Validation of Numerical Analysis Based on Mode Decomposition

Yu Otake[1], Yosuke Higo[2], Kyohei Shigeno[1], and Shinya Watanabe[1]

[1]Department of Civil Engineering and Architecture, Niigata University, 8050, Ikarashi 2-no-cho, Nishi-ku, Niigata, 950-2181, Japan. E-mail: y_ohtake@niigata-u.ac.jp
[2] Department of Urban Management, Kyoto University, C1-211 Kyotodaigaku-Katsura, Nishikyo-ku, Kyoto, 615-8540, Japan. E-mail: higo.yohsuke.5z@kyoto-u.ac.jp

Abstract: Recently, the method of verification and validation (V&V) for sophisticated numerical analyses has received considerable attention in geotechnical engineering in the evaluation of the performance of structures. We are currently developing a V&V scheme to implement an actual structural design based on reliability analysis. The feature of the scheme involves quantifying uncertainties related to the accuracy of the parameter settings. This paper describes the deformation calculation procedure for the case of a simple embankment on a liquefiable sand layer. We conducted a dynamic effective stress analysis to evaluate the liquefaction phenomenon in the sand layer. This numerical analysis is required to define many input parameters (the number of input parameters was 14) in advance. The purpose of this study was to derive an alternative numerical analysis model that takes into account the space and time required for calculating the aforementioned deformed embankment.

Keywords: Proper orthogonal decomposition; principal component analysis; Bayesian inference; mode decomposition.

1 Background and Purpose

Recently, as a framework to evaluate the reliability of numerical analysis, some specifications related to V&V (verification and validation) have been established in Europe and in the U.S. However, details of the validation method in V&V have not been clarified presently. In this paper, we propose a validation scheme that combines the reliability-based design concept with V&V. Any numerical analysis in which reliability is quantified can be utilized for practical design in the framework of a reliability-based design concept. Therefore, we propose to quantify the reliability of numerical analysis as validation in V&V in this study. Specifically, we propose a framework to quantify the numerical analysis accuracy (reliability) by setting the accuracy of input parameters.

2 Basic Analysis Model and Data Used in This Study

The governing equations for coupling problems of the soil skeleton and pore water were obtained based on two-phase mixture theory, and the u-p (displacement of the whole mixture-pore water pressure) formulation was adopted for the two-dimensional analysis (LIQCA, Oka et al.1999). Table 1 shows an example of the input parameter sets for this analysis. In this parameter set, there are parameters determined by experiments and fitting parameters whose physical meanings are not sufficiently clear. Fitting parameters are determined based on a liquefaction strength curve obtained from laboratory tests via element simulation and the validation regarding the reproducibility of the FEM analysis. The condition of FE-analysis were determined by a centrifuge test (50G) that was previously conducted. The input wave was a sine wave with a frequency of 1 Hz, a 5 G main wave, and a 3.5 G subsequent wave in the prototype scale. The performance function, i.e., $g(\cdot)$, is denoted as follows:

$$g(\mathbf{z_p}) = D_y^{cal}(\mathbf{z_p}) \cdot \delta_m - R \cdot \delta_R \tag{1}$$

where $\mathbf{z_p} \in \mathbb{R}^{14}$ is a certain parameter set for the analysis; $D_y^{cal}(\cdot)$ is the settlement of the crest of embankment calculated by FEM analysis; R is a deterministic value representing the limit state (criteria) in the analysis, which is half of the height of the embankment; H. δ_m and δ_R represent the modeling error for the analysis and the limit state (criteria), respectively. The uncertainties of both are the same, given a variation of 10% as the coefficient of variation.

Table 1. Parameter set for the $\mathbf{z_p}$ configuration model (an example out of 19 cases; Toyoura sand).

	ρ	e	λ	K	G_0/σ_{m0}	M_f	M_m	B_0	B_1	C_f	γ_r^P	γ_r^E	D_0	n
Toyoura–Sand	0.754	0.00910	0.00052	1.2	2023.6	0.990	0.707	4089	54.5	-	0.002	0.012	0.60	5.1
Parameter Type	physical parameter							fitting parameter						

Proceedings of the 7th International Symposium on Geotechnical Safety and Risk (ISGSR)
Editors: Jianye Ching, Dian-Qing Li and Jie Zhang

Published by Research Publishing, Singapore.
ISBN: 978-981-11-2725-0; doi:10.3850/978-981-11-2725-0_IS7-9-cd

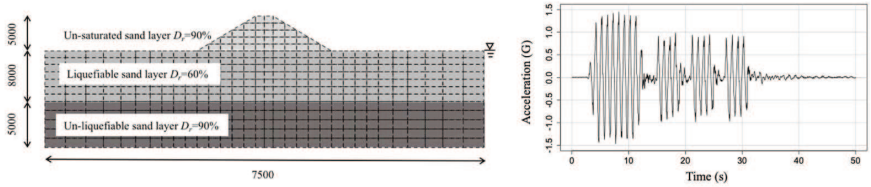


Figure 1. FE model diagram and input wave (These were modeled based on the centrifugal model experiment conducted by PWRI considering the similarity law).

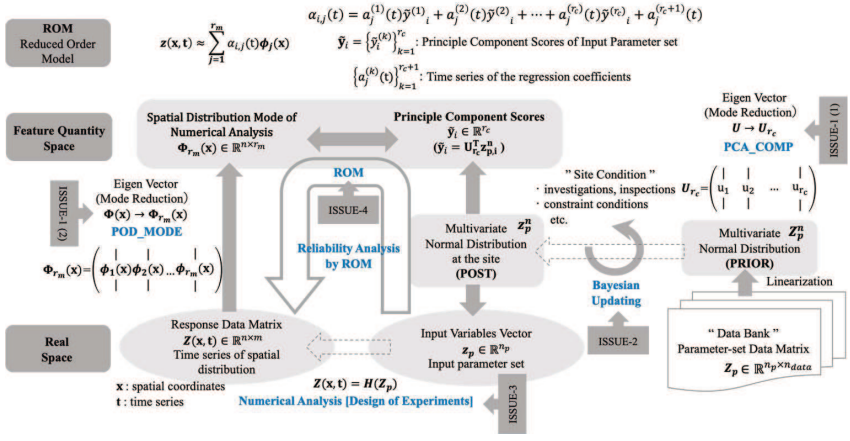


Figure 2. Validation scheme for the numerical analysis and research subjects proposed in this study.

3 Methodology

3.1 Proposed validation scheme

Figure 2 shows the validation scheme proposed in this study. The input parameters and FEM analysis responses are transformed to the eigenvalue space (feature space). The projection scores in the eigenvalue space of the input parameter and of the FEM analysis response are associated with each other to construct an alternative model that takes into account both time and space. Here the marginal distribution of the input parameters is converted to a normal distribution, thereby making the principal component analysis function effective. Then, a Bayesian update by adding observational information (laboratory physical test, liquefaction strength test, physical constraint, etc.) can be conducted immediately using the theoretical solution. Issue-1 and Issue-4 in Figure 2 show the formulations described in detail in the following sections.

3.2 ISSUE-1: Proper orthogonal decomposition (PCA, POD)

Proper orthogonal decomposition (POD) is the central method for dimension reduction or extraction of characteristics of a time series of data. POD is an expression mainly used in the field of fluid dynamics; however, it is also called principal component analysis (PCA) in the field of statistics. In this study, when this dimension reduction method is applied to the FEM analysis response space, it is considered as POD, and when applied to the parameter set space, it is considered as PCA. In addition, the eigenvectors of the FEM analysis response space are called "MODE" in POD and the eigenvectors in the parameter set space are called "COMP" in PCA. First, we denoted the theory of POD analysis using response snapshots of FEM analysis as an example.

A snapshot at time t of the distribution for the displacement or extra-pore water pressure calculated by the FEM analysis is defined as $\mathbf{z}(\mathbf{x}) \in \mathbb{R}^{\boldsymbol{n}}$, where $\boldsymbol{n}$ is the number of the grid of spatial location of FEM in POD, $\mathbf{x}$ is the planar 2D-coordinate matrix of the grid of spatial locations, $\boldsymbol{t}$ is the coordinate vector of time axis at the same interval, i.e., $\boldsymbol{\Delta t}$. In FEM analysis, $\boldsymbol{m}$ snapshots are output from time $\boldsymbol{t_1}$ to $\boldsymbol{t_m}$ at time intervals of $\boldsymbol{\Delta t}$. Then, the

relationship between time $\boldsymbol{t_k}$ and the next time, i.e., $\boldsymbol{t_{k+1}}$, is $\boldsymbol{t_{k+1}} = \boldsymbol{t_k} + \Delta\boldsymbol{t}$, and the snapshots at each time are synthesized into one data matrix, which is $\mathbf{Z} \in \mathbb{R}^{\boldsymbol{n\times m}}$.

$$\mathbf{Z} = \begin{bmatrix} | & | & | \\ \mathbf{z}(\mathbf{x}, t_1) & \mathbf{z}(\mathbf{x}, t_2) \cdots & \mathbf{z}(\mathbf{x}, t_m) \\ | & | & | \end{bmatrix} \tag{2}$$

For easy calculation of the covariance matrix of a data matrix, that is $\boldsymbol{Z}$, $\boldsymbol{Z}$ is transformed to a mean centering data matrix, $\mathbf{Z}^*$ The covariance matrix of $\mathbf{Z}$ may be denoted as $\mathbf{C_Z} = \mathbf{Z^*Z^{*T}} \in \mathbb{R}^{n\times n}$. Then, the eigenvalues of $\mathbf{C_Z}$ are denoted by $\{\lambda_1, \lambda_2, \cdots \lambda_n\}$ in descending order, and the corresponding eigenvectors are denoted by $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_\mathrm{n}\}$. Eigenvalue decomposition is performed on the covariance matrix, $\mathbf{C_Z} = \mathbf{U\Lambda U^T}$. Where, $\boldsymbol{\Lambda}$ is a diagonal matrix with $\{\lambda_1, \lambda_2, \cdots \lambda_n\}$ in diagonal elements and $\mathbf{U}$ is an orthogonal matrix with $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_\mathrm{n}\}$ in a column. Here, considering the projection $\mathbf{Y} = \mathbf{U^T Z^*} \in \mathbb{R}^{n\times m}$ to the eigenspace, $\mathbf{C_Y} \in \mathbb{R}^{n\times n}$ represents the variance of $\mathbf{Z}$ in the eigenspace.

$$\mathbf{C_Y} = \mathbf{YY^T} = \mathbf{U^T Z^* Z^{*T} U} = \boldsymbol{\Lambda} \tag{3}$$

We can easily capitalize on the projection score,$\widetilde{\mathbf{Y}}$, with a low-dimensional structure as shown below.

$$\widetilde{\mathbf{Y}} = \mathbf{U}_r^T \mathbf{Z}^* \ (r < n) \tag{4}$$

where $\mathbf{U}_r$ is a matrix having r eigenvectors $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_r\}$ in a column among $\mathbf{U}$ eigenvectors $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_\mathrm{n}\}$. Typically, r is chosen to be large enough such that the approximation is accurate and yet small enough such that $r << n$. When applying this theory to the input parameter set, it is sufficient to replace the snapshot of the numerical analysis result obtained for each time step with the input parameter set.

3.3 ISSUE-4: Regression analysis for deriving the ROM model with DE

A time series of data matrix containing displacement, extra-pore water pressure, etc., calculated via FEM at a grid of spatial locations is denoted as $\mathbf{Z_i}$. Here, $\boldsymbol{i}$ means the case number of FEM in the DE. Then, $\mathbf{Z_i^*}$ is a mean centering data matrix at each discretized FEM analysis time step. Here, the symbol of POD modes, $\mathbf{U} = [\mathbf{u_1} \quad \mathbf{u_2} \cdots \quad \mathbf{u}_{r_m}]$, is replaced by $\boldsymbol{\Phi} = [\boldsymbol{\phi_1} \quad \boldsymbol{\phi_2} \cdots \quad \boldsymbol{\phi}_{r_m}]$. $\boldsymbol{r_m}$ is the rank of the reduced-order ROM model.

Here, the FEM response spatial distribution vector, $\mathbf{z_i^*}(\mathbf{x}, \boldsymbol{t})$ at a certain time, $\boldsymbol{t}$, is defined. $\mathbf{x}$ is the planar coordinates of a grid of spatial locations in the FEM model, t is the coordinates on the time axis discretized by a time step, $\Delta\boldsymbol{t}$. In FEM analysis, because $\boldsymbol{m}$ snapshots are output from time $\boldsymbol{t_1}$ to $\boldsymbol{t_m}$ at time interval $\boldsymbol{\Delta t}$, $\mathbf{Z_i^*}$ is defined as follows:

$$\mathbf{Z_i^*} = \begin{bmatrix} | & | & | \\ \mathbf{z_i^*}(\mathbf{x}, t_1) & \mathbf{z_i^*}(\mathbf{x}, t_2) \cdots & \mathbf{z_i^*}(\mathbf{x}, t_m) \\ | & | & | \end{bmatrix} \tag{5}$$

$$\mathbf{z_i^*}(\mathbf{x}, t) \approx \alpha_{i,1}(t)\boldsymbol{\phi}_1(\mathbf{x}) + \alpha_{i,2}(t)\boldsymbol{\phi}_2(\mathbf{x}) + \cdots + \alpha_{i,r_m}(t)\boldsymbol{\phi}_{r_m}(\mathbf{x}) = \textstyle\sum_{j=1}^{r_m} \alpha_{i,j}(t)\boldsymbol{\phi}_j(\mathbf{x}) \tag{6}$$

where $\alpha_{i,j}(t)$ represents the expansion coefficients corresponding to the spatial displacement mode, $\boldsymbol{\phi}_j(\mathbf{x})$. Furthermore, $\alpha_{i,j}(t)$ is associated with the principle components of $\tilde{\mathbf{y}}_i = (\tilde{y}^{(1)}{}_i, \tilde{y}^{(2)}{}_i, \cdots \tilde{y}^{(r_c)}{}_i, 1)^T \in \mathbb{R}^{r_c+1}$ by principle component regression based on the linear regression model as provided below. r_c is the rank of reduced input variables (i.e., input parameter set for the FEM analysis).

$$\alpha_{i,j}(t) = a_j^{(1)}(t)\tilde{y}^{(1)}{}_i + a_j^{(2)}(t)\tilde{y}^{(2)}{}_i + \cdots + a_j^{(r_c)}(t)\tilde{y}^{(r_c)}{}_i + a_j^{(r_c+1)}(t) \tag{7}$$

where $\tilde{\mathbf{y}}_i$ is the projection score for the input variables approximated to the low-dimensional structure. Therefore, the POD expansion equation can be described as follows:

$$\mathbf{z_i^*}(\mathbf{x}, t) \approx \textstyle\sum_{j=1}^{r_m} \alpha_{i,j}(t)\boldsymbol{\phi}_j(\mathbf{x}) = \sum_{j=1}^{r_m} \left\{ a_j^{(1)}(t)\tilde{y}^{(1)}{}_i + a_j^{(2)}(t)\tilde{y}^{(2)}{}_i + \cdots + a_j^{(r_c)}(t)\tilde{y}^{(r_c)}{}_i + a_j^{(r_c+1)}(t) \right\} \boldsymbol{\phi}_j(\mathbf{x}) \tag{8}$$

Here, using $\left\{\tilde{y}^{(1)}{}_i, \tilde{y}^{(2)}{}_i, \cdots \tilde{y}^{(r_c)}{}_i\right\}$ and $\boldsymbol{\phi}_j(\mathbf{x})$, we define the following two matrices, $\boldsymbol{\phi}_{i,j}(\mathbf{x}) \in \mathbb{R}^{n\times(r_c+1)}$ and $\boldsymbol{\Phi}_i(\mathbf{x}) \in \mathbb{R}^{n\times\{(r_c+1)\times r_m\}}$ as follows:

$$\boldsymbol{\phi}_{i,j}(\mathbf{x}) = \begin{bmatrix} | & | & & | & | \\ \tilde{y}_{i,1}\boldsymbol{\phi}_j(\mathbf{x}) & \tilde{y}_{i,2}\boldsymbol{\phi}_j(\mathbf{x}) & \cdots & \tilde{y}_{i,r_c}\boldsymbol{\phi}_j(\mathbf{x}) & \boldsymbol{\phi}_j(\mathbf{x}) \\ | & | & & | & | \end{bmatrix};\ \boldsymbol{\Phi}_i(\mathbf{x}) = \begin{bmatrix} | & | & & | \\ \boldsymbol{\phi}_{i,1}(\mathbf{x}) & \boldsymbol{\phi}_{i,2}(\mathbf{x}) & & \boldsymbol{\phi}_{i,r_m}(\mathbf{x}) \\ | & | & & | \end{bmatrix} \tag{9}$$

In addition, we define the following vector, $\mathbf{a}_j^{(l)} \in \mathbb{R}^m$, and two matrices, $\mathbf{A}_j \in \mathbb{R}^{(r_c+1)\times m}$ and $\mathbf{A} \in \mathbb{R}^{\{(r_c+1)\times r_m\}\times m}$.

$$\mathbf{a}_j^{(l)} = \left(a_j^{(l)}(t_1), a_j^{(l)}(t_2), \cdots, a_j^{(l)}(t_m)\right)^T \tag{10}$$

$$\mathbf{A}_j = \begin{bmatrix} - & \mathbf{a}_j^{(1)T} & - \\ - & \mathbf{a}_j^{(2)T} & - \\ & \vdots & \\ - & \mathbf{a}_j^{(r_c)T} & - \\ - & \mathbf{a}_j^{(r_c+1)T} & - \end{bmatrix};\ \mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_{r_m} \end{bmatrix} \tag{11}$$

Using $\boldsymbol{\Phi}_i(\mathbf{x})$ and $\mathbf{A}$ defined above, $\mathbf{Z}_i^* \in \mathbb{R}^{n\times m}$ is written as follows:

$$\mathbf{Z}_i^* = \boldsymbol{\Phi}_i(\mathbf{x})\,\mathbf{A} \tag{12}$$

By the above preparation, the time series of regression coefficients can be easily obtained by the least squares method.

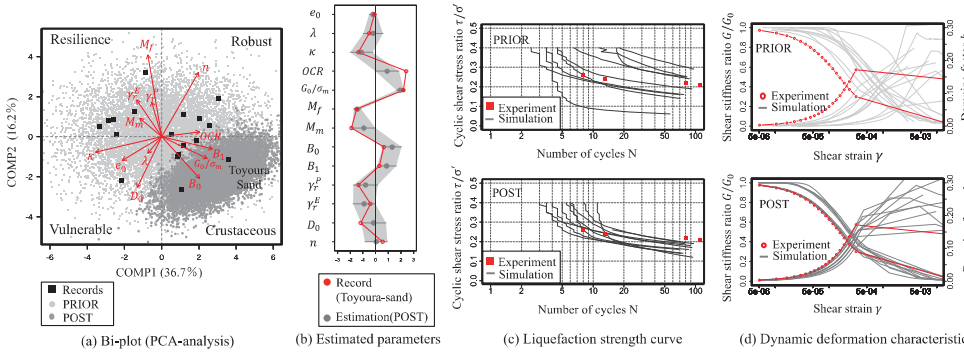


Figure 3. Bayesian inference results and element simulation results based on the known parameters of e_0, G_0, and M_f.

4 Results of the Analysis

4.1 ISSUE-1 (1); PCA analysis of the input parameter set (PCA-COMP)

The PCA results of the input parameter data matrix are shown in Figure 3. The accumulated contribution ratio until the third principle component is 70%. The arrows in the figure show the projection to each principle component (i.e., the contribution ratio to each principle component). The black plots in the figure represent the principle component score of the original data on 19 sites used for this analysis. Upon examining the meaning of each principle component based on the contribution of each parameter to the principle component, the first principle component was interpreted as "Comprehensive Rigidity," and the second principle component was interpreted as "Comprehensive Strength." Specifically, we made following assignments: (i) the first quadrant is "ROBUST," (ii) the second quadrant is "RESILIENCE," (iii) the third quadrant is "VULUNERABLE," and (iv) the forth quadrant is "CRUSTACEOUS." Because of space limitations, here we focus on the description until the second principal component. However, after examining the characteristics of each principal component from an engineering stand-point, we decided to perform mode reduction until the third principal component (i.e., r_m = 3). Data matrix **Z** is defined as follows:

$$\mathbf{Z} = \begin{bmatrix} | & | & & | \\ z_{p,1} & z_{p,2} & \cdots & z_{p,19} \\ | & | & & | \end{bmatrix} = [\mathbf{z_{p,1}} \quad \mathbf{z_{p,2}} \cdots \quad \mathbf{z_{p,19}}] \in \mathbb{R}^{n_p \times 19} \tag{13}$$

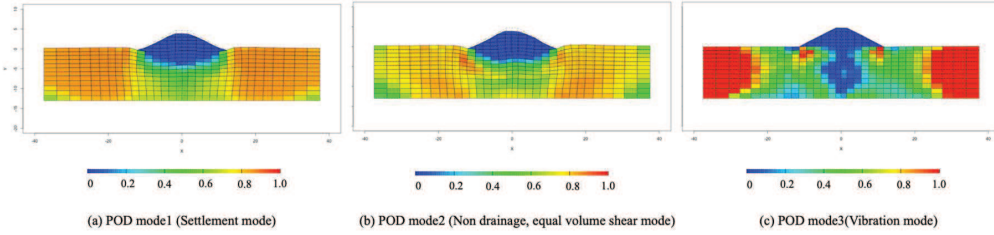
4.2 ISSUE-1 (2); POD analysis of the results (POD-MODE)

Figure 4 shows the 3-POD modes obtained by POD analysis. The accumulated contribution rate until the three modes is approximately 85%. In the fourth and subsequent modes, it is a mode that is difficult to interpret from an engineering stand-point, and mode reduction is conducted by using three modes. The first POD mode was interpreted as a "Settlement Mode," the second POD mode was interpreted as a "Non-drainage, equal volume shear mode," and the third mode as a "Vibration mode." Figure 4 shows the projection scores for each mode. Thus, the first POD mode is dominant. However, in the major vibration domain the non-drainage shear mode and vibration mode are also dominant. However, after the foundation ground has reached liquefaction, these modes tend to sharply dampen. Thus, POD mode decomposition is useful not only for constructing alternative models but also for understanding complex phenomena found in the numerical analysis results. Data Matrix **Z** is defined as follows:

$$\mathbf{Z} = \left[\mathbf{Z_1}(\mathbf{x}, \mathbf{t}) \; \mathbf{Z_2}(\mathbf{x}, \mathbf{t}) \; \ldots \; \mathbf{Z_{n_{cal}}}(\mathbf{x}, \mathbf{t})\right] \in \mathbb{R}^{n \times (m \times n_{cal})} \tag{14}$$

$$\mathbf{Z_i} = \begin{bmatrix} | & | & & | \\ \mathbf{z_i}\ (\mathbf{x}, t_1) & \mathbf{z_i}\ (\mathbf{x}, t_2) & \cdots & \mathbf{z_i}\ (\mathbf{x}, t_m) \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{n \times \mathrm{m}} \; ; \mathbf{z_i} = \begin{bmatrix} \boldsymbol{Dip_x} \\ \boldsymbol{Dip_y} \\ \boldsymbol{EPWP} \end{bmatrix} \in \mathbb{R}^n ; n = 3 \times n_{grid} \tag{15}$$

where $\boldsymbol{Dip_x}$, $\boldsymbol{Dip_y}$, and **EPWP** represent the horizontal displacements, the vertical displacements, and the extra pure water pressure of a grid of spatial locations, respectively; n_{cal} is the number of FEM analysis calculation cases; n_{grid} is the number of grids for FEM analysis.



(a) POD mode1 (Settlement mode) (b) POD mode2 (Non drainage, equal volume shear mode) (c) POD mode3(Vibration mode)

Figure 4. POD modes calculated via POD analysis. (L: POD mode1, Settlement mode, M: POD mode2, Non-drainage, equal volume shear mode, R: POD mode3, Vibration mode)

4.3 ISSUE-2; Bayesian inference based on observations

The parameters that were measured by the laboratory physical test of the 14 parameters used for the LIQCA analysis were e_0, G_0, and M_f. Therefore, assuming that these parameters are obtained, the Bayesian inference of other parameter sets was performed to confirm the validity of the Bayesian update. As an example, Figure 2 (c) shows the estimated parameters of Toyoura sand. The liquefaction strength curve and the dynamic deformation characteristics obtained from the element simulation (100 cases) via MCS with the updated parameters are shown in Figure 2 (c). The updated parameters are considered to be appropriate because the calculation results from the LICA simulation are roughly consistent with the experimental results.

4.4 ISSUE-4; Regression analysis connecting COMP and MODE

Figure 5(a) shows the comparison between FEM analysis and the proposed ROM analysis with a focus on the deformation of the crest of the embankment and the extra-pore water pressure under the embankment. The ROM analysis can be reproduced with a high accuracy when associated with the deformation distribution and excess pore water pressure distribution calculated via FEM. Figure 5(b)(c) shows a comparison diagram of FEM and ROM analyses for the time series and the final step of constituent coefficients of each mode. The coloration in the figure shows the magnitude of the excess pore water pressure. Figure 6 shows the accuracy of ROM simulation. This represents the calculation accuracy, which reflects the estimation accuracy of the input parameter set according to the type and content of the preliminary soils that were investigated.

In this study, we defined "Quantification of calculation accuracy = Validation." With the proposed frame-

work described above, the calculation accuracy of numerical analysis was effectively quantified. As a result, this analysis method may be practically used within the framework of reliability design.

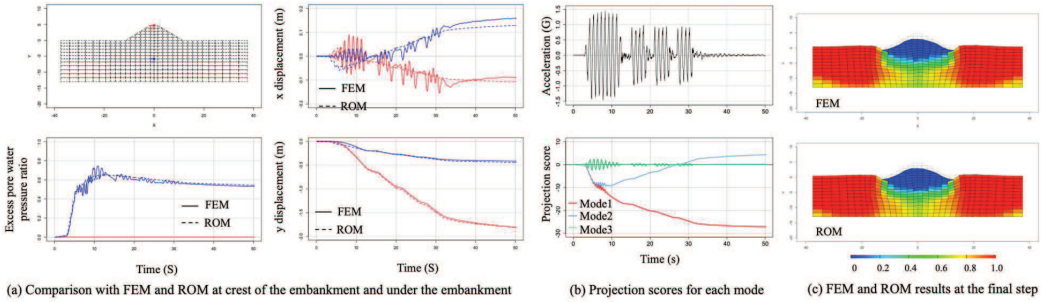


Figure 5. (a) Comparison between FEM analysis and ROM analysis results at the crest of the embankment and under the embankment (time series data of the settlement and excess pore water pressure ratio) (b) relationship between the input wave and the time series of projection scores and (c) the displacement special distribution with the excess pore water pressure ratio at the final step.

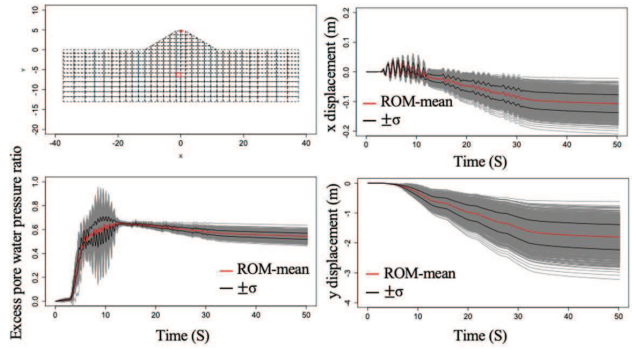


Figure 6. Calculation accuracy based on ROM analysis and PCA analysis in the parameter set space.

5 Conclusion

In this study, we proposed a validation scheme for numerical analysis of liquefaction-induced deformation of soil. We demonstrated a method to quantify the analysis accuracy of a sophisticated numerical analysis using the content of the survey by linking the mode reduction result of the numerical analysis with the regression analysis. The scheme proposed may be utilized in sophisticated numerical analysis for practical design in combination with the reliability-based design concept.

References

Oka F., Yashima A., Tateishi A., Taguchi Y., and Yamashita S. (1999). A cyclic elasto-plastic constitutive model for sand considering a plastic-strain dependence of the shear modulus. *Geotechnique,* 49(5), 661–680.

PWRI (Public Works Research Institute) (2000). *An Experiment Study on the Effects of Remedial Measure for Embankment Failure Due to Earthquake-Induced Foundation Liquefaction*. Technical Memorandum of PWRI, No.3688.