8 and Fig. 6.9 present the spatial distributions of the statistical values inside the dam derived from the conditional simulation. Fig. 6.8 (a) and Fig. 6.9 (a) correspond to the spatial distributions of the expected values for  $N_{\text{SPT}}$  and  $F_c$ , Fig. 6.8 (b) and Fig. 6.9 (b) correspond to the spatial distributions of the standard deviations for

Table 6.3 Statistical models for  $\log N_c$  and  $\log F_{c/c}$  introduced into simulation

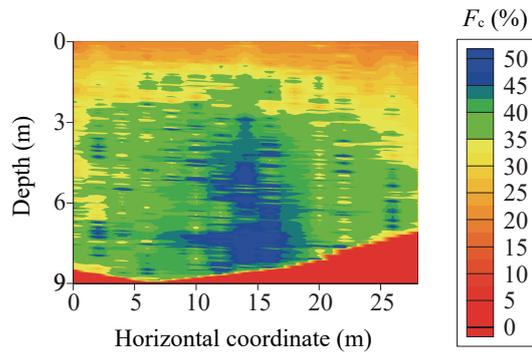
	Mean function	S.D.
$\log N_c$	$m = 0.370 - 0.012x - 0.030z + 0.001x^2 + 0.005z^2 - 0.001xz$	$\sigma = 0.246$
$\log F_{c/c}$	$m = 0.874 + 0.021x + 0.217z - 0.001x^2 - 0.019z^2 - 0.0001xz$	$\sigma = 0.123$
Covariance function ( $i, j = 1, 2, \dots, M$ )		
$\log N_c$	$C = [C_{ij}] = \sigma^2 \cdot N_e \cdot \exp\left(- x_i - x_j /2.42 -  z_i - z_j /0.29\right)$ ( $i \neq j$ ) $N_e = 0.475$ ( $ x_i - x_j  \neq 0,  z_i - z_j  \neq 0$ ) $N_e = 0.519$ ( $ x_i - x_j  \neq 0,  z_i - z_j  = 0$ ) $N_e = 0.917$ ( $ x_i - x_j  = 0,  z_i - z_j  \neq 0$ ) $C = 0.246^2$ ( $i = j$ )	
$\log F_{c/c}$	$C = [C_{ij}] = \sigma^2 \cdot N_e \cdot \exp\left(- x_i - x_j /9.17 -  z_i - z_j /0.82\right)$ ( $i \neq j$ ) $N_e = 0.256$ ( $ x_i - x_j  \neq 0,  z_i - z_j  \neq 0$ ) $N_e = 0.429$ ( $ x_i - x_j  \neq 0,  z_i - z_j  = 0$ ) $N_e = 0.598$ ( $ x_i - x_j  = 0,  z_i - z_j  \neq 0$ ) $C = 0.123^2$ ( $i = j$ )	

$m$  :Mean function, S.D.:Standard deviation,  $x$  :Horizontal coordinate(m),  $z$  :Depth(m),  
 $M$  :The number of test points,  $N_e$  :Nugget effect.

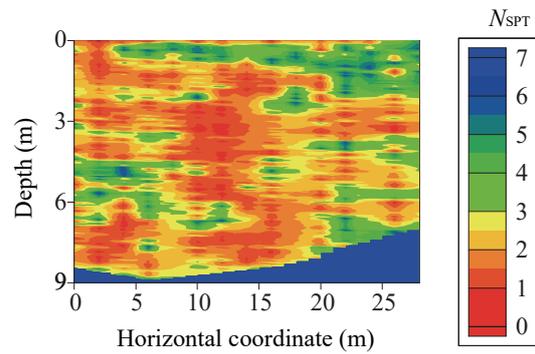
$N_{SPT}$  and  $F_c$ , and Fig. 6.8 (c) and Fig. 6.9 (c) correspond to the spatial distributions of the probability of  $F_c < 35\%$  and  $N_{SPT} < 2$ , respectively.

According to Fig. 6.8 (a), soil which comprises about 35-40%  $F_c$  is distributed in the dominant area of the dam. The soil profile for this area is categorized as intermediate soil; it consists of sandy and fine materials. The area of  $x = 10-17$  m and  $z = 3-9$  m has a relatively high  $F_c$  (about 50%), because it is located above the sluiceway inside the dam. Fig. 6.8 (b) shows the spatial distribution of the standard deviation of  $F_c$ , including the conversion error given in Eq. (2.8). Since the conversion error is considered, the standard deviation is large in the dominant area inside the dam, and the value is up to about 30%. In Fig. 6.8 (c), the probability of  $F_c < 35\%$  is more than 0.5 in the dominant area inside the dam, because the standard deviation is large.

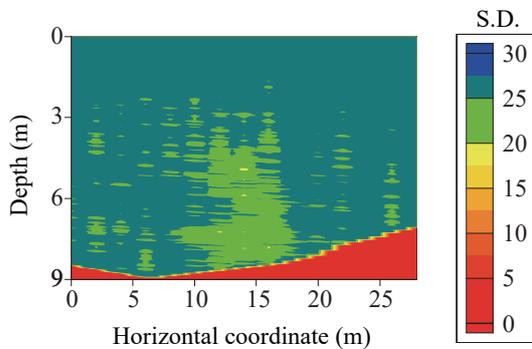
According to Fig. 6.9 (a), the overall  $N$ -value inside the dam is relatively small. In particular, there are weak areas around  $x = 0-28$  m and  $z = 3$  m and  $x = 10-17$  m and  $z = 3-9$  m. Fig. 6.9 (b) shows the spatial distribution of the standard deviation of  $N_{SPT}$ , including the conversion error, as shown in Eq. (2.7). The standard deviations are



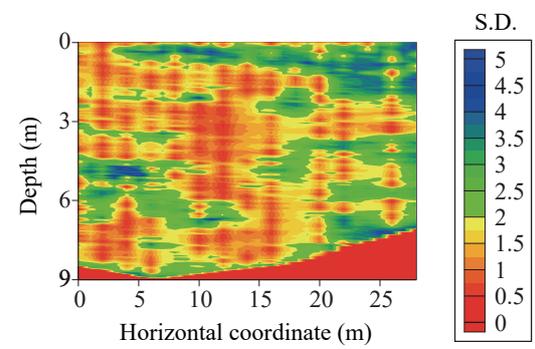
(a) Mean value ( $F_c$ )



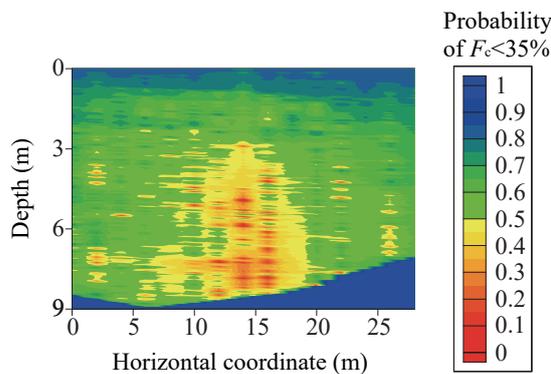
(a) Mean value ( $N_{SPT}$ )



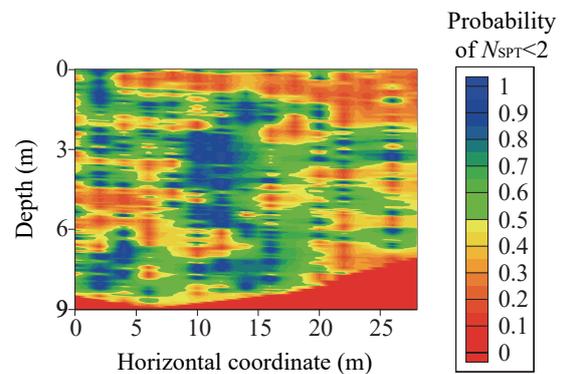
(b) Standard deviation ( $F_c$ )



(b) Standard deviation ( $N_{SPT}$ )



(c) Probability of  $F_c < 35\%$



(c) Probability of  $N_{SPT} < 2$

Fig. 6.8 Spatial distribution of statistical values of  $F_c$

Fig. 6.9 Spatial distribution of statistical values of  $N_{SPT}$

relatively small at the test points where the expected values are small. Based on Fig. 6.9 (c), the particularly weak areas inside the dam are identified. These areas are located around  $x = 10-15$  m and  $z = 3-6$  m and  $x = 0-15$  m and  $z = 7-8$  m.

## 6.5 Liquefaction resistance factor $F_L$

Liquefaction resistance factor  $F_L$  (Iwasaki et al., 1984) is calculated by the following equation:

$$F_L = \frac{R(N_{\text{SPT}}, F_c)}{L(a_{\text{max}})} \quad (6.1)$$

where  $R$  is the liquefaction resistance and  $L$  is the dynamic load.

Dynamic load  $L$ , caused by earthquakes, is defined by the following equation:

$$L = \frac{a_{\text{max}} \sigma_v}{980 \sigma'_v} r_d \quad (6.2)$$

where  $a_{\text{max}}$  is the peak ground acceleration of the dam (gal), and  $\sigma_v$  and  $\sigma'_v$  are the total and the effective vertical overburden stresses (kPa), respectively. The reduction factor  $r_d$  of dynamic load  $L$  in the depth direction is estimated, as shown in Eq. (6.3). Reduction factor  $r_d$  is obtained from a seismic response analysis inputting the time histories of bedrock acceleration at the studied site, as shown Fig. 6.10. The wave form was determined by the Japanese Cabinet Office assuming a Nankai Trough earthquake. To calculate the  $r_d$ , the validated program code “LIQCA” (e.g., Oka et al., 2016; Development Group of Liquefaction Analysis Code LIQCA, 2015) is utilized assuming the linear elastic model.

$$r_d = 0.0004z^3 - 0.0062z^2 - 0.0036z + 1.0 \quad (6.3)$$

where  $z$  is the depth (m) from the crest of the dam.

The exceedance probability of the dam depends on the peak ground acceleration of an earthquake and the review period. First, the useful lives of soil structures are assumed to be about 50 years. Second, since large earthquakes occur periodically around the Nankai Trough, namely, about every 100-200 years, seismic hazards shorter than the return period of the Nankai Trough earthquake should be considered. Therefore, the seismic hazard curve over the next 50 years, taken from J-SHIS (Japan Seismic Hazard Information Station, 2014), is employed. Fig. 6.11 shows

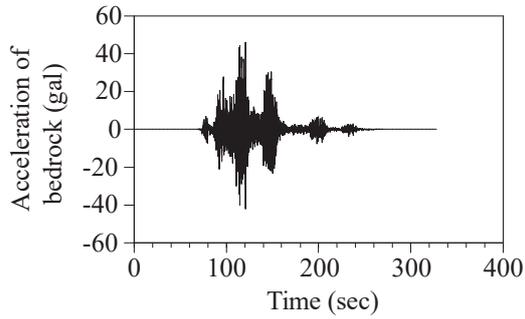


Fig. 6.10 Time histories of bedrock acceleration at studied site assuming Nankai Trough earthquake

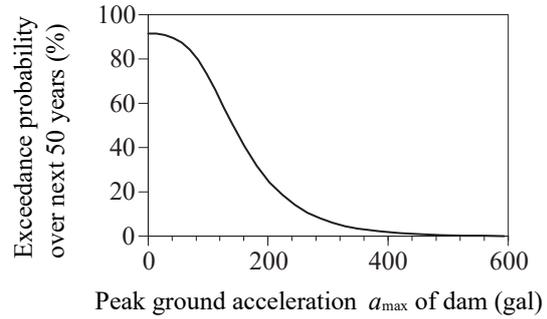


Fig. 6.11 Seismic hazard curve of Nankai Trough earthquake for studied dam

the seismic hazard curve over the next 50 years at the top of the dam,  $H$ . In other words, in order to obtain seismic hazard curve  $H$ , the maximum velocity on the bedrock at the site, provided by J-SHIS, is converted into the maximum acceleration at the top of the dam through the use of the seismic response analysis.

Liquefaction resistance  $R$  is given by the following empirical equations (Japan Road Association, 2012). The idea for the equations has two modified points compared to Iwasaki's method. One is that liquefaction resistance  $R$  is calculated from fines content  $F_c$  instead of mass median diameter  $D_{50}$ . The other is that the effects of the particle size distribution are considered using  $F_c$  in order to correct the SPT  $N$ -values,  $N_{SPT}$ . The  $N_{SPT}$  values employed here do not include energy correction, because they are derived from CPTs through the use of the conversion formulas.

$$R = \begin{cases} 0.0882\sqrt{N_a/1.7} & (N_a < 14) \\ 0.0882\sqrt{N_a/1.7} + 1.6 \times 10^{-6}(N_a - 14)^{4.5} & (14 \leq N_a) \end{cases} \quad (6.4)$$

$$\left. \begin{aligned} N_a &= c_1 \cdot N_1 + c_2 \\ N_1 &= 170N_{SPT} / (\sigma'_v + 70) \end{aligned} \right\} \quad (6.5)$$

$$c_1 = \begin{cases} 1.0 & (0\% \leq F_c < 10\%) \\ (F_c + 40) / 50 & (10\% \leq F_c < 60\%) \\ F_c / 20 - 1 & (60\% \leq F_c) \end{cases} \quad (6.6)$$

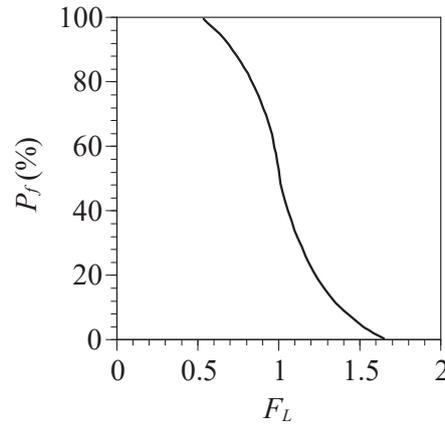


Fig. 6.12 Relationship between liquefaction probability  $P_f$  and liquefaction resistance factor  $F_L$  (Relationship was derived from data included in Iwasaki et al., 1984)

$$c_2 = \begin{cases} 0 & (0\% \leq F_c < 10\%) \\ (F_c - 10)/18 & (10\% \leq F_c) \end{cases} \quad (6.7)$$

Here,  $N_a$  is the corrected  $N$ -value including the effects of the particle size distribution,  $N_1$  is the converted  $N$ -value as the effective overburden stress equivalent to 100 kPa (1 atm),  $\sigma'_v$  is the effective vertical overburden stress (kPa), and  $F_c$  is the fines content (%).

## 6.6 Calculation of liquefaction probability

The effects of liquefaction damage on structures must be quantitatively clarified in order to establish the proper seismic design for structures. Liquefaction resistance factor  $F_L$  is used to evaluate liquefaction probability  $P_f$ . The relationship between factor  $F_L$  and  $P_f$  is given based on many in-situ investigations (Iwasaki et al., 1984), as shown in Fig. 6.12. The relationship has also been examined in other articles, such as those by Lai et al. (2005) and Juang et al. (2008). However, their relationships were derived from the cyclic resistance ratio (CRR) and the cyclic stress ratio (CSR) through the use of the SPT. Since liquefaction resistance  $R$  and dynamic load  $L$  are employed to calculate  $F_L$ , as shown in Eq. (6.8), Iwasaki's relationship is utilized to transform  $F_L$  into  $P_f$  by function  $g$ , as shown in Eq. (6.9).

$$F_L = F_L(x, z, a_{\max}, N_{\text{SPT}}, F_c) = R(x, z, N_{\text{SPT}}, F_c) / L(z, a_{\max}) \quad (6.8)$$

$$P_f(x, z, a_{\max}) = g(F_L) = g(x, z, a_{\max}, N_{\text{SPT}}, F_c) \quad (6.9)$$

for which function  $g$  is given in Fig. 6.12. In the numerical analysis, function  $g$  is introduced as a digital value.

In the analysis, random variables generated from the conditional simulation tool SGSIM, included in GSLIB (Deutsch and Journel, 1992), are assigned to  $N_{\text{SPT}}$  and  $F_c$ . Since the Monte Carlo method is repeated 2000 times in the analysis, the expected value for liquefaction probability  $P_{fE}$  as the fragility is obtained at the grid points of the analyzed cross section by the following equation. Herein, fragility means the liquefaction probability over the next 50 years as corresponds to  $a_{\max}$ .

$$P_{fE}(x, z, a_{\max}) = E\left\{g\left[F_L(x, z, a_{\max}, N_{\text{SPT}}, F_c)\right]\right\} \quad (6.10)$$

Considering all the probabilistic variables, namely,  $N_{\text{SPT}}$ ,  $F_c$ , and the peak ground acceleration of the dam over the next 50 years,  $a_{\max}$ , the liquefaction probability over the next 50 years is calculated by Eq. (6.11) at the output points of the simulation  $(x, z)$  inside the dam.

$$P_{fE50}(x, z) = -\int_0^{\infty} \frac{dH(a_{\max})}{da_{\max}} \cdot P_{fE}(x, z, a_{\max}) da_{\max} \quad (6.11)$$

where  $P_{fE50}$  is the liquefaction probability over the next 50 years at  $(x, z)$  inside the dam and  $H$  is the seismic hazard curve over the next 50 years at the top of the dam, as shown in Fig. 6.11.

To compare the degree of vulnerability against liquefaction among many dams all over Japan, the spatial average of  $P_{fE}$  and  $P_{fE50}$  of a dam could be useful. This is because the local liquefaction is difficult to compare quantitatively. The spatial average of  $P_{fE}$  is obtained for evaluating the fragility of the whole dam by Eq. (6.12).

$$\bar{P}_{fE}(a_{\max}) = \frac{1}{L_x} \int_{L_x} \frac{\int_0^{20} P_{fE}(x, z, a_{\max})(10 - 0.5z) dz}{\int_0^{20} (10 - 0.5z) dz} dx \quad (6.12)$$

$\bar{P}_{fE}$  is the spatial average of the expected value for the liquefaction probability of the whole dam, and  $L_x$  is the horizontal length of the site investigation at the studied site, namely, 28 m. It is reasonable that the effect of liquefaction near the surface is more important than that in deeper areas. Therefore,  $P_{fE}$  is weighted based on the linear function of the depth direction. The weight decreases in the direction of depth and corresponds to 0 at  $z=20$  m. This assumption is determined based on the results of the site investigations shown in Iwasaki et al. (1984).

The spatial average of  $P_{fE50}$  is calculated with the following equation:

$$\bar{P}_{fE50} = -\int_0^{\infty} \frac{dH(a_{\max})}{da_{\max}} \cdot \bar{P}_{fE}(a_{\max}) da_{\max} \quad (6.13)$$

where  $\bar{P}_{fE50}$  is the liquefaction probability of the whole dam over the next 50 years.

## 6.7 Evaluation of liquefaction probability over next 50 years

In Fig. 6.13, the seismic hazard curve at the crest of the dam, shown in Fig. 6.11, and the fragility curve of the whole dam, calculated by Eq. (6.12), are summarized. The seismic hazard shown in Fig. 6.13 is relatively sensitive to the span of  $a_{\max}$  from 111 to 223 gal. In addition, the fragility is relatively sensitive to the span of  $a_{\max}$  from 70 to 203 gal. For  $a_{\max}$  greater than 265 gal, the fragility becomes greater than 95%.

Fig. 6.14 shows the cumulative distribution function of the liquefaction probability of the whole dam over the next 50 years,  $\bar{P}_{fE50}$ , derived from Eq. (6.13). In the calculation of  $\bar{P}_{fE50}$ , the convolution integral of the seismic hazard,  $H$ , and the fragility curve, shown in Fig. 6.13, is executed. As a result,  $\bar{P}_{fE50}$ , which is the spatial average of the  $P_{fE50}$  of the dam, is evaluated as 61%. The value of the vertical axis in Fig. 6.14 shows the exceedance probability over the next 50 years; it increases about 32% between the peak ground acceleration,  $a_{\max}$ , of 160 and 224 gal. It is derived from the condition whereby not only is the fragility high, but the probability of the occurrence of an earthquake is high in that range of  $a_{\max}$ , as shown in Fig. 6.13. As  $\bar{P}_{fE50}=61\%$  is not small, the liquefaction probability of the dam against a Nankai Trough earthquake cannot be ignored.

Fig. 6.15 shows the spatial distribution of the expected value of the liquefaction

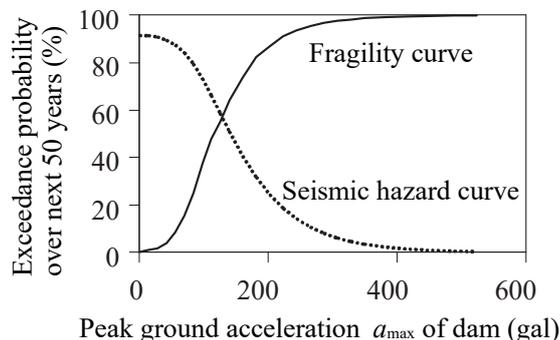


Fig. 6.13 Seismic hazard curve and fragility curve of dam

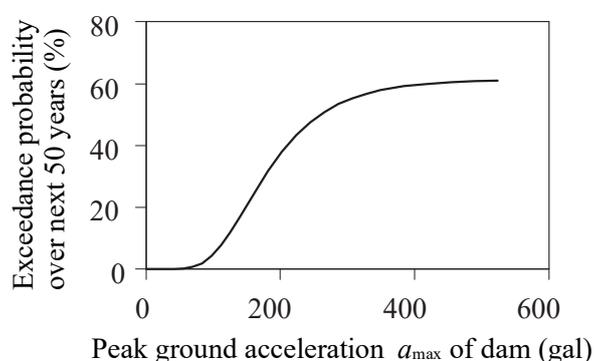


Fig. 6.14 Cumulative distribution function of spatial average of liquefaction probability of dam over next 50 years  $\bar{P}_{fE50}$

probability when  $a_{\max}$  is equivalent to 140 gal. It is predicted that the value of  $a_{\max} = 140$  gal will rise to nearly 50% over the next 50 years at the site.  $N_{\text{SPT}}$  is relatively small at  $x = 0-16$  m,  $z = 7-8$  m and at  $x = 10-13$  m,  $z = 3-6$  m, as shown in Fig. 6.9 (a). Corresponding to the tendency of  $N_{\text{SPT}}$ , the liquefaction probability in these areas is relatively high.

Fig. 6.16 shows the spatial distribution of  $P_{fE50}$  inside the dam calculated by Eq. (6.11). According to this figure, the liquefaction probability over the next 50 years is seen to be relatively high at similar locations to those in Fig. 6.15. Based on the results shown in Fig. 6.15 and Fig. 6.16, it seems that the  $N_{\text{SPT}}$  values have a greater influence on the liquefaction probability.

Since the CPTs were conducted at short intervals for examining the spatial variability of  $N_{\text{SPT}}$  and  $F_c$ , it was possible to evaluate in detail the location and the scale of the weak areas against the liquefaction inside the dam over the next 50 years. Based on the spatial distribution of the liquefaction probability, the diagnosis of an earth-fill dam against liquefaction was properly made. In addition, introducing the seismic hazard

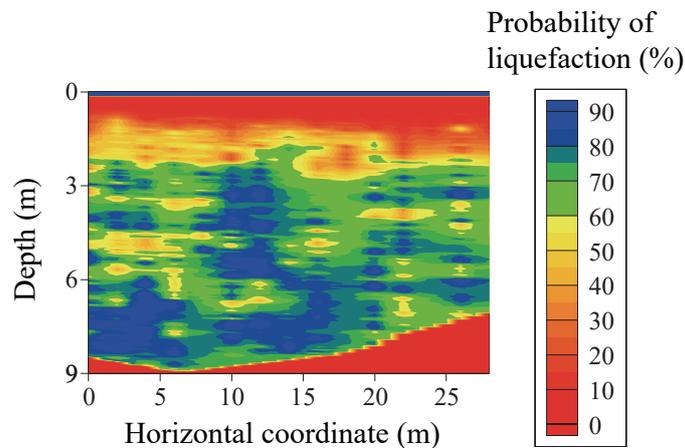


Fig. 6.15 Spatial distribution of expected values of liquefaction probability when  $a_{\max}$  is equivalent to 140 gal

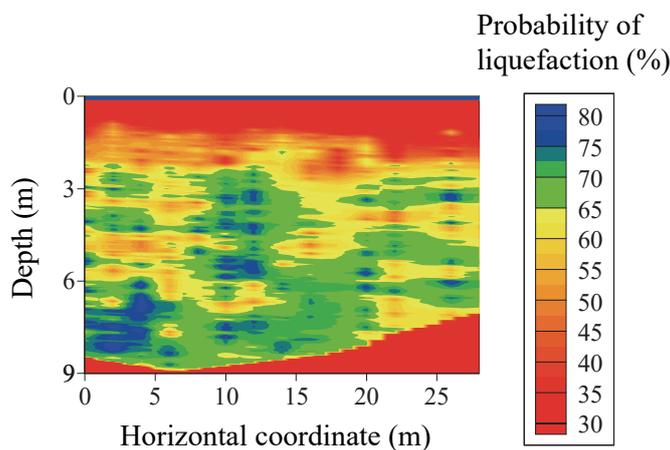


Fig. 6.16 Spatial distribution of expected values of liquefaction probability over next 50 years

and the fragility curve into the calculation of the liquefaction probability of an earth-fill dam, the liquefaction hazard during a selected time period could be incorporated. Using the seismic hazard and the fragility curve of the dams, if the dams are located in different parts of Japan, the priority of the improvements to be made to them against the probable liquefaction damage of each dam can be compared based on the proposed method. Therefore, the proposed method for evaluating the liquefaction probability of an earth-fill dam can provide useful information for determining the order of priority among the many dams in Japan in terms of making improvements to them.

## 6.8 Conclusion

The spatial distribution of the liquefaction probability inside the studied dam has been calculated in detail from CPTUs including conversion errors and the spatial variability of the soil parameters. In addition, based on the seismic hazard at the studied site and the fragility of the dam, the spatial average of the liquefaction probability of the dam over the next 50 years has been evaluated. The concluding remarks are summarized below.

- 1) The conversion errors, from the CPT  $N_c$ -value,  $N_c$ , to the SPT  $N$ -value,  $N_{SPT}$ , and from the fines content,  $F_{cIc}$ , obtained from  $I_c$ , to the proper fines content,  $F_c$ , have been quantified, respectively, in the conversion formulas proposed in Chapter 2. According to the results of a conditional simulation with the conversion error, the standard deviation of  $F_c$  has been calculated as nearly 30% in the dominant area of the studied dam. The value seems large compared with the mean value.
- 2) The spatial structures of  $\log N_c$  and  $\log F_c$  have been determined based on the results obtained from the CPTUs. Since the CPTUs were conducted at 2-m intervals in the horizontal direction, a sufficient amount of  $F_c$  data has been estimated from the CPTUs through the use of conversion formulas. As a result, the correlation distances of  $\log N_c$  and  $\log F_c$  have been reasonably identified, because the horizontal value was about 10 times that of the vertical one.
- 3) Utilizing the conditional simulation, the spatial distribution of the expected values for  $F_c$  and  $N_{SPT}$ , the standard deviation of  $F_c$  and  $N_{SPT}$ , and the probability that the simulated values would be lower than the prescribed threshold values for  $F_c$  and  $N_{SPT}$ , respectively, inside the dam have been visualized in detail. These results clarify the scale of the weak areas and the spatial continuity of the soil profile inside the dam.
- 4) The seismic hazard at the studied site and the fragility against the liquefaction of the dam have been introduced into the evaluation of the liquefaction probability of the dam over the next 50 years. As a result, it has been confirmed that the spatial average of the liquefaction probability of the dam over the next 50 years has been calculated as 61%. Although the desired value of the liquefaction probability of a dam depends on the situation, the liquefaction probability of dams calculated by the proposed method can offer supporting evidence for decision-making in terms of improving the dams.

Nevertheless, a large quantity of data is required to identify the spatial structure of the soil parameters in order to evaluate the liquefaction probability of a dam, and the

amount of test data is not enough for the statistical modeling. Therefore, to calculate the liquefaction probability with limited information, a database for the spatial structure of the soil parameters needs to be developed. Although the data obtained from SPTs or laboratory tests are generally sparse for characterizing the statistical model, a sufficient amount of data on the soil parameters can be obtained by applying conversion formulas to the measured data from CPTs. Thus, the proposed method could be useful for properly evaluating the spatial structure of the soil parameters and for developing a database for the spatial structure of the soil parameters.

A common problem for evaluating the liquefaction probability of an earth-fill dam is the limited time and the limited budget for site investigations. It has been shown that the proposed method can accurately evaluate the liquefaction probability of a dam based on  $N$ -value and  $F_c$ ; and thus, the accumulated soil data from past site investigations, such as SPT  $N$ -value and  $F_c$ , can be easily introduced into the proposed method. Several methods to synthesize the different sources of the investigation data have been proposed (e.g., Nishimura et al., 2016a; Cardarelli et al., 2014), and their applicability to the evaluation of the liquefaction probability should be studied.

## Chapter 7

### Concluding remarks and future work

The following three sections are arranged in this chapter. First, the highlights of each chapter are summarized. Second, the whole conclusion of this paper is presented. Lastly, future works related to this research are mentioned.

#### 7.1 Summary of each chapter

This paper mainly addressed the evaluation of the spatial variability of the soil parameters inside earth-fill dams based on the results obtained from Piezometer Cone Penetration Tests (CPTUs). The results in the preceding chapters will be summarized once again.

Chapter 1 showed the objective of this thesis, a review of past literature, and the composition of this thesis. By reviewing the past literature, the unique point of this thesis has been clarified.

In Chapter 2, the advantages of the cone penetration test for examining the spatial variability of the soil properties inside earth-fill dams were explained. The CPT has the characteristics of often providing more continuous and precise data in the depth direction at a better speed and a lower cost. In addition, static testing, like the CPT, can be performed more quickly than dynamic testing, like the SPT. Therefore, the CPTs conducted at short intervals could be performed in the horizontal direction. The results obtained from the CPT were transformed into several soil properties using the conversion formulas, for instance, the soil types, the  $N$ -value, and the fines content,  $F_c$ .

Chapter 3 summarized several methods for modeling the soil properties, and the fundamental theories were explained. The optimum statistical model for the soil properties, based on the random field theory, was selected using the maximum likelihood method, and the spatial distribution of the soil properties was visualized through the use of a geostatistical method. To select the optimum statistical model, Akaike's Information Criterion (AIC), which is one of the MLEs was employed. In the modeling of soil for practical problems, the weakly stationary of the random field theory was assumed, and information on the locations of the in-situ data was considered. To deal with evaluations of the spatial distributions of the soil properties of existing soil

structures, one of the geostatistical methods, called kriging, was used. The spatial distribution of the soil properties was evaluated here by the conditional simulation, which includes the application of kriging.

In Chapter 4, the correlation distances of the soil strength, used to model the spatial variability of them inside earth-fill dams, were summarized. The correlation distances at five earth-fill dams were examined based on CPTUs carried out at short intervals. Since the  $N$ -value was calculated from the results obtained from the CPTUs considering the quantity of the conversion errors, the statistical model was appropriately determined using a sufficient number of  $N$ -values. To secure the stationarity of the data, the outliers of the data were removed from the calculation of the semi-variogram, and the correlation distances in the horizontal direction and the depth direction were estimated, respectively. The obtained correlation distances reasonably corresponded with the values shown in past literature.

In Chapter 5, an evaluation method for the spatial distribution of the soil strength inside earth-fill dams, composed of materials with different particle size distributions, was proposed using a geostatistical method based on the results obtained from CPTUs. It was seen that, in the CPTUs, the spatial distributions of the weak areas and the strong areas originated from the amount of gravel mixed into the soil and that they affected the soil strength. The novelty of the proposed method is in the re-composition of the simulation values of the three groups, namely, high,  $Y_H$ , middle,  $Y_M$ , and low,  $Y_L$ . The  $Y_H$  and  $Y_L$  groups model the outliers of high strength and low strength, respectively. In the proposed method, the rate of outliers is determined from the measured values. The simulated values for each of the three groups are re-composed so as to follow the determined rate, and the locations of the outliers are determined based on the simulated values of the middle range,  $Y_M$ .

In Chapter 6, the spatial distribution of the liquefaction probability inside the studied dam was calculated in detail from CPTUs including conversion errors and the spatial variability of the soil parameters. In addition, based on the seismic hazard at the studied site and the fragility of the dam, the spatial average of the liquefaction probability of the dam over the next 50 years was evaluated. Utilizing the conditional simulation, the spatial distribution of the expected values for  $F_c$  and  $N_{SPT}$ , the standard deviation of  $F_c$  and  $N_{SPT}$ , and the probability that the simulated values would be lower than the prescribed threshold values for  $F_c$  and  $N_{SPT}$ , respectively, inside the dam were visualized in detail. These results clarify the scale of the weak areas and the spatial continuity of the soil profile inside the dam.

## 7.2 Summary of research

This paper mainly discussed the evaluation of the spatial variability of soil parameters inside earth-fill dams based on the results obtained from Piezometer Cone Penetration Tests (CPTUs). As a result, it was confirmed that the CPTUs conducted at short intervals were able to appropriately detect the spatial variability of the soil properties inside earth-fill dams for the evaluation of the stability. A database of the correlation distances of the soil properties obtained from five earth-fill dams was presented, an evaluation method for the spatial distribution of the soil strength considering the outliers was given, and an evaluation method for the liquefaction probability over the next 50 years was described. They will all provide useful information for effectively performing the diagnosis of earth-fill dams.

## 7.3 Future works

To diagnose the health of an earth-fill dam and other soil structures, the limited time and the limited budget for the site investigations are common problems. Therefore, the precise identification of the weak areas inside soil structures is generally difficult by the shortage of the information. In addition, the local weak areas which cannot be identified, could be one of the reasons cause the piping phenomena. Several challenges remain to be overcome for efficiently diagnose the health of earth-fill dams considering the spatial variability, and the three points are summarized below.

First, a database of the soil properties of the correlation structure should be enhanced. This is because there are 200 thousand of earth-fill dams in Japan, and attempting to investigate each dam would be difficult in terms of time and economic factors. The database would provide useful information to assume the spatial variability of the soil properties of each dam. For example, the general values of the correlation length of earth-fill dams could be used for the dams which have insufficient investigations to incorporate the spatial variability. In the similar manner, river dikes would be evaluated considering the spatial variability by using the general values of the correlation lengths of river dikes. Based on the database, the spatial variability of the soil properties would be easily incorporated into the diagnosis of soil structures.

Second, in Chapter 6, it was shown that a proposed method can evaluate the liquefaction probability inside a dam based on  $N$ -value and  $F_c$ . Since  $N$ -value and  $F_c$  have been used for the several design analyses, and accumulated for a long time in

Japan, and the data are utilized as the data base of the geotechnical parameters. The proposed method to evaluate the liquefaction resistance in Chapter 6, has the capability to easily introduce such the accumulated data of  $N$ -value and  $F_c$ . In Chapter 2, Eqs. (2.5), (2.6), (2.7), and (2.8) present the conversion formulas to derive  $N$ -value and  $F_c$  based on the results obtained from CPTUs. By synthesizing the accumulated data of  $N$ -value and  $F_c$  obtained from boring tests and the converted values of  $N$ -value and  $F_c$  derived from CPTUs, the information of the soil properties in studied sites can be enhanced. The further study about the synthesis method of accumulated data and converted data should be examined to make the diagnosis sufficiently reflect actual situation of soil structures.

Third, since the proposed approach requires site investigations in short interval to estimate the spatial variability, this approach has difficulty to conduct at the earth-fill dams or the river dikes, which have long length. It is because the number of testing points is generally limited by economic factor. Therefore, an efficient procedure to evaluate the spatial variability at the soil structures, which have long length is required. For example, as a one of efficient ways, the investigation plan as shown in Fig. 4.1 (e) was conducted. Fig. 4.1 (e) shows that the CPTUs were conducted in two different intervals at E dam. The one is 50m and the other is 5m. The tests with long interval aims to comprehend the trends of the soil properties in long length and to comprehend the areas where should be diagnosed precisely. On the other hand, the tests with short interval aims to grasp the parameters of the spatial variabilities and to visualize the weak areas inside the soil structures in question. In addition, as the efficient investigation method instead of sounding tests, geophysical exploration can be used. As an example of efficient methods to spatially evaluate the soil strength, Nishimura et al., (2016a) synthesized the results of SWS and Surface wave method (SWM) using a geostatistical method. The syntheses of the results obtained from efficient method like the geophysical exploration are one of the useful solutions to efficiently perform the diagnosis of soil structures.

## Chapter 8

### References

- Akaike, H., 1974. A New Look at the Statistical Model Identification. *IEEE Trans. Automat. Contr.* 19, 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Baecher, G.B., Christian, J.T., 2003. *Reliability and statistics in geotechnical engineering*. Wiley.
- Banerjee, S., Carlin, B.P., Gelfand, A.E., 2014. *Hierarchical modeling and analysis for spatial data*. Chapman and Hall/CRC.
- Bombasaro, E., Kasper, T., 2016. Evaluation of spatial soil variability in the Pearl River Estuary using CPTU data. *Soils Found.* 56, 496–505.
- Bong, T., Stuedlein, A.W., 2017. Spatial variability of CPT parameters and silty fines in liquefiable beach sands. *J. Geotech. Geoenvironmental Eng.* 143, 4017093.
- Cardarelli, E., Cercato, M., De Donno, G., 2014. Characterization of an earth-filled dam through the combined use of electrical resistivity tomography, P- and SH-wave seismic tomography and surface wave data. *J. Appl. Geophys.* 106, 87–95. <https://doi.org/https://doi.org/10.1016/j.jappgeo.2014.04.007>
- Cetin, K.O., Seed, R.B., Der Kiureghian, A., Tokimatsu, K., Harder, L.F., Kayen, R.E., Moss, R.E.S., 2004. Standard Penetration Test-Based Probabilistic and Deterministic Assessment of Seismic Soil Liquefaction Potential. *J. Geotech. Geoenvironmental Eng.* 130, 1314–1340. [https://doi.org/10.1061/\(ASCE\)1090-0241\(2004\)130:12\(1314\)](https://doi.org/10.1061/(ASCE)1090-0241(2004)130:12(1314))
- Chen, Q., Wang, C., Hsein Juang, C., 2015. CPT-Based Evaluation of Liquefaction Potential Accounting for Soil Spatial Variability at Multiple Scales. *J. Geotech. Geoenvironmental Eng.* 142, 04015077. [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0001402](https://doi.org/10.1061/(ASCE)GT.1943-5606.0001402)
- Dawson, K.M., Baise, L.G., 2005. Three-dimensional liquefaction potential analysis using geostatistical interpolation. *Soil Dyn. Earthq. Eng.* 25, 369–381. <https://doi.org/10.1016/j.soildyn.2005.02.008>
- DeGroot, D.J., Baecher, G.B., 1993. Estimating autocovariance of in-situ soil properties. *J. Geotech. Eng.* 119, 147–166.
- Deutsch, C. V., Journel, A.G., 1992. *Geostatistical Software Library and User's Guide*. Oxford, England: Oxford University Press.

## References

- Development Group of Liquefaction Analysis Code LIQCA, 2015. Data of LIQCA2D15.
- Fenton, G.A., Griffiths, V.D., 2008. Risk assessment in geotechnical engineering. John Wiley & Sons.
- Griffiths, D. V., Huang, J., Fenton, G.A., 2009. Influence of Spatial Variability on Slope Reliability Using 2-D Random Fields. *J. Geotech. Geoenvironmental Eng.* 135, 1367–1378. [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0000099](https://doi.org/10.1061/(ASCE)GT.1943-5606.0000099)
- Honjo, Y., 2011. Challenges in geotechnical reliability based design, in: N.Vogt et al. (Ed.), *Proc. of the 3rd International Symposium on Geotechnical Safety and Risk*. pp. 11–27.
- Honjo, Y., 2008. General vs. local reliability based design in geotechnical engineering, in: Katafygiotls, L.S., Zhang, L., Tang, W.H., Cheung, M. (Eds.), *Proceedings of the 4th Asian-Pacific Symposium on Structural Reliability and Its Applications (APSSRA'08)*. pp. 41–52.
- Honjo, Y., Otake, Y., 2012. Verification of statistical estimation error evaluation theory of local averages of geotechnical parameters. *J. Japan Soc. Civ. Eng. Ser. C (Geosph. Eng.* 68, 491–507. <https://doi.org/10.2208/jscejge.68.491>
- Honjo, Y., Setiawan, B., 2007. General and local estimation of local average and their application in geotechnical parameter estimations. *Georisk* 1, 167–176.
- Imaide, K., Nishimura, S., Shibata, T., Shuku, T., 2018. Evaluation of the spatial variability of cone penetration resistance inside an earth-fill dam composed of materials with different particle size distribution. *J. Japan Soc. Civ. Eng. Ser. C (Geosph. Eng.* 74, 213–224. <https://doi.org/10.2208/jscejge.74.213>
- Iwasaki, T., Arakawa, T., Tokida, K., 1984. Simplified Procedures for Assesing Soil Liquefaction During Earthquakes. *Soil Dyn. Earthq. Eng.* 3, 49–58.
- Japan Geotechnical Society, 2016. Japanese standards and explanations of geotechnical and geoenvironmental investigation methods. Maruzen, Tokyo.
- Japan Road Association, 2012. Specification for highway bridge, Part 5 Seismic Design. Maruzen, Tokyo.
- Japan Seismic Hazard Information Station [WWW Document], 2014. URL <http://www.j-shis.bosai.go.jp/en/> (accessed 9.19.18).
- Journel, A.G., Huijbregts, C.J., 1978. *Mining geostatistics*. Academic Press, London, England, United Kingdom.
- Juang, C.H., Chen, C.J., Rosowsky, D. V., Tang, W.H., 2000. CPT-based liquefaction analysis, Part 2: Reliability for design. *J. Mater.* 15, 2321–2324.

## References

- Kasama, K., Zen, K., 2010. Reliability-Based Analysis on the Bearing Capacity of Cement-Treated Ground Considering the Spatial Variability of Shear Strength. *J. Soc. Mater. Sci. Japan* 59, 44–49. <https://doi.org/10.2472/jsms.59.44>
- Kulhawy, F.H., Mayne, P.W., 1990. Manual on estimating soil properties for foundation design. Electric Power Research Inst., Palo Alto, CA (USA).
- Kullback, S., Leibler, R.A., 1951. On Information and Sufficiency. *Ann. Math. Stat.* 22, 79–86.
- Lai, S.-Y., Lin, P.-S., Hsieh, M.-J., Jim, H.-F., 2005. Regression model for evaluating liquefaction potential by discriminant analysis of the SPT  $N$  value. *Can. Geotech. J.* 42, 856–875. <https://doi.org/10.1139/t05-020>
- Lark, R.M., 2000. A comparison of some robust estimators of the variogram for use in soil survey. *Eur. J. Soil Sci.* 51, 137–157. <https://doi.org/10.1046/j.1365-2389.2000.00280.x>
- Lenz, J.A., Baise, L.G., 2007. Spatial variability of liquefaction potential in regional mapping using CPT and SPT data. *Soil Dyn. Earthq. Eng.* 27, 690–702. <https://doi.org/10.1016/j.soildyn.2006.11.005>
- Lloret-Cabot, M., Fenton, G.A., Hicks, M.A., 2014. On the estimation of scale of fluctuation in geostatistics. *Georisk* 8, 129–140. <https://doi.org/10.1080/17499518.2013.871189>
- Lumb, P., 1974. Application of statistics in soil mechanics, in: Lee, I. (Ed.), *Soil Mechanics New Horizons*. Newnes-Butterworth.
- Lumb, P., 1966. The variability of natural soils. *Can. Geotech. J.* 3, 74–97.
- Matheron, G., 1963. Principles of geostatistics. *Econ. Geol.* 58, 1246–1266.
- Matsuo, M., Asaoka, A., 1977. Probability models of undrained strength of marine clay layer. *Soils Found.* 17, 53–68. [https://doi.org/10.3208/sandf1972.17.3\\_53](https://doi.org/10.3208/sandf1972.17.3_53)
- Matsuo, M., Asaoka, A., 1976. A statistical study on a conventional " safety factor method". *Soils Found.* 16, 75–90.
- Matsuo, M., Kuroda, K., 1974. Probabilistic Approach To Design of Embankments. *Soils Found.* 14, 1–17.
- Ministry of Agriculture Forestry and Fisheries, 2015. The Design Guideline of Land Improvement Project (Improvement of an Earth-Fill Dam). Japanese Society of Irrigation, Drainage and Rural Engineering.

## References

- Murakami, A., Nishimura, S., Suzuki, M., Mori, M., Kurata, T., Fujimura, T., 2009a. Determination of Partial Factors for Verification of Bearing Capacity of Shallow Foundations beneath Open Channels. *Trans. Japanese Soc. Irrig. Drain. Rural Eng.* 77, 71–78. <https://doi.org/10.11408/jsidre.77.71>
- Murakami, A., Nishimura, S., Suzuki, M., Mori, M., Kurata, T., Fujimura, T., 2009b. Reliability Analysis for Bearing Capacity of Open Channel. *Trans. Japanese Soc. Irrig. Drain. Rural Eng.* 77, 175–181. <https://doi.org/10.11408/jsidre.77.175>
- Nishimura, S., 2007. The investigation of the establishment method of soil parameters. *Proc. Symp. Perform. Des. Deform. Predict. Geotech. Struct. Dur. Earthq.* 121–126.
- Nishimura, S., Shibata, T., Shuku, T., 2016a. Diagnosis of earth-fill dams by synthesised approach of sounding and surface wave method. *Georisk* 10, 312–319. <https://doi.org/10.1080/17499518.2016.1197406>
- Nishimura, S., Shimizu, H., 2011. Reliability-Based Design for Liquefaction Mitigation by Using Conditional Simulation. *Trans. Japanese Soc. Irrig. Drain. Rural Eng.* 79, 329–337. <https://doi.org/10.11408/jsidre.79.329>
- Nishimura, S., Shimizu, H., 2008. Reliability-based design of ground improvement for liquefaction mitigation. *Struct. Saf.* 30, 200–216. <https://doi.org/10.1016/j.strusafe.2006.11.002>
- Nishimura, S., Shuku, T., Shibata, T., 2016b. Reliability-based design of earth-fill dams to mitigate damage due to severe earthquakes. *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* 10, 83–90. <https://doi.org/10.1080/17499518.2015.1124123>
- Nishimura, S., Takayama, Y., Suzuki, M., Murakami, A., Fujisawa, K., 2011. Prediction of spatial distribution for n-values in earth-fill embankments. *J. Japan Soc. Civ. Eng. Ser. C (Geosph. Eng.* 67, 252–263. <https://doi.org/10.2208/jscejge.67.252>
- Oka, F., Yui, H., Kimoto, S., Kinugawa, T., 2016. New Evaluation Method for Liquefaction of Ground Using Dynamic Liquefaction Analysis Method and Its Application. *Int. J. Geomech.* 16, C4016002. [https://doi.org/10.1061/\(ASCE\)GM.1943-5622.0000557](https://doi.org/10.1061/(ASCE)GM.1943-5622.0000557)
- Ono, K., Kazama, S., Kawagoe, S., Yokoo, Y., Gunawardhana, L., 2011. Possible earthen dam failure mechanisms of Fujinuma reservoir due to the Great East Japan Earthquake of 2011. *Hydrol. Res. Lett.* 5, 69–72. <https://doi.org/10.3178/hrl.5.69>
- Otake, Y., Honjo, Y., 2014. Characterization of transformation error in geotechnical structural design. *J. Japan Soc. Civ. Eng. Ser. C (Geosph. Eng.* 70, 186–198. <https://doi.org/10.2208/jscejge.70.186>

## References

- Otake, Y., Honjo, Y., 2012. Verification of a simplified geotechnical reliability analysis scheme of spatial variability based on local average of geotechnical parameters. *J. Japan Soc. Civ. Eng. Ser. C (Geosph. Eng.* 68, 475–490.  
<https://doi.org/10.2208/jscejge.68.475>
- Otake, Y., Honjo, Y., Hiramatsu, Y., Yoshida, I., Sako, S., Nakayama, O., Nagano, T., 2014. Verification of a reliability analysis for long continuous structure considering investigation sites on the 20km-long river dike incurred the earthquake disaster. *Japanese Geotech. J.* 9, 203–217. <https://doi.org/10.3208/jgs.9.203>
- Phoon, K.-K., Kulhawy, F.H., 1999. Characterization of geotechnical variability. *Can. Geotech. J.* 36, 612–624. <https://doi.org/10.1139/t99-038>
- Phoon, K.K., Retief, J. V., 2016. *Reliability of Geotechnical Structures in ISO2394*. CRC Press. <https://doi.org/10.1201/9781315364179>
- Robertson, P.K., 1990. Soil classification using the cone penetration test. *Can. Geotech. J.* 27, 151–158. <https://doi.org/10.1139/t90-014>
- Robertson, P.K., Fear, C.E., 1995. Liquefaction of sands and its evaluation. *Proc. First Int. Conf. Earthq. Geotech. Eng. IS-Tokyo'95* Vol. 3, 1253–1289.
- Stuedlein, A.W., Kramer, S.L., Arduino, P., Holtz, R.D., 2012. Geotechnical characterization and random field modeling of desiccated clay. *J. Geotech. Geoenvironmental Eng.* 138, 1301–1313.
- Suzuki, M., Ishii, K., 1988. Stochastic finite element method using estimations of spatial distributions of soil properties. *Doboku Gakkai Ronbunshu* 1988, 97–104. [https://doi.org/10.2208/jscej.1988.394\\_97](https://doi.org/10.2208/jscej.1988.394_97)
- Suzuki, Y., Tokimatsu, K., Sanematsu, T., 2003. Correlations between CPT data and soil characteristics obtained from SPT. *J. Struct. Constr. Eng. (Transactions AIJ)*. [https://doi.org/10.3130/aijs.68.73\\_2](https://doi.org/10.3130/aijs.68.73_2)
- Tatsuoka, F., Koseki, J., Takahashi, A., 2017. Earthquake-induced damage to earth structures and proposal for revision of their design policy -based on a case history of the 2011 off the Pacific coast of Tohoku Earthquake-. *J. JSCE* 5, 101–112. [https://doi.org/10.2208/journalofjsce.5.1\\_101](https://doi.org/10.2208/journalofjsce.5.1_101)
- Uzielli, M., Vannucchi, G., Phoon, K.K., 2005. Random field characterisation of stress-normalised cone penetration testing parameters. *Géotechnique* 55, 3–20. <https://doi.org/10.1680/geot.2005.55.1.3>
- Vanmarcke, E., 1983. *Random Fields: Analysis and Synthesis*, MIT Press Classics. MIT Press.
- Vanmarcke, E.H., 1977. Probabilistic modeling of soil profiles. *J. Geotech. Eng. Div.* 103, 1227–1246.

## *References*

- Vivek, B., Raychowdhury, P., 2014. Probabilistic and spatial liquefaction analysis using CPT data: A case study for Alameda County site. *Nat. Hazards* 71, 1715–1732. <https://doi.org/10.1007/s11069-013-0976-4>
- Zimmerman, D.L., 1993. Another look at anisotropy in geostatistics. *Math. Geol.* 25, 453–470. <https://doi.org/10.1007/BF00894779>