23rd International Conference
ENGINEERING MECHANICS 2017
Svratka, Czech Republic, 15 – 18 May 2017

ROLE OF RANDOM FACTORS IN NONLINEAR REGRESSION: A CASE STUDY FOR ESTIMATION OF THERMOPHYSICAL PARAMETERS

E. Janouchová*, A. Kučerová**, J. Sýkora***

Abstract: *In order to design an optimal experimental setup the designers have to take into account uncertainties connected to the investigated system. The input random factors associated with for example values of loading, specimen dimensions or measurement errors influence behaviour of the system, which thus becomes also uncertain. From this point of view, the experiment design is a very important because it effects amount of information which can be obtained from the experiment. More specifically, accuracy of the identified parameters from indirect experimental measurements depends on experimental settings. In this contribution we demonstrate a role of random factors in a nonlinear model calibration on an illustrative example of one dimensional heat conduction. The thermophysical parameters such as thermal capacity and thermal conductivity are identified on a basis of noisy measurements from experiments with different setup. The experiments vary in a number of sensors and number of observed time steps. The presented statistical analysis shows dependence of the parameter estimation on the choice of measured quantities involving different uncertainties.*

Keywords: Random factors, Uncertainty quantification, Inverse problems, Nonlinear regression, Heat conduction.

1. Introduction

Thanks to extensive developments in the field of uncertainty quantification in the last decades we are enabled to simulate the nonlinear systems with uncertain input parameters. It brings advantages as for example opportunity to design optimized and robust experiments for calibrating the models of such systems (Jarušková et al., 2016).

Involving random factors into the analysis of the system requires an appropriate formulation of the problem with respect to the considered source of uncertainty. Uncertainties can be separated into two principal categories according to whether a source of nondeterminism is irreducible or reducible (Oberkampf et al., 2002; Der Kiureghian et al., 2009). This contribution focuses on epistemic (reducible, subjective, cognitive) uncertainties arising from our lack of knowledge which is supposed to be reduced by any new measurement according to the coherence of learning (Mantovan et al., 2006; Beven et al., 2007). While aleatory uncertainty expresses the inherent randomness which cannot be reduced.

An efficient experiment design provides enough amount of suitable information which allows successful estimation of the model parameters with maximal reduction of the epistemic uncertainties. The corresponding inverse problems can be solved in deterministic or stochastic way (Tarantola, 2005) but in the both cases the identification approach has to be chosen appropriately considering the involved uncertainties.

* Ing. Eliška Janouchová: Faculty of Civil Engineering, Czech Technical University in Prague; Thákurova 7/2077; 166 29, Prague; CZ, eliska.janouchova@fsv.cvut.cz

** Ing. Anna Kučerová, PhD: Faculty of Civil Engineering, Czech Technical University in Prague; Thákurova 7/2077; 166 29, Prague; CZ, anicka@cml.fsv.cvut.cz

*** Ing. Jan Sýkora, PhD: Faculty of Civil Engineering, Czech Technical University in Prague; Thákurova 7/2077; 166 29, Prague; CZ, jan.sykora.1@fsv.cvut.cz

This contribution concentrates on comparison of random factor influence on parameter estimation based on nonlinear regression (Seber et al., 1989). Distinct sources of uncertainty are included as well as different experimental conditions.

2. Numerical study

The role of uncertain parameters in estimation of the system parameters is demonstrated on an illustrative example of one dimensional heat conduction, where observable is temperature in the location x and time step t. The thermal diffusion equation

$$\rho c_p \frac{\partial T}{\partial t} = \lambda \frac{\partial^2 T}{\partial x^2} + \dot{q}, \tag{1}$$

where λ is thermal conductivity, ρ is density and c_p is thermal capacity, can be solved with using thermal diffusivity defined as $\alpha = \frac{\lambda}{\rho c_p}$ for initial condition $T = T_\infty = 10$ °C and constant flux boundary condition at $x = 0$ $q_x = -\lambda \frac{dT}{dx} = q_0 = 100\ \mathrm{W/m^2}$ as

$$T = T_\infty + \frac{q_0}{\lambda}\left[2\sqrt{\frac{\alpha t}{\pi}}\exp\left(-\frac{x^2}{4\alpha t}\right) - x\left(1 - \mathrm{erf}\frac{x}{2\sqrt{\alpha t}}\right)\right]. \tag{2}$$

The parameters to be identified are the thermal conductivity and the thermal capacity. Their true values are $\lambda = 2\ \mathrm{W.m^{-1}K^{-1}}$ and $c_p = 1000\ \mathrm{J.m^{-3}K^{-1}}$. The considered domains of the parameters in the identification procedure are intervals $[1.6, 2.4]\ \mathrm{W.m^{-1}K^{-1}}$ for λ and $[750, 1700]\ \mathrm{J.m^{-3}K^{-1}}$ for c_p.

2.1. Uncertainties

In our particular example, there are three types of uncertain inputs. The first one is connected to loading, when the heat flux is burdened by an error $\varepsilon_1$which is a normal random variable with zero mean value and standard deviation $1\ \mathrm{W/m^2}$. This uncertain input is a constant for the whole experiment regardless of the number of sensors and time steps. The second uncertainty is a systematic error ε_2 of sensors, which is an additive error modeled as a normal random variable with zero mean value and standard deviation 0.005 °C. This uncertain input is a constant for every sensor in the experiment regardless of the number of time steps. The last source of uncertainty is a measurement random error, which is also an additive error modeled as a normal random variable with zero mean value and standard deviation 0.1 °C, but its value differs for each measured temperature.

2.2. Estimation of thermophysical parameters

Uncertainty quantification is based on statistical analysis of 10^4 repetitions of the experiments simulated numerically. The synthetic experimental data are generated according to the following equation of the model modified by involving the uncertainties:

$$T_{Exp}(x,t) = T_\infty + \frac{q_0+\varepsilon_1}{k}\left[2\sqrt{\frac{\alpha t}{\pi}}\exp\left(-\frac{x^2}{4\alpha t}\right) - x\left(1 - \mathrm{erf}\frac{x}{2\sqrt{\alpha t}}\right)\right] + \varepsilon_2(x) + \varepsilon_3(x,t). \tag{3}$$

The parameters are estimated from each experiment separately by using the method of nonlinear least squares regression. The identification is based on fitting the response of the numerical model to the experimental data. This deterministic approach leads to optimising parameters so as to minimise the objective function

$$f(\lambda, c_p) = \sum_{i=1}^{nx}\sum_{j=1}^{nt}\left(T_{Exp,i,j} - T_{i,j}(\lambda, c_p)\right)^2. \tag{4}$$

The parameter estimation is provided for different numbers of sensors nx from one to 40 sensors uniformly distributed on the interval $x \in [0.02, 0.2]$ m (and different numbers of observed time steps nt from 5 to 160 steps uniformly distributed on the interval $t \in [10, 240]$ min.

Comparison of standard deviations of the parameters obtained for different combinations of nx and nt is shown in Tab. 1. As expected the standard deviations decrease for the specific nx with increasing nt as well as for the specific nt with increasing nx. The efficiency of the estimation does not dependent only on the total number of measurements $nt \cdot nx$, but the specific experiment setup plays an important role.

Tab. 1: Identified standard deviation of the parameters depending on nx and nt.

	STD λ $[\mathrm{W.m^{-1}K^{-1}}]$					STD c_p $[\mathrm{J.m^{-3}K^{-1}}]$				
$nt\backslash nx$	**1**	**5**	**10**	**20**	**40**	**1**	**5**	**10**	**20**	**40**
5	0.3106	0.0634	0.0515	0.0409	0.0325	144.63	30.22	26.92	20.09	16.24
10	0.2793	0.0477	0.0396	0.0321	0.0276	126.12	24.64	19.92	16.19	13.89
20	0.2362	0.0369	0.0318	0.0274	0.0241	135.59	18.39	15.50	13.85	11.96
40	0.1864	0.0302	0.0265	0.0241	0.0220	94.69	14.84	13.33	12.11	10.99
80	0.1377	0.0258	0.0236	0.0223	0.0214	71.64	13.08	11.67	11.26	10.67
160	0.0997	0.0234	0.0224	0.0214	0.0207	51.73	11.60	11.14	10.73	10.30

For the case of measuring by only one sensor but in different numbers of time steps the obtained marginal distributions of the parameters are depicted in Fig. 1. For $nt \leq 40$ the influence of uncertain inputs is so significant that the optimisation algorithm pulls the optimum to the bounds of the domain. In the case of the higher numbers of time steps the mean value of the thermal conductivity λ is identified well whereas the mean value of the thermal capacity c_p is not a good match. This phenomenon can be explained by the fact that the influence of the random measurement error is reduced by increasing nt but the systematic sensor error and random loading factor have the undiminished impact on the parameter estimation.

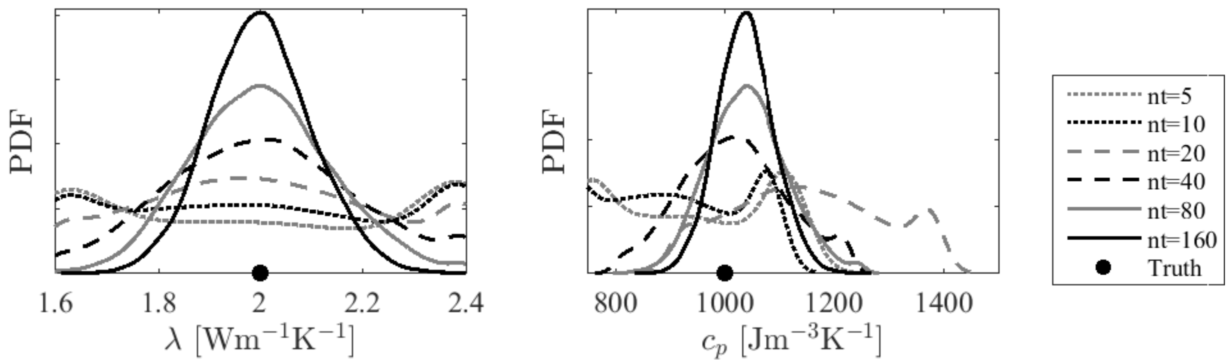

Fig. 1: Marginal distributions of λ and c_p identified from measurements obtained by only one sensor in different numbers of time steps nt.

Fig. 2 shows the marginal distributions of the parameters identified from measurements obtained by different numbers of sensors nx in 40 time steps. In this case the mean value of the thermal capacity obviously converges. Usage of several sensors enables to reduce the influence of the experimental random as well as systematic error.

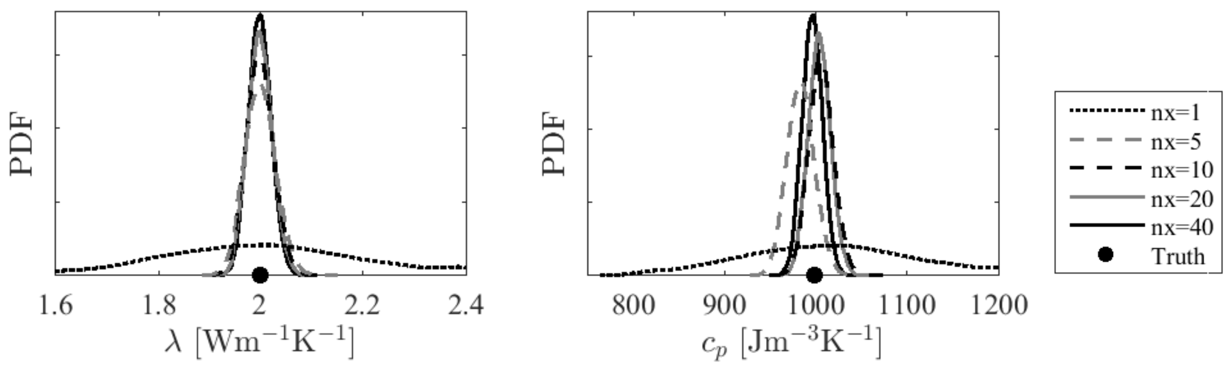

Fig. 2: Marginal distributions of λ and c_p identified from measurements obtained by different numbers of sensors nx in 40 time steps.

The error associated with loading is not reduced at all because it has one constant value for the whole experiment regardless of the number of sensors and time steps. This fact causes that the variance of the parameters asymptotically converges with an increasing number of measurements to a non zero value.

3. Conclusions

The presented analysis of random factor impact on the parameter estimation compares different sources of epistemic uncertainties and corresponding possibilities to reduce them. The uncertain system under the study is one dimensional heat conduction with random heat flux and two types of experimental errors. The study shows efficiency of the identification of thermal conductivity and thermal capacity on a basis of the different experimental setups. The temperature measured by only one sensor does not provide enough information to allow reduction of the systematic experimental error but the random experimental error is reduced increasingly with a higher number of observed values of temperature in different time steps. In order to diminish the systematic error more sensors have to be employed. The uncertainty caused by randomness of loading cannot be reduced in any considered settings. The mentioned conclusions have to be taken into account in designing an optimal experiment whose purpose is to get sufficient information for parameter identification from noisy indirect experimental data.

Acknowledgement

The financial support of the Czech Science Foundation (Projects No. 16-11473Y and 15-07299S) and the Grant Agency of the Czech Technical University in Prague (Project No. SGS17/042/OHK1/1T/11) is gratefully acknowledged.

References

Beven, K., Smith, P. and Freer, J. (2007) Comment on "Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology" by Pietro Mantovan and Ezio Todini. Journal of Hydrology, 338, pp. 315-318.

Der Kiureghian, A. and Ditlevsen, O. (2009) Aleatory or epistemic? Does it matter? Structural Safety, 31(2), pp. 105-112.

Jarušková, D. and Kučerová, A. (2016) Estimation of thermophysical parameters revisited from the point of view of nonlinear regression with random parameters. International Journal of Heat and Mass Transfer, 106, pp. 135-141.

Mantovan, P. and Todini, E. (2006) Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. Journal of Hydrology, 330, pp. 368-381.

Oberkampf, W.L., DeLand, S.M., Rutherford, B.M., Diegert, K.V. and Alvin, K.F. (2002) Error and uncertainty in modeling and simulation. Reliability Engineering and System Safety, 75, pp. 333-357.

Seber, G.A.F. and Wild, C.J. (1989) Nonlinear regression. John Wiley and Sons.

Tarantola, A. (2005) Inverse problem theory and methods for model parameter estimation. Philadelphia, PA: Society for Industrial and Applied Mathematics, 348 p.