Uncertainty and Sensitivity Analysis to Quantify the Accuracy of an Analytical Model

Anurag Goyal and R. Srinivasan

Abstract—**Accuracy of results from mathematical model describing a physical phenomenon is complicated to infer due toseveral parameters that affect the model. With the ever increasing complexity of the models, the uncertainty in model development and parameter values are increased.For an analytical model having various input variables, only a few of the parametric values are known and the remaining values are assumed as the best case values. A quantitative value for each parameter in the analytical model, ranking in importance, is required to validate the model output. In this paper, the accuracy of an analytical model is estimated quantitatively using the uncertainty and sensitivity analysis. The developed methodology was applied and analyzed for two cases, a fluid flow equation and a heat transfer model. It is shown in this paper that the accuracy can be quantitatively predicted for an analytical model and the input parameters intheir range can be effectively judged.**

Index Terms—**Model accuracy, parametric importance, sensitivity and uncertainty analysis, weight percentage.**

I. Introduction

Analytical models are used in various areas of science and engineering to simulate a process or assess the performance of a system. These models can have numerous input variables having complex relationship with each other.The accuracy of the model output depends on the accuracy of the input parameters and its range. The individual input parameters can be exact, measured, predicted or assumed.Each of these methods introduces some uncertainty in the values because of natural variation, measurement errors or lack of measurement techniques. Accuracy of results from mathematical and computer models of a system is often complicated by the presence of uncertainties in experimental data that are used to estimate parameter values [1]. It is vital to know which input parameters have significant impact on the output value of a model and these individual parameters require a closer examination. Model accuracy is also a critical parameter for robust process modeling and design calculations. It helps to ensure that the model works according to the specifications and the output is physically reasonable with a known amount of variation in the input parameters. The analysis can suggest research priorities among various uncertain parameters [2]. The model accuracy is highly correlated to the accuracy of different parameters involved in the formulation of the model and their effect on the output of the model. While modeling real physical systems, a set of assumptions and validations are often made without knowing the exact quantified impact of the individual parameter [3].

Numerous methodologies are available for estimating the parametric importance for a model qualitatively. Saltelli and Tarantola [4] and Saltelli et al. [5] discussed about the assessment of relative importance of input factors from probability distribution of output given the probability distribution of the inputs using Monte Carlo analysis. Benkeet. al. [6] combined differential erroranalysis and Monte Carlo simulation with stochastic and deterministic sensitivity analysis to determine the parameters having larger impact on model output.Principal component analysis (PCA) was used to qualitatively identify the main parameters from a multivariate data set [7]. Larsen [8] discusses some of the qualitative and quantitative parameters for data quality.

To estimate the effect of input parameters, uncertainty and sensitivity analysis provides a good platform.Bushnell [9] has put forth a simplistic “one parameter at a time”: derivative-based approach towards uncertainty and sensitivity analysis for predicting the parametric importance with respect to each individual parameter.Sensitivity analysis of a model output aims toquantify the relative importance of each input modelparameter in determining the value of an assignedoutput variable.Sensitivity and uncertaintyanalyses methods for computer models are being applied in performanceassessmentmodeling in the geologic high-level radioactive-waste repository program [10].Homma and Saltelli [11] introduced the 'total effect'parameter index while giving a methodology for global sensitivity analysis. This index provides a measure of the total effect of a givenparameter, including all the possible synergetic terms between that parameterand all the others. Rank transformation of the data is also introduced in orderto increase the reproducibility of the method.Zu and Gertner [12] proposed a regression-based method,for models with correlated inputs, toquantitatively decompose the total uncertainty in model output into partial variances contributed by the correlated variations and partialvariances contributed by the uncorrelated variations.It was also shown that the sensitivity analysis is easier for linear models than for non-linear ones, and for monotonic than non-monotonic ones [13].

However, very little literature is available to estimate the impact of input parameters quantitatively and no study was done to describe the representation of model accuracy based on the quantified input parametric data. In this paper, a methodology was developed to estimate the accuracy based on the quantified input parametric estimation and brings out the importance of each individual parameter. The parameters with the highest importance or the ones that affect the output

Manuscript received October 7, 2012; revised November 8, 2012.

A. Goyal is with the Department of Mechanical Engineering, Indian Institute of Technology, Delhi, New Delhi-110016, India (e-mail: me1080494@mech.iitd.ernet.in).

R. Srinivasan is with Nalco Technology Center, Pune-411028, India. (e-mail: sramanathan@nalco.com).

DOI: 10.7763/IJMO.2012.V2.202

of the model, the most, contribute more to model accuracy. The inputs to the model, most generally, are the input parameters at design condition values, and at maximum and minimum range values. The model accuracy prediction based on defined parameters can serve as a preliminary measure to process design and parameter control. The parametric importance and its quantitative contribution towards the model were estimated using the uncertainty and sensitivity model developed by Bushnell [9]. The methodology for estimating the model accuracy was discussed and it was applied to two test cases one in Taylor Couette flow model and the other in a heat transfer modelto analyse the model quantitatively.

II. Methodology

A. Model Accuracy

In order to bring out the significance of the model accuracy, it is vital to estimate the weight percentage of each individual design parameters contribution for a mathematical model. The model accuracy will always be a function of the weight percentage and the known value of the design parameter, and is represented as,

$$A = \sum_{i=1}^{n} a_i \times w_i \tag{1}$$

where A is the model accuracy, a_i is a constant which would be 1 if the parameter value is known else would be 0 and w_i is the weight percentage of a parameter. The significance of the constant attributes to the cases where the model prediction is based on the actual known parameter value or the assumed case values. In many design cases, it is required to estimate model calculations based on assumptions when some of the values are not known. The constant, a_i, is used for estimating the model accuracy taking into account for each parameter whether it is assumed or known. The following sections describe about the weight percentage and its estimation.

B. Weight Percentage

The weight percentage, which is significantly important to the model accuracy, is basically a measure of each design parameters contribution to the mathematical model prediction. Hence, it is an essential criterion to estimate the parametric weight percentages for the development of any model. Sensitivity and uncertainty analysis methods are numerical methods for determining the relationships among the input and output variables of a mathematical model. Most of the mathematical models have a number of input variables and the governing relationship between input and output is complex. The uncertainty and sensitivity analysis developed by Bushnell [9] was used to estimate the weight percentage with respect to each parameter. The parameter uncertainty and sensitivity are discussed in the following sections.

C. Parameter Uncertainty

Parameter uncertainty can be defined as the range of possible values for an individual parameter for that mathematical model with valid assumptions. This uncertainty can come from measurement, sampling, or estimation errors. The relative range or the uncertainty is calculated using the following equation. The relative range or the uncertainty is calculated using the following equation:

$$N_i = \frac{R_i}{\xi_{b,i}} \tag{2}$$

where N_i is uncertainty of i^{th} parameter, R_i is equal to the expected range of i^{th} parameter which is the difference between the parametric range and ξ_{bi} is equal to the base case value of i^{th} parameter. To compare parameters with widely different magnitudes, the range of each parameter was normalized with a determined base-case value of the parameter. The base-case value is the best estimate of the parameter value. In general the base case value happens to be the median of the parametric range.

D. Parameter Sensitivity

Another important parameter for estimating the weight percentage is the sensitivity. Parameter sensitivity is the amount of variation in the model output in response to changes in the parameter inputs. Minor changes in some input parameters may make considerable changes to the model results, while larger changes to other parameters may have insignificant effects on results. It can identify the parameters that have largest effect on output for model calibration by linearizing the analytical model and also which deserve the most attention, accuracy, or research during data collection. The equation for normalized sensitivity is:

$$S_i = \frac{\xi_{b,i}}{F_b}\frac{\partial F}{\partial \xi_i} \tag{3}$$

where, S_i is the normalized sensitivity of i^{th} parameter, F_b is the objective function at base case, F is the objective function/ governing equation, $\partial F/\partial \xi_i$is the partial derivative of the objective function for the parameter at base case value. The resulting magnitude of the sensitivity (S_i) indicates the effect of the input parameter on the model prediction. Positive sensitivity shows that an increase in the input value will increase the model prediction value, while a negative sensitivity shows that an increase in the input value will decrease the model prediction value. In general, the sensitivity calculated using this technique is defined as the ratio of the change in output to the change in input. Since the sensitivity equation uses differential techniques, more care should be taken for formulating the derived form and evaluating the numerical value for complex models

E. Weight Percentage Calculation

Weight percentage assessment identifies individual parameter’s contribution towards variance of the output. Parametric importance is the combined effect of uncertainty and sensitivity. A parameter that is not sensitive will not cause variance in the output even with large uncertainty, and a parameter that is highly sensitive but known precisely also will not cause variance in the output. By including both uncertainty and sensitivity, importance assessment identifies the parameters that can best reduce the output variability with better measurements, increasing the effectiveness of sensitivity analysis in all its uses. To combine uncertainty and sensitivity into a dimensionless gauge of importance (Ii), the absolute value of the product of the relative range and normalized sensitivity is taken:

$$I_i = |N_i S_i| = \left|\frac{R_i}{\xi_{b,i}}\frac{\xi_{b,i}}{F_b}\frac{\partial F}{\partial \xi_i}\right| = \left|\frac{R_i}{F_b}\frac{\partial F}{\partial \xi_i}\right| \tag{4}$$

Greater values of importance would indicate where efforts to better estimate parameters would have the most effect on providing a more accurate model prediction and risk assessment and, indicate where resources should be focused. Finally, to ascertain the weight percentage associated with each parameter, the importance value was used and weight percentage was formulated as:

$$w_i = \frac{I_i}{\sum_{i=1}^{n} I_i} \quad (5)$$

where n is equal to the number of parameters and w_i is equal to the weight percentage associated with each parameter.

III. Results

To analyze the developed model for estimating the accuracy, two case studies were performed. One on Taylor Couette flow between rotating concentric cylinders and the other, a rigorous heat transfer model, for estimating the convective heat transfer coefficient from a circular cylinder in a cross flow to liquid. Details of the case studies and the results are discussed in the following sections.

A. *Case Study A: Taylor Couette Flow Between Rotating Concentric Cylinders*

The model developed by Srinivasan et.al [14] was used for estimating the Power Number for a Taylor Couette flow between two concentric cylinders. The Power Number can be represented a

$$Po = 2\pi/(\rho N^2 D_i^4) \quad (6)$$

where Po is Power Number, Di is the internal diameter of cylinder in meters, ρ is the fluid density in kg/m3 and N is the rotational speed in revolutions per second. Water and oil are assumed to be fluid of interest for this case study and the parametric ranges are summarized in Table I. The parametric range (min and max) values are chosen with consideration of standard design conditions and properties.

TABLE I: Parametric Ranges, Uncertainty and Sensitivity Values for Case Study A

Parameter	Min	Max	N_i	S_i
Cylinder Diameter, D_i, (m)	0.01	0.025	1.00	-4.23
Density of fluid, ρ (m3/s)	912	1050	0.13	-1.05
Rotational Speed, N (rps)	0.7	142	28.26	-2.11

Based on the parametric range tabulated, the uncertainty and sensitivity was calculated using Equation 2 and 3 respectively. Table I shows the uncertainty and sensitivity values derived for each parameter using the minimum and maximum range values. As discussed in the Section IIB, the fluid density, cylinder diameter and rotational speed have a negative magnitude for sensitivity and it is due to the effect of the differential analysis techniques. The uncertainty values depend on the parametric range and the base-case values which are chosen to be the median. The uncertainty values are strongly dependent on the parametric range applied to that specific analytical model. It is not possible to generalize the uncertainty for a particular model without the known range of minimum and maximum values.

The estimation of weight percentage for different parameter is shown in Figure 1. The combined effect of the sensitivity and uncertainty can be seen in the weight percentage plot. The results show that the rotational speed has the maximum weightage and is the critical parameter for minimum and maximum range considered for this study. The dependence will change for different set of parametric range values.Figure 2shows the estimation of accuracy for each parameter assuming some base-case value for that parameter and estimating the model accuracy, which is the Power Number.

The effect of a parameter becoming constant can be seen with the weight percentage of fluid density. It was observed that since the variation in density is very less, it is almost constant and therefore contributes very less towards model accuracy. On the other hand, the parameter rotational speed varies over a large range and therefore has high levels of uncertainty and contributes more to the accuracy of the model.

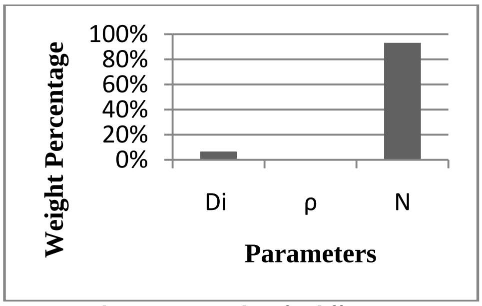

Fig. 1. Weight percentage values for different parameters

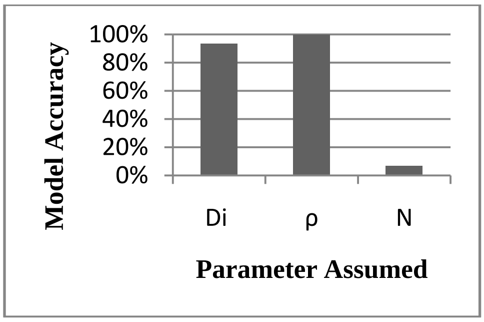

Fig. 2. Model accuracy for assumed base case values

B. *Case Study B: Heat Transfer Model*

To extend the applicability of the developed methodology to a more rigorous analytical model, heat transfer equation for predicting the convective heat transfer coefficient for cross flow over a cylinder for liquid, as given by Sanitjai and Goldstein [15], was used for the study. The heat transfer coefficient for a cross flow over a cylinder for liquid in a turbulent condition can be represented as below,

$$h = 0.037 * \rho^{0.8} * v^{0.8} * D^{-0.2} * \mu^{-0.38} * k^{0.58} * C_p^{0.42} \quad (7)$$

where ρ is the density of fluid in kg/m3, v is the velocity of fluid in m/s, D is the diameter of cylinder in meter, Cp is the specific heat capacity of fluid in J/kgK, k is the thermal conductivity in W/mK, h is the convective heat transfer coefficient in W/m^2K and μ is the coefficient of viscosity in $N\text{-}s/m^2$. The minimum and maximum parameter range values were calculated based on Reynolds number range (between 10^4 and 10^5) and Prandtl number range (between 6.5 and 176) and tabulated in Table 2. Pure component liquid properties of water and ethylene glycol at 23°C were used to estimate the parametric range values. For the base case values, the liquid mixtures are considered and the pure component properties are not considered. When the pure component properties are used as base case values, the model resulted in strong dependence for velocity constructing all the other parameters to negligible contribution. Hence based on the test analysis, mixture properties were used for base case values.

TABLE II: PARAMETRIC RANGES, UNCERTAINTY AND SENSITIVITY VALUES FOR CASE STUDY B

Parameter	Min	Max	N_i	S_i
Specific heat capacity, Cp (J/kgK)	2401.5	4181	0.51	0.43
Coefficient of viscosity, μ (N.s/m^2)	9.59E-04	1.84E-02	1.74	-0.39
Thermal Conductivity, k (W/mK)	2.51E-01	6.06E-01	1.41	0.59
Diameter of cylinder , D (m)	0.028	0.029	0.04	-0.20
Density of fluid, ρ (m^3/s)	998	1113.2	0.11	0.82
Velocity of fluid, v (m/s)	0.344	57	1.89	0.82

Based on the parametric range tabulated, the uncertainty and sensitivity values were calculated using Equation 2 and 3 respectively, and tabulated in Table II. It shows the uncertainty and sensitivity values derived for each parameter using the selected range of values. As discussed in the Section IIB, the cylinder diameter and coefficient of viscosity show a negative magnitude and it is due to the effect of the differential analysis techniques and the powers raised to negative magnitude. Also the fluid velocity and density show a similar positive magnitude compared to other parameters. The base-case values for this case are chosen to be the mean for 50% liquid mixtures. Again, it is not possible to generalize the uncertainty for a particular model without the known range of values.

The estimation of weight percentage is shown in Figure 3. The combined effect of the sensitivity and uncertainty can be seen in the weight percentage plot. The results show that the fluid velocity has more weight percentage and is a key contributing factor for the heat transfer coefficient within the specified range. Also, it shows some dependence for the model on thermal conductivity, and viscosity quantitatively. The accuracy values are estimated by using the base case values when the parameter is considered to be unknown. It can be noted that the dependence will change for different set of range of values for that specific analytical model. Figure4 shows the model accuracy variation for each parameter when it is not considered for the evaluation. The effect of a parameter becoming constant can be seen with the weight percentage of cylinder diameter.

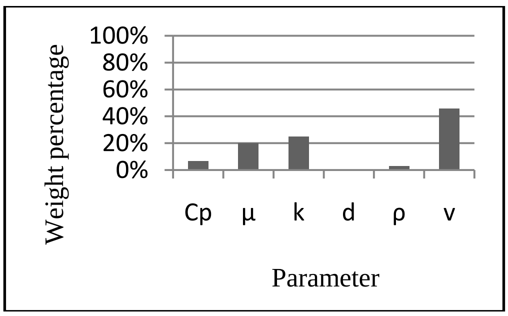

Fig. 3. Weight percentage values for different parameters

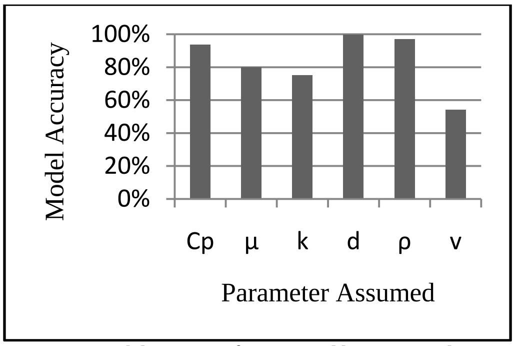

Fig. 4. Model accuracy for assumed base case values

IV. CONCLUSIONS

A methodology, based on uncertainty and sensitivity analysis, was proposed for quantitative estimation of model accuracy. The developed methodology was applied for two case studies, one for the Taylor Couette flowbetween rotating concentric cylindersand the other, a rigorous heat transfer model. It has been shown that the developed technique predicts the model accuracy quantitatively and can be used for estimating the model output when some of the parameters are assumed.

ACKNOWLEDGMENT

This work was carried out at the Nalco Technology Center, Pune, as a part of an industry-university collaborative research internship project. The authors would like to acknowledge Dr. Hari Reddy, Director, Nalco Technology Center, Pune, India for providing support to carry out this project.

REFERENCES

[1] S. Marino, I. B. Hogue, C. J. Ray, and D. E. Kirschner, “A methodology for performing global uncertainty and sensitivity analysis in systems biology,” *Journal of Theoretical Biology*, vol. 254, pp. 178-196, 2008.

[2] A. Saltelli and T. Homma, “Sensitivity analysis for model output: Performance of black box techniques on three international benchmark exercises,” *Computational Statistics and Data Analysis*, vol. 13, pp. 73-1994, 1992.

[3] A. Saltelli, “Sensitivity Analysis for Importance Assessment. Risk Analysis,” *Joint Research Centre of the European Communities in ISPRA*, 2002, vol. 22, no. 3, pp. 579-590.

[4] A. Saltelli and S. Tarantola,“On relative importance of input factors in mathematical models: the case of Level E,” *Joint Research Centre of the European Communities in ISPRA*, 2002.

[5] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana and S. Tarantola, “Global Sensitivity Analysis,” *The Primer*, John Wiley and Sons, Ltd., 2008.

[6] K. A. Benke, K. E. Lowell, and A. J. Hamilton, “Parameter uncertainty, sensitivity analysis and prediction error in a water-balance hydrological model,” *Mathematical and Computer Modelling*, vol. 47, pp. 1134-1149, 2008.

[7] N. Kuhnert, R. Jaiswal, P. Eravuchira, R. M. El-Abassy, B. von der Kammer, and A. Materny, “Scope and Limitations of principal component analysis of high resolution LC-TOF-MS data: the analysis of the chlorogenic acid fraction in green coffee beans as a case study,” *Anal. Methods*, vol. 3, pp. 144-155, 2011.

[8] A. H. Larsen, “Data Quality for Process Design,” *Fluid Phase Equilibria*, vol. 29, pp. 47-58, 1986.

[9] T. H. Bushnell, *Parameter Importance of an analytical model for transport in the vadose zone*, M. S. Thesis, Brigham Young University 2007.

[10] P. G. Doctor, “Sensitivity and uncertainty analyses for performance assessment modeling,” *Engineering Geology*, vol. 26, pp. 411-429, 1989.

[11] T. Homma and A. Saltelli, "Importance measures in global sensitivity analysis of nonlinear models," *Reliability Engineering and System Safety*, vol. 53, pp. 1-17, 1996.

[12] C. Zu and G. Z. Gertner, "Uncertainty and sensitivity analysis for models with correlated parameters," *Reliability Engineering and System Safety*, vol. 93, pp. 1563-1573, 2008.

[13] A. Saltelli, T. H. Andres, and T. Homma, "Sensitivity analysis of model output: An investigation of new techniques," *Computational Statistics and Data Analysis*, vol. 15, pp. 211-238, 1993.

[14] R. Srinivasan, S. Jayanti, and A. Kannan, "Power Number Calculations for a Taylor Couette Flow," in *Proc. 3rd IASME Int. Conf. on Fluid Dynamics and Aerodynamics*, Corfu, Greece. 20-22 August, 2005, pp. 218-223.

[15] S. Sanitjai and R. J. Goldstein, "Forced convection heat transfer from a circular cylinder in cross flow to air and fluids," *International Journal of Heat and Mass Transfer*, vol. 47, pp. 4795-4805, 2004.

Anurag Goyal is a senior undergraduate student, in the Department of Mechanical Engineering at IIT Delhi. His areas of interest are broadly in computer simulation of thermal and fluid systems and computational mechanics. Specifically, he is interested in combustion and flame characterization, heat transfer and microfluidics and, finite element methods (FEM).

Dr. Srinivasan Ramanathan is currently working as a Senior Research Engineer at Nalco Technology center, Pune, India for the Modeling Group. His areas of interest are Process Modeling and Design, Mathematical Modeling of Heat and Mass Transfer systems, CFD, FEA, Optimization and Numerical techniques. Srinivasan has a PhD in Chemical Engineering with 7 years of industrial experience in Oil & Gas, Chemical and Water treatment industries with extensive modeling and simulation experience.