ELSEVIER

Computers and Chemical Engineering 23 (2000) 1825–1838

Computers & Chemical Engineering

www.elsevier.com/locate/compchemeng

Uncertainty and sensitivity analysis of thermodynamic models using equal probability sampling (EPS)

Victor R. Vasquez, Wallace B. Whiting *

Chemical Engineering Division, University of Nevada-Reno, Reno, NV 89557-0136, USA

Received 22 May 1998; received in revised form 7 January 2000; accepted 7 January 2000

Abstract

A novel approach called equal probability sampling (EPS) is used for analyzing uncertainty and sensitivity in thermodynamic models. Uncertainty and sensitivity analysis for simulation and design of industrial processes are becoming increasingly important. The (EPS) method produces more realistic results in uncertainty analysis than methods based on other sampling techniques such as Latin hypercube sampling (LHS) or shifted Hammersley sampling (SHS), when parameters are highly correlated. When parameters are not correlated, EPS reduces to the LHS method. The EPS method is based on resampling to obtain uniform coverage over level sets of the objective function used to obtain the parameters of the model. The existence of unfeasible situations is substantially reduced with EPS. It can be extended to any regression model describing other kinds of physical applications and can be used as a better tool to estimate more reliable safety factors in the design and simulation of industrial chemical processes.

Keywords: Thermodynamics models; Non-linear regression; Sampling technique; Monte Carlo; Sensitivity; Uncertainty

1. Introduction

Process design and the simulation of chemical processes are strongly dependent on thermodynamic models. The uncertainty associated with thermodynamic models is an important factor for risk analysis and performance studies, and consequently this topic is very important for decision making in process safety and economic profitability analysis.

Normally, in actual process design and simulation operations, the associated uncertainty is covered using safety factors, which can increase costs and investment without a quantitative measure of the avoided risk. Uncertainty analysis can be used for studying the safety factors involved in the design. Decreasing their magnitude increases the efficiency and financial attractiveness of the global process. Alternately, the safety factor may be increased to attain a specific quantitative level of safety.

Specifically, the uncertainty associated with thermodynamic models derives from their parameters and modeling errors. The parameter uncertainty depends on the parameterization chosen and the random and systematic errors present in the experimental data (Tarantola, 1987).

Previous work has shown that the uncertainty associated with thermodynamic models can be very significant. Whiting, Tong and Reed (1993) showed how the uncertainty in the Soave–Redlich–Kwong equation of state affects the reflux ratio and reboiler heat duty in a superfractionator. A similar analysis for the number of stages in a distillation column is presented by Reed and Whiting (1993).

Additionally, surprising results have been obtained studying the effect of regressing parameters for thermodynamic models using different sources of experimental data (Vasquez & Whiting, 1998b) and regressing them from vapor-liquid and liquid–liquid data for liquid–liquid extraction operations (Vasquez & Whiting, 1998a), showing that the experimental data source and the approach followed for the parameter regression are very significant for uncertainty analysis.

---

* Corresponding author. Tel.: + 1-775-7844307; fax: + 1-775-7844764.

E-mail address: wwhiting@unr.edu (W.B. Whiting)

PII: S0098-1354(00)00297-0

A special concern when performing uncertainty analyses for regression models is the sampling approach used to obtain representative values from the parameter space. Traditional methods are based on stratified sampling over individual parameter distributions, and their correlation structures are approximated through the use of pairing procedures. Usually, the parameter distributions are estimated as normal with mean μ and standard deviation σ. An appropriate estimation of or can be difficult to calculate when dealing with highly nonlinear models. Examples of these methods are Latin hypercube sampling (LHS) (Iman & Conover, 1982) and Shifted Hammersley sampling (Kalagnanam & Diwekar, 1997). In terms of the objective function used to regress the parameters, these techniques have been shown (Vasquez, Whiting & Meerschaert, 1999) to be functionally equivalent to sampling from an approximation of a first-order in a Taylor's series expansion for the level sets. For nonlinear models (thermodynamic models belong to this category), the level sets are poorly estimated using first-order approximations when the parameter effects and intrinsic non-linearities as described by Bates and Watts (1980, 1981) are significant. Cook and Witmer (1985), and Seber and Wild (1989) present several examples, for simple nonlinear regression models, where the first-order approximation methods are inadequate. Mathematically speaking, these techniques are an approximation of the level sets defined by the right-hand-side of Eq. (1), which is the objective function approximated by an ellipsoid around the optimum using a Taylor development. The vector of unknown parameters is denoted by $\boldsymbol{\theta}$; $\boldsymbol{\theta}^*$ are the optimum parameters; $\mathbf{H}^*$ is the Hessian of $S(\boldsymbol{\theta})$ at $\boldsymbol{\theta}$; and $\mathbf{V}_{\boldsymbol{\theta}}$ is the covariance matrix of $\boldsymbol{\theta}$

$$S(\boldsymbol{\theta}) = S^* + (\boldsymbol{\theta} - \boldsymbol{\theta}^*)^T \frac{1}{2} \mathbf{H}^* (\boldsymbol{\theta} - \boldsymbol{\theta}^*)^T \frac{1}{2} \mathbf{V}_{\boldsymbol{\theta}}^{-1} (\boldsymbol{\theta} - \boldsymbol{\theta}^*) \tag{1}$$

If the parameter space is not sampled properly, the values for the stochastic variables generated may not correctly represent the physical problem, producing either unlikely results or infeasible situations.

The main goal of this work is to apply the sampling approach proposed by Vasquez et al. (1999), which combines an improved level set estimation method developed by Potocký and Ban (1992) with a new resampling method along the estimated level set, to obtain more realistic output distributions of the performance or design of chemical industrial processes. The results show narrower cumulative frequency distributions than the ones obtained using traditional sampling methods, a consequence of sampling over the improved level sets obtained from the EPS technique. The NRTL and UNIQUAC models are used for the uncertainty and sensitivity analysis of several practical illustrative cases.

2. Equal probability sampling

The major concern in uncertainty analysis is to obtain reliable results for the output distributions of the variable being studied or analyzed. This problem mainly relies on the sampling technique used to get the samples and the criteria employed to assign the probability distributions of the input parameters in a given stochastic model. For general purposes, the latter is a topic of wide discussion because the complexity involved is specific for each particular case. A more detailed discussion about this topic can be found in Refs. Haimes, Barry and Lambert (1994), Lambert Matalas, Ling and Haines (1994) and Seiler and Alvarez (1996).

From the engineering point of view, the sampling method used for uncertainty analysis should take into account only parameter sets with physical or 'practical' meaning. For example, parameter sets that provide an extremely poor fit of the equilibrium prediction to experimental data do not have to be considered for design and simulation purposes for obvious reasons. Also, the strong dependence among the parameters has to be maintained in cases like the activity coefficient models in order to keep the physical meaning of the regression. Thus, the sampling technique must provide an unbiased set of parameter samples with high fidelity of the nonlinear, highly correlated parameter space. The equal probability sampling (EPS) approach stratifies the parameter space into equal probability shells using the level sets of the objective function, automatically taking into account the geometry of the parameter near the optimum. Therefore, the EPS approach yields more realistic uncertainty analyses when compared with the results of methods based on traditional sampling techniques of nonlinear parameter spaces.

2.1. Method description

The equal probability sampling (EPS) method stratifies the parameter space according to the level sets of the objective function $S(\boldsymbol{\theta})$ used to regress the parameters. The probability distribution of $S(\boldsymbol{\theta})$ is stratified into N intervals of equal probability, and then the inverse image of these intervals form N shells of equal probability in the parameter space. From each of the N shells, a resampling scheme is used to obtain uniform coverage.

Fig. 1 shows a hypothetical level set of a simple arbitrary quadratic objective function $S(\alpha, \beta)$ defining an ellipsis of equal probability for the parameters α and