Comparison Between Gaussian Processes and DMD Surrogates for Isotopic Composition Prediction

Rabab Elzohery, Mohammad Abdo, Jeremy Roberts

Department of Mechanical & Nuclear Engineering, Kansas State University, Manhattan, KS 66506

INTRODUCTION

Ongoing work at Kansas State University aims to develop a high fidelity computational model for the KSU TRIGA Mark II research reactor. One objective of this work is to estimate the current isotopic compositions of the fuel elements with uncertainties. The challenge is that most of the fuel in the reactor core came to KSU with previous burnup with significant uncertainties. Hence a formal analysis must be done to propagate the uncertainties of the initial compositions to those of the current fuel inventory. Related work has been done in [1]. In order to compute the current fuel compositions, a depletion sequence should be established throughout the 40 years operational period of the reactor. There have been 26 different core configurations, and each has to be simulated individually to account for changes in the power history and the core loading pattern. Moreover, each configuration has to be run hundreds of times, which leads to an enormous computational cost. To reduce this computational burden, a reduced-order model that well approximates the original model could be used in place of the expensive simulation, e.g, Serpent [2] or Triton module in SCALE code [3].

In this work, two reduced order modeling methods were investigated. The first method is Gaussian Process (GP) with Principal Component analysis (PCA). GP is a widely used regression method that assumes a prior correlation between the data points. For high-dimensional input and output space, a reduction technique such as PCA can be combined with GP. Recent works in the nuclear engineering field used this technique in different applications[4].

The second method explored in this work is Dynamic Mode Decomposition (DMD). DMD is used to explore the dynamic behavior of nonlinear systems. It has been first used in fluid dynamic by Schmid [5]. It can be viewed as a combination of a PCA in the spatial domain and Discrete Fourier Transform (DFT) in the frequency domain [6]. Unlike the first method which does not provide any information about the system dynamics, DMD is capable of extracting the spatio-temporal basis along with the associated growth/decay rates. The two methods were applied to a standard TRIGA fuel element discretized into three axial zones to account for non uniform burnup. This work aims to justify the use of DMD in a previous work [7], explore other surrogates and compare them.

THEORY

Gaussian Process Based Surrogate

The first method applied in this work is the Gaussian Process (GP) regression [8]. Gaussian Process (a.k.a, Kriging) is a surface fitting method used to build an approximate function given a set of training points generated by the original model. It provides the predictions and their associated uncertainties at unknown points. The GP emulator at any point $\mathbf{x}$ has the form:

$$f^{\mathcal{GP}}(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T\boldsymbol{\beta} + \mathbf{k}(\mathbf{x})^T\mathbf{K}^{-1}(\mathbf{f} - \mathbf{G}\boldsymbol{\beta}), \quad (1)$$

where $\mathbf{g}(\mathbf{x})^T\boldsymbol{\beta}$ is the hypothesis (trend) function whose vector of regression coefficients $\boldsymbol{\beta}$ is a least-squares fit of the data, $\mathbf{k}(\mathbf{x})$ is the correlation vector between current $\mathbf{x}$ and the data points, $\mathbf{f}$ is the vector of responses of the high fidelity code under inspection, and $\mathbf{G}$ contains the evaluation of the trend basis functions at all data points. The symmetric matrix $\mathbf{K}$ is the correlation matrix of all data defined as,

$$\mathbf{K}_{i,j} = k(\mathbf{X}_i, \mathbf{X}_j) = exp(-\sum_{k=1}^{m} \theta_k |\mathbf{X}_{i,k} - \mathbf{X}_{j,k}|^{\gamma}), \quad (2)$$

where γ and θ's are hyper parameters that are computed via Maximum Likelihood Estimation (MLE). Equation 2 is interchangeably referred to as the powered-exponential covariance (kernel or correlation) function.

Construction of a surrogate model with a relatively high-dimensional input/output space may be prohibitively expensive. Rather a dimensionality reduction technique can be used with the surrogate when the data has a potential for reduction. In this work, both the inputs and outputs are the isotopic compositions, and it is desired to track a large number of isotopes with sufficient spatial resolution which increases the problem space. Also, strong correlations exist between the different isotopic concentrations and hence it should be possible to apply a dimension reduction method in conjunction with the surrogate.

One reduction method that is commonly used with GP is Principal Component Analysis (PCA). The basic idea of PCA relies on the transform of the basis on which the data is represented. In other words, the data will be re-expressed in terms of a different coordinate system that better explains the variance of the data. Consider a data matrix $\mathbf{X} \in \mathbb{R}^{n\times m}$, where n is the dimension of the feature space and m is the number of snapshots, new orthonormal basis can be identified by computing the Singular Value Decomposition (SVD) of this data matrix, i.e,

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H, \quad (3)$$

where $\mathbf{U}$ is a unitary matrix whose columns span the new space onto which the data are projected, $\boldsymbol{\Sigma}$ is a diagonal matrix containing the singular values, which can be interpreted as the variance of the data along the corresponding left singular vector, $\mathbf{V}$ is the right singular matrix, and $(.)^H$ denotes the hermitian conjugate. Assuming the data has a potential of reduction, the first r vectors are used to construct an active subspace onto which the data will be projected. The decorrelated reduced basis can now be re-expressed as a linear combination

of the original basis, each weighted according to its contribution to the data variance. A surrogate model can then be trained using data mapped from the projected data $X^{(r)}$,

$$\mathbf{X}^{(r)} = \mathbf{U}_r^T \mathbf{X}. \tag{4}$$

In this study, the DAKOTA framework was used to build the GP surrogate [8].

Dynamic Mode Decomposition

Dynamic Mode Decomposition (DMD) is a model decomposition that can serve as another ROM technique. DMD is an emerging tool that has been used recently for analyzing the behavior of dynamic systems. It is capable of providing a spatio-temporal decomposition of the data of interest into the dominant dynamic modes, i.e, DMD modes, using a set of data snapshots. Consider the same data matrix in the previous subsection $\mathbf{X}$, but now splitting it into two data matrices, whose columns represent snapshots equally spaced by a step Δt as follows,

$$\mathbf{X_1} = \begin{bmatrix} | & | & | \\ \mathbf{x_1} & \dots & \mathbf{x_{m-1}} \\ | & | & | \end{bmatrix} \tag{5a}$$

$$\mathbf{X_2} = \begin{bmatrix} | & | & | \\ \mathbf{x_2} & \dots & \mathbf{x_m} \\ | & | & | \end{bmatrix} \tag{5b}$$

The basic assumption here is that there is a linear operator $\mathbf{A}$ that approximates the dynamic behavior of the system such that

$$\mathbf{X_2} \approx \mathbf{A X_1}. \tag{6}$$

The problem now reduces to finding the eigenpairs of the operator $\mathbf{A}$. As a first step, the PCA modes need to be estimated in a way similar to what has been discussed above. By using those PCA modes, an approximation of the operator $\mathbf{A}$ can be defined as

$$\tilde{\mathbf{A}} = \mathbf{U_r^H A U_r}. \tag{7}$$

Since the matrix $\tilde{A}$ is related to A via a similarity transformation, they have a common set of eigenvalues, which can be found by evaluating the eigenvalue decompositions

$$\tilde{\mathbf{A}}\mathbf{W} = \mathbf{W}\mathbf{\Lambda}. \tag{8}$$

The DMD modes can then be expressed as

$$\mathbf{\Phi}^{DMD} = \mathbf{X_2 V_r \Sigma_r^{-1} W}. \tag{9}$$

Finally, the predicted response at any time is

$$\mathbf{X}^{DMD}(t) = \sum_{i=1}^{r} b_i \phi_i^{DMD} e^{w_i t} = \mathbf{\Phi}^{DMD} \mathbf{diag}(e^{\mathbf{w}t})\mathbf{b}, \tag{10}$$

where $\mathbf{b}$ is the amplitude of the initial condition projected on the DMD modes, and w_i is the discrete characterization of the continuous eigen values computed in (8).

$$w_i = log(\lambda_i)/\Delta t, \tag{11}$$

where λ_i is the i^{th} diagonal element in $\mathbf{\Lambda}$.

NUMERICAL EXPERIMENTS AND RESULTS

To examine the two methods demonstrated above, two numerical experiments were conducted in which a standard TRIGA fuel element placed in a square-pitch cell was depleted using Serpent to provide training data for surrogate construction and testing points to assess their predictions. The fuel element was discretized axially into three equal segments. A total of 61 isotopes were tracked, leading to a 183-dimensional space.

Case study 1

The first case study depleted a fresh TRIGA fuel element at 600° K to 55 GWD/MTHM which is approximately the highest burnup of all the fuel elements in the KSU core. A total of 150 snapshots were produced, of which 75 were used for training and 75 were used for testing. By inspecting the singular values computed by (3) and shown in Fig1, it was inferred that 3 dimensional subspace would be sufficient to capture most of the data variance. To compare the two surrogates and assess their performance, the relative errors in their predictions were estimated for all of the isotopic compositions involved in the study, but results for only ^{238}U, ^{239}Pu, ^{235}U, ^{137}Cs and ^{241}Am of one segment are shown here for brevity. As shown in figs 2 and 3, both methods were able to predict the compositions of ^{238}U and ^{235}U with a reasonable accuracy,i.e, the maximum relative error in ^{235}U predictions are 0.0175% and 0.0258% for GP and DMD surrogates respectively, while for ^{238}U, the errors are 0.00023% and 0.0026%. For the other nuclides shown in figs 4 through 6, neither method offered good predictions at early irradiation times, and this was also the case for most of the nuclides that were present with low concentrations at the beginning of the irradiation history. The predictions of both surrogates improved for later times, which indicates that the surrogates treated those isotopes as a background at the beginning since they did not contribute much to the dynamics because of their minor concentrations which in turn made their representations in the leading vectors computed by the SVD weak, but as soon as they start to evolve and make contribution, the surrogates were able recognize their presence and gave better predictions. In case of DMD, the predicted isotopic compositions were then used to estimate

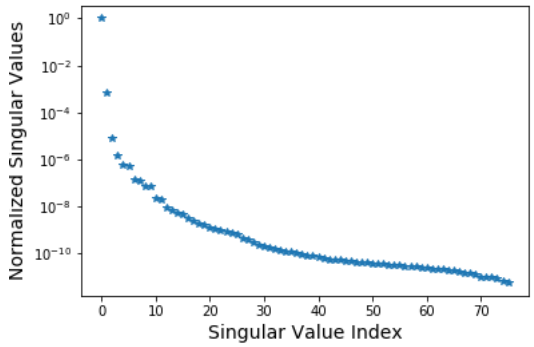

Fig. 1. Singular Values.

the corresponding k-eigenvalue in order to assess the predictions based on a more global quantity. It was found that the maximum observed reactivity bias is about 0.1%,i.e, less than $ 1.00 of reactivity.

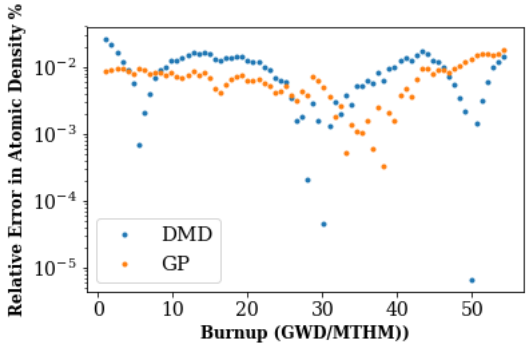

Fig. 2. Relative Errors in ^{235}U.

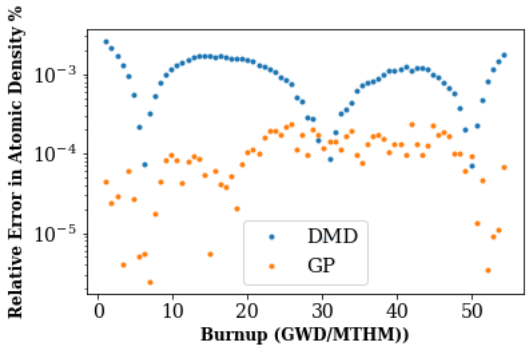

Fig. 3. Relative Errors in ^{238}U.

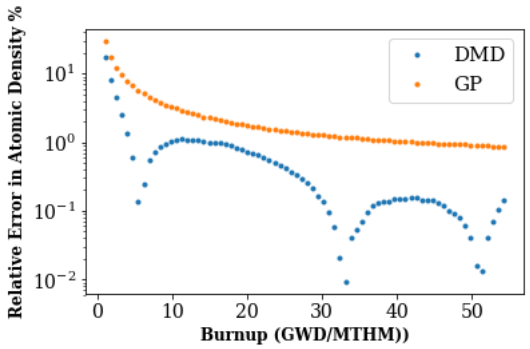

Fig. 4. Relative Errors in ^{239}Pu.

Case study 2

To examine the performance of the two surrogates more thoroughly, another case with a perturbed initial condition was studied. Here, the fuel was depleted to 25 GWD/MTHM under an elevated temperature of 800°K and then depleted another

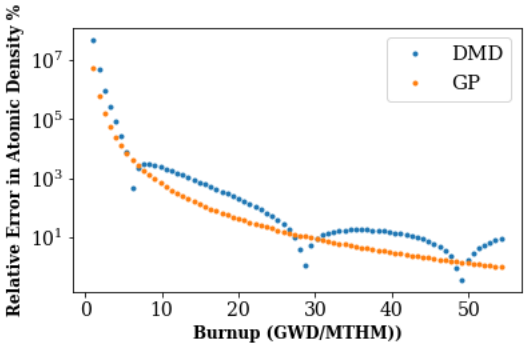

Fig. 5. Relative Errors in ^{241}Am.

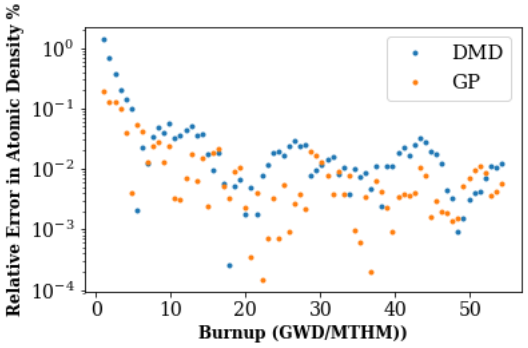

Fig. 6. Relative Errors in ^{137}Cs.

10 GWD/MTHM at 600 °K, which is similar to what happens when a hot element is shuffled to another location at lower temperature inside the reactor core. This could introduce a nonlinear effect and hence, produce nuclide concentrations that deviate from the original training domain. For this case study, the GP surrogate was trained using the same dataset used in the previous study, and for DMD, the eigenpairs are the ones obtained using this dataset, too. The figures below show that DMD outperforms GP. This is not surprising since DMD is a physics revealing technique that provides the spatial-temporal modes which demonstrate the physical state of the system so they can still be representative of the perturbed system to some extent, while the GP method used here is an interpolation method that performs well if the unknown points lie inside the training domain but performs poorly if the points deviate from this space.

The reactivity bias resulting from the DMD predictions was found to be no larger than 0.2%.

Conclusion

GP and DMD are two reduced-order models that were used here to predict the isotopic compositions of a TRIGA fuel element at different burnups. It was shown that both methods perform well when the desired prediction lies inside the space used to train the surrogates. When a perturbation was introduced to the system, DMD was still able to provide an

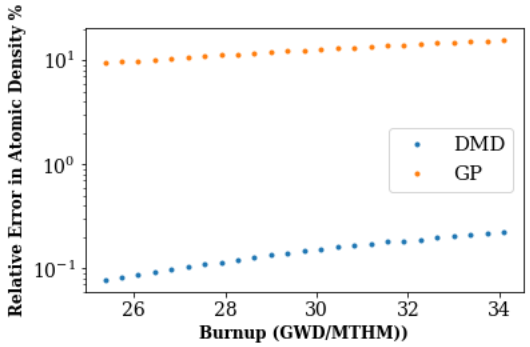

Fig. 7. Relative Errors in ^{235}U.

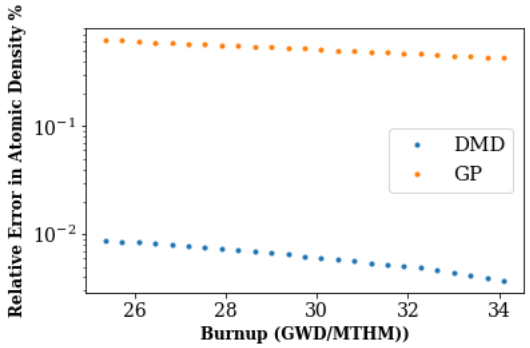

Fig. 8. Relative Errors in ^{238}U.

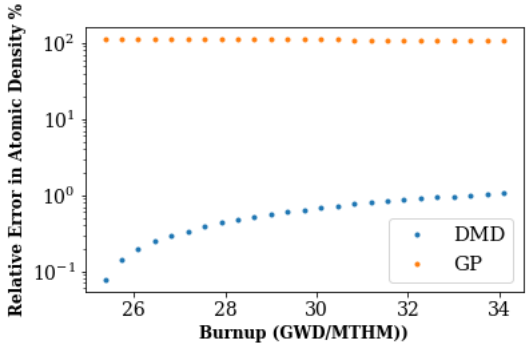

Fig. 9. Relative Errors in ^{239}Pu.

acceptable prediction. However, GP failed, and there was a noticeable difference in its performance compared to the original case. It is expected that the GP surrogate would be improved by widening the sampling space to include more snapshots from several perturbed cases. However, improvement in DMD could be introduced without requiring more computational burden by using more advanced DMD algorithm such as Multi-Resolution DMD and optimized Koopman DMD. Future work will include an extension of these methods to a full TRIGA core and relate uncertainty and sensitivity analysis.

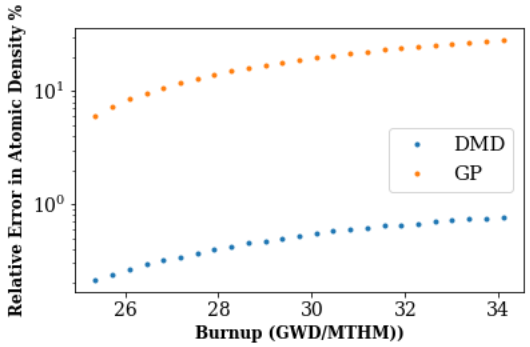

Fig. 10. Relative Errors in ^{137}Cs.

ACKNOWLEDGMENTS

The material presented is based on work supported in part by the U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, under award number NRC-HQ-60-15-G-0004.

REFERENCES

1. A. GAIROLA, S. ALSHOGEATHRI, Y. CHENG, D. LARAMORE, H. BINDRA, and J. ROBERTS, "Estimating burnup uncertainties in KSU TRIGA Mark II reactor and their impact on reactor safety," *Progress in Nuclear Energy*, **99**, 177–184 (2017).
2. J. LEPPÄNEN, "Serpent–a continuous-energy Monte Carlo reactor physics burnup calculation code," *VTT Technical Research Centre of Finland*, **4** (2013).
3. S. M. BOWMAN, "SCALE 6: comprehensive nuclear safety analysis code system," *Nuclear technology*, **174**, *2*, 126–148 (2011).
4. T. N. NGUYEN and T. J. DOWNAR, "Surrogate-based multi-experiment calibration of the BISON fission gas behavior model," *Nuclear Engineering and Design*, **320**, 409–417 (2017).
5. P. J. SCHMID, "Dynamic mode decomposition of numerical and experimental data," *Journal of fluid mechanics*, **656**, 5–28 (2010).
6. J. N. KUTZ, S. L. BRUNTON, B. W. BRUNTON, and J. L. PROCTOR, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*, SIAM (2016).
7. M. ABDO, R. ELZOHERY, and J. ROBERTS, "Data-Driven Surrogate Model to Predict Isotopic Composition Using Dynamic Mode Decomposition," *Submitted to PHYSOR 2018*.
8. M. S. ELDRED, A. A. GIUNTA, B. G. VAN BLOEMEN WAANDERS, S. F. WOJTKIEWICZ, W. E. HART, and M. P. ALLEVA, "DAKOTA, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis," Tech. rep., Tech. Rep. SAND2006-6337, Sandia National Laboratories (2006).