Risk Subcommittee Session

Bayesian Approach to Risk Assessment

(4) Application of Bayesian Statistics to Source Term Analysis

*Xiaoyu Zheng[1], Tomoyuki Sugiyama[1] and Yu Maruyama[1]

[1]Japan Atomic Energy Agency

1. Introduction

Bayesian statistical methods use Bayes' theorem to compute or update probabilities after obtaining new data. The methods are widely used in nuclear probabilistic risk assessment for updating, for example, failure probabilities of system or components.

Bayesian methods are also powerful to build predictive models. If there is an unknown model $\mathcal{M}$ with a set of input variables $\mathbf{x}$, we need to find the best set of parameters $\boldsymbol{\theta}$, which map the input variables $\mathbf{x}$ to output results $\mathbf{y}$. The model can be written as: $\mathcal{M}: \mathbf{y} = f(\boldsymbol{\theta}, \mathbf{x})$. The predictive Bayesian model aims to find an appropriate model $\widehat{\mathcal{M}}$, which fits best to the observed data $\mathcal{D}$. When the form of the model is clear, we compute the Maximum a Posteriori Estimation (MAP) of all parameters ($\boldsymbol{\theta}$), and this process is known as the Bayesian parametric approach. When the form of the model is unclear, we need to find a best-fitted model for the data, which is based on a prior distribution of all possible models $p(\mathcal{M})$. If the prior of unrestricted shapes is constructed on the space of functions, this process is known as the Bayesian nonparametric approach [1].

Bayesian predictive models are useful in some traditional nuclear research fields such as nuclear reactor severe accident. Even though most models in severe accident simulation are deterministic and physical/chemical-rule-based, statistical models can still help to reduce modeling complexity and give us insights from a probabilistic perspective. At JAEA, we applied Bayesian approaches to severe accident source term simulation, including efforts on uncertainty and sensitivity analyses, optimization analysis and prediction of chemical forms of fission products (FP). The key step is that we build statistical surrogate models to assist numerical simulations, which are generally performed using mechanistic codes, for example, MELCOR [2] and THALES2/KICHE [3], etc. The scientific simulation of source terms reveals the consequence of a severe accident, and Bayesian statistics provides supports by creating more simplified predictive models.

2. Bayesian Statistics and Surrogate Model

The best model fitting the available database $\mathcal{D}$ is the one which maximizes the posterior distribution $p(\mathcal{M}|\mathcal{D})$, and the computation of the posterior distribution can be written in the form of Bayes' rule [4].

$$p(\mathcal{M}|\mathcal{D}) \propto \mathrm{p}(\mathcal{D}|\mathcal{M})p(\mathcal{M}) = \int p(\mathcal{D}|\boldsymbol{\theta})p(\boldsymbol{\theta}|\mathcal{M})p(\mathcal{M})d\boldsymbol{\theta} \quad (1)$$

By choosing an appropriate form of the prior distribution of possible models (may be an infinite number of models) and a likelihood function, an optimal model can always be found. Equation (1) explains the Bayes' rule from the perspective of models (instead of parameters), and all parameters $\boldsymbol{\theta}$ of each model are integrated out for the model selection process. We use this method to find surrogate models (or reduced order models) for mechanistic severe accident codes.

Figure 1 illustrates the process of how to train and validate a statistical surrogate model. A surrogate model is equivalent to a mechanistic model regarding to the mapping between inputs and outputs. The main difference between two models is that a surrogate model is statistical and there is no physical/chemical rule inside. The advantage allows us to build a much simpler model and such a model is generally fast-running. At first, we perform multiple computation of the mechanistic modes based on random sampled inputs (A). The according input/output database is used for training a surrogate model, with the aid of Bayesian methods. The correctness of the surrogate model can be validated by comparing with the original code against new inputs.

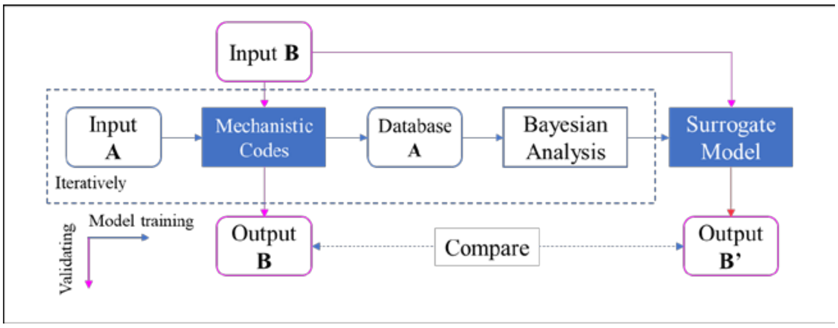

Figure 1 Training of a surrogate model using Bayesian approaches

3. Application of Bayesian Statistics to Severe Accident Source Term at JAEA

At JAEA, we apply Bayesian approaches to source term analysis. Source term is evaluated using integrated codes such as MELCOR and THALES2/KICHE. Execution of such codes is time-consuming, and because a model of a whole nuclear power plant includes too many sub-models, the relationship between inputs (plant parameters) and outputs (source term) is extremely unclear. Bayesian approaches help us optimize a statistical model to predict the outputs and then avoid direct execution of the severe accident codes. As the previous work, we introduce three fields of application in Table 1.

(1) Uncertainty and sensitivity analyses [5]: a Bayesian nonparametric method (Dirichlet process [6]) is used for training a surrogate model to predict the simulation results of MELCOR and also to plot the probability density function of uncertainty analysis results.

(2) Optimization analysis [7]: another Bayesian nonparametric method (Gaussian process [8]) is used for predicting the probable optimum for the timing of containment venting.

(3) Prediction of FP chemical forms [9]: to simplify the models in severe accident codes, we applied both Bayesian and non-Bayesian methods to train surrogate models of VICTORIA [10] and CHEMKEq [11]. The statistical models are integrated into severe accident code, THELAS2/KICHE.

To perform the Bayesian analysis, there are open-sourced libraries for programming languages such as R (DPpackage for Dirichlet process) and Python (scikit-learn for most of Bayesian approaches including Gaussian process, Dirichlet process, etc.).

Table 1 Previous efforts at JAEA relating to Bayesian analysis

	Research Topic	Statistical Algorithms	Mechanistic Codes	Usage
1	Source term uncertainty and sensitivity analyses	Dirichlet process (Bayesian nonparametric)	MELCOR (SNL)	Surrogate model, Probability density estimation
2	Optimization of severe accident consequence-mitigation measures	Gaussian process (Bayesian nonparametric)	THALES2/KICHE (JAEA)	Surrogate model, Global optimization
3	Prediction of chemical forms of FP	K-nearest-neighbors regression (nonparametric) Dirichlet process (Bayesian nonparametric)	VICTORIA (SNL) CHEMKEq (JAEA)	Surrogate model

3-1. Example: source term uncertainty and sensitivity analyses

MELCOR is widely used for source term analysis, but it is still necessary to estimate the uncertainties during simulation, and when try to reduce the uncertainties, it is also required to estimate the sensitivity of input parameters. Random sampling is an effective way to observe the parametric uncertainties in simulation, and it usually needs only hundreds of code executions to generate a stable probability density function (according to Wilks' formula). The sensitivity analysis (e.g. Sobol' global sensitivity index), however, needs a great number of code executions to reach reasonable results, because the sensitivity of a parameter will be affected by the setting of other parameters. It is laborious to run MELCOR directly for thousands of times. Instead, we train a surrogate model, run the surrogate with new inputs for numerous times, and evaluate the global sensitivity of all uncertain inputs. The detailed process is shown in Figure 2. After a raw screening of parameters, we use random sampling (Monte-Carlo or LHS) to generate inputs for MELCOR simulation. A probability density function of source term can be concluded based on the simulation data. Then, a surrogate model is trained based on the database. If the predictability of the surrogate model is not good enough, we execute MELCOR more until the prediction shows agreement with the real simulation. At last, we run the surrogate model to obtain the quantitative sensitivity measure of all uncertain inputs from the viewpoint of uncertainty reduction. Figure 2 shows the example results of (a) the predictability of a surrogate model and (b) global sensitivity analysis using Sobol' index. Evaluated via the surrogate model, the probability density function predicts the most probable released amount of CsI to the environment, which agrees with the MELCOR simulation result (the red dot). By iteratively running the surrogate model, as an example, all three uncertain input parameters can be ranked according to their contribution to the source term uncertainties.

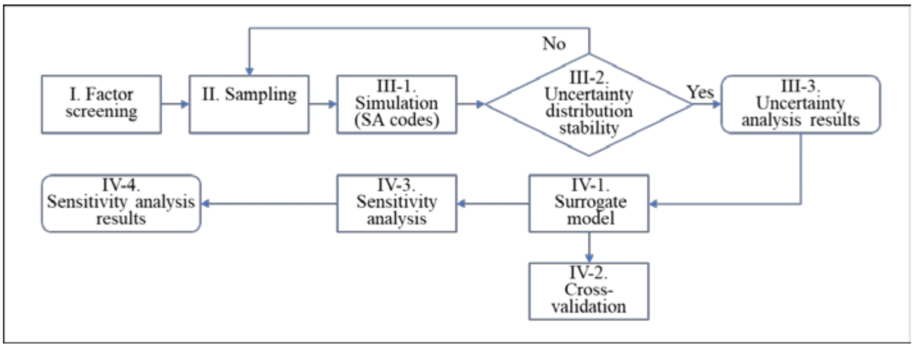

Figure 2 The process uncertainty and sensitivity analyses of severe accident source term

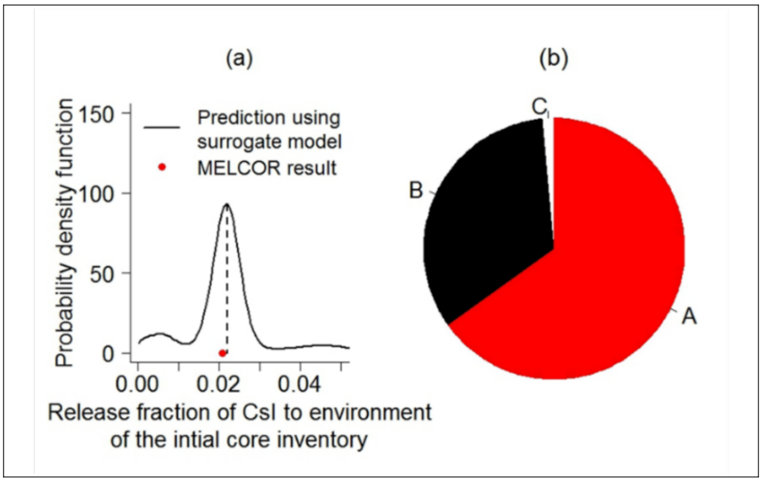

Figure 3 Sensitivity analysis results: (a) the validation of surrogate model by comparing with MELCOR simulation (b) global sensitivity measure of parameters A, B and C

4. Conclusions

We are applying machine learning, especially Bayesian approaches, to severe accident analysis. The predictive Bayesian models greatly saved computational costs, and Bayesian approaches also show the potential to other on-going researches at JAEA, for example, simulation-based risk assessment.

Acknowledgements

The development of surrogate models to predict fission products chemical forms was performed under the support by the Nuclear Regulation Authority of Japan.

References

[1] Hjort N.L., et al., Bayesian nonparametrics. Cambridge University Press (2010)

[2] Gauntt R.O., et al., MELCOR computer code manuals: Vol.1&Vol.2, Version 1.8.5, Rev.2. NUREG/CR-6110 (2000)

[3] Ishikawa J., et al., Analysis for iodine release from unit 3 of Fukushima Dai-ichi nuclear power plant with consideration of water phase iodine chemistry. J Nucl Sci Technol, 52(3):308-314 (2015)

[4] Murphy K.P., Machine learning: a probabilistic perspective. The MIT Press (2012)

[5] Zheng X., et al., An integrated approach to source term uncertainty and sensitivity analyses for nuclear reactor severe accidents. J Nucl Sci Technol, 53(3): 333-344 (2016)

[6] Gershman S.J. & Blei D.M., A tutorial on Bayesian nonparametric models. J Math Psychol, 56:1-12 (2012)

[7] Zheng X., et al., Bayesian optimization analysis of containment-venting operation in a boiling water reactor severe accident, Nucl Eng Technol, 49:434-141 (2017)

[8] Rasmussen C.E. & Williams C.K., Gaussian processes for machine learning. The MIT Press (2006)

[9] Zheng X., et al., Application of Bayesian approaches to nuclear reactor severe accident analysis. In: Proceedings of ASRAM2017, Yokohama, Japan (2017)

[10] Bixler NE. VICTORIA 2.0: a mechanistic model for radionuclide behavior in a nuclear reactor coolant system under severe accident conditions. NUREG/CR-6131 (1998)

[11] Ito H., et al., CHEMKEq: Evaluation code for chemical composition based on partial mixed model with chemical equilibrium and reaction kinetics. JAEA-Data/Code 2018-012 (2018)