2011 International Conference on Modeling, Simulation and Control
IPCSIT vol.10 (2011) © (2011) IACSIT Press, Singapore

Quantitative Analysis of an Analytical Model for Estimating the Model Accuracy

Anurag Goyal[1] and R.Srinivasan[2]

[1] Department of Mechanical Engineering, Indian Institute of Technology, Delhi, India

[2] Nalco Technology Center, Pune, India

Corresponding author: R.Srinivasan Tel.: + 91 20 39394089; Fax: +91 20 39394381

E-mail address: sramanathan@nalco.com

Abstract. Process modeling and design calculations require predicting the accuracy of analytical model output values. For any analytical model which requires many input parameters, only few parametric values are known and the rest, in most cases, are assumed as 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 in their range can be effectively judged.

Keywords: Model Accuracy, Weight Percentage, Parameter Importance, Sensitivity and Uncertainty Analysis

1. Introduction

Analytical models are used in numerous applications in science and engineering to predict and analyse the behaviour of a system or equipment. There are several input parameters to an analytical model and the accuracy of model output or value depends on the accuracy of the input parameters. The individual input parameters can be exact, measured, predicted or assumed. It is important to know which input parameters have a significant impact on the output value of a model and those individual parameters require a closer examination. Model accuracy is also a critical parameter for robust process modeling and design calculations. 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 [1].

Numerous methodologies are available for estimating the parametric importance for a model qualitatively. Saltelli and Tarantola [2] and Saltelli et al. [3] 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. Kuhnert et al. [4] used principal component analysis (PCA) to qualitatively identify the main parameters from a multivariate data set. Larsen [5] discusses some of the qualitative and quantitative parameters for data quality. 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 parametric importance and its quantitative contribution towards the model were estimated using the uncertainty and sensitivity model developed by Bushnell [6].

The model accuracy is highly correlated to the accuracy of measurement of different parameters involved in the formulation and their effect on the output of the model. To estimate the effect of input parameters, uncertainty and sensitivity analysis provides a good platform.

Bushnell [6] 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. The parameters with the highest importance or the ones that affect the output 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 methodology for developing the model accuracy and its application to two test cases, one in fluid flow equation and the other in the heat transfer model, are discussed in the following sections.

2. Methodology

2.1. 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,

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

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.

2.2. 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. The uncertainty and sensitivity analysis developed by Bushnell [6] was used to estimate the weight percentage with respect to each parameter. The parameter uncertainty and sensitivity are discussed in the following sections.

2.2.1. 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:

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

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 $\xi_{b,i}$ 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.

2.2.2. 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:

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

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.

2.2.3. 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 (I_i), the absolute value of the product of the relative range and normalized sensitivity is taken:

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

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.

3. Results

To analyse the developed model for estimating the accuracy, two case studies were performed. One on the simple, classic equation of predicting the flow rate using a venturi flow meter and the other 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 study and the results are desired in the following sections.

3.1. Case Study A: Fluid flow

The classic equation used for estimating the flow rate in a venturi flow meter is analysed using the developed model. The volumetric flow rate [7] for a venturi meter can be represented as

$$Q = \frac{\pi}{4} * C_d * D_t^2 \sqrt{2 * \rho_b * g * \Delta h/\rho_f} \quad (6)$$

where C_d is the coefficient of discharge, D_t is the throat diameter of the venturi in metres, ρ_b is the barometric fluid density (mercury) in kg/m^3, ρ_f is the fluid density in kg/m^3, g is the acceleration due to gravity, 9.8m/s^2 and Δh is the height difference in columns of barometer in m. Water is assumed to be fluid of interest for this case study and the parametric range are summarized in Table 1. The parametric range (min and max) values are chosen with consideration of standard design conditions and properties.

Table 1: Parametric ranges, uncertainty and sensitivity values for Case study A

Parameter	Minimum Value	Maximum Value	Uncertainty	Sensitivity
Height difference in barometer, Δh (cm)	4	12	1.00	0.53
Barometric fluid density, ρ_b) (kg/m^3)	13000	14000	0.07	0.53
Throat diameter, D_t (cm)	2	4	0.67	2.12
Fluid density, ρ_f (kg/m^3)	950	1050	0.10	-0.53
Discharge coefficient, C_d	0.6	0.9	0.33	1.06

Based on the parametric range tabulated, the uncertainty and sensitivity was calculated using Equation 2 and 3 respectively. Table 1 shows the uncertainty and sensitivity values derived for each parameter using the minimum and maximum range values. As discussed in the section 2.2.2, the fluid density has a negative magnitude for sensitivity and is due to the effect of the differential analysis techniques. Also the throat diameter shows a high positive magnitude compared to other parameters. It is obvious from the model equation; the parameters fluid density and throat diameter have a variance in their powers which directly attributes for the negative or positive magnitude. 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 generalise 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 throat diameter has more 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 for that analytical model. For estimating the accuracy, it is required to know whether the value is known (measured or calculated) or assumed and the assumption for estimating the accuracy for any parametric value would be the base-case value when the value is not known. Figure 2 shows the estimation of accuracy for each parameter assuming some base-case value for that parameter and estimating the model accuracy, which is the flow rate.

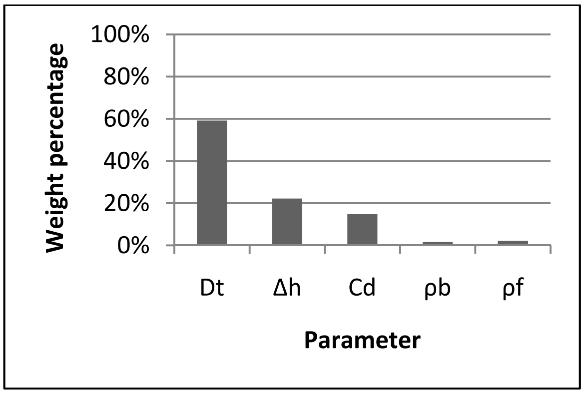

Fig. 1: Weight percentage values for different parameters

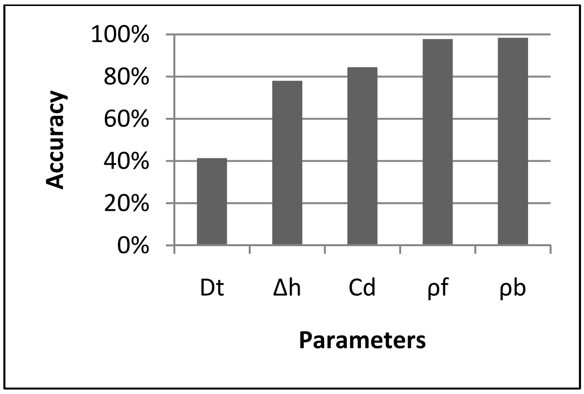

Fig. 2: Model accuracy for assumed base-case values

3.2. Case Study B: Heat transfer model

To extend the applicability of the developed model to a more rigorous analytical model, heat transfer equation for predicting the convective heat transfer coefficient for cross flow over a cylinder for liquid [8] 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/m^3, 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^{2}K and μ is the coefficient of viscosity in N-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^{o}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 2: Parametric ranges, uncertainty and sensitivity values for Case study B

Parameter	Minimum	Maximum	Uncertainty	Sensitivity
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 was calculated using Equation 2 and 3 respectively, and tabulated in Table 2. It shows the uncertainty and sensitivity values derived for each parameter using the selected range of values. As discussed in the section 2.2.2, the cylinder diameter and coefficient of viscosity shows a negative magnitude and 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 generalise the uncertainty for a particular model without the known range of values.

The estimation of weight percentage and accuracy is shown in Figure 3 and 4 respectively. 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. As discussed in Section 3.1, the dependence will change for different set of range of values for that specific analytical model.

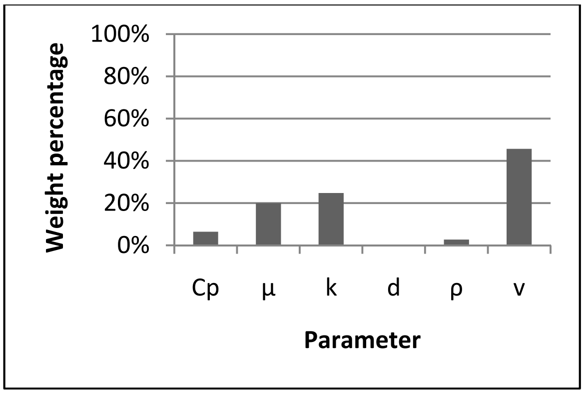

Fig. 3: Weight percentage values for different parameters

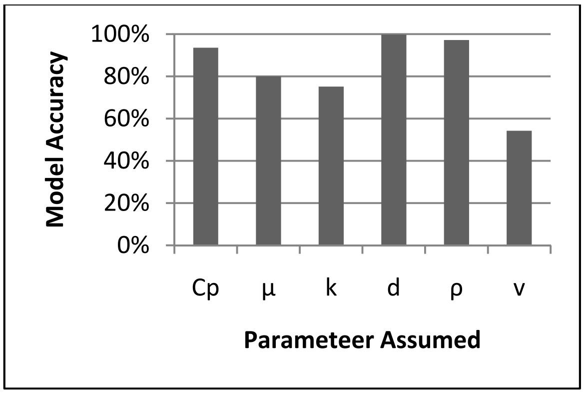

Fig. 4: Model accuracy for assumed base case values

4. Conclusion

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 simple flow equation and the other, 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.

5. Acknowledgements

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 and Dr. J.P. Subrahmanyam, Head, Department of Mechanical Engineering, Indian Institute of Technology, Delhi, India for providing their support to carry out this project.

6. References

[1] A. Saltelli. Sensitivity Analysis for Importance Assessment. *Risk Analysis, Joint Research Centre of the European* Communities *in* ISPRA*,* 2002, 22(3): 579-590.

[2] A.Saltelli and Stefano 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.

[3] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana and Stefano Tarantola. *Global Sensitivity Analysis. The Primer.* John Wiley and Sons, Ltd., 2008.

[4] 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.2011, 3:144-155.

[5] Alvin H. Larsen. Data Quality for Process Design. *Fluid Phase Equilibria.* 1986, 29: 47-58.

[6] Tanner H Bushnell. Parameter Importance of an analytical model for transport in the vadose zone. *Thesis for fulfilment of Master of Science.* Brigham Young University 2007

[7] Robert W. Fox, Alan T. McDonald and Philip J. Pritchard. *Introduction to Fluid Mechanics, Sixth Edition.* John Wiley and Sons, Ltd., 2004.

[8] S. Sanitjai, 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. 2004*, 47: 4795-4805.