Advanced methodology for uncertainty propagation in computer experiments with large number of inputs

B. Iooss, A. Marrel

ADVANCED METHODOLOGY FOR UNCERTAINTY PROPAGATION IN COMPUTER EXPERIMENTS WITH LARGE NUMBER OF INPUTS

B. Iooss[1] and A. Marrel[2]
[1]: EDF R&D, 6 Quai Watier, 78401, Chatou, France
[2]: CEA, DEN, DER, F-13108, Saint-Paul-lez-Durance, France

bertrand.iooss@edf.fr, amandine.marrel@cea.fr

ABSTRACT

Complex computer codes, as the ones used in thermal-hydraulic accident scenario simulations, are often too time expensive to be directly used to perform uncertainty propagation. A solution to cope with this problem consists in replacing the cpu-time expensive computer model by a cpu inexpensive mathematical function, called metamodel. Among the metamodels classically used in computer experiments, the Gaussian process model has shown strong capabilities to solve practical problems. However, in case of high dimensional experiments (with typically several tens of inputs), the Gaussian process metamodel building process remains difficult. To face this limitation, we propose a methodology which combines several advanced statistical tools: initial space-filling design, screening to identify the non-influential inputs, Gaussian process metamodel building with the group of influential inputs as explanatory variables. The residual effect of the group of non-influential inputs is captured by another Gaussian process metamodel. From this joint metamodel, uncertainty propagation (here 95%-quantile estimation) can be performed. The efficiency of the methodology is illustrated on a thermal-hydraulic calculation case simulating accidental scenario in a Pressurized water Reactor. More precisely, a Loss Of Coolant Accident (LOCA) is considered, which takes into account a double-ended guillotine break with a specific size piping rupture.

1. INTRODUCTION

Quantitative assessment of the uncertainties tainting the results of computer simulations is nowadays a major topic of interest in both industrial and scientific communities. One of the key issues in such studies is to get information about the output when the numerical simulations are expensive to run. For example, in nuclear engineering problems, one often faces up with cpu time consuming numerical models and, in such cases, uncertainty propagation, sensitivity analysis, optimization processing and system robustness analysis become difficult tasks using such models. In order to circumvent this problem, a widely accepted method consists in replacing cpu-time expensive computer models by cpu inexpensive mathematical functions, called metamodels [6]. This solution has been applied extensively and has shown its relevance especially when simulated phenomena are related to a small number of random input variables (see [7] for example).

However, high dimensional numerical experiments are often present in industrial practice (with typically several tens of inputs). The current methodology consists in a first screening step in

order to reduce the dimension of the inputs, then in a quantitative sensitivity analysis allowing to prioritize the inputs as a function of their influence on the output uncertainty [9]. A following Monte Carlo sampling step is applied in order to estimate the desired statistical quantity of interest, such as a 95%-quantile with a 95% confidence level [4].

Recent works [11, 10] have proposed a more efficient robust method where a Gaussian process (Gp) metamodel [17, 12] is built upon the same calculations sample as the screening step, and using the results of the screening step. In this paper, we pursue these efforts by showing how to use the metamodel to solve an uncertainty propagation problem (high-quantile estimation), and applying the methodology to a thermal-hydraulic accident scenario.

The system under study is denoted

$$Y = g\left(X_1, \ldots, X_d\right) \tag{1}$$

where $g(\cdot)$ is the numerical model (also called the computer code), $Y \in \mathbb{R}$ is the output and $\mathbf{X} = (X_1, \ldots, X_d) \in \mathbb{R}^d$ are the input variables. Our approach consists in three steps:

1. **Step 1: Initial experimental design.** Knowing the variation domain of the input variables, a design of n experiments is firstly performed and yields n model output values. To constitute this learning sample, we use a space-filling design (SFD) of experiments, providing a full coverage of the high-dimensional input space [6, 18].

2. **Step 2: Screening.** From the learning sample, a screening technique is performed in order to identify the primary influential inputs (PII) on the model output variability. It has been recently shown that screening based on dependence measures [3, 5, 10] are very efficient methods which can be directly applied on a SFD. One of their great interest is that, additionally to their screening job, the sensitivity indices they provide can be quantitatively interpreted. From these screening results, the inputs are then ordered by decreasing PII, for the purpose of the metamodeling step.

3. **Step 3: Joint metamodeling and metamodel validation.** The sorted inputs are successively included in the group of explanatory inputs while the other inputs are considered as a global stochastic (*i.e. unknown*) input and a joint Gp metamodel is built [11]. At each iteration, a first Gp model [12], only depending on the explanatory inputs, is built to approximate the mean component of the metamodel. The residual effect of the other inputs is captured using a second Gp model which approximates the variance component as a function of the explanatory inputs. The accuracy and prediction capabilities of the metamodel are controlled on a test sample or by cross-validation.

4. **Step 4: Use of the metamodel**. A quantitative sensitivity analysis or an uncertainty propagation step can be performed using the joint metamodel instead of the computer model, leading to a large gain of computation time [19, 11]. In this work, we are interested by a the estimation of the 95%-quantile of the model output.

The next section of this paper presents each step of the general methodology. The third section shows an application of this work on a thermal-hydraulic calculation case simulating accidental scenario in a nuclear reactor. The last section gives some conclusions of this work.

2. GENERAL METHODOLOGY

2.1 Step 1: Initial design of experiments

The objective of the initial sampling step is to investigate the whole variation domain of the uncertain parameters in order to fit a predictive metamodel which approximates as accurately as possible the code in the whole domain of variation of the uncertain parameter. For this, we use a space-filling design (SFD) of a certain number n of experiments, providing a full coverage of the high-dimensional input space [6]. This design provides a learning sample. Mathematically, this corresponds to the sample $\left(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)}\right)$ which is performed on the model g. This yields n model output values denoted $\left(y^{(1)},\ldots,y^{(n)}\right)$ with $y^{(i)}=g(\mathbf{x}^{(i)})$. The obtained learning sample is denoted (X_s,Y_s) with $X_s=\left[\mathbf{x}^{(1)^T},\ldots,\mathbf{x}^{(n)^T}\right]^T$ and $Y_s=\left[y^{(1)},\ldots,y^{(n)}\right]^T$. The goal is to build an approximating model of g using the n-sample (X_s,Y_s).

The number n of simulations is a compromise between the CPU time required for each simulation and the number of input parameters. Some thumb rules propose to choose n at least as large as 10 times the dimension d of the input vector [12]. For the SFD type, a Latin Hypercube Sample (LHS) with optimal space-filling and good projection properties [6, 18] would be well adapted. In particular, ensuring good low-order sub-projection properties is important. Maximum projection designs or low-centered L^2 discrepancy LHS are then particularly well-suited.

2.2 Step 2: Initial screening

From the learning sample, an initial screening procedure [9, 18] is performed in order to identify the PII and sort them by decreasing order of influence. For this, the approach based on dependence measures is particularly relevant.

[3] and more recently [5] have proposed to use dependence measures for screening purpose, by applying them directly on a SFD. These sensitivity indices are not the classical ones variance-based measures as they consider higher order information about the output behavior in order to provide more detailed information. Among them, the Hilbert-Schmidt independence criterion (HSIC) introduced by [8] builds upon kernel-based approaches for detecting dependence, and more particularly on cross-covariance operators in reproducing kernel Hilbert spaces (RKHS). [3] has introduced a normalized version of the HSIC which provides a sensitivity index of X_k, that can be estimated by using a simple Monte Carlo estimator. As in [5, 10], from the estimated R^2_{HSIC}, we can then use independence tests for the screening purpose. The objective is to separate the inputs into two sub-groups, the significant ones and the non-significant ones. At the end of the screening step, the inputs selected as significant are also ordered by decreasing R^2_{HSIC}. This order will be used for the sequential metamodel building in step 3.

2.3 Step 3: Joint GP metamodel with sequential building process

Among all the metamodel-based solutions (polynomials, polynomial chaos, neural networks, etc.), we focus our attention on the Gp regression, which extends the kriging principles of geostatistics to computer experiments by considering the correlation between two responses of a computer code depending on the distance between input variables [17]. The Gp-based metamodel presents

some advantages compared to other metamodels: exact interpolation property, simple analytical formulations of the predictor, availability of the mean squared error of the predictions and the proved efficiency of the model [7].

However, fitting a predictive Gp on high dimensional numerical codes (several tens of inputs) remains an issue. Indeed, it often implies the estimation of several hyperparameters involved in the covariance function of the Gp (e.g. usual case of anisotropic stationary covariance function). Therefore, some difficulties arise from the parameter estimation procedure (instability, high number of hyperparameters, see [12] for example). To tackle this issue, we propose a progressive estimation procedure which combines the result of the previous screening step and a joint Gp approach introduced in [11]. The interest of the initial screening step is twofold: to reduce the input space on which each component of the joint GP is built and to sort the selected explanatory inputs in order to enable a sequential building process. It is expected that these two uses of screening results could significantly make the joint Gp building easier and more efficient.

This approach is described in details in [10]. To summarize, it consists in:

1. **A successive inclusion of explanatory variables**

   At the end of the screening step, the inputs selected as significant are ordered by decreasing influence. The sorted inputs thus obtained are successively included in the group of explanatory inputs that are taken into account in the Gp metamodel. At the j^{th} iteration, only the j first sorted inputs are considered as explanatory input variables in the Gp metamodel, while all the remaining inputs are included in a single macro-parameter. This macro-parameter is considered as an uncontrollable parameter (i.e. a stochastic parameter, notion detailed in the next subsection). In order to improve the robustness of the optimization process, the estimated hyperparameters obtained at the $(j-1)^{th}$ iteration are used, as starting points for the optimization algorithm. This procedure is repeated until the inclusion of all the significant input variables. Note that this sequential process is directly adapted from the one proposed by [12].

2. **Building a joint Gp metamodel**

   [11] has proposed to both model the mean and dispersion of the output of a stochastic code by two interlinked Gp models, called "joint Gp". [10] has applied this approach by considering the group of non-explanatory variables as the uncontrollable parameter denoted $\mathbf{X}_\varepsilon$. More precisely, the input variables $\mathbf{X} = (X_1, \ldots, X_d)$ are divided in two subgroups: the explanatory ones denoted $\mathbf{X_{exp}}$ and the others denoted $\mathbf{X}_\varepsilon$. The output is thus defined by $y = g(\mathbf{X_{exp}}, \mathbf{X}_\varepsilon)$. Under this hypothesis, the joint metamodeling approach yields building two metamodels, one for the mean Y_m and another for the dispersion component Y_d:

   $$Y_m(\mathbf{X_{exp}}) = \mathbb{E}(Y|\mathbf{X_{exp}}) \,, \tag{2}$$

   $$Y_d(\mathbf{X_{exp}}) = \mathrm{Var}(Y|\mathbf{X_{exp}}) = \mathbb{E}\left[(Y - Y_m(\mathbf{X_{exp}}))^2|\mathbf{X_{exp}}\right] . \tag{3}$$

   To fit these mean and dispersion components, we use the methodology proposed by [11, 10].

3. **Assessing the metamodel accuracy**

   To evaluate the accuracy of a metamodel, we use the predictivity coefficient Q^2:

$$Q^2 = 1 - \frac{\sum_{i=1}^{n_{\text{test}}} \left(y^{(i)} - \hat{y}^{(i)} \right)^2}{\sum_{i=1}^{n_{\text{test}}} \left(y^{(i)} - \frac{1}{n_{\text{test}}} \sum_{i=1}^{n_{\text{test}}} y^{(i)} \right)^2} \tag{4}$$

   where $(x^{(i)})_{1 \leq i \leq n_{\text{test}}}$ is a test sample, $(y^{(i)})_{1 \leq i \leq n_{\text{test}}}$ are the corresponding observed outputs and $(\hat{y}^{(i)})_{1 \leq i \leq n_{\text{test}}}$ are the metamodel predictions. Q^2 corresponds to the coefficient of determination in prediction and can be computed on a test sample independent from the learning sample or by cross-validation on the learning sample. The closer to one the Q^2, the better the accuracy of the metamodel.

   To evaluate the accuracy of the dispersion part of a joint metamodel, we use the graphical tool introduced in [11]. It consists in evaluating the proportions of observations that lie within the α-theoretical confidence intervals which are built from the Gp model on the mean component Y_m while including a heteroscedastic nugget effect. This heteroscedastic nugget effect is modeled and predicted by the Gp metamodel built on the dispersion component Y_d. We can visualize the proportions (i.e. the observed confidence intervals) against the α-theoretical confidence interval. By definition, if a model is suited for both mean and dispersion modeling, the points should be located around the $y = x$ line. As a consequence, this plot is useful to compare the goodness of fit for the different models.

2.4 Step 4: Quantile estimation using the Gp metamodel

As already said in the introduction, a well known approach for the uncertainty analysis of cpu-time expensive computer models consists in replacing the complex computer code by a a metamodel [6]. However, the estimation of a low or high quantile from a direct use of the metamodel predictor tends to be substantially different from the full computer model quantile because the metamodel is usually constructed by smoothing the computer model output values (see an analysis of this phenomenon in [1]). For this reason, [15] has proposed to use the Gp metamodel for the quantile estimation problem, as it gives a probabilistic-type metamodel. It includes the mean which gives a predictor, the variance which gives a local indicator of prediction accuracy and the covariance which defines the full law of the Gp. Stochastic simulations (conditionally to the learning sample) of the Gp metamodel are then possible. [16] has proved the efficiency of this approach in order to estimate high quantile (of the order of 95%). A great advantage of this approach is that it provides confidence intervals for the quantile estimate, in addition to the quantile estimation.

In this paper, we will compare this quantile estimation process on two Gp metamodels:

- The simple Gp metamodel (noted sGp) built for Y_m and considering a homoscedastic nugget effect. Gp conditional simulations are then done using the standard method [2];
- The joint Gp metamodel (noted jGp) which uses a heteroscedastic nugget effect predicted by the Gp metamodel on Y_d. In this case, a new technique is required:

1. The Gp for Y_m (noted jGpm) is predicted,
2. The Gp for Y_d (noted jGpd) is predicted and provides the heteroscedastic nugget effect which is added to the diagonal of the covariance matrix of jGpm,
3. The conditional simulations are then done using the standard method [2].

3. LOCA APPLICATION

3.1 Description of the use-case

Our use-case consists in thermal-hydraulic computer experiments, typically used in support of regulatory work and nuclear power plant design and operation. Indeed, some safety analysis considers the so-called "Loss Of Coolant Accident" (LOCA), which takes into account a double-ended guillotine break with a specific size piping rupture. It is modeled with code CATHARE2 (V2.5_3mod3.1) which simulated the thermal-hydraulic responses during a LOCA in a Pressurized water Reactor [13].

In this use-case, 27 scalar input variables of CATHARE2 are uncertain, defined by their probability density function (uniform, log-uniform, normal or log-normal). They correspond to various system parameters as initial conditions, boundary conditions, some critical flowrates, interfacial friction coefficients, condensation coefficients, ... The output variable of interest is a single scalar which is the maximal peak cladding temperature (PCT) during the accident transient. Our objective with this use-case is to provide a good metamodel to the safety engineers. Indeed, the cpu-time cost of this computer code is too important to develop all the statistical analysis required in a safety study only using direct calculations of the computer code. A metamodel would allow to develop more complete and robust demonstration.

3.2 Design of experiments

$n = 500$ CATHARE2 simulations of this test case are performed following a space-filling LHS with good projection properties as the design of experiments. The input values have been sampled following their prior distributions defined on their variation ranges. Indeed, as we are not ensured to be able to build a sufficiently accurate metamodel, we prefer to sample the inputs following the probabilistic distributions in order to have at least a probabilized sample of the uncertain output, on which statistical characteristics could be estimated. Moreover, dependence measures can be directly estimated on this sample, providing first usable results of sensitivity analysis.

The obtained inputs-output sample constitutes the learning sample. The figure 1 contains the histogram of the PCT (temperature is in °C).

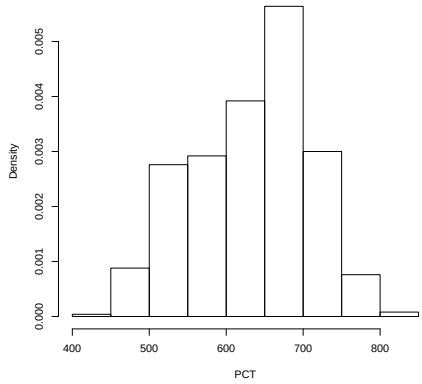

Figure 1 – Histogram of the PCT from the learning sample.

For this exercise, 600 remaining simulations, not included in the learning sample, are available. Therefore, this sample of $n_{\text{test}} = 600$ point will be our test sample, used to validate the results. In practical cases, if no test sample is available, it is possible to use a cross-validation method to perform the validation procedure [12].

3.3 Screening step with HSIC

From the learning sample of $n = 500$ simulations, R^2_{HSIC} dependence measures are estimated and bootstrap tests with $\alpha = 0.1$ are performed. 11 inputs are selected as significantly influential. Ordering them by decreasing R^2_{HSIC} reveals the predominance influence of X_{10} ($R^2_{\text{HSIC}} \approx 0.39$), followed by X_2, X_{12} and X_{22}, ($R^2_{\text{HSIC}} \approx 0.04$, 0.02 and 0.02 respectively). X_{15}, X_{13}, X_9, X_5, X_{14}, X_{26} and X_{27} have a lower influence (R^2_{HSIC} around 0.01) and the others variables are considered as negligible by statistical tests.

In the next steps, the 11 significant inputs are considered as the explanatory variables, denoted PII, in the joint metamodel and will be successively included in the building process. The other 16 variables will be joined in a so-called *uncontrollable* parameter.

3.4 Joint Gp fitting

As described in Section 2.4, two Gp metamodels are built from the learning sample of $n = 500$:

- The simple Gp metamodel sGp (just the mean component) which takes into account the residual noise in a homoscedastic nugget effect (see [12] for example);
- The joint GP metamodel jGp, using the method described in Section 2.3.

The 11 PII identified at the end of the screening step are considered as the explanatory variables while the 16 others are considered as an unique uncontrollable parameter. The Gp on the mean component (and on the dispersion component for the joint Gp) are built using the sequential building process where PII ordered by decreasing R^2_{HSIC} are successively included in Gp.

The Q^2 predictivity coefficient is computed thanks to the test sample. We obtain the same values for sGp and for the mean component jGpm of jGp:

$$Q^2 \approx 0.90 \, .$$

Thus, only 10% of the output variability remains not explained by jGpm, this includes both the inaccuracy of the jGpm (part of Y_m not fitted by Gp) and the total effect of the uncontrollable parameter, i.e. the group of non-selected inputs.

Figure 2 gives the results obtained with the confidence interval predicted by the simple Gp (homoscedatic nugget effect) and the joint Gp (heteroscedatic nugget effect) modeling. It clearly illustrates both the interest of considering a heteroscedatic nugget effect and the efficiency of using a joint Gp model to fit and predict this nugget. Indeed, it can be seen that the joint Gp is the most accurate model: all its points are close to the theoretical $y = x$ line, while the simple Gp tends to give too large confidence intervals. Thus, in this case, the heteroscedasticity hypothesis is justified and, consequently, the proposed joint Gp model is clearly more competitive than the simple one.

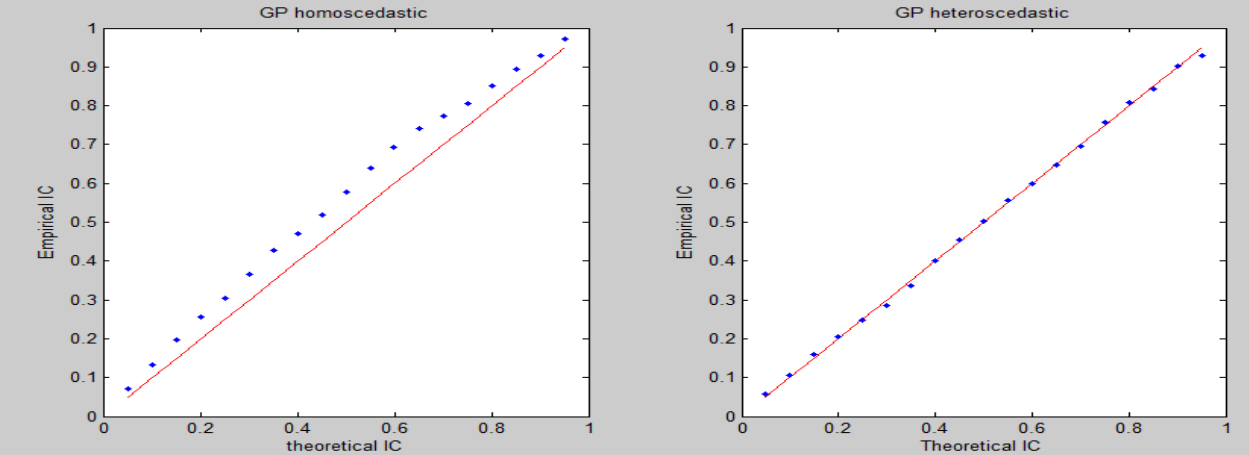

Figure 2 – Proportion of observations that lie within the α theoretical confidence interval as a function of the confidence level α for the data.

3.5 Quantile estimation results

In this section, we estimate the 95%-quantile of the PCT by our Gp metamodel and compare the results with other methods. The reference value of the 95%-quantile is obtained from a the full sample of size $n = 1100$ (learning sample plus test sample) from the classical empirical quantile estimator:

$$\hat{q}_{95}^{\text{ref}} = 742.28 \ ^\circ\text{C} \, .$$

By just using the learning sample of size $n = 500$, we obtain

$$\hat{q}_{95}^{\text{emp}} = 746.80 \ ^\circ\text{C} \, .$$

The bootstrap method (see for example [20]) allows to obtain a 90%-confidence interval for the true quantile: $[736.706; 747.4076]$.

We estimate now the 95%-quantile by a direct use of the two metamodels obtained previously. This is called the plug-in (noted pi) estimators. We obtain:

$$\hat{q}_{95}^{\text{sGp-pi}} = 736.26\ ^{\circ}\text{C}\ ,$$

and

$$\hat{q}_{95}^{\text{jGp-pi}} = 735.83\ ^{\circ}\text{C}\ .$$

No confidence intervals are obtained using this estimation method. As expected, the plug-in method provides nasty estimations of the quantile, which is here strongly underestimated.

We apply now the Gp conditional simulation method described in Section 2.4. We obtain:

$$\hat{q}_{95}^{\text{sGp}} = 747.11\ ^{\circ}\text{C}$$

with $[742.93; 751.32]$ as a 90%-confidence interval for the true quantile, and

$$\hat{q}_{95}^{\text{jGp}} = 741.46\ ^{\circ}\text{C}$$

with $[738.76; 744.17]$ as a 90%-confidence interval for the true quantile. This last result shows that the best precision of the quantile estimation is obtained with conditional simulation method based on the joint Gp metamodel (heteroscedastic Gp). This approach yields a more accurate prediction than the simple Gp metamodel based on homoscedatic hypothesis and outperforms the empirical estimator in terms of confidence interval.

4. CONCLUSION

Using an efficient sequential building process, we fitted a predictive joint Gp metamodel on a high dimensional thermal-hydraulic test case simulating accidental scenario in a Pressurized Water Reactor (LOCA test case). An initial screening step based on advanced dependence measures and associated statistical tests enabled to identify a group of influential inputs, allowing dimension reduction. The efforts of optimization when fitting the metamodel fitting can be concentrated on the main influential inputs and the robustness of metamodeling is thus increased. Moreover, thanks to the joint metamodel approach, the non-selected inputs are not completely removed: the residual uncertainty due to dimension reduction is integrated in the metamodel and the global influence of non-selected inputs is so controlled.

In nuclear safety, methods of conservative computation of quantiles [14] have been largely studied. However, several complementary information are often useful and are not accessible in a high-dimensional context. Then, we expect that the joint Gp metamodel could help to access this information by complementary sensitivity analysis. Moreover, the uncertainty of the influential inputs will be directly and accurately propagated through the mean component of the joint metamodel while a confidence bound could be derived from the dispersion component in order to take into account the residual uncertainty of the other inputs.

From this joint Gp metamodel, several statistical analyses, not feasible with the numerical model due to its computational cost, become accessible. For example, a sensitivity analysis based on

variance decomposition can be performed using the joint Gp [11]. In this paper, we have used the joint Gp metamodel to perform uncertainty propagation for the estimation of a 95%-quantile. Indeed, in LOCA test cases, we are particularly interested by the estimation of high quantile of the model output temperature. We have demonstrated the interest and advantages in terms of precision of a heteroscedastic joint Gp metamodel compared to simpler model which does not accurately capture the residual errors.

5. ACKNOWLEDGMENTS

We are grateful to Henri Geiser and Thibault Delage who performed the computations of the CATHARE code.

6. REFERENCES

[1] C. Cannamela, J. Garnier, and B. Iooss. Controlled stratification for quantile estimation. *Annals of Apllied Statistics*, 2:1554–1580, 2008.

[2] J-P. Chilès and P. Delfiner. *Geostatistics: Modeling spatial uncertainty*. Wiley, New-York, 1999.

[3] S. Da Veiga. Global sensitivity analysis with dependence measures. *Journal of Statistical Computation and Simulation*, 85:1283–1305, 2015.

[4] A. de Crécy, P. Bazin, H. Glaeser, T. Skorek, J. Joucla, P. Probst, K. Fujioka, B.D. Chung, D.Y. Oh, M. Kyncl, R. Pernica, J. Macek, R. Meca, R. Macian, F. D'Auria, A. Petruzzi, L. Batet, M. Perez, and F. Reventos. Uncertainty and sensitivity analysis of the LOFT L2-5 test: Results of the BEMUSE programme. *Nuclear Engineering and Design*, 12:3561–3578, 2008.

[5] M. De Lozzo and A. Marrel. New improvements in the use of dependence measures for sensitivity analysis and screening. *Journal of Statistical Computation and Simulation*, 86:3038–3058, 2016.

[6] K-T. Fang, R. Li, and A. Sudjianto. *Design and modeling for computer experiments*. Chapman & Hall/CRC, 2006.

[7] R. Ghanem, D. Higdon, and H. Owhadi, editors. *Springer Handbook on Uncertainty Quantification*. Springer, 2017.

[8] G. Gretton, O. Bousquet, A. Smola, and B. Schölkopf. Measuring statistical dependence with hilbert-schmidt norms. In *Proceedings Algorithmic Learning Theory*, pages 63–77. Springer-Verlag, 2005.

[9] B. Iooss and P. Lemaître. A review on global sensitivity analysis methods. In C. Meloni and G. Dellino, editors, *Uncertainty management in Simulation-Optimization of Complex Systems: Algorithms and Applications*, pages 101–122. Springer, 2015.

[10] B. Iooss and A. Marrel. An efficient methodology for the analysis and metamodeling of computer experiments with large number of inputs. In *Proceedings of UNCECOMP 2017 Conference*, Rhodes Island, Greece, june 2017.

[11] A. Marrel, B. Iooss, S. Da Veiga, and M. Ribatet. Global sensitivity analysis of stochastic computer models with joint metamodels. *Statistics and Computing*, 22:833–847, 2012.

[12] A. Marrel, B. Iooss, F. Van Dorpe, and E. Volkova. An efficient methodology for modeling complex computer codes with Gaussian processes. *Computational Statistics and Data Analysis*, 52:4731–4744, 2008.

[13] P. Mazgaj, J-L. Vacher, and S. Carnevali. Comparison of CATHARE results with the experimental results of cold leg intermediate break LOCA obtained during ROSA-2/LSTF test 7. *EPJ Nuclear Sciences & Technology*, 2(1), 2016.

[14] W.T. Nutt and G.B. Wallis. Evaluation of nuclear safety from the outputs of computer codes in the presence of uncertainties. *Reliability Engineering and System Safety*, 83:57–77, 2004.

[15] J.E. Oakley. Estimating percentiles of uncertain computer code outputs. *Applied Statistics*, 53:83–93, 2004.

[16] B. Rutherford. A response-modeling alternative to surrogate models for support in computational analyses. *Reliability Engineering and System Safety*, 91:1322–1330, 2006.

[17] T. Santner, B. Williams, and W. Notz. *The design and analysis of computer experiments*. Springer, 2003.

[18] D. Woods and S. Lewis. Design of experiments for screening. In R. Ghanem, D. Higdon, and H. Owhadi, editors, *Springer Handbook on Uncertainty Quantification*, pages 1143–1185. Springer, 2017.

[19] I. Zabalza, J. Dejean, and D. Collombier. Prediction and density estimation of a horizontal well productivity index using generalized linear models. In *ECMOR VI, Peebles*, september 1998.

[20] E. Zio and F. Di Maio. Bootstrap and order statistics for quantifying thermal-hydraulic code uncertainties in the estimation of safety margins. *Science and Technology of Nuclear Installations*, 2008, Article ID 340164, 9 pages, DOI:10.1155/2008/340164, 2008.