Quantification of Modeling Approximation Error of Pin-Cell Calculation Using Kriging and Principal Component Analysis

Tomomi Hanai,[1] Akio Yamamoto,[1] Tomohiro Endo,[1] Kento Yamamoto,[2] Yasunori Ohoka,[2] Hiroaki Nagano[2]

1) *Nagoya University: Furo-Cho, Chikusa-Ku Nagoya, Japan, 464-8603,*
t-hanai@fermi.nucl.nagoya-u.ac.jp, a-yamamoto@nucl.nagoya-u.ac.jp, t-endo@nucl.nagoya-u.ac.jp
2) *Nuclear Fuel Industries, Ltd.: 950-1, Asashironishi, kumatori-Cho, Sennan-Gun, Osaka, Japan, 590-0481*
ken-yamamoto@nfi.co.jp, ya-ohoka@nfi.co.jp, nagano@nfi.co.jp

INTRODUCTION

In this summary, the uncertainty of neutronics characteristics parameters due to the numerical modeling is estimated using a multivariate analysis with Kriging and the principal component analysis.

In core analysis, behavior of neutrons in a reactor is estimated by numerical simulations and safety of a reactor is confirmed by simulations. However, the results of core analysis inevitably contain uncertainties originating from a numerical simulation. The uncertainties in core analysis are caused by approximations in the numerical modeling (*i.e.*, spatial homogenization, discretization of neutron flight direction and energy) and the uncertainties of input data (*i.e.*, manufacturing tolerance, nuclear data, operating conditions, measurement data). Quantification of these uncertainties is desirable from the viewpoint safety and efficiency of reactor operation.

Uncertainty quantification of the neutronics characteristics parameters caused by the cross-section covariance has been extensively carried out in recent years [1, 2]. Meanwhile, the uncertainty quantification caused by the approximations in numerical modeling is ongoing but limited [3]. In this study, the uncertainty quantification of modeling approximation is discussed.

In order to estimate the modeling approximation error due to numerical modeling, comparison with high-fidelity results obtained by a detailed core analysis code (e.g. a continuous energy Monte-Carlo code) is necessary. However, the detailed modeling calculation takes longer computation time. Number of state-points of a reactor during an operating cycle is enormous, thus exhaustive estimation for all state-points with the detailed calculation is impractical. Namely, reducing calculation cost and increasing coverage of the analysis conditions are necessary for the quantification of uncertainty due to the numerical modeling approximation.

In the previous study [4], a surrogate model [5] is developed with Kriging to estimate the uncertainty due to numerical modeling approximation. The neutronics characteristic parameters obtained from a simplified calculation (i.e. cross sections and neutron fluxes) are used as the input variables of the surrogate model to estimate modeling approximation error. The surrogate model interpolates the modeling approximation error in typical core conditions. Thus the modeling approximation error for arbitrary calculation condition can be estimated from the surrogate model. The previous result reveals that calculation error (defined as the difference between a continuous energy Monte Carlo calculation and a deterministic lattice physics calculation) in the UO_2 pin cell geometry can be predicted by the surrogate model. However, in this method, there is ambiguity in the choice of input variables of the surrogate model. Since there are many neutronics characteristics parameters that can be used for input variables, optimum selection of input parameters for the surrogate model is very difficult. This issue implies that direct application of this approach would be difficult for practical and more complicated problems.

In order to address this issue, the principal component analysis [6] (PCA) is applied in order to reduce dimension of input variables of the surrogate model. The surrogate model for estimation of modeling approximation error is constructed though multivariate analyses with Kriging [7] and the principal component analysis.

In the next section, a brief description on Kriging is provided. Then, application of the principal component analysis for the dimensional reduction of input data is explained. Finally, the verification method of the developed surrogate model and verification results are given. In the verification calculation, the difference of k-effective between a deterministic lattice physics calculation and a continuous energy Monte Carlo calculation is tried to quantify by the surrogate model, which is created using Kriging and the principal component analysis.

KRIGING METHOD

Kriging is a nonparametric interpolation method, different from a parametric interpolation method such as the polynomial spline that utilizes predetermined functions. In the Kriging, the predicted value is modeled by the Gaussian process having spatial covariance.

Kriging performs interpolation in a D-dimensional space by learning N training data consisted of D-dimensional input variables and teach signal (modeling approximation error in this study). The interpolated (estimated) value can be obtained by the weighted sum of the measured values:

$$\hat{y}(\boldsymbol{x}_0) = \sum_{i=1}^{N} w_i y(\boldsymbol{x}_i), \qquad (1)$$

where $y(\boldsymbol{x}_i)$ is the modeling approximation error at a state-point condition $\boldsymbol{x}_i$, which is a vector of the input variables for the surrogate model given by $[x_{i1}, x_{i2}, \cdots, x_{iD}]$; x_{id} is a d-th input variable ($d = 1, \cdots, D$); $\hat{y}(\boldsymbol{x}_0)$ is an interpolated (estimated) value of the calculation error at an unmeasured state-point condition $\boldsymbol{x}_0$; w_i is a weight for $y(\boldsymbol{x}_i)$ and N = the number of training data.

The value of w_i is determined so as to minimize the variance of $\hat{y}(\boldsymbol{x}_0)$,: the detail is shown in Ref. [4].

PRINCIPAL COMPONENT ANALYSIS

The principal component score (PC score) obtained from PCA in order to reduce the dimension of input variables is used for the Kriging model. PCA finds the "major" axes of a data set that is consisted of correlated components. The "major" axes represent characteristics of components distribution in the data set. The major axes for the components are called as the principal components. The PC score represents component of each data along the major axis.

In Kriging, no correlation is assumed among input variables. However, neutronics characteristics parameters, which are used for input data for Kriging, are usually highly correlated. By applying PCA to a data set, components in the data set can be transformed into uncorrelated ones using the principal components that are orthogonal.

Then, the estimation method of PC score from neutronics characteristics parameters is described below.

PCA needs a sample covariance matrix **S** obtained from the input data of Kriging (training data for Kriging), which consists of N samples of D-dimensional neutronics characteristics parameters. (d, d')th element of the matrix **S** is the covariance between the d-th and d'-th elements of $\boldsymbol{x}$ $(d \neq d')$, or the variance of the d-th element of $\boldsymbol{x}$ $(d = d')$. Using the data set **X** of the D × N (N > D) matrix composed of D-dimensional N data, **S** can be obtained as follows:

$$\mathbf{S} = \mathrm{var}(\mathbf{X}) = \frac{1}{N-1}\mathbf{X}^{\mathbf{T}}\mathbf{X}\,. \qquad (2)$$

Note that **X** is a D × N (N > D) matrix composed of N-set D-dimensional data. By the eigenvalue decomposition of **S**, D sets of eigenvalues and eigenvectors are obtained. d-th PC score z_{id} in the i-th sample is calculated as follows:

$$z_{id} = \boldsymbol{x}_i \boldsymbol{\alpha}_d, \qquad (3)$$

where $\boldsymbol{\alpha}_d$ is an eigenvector of **S** corresponding to its d-th largest eigenvalue λ_d. Furthermore, if $\boldsymbol{\alpha}_d$ is chosen to have unit length ($\boldsymbol{\alpha}_d^T\boldsymbol{\alpha}_d = 1$), then λ_d means the variance of the data of d -th principal components.

In this study, part of the principal component scores obtained from Eq.(3) is used as the input parameters for the Kriging model. A PC score correspond to a larger eigenvalue has larger importance, thus it can express the larger amount of data set. From these eigenvalues, the cumulative percentage of total variation [6], which is an index for selecting the number of PC scores, can be obtained. The index means the percentage of variation accounted by the first k-th PC score, defined as follows:

$$t_k = 100\sum_{d=1}^{k}\lambda_d \Big/ \sum_{d=1}^{D}\lambda_d \ [\%]. \qquad (4)$$

Using the obtained PC score for inputs variables of Kriging, the surrogate model for the estimation of modeling approximation error is constructed.

Note that before the PC score estimation, each neutronics characteristic parameter is normalized so that the average and standard deviation are 0 and 1.

LEAVE-ONE-OUT CROSS-VALIDATION

A good surrogate model satisfies two conditions. Firstly, the predicted value of the model becomes closer to the true value than non-corrected value in the untrained condition. Secondly, the surrogate model provides adequate uncertainty on the predicted value and true value should be within the range of uncertainty. In order to verify these points, the leave-one-out cross-validation (LOOCV) [7] was carried out for the constructed surrogate model. In this verification, one sample of the data set is extracted and defined as the verification reference data, and the rests are used as training data to construct the surrogate model. These model training and verification are repeated for all sample.

As an index of the prediction accuracy, the maximum prediction error is used, which is defined as follows:

$$\Delta \mathrm{y}_{\max} = \max_{i}|y(\boldsymbol{x}_i) - \hat{y}(\boldsymbol{x}_i)| \quad (i = 1, \cdots, \mathrm{N}). \qquad (5)$$

The smaller index indicates higher prediction accuracy of the surrogate model.

Fidelity of the surrogate model is evaluated by 95% coverage probability. In this summary, 95% coverage probability is defined as the probability that the 95% confidence interval contains the true value of interest. The 95% confidence interval is theoretically given by:

$$a = 1.96\sqrt{\sigma_u^2 + \sigma_K^2}, \qquad (6)$$

where σ_K is uncertainty on the predicted value and σ_u corresponds to the statistical uncertainty of a continuous-energy Monte-Carlo code in this summary.

When the constructed surrogate model is appropriate, the probability of an actual modeling approximation error within the range of $\pm 1.96\sigma$ should be comparable to 95%. Thus, adequacy of the surrogate model is measured by the above probability.

CALCULATION GEOMETRY AND CONDITIONS

As a target of modeling approximation error, the relative difference of k-effective δ_k (unit: pcm = 10^{-3} %) obtained by 1) a deterministic lattice physics code and 2) a continuous-energy Monte Carlo code is used. δ_k is obtained as follows:

$$\delta_k = \frac{k_d - k_m}{k_m} \times 10^5 \ [\mathrm{pcm}], \qquad (7)$$

where k_d is k-effective from a deterministic lattice physics code and k_m is k-effective from a Monte Carlo code.

The calculation is carried out in a pin cell geometry consisting of the MOX fuel and moderator regions as shown in Fig. 1.

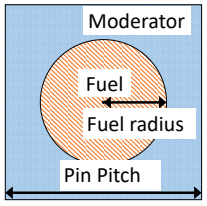

Fig. 1. Example of calculation geometry (reflective boundary condition).

In this summary 8 parameters shown in Table I and Pu composition in Table II are considered. Except for Pu-239 composition, each parameter is assumed to be uniformly distributed within the range shown in Tables I and II without correlation.

Table I. Variation of fuel cell specifications

Core parameters	Max.	Min.
Fuel radius [cm]	0.5000	0.3500
Pin pitch [cm]	1.5000	1.2000
Temperature of fuel [K]	1000	300
Temperature of moderator [K]	600	300
Concentration of boron [ppm]	2400.0	100.0
Pu concentration [wt%]	13.000	1.250
Enrichment of U-235 [wt%]	0.400	0.200
Fuel density [g/cm^3]	10.950	10.000

Table II. Variation of Pu composition (unit: wt%)

Nuclide	Max.	Min.
Pu-238	2.50	0.03
Pu-239	*	
Pu-240	30.00	7.00
Pu-241	14.00	0.70
Pu-242	7.00	0.00
Am-241	3.00	0.00

* Pu-239 composition is adjusted so that sum of all Pu composition is equal to 100 wt%

In the uncertainty evaluation, the CASMO4 and MVP are used as the deterministic lattice physics and the Monte Carlo codes, respectively [8, 9]. The construction and verification procedures of the surrogate model for the difference of k-effective are summarized as follows.

(1) Perform random sampling of the fuel cell specifications and Pu composition for 200 samples.
(2) Execute the CASMO4 and MVP codes for each sample. The total number of neutron histories in MVP is 30 million.
(3) Obtain the k-effective, the cell homogenized neutron flux and the cross-section from the calculation result of CASMO4 as shown in Table III.
(4) Obtain k-effective from the calculation result of MVP and estimate the modeling approximation error δ_k from Eq. (7). The median of statistical uncertainty of δ_k (σ_u) is 9.6 pcm.
(5) Perform PCA for the neutronics characteristics parameters shown in Table III.
(6) Select the input variable(the Pc scores) used for the input variables $\boldsymbol{x}$ in the surrogate model from the results of PCA
(7) Determine the covariance function for the Kriging by the weighted least square fitting.
(8) Develop the surrogate model to estimate the modeling approximation error δ_k and carry out LOOCV.
(9) Repeat procedure (7) and (8), and choose the best number of input variables for Kriging.

In procedure (3), 37 parameters, as shown in Table III, are obtained for the neutronics characteristic parameters. Neutron fluxes Φ_i and macroscopic cross-sections (total Σ_{ti}, fission $\nu\Sigma_{fi}$ and capture Σ_{ci}) calculated by CASMO4 are homogenized within the pin cell and collapsed to 9 groups: for which the energy boundary is shown in Table IV.

Table III. Neutronics characteristic parameters used for PCA analysis

Neutronics characteristic parameters
$k_{\text{eff}}, \Phi_i, \Sigma_{ti}, \ \nu\Sigma_{fi}, \Sigma_{ci}$ $(i = 1, \cdots, 9)$,

Table IV. Energy boundary for neutronics characteristic parameters (unit: eV)

Energy group	Upper boundary
1	1.00E+07
2	8.21E+05
3	5.53E+03
4	4.00E+00
5	1.15E+00
6	9.72E-01
7	6.25E-01
8	1.40E-01
9	5.80E-02

RESULTS

Firstly, the result of PCA is shown in Fig. 2. And Fig. 3 shows the 95% coverage probability and maximum prediction error using different number of input variables.

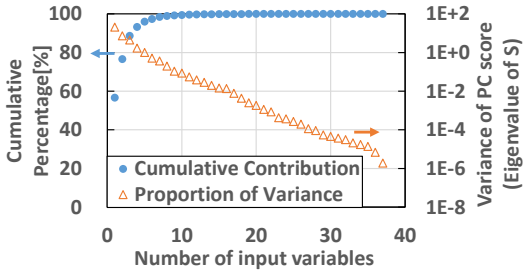

Fig. 2. Results of PCA.

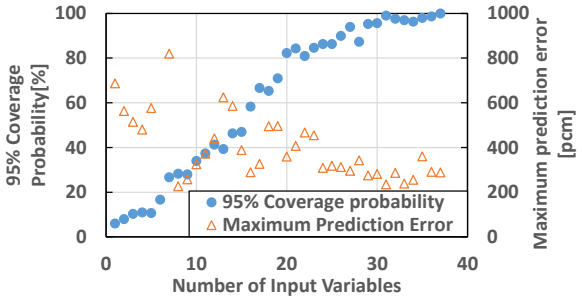

Fig. 3. Fidelity of the surrogate model with various combinations of input variables.

From Fig. 3, the 95% coverage probability increases as increasing the number of parameters. The maximum estimation error becomes minimum (approximately 230 pcm) when the number of input variables is 8, and the 95% coverage probability is about 30%. Therefore, the model is acceptable. From Fig. 2, the variance of the 8th PC is about 4×10^{-3} and, cumulative percentage approximately 99%.

Next, the verification results of LOOCV of the surrogate model are shown in Fig. 4. The plots show the true value of the modeling approximation error δ_k (horizontal) and the estimated value of δ_k (vertical) from the surrogate model. Summary of the modeling approximation error is shown in Table V.

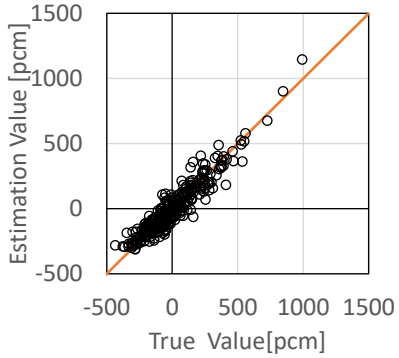

Fig. 4. Result of estimation of modeling approximation error from surrogate model with 8 input variables.

Table V. Summary of modeling approximation error δ_k obtained by the surrogate model with 8 input variables (unit: pcm)

	True value	Estimated value
Average	-9.7	0.4
Standard deviation	207.0	202.0
Maximum absolute	992.4	1145.8

As shown in Fig. 4, the surrogate model can appropriately estimate the modeling approximation error. Fig. 4 indicates that the modeling approximation error can be reduced to approximately 230 pcm when the correction by the surrogate model is used. Note that the uncorrected value of the modeling approximation error was 1000 pcm at the maximum. In other words, the modeling approximation error of a deterministic lattice physics code can be reasonably estimated by the present method.

CONCLUSIONS

In the present study, an estimation method of modeling approximation error using Kriging and PCA is developed and verified. The difference between the results obtained by a deterministic lattice physics code and a continuous-energy Monte Carlo code can be estimated by the proposed surrogate model.

In the present study, the accuracy of the surrogate model is verified in a simple geometry. Verifications with more complicated and practical conditions are the future tasks.

ACKNOWLEDGMENTS

This work was supported in part by JSPS KAKENHI, Grant-in-Aid for Scientific Research (C) (16K06956).

REFERENCES

1. H S. ABDEL-KHALIK, *Adaptive Core Simulation*, Ph.D. thesis, North Carolina State University, (2004).
2. A. YAMAMOTO, K. KINOSHITA, T. WATANABE et al., "Uncertainty Quantification of LWR Core Characteristics Using Random Sampling Method," *Nucl. Sci. Eng.*, **181**, 160 (2015).
3. A. YAMAMOTO, K. KINOSHITA, T. ENDO, "Estimation of Modeling Approximation Errors using Data Assimilation with the Minimum Variance Approach," *J. Nucl. Sci. Technol.*, **54**, 549 (2017).
4. T. HANAI, T. ENDO, Y. KODAMA et al, "Estimation of Modeling Approximation Error of Core Analysis Using the Surrogate Model with Kriging" *Trans.* ANS Winter Mtg., Washington, DC, Oct.29-Nov.2, 2017, American Nuclear Society (2017) (USB).
5. Y. BANG, H. ABDEL-KHALIK, "A Surrogate-based, Uncertainty Quantification with Quantifiable Errors," *Proc. PHYSOR2012*, Knoxville, TN, Apr.15-20, 2012, American Nuclear Society (2012) (CD-ROM).
6. I. T. JOLLIFFE, "Principal Component Analysis, Second Edition," Chapter 2 and 3 ,Springer-Verlag, New York, (2002)
7. N. CRESSIE, *Statistics for Spatial Data Revised Edition*, pp.58-72, 151-161, John Wiley & Sons, New York, (2015).
8. E. MALTE, E. KIM, B. H. FORSSEN et al., "CASMO-4, A Fuel Assembly Burnup Program, User's Manual," SSP-09/443-U, Studsvik of America, Inc., (2009).
9. Y. NAGAYA, K. OKUMURA, T. MORI et al., "MVP/GMVP II: General Purpose Monte Carlo Codes for Neutron and Photon Transport Calculations Based on Continuous Energy and Multigroup Methods," JAERI-1348, JAERI (2005).