On the Connection Between Sensitivity and Identifiability for Inverse Uncertainty Quantification

Xu Wu,[a] Koroush Shirvan,[a] Tomasz Kozlowski[b]

[a]*Department of Nuclear Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, xuwu@mit.edu, kshirvan@mit.edu*

[b]*Department of Nuclear, Plasma and Radiological Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA, txk@illinois.edu*

INTRODUCTION

Inverse Uncertainty Quantification (UQ) [1-3], or Bayesian calibration [4], is the process to quantify the uncertainties of random input parameters that are most consistent with the physical observations, the computer model and any prior beliefs gained through previous experiments or expert judgments. Inverse UQ is crucial to forward uncertainty propagation, sensitivity and validation studies which all suffer from the "lack of input uncertainty information" issue.

Most inverse UQ related work follow the seminal work of Kennedy and O'Hagan [4], which is comprehensive in that it accounts for many sources of uncertainties, especially model discrepancy. **Model discrepancy**, also referred to as *model inadequacy*, *model uncertainty* or *model bias/error*, is due to insufficient or inaccurate underlying physics, numerical approximation errors, and/or other inaccuracies that would exist even if all the parameters in the computer model were known [5, 6].

The introduction of model discrepancy term is significant because "over-fitting" can theoretically be avoided [5, 7]. But it also poses challenges in the practical applications. One of the mostly concerned and unresolved problem is the "lack of identifiability" issue [8]. **Identifiability** refers to whether the true values of calibration parameters can theoretically be inferred based on given data [6]. With the presence of model discrepancy, inverse UQ becomes "non-identifiable" in the sense that it is difficult to precisely estimate the calibration parameters and to distinguish between the effects of parameter uncertainties and model discrepancy.

The identifiability issue is caused by the **confounding** [5, 6, 9, 10] between the calibration parameters and the model discrepancy. Several different combinations of the computer model (with different values for the calibration parameters) and its corresponding model discrepancy might result in equally good agreement with the measurement data and equally high values for the likelihood function during Markov chain Monte Carlo (MCMC) sampling. Previous research to alleviate the non-identifiability issue focused on using informative priors [5, 9, 10] for the calibration parameters and the model discrepancy, which is not a viable solution because one rarely has such accurate and informative prior knowledge.

In this work, we will show that *identifiability is largely related to the sensitivity of the calibration parameters w.r.t. the chosen Quantity-of-Interest (QoI)*. We used an improved modular Bayesian approach developed in [7, 11] for inverse UQ. The relationship between sensitivity and identifiability has been demonstrated with a practical example in nuclear engineering, which is the inverse UQ of TRACE physical model parameters using OECD/NEA NUPEC BWR Full-size Fine-Mesh Bundle Tests (BFBT) benchmark steady-state void fraction data. It is proved that, in order for a certain calibration parameter to be statistically identifiable, it should be significant to at least one of the QoIs used for inverse UQ.

INVERSE UQ FORMULATION

In an earlier ANS summary [7], we have introduced the theory and application of inverse UQ. In this section, we will briefly review the Bayesian formulation for inverse UQ. Consider a forward computer model $\mathbf{y}^{\mathrm{M}} = \mathbf{y}^{\mathrm{M}}(\mathbf{x}, \boldsymbol{\theta})$ where $\mathbf{y}^{\mathrm{M}}$ is the QoI, $\mathbf{x}$ is the vector of design variables, and $\boldsymbol{\theta}$ is the vector of calibration parameters. See [11] for detailed discussion on the classification of input parameters. If we use $\mathbf{y}^{\mathrm{E}}(\mathbf{x})$ to denote the physically measured QoI, we have the following "model updating equation" [4, 6]:

$$\mathbf{y}^{\mathrm{E}}(\mathbf{x}) = \mathbf{y}^{\mathrm{M}}(\mathbf{x}, \boldsymbol{\theta}^*) + \delta(\mathbf{x}) + \boldsymbol{\epsilon} \tag{1}$$

where $\boldsymbol{\theta}^*$ is the "true" but unknown value for the calibration parameters, the learning of which is the goal of inverse UQ process. $\delta(\mathbf{x})$ is the model discrepancy/bias. Finally $\boldsymbol{\epsilon}$ represents the measurement error.

The model updating equation serves as the starting point of inverse UQ. The model discrepancy term was first addressed in the seminal work of Kennedy and O'Hagan [4]. It is important to consider model discrepancy as otherwise we would have an unrealistic level of confidence in the computer simulations [5]. Based on the model updating equation and the Gaussian assumption of the experiment uncertainty, $\boldsymbol{\epsilon} = \mathbf{y}^{\mathrm{E}}(\mathbf{x}) - \mathbf{y}^{\mathrm{M}}(\mathbf{x}, \boldsymbol{\theta}^*) - \delta(\mathbf{x})$ follows a multi-dimensional Gaussian distribution. The posterior can be written as:

$$p\left(\boldsymbol{\theta}^*|\mathbf{y}^{\mathrm{E}}, \mathbf{y}^{\mathrm{M}}\right) \propto p\left(\boldsymbol{\theta}^*\right) \cdot \frac{1}{\sqrt{|\boldsymbol{\Sigma}|}} \cdot \exp\left[-\frac{1}{2}\left[\mathbf{y}^{\mathrm{E}} - \mathbf{y}^{\mathrm{M}} - \delta\right]^{\top} \boldsymbol{\Sigma}^{-1} \left[\mathbf{y}^{\mathrm{E}} - \mathbf{y}^{\mathrm{M}} - \delta\right]\right] \tag{2}$$

where $\boldsymbol{\Sigma}$ is the covariance of the likelihood. For a complete and detailed discussion of the inverse UQ formulation, see [7, 11]. The posterior function can be explored by MCMC sampling to obtain the posterior distributions.

Due to the page limit of this summary, the detailed solution process for inverse UQ will not be discussed. Interested readers can refer to [11] for detailed introduction, comparisons and limitations of full Bayesian and modular Bayesian approaches. Both full and modular Bayesian approaches suffer greatly from the "identifiability" issue. The most concerned and unresolved problem is to select an appropriate prior for model discrepancy. In the same reference [11] we also proposed an improved modular Bayesian approach that does not require priors for the model discrepancy function. But this does not mean that it is capable of bypassing the identifiability issue. In this work, we will show that identifiability is largely related to the sensitivity (or significance) of the calibration parameters w.r.t. the chosen QoIs.

To justify this claim, we designed an inverse UQ numerical tests whose calibration parameters have distinct sensitivities for the QoIs. This test consists of five calibration parameters and four QoIs. Each QoI has different significant contributors. By performing inverse UQ with different combinations of the QoIs and looking at the corresponding posteriors, a connection between sensitivity and identifiability can be established. We will detect the goodness of identifiability by the posterior standard deviations following previous work [6, 8], with small standard deviations indicating good identifiability.

NUMERICAL TEST

The numerical test is the inverse UQ of TRACE physical model parameters using BFBT benchmark steady-state void fraction data. This test has been used for demonstration of inverse UQ [2, 7]. Details of BFBT benchmark can be found in [12]. In this summary, we will not provide such details to avoid detour in the flow of the narrative. In brief, the QoIs are `VoidF1`, `VoidF2`, `VoidF3` and `VoidF4`, which are cross-sectional averaged void fractions measured at lower to upper axial locations. The design variables $\mathbf{x}$ consists of four parameters: pressure, mass flow rate, power and inlet temperature. The calibration parameters $\boldsymbol{\theta}$ represents five uncertain physical model parameters, including `P1008`, `P1012`, `P1022`, `P1028` and `P1029`, as shown in Table I.

Table I. Selected TRACE Physical Model Parameters

Parameter (multiplication factors)	Symbol
Single phase liquid to wall HTC	`P1008`
Subcooled boiling HTC	`P1012`
Wall drag coefficient	`P1022`
Interfacial drag (bubbly/slug Rod Bundle - Bestion) coefficient	`P1028`
Interfacial drag (bubbly/slug Vessel) coefficient	`P1029`

All of the five calibration parameters are multiplicative factors with nominal values of 1.0. Note that TRACE includes options for user access to 36 physical model parameters [13]. Those reported in Table I are selected after dimensional reduction using local/global Sensitivity Analysis (SA). The removed 31 physical model parameters are either inactive or have negligible importance for BFBT benchmark. See [2] for the detailed dimensional reduction process. The priors are chosen as uniform distributions over the range of $(0, 5)$ for all the parameters.

Results for Sensitivity Analysis

In this work we use Sobol' indices to represent sensitivity. Sobol' indices [14] measures the percentage of the variances (or uncertainties) in QoIs that can be apportioned to each of the input parameters or their combinations, which a very straightforward measure of sensitivity. Table II shows the Sobol' indices of each parameter for each QoI and Figure 1 visualizes the results. Main effect represents the standalone influence of a certain input on the QoI, while total effect also accounts for its interaction with other inputs. Taking into account the physical meaning of each parameter, the observed sensitivity can be described as below:

1. The significance of `P1008` decreases to almost zero at higher elevations. This is because single-phase liquid exists only in the lower elevations of the bundle.
2. Similarly, `P1012` is more important at lower elevations because that is where subcooled boiling occurs.
3. `P1022` increases at higher elevations, while `P1028` dominates at intermediate elevations.
4. `P1029` is only important for `VoidF4`.

The fact that a certain input has close main and total effects means that this input has no interaction with others (e.g. `P1022`, `P1028` and `P1029`). If the total effect is larger than the main effect, this input has interaction with others (e.g. `P1008` and `P1012`), the degree of which depends on the difference between main and total effects. Another convenient way to detect possible interactions between inputs is to look at the sum of main or total effects. If main effects of all the inputs sums to a value less than but close to 1.0, or if the sum of total effects results in a value greater than but close to 1.0, it can be concluded that the interactions are negligible. Table II (last row) and Figure 1 indicate that noticeable interactions only exist between `P1008` and `P1012` for `VoidF1` and `VoidF2`.

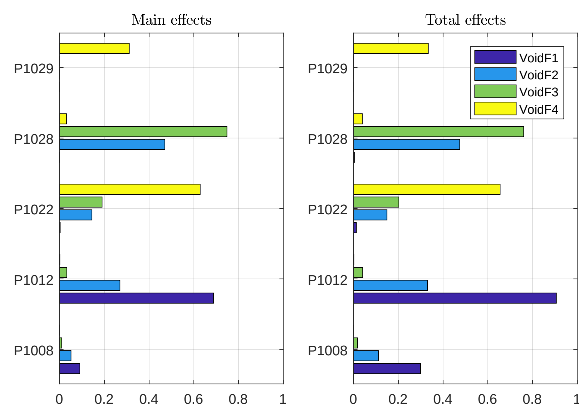

Fig. 1. Sobol' indices (main and total effects) for selected physical model parameters.

Based on the SA results, we can see that the current numerical test is ideal for the investigation of the relationship between sensitivity and identifiability. There are multiple inputs and QoIs, and each QoI has different significant contributors (calibration parameters). For instance, (1) `VoidF1` has `P1008` and `P1012`, (2) `VoidF2` has `P1008`, `P1012`,`P1022` and `P1028`, (3) `VoidF3` has `P1022` and `P1028` (`P1012` is negligible), (4) `VoidF4` has `P1022` and `P1029` (`P1028` is negligible). In the inverse UQ study, we can use data from different combinations of the QoIs and see if the corresponding significant inputs are identifiable. In this way we can establish a connection between sensitivity and identifiability.

Table II. Main and Total Effects Sobol' Indices From Sensitivity Analysis.

Output	Main effect Sobol' indices				Total effect Sobol' indices			
	`VoidF1`	`VoidF2`	`VoidF3`	`VoidF4`	`VoidF1`	`VoidF2`	`VoidF3`	`VoidF4`
`P1008`	0.0903	0.0507	0.0093	0.0008	0.2989	0.1107	0.0176	0.0009
`P1012`	0.6875	0.2696	0.0320	0.0012	0.9064	0.3307	0.0404	0.0013
`P1022`	0.0022	0.1442	0.1896	0.6285	0.0119	0.1490	0.2022	0.6551
`P1028`	0.0006	0.4705	0.7483	0.0299	0.0031	0.4747	0.7607	0.0388
`P1029`	0.0000	0.0000	0.0000	0.3112	0.0000	0.0000	0.0000	0.3338
Sum	0.7806	0.9350	0.9792	0.9715	1.2203	1.0650	1.0208	1.0298

Results for Identifiability

The detailed step-by-step implementation of the improved modular Bayesian approach for this problem can be found in Chapter 7 of [13]. Therefore, in this paper we will directly present the posterior results pertinent to the identifiability issue. There are 15 different combinations of data from the four QoIs, from using only single QoI to using all of the four QoIs. Table III shows the posterior statistical moments for the five physical model parameters and Figure 2 visualizes the results. The "output" column means the combinations of outputs whose data are used for inverse UQ. For instance, output "124" means data from `VoidF1`, `VoidF2` and `VoidF4` are used.

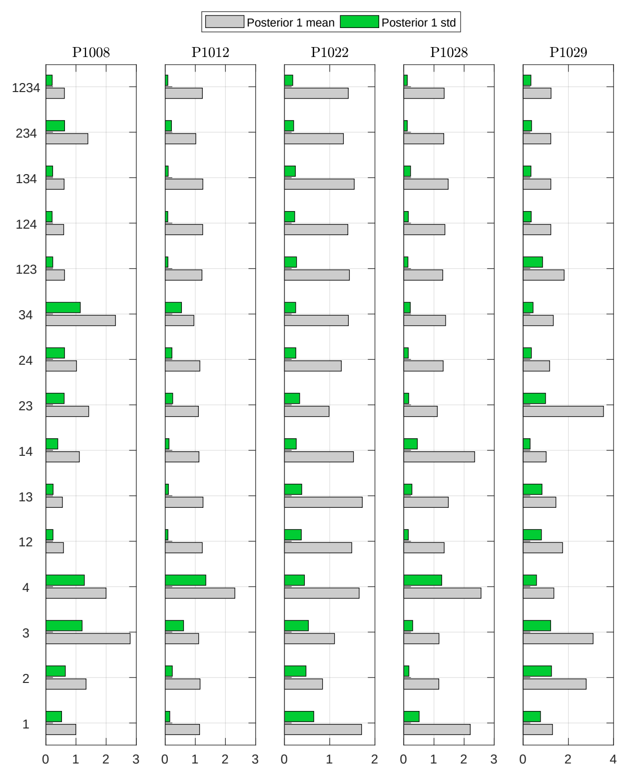

Fig. 2. Posterior statistical moments for TRACE physical model parameters.

Combining the SA results, Table III and Figure 2 demonstrate that:

1. `P1008` only has small significance for `VoidF1` (mostly through interaction with `P1012`) and even smaller importance for `VoidF2`. It can be observed that whenever `VoidF1` and `VoidF2` are included, the posterior standard deviations for `P1008` are small. For outputs "3", "4" and "34", the posterior standard deviations are large.

2. `P1012` is significant for `VoidF1` and `VoidF2`. Similar to `P1008`, the posterior standard deviations for `P1012` are small whenever `VoidF1` and `VoidF2` are present, and large with outputs "3", "4" and "34".

3. `P1022` is significant for `VoidF4`, and relatively important for `VoidF2` and `VoidF3`. As expected, its posterior standard deviation without any one of these outputs (i.e. outputs "1") is the largest.

4. `P1028` is significant for `VoidF2` and `VoidF3`. When either `VoidF2` or `VoidF3`, or both are used, the posterior standard deviations are small. Posterior standard deviations with neither of them (i.e. outputs "1", "4" and "14") are among the largest.

5. `P1029` is only significant for `VoidF4`. Therefore, its posterior standard deviations are only relatively smaller when `VoidF4` is included.

This numerical test provides solid evidence that identifiability is closely related to sensitivity, which was ignored in previous research. Good identifiability cannot be achieved for a certain calibration parameter if it is not significant to any of the QoIs whose data are used for inverse UQ. The posteriors obtained with output "1234" are considered as the most appropriate results because all four QoIs are considered, and all the five calibration parameters are close to be the most "identifiable" (with smallest standard deviations) among all the 15 combinations.

The remarkable differences in the posterior statistical moments achieved using different combinations of QoIs warns us about the risk of obtaining improper posteriors if the physical data is not properly chosen. The first risk is the bad identifiability, which is unlikely to be alleviated by numerical techniques (e.g. providing informative priors for model discrepancy) concerning inverse UQ. The second risk, however, is "fake identifiability", which refers to the case when posterior standard deviations are small enough but mean values are inappropriate. For example, with output "234", all the posterior standard deviations are small or acceptable, but the mean value of `P1008` is twice as large as the value obtain with output "1234".

Table III. Posterior Statistical Moments

Output	Mean values					Standard deviations				
	P1008	P1012	P1022	P1028	P1029	P1008	P1012	P1022	P1028	P1029
1	0.9943	1.1422	1.7042	2.2032	1.2974	0.5241	0.1539	0.6474	0.5093	0.7677
2	1.3339	1.1582	0.8413	1.1614	2.7884	0.6469	0.2378	0.4759	0.1691	1.2545
3	2.7964	1.1109	1.1060	1.1660	3.0939	1.2038	0.6112	0.5287	0.2944	1.2177
4	1.9979	2.3113	1.6510	2.5607	1.3588	1.2781	1.3494	0.4418	1.2574	0.5918
12	0.5854	1.2321	1.4871	1.3385	1.7452	0.2374	0.0921	0.3723	0.1500	0.8096
13	0.5495	1.2563	1.7204	1.4770	1.4485	0.2414	0.1091	0.3818	0.2698	0.8348
14	1.1119	1.1215	1.5256	2.3491	1.0182	0.3971	0.1285	0.2604	0.4491	0.3081
23	1.4224	1.1041	0.9852	1.1117	3.5537	0.6084	0.2510	0.3356	0.1621	0.9919
24	1.0176	1.1510	1.2568	1.3054	1.1722	0.6200	0.2249	0.2509	0.1476	0.3651
34	2.3102	0.9528	1.4135	1.3864	1.3342	1.1432	0.5437	0.2477	0.2148	0.4394
123	0.6201	1.2210	1.4362	1.2921	1.8126	0.2305	0.0920	0.2668	0.1376	0.8596
124	0.5927	1.2450	1.4012	1.3634	1.2245	0.2088	0.0873	0.2259	0.1498	0.3580
134	0.6062	1.2497	1.5421	1.4711	1.2252	0.2274	0.1007	0.2412	0.2239	0.3467
234	1.3945	1.0158	1.3036	1.3258	1.2231	0.6233	0.2099	0.2029	0.1151	0.3735
1234	0.6162	1.2358	1.4110	1.3385	1.2340	0.2113	0.0890	0.1833	0.1155	0.3453

CONCLUSIONS

In this work, we have shown that identifiability is largely related to the sensitivity of the calibration parameters w.r.t. the chosen QoIs whose data are used for inverse UQ. The numerical test demonstrated that, in order for a certain calibration parameter to be statistically identifiable, it should be significant to at least one of the QoIs. Good identifiability cannot be achieved for a certain calibration parameter if it is not significant to any of the QoIs. Therefore, to obtain good identifiability in inverse UQ activities, we recommend that the users first conduct a sensitivity study to identify the important calibration parameters for the QoIs whose data are available. Such sensitivity analysis step can provide guidance to reduce the risk of bad identifiability or fake identifiability.

REFERENCES

1. X. WU and T. KOZLOWSKI, "Inverse uncertainty quantification of reactor simulations under the Bayesian framework using surrogate models constructed by polynomial chaos expansion," *Nuclear Engineering and Design*, **313**, 29–52 (2017).
2. X. WU, T. MUI, G. HU, H. MEIDANI, and T. KOZLOWSKI, "Inverse uncertainty quantification of TRACE physical model parameters using sparse gird stochastic collocation surrogate model," *Nuclear Engineering and Design*, **319**, 185–200 (2017).
3. X. WU, T. KOZLOWSKI, and H. MEIDANI, "Kriging-based Inverse Uncertainty Quantification of Nuclear Fuel Performance Code BISON Fission Gas Release Model using Time Series Measurement Data," *Reliability Engineering & System Safety*, **169**, 422–436 (2018).
4. M. C. KENNEDY and A. O'HAGAN, "Bayesian calibration of computer models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **63**, *3*, 425–464 (2001).
5. J. BRYNJARSDÓTTIR and A. O'HAGAN, "Learning about physical parameters: The importance of model discrepancy," *Inverse Problems*, **30**, *11*, 114007 (2014).
6. P. D. ARENDT, D. W. APLEY, and W. CHEN, "Quantification of model uncertainty: Calibration, model discrepancy, and identifiability," *Journal of Mechanical Design*, **134**, *10*, 100908 (2012).
7. X. WU and T. KOZLOWSKI, "Inverse Uncertainty Quantification of TRACE Physical Model Parameters with Model Discrepancy," in "Transactions of American Nuclear Society," Washington, DC, USA, Oct. 29 - Nov. 2, 2017. (2017).
8. P. D. ARENDT, D. W. APLEY, W. CHEN, D. LAMB, and D. GORSICH, "Improving identifiability in model calibration using multiple responses," *Journal of Mechanical Design*, **134**, *10*, 100909 (2012).
9. M. J. BAYARRI, J. O. BERGER, R. PAULO, J. SACKS, J. A. CAFEO, J. CAVENDISH, C.-H. LIN, and J. TU, "A framework for validation of computer models," *Technometrics*, **49**, *2*, 138–154 (2007).
10. F. LIU, M. BAYARRI, J. BERGER, ET AL., "Modularization in Bayesian analysis, with emphasis on analysis of computer models," *Bayesian Analysis*, **4**, *1*, 119–150 (2009).
11. X. WU, T. KOZLOWSKI, H. MEIDANI, and K. SHIRVAN, "Inverse Uncertainty Quantification using the Modular Bayesian Approach based on Gaussian Process, Part 1: Theory," *(in review, submitted to Nuclear Engineering and Design, arXiv preprint arXiv:1801.01782)* (2018).
12. B. NEYKOV, F. AYDOGAN, L. HOCHREITER, K. IVANOV, H. UTSUNO, F. KASAHARA, E. SARTORI, and M. MARTIN, *NUPEC BWR full-size fine-mesh bundle test (BFBT) benchmark*, OECD/NEA, NEA/NSC/DOC(2005)5 (2005).
13. X. WU, *Metamodel-based Inverse Uncertainty Quantification of Nuclear Reactor Simulators under the Bayesian Framework*, Ph.D. thesis, University of Illinois at Urbana-Champaign (2017).
14. I. M. SOBOL, "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates," *Mathematics and Computers in Simulation*, **55**, *1*, 271–280 (2001).