Automatic Selection of High-Fidelity Models and Surrogates for Uncertainty Analysis

Congjian Wang*, Andrea Alfonsi*, Diego Mandelli*, Paul Talbot*, Cristian Rabiti*

*Idaho National Laboratory, 2525 Fremont Ave., Idaho Falls, ID 83415
{congjian.wang, andrea.alfonsi, diego.mandelli, paul.talbot, cristian.rabiti}@inl.gov

INTRODUCTION

Surrogate models (SMs) with distinct features developed in the last decade provide an efficient substitute to the high-fidelity physical simulations [1, 2]. However, it is generally challenging to quantify the accuracy of surrogate models in a region of interest and examine reliability in the overall domain without additional system evaluations. This paper introduces a new scheme, automated adaptive model selection, to adaptively construct and quantify a set of surrogate models and automatically switch between the surrogates and a high-fidelity code. Cross-validation approaches [3], such as KFold, Leave One Out (LOO) and Leave P Out (LPO), are employed to assess the global accuracy of the surrogates. The switching strategy is based on local validation metrics, crowding distance [4], that are able to provide an assessment on the goodness of the prediction made by the surrogates. This decision metric decides whether the surrogates are good enough for a certain realization or whether the high-fidelity code needs to be executed. This new scheme has been implemented in **R**isk **A**nalysis in **V**irtual **EN**vironment (**RAVEN**) software [5], which can be employed to accelerate the Probabilistic Risk Assessment (PRA) and Uncertainty Quantification (UQ) processes.

AUTOMATIC MODEL SELECTION

Any high-fidelity model is composed of a vector $\boldsymbol{x}$ of input parameters, state variable $\boldsymbol{u}$, and state equations $\boldsymbol{\Gamma}(\boldsymbol{x},\boldsymbol{u}) = 0$. Once the solution of state variable $\boldsymbol{u}$ is available, the high-fidelity model may also include the extra functions $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{u})$ to compute the system responses $\boldsymbol{y} = \boldsymbol{F}(\boldsymbol{x},\boldsymbol{u})$. SMs, generally not based on the physics of the system, are purely mathematical models that could be used to capture the relationships between input parameters and system responses of the high-fidelity model, e.g. $\boldsymbol{R}(\boldsymbol{x},\boldsymbol{\beta}) = \boldsymbol{y}$, where $\boldsymbol{\beta}$ are the unknown coefficients that can be determined when the SMs are trained with a batch of input parameters $\boldsymbol{x}$ and system responses $\boldsymbol{y}$. Once the SMs are trained, the prediction $\boldsymbol{p} = \boldsymbol{R}(\boldsymbol{x},\boldsymbol{\beta})$ from the SMs for the new set of input parameters $\boldsymbol{x}$ can be computed. Fig. 1 represents the proposed automatic model selection scheme that has been recently implemented in RAVEN [6]. For a single evaluation request $\boldsymbol{x}$:
1) The global accuracy $\bar{d}$ in the SMs is computed via the cross validation methods [3]
2) If the SMs are not converged based on step 1, the high-fidelity model is inquired and the new point is used to update the surrogate training set. The steps 1 and 2 are repeated for all the new evaluations until the convergence test via the cross validation is satisfied
3) SMs are used to predict the outcome $\boldsymbol{p}$ based on new $\boldsymbol{x}$;
4) The local prediction validation metric is inquired in order to estimate the validity of the prediction α_p;
5) If the prediction meets a user-defined acceptability criterion, the outcome $\boldsymbol{p}$ is kept and used as actual "true" value $\boldsymbol{y}$;
6) If the prediction does not meet the user-defined acceptability criterion, the outcome $\boldsymbol{p}$ is discarded and the high-fidelity model is inquired producing the response $\boldsymbol{y}$. The new evaluation is added to the current training set and the SMs are then reconstructed.

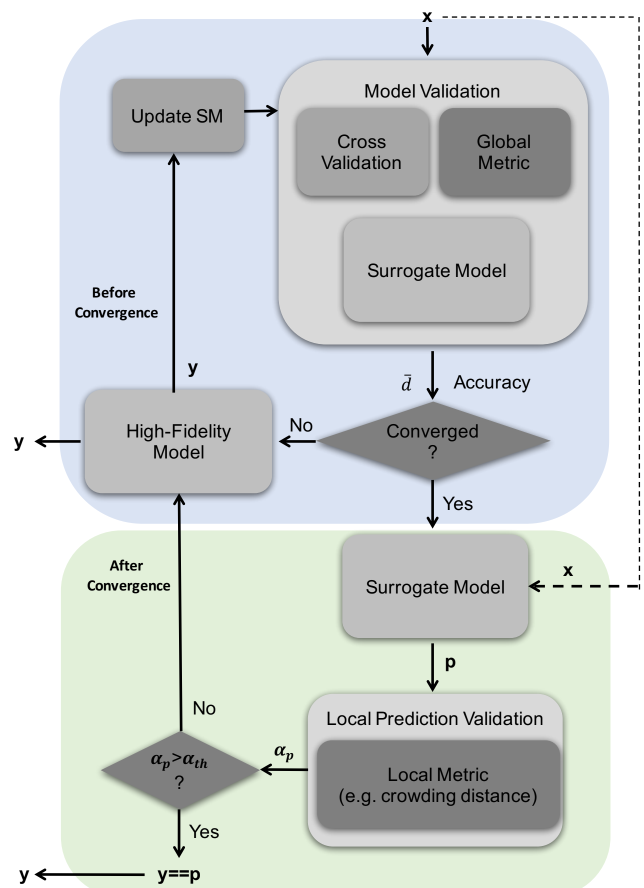

Fig. 1. Automatic model selection scheme in RAVEN.

In general, cross validation is a statistical methodology that partitions the training space and tests the SMs against the real data to assess the global accuracy in the SMs without additional high-fidelity code evaluations. This technique is crucial to identify when SMs are converged, which represents the first trigger that needs to be satisfied in order to start using the SMs instead of the high-fidelity model. Even if the convergence test performed via cross validation is the key component, an additional trigger needs to be identified to get information about the "confidence" of a single prediction; indeed, the cross validation is a global metric that does not provide any information about the "confidence" of a certain prediction. For this reason, a local validation metric based on *crowding distance* has been implemented. This decision metric decides if the SMs are good enough for a certain realization or if the high-fidelity model needs to be used.

The *crowding distance* as defined in Eq. 1 is used in [4] to formulate a reliable trust region for SMs. As mentioned in [4], the *crowding distance* of a point represents the density of training points around that point:

- Small values indicate more training points around that point, and the error measures of the SMs are expected to be relatively smaller around that point. Therefore, the confidence of the prediction of SMs is much higher.
- Large values indicate fewer training points around that point, and the error measures of the SMs are expected to be relatively larger around that point. Therefore, the ability of the SMs to predict the outcomes of interest with a reasonable accuracy is not guaranteed

$$\boldsymbol{CD}_p = \sum_{j=1}^{n_{train}} \left\| x_j - x_p \right\|^2 \quad (1)$$

where $\boldsymbol{CD_p}$ is the crowding distance of point x_p with $\boldsymbol{n_{train}}$ training data points. In addition, parameter α_p (normalized between 0 and 1), defined in Eq. 2, represents the decision coefficient for determining if the SMs evaluations should be accepted or rejected.

$$\alpha_p = \frac{\max(\boldsymbol{CD}) - \boldsymbol{CD}_p}{\max(\boldsymbol{CD}) - \min(\mathbf{CD})} \quad (2)$$

Practically, a threshold (α_{th}) of 0.2 is generally acceptable. In other words, the crowding distance metric reflects that SMs are valid within the global accuracy if new generated inputs are inside the training points, and invalid if new generated inputs are outside the training points.

RESULTS

In order to test the correct implementation of the newly developed capability in RAVEN, a single base accident scenario case has been considered: Station Black Out (SBO) accident of a Pressurized Water Reactor (PWR) designed in Ref. [7]. It is important to notice that the main goal of the results presented is to show the potential of the new scheme only, and the detailed description about the model can be found in [7]. The accident sequence and, consequentially, the system have been modeled using RELAP5-3D [8]. The PWR model is based on the so-called INL Generic PWR (IGPWR) model [9]. The input deck is modeling a 2.5 GWth Westinghouse 460 3-loop PWR, including the reactor pressure vessel (RPV), the 3 loops and the primary and secondary sides of the steam generators (SGs) as shown in Fig. 2. Four independent channels are used for representing the reactor core. Three channels model the active core and one channel models the core bypass. Different power values are assigned to the three core channels in order to take into account the radial power distribution. Passive and active heat structures simulate the heat transfer between the coolant and fuel, the structures and the secondary side of the IGPWR. The possible operator actions that can occur during a SBO event with the reactor at full power are implemented through the RELAP5-3D control logic. These actions include SG cool-down, feed and bleed, flow control, primary/secondary side emergency injection, etc.

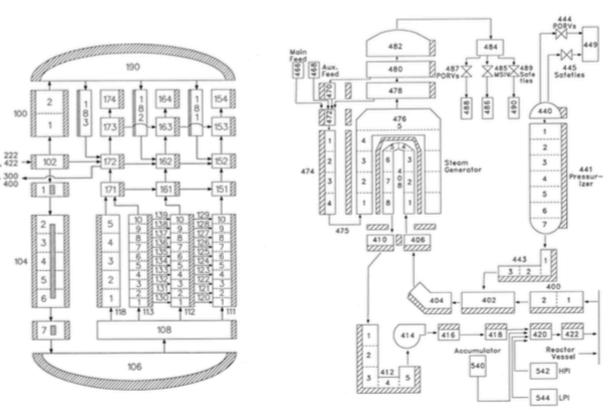

Fig. 2. PWR Model

For the scope of this analysis, 11 stochastic parameters have been identified, focused on two main elements: lifetime of the batteries and the LOCA associated to the seal of the Reactor Coolant Pumps (RCPs). Battery systems provide DC power to I&C systems of the PWR such as the control of the Pilot Operated Relief Valves (PORVs). Regarding the recovery of the unit through the EPEs, it has been modeled with a single stochastic parameter which represents the time to connect the EPE to its own unit. Lastly, the recovery plan followed by the plant crew has been modeled using a single parameter: recovery strategy. Considering that the main goal of this summary is to show the newly developed surrogate capabilities only, the details of the stochastic modeling (i.e. distributions) is not reported but it can be found in Ref. [7].

In order to accelerate the PRA and UQ analysis, the *HybridModel* based on the proposed automatic model selection scheme has been developed. The support vector regression (SVR) SM from RAVEN has been used in parallel with RELAP5-3D. In Table I the settings for the cross-validation scheme (the one used to assess the global convergence of the SM) is reported. Once the SM is converged, the *crowding distance* local metric has been used with a threshold of 0.2. In order to compare the results with and without the surrogate model, 1000 Monte Carlo samples have been generated on the scenario described above. When the *HybridModel* is used, of the 1000 Monte Carlo samples, only 200 were run using the RELAP5-3D model, resulting in a computational time reduction of 800*4 = 3200 CPU-hrs. Fig. 7 illustrates the automatic model selection of the *HybridModel*.

TABLE I. Cross Validation Parameters

Parameter	Short description	Value used
Algorithm	Partition Algorithm	KFold
N_splits	Number of folds	10
Metric	The score that is used in the CV process	Mean absolute error (0.005)

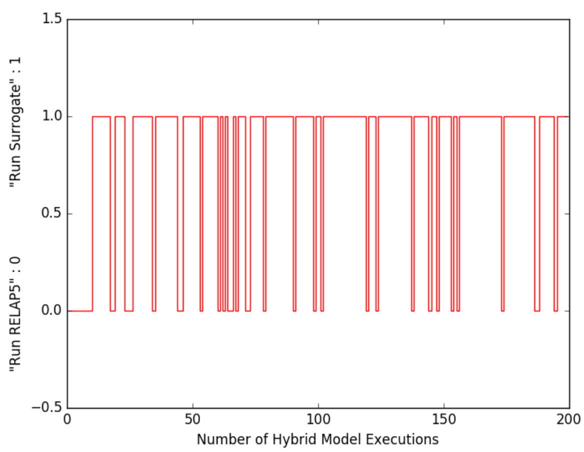

Fig. 7. Automatic model selection for the first 200 runs (The value "1" indicates that we will execute the surrogate while the value "0" indicates that we will execute RELAP5 model)

CONCLUSIONS

The objective of this summary is to present a novel scheme, automatic adaptive model selection within RAVEN framework, to assess the validity of the predictive capabilities of surrogate models. Both cross validations and crowding distance are used to assess the global accuracy and local validity of the surrogate models, respectively. The effectiveness of the new scheme is tested on a Station Black Out scenario for a prototypical Pressurized Water Reactor (modeled with RELAP5-3D). In the future, other local validation metrics that might better validate the surrogates will be investigated.

REFERENCES

1. R. Myers and D. Montgomery, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments,* 2nd edn. Wiley-Interscience, New York (2002)
2. H. S. Abdel-Khalik, Y. Bang and C. Wang, "Overview of Hybrid Subspace Methods for Uncertainty Quantification, Sensitivity Analysis," *Annals of Nuclear Energy*, **Vol. 52**, pp. 28-46 (2013)
3. R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*. San Mateo, CA (1995)
4. J. Zhang, S. Chowdhury, and A. Messac, "An Adaptive Hybrid Surrogate Model," *Struct Multidisc Optim*, **Vol. 46**, pp. 223-238 (2012).
5. A. Alfonsi et al., "RAVEN Theory Manual," INL/EXT-16-38178 (2016)
6. A. Alfonsi et al., "Status of Adaptive Surrogates within the RAVEN Framework," INL/EXT-17-43438.
7. D. Mandelli et al., "Multi-Unit PRA: a Simulation-Based Approach", submitted to Reliability Engineering & System Safety (2017)
8. C. B. Davis, "Assessment of RELAP5-3D Using Data from Two-Dimensional RPI Flow Tests," Proceedings from the 1998 RELAP5 International Users Seminar, College Station, Texas, May 17-21, 1998.
9. C. Parisi et al., "Demonstration of External Hazards Analysis", Technical Report, Idaho National Laboratory Technical Report: INL/EXT-16-39353, 2016.