M&C 2017 - International Conference on Mathematics & Computational Methods Applied to Nuclear Science & Engineering, Jeju, Korea, April 16-20, 2017

Kriging-based surrogate models for sensitivity analysis, uncertainty quantification and parameter calibration of the fuel performance code BISON

Tat Nghia Nguyen, Thomas Downar

University of Michigan, Nuclear Engineering & Radiological Sciences 2355 Bonisteel Blvd. Ann Arbor, MI 48105
nghiant@umich.edu, downar@umich.edu

Abstract - *In this study, a surrogate-based approach was used for the sensitivity analysis, uncertainty quantification and parameter calibration of the fission gas release (FGR) model in the fuel performance code BISON. These analyses are useful for the development and validation of the code but performing them using BISON directly is impractical because of the computation cost. A framework was developed to construct surrogate models for both single-valued and high-dimensional BISON outputs. By employing surrogate models, the computational cost for these analyses was reduced by more than two orders of magnitude. The sensitivity analysis results show that the initial fuel grain radius is the most important parameter of the fission gas release model. Uncertainties in fission gas release prediction and other BISON outputs were quantified for the given intrinsic uncertainties of the FGR model parameters. A set of FGR parameters to optimize BISON performance for the Riso-GE7 experiment was also determined by fitting the simulation output to the experimental data.*

I. INTRODUCTION

This study focused on analyses of the fission gas release (FGR) model in BISON, a fully-coupled nuclear fuel performance code under development at the Idaho National Laboratory[1]. FGR models are employed in BISON to simulate the processes induced by the build-up and release of fission gases in nuclear fuel rods. The latest physics-based FGR model in BISON, introduced by Pastore et al. [2], has shown clear improvements over the previously used empirical and semi-empirical models[3] but still has high output uncertainties due to the intrinsic uncertainties of various model parameters. Understanding the effects of these uncertain parameters on the performance of the FGR model is crucial to the development of the model and of the BISON code in general. The paper by Pastore et al. in 2015 [4] has pointed out the significance of these analyses and provided initial assessment of the uncertainty and sensitivity analysis in fission gas behavior. However, the number of parameters considered, the complexity of the simulation and ultimately the uncertainty assessment were limited because of the computational burden shared by these analyses. Specifically, the sensitivity analysis was carried out only for a simple 2D model of a single fuel pellet. The uncertainty analysis performed for the real fuel rod model from the Riso-GE7 test[5] stopped at comparing the difference between the reference outputs and those at the upper and lower bounds of the model parameters.

The objective of this research is to develop a comprehensive framework for the sensitivity analysis, uncertainty quantification and parameter calibration of computational intensive engineering codes such as BISON. Since these analyses require numerous runs of the code, the computational burden is considerable if BISON is to be run directly. To overcome this limitation, a surrogate-based approach is employed in which surrogate models are constructed for both single-valued and high-dimensional outputs of BISON. These surrogate models have the capability of providing fast mapping between the FGR model parameters (as inputs) and the BISON outputs and will be used instead of BISON in the analyses of the FGR model.

II. METHODOLOGY

The Kriging surrogate method and the principal component analysis (PCA) method are the principle mathematical tools used in this research to construct the surrogate models in this study. The Kriging surrogate is a Gaussian process regression method commonly used in many fields of machine learning and statistics to provide fast mappings between multiple input variables and a single-valued output of a target function or computer code [6, 7]. The method is directly applicable for computer codes with one or a few number of single-valued outputs. For single-valued BISON outputs, a Kriging surrogate model can be constructed for each of the outputs.

For high dimensional outputs, for instance a time series of fission gas release fraction from a BISON simulation which could consist of hundreds of time steps, it is inefficient to construct surrogate models separately for each dimension in order to predict one output series. Instead, to reduce the number of surrogate models needed, a dimension reduction technique such as the Principal Component Analysis (PCA) should be applied to the output space before a surrogate model is constructed for each of the reduced dimension[7]. The principal component analysis method is a pattern identification technique that can be used to reduce the effective dimension of a data set by projecting it into a small number of orthogonal component vectors which are the common patterns among the data set [8, 7]. By applying PCA and keeping only the first few principal components that capture majority of the variance (defined by a truncation parameter), each instance of the high dimensional output can be decomposed as a linear combination of the determined orthogonal components. A surrogate model can be constructed for each of the principal components and the combination of those sub-models provides the surrogate model for the high dimensional output. This Kriging-based surrogate method is used for high dimensional BISON outputs.

The process to construct a surrogate model for a BISON

output, as suggested in [6] for typical surrogate models, consists of the following steps:

- Variable screening: The first step is to determine the relevant input variables. For the specific application in this study, the relevant variables are uncertain parameters of the fission gas release model.
- Sampling plan: The next step is to design a sampling plan, which specifies a set of input variables for each data point in the training data set to be evaluated. For a given number of data points, an optimized sampling plan is the one covers the input space most thoroughly. The most commonly used sampling plan is the optimized Latin hypercube[6].
- Data sampling: Following the sampling plan constructed in the previous step, BISON is executed using each set of input variables. BISON outputs are collected and stored as a training data set for the surrogate models.
- Surrogate constructing: If the output of interest is single-valued, a Kriging surrogate model can be constructed directly using the training data set. If the output is high-dimensional, the PCA method is applied to the training data set to determined the number of principal components and the common eigen-modes.

For single-valued outputs, the constructed surrogate model can be used to directly predict the output given a new set of input variables. For the case with high-dimensional output, the surrogate models, one for each principal component, can be used to predict the principal components (expansion coefficients) of the new output. By combining them with the determined eigenmodes, a prediction for the high-dimensional output can be constructed. The combination of these separate surrogate models of the principal components and the eigen-modes is referred in the rest of this paper as a Kriging-based surrogate model for the high-dimensional output. A more detailed description of the these mathematical tools are provided in the following sections.

1. Kriging surrogate method

The Kriging surrogate method can be described by considering a target function $y(\mathbf{x})$, with $\mathbf{x} \in \mathbb{R}^d$ for which the objective is to build a Kriging predictor from a training data set $\mathbf{y} = (y_1, y_2, ..., y_n)^\top$ of the function values at $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$[9]. First, the prediction value at any point $\mathbf{x}$ is assumed to be the realization of a Gaussian random variable $Y(\mathbf{x})$ with mean μ and variance σ^2. In any continuous region of the target function, it is reasonable that the function values at any two points $\mathbf{x}_i$ and $\mathbf{x}_j$ are close if the distance between those two points $\left\|\mathbf{x}_i - \mathbf{x}_j\right\|$is small, which suggests the correlation between the two random variables at those two points is small. The correlation between the two random variables is given by:

$$\mathrm{Corr}\left[Y(\mathbf{x}_i), Y(\mathbf{x}_j)\right] = \exp\left(-\sum_{l=1}^{d} \theta_l \mid \mathbf{x}_{i,l} - \mathbf{x}_{j,l} \mid^{p_l}\right) \quad (1)$$

in which θ_l and p_l, $l = 1, 2, ..., d$, are the parameters of the model: θ_l controls how fast the correlation decrease and p_l controls the smoothness of the function in the l direction of $\mathbf{x}$.

The predicted values at the n points of the training set is given by:

$$\mathbf{Y} = \begin{pmatrix} Y(\mathbf{x}_1) \\ . \\ . \\ . \\ Y(\mathbf{x}_n) \end{pmatrix}$$

Each of these random variables has mean μ and variance σ, and therefore the mean vector of $\mathbf{Y}$ is $\mathbf{1}\mu$ and the covariance matrix is $\mathrm{Cov}(\mathbf{Y}) = \sigma^2\mathbf{R}$, with $\mathbf{R}(i, j) = \mathrm{Corr}\left[Y(\mathbf{x}_i), Y(\mathbf{x}_j)\right]$.

The likelihood of this data set is given by:

$$L(\mathbf{Y} \mid \mu, \sigma, \{\theta_1, ..., \theta_d\}, \{p_1, ..., p_d\}) = \frac{1}{(2\pi\sigma^2)^{n/2}|\mathbf{R}|^{0.5}} \exp\left[-\frac{(\mathbf{y} - \mathbf{1}\mu)^\top \mathbf{R}^{-1}(\mathbf{y} - \mathbf{1}\mu)}{2\sigma^2}\right] \quad (2)$$

The problem now becomes the selection of the model parameters (μ, σ, θ and p) to maximize the likelihood or the log of the likelihood. Equivalently we want to minimize the function:

$$-\ln(L) = n\ln(\sigma) + 0.5\ln(|\mathbf{R}|) + \frac{-(\mathbf{y} - \mathbf{1}\mu)^\top \mathbf{R}^{-1}(\mathbf{y} - \mathbf{1}\mu)}{2\sigma^2} \quad (3)$$

Take partial derivatives of the right hand side of Eq. 3 with respect to μ and σ and set them to zero, the optimal values of μ and σ are given by:

$$\tilde{\mu} = \frac{\mathbf{1}^\top \mathbf{R}^{-1}\mathbf{y}}{\mathbf{1}^\top \mathbf{R}^{-1}\mathbf{1}}$$

$$\tilde{\sigma} = \sqrt{\frac{(\mathbf{y} - \mathbf{1}\tilde{\mu})^\top \mathbf{R}^{-1}(\mathbf{y} - \mathbf{1}\tilde{\mu})}{n}}$$

Using these optimal values in Eq. 3 we get a expression of the negative log likelihood that depends only on $\mathbf{R}$, which depends on the θ's and p's:

$$f(\mathbf{R} \mid \{\theta_1, ..., \theta_d\}, \{p_1, ..., p_d\}) = n\ln(\tilde{\sigma}) + 0.5\ln(|\mathbf{R}|) \quad (4)$$

By using a global search optimization algorithm, the optimal values for θ and p can be estimated. After the model has been trained using the training set (for which the optimal values for the model parameters are found), the predicted value for the function at any new point $\mathbf{x}$ is the value that maximizes the likelihood of the $(n + 1)$ points data set (1 new point $\mathbf{x}$ combining with the training data set). With $\mathbf{r}^\top = (\mathrm{Corr}[Y(\mathbf{x}), Y(\mathbf{x}_1)], ..., \mathrm{Corr}[Y(\mathbf{x}), Y(\mathbf{x}_n)])$, the prediction is given by:

$$\tilde{y}(\mathbf{x}) = \tilde{\mu} + \mathbf{r}^\top \mathbf{R}^{-1}(\mathbf{y} - \mathbf{1}\tilde{\mu})$$

2. Principal component analysis

The Principal Component Analysis method can be described by considering a data matrix X of size $n \times k$, where each of the n rows represents an observation (or evaluation) of the experiment[10]. Each of the k columns provides data for a particular feature or the measurement results at a specific time or spatial step. Each observation is specified by a row vector of k elements, or in other words the problem is k-dimensional. The goal of PCA is to transform each row vector $\mathbf{x}_i \in X$ into a linear combination of m orthogonal k-dimensional unit vectors of weights $\mathbf{w}_j$ ($j = 1, 2..., m$):

$$\mathbf{x}_i = \sum_{j=1}^{m} t_{ij}\mathbf{w}_j$$

where $t_{ij} = \mathbf{x}_i \cdot \mathbf{w}_j$ is the projection of $\mathbf{x}_i$ on $\mathbf{w}_j$. The vectors of weights $\mathbf{w}_j$ are sorted in the order so that the first vector $\mathbf{w}_1$ captures the maximum possible variance of the data set:

$$\mathbf{w}_1 = \arg\max_{\|\mathbf{w}\|=1} \left\{ \sum_{i=1}^{n} (t_{i1})^2 \right\}$$

Similarly the second vector captures the second most variance, etc.
The vectors of weights $\mathbf{w}_j$ are the eigenvectors of the covariance matrix of X, ranking from the highest eigenvalue to the lowest. Assuming X with column-wise zero empirical mean, the covariance matrix of X is a $k \times k$ matrix given by $X^\top X$. This covariance matrix has k eigenvectors, i.e $m = k$.
However, not all the components need to be retained. A truncation parameter $\rho_m < 100\%$ can be used to keep only the first p eigenvectors which capture ρ of the variance, and each data vector $\mathbf{x}_i \in X$ can be transformed from the original k-dimensional space into the p-dimensional space. The amount of variance captured by the first p eigenvectors can be estimated by the cumulative eigenvalue[7, 10]:

$$\rho = \frac{\sum_{j=1}^{p} \alpha_j}{\sum_{j=1}^{k} \alpha_j} \geq \rho_m$$

III. THE RISO-GE7 IRRADIATION EXPERIMENT

The Riso-GE7 experiment is a bump test performed in the water-cooled HP-1 rig under BWR conditions in the DR3 test reactor.The test fuel pin ZX115 was the lower middle segment of four approximately 0.975 m long segments assembled to a stringer. The fuel segment was base irradiated in the Quad Cities-1 boiling water reactor (BWR) over four reactor cycles. The bump test was performed in the water-cooled HP-1 rig under BWR conditions in the Riso DR3 test reactor[5].

The GE7 experiment was modeled in BISON and is used as an assessment problem. The experiment provided two measured outputs: the fission gas release fraction and the rod outer diameter at various axial location by the end of the bump test. The simulation results from the latest BISON version compared to the experimental data are shown in Figure 1 and 2[11].

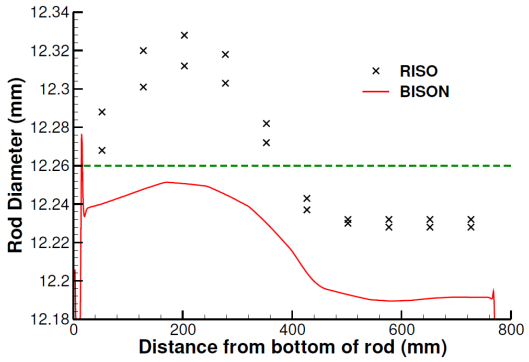

Fig. 1. BISON reference result for the outer rod diameter [11]

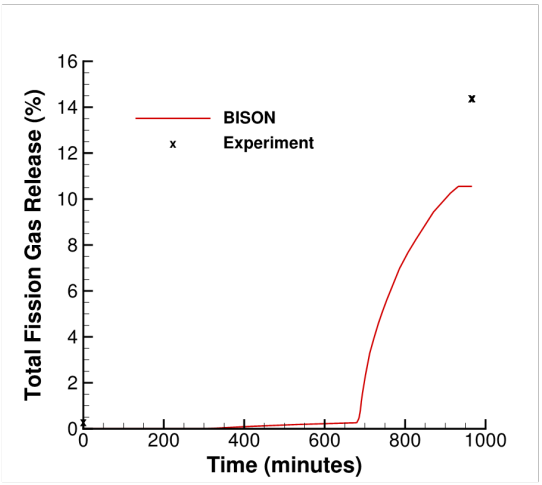

Fig. 2. BISON reference result for fission gas release fraction[11]

Six FGR model parameters were considered in this study. The uncertainty band of these parameters as analyzed in [4] are listed in Table I with the reference BISON inputs.

The topic of this research focused on analyzing BISON performance in the simulation of this experiment. In the next section, a surrogate model is constructed and validated for each the output of the GE7 experiment.

IV. CONSTRUCTION AND VALIDATION OF THE SURROGATE MODELS

In order to build the training data for the Kriging-based surrogate method, a sampling plan consisting of 100 data points was created using the optimized Latin hypercube sampling plan in DAKOTA, the uncertainty quantification and sensitivity analysis tool developed by Sandia National Laboratories[12] with each point specified a unique combination of the FGR parameters. BISON was then executed following this sampling plan and the results for the rod outer diameter prediction from 98 runs are shown in Figure 3(2 out of 100 runs failed to converge). Firstly, A Kriging surrogate model was constructed for the single-valued output of the fission gas release fraction at the end of the power ramp. For the high dimensional output of rod outer diameter, by applying

TABLE I. Fission gas release model parameters.

	Lower Bound	Upper Bound	Reference
Initial Fuel Grain Radius	2.0E-6	15.0E-6	4.68E-6
Temperature	0.95	1.05	1.0
Fuel Grain Radius Scaling Factor	0.4	1.6	1.0
Vacancy Diffusion Coef. Scaling Factor	0.1	10	1.0
Resolution Parameter	0.1	10	1.0
Intra-granular Diffusion Coef. Scaling Factor	0.316	3.162	1.0

the principal component analysis to the data set, the results of the eigenvalues and the first 3 eigenvectors, which cover more than 99.9% of the variance, are shown in Figure 4 and 5, respectively. Using these results, only the first three principal components were kept. Three Kriging surrogate models were then constructed for these three principal components, or in other words expansion coefficients.

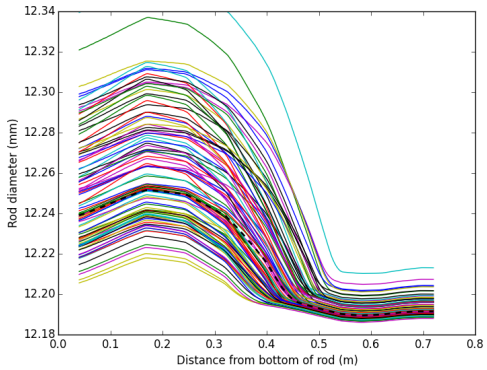

Fig. 3. Rod diameter predictions of the 98 BISON runs

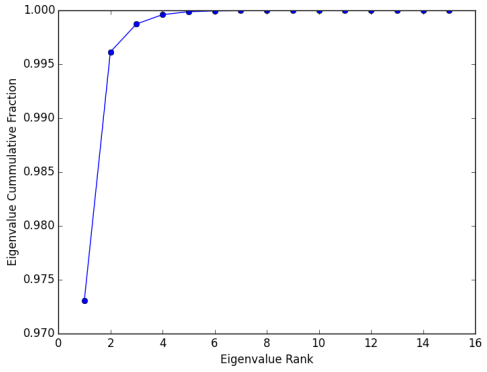

Fig. 4. Cumulative variance carried by successive eigenvalues

1. Validation of the Kriging single-valued surrogate model

A Kriging surrogate model was constructed using the data from 100 BISON runs to predict the fission gas release fraction at the end of the GE7 bump test. The prediction accuracy of that surrogate model is validated here. 50 additional BISON runs were used for this analysis, and the results of the fission gas release fraction from Bison and the corresponding predicted values are plotted in Figure 6. As indicated the relative prediction error falls well within the ±10% range and the overall RMS error from these 50 data points is 2.23%. Considering the high uncertainty in BISON simulation results, the uncertainty added by the surrogate model is considered acceptable.

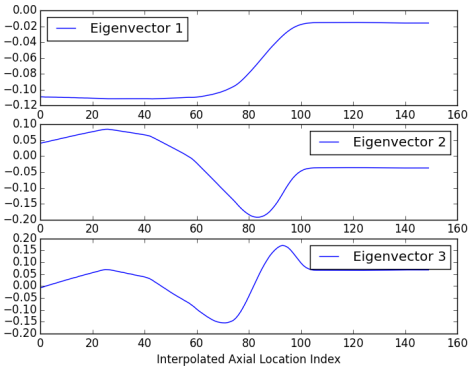

Fig. 5. The first three eigenvectors

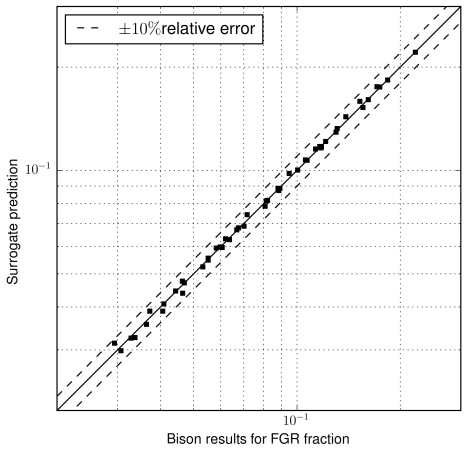

Fig. 6. Bison results vs. surrogate predictions for fission gas release

2. Validation of the high dimensional Kriging-based surrogate model

Similarly, a Kriging-based surrogate model were constructed from three separate Kriging models created for the first three expansion coefficients. This high dimensional Kriging-based surrogate model is used to predict the radial displacement at ten axial locations, ultimately provides the rod diameter predictions. Figure 7 shows a plot of the surrogate predictions in comparison with Bison results. As shown, the relative error is within the ±10% , except for a few data points around the zero-displacement region.

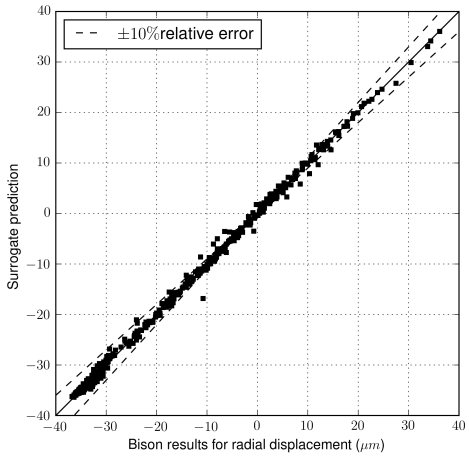

Fig. 7. Bison results vs. surrogate predictions for the radial displacement

V. APPLICATIONS OF THE KRIGING-BASED SURROGATE MODELS

This section presents three applications of the surrogate models of the BISON fission gas release model, namely: sensitivity analysis, uncertainty quantification and parameter calibration.

1. Application 1: Sensitivity analysis

By using surrogate models, one can perform global sensitivity analyses which were otherwise impractical because of the computational cost. In a global sensitivity analysis, the parameters are varied simultaneously and the relative importance of each parameter, in term of contribution to the variance of the output, can be computed. In this study, two global sensitivity analysis methods were considered, namely the Sobol's method and Fourier amplitude sensitivity analysis (FAST) method.

The surrogate models can be stored as an input function for sensitivity analysis tools. In the sections below, the calculations were performed using the Sensitivity Analysis Library in Python (SALib)[13].

A. Fission gas release fraction output

Considering the fission gas release fraction as the output variable, the Sobol's and FAST sensitivity analyses were performed for the six fission gas release parameters. The results are listed in Table II for $N = 10000$ samples. The first order Sobol index uncertainty values are the bootstrap confidence intervals provided by the calculations. The FAST library does not provide uncertainty estimations.

Results of both Sobol's and FAST methods are in good agreement, indicating that the initial fuel grain radius is the most important variable and contributes the most to the variance of the fission gas release fraction.

B. Maximum outer diameter output

Similarly for the maximum rod outer diameter as the output of interest, results for the first order Sobol index and FAST index are listed in Table III for $N = 10000$ samples.

The first order Sobol index and FAST index agree well with each other. For the maximum outer diameter as output of interest, the initial fuel grain radius and the fuel grain radius scaling factor are the two variables with highest indices. The temperate and resolution parameter scaling factor also contribute significantly to the variance of the output. As the deformation of the rod depends on the total amount of fission gas release (not the fission gas release fraction,i.e the ratio of fission gas release to fission gas generation) and other variables, it is reasonable that the initial fuel grain radius, determined in the last section to be the most important parameter to the fission gas release fraction, does not have the highest importance.

2. Application 2: Uncertainty quantification

The uncertainty band (lower-upperbound range) for each parameter listed in Table I is the 95% confidence interval[4]. Assuming a normal (Gaussian) distribution for each of the parameters, the uncertainty band for each FGR parameter is equivalent to 4 times the standard deviation ($\pm 2\sigma$). Specifically, the standard deviation for the fuel grain radius (including the initial value) is 0.3 of the mean value and the scaling factor for the temperature has a mean value of 1 and standard deviation of 0.025. Since the scaling factors for the diffusion coefficients and the resolution parameters were considered within a factor of 100 (or 10), normal distributions in the log scale were assumed for these factors. Parameters for the normal distribution of the fission gas release model are summarized in Table IV.

By using surrogate models in a Monte Carlo sampling method, the uncertainty of the outputs can be quantified given the distribution of the input variables given in Table IV. The uncertainty quantification results for the rod outer diameter prediction, analyzed with 10000 MC samples, are summarized in Table V and plotted in Figure 8. The uncertainty of the fission gas release fraction is 2.73% with the reference BISON value at 10.09%.

It should be noted that these results include the uncertainties added by the prediction uncertainty of the surrogate models and the statistical uncertainty of the Monter Carlo sam-

TABLE II. First order Sobol index and FAST index with the fission gas release fraction as the output of interest.

	Sobol index	FAST index
Initial Fuel Grain Radius	0.971 ± 0.028	0.9606
Temperature	0.0092 ± 0.0025	0.0093
Fuel Grain Radius Scaling Factor	0.0032 ± 0.0022	0.0030
Vacancy Diffusion Coef. Scaling Factor	0.0013 ± 0.0012	0.0012
Resolution Parameter	0.0006 ± 0.0008	0.0006
Intra-granular Diffusion Coef. Scaling Factor	0.0107 ± 0.0033	0.0102

TABLE III. First order Sobol index and FAST index with maximum rod diameter as the output of interest.

	Sobol index	FAST index
Initial Fuel Grain Radius	0.351 ± 0.018	0.3278
Temperature	0.105 ± 0.009	0.0951
Fuel Grain Radius Scaling Factor	0.369 ± 0.019	0.3672
Vacancy Diffusion Coef. Scaling Factor	0.0000 ± 0.0007	0.0002
Resolution Parameter	0.129 ± 0.010	0.1273
Intra-granular Diffusion Coef. Scaling Factor	0.004 ± 0.002	0.0063

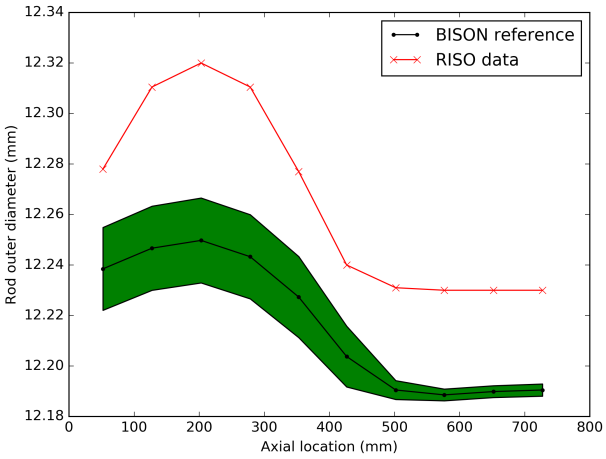

Fig. 8. UQ result by Monte Carlo sampling with $\pm\sigma$ uncertainty band

pling method. The statistical uncertainty of the results from MC sampling should can be reduced by simply increasing the number of samples. With 10000 samples, this uncertainty should be negligible. The uncertainty of the surrogate models can be reduced by improving the training data set, i.e increasing the number of data points and optimizing the sampling plan.

3. Application 3: Parameter calibration

The objective a parameter calibration is to select a set of FGR parameters that optimizes the BISON simulation performance. If the rod outer diameter is the output of interest, the root-mean-square error, in comparison with experimental data, can be used to quantify the simulation performance. The optimal set of FGR parameters is the one that minimizes the RMS error. By using the Kriging-based surrogate model for the rod diameter prediction, an objective function can be created which provides the RMS error of the simulation result given a new set of input variables. The optimal set of FGR parameters can then be selected by performing a global optimization on that objective function. Results using the SciPy L-BFGS-B optimizer[14] are listed in Table VI. BISON prediction in this optimal case is plotted in Figure 9, which shows better agreement with the experimental data in comparison with the reference case.

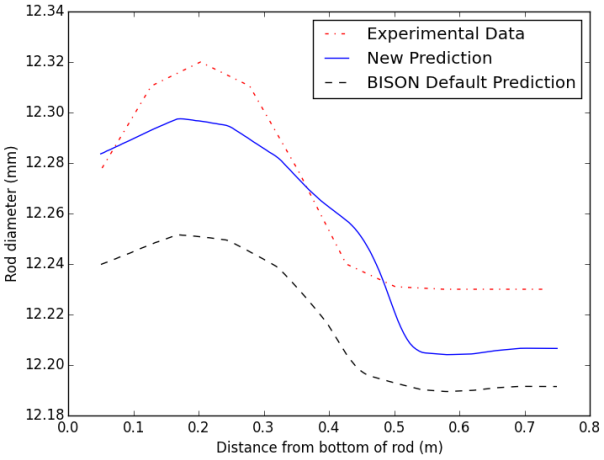

Fig. 9. Optimized surrogate prediction for GE7 rod diameter

VI. CONCLUSIONS

Comprehensive assessment analyses like sensitivity analysis, uncertainty quantification and parameter calibration are

TABLE IV. Normal distribution of the FGR parameters

	Mean	Std. Dev.	Ref. Value
Initial Fuel Grain Radius	$6.0E-6$	$1.8E-6$	$6.0E-6$
Temperature	1.0	0.025	1.0
Fuel Grain Radius Scaling Factor	1.0	0.3	1.0
Vacancy Diffusion Coef. Scaling Factor (log)	0.0	0.5	0.0
Resolution Parameter (log)	0.0	0.5	0.0
Intra-granular Diffusion Coef. Scaling Factor (log)	0.0	0.25	0.0

TABLE V. Uncertainties of the radial displacement prediction at different axial locations.

Axial location (mm)	Ref. value (mm)	Standard deviation (mm)
52	12.23848	0.01649
128	12.24667	0.01674
203	12.24975	0.01691
278	12.24328	0.01676
353	12.22731	0.01614
426	12.20376	0.01212
502	12.19050	0.00379
576	12.18856	0.00235
652	12.18987	0.00236
727	12.19049	0.00244

often very useful for the development and validation of simulation codes but performing these analyses can be challenging for high-fidelity engineering codes such as BISON because of the computational burden. These analyses requires numerous executions of the code, each of which is computationally expensive. In this case, a surrogate-based approach is recommended to reduce the computational cost. Specifically, surrogate models with the capability of providing efficient mapping between the input parameters and output of a simulation tool can be built and used to replace the original code in intensive analyses. In this study, a framework to construct surrogate models was developed for both single-valued and high-dimensional BISON simulation outputs, which consists of the following steps:

- Variable screening: Determine relevant parameters.
- Sampling plan: Design a sampling plan for the training data set.
- Data sampling: Execute the code and collect data for the training data set.
- Surrogate constructing: Construct the needed surrogate models.

For single-valued BISON outputs, a Kriging surrogate model is constructed for each output. For high-dimensional outputs, a high-dimensional Kriging-based surrogate model is constructed by combining the Kriging surrogate method with the principal component analysis method. The resultant Kriging-based surrogate models for the Riso-GE7 were validated using a set of test data and were used in various applications including sensitivity analyses, uncertainty quantification and parameter calibration of the BISON fission gas release model.

To illustrate the speed-up by the surrogate-based approach: each BISON run with three computer processors for the Riso-GE7 simulation took roughly nine hours; 100 BISON runs were performed to build a training data set for the surrogate models. Since the surrogate models can provide predictions of the outputs in under a second, the computation time needed to sample the surrogate models for all three assessment analyses is negligible. Therefore the total computational cost for all three analyses by the surrogate-based approach is 100 BISON runs. Considering that tens of thousands of BISON runs would be needed for each analysis if BISON was to be used directly, the computational cost was reduced by more than two orders of magnitude by the surrogate-based approach.

As a trade-off, predictions provided by the surrogate models can be close to the actual simulation outputs but are not 100% exact, hence an additional uncertainty is added to the results by the surrogate models. The prediction accuracy of BISON surrogate models was validated in Section IV and the uncertainty of a surrogate model can be reduced by improving the training data set, i.e increasing the number of data points and optimizing the sampling plan. Finally, although the method in this study was developed specifically for the fuel performance code BISON, it can be generalized for applications with any other engineering code.

VII. ACKNOWLEDGMENTS

The authors would like to thank Artem Yankov for his original work on combining the Kriging surrogate method and

TABLE VI. Optimal FGR parameters for the rod diameter prediction.

	Default value	Optimal value
Initial Fuel Grain Radius	4.68E-6	3.00E-6
Temperature	1.00	1.05
Fuel Grain Radius Scaling Factor	1.0	0.46
Vacancy Diffusion Coef. Scaling Factor	1.0	0.1
Resolution Parameter	1.0	0.1
Intra-granular Diffusion Coef. Scaling Factor	1.0	3.0

PCA method for high-dimensional surrogate models. This work was funded by the U.S Department of Energy Nuclear Engineering University Programs (NEUP).

REFERENCES

1. R. WILLIAMSON, J. HALES, S. NOVASCONE, M. TONKS, D. GASTON, C. PERMANN, D. ANDRS, and R. MARTINEAU, "Multidimensional multiphysics simulation of nuclear fuel behavior," *Journal of Nuclear Materials*, **423**, 149–163 (2012).
2. G. PASTORE, L. LUZZI, V. D. MARCELLO, and P. V. UFFELEN, "Physics-based modelling of fission gas swelling and release in UO2 applied to integral fuel rod analysis," *Nuclear Engineering and Design*, **256**, 75 – 86 (2013).
3. G. PASTORE, J. HALES, S. NOVASCONE, D. PEREZ, B. SPENCER, and R. WILLIAMSON, "Analysis of fission gas release in LWR fuel using the BISON code," in "Proc. of the LWR Fuel Performance Meeting, Charlotte, NC, USA," (2013).
4. G. PASTORE, L. SWILER, J. HALES, S. NOVASCONE, D. PEREZ, B. SPENCER, L. LUZZI, P. V. UFFELEN, and R. WILLIAMSON, "Uncertainty and sensitivity analysis of fission gas behavior in engineering-scale fuel modeling," *Journal of Nuclear Materials*, **456**, 398–408 (2015).
5. "The Third Risø Fission Gas Project: Bump Test GE7 (ZX115)," Tech. Rep. RISØ-FGP3-GE7, RISØ Materials Department (1990).
6. A. I. J. FORRESTER, A. SOBESTER, and A. J. KEANE, *Engineering Design via Surrogate Modelling*, University of Southampton, UK (2008).
7. A. YANKOV, *Analysis of Reactor Simulations Using Surrogate Models*, Ph.D. thesis, University of Michigan (2015).
8. J. SHLENS, "A Tutorial on Principal Component Analysis," *Computing Research Repository*, **abs/1404.1100** (2014).
9. D. R. JONES, "A Taxonomy of Global Optimization Methods Based on Response Surfaces," *Journal of Global Optimization*, **21**, 345–383 (2001).
10. I. T. JOLLIFFE, *Principal Component Analysis*, 978-0-387-95442-4, Springer New York, 2nd ed. (2002).
11. D. M. PEREZ, R. W. WILLIAMSON, S. R. NOVASCONE, R. J. GARDNER, K. A. GAMBLE, A. T. RICE, G. PASTORE, J. D. HALES, and B. W. SPENCER, "Assessment of BISON: A Nuclear Fuel Performance Code," Tech. rep., Idaho National Laboratory, Fuel Modeling and Simulation Department (2015).
12. B. M. ADAMS, L. E. BAUMAN, W. J. BOHNHOFF, K. R. DALBEY, J. P. EDDY, M. S. EBEIDA, M. S. ELDRED, P. D. HOUGH, K. T. HU, J. D. JAKEMAN, L. P. SWILER, S. J. A., D. M. VIGIL, and T. M. WILDEY, "Dakota, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis: Version 6.0 reference manual," Tech. rep., Sandia National Laboratories, Albuquerque, NM (2014).
13. J. HERMAN and W. USHER, "SALib: An open-source Python library for Sensitivity Analysis," *The Journal of Open Source Software*, **2**, *9* (jan 2017).
14. E. JONES, T. OLIPHANT, P. PETERSON, ET AL., "SciPy: Open source scientific tools for Python," (2001–), [Online; accessed 2016-10-10].