Proc. of the 9th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Zhongqiang Liu, Jian Dai and Kate Robinson
Published by Research Publishing, Singapore.
doi: 10.3850/GRF-25280825_isgsr-097-P396-cd

ISGSR2025

PROPOSAL FOR AIC IN A REDUCED-ORDER MODEL BASED ON PROPER ORTHOGONAL DECOMPOSITION

Yusuke Fukunaga[1], Naoki Sumioka[2], Yu Otake[3], Noriki Sugahara[4], and Masafumi Miyata[5]

[1] *Engineering Seismology Group, Earthquake Disaster Prevention Engineering Department, Port and Airport Research Institute (PARI), National Institute of Marine, Port and Aviation Technology (MPAT), Japan*
E-mail: fukunaga-y@p.mpat.go.jp
[2] *Structural Design Division, ECOH CORPORATION, Japan E-mail: sumioka@ecoh.co.jp*
[3] *AIS Lab. (Advanced Infrastructure Systems), Department of Civil Environmental Engineering, Tohoku University, Japan*
[4] *Big Data Technology Group, Infrastructure Digital Transformation Engineering Department, Port and Airport Research Institute (PARI), National Institute of Marine, Port and Aviation Technology (MPAT), Japan*
[5] *National Institute of Land and Infrastructure Management (NILIM), Ministry of Land, Infrastructure, Transport and Tourism (MLIT), Japan*

In this study, we propose an equation for the Akaike Information Criterion (AIC) as an index to quantitatively evaluate the generalization performance of a reduced-order model (ROM) based on proper orthogonal decomposition (POD). The ROM proposed by Otake et al. (2021) has primarily been validated using known input data, and the evaluation of its generalization performance for unknown input data has been an issue. In this study, we incorporate a linear Gaussian state-space model into the ROM by Otake et al. (2021) and present a method to quantitatively evaluate its generalization performance using AIC through parameter estimation with the Expectation-Maximization algorithm.

Keywords: Akaike Information Criterion (AIC), Proper Orthogonal Decomposition (POD), Reduced-Order Model (ROM), Lenear Gaussian State-Space Model (LGSSM).

1. Introduction

Proper orthogonal decomposition (POD), based on singular value decomposition (SVD), is a widely used dimensionality reduction technique for analyzing complex spatiotemporal systems governed by partial differential equations (PDEs) (e.g., Brunton et al., 2022). Typically, solution functions and their time derivatives are obtained numerically using methods such as the finite element method (FEM) under various initial and boundary conditions. SVD is then applied to the snapshot matrix composed of these solutions to extract the POD basis. A reduced-order model (ROM), often constructed using the POD-Galerkin expansion, expresses the system dynamics as a low-dimensional ordinary differential equation (ODE), which can be solved numerically.

POD was originally introduced in the context of turbulence analysis governed by the Navier–Stokes equations (Berkooz et al., 1993; Holmes et al., 2012), where the expansion was typically applied to the velocity field. However, applying POD to geotechnical materials presents additional challenges. For instance, saturated soils consist of two phases—soil skeleton and pore water—and their mechanical behavior is often modeled with incremental elastoplastic constitutive laws. This leads to more complex governing equations than those encountered in incompressible fluid mechanics.

To overcome these difficulties, Otake et al. (2021) proposed a regression-based ROM (hereinafter referred to as the Otake ROM), in which the POD basis is constructed from numerical outputs, and the expansion coefficients are regressed on principal components of input parameters such as the initial void ratio and hardening parameters. Their model successfully reproduced time-series responses, such as displacement and excess pore water pressure, with good agreement compared to FEM results. However, the generalization performance of the model under unseen input conditions remains unverified.

In this study, we incorporate a probabilistic structure into the Otake ROM by modeling the regression coefficients as latent variables in a linear Gaussian state-space model (LGSSM), where the observed variables correspond to the POD expansion coefficients. Based on this formulation, we derive an expression for the Akaike Information Criterion (AIC; Akaike, 1973), as an index for quantitatively evaluating the generalization performance of the resulting probabilistic model.

Here, we describe the notation used throughout this paper: $\mathbb{N}$ denotes the set of natural numbers. $\mathbb{R}$ denotes the set of real numbers, $\mathbb{R}_{>0}$ denotes the set of positive real numbers. $\mathbb{R}_{\geq 0}$ $\left(:= \mathbb{R}_{>0} \cup \{0\}\right)$ denotes the set of non-negative real numbers. I_J $\left(\in \mathbb{R}^{J\times J}\right)$ denotes J-dimensional identity matrix. $X_{j_1:j_2,k_1:k_2}$ $\left(\in \mathbb{R}^{(j_2-j_1+1)\times(k_2-k_1+1)}\right)$ denotes the submatrix of X $\left(\in \mathbb{R}^{J\times K}\right)$ consisting of rows j_1 to j_2 and columns k_1 to k_2. When $j_2 = j_1$ or $k_2 = k_1$, the submatrix is denoted as $X_{j_1,k_1:k_2}$ or $X_{j_1:j_2,k_1}$, respectively. X^{T} $\left(\in \mathbb{R}^{K\times J}\right)$ denotes the transpose of X. $A : B := \sum_{j=1}^{J}\sum_{k=1}^{K} A_{jk} B_{jk}$ $(\in \mathbb{R})$ denotes the Frobenius inner product of matrices A, B $\left(\in \mathbb{R}^{J\times K}\right)$.

$A \otimes B := \begin{bmatrix} A_{1,1} B & \cdots & A_{1,K} B \\ \vdots & \ddots & \vdots \\ A_{J,1} B & \cdots & A_{J,K} B \end{bmatrix}$ $\left(\in \mathbb{R}^{JL \times KS}\right)$ denotes the Kronecker product of matrices A $\left(\in \mathbb{R}^{J \times K}\right)$, B $\left(\in \mathbb{R}^{L \times S}\right)$.

$\mathbb{E}_{Y|X}\left[g(Y) \middle| X\right]$, $\mathrm{Cov}_{Y|X}\left[g(Y) \middle| X\right]$ $\left(:= \mathbb{E}_{Y|X}\left[\left(g(Y) - \mathbb{E}_{Y|X}\left[g(Y) \middle| X\right]\right)\left(g(Y) - \mathbb{E}_{Y|X}\left[g(Y) \middle| X\right]\right)^{\mathrm{T}} \middle| X\right]\right)$ denote the conditional expectation and the conditional variance-covariance matrix of the random variable $g(Y)$ given the random variable *X*, respectively. The indices of elements in a finite set of size J $(\in \mathbb{N})$ are denoted by the corresponding lowercase letter *j* (e.g., $j \in \{1, \cdots, J\}$) for the uppercase letter *J*. Furthermore, other notations will be explained as needed.

2. Proposal of AIC for the ROM Based on POD

2.1. *Overview of the Otake ROM*

To derive the AIC for the ROM, we begin by outlining the Otake ROM.

Finite element analyses (FEAs) are conducted for a total of *C* cases, and the output data for each case are obtained. A calculation case refers to a combination of input parameters such as the initial void ratio and soil constitutive model parameters, and is indexed by c $\left(\in \{0, \cdots, C-1\}\right)$. For each case, the input variables are mapped to a feature vector $\boldsymbol{\varphi}^{(c)}$ $\left(:= \boldsymbol{\varphi}\left(\boldsymbol{\xi}^{(c)}\right)\right)$ $\left(\in \mathbb{R}^{D}\right)$, based on the principal component scores, $\boldsymbol{\xi}^{(c)}$, which are obtained using another ground database.

The FEA output for each case consists of horizontal displacement u_x, vertical displacement u_y, and excess pore water pressure ratio *p*. At each time step, these are stacked row-wise to form a column vector. The time series of these vectors is arranged column-wise to construct the snapshot matrix, $Y^{(c)}$ $\left(\in \mathbb{R}^{N \times M}\right)$, which is centered in the time (column) direction:

$$\forall c \in \{0, \cdots, C-1\}\left(Y^{(c)} := \left[\left(Y_{u_x}^{\mathrm{orig},(c)}\right)^{\mathrm{T}} \quad \left(Y_{u_y}^{\mathrm{orig},(c)}\right)^{\mathrm{T}} \quad \left(Y_{p}^{\mathrm{orig},(c)}\right)^{\mathrm{T}}\right]^{\mathrm{T}} \left(I_M - M^{-1} \mathbf{1}_M \mathbf{1}_M^{\mathrm{T}}\right)\right) \tag{1}$$

where $Y_{\alpha}^{\mathrm{orig},(c)}$ denotes the snapshot matrix for the physical quantity α $\left(\in \{u_x, u_y, p\}\right)$ before centering in the column direction, and $\mathbf{1}_J := [1, \cdots, 1]^{\mathrm{T}}$ $\left(\in \mathbb{R}^{J}\right)$.

Among all *C* cases, a representative calculation case is designated as case 0, and the POD basis is constructed from the snapshot matrix $Y^{(0)}$ using SVD:

$$\Psi := U_{1:R}^{(0)} = Y^{(0)} V_{1:R}^{(0)} \left(\Gamma_R^{(0)}\right)^{-1} \tag{2}$$

where $Y^{(0)} = U^{(0)} \Sigma^{(0)} \left(V^{(0)}\right)^{\mathrm{T}}$, and $U^{(0)}$ $\left(\in \mathbb{R}^{N \times N}\right)$, $V^{(0)}$ $\left(\in \mathbb{R}^{M \times M}\right)$, and $\Sigma^{(0)}$ $\left(\in \mathbb{R}_{\geq 0}^{N \times M}\right)$ denote the left singular vector matrix, right singular vector matrix, and singular value matrix of $Y^{(0)}$, respectively. Additionally, $\Gamma_R^{(0)}$ $\left(:= \Sigma_{1:R,1:R}^{(0)}\right)$ $\left(\in \mathbb{R}_{>0}^{R \times R}\right)$ denotes the leading principal submatrix of order *R* of $\Sigma^{(0)}$.

The snapshot matrix is projected onto the POD basis to obtain the expansion coefficients, A $\left(:= \Psi^{\mathrm{T}} Y\right)$ $\left(\in \mathbb{R}^{R \times M}\right)$. Each row, $A_{r,1:M}$ $\left(\in \mathbb{R}^{1 \times M}\right)$, contains the POD expansion coefficients across all snapshots for mode r $\left(\in \{1, \cdots, R\}\right)$, and is regressed on the input feature vectors. For each mode *r*, the regression model is:

$$\left[\forall r \in \{1, \cdots, R\}\left(A_{r,1:M} = \boldsymbol{\varphi}^{\mathrm{T}} B_r + E_{r,1:M}\right)\right] \Leftrightarrow A = \Phi B + E \tag{3}$$

where $B := \left[B_1^{\mathrm{T}} \quad \cdots \quad B_R^{\mathrm{T}}\right]^{\mathrm{T}}$ $\left(\in \mathbb{R}^{RD \times M}\right)$ denotes the regression coefficient parameter, $E := \left[E_{1,1:M}^{\mathrm{T}} \quad \cdots \quad E_{R,1:M}^{\mathrm{T}}\right]^{\mathrm{T}}$ $\left(\in \mathbb{R}^{R \times M}\right)$ denotes the error term, and $\Phi := I_R \otimes \boldsymbol{\varphi}^{\mathrm{T}}$ $\left(\in \mathbb{R}^{R \times RD}\right)$.

The regression coefficients are estimated using ordinary least squares (OLS). For a new input feature vector, $\boldsymbol{\varphi}^{*}$ $\left(\in \mathbb{R}^{D}\right)$, the expansion coefficients are predicted as:

$$\hat{A} = \Phi^{*} \hat{B}^{\mathrm{OLS}} = \left[I_R \otimes \left(\boldsymbol{\varphi}^{*}\right)^{\mathrm{T}}\right] \hat{B}^{\mathrm{OLS}} \tag{4}$$

where $\hat{A}$ $\left(\in \mathbb{R}^{R\times M}\right)$ denotes the predicted value of A, $\left\{\left(\boldsymbol{\varphi}^{(c)},\, A^{(c)}\right)\right\}_{c=0}^{C-1}$ denotes the training dataset for the regression model where each pair of the feature vector, $\boldsymbol{\varphi}^{(c)}$ $\left(\in \mathbb{R}^{D}\right)$, obtained as the principal component scores of the input variables and the POD expansion coefficients, $A^{(c)}$ $\left(\in \mathbb{R}^{R\times M}\right)$, of the snapshot matrix,

$\hat{B}^{\mathrm{OLS}} := \underset{B\in\mathbb{R}^{RD\times M}}{\operatorname{argmin}} \sum_{c=0}^{C-1}\left\|E^{(c)}\right\|_{\mathrm{F}}^{2} = \left[\sum_{c=0}^{C-1}\left(\Phi^{(c)}\right)^{\mathrm{T}}\Phi^{(c)}\right]^{-1}\left[\sum_{c'=0}^{C-1}\left(\Phi^{(c')}\right)^{\mathrm{T}}A^{(c')}\right] = \left\{\sum_{c=0}^{C-1} I_R \otimes \left[\boldsymbol{\varphi}^{(c)}\left(\boldsymbol{\varphi}^{(c)}\right)^{\mathrm{T}}\right]\right\}^{-1}\left[\sum_{c'=0}^{C-1}\left(I_R\otimes\boldsymbol{\varphi}^{(c')}\right)A^{(c')}\right]$,

$\Phi^{*} := I_R \otimes \left(\boldsymbol{\varphi}^{*}\right)^{\mathrm{T}}$ $\left(\in \mathbb{R}^{R\times RD}\right)$, and $\left\|X\right\|_{\mathrm{F}} := \sqrt{X : X}$ $\left(\in \mathbb{R}_{\geq 0}\right)$ denotes the Frobenius norm of the matrix of X.

2.2. *Modification of the Otake ROM to Incorporate a Probabilistic Structure*

When the snapshot matrix Y is constructed as described above, that is, without applying spatiotemporal delay embedding (e.g., Brunton et al., 2022) in its creation, the SVD does not take into account any autocorrelation that may exist among the components in the row and column directions of Y. This is because, in SVD, the set of left singular vectors (i.e., the column vectors of U) is determined such that any two distinct row vectors of the projection components of Y $\left(= U\,\Sigma\,V^{\mathrm{T}}\right)$ onto the image of U are orthogonal (i.e., their standard inner product is zero). Similarly, the set of right singular vectors (i.e., the column vectors of V) is determined such that any two distinct column vectors of the projection components of Y onto the row space of V are orthogonal. However, since Y has been centered in the column direction as described above, any two distinct row vectors of the projection components of Y onto the image of Ψ, denoted as A $\left(= \Psi^{\mathrm{T}}\,Y\right)$, are uncorrelated. This does not imply that there is no spatial autocorrelation among the row vectors of A.

In this study, for the sake of simplicity, the consideration of spatial autocorrelation is deferred as a subject for future work. Instead, the regression model in Eq. (3) is modified to account for temporal autocorrelation, and the corresponding AIC is derived.

In Eq. (3), the m-th column of B, denoted as $\boldsymbol{\beta}_m$ $\left(:= B_{1:RD,m}\right)$ $\left(\in \mathbb{R}^{RD}\right)$, corresponds to the regression coefficient parameter at time step m. In other words, since the regression coefficient parameters vary with each time step, the model is formulated as a time-varying coefficient regression model. To account for the autocorrelation of the column vectors of B, B is regarded as a state variable and formulated using a linear Gaussian state-space model (LGSSM) (e.g., Bishop, 2006) as follows. Note that G, P, and Q are assumed to be time-invariant unknown parameters. The LGSSM is adopted because it is the most standard type of state-space model (SSM). Hereafter, the model in which the LGSSM is incorporated into the Otake ROM is referred to as the modified Otake ROM.

$$\left.\begin{aligned} \boldsymbol{\beta}_m &= G\,\boldsymbol{\beta}_{m-1} + \boldsymbol{w}_m & \boldsymbol{w}_m \mid \boldsymbol{\beta}_{m-1} &\sim \mathcal{N}\left(\mathbf{0}_{RD}, Q\right) \\ A^{(c)}_{1:R,m} &= \Phi^{(c)}\,\boldsymbol{\beta}_m + \boldsymbol{v}^{(c)}_m & \boldsymbol{v}^{(c)}_m \mid \boldsymbol{\beta}_m &\sim \mathcal{N}\left(\mathbf{0}_R, P\right) \quad \left(\forall c \in \{0, \cdots, C-1\}\right) \end{aligned}\right\} \quad \left(\forall m \in \{1, \cdots, M\}\right) \tag{5}$$

where G $\left(\in \mathbb{R}^{RD\times RD}\right)$ denotes the state transition matrix, P $\left(\in \mathbb{R}^{R\times R}\right)$ and Q $\left(\in \mathbb{R}^{RD\times RD}\right)$ denote the variance-covariance matrices of the system noise and observation noise, respectively, and $\mathbf{0}_J := [0, \cdots, 0]^{\mathrm{T}}$ $\left(\in \mathbb{R}^{J}\right)$.

The lower equation in Eq. (5) can be expressed as follows by stacking in the row direction with respect to c.

$$\tilde{\boldsymbol{a}}_m = \tilde{\Phi}\,\boldsymbol{\beta}_m + \boldsymbol{v}_m, \quad \boldsymbol{v}_m \mid \boldsymbol{\beta}_m \sim \mathcal{N}\left(\mathbf{0}_{CR}, I_C \otimes P\right) \quad \left(\forall m \in \{1, \cdots, M\}\right) \tag{6}$$

where $\tilde{\boldsymbol{a}}_m := \left[\left(A^{(0)}_{1:R,m}\right)^{\mathrm{T}} \;\cdots\; \left(A^{(C-1)}_{1:R,m}\right)^{\mathrm{T}}\right]^{\mathrm{T}}$ $\left(\in \mathbb{R}^{CR}\right)$ and $\tilde{\Phi} := \left[\left(\Phi^{(0)}\right)^{\mathrm{T}} \;\cdots\; \left(\Phi^{(C-1)}\right)^{\mathrm{T}}\right]^{\mathrm{T}}$ $\left(\in \mathbb{R}^{CR\times RD}\right)$.

2.3. *Derivation of the AIC for the Modified Otake ROM*

When the explanatory variable Φ in Eq. (3), by extension $\tilde{\Phi}$ in Eq. (6) are treated as a non-random variables, the log-likelihood of the modified Otake ROM is expressed as follows.

$$\ell(\theta) := \ln\prod_{m=1}^{M} f\left(\tilde{\boldsymbol{a}}_m \mid \tilde{\boldsymbol{a}}_{1:m-1}; \theta\right) = -\frac{1}{2}\left[M\,C\,R\ln(2\pi) + \sum_{m=1}^{M}\ln\left|\Sigma_m\right| + \sum_{m=1}^{M}\left(\tilde{\boldsymbol{a}}_m - \tilde{\Phi}\,\boldsymbol{\beta}_{m|m-1}\right)^{\mathrm{T}}\Sigma_m^{-1}\left(\tilde{\boldsymbol{a}}_m - \tilde{\Phi}\,\boldsymbol{\beta}_{m|m-1}\right)\right] \tag{7}$$

where $S_{j|k} := \operatorname{Cov}_{\boldsymbol{\beta}_j \mid \tilde{\boldsymbol{a}}_{1:k}}\left[\boldsymbol{\beta}_j \mid \tilde{\boldsymbol{a}}_{1:k}\right]$ $\left(\in \mathbb{R}^{RD\times RD}\right)$, $\boldsymbol{\beta}_{j|k} := \mathbb{E}_{\boldsymbol{\beta}_j \mid \tilde{\boldsymbol{a}}_{1:k}}\left[\boldsymbol{\beta}_j \mid \tilde{\boldsymbol{a}}_{1:k}\right]$ $\left(\in \mathbb{R}^{RD}\right)$, $\Sigma_m := I_C \otimes P + \tilde{\Phi}\,S_{m|m-1}\,\tilde{\Phi}^{\mathrm{T}}$ $\left(\in \mathbb{R}^{CR\times CR}\right)$, and $\theta := \{G, P, Q\}$. Since there is no prior information on the initial state, it is assumed that $S_{1|0} = \kappa I_{RD}$ (where κ is a sufficiently large positive real number) and $\boldsymbol{\beta}_{1|0} = \mathbf{0}_{RD}$.

Based on this log-likelihood, parameter estimation is performed using maximum likelihood (ML). Since the SSM includes latent variables (in the modified Otake ROM, B), each parameter is updated using the Expectation-Maximization (EM) algorithm. In the E-step and M-step, the Kalman filter and the maximum likelihood method are used, respectively.

2.3.1. *E-step (Kalman Filter)*

$$\text{prediction step: } \begin{cases} \boldsymbol{\beta}_{m|m-1} = G\,\boldsymbol{\beta}_{m-1|m-1} \\ S_{m|m-1} = Q + G\,S_{m-1|m-1}\,G^{\mathrm{T}} \end{cases}, \quad \text{update step: } \begin{cases} K_m = S_{m|m-1}\tilde{\boldsymbol{\Phi}}^{\mathrm{T}}\,\Sigma_m^{-1}; \text{Kalman gain} \\ \boldsymbol{\beta}_{m|m} = \boldsymbol{\beta}_{m|m-1} + K_m\left(\tilde{\boldsymbol{a}}_m - \tilde{\boldsymbol{\Phi}}\,\boldsymbol{\beta}_{m|m-1}\right) \\ S_{m|m} = \left(I_{RD} - K_m\tilde{\boldsymbol{\Phi}}\right)S_{m|m-1} \end{cases} \tag{8}$$

2.3.2. *M-step (ML estimation)*

The Q function (note that this is not the same as the system noise variance-covariance matrix denoted by the same symbol) of EM algorithm is expressed as follows.

$$\begin{aligned} Q\left(\theta,\theta^{\mathrm{old}}\right) &:= \mathbb{E}_{\boldsymbol{\beta}_{1:M}|\tilde{\boldsymbol{a}}_{1:M}}\left[\ln f\left(\tilde{\boldsymbol{a}}_{1:M},\boldsymbol{\beta}_{1:M};\theta\right)\middle|\tilde{\boldsymbol{a}}_{1:M};\theta^{\mathrm{old}}\right] \\ &= -\frac{1}{2}\left[M\,C\,R\ln(2\pi) + M\ln\left|I_C\otimes P\right| + \sum_{m=1}^{M}\left(I_C\otimes P\right)^{-1} : \mathbb{E}_{\boldsymbol{\beta}_{1:M}|\tilde{\boldsymbol{a}}_{1:M}}\left[\left(\tilde{\boldsymbol{a}}_m - \tilde{\boldsymbol{\Phi}}\,\boldsymbol{\beta}_m\right)\left(\tilde{\boldsymbol{a}}_m - \tilde{\boldsymbol{\Phi}}\,\boldsymbol{\beta}_m\right)^{\mathrm{T}}\middle|\tilde{\boldsymbol{a}}_{1:M};\theta^{\mathrm{old}}\right]\right] \\ &\quad -\frac{1}{2}\left[(M-1)\,R\,D\ln(2\pi) + (M-1)\ln|Q| + \sum_{m=2}^{M}Q^{-1} : \mathbb{E}_{\boldsymbol{\beta}_{1:M}|\tilde{\boldsymbol{a}}_{1:M}}\left[\left(\boldsymbol{\beta}_m - G\,\boldsymbol{\beta}_{m-1}\right)\left(\boldsymbol{\beta}_m - G\,\boldsymbol{\beta}_{m-1}\right)^{\mathrm{T}}\middle|\tilde{\boldsymbol{a}}_{1:M};\theta^{\mathrm{old}}\right]\right] \end{aligned} \tag{9}$$

By taking the partial derivative of this with respect to θ and applying the first-order optimality condition, the optimal solution of θ, denoted as $\theta^* = \{G^*, P^*, Q^*\}$, is obtained as follows. In other words, θ is estimated using the maximum likelihood method.

$$\left.\begin{aligned} G^* &= \left(\sum_{m=2}^{M}\mathbb{E}_{\boldsymbol{\beta}_{m-1:m}|\tilde{\boldsymbol{a}}_{1:M}}\left[\boldsymbol{\beta}_m\,\boldsymbol{\beta}_{m-1}^{\mathrm{T}}\,|\,\tilde{\boldsymbol{a}}_{1:M}\right]\right)\left(\sum_{m=2}^{M}\mathbb{E}_{\boldsymbol{\beta}_{m-1}|\tilde{\boldsymbol{a}}_{1:M}}\left[\boldsymbol{\beta}_{m-1}\,\boldsymbol{\beta}_{m-1}^{\mathrm{T}}\,|\,\tilde{\boldsymbol{a}}_{1:M}\right]\right)^{-1} \\ Q^* &= (M-1)^{-1}\sum_{m=2}^{M}\mathbb{E}_{\boldsymbol{\beta}_{m-1:m}|\tilde{\boldsymbol{a}}_{1:M}}\left[\left(\boldsymbol{\beta}_m - G^*\,\boldsymbol{\beta}_{m-1}\right)\left(\boldsymbol{\beta}_m - G^*\,\boldsymbol{\beta}_{m-1}\right)^{\mathrm{T}}\middle|\tilde{\boldsymbol{a}}_{1:M}\right] \\ P^* &= (C\,M)^{-1}\sum_{m=1}^{M}\sum_{c=0}^{C-1}\mathbb{E}_{\boldsymbol{\beta}_m|\tilde{\boldsymbol{a}}_{1:M}}\left[\left(\boldsymbol{a}_m^{(c)} - \boldsymbol{\Phi}^{(c)}\,\boldsymbol{\beta}_m\right)\left(\boldsymbol{a}_m^{(c)} - \boldsymbol{\Phi}^{(c)}\,\boldsymbol{\beta}_m\right)^{\mathrm{T}}\middle|\tilde{\boldsymbol{a}}_{1:M}\right] \end{aligned}\right\} \tag{10}$$

Hereafter, following the standard procedure of the EM algorithm, the E-step and M-step are repeated until the $\ell(\theta)$ converges, and the converged value of θ^* is obtained.

The number of parameters in this model is $(RD)^2$ for G $\left(\in\mathbb{R}^{RD\times RD}\right)$, $RD(RD+1)/2$ and $R(R+1)/2$ for Q $\left(\in\mathbb{R}^{RD\times RD}\right)$ and P $\left(\in\mathbb{R}^{R\times R}\right)$, respectively, since Q and P are symmetric matrices. Therefore, the AIC of the modified Otake ROM is expressed as follows.

$$\mathrm{AIC} = -2\,\ell\left(\theta^*\right) + 2\left[(RD)^2 + \frac{RD(RD+1)}{2} + \frac{R(R+1)}{2}\right] \tag{11}$$

3. Conclusion and Future Work

In this study, we proposed the AIC, as shown in Eq. (11), as an indicator for quantitatively evaluating the generalization performance of the modified Otake ROM, which incorporates a probabilistic structure into the original Otake ROM. In constructing the modified Otake ROM, we assumed a LGSSM. As future work, it is necessary to verify the generalization performance of this model using real data. If the performance is found to be inadequate, the model should be extended to other types of SSMs, such as nonlinear and non-Gaussian state-space models. In addition, since the explanatory variables are assumed to be non-random variables in this study, it is necessary to consider modeling them as random variables. Furthermore, as spatial autocorrelation was not considered in the modeling in this study, it is also necessary to explore modeling that takes it into account.

References

H. Akaike (1973). Information theory and an extension of the maximum likelihood principle. In International Symposium on Information Theory, Ed. B. N. Petrov and F. Csaki, pp. 267–81. Budapest: Akademia Kiado.

C. M. Bishop (2006). Pattern Recognition and Machine Learning, Springer.

G. Berkooz, P. Holmes, and J. L. Lumley (1993). The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows, Annu. Rev. Fluid Mech. 1993.25: 539-75.

P. Holmes, J. L. Lumley, G. Berkooz, and C. W. Rowley (2012). Turbulence, Coherent Structures, Dynamical Systems and Symmetry (Second Edition), Cambridge University Press.

S. L. Brunton, J. N. Kutz (2022). Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control (Second Edition), Cambridge University Press.

Y. Otake, K. Shigeno, Y. Higo, and S. Muramatsu (2021). Practical dynamic reliability analysis with spatiotemporal features in geotechnical engineering, Georisk Assessment and Management of Risk for Engineered Systems and Geohazards 16(7):1-16.