Computers and Chemical Engineering 34 (2010) 298–305

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Confidence interval estimation under the presence of non-Gaussian random errors: Applications to uncertainty analysis of chemical processes and simulation

V.R. Vásquez[a,*], W.B. Whiting[a], M.M. Meerschaert[b]

[a] Chemical and Metallurgical Engineering Department, University of Nevada Reno, Reno, NV 89557, USA
[b] Department of Statistics and Probability, Michigan State University, E. Lansing, MI 48823, USA

ARTICLE INFO

Article history:
Received 26 January 2007
Received in revised form 24 October 2009
Accepted 5 November 2009
Available online 22 November 2009

Keywords:
Confidence intervals
Uncertainty analysis
Heavy tails
Error propagation
Monte Carlo

ABSTRACT

Confidence intervals (CIs) are common methods to characterize the uncertain output of experimental measurements, process design calculations and simulations. Usually, probability distributions (pdfs) such as Gaussian and *t*-Student are used to quantify them. There are situations where the pdfs have anomalous behavior such as heavy tails, which can arise in uncertainty analysis of nonlinear computer models with input parameters subject to different sources of errors. We present a method for the estimation of CIs by analyzing the tails of the pdfs regardless of their nature. We present case studies in which heavy tail behavior appears due to the systematic errors in the input variables of the model. Taking into account the probability distributions behavior to estimate appropriate CIs is a more realistic approach to characterize and analyze the effect of random and systematic errors for uncertainty analysis of computer models.

1. Introduction

Confidence intervals are widely used to characterize random variables with a given probabilistic confidence. Typical methods of estimating confidence intervals are using standard normal or Student-*t* quantiles. They are very common in many industrial applications, experimentation, and product quality control. In uncertainty and robustness analysis of computer models, confidence intervals are also important, but traditional methods may not be practical or applicable due to the probabilistic *nature* of the output random variables. Sometimes these cannot be characterized using standard normal or *t*-Student distributions. We find that in uncertainty analysis of process design and simulation, the presence of uncertainty sources of different natures such as random and systematic errors can cause heavy tail behavior on the probability distributions of output random variables.

In this work, we present a simple method to estimate confidence intervals when the probability distribution of output random variables tend to have heavy tail behavior caused by uncertainty in the input parameters of the model. We explore the effect that systematic and random errors have on the inter-quantile range estimation using Monte Carlo simulation with simple case studies. We compare the confidence interval estimation using traditional techniques and the proposed methodology. The results show that the differences can be significant when the nature of the probability distributions deviates in the tails from standard normal or Student-*t*.

In general, appropriate estimation of confidence intervals is very important to evaluate the robustness of mathematical models and experimental data. It is common to assume that errors are Gaussian, particularly for experimental measurements. The presence of systematic errors can have significant effects, but it is more difficult to evaluate and it is usually not reported even for highly accurate measurements. One can find evidence of systematic errors in measurements by examining experimental data for the same system from different literature sources or by analysis of historical measurements (Henrion & Fischoff, 1986; Vasquez & Whiting, 1999).

Complex nonlinear computer models can induce heavy tail behavior in the output variables. In addition to computer models, there are many other situations such as anomalous diffusion (Sokolov & Klafter, 2005) where the presence of heavy tails are common. Appropriate methods to analyze the probability distributions are necessary in order to capture the significance of unlikely events. For example, in the analysis of process conditions and design calculations where catastrophic events are likely, detecting

* Corresponding author at: Chemical and Metallurgical Engineering Department, University of Nevada, Reno, Mail Stop 170, Reno, NV 89557, USA.
Tel.: +1 775 784 6060; fax: +1 775 784 4764
E-mail addresses: victor.vasquez@unr.edu (V.R. Vásquez), wwhiting@unr.edu (W.B. Whiting), mcubed@stt.msu.edu (M.M. Meerschaert).

0098-1354/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2009.11.004

these become important for design and selection of safety factors in equipment and operation conditions.

2. Random and systematic errors

Random errors can be defined as the uncertainty detected by repeating measuring procedures under the same conditions. On the other hand, *systematic errors* are those that cannot be detected this way and are usually associated with bias in experimental data. Experimental random errors can be caused by any random event occurring during the measuring procedure. These include any random variation in the environmental conditions where the measurements are taken and inherent random variations of the measuring instruments. Random errors are usually characterized using the statistics of the instrument or by replication. Under some circumstances obtaining a statistical number of replicates can be costly and often only conservative error quantification estimates are reported using Gaussian or uniform probability distributions.

A classical example of systematic error is the effect that a bias in a calibration procedure would have on experimental measurements. For example, if during the calibration of a thermometer there is a calibration error of half of a degree, all the measurements performed with that instrument will be off by that amount. This is normally called a constant bias or calibration error, and it can be positive or negative. There are situations where the magnitude of the systematic and random errors can be a function of one or more of the measured variables involved (Vasquez & Whiting, 1999). In general, systematic errors can be defined as those that cannot be eliminated through repeated experiments. The main problem when dealing with systematic errors is their proper identification. There is no standard procedure to identify and report them and subjective judgment in their identification and quantification is common.

As a consequence, procedures to quantify this type of uncertainty and its effects are difficult to find and are restricted mostly to the analysis of *constant bias*. The standard ANSI/ASME PTC 19.1 (ANSI, 1998) and the *Guide to the Expression of Uncertainty in Measurement* (ISO, 1995) of the International Organization of Standardization (ISO) point out that uncertainty analyses should be able to identify several sources of systematic or bias errors. Modeling approaches that take into account the effect of random and systematic errors on process design and simulation are important to study the reliability of models and to the rational design of safety factors.

According to ANSI/ASME PTC 19.1 (ANSI, 1998) and ISO (ISO, 1995), the total uncertainty for an observable or measurable event μ is expressed as

$$u_\mu = \beta + \varepsilon_\mu \tag{1}$$

where β is a fixed bias error and ε_μ is the random error. When several sources of systematic errors are identified, a representative value of β is estimated as

$$\tilde{\beta} \approx \left[\sum_{i=1}^{m} \beta_i^2\right]^{1/2} \tag{2}$$

where i defines the sources of bias errors, β_i is the bias limit within the error source i. Similarly, the same approach is used to define the total random error based on individual standard deviation estimates (ANSI, 1998; ISO, 1995),

$$\tilde{\varepsilon}_\mu \approx \left[\sum_{i=1}^{n} s_i^2\right]^{1/2} \tag{3}$$

As mentioned earlier, Vasquez and Whiting (1998) presented evidence of systematic errors in experimental data that vary with one or more of the measured quantities. Figs. 4 and 6 show examples of this behavior for ternary liquid–liquid equilibria (LLE) systems. We can see that there are systematic errors on the data (see the systematic differences among the different laboratories) that are function of the composition of one of the components. Similar evidence was reported by Henrion and Fischoff (1986) for fundamental physical constants. This suggests that methods for the analysis of systematic uncertainty based upon the use of Eq. (1) may not be appropriate under this type of situation.

Understanding how these errors propagate through computer models is important for uncertainty analysis. Analytical approaches are very limited because of the usual nonlinear complexity of engineering models. Current analytical methods are based on error propagation analysis using the Taylor's series expansion (ANSI, 1998; ISO, 1995; Taylor, 1982). They consist of estimating the combined uncertainty $u_c(y)$ on the estimate of y, where y is given by the model $y = f(x_1, x_2, \ldots, x_N)$ with uncertain input variables $x_1, x_2, \ldots, x_N$. Then, the combined uncertainty $u_c(y)$ is estimated as

$$u_c^2(y) = \sum_{i=1}^{N} \left[\frac{\partial f}{\partial x_i}\right]^2 u^2(x_i) \tag{4}$$

A more practical approach is to use Monte Carlo methods (Helton, 1993). These can be used in highly complex nonlinear models with many input and output variables. The main limitation is the computing cost associated with the simulations and post-processing the results.

3. Confidence intervals estimation

To estimate confidence intervals of random variables, we need to take into account the probability distribution characteristics. Once the probability density is known, upper confidence bounds for the inter-quantile ranges are easily estimated by applying:

$$\int_{x_o}^{\infty} f(x)dx = q \tag{5}$$

where $f(x)$ is the probability density, q the desired significance level, and x_o corresponds to the upper bound of the confidence interval (CI). Estimation of lower confidence bounds can be obtained in a similar manner. By looking at Eq. (5), we can see that information about the probability tails is required in order to have good estimates of the confidence intervals. This information is contained in the extreme order statistics. Analysis of just the tails provides a very robust model which can be applied without knowledge of the underlying distribution (Leadbetter, Lindgren, & Rootz, 1980). In fact, extreme value theory predicts that for an unbounded model, the extreme order statistics will (at least approximately) follow one of two general distributional forms (Leadbetter et al., 1980). These are exponential $f(x) = \lambda \exp(-\lambda(x-d))$ or Pareto $f(x) = C\alpha x^{-\alpha-1}$, where $f(x)$ is the probability density (Leadbetter, Lingreen, & Rootzén, 1983) (Fig. 1). This fact will affect clearly the estimation of confidence intervals for probability distributions presenting this type of behavior. As shown later, random and systematic uncertainty can cause *heavy* character on the tails of probability distributions for output variables of computer models. Thus, appropriate methods to compute confidence intervals involving the tails are required under these circumstances. The *nature* of the model can also be the cause of heavy or exponential tails.

The main idea is to characterize the probability distribution tails by using either an exponential or Pareto model. To find out which model is more appropriate one can plot the information on the tails using log–log and semi-log plots (see the densities $f(x)$ for the exponential and Pareto models defined above). The empirical probability distribution is obtained from an independent random sample of

size n using the order statistics $X_{(1)} \geq X_{(2)} \geq \cdots \geq X_{(n)}$, where the random variable X follows an exponential probability distribution. Since the fraction of the data at or above $X_{(i)}$ is (i/n) we estimate $P[X \geq X_{(i)}] = i/n$. This simple idea can be used to investigate the nature of the extreme order statistics or tail, to determine which probability model is more appropriate for the tail. For exponential tails $q = P[X > x] = \exp(-\lambda(x-d))$ so $\ln q = -\lambda x + \lambda d$. Substituting the empirical distribution function values, we see that a plot of $\log(i/n)$ versus $X_{(i)}$ should appear linear with slope $-\lambda/\ln(10)$. For a random variable Y following a Pareto distribution $q = P[Y > y] = Cy^{-\alpha}$ and $\log q = \log C - \alpha \log y$ so that a plot of $\log(i/n)$ versus $\log Y_{(i)}$ would appear linear with slope $-\alpha$. In other words, a semi-log plot of the empirical tail distribution function will appear linear for exponential tails, and a log–log plot will appear linear for heavy tails. This reflects the fact that if Y is Pareto then $\ln Y$ is shifted exponential. For many practical problems there is no prior indication to guide our choice of probability tail model, and in this case we recommend examining both plots to determine the best fit.

Once a tail probability model has been chosen, simple linear regression can be used to obtain approximate parameter estimates. A more accurate method is to employ the conditional maximum likelihood estimates developed by Hill (1975). For a heavy tail Pareto model

$$\hat{\alpha}_r = \left[-\ln Y_{(r+1)} + r^{-1} \sum_{i=1}^{r} \ln Y_{(i)} \right]^{-1} \tag{6}$$

$$\hat{C}_r = \left(\frac{r+1}{n} \right) (Y_{(r+1)})^{\hat{\alpha}_r} \tag{7}$$

are the conditional maximum likelihood estimators for $\alpha > 0$ and $C > 0$ based on the $r+1$ largest order statistics. For data which are approximately Pareto in the tail, one should choose r small enough so that only the Pareto-like tail is represented. Inspection of the tail empirical distribution function will give some idea about an appropriate value of r. Choose r as large as possible but small enough so that the log–log plot remains linear. Another method is to plot the parameter estimates versus r (some authors call this a Hill plot) and make r large enough that the plot stabilizes. A typical application of Hill's estimator in the literature involves 1000–3000 observations and estimates α using the upper 10 percent of the data or less (Aban & Meerschaert, 2004; Anderson & Meerschaert, 1998; Hill, 1975; Jansen & de Vries, 1991; Loretan & Phillips, 1994; Meerschaert & Scheffler, 1998; Resnick & Stărică, 1995).

For an exponential tail model, it is permissible to apply the above methods along with the relations $X = \ln Y$, $\lambda = \alpha$, and $d = \ln C/\lambda$. Equivalently, it is possible to show directly by an extension of the arguments used by Hill (1975) that

$$\hat{\lambda}_r = \left[-X_{(r+1)} + r^{-1} \sum_{i=1}^{r} X_{(i)} \right]^{-1} \tag{8}$$

$$\hat{d}_r = X_{(r+1)} + \hat{\lambda}_r^{-1} \ln \left(\frac{r+1}{n} \right) \tag{9}$$

are the conditional maximum likelihood estimators for $\lambda > 0$ and d based on the $r+1$ largest order statistics. As with the heavy tail model, r is typically chosen by graphical methods.

Upper confidence bounds are easily obtained from the tail probability model once the parameters have been estimated. For Pareto tails the $100(1-q)\%$ upper confidence bound is just the $1-q$ percentile y_q defined by $q = P[Y > y_q] = C(y_q)^{-\alpha}$. Solving this equation yields

$$y_q = (C/q)^{1/\alpha} \tag{10}$$

For the exponential model the $100(1-q)\%$ upper confidence bound is

$$x_q = d + \ln q/\lambda \tag{11}$$

which is the solution of the equation $q = P[X > x_q] = \exp(-\lambda(x_q - d))$.

It is important to point out that the estimators for α and λ defined by Eqs. (6) and (8) can be used for upper tails or largest values (also called maxima). The application of these estimators to smallest values or lower tails (also called minima) can be achieved similarly by taking the inverse of the data or by multiplying by -1 for cases where the lower tail has negative values.

4. Methodology

The methodology used in this work consists of three general steps:

1. Monte Carlo simulation of random and systematic error effects.
2. Characterization of tails for the cumulative probability distribution of output variables.
3. Estimation of confidence intervals taking into account probability distribution tail characteristics.

To analyze random and systematic error effects using Monte Carlo simulation, the approach proposed by Vasquez and Whiting (1999) is used, which consists of, first, defining appropriate prob-

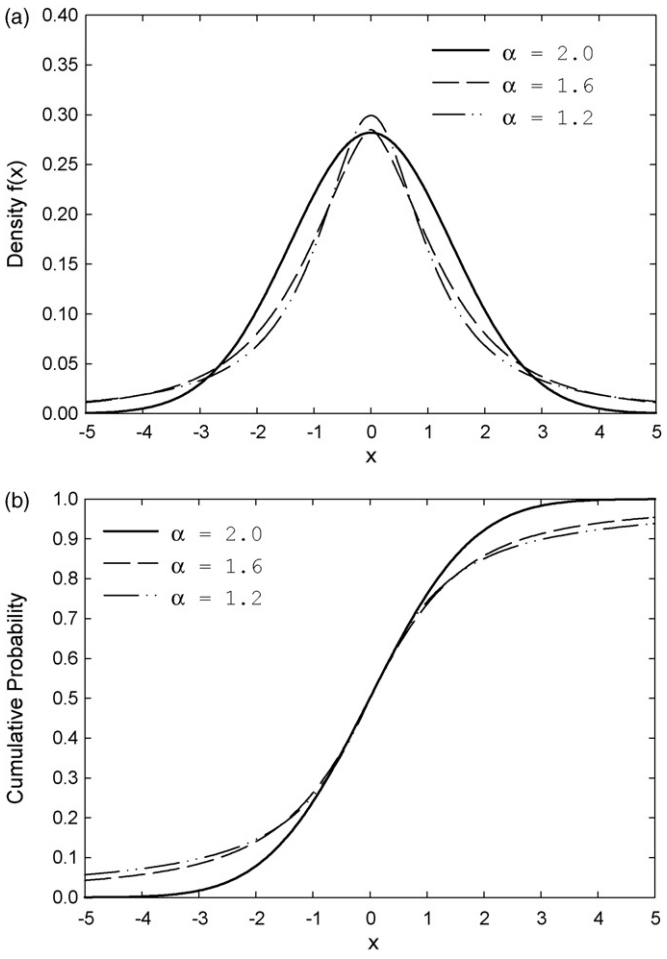

Fig. 1. Symmetric α-stable densities showing *heavy* tailed character. (a) Probability density and (b) cumulative probability.

Table 1
Parameter definitions for the ILCR model.

Variable	Definition	Type	Mean μ	Standard deviation σ
C_s	Benzene concentration (mg/kg)	Log-normal	0.84	0.77
S_r	Soil ingestion rate (mg/day)	Log-normal	3.44	0.80
RBA	Relative bio-availability	Parameter	1.00	–
E_w	Exposure days per week	Parameter	1.00	–
E_y	Exposure weeks per year	Parameter	20.0	–
E_{Lf}	Exposure years per life	Parameter	10.0	–
cf	Conversion factor (kg/μg)	Parameter	10^{-6}	–
bw	Body weight (kg)	Normal	47.0	8.3
dy	Days per year	parameter	364	–
ylf	Years per lifetime	Parameter	70	–
CPF	Cancer potency factor (kg day/μg)	Log-normal	−4.33	0.67

ability distributions for the random and systematic errors based on evidence found from different data sources. Then bias limits are defined for the systematic errors of the input variables of the model. Samples are drawn using an appropriate probability distribution for the systematic errors if *a priori* information is available. Otherwise, a uniform distribution is used. For the random errors, samples are taken from each of the probability distributions characterizing the random error component of the input variables and then the samples are passed through the computer model. Gaussian probability distributions are usually used to characterize variables with random errors. The characterization of the output probability distribution tails is accomplished using the Hill's estimator for the thickness of heavy and exponential tails (*i.e.* determination of either α or λ using Eq. (6) or Eq. (8)). The last step is the estimation of the confidence intervals. We use Eqs. (10) and (11) to compute the $100(1-q)\%$ upper confidence bound for heavy and exponential tails, respectively.

4.1. Case study: effect of random and systematic errors on Incremental Lifetime Cancer Risk (ILCR) calculations

An example adapted from Shlyakhter (1994) is used to illustrate the proposed approach for confidence intervals estimation under the presence of random and systematic errors. The model is simple and easy to understand and it consists of Incremental Lifetime Cancer Risk (ILCR) calculations for children, produced by benzene ingestion with soil. This problem was first addressed by Thompson, Burmaster, and Crouch (1992) in public health risk assessments. The model for the ILCR is of the form:

$$\text{ILCR} = \frac{C_s \times S_r \times RBA \times E_w \times E_y \times E_{Lf} \times cf}{bw \times dy \times ylf} \times CPF \tag{12}$$

The parameters definitions and their values for the ILCR model are given in Table 1. Random error effects were studied by performing 1000 Monte Carlo simulations using the provided probability distributions for the input variables. The effect of systematic errors was studied by adding systematic changes to the statistical means of the stochastic input variables involved. The systematic errors are assigned in every trial by drawing pseudo-mean values for the input variables from a uniform distribution defined as $\boldsymbol{U}[\mu(1-p), \mu(1+p)]$, where p is the maximum fractional change allowed in μ due to systematic errors. Once the new pseudo-mean (μ^{new}) is defined then a normal random sample is chosen from $\mathbf{N}[\mu^{new}, \sigma]$. Six levels of systematic error were analyzed: 10%, 30%, 50%, 70%, 85%, and 90%. The higher levels of systematic error may not be realistic for this situation, but are included to demonstrate the possible range of behavior for simple nonlinear models.

Fig. 2 shows the cumulative probability curves of the Monte Carlo simulations for the ILCR model outputs involving both random and systematic error effects for different levels of systematic error. We can see that the random error is mainly controlling the error propagation. Only for significant levels of systematic error the cumulative probability distributions tend to be different from the one for the random error only. Notice that all the cumulative probability distributions seem Gaussian. Fig. 3 show the upper tails using a log–log plot. We can see the heavy character of the tails (*i.e.* linear trends). The lower tails present a similar behavior. The characterization of the tails was performed using the Pareto model for the upper tail of the cumulative probability distributions of Fig. 2 and the results are presented in Table 2.

Table 3 summarizes the 99% confidence interval calculations using Gaussian, log-normal, and Pareto probability distributions. It is clearly observed that the use of Gaussian distributions is not appropriate in this case. On the other hand, the log-normal distribution is a good representation for the ILCR values, which is in very close agreement with the bounds estimated using the Pareto distribution. Also, it is observed that the higher the level of uncertainty the heavier the tails of the probability distributions become, so that the lower and upper bounds estimation using the normal distributions becomes too narrow. This observation is expected due to the fact that the Pareto model produces a better representation of the distribution tails than the normal model does.

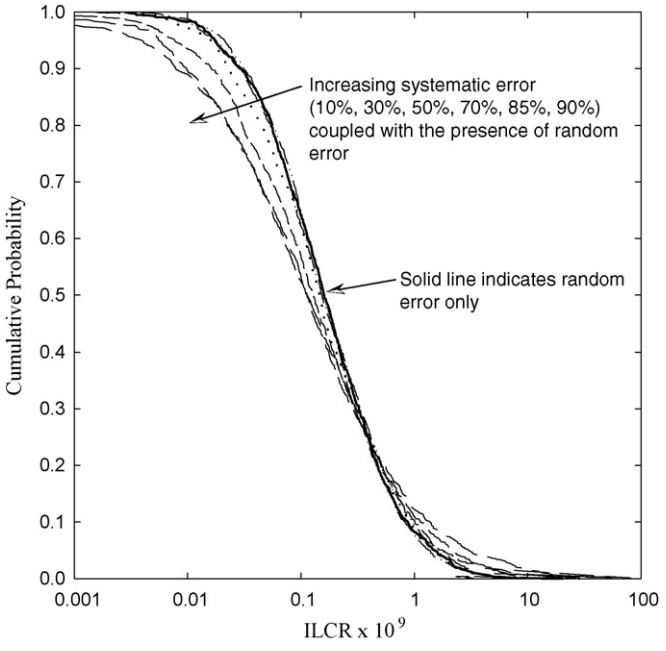

Fig. 2. Monte Carlo simulations of the Incremental Lifetime Cancer Risk (ILCR) for children caused by ingestion of benzene in soil (Thompson et al., 1992). Cumulative probability curves are presented for different combined error effects (random and systematic) of the random variables in Eq. (12). The percentage values indicated correspond to the systematic error used on the statistical mean of the input variables. For comparison purposes, the Monte Carlo simulation produced from having only random error effects (sampling performed using Gaussian distributions $\mathbf{N}[\mu, \sigma]$) is included ("Random error only").

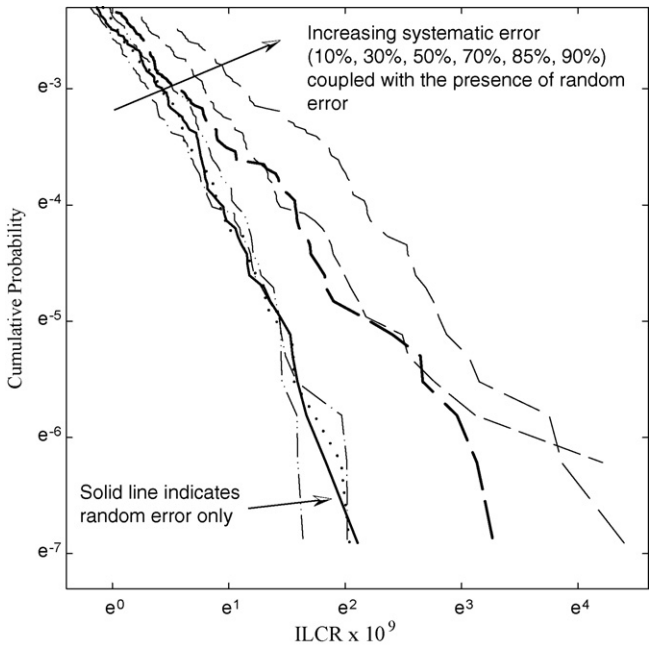

Fig. 3. Monte Carlo simulations of the Incremental Lifetime Cancer Risk (ILCR) for children caused by ingestion of benzene in soil (Thompson et al., 1992). Upper 10% tails of the cumulative probability curves are plotted (semi-log scales) for combined effects of systematic and random errors in the random variables of Eq. (12). The percentage values indicated correspond to the systematic error used in the statistical mean characterizing the input variables. For instance, for a given random variable with mean μ, its new mean after a systematic error of 10% is introduced is defined by a random value chosen from the distribution $\mathbf{U}[0.9\mu, 1.1\mu]$. Then a normal random sample is chosen from $\mathbf{N}[\mu^{new}, \sigma]$. For comparison purposes, the Monte Carlo simulation produced from having only random error effects is included ("Random error only"). See Fig. 2 for more details.

Table 2
Parameters α and C for different levels of systematic uncertainty in the estimation of ILCR. The parameters are for the lower and upper tails of the probability distributions presented in Fig. 2, respectively. The parameter r represents the order statistics at which α and C were determined for the lower and upper 10% of the distributions. For the lower tails, the data were transformed by taking the inverse and then the Hill estimator was used.

S&R	$\hat{\alpha}_r$				$\hat{C}_r$	
Level	r	Lower	r	Upper	Lower	Upper
10%	50	1.86	50	2.16	66.16	0.1540
20%	50	1.60	50	1.82	32.36	0.0857
50%	50	1.63	50	2.10	42.26	0.1209
70%	70	1.28	50	1.15	20.80	0.0916
85%	50	1.52	40	1.34	115.5	0.1308
90%	50	1.17	50	1.04	23.70	0.1585
Random	50	1.97	50	1.92	148.2	0.1054

Table 3
Lower and upper bounds for the 99% confidence intervals of the ILCR calculations under the influence of both random and systematic errors. Confidence intervals are compared with Gaussian normal and log-normal estimates.

S&R	Normal		Log-normal		Pareto	
	Lower	Upper	Lower	Upper	Lower	Upper
10%	−1.44	2.25	0.0057	5.07	0.0061	4.88
20%	−15.2	2.26	0.0051	5.64	0.0043	4.78
50%	−2.63	3.44	0.0040	5.61	0.0039	4.56
70%	−3.75	4.79	0.0019	9.14	0.0014	12.44
85%	−6.06	7.30	0.0011	12.88	0.0013	11.52
90%	−9.12	11.05	0.0008	22.08	0.0007	28.17
Random	−1.97	2.76	0.0053	4.94	0.0054	4.89

5. Engineering computational studies

To illustrate the application of the confidence interval estimation method proposed, two liquid–liquid extraction cases using the UNIQUAC activity coefficient model (Abrams & Prausnitz, 1975) are studied involving liquid–liquid equilibria predictions for the ternary systems diisopropyl ether + acetic acid + water and chloroform + acetone + water.

The binary interaction parameters of the UNIQUAC model were regressed using an objective function based on the minimization of the distances between experimental and estimated mole fractions, using an inside–variance estimation regression method proposed by Vasquez and Whiting (1998).

5.1. Liquid–liquid extraction of acetic acid

This is an example selected from Treybal (1981), and it consists of 8000 kg/h of an acetic acid–water solution, containing 30% (mass) acid, which is to be countercurrently extracted with diisopropyl ether to reduce the acid concentration in the solvent-free raffinate product. The solvent flowrate is 12,500 kg/h, and the column has 8 equilibrium stages. The column operates at 25 °C. The output variable is the percentage of acetic acid extracted at steady state conditions in the extractor. The experimental data for this system were taken from Treybal (1981), Othmer, White, and Trueger (1941), and Hlavaty and Linek (1973). The last two data sets are also reported by Sørensen and Arlt (1980). Fig. 5 presents the experimental tie lines for the different data sets mentioned. Clear systematic trends are observed in the left phase of the system, as well as systematic variations on the tie-line slopes.

To simulate both random and systematic errors, the first step consists of setting approximate bias limits based on experimental evidence. To do this, we plotted the different experimental data sets describing the equilibrium of the system in Fig. 4. If we look at the left side of the plot (the left phase), we note substantial variation between the experimental data. Thermodynamically, the data should all fall on one well defined curve, like the right phase on

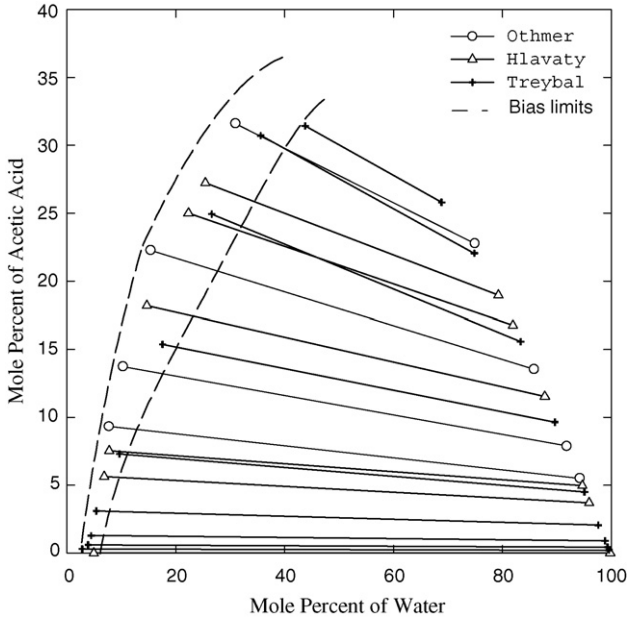

Fig. 4. Systematic error limits defined for the left phase in the liquid–liquid equilibria experimental of the diisopropyl ether(1) + acetic acid(2) + water(3) ternary system at 25 ° C based on evidence of systematic uncertainty found using experimental data from different sources. The experimental data are from Treybal (1981), Othmer et al. (1941), and Hlavaty and Linek (1973).

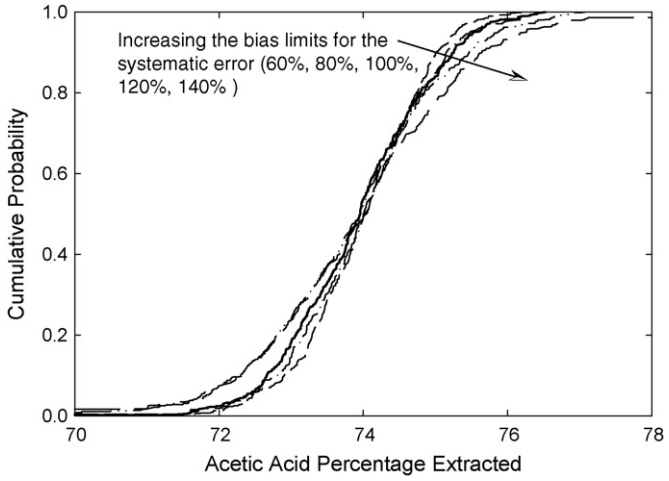

Fig. 5. Uncertainty of percentage of acetic acid extracted in the liquid–liquid extractor under the effect of systematic and random errors. The percent values indicate how much of the bias limits defined in this figure were used to perform the analysis.

the same plot. Next we choose boundaries (dashed lines) that we believe bracket the true values of the left boundary of the system. Since we are choosing these limits purely on a visual inspection of the experimental data, we will also perform a sensitivity analysis, varying the range of these bias limits. The dashed lines in Fig. 4 define what we call the 100% bias limits. For uncertainty analysis purposes, we proportionally expand or contract these limits. For instance, the horizontal distances within the 100% bias limits (x-axis) can be expanded away from center by a factor of 1.20. This procedure will produce a new set of dashed lines, which we call the 120% bias limits. Similarly, we can contract the 100% base case to produce other sets of bias limits that are smaller than the one presented in the figure.

The data set of Hlavaty and Linek (1973) was used as the basis for the simulations. For each tie line pair from that data set, for each set of bias limits, we simulate a systematic error in the left phase by choosing a pseudo-random point uniformly distributed between the bias limits along the corresponding tie line. Then we add a random experimental error following a normal distribution with mean zero and standard deviation equal to 0.3236%, which yields a maximum error of 1%, typical of measurement techniques to quantify concentration. Both the systematic and random error are drawn again for each replication of the simulation. The final step is to use these data sets to regress the binary interaction parameters and evaluate the performance of the liquid–liquid extraction unit operation.

All 1000 random pseudo-data sets were generated including both types of errors following the procedure described above, which are then used to regress all 1000 sets of binary interaction parameters, which are passed through the ASPEN Plus process simulator to calculate the performance of the extraction operation. All 1000 performance evaluations were carried out for each level of systematic error considered. In this case study, 5 levels of systematic uncertainty were analyzed. They are 60%, 80%, 100%, 120%, and 140% of the bias limits defined in Fig. 4. The results of the simulations are presented in Fig. 5 as cumulative probability distributions for the percentage of acetic acid extracted. We can see clearly that when the error level increases there is a tendency on the tails of the probability distributions to become heavier. The upper 5% tails of these distributions were used to characterize the tails using the parameter λ (*i.e.* exponential tails) and the results are presented in Table 4.

Then 99% confidence intervals were calculated for the percentage of acetic acid extracted using the Gaussian normal and exponential models. The results are presented in Table 5. By looking at Fig. 5, we can observe that the confidence intervals estimated using the exponential model for the tails is significantly more reasonable than the ones obtained from Gaussian normal estimates. This behavior is of particular importance when the level of uncertainty increases. This can be easily explained by the fact that Gaussian estimates are based on the assumption of perfect symmetry in the empirical probability distributions and the tail exponential model does not use such assumption (*i.e.* each tail is analyzed independently), but just looking at the shape of the distributions in Fig. 5 such conclusion is not obvious. One way of overcoming this problem is to find a suitable probability density to fit the empirical distribution, but it can be a cumbersome task and unnecessary in this case.

Table 4
Parameters λ and d for different levels of systematic uncertainty in the percentage of acetic acid extracted. The parameters are for the lower and upper tails of the probability distributions presented in Fig. 5, respectively. The parameter r represents the order statistics at which λ and d were determined for the lower and upper 5% of the distributions. For the lower tails, the data were transformed by taking the inverse and then the Hill-like estimator was used. Notice from Eq. (8) that λ and d are unit dependent. Therefore, the inverse of the percentage of acetic acid extracted was re-scaled by a factor of 1000.

Bias limits	$\hat{\lambda}_r$				$\hat{d}_r$	
Level	r	Lower	r	Upper	Lower	Upper
60%	70	2.53	50	3.43	12.71	74.41
80%	50	11.09	50	2.12	13.53	74.07
100%	50	2.56	60	2.50	12.66	74.32
120%	50	0.815	50	2.02	10.29	74.39
140%	50	0.865	50	1.33	10.49	73.82

5.2. Liquid–liquid extraction of acetone

For this system, an example of a liquid–liquid extraction operation reported by Smith (1963) is used. Water is used to separate a chloroform–acetone mixture in a simple countercurrent extraction column with two stages. The feed contains equal amounts of chloroform and acetone on a weight basis. The column operates at 25 ° C and 1 atm. A solvent/feed mass ratio of 1.565 is used. The output variable for this case is the percentage of acetone extracted. The experimental data used in this example are from Bancroft and Hubard (1942), Brancker, Hunter, and Nash (1940), Reinders and De Minjer (1947), and Ruiz and Prats (1983), all of them also reported by Sørensen and Arlt (1980). Fig. 6 presents experimental LLE data sets showing evidence of systematic errors on the left phase of the system. For this case study, 4 levels of systematic uncertainty were studied. The 100% bias limits corresponds to the systematic limits presented in Fig. 6, then levels at 60%, 80%, and 120% were also analyzed. All 1000 pseudo-data sets were generated randomly including the effects of random and systematic errors as explained in the previous case study. The reference experimental data set used to define the systematic errors is the one provided by Reinders and De Minjer (1947). Then, with the binary interaction parame-

Table 5
Lower and upper bounds for the 99% confidence intervals of the percentage of acetic acid extracted under the influence of both random and systematic errors. Confidence intervals are compared with Gaussian normal estimates.

Bias limits	Normal		Exponential	
Level	Lower	Upper	Lower	Upper
60%	63.33	84.48	67.56	75.95
80%	72.98	75.02	71.37	76.57
100%	64.45	83.33	67.94	76.43
120%	56.65	90.70	59.58	77.01
140%	56.41	91.87	60.26	77.81

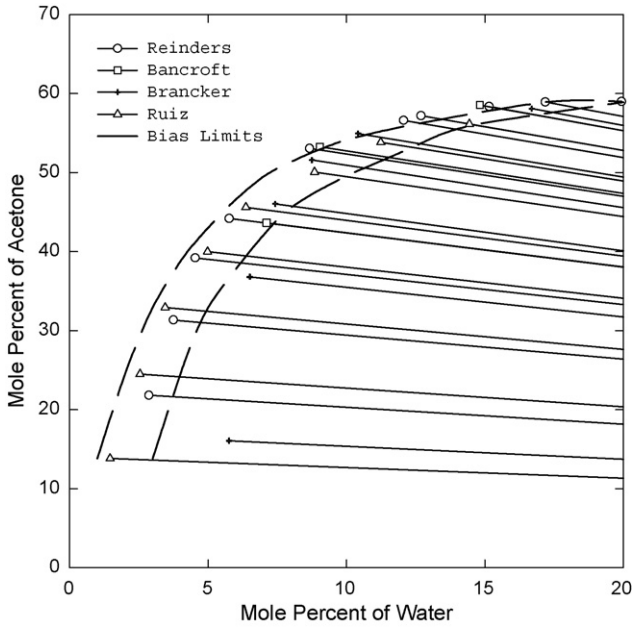

Fig. 6. Systematic error limits defined for the left phase in the liquid–liquid equilibria of the chloroform(1) + acetone(2) + water(3) ternary system at 25 ° C based on evidence of systematic uncertainty found using experimental data from different sources. The data are from Reinders and De Minjer (1947), Bancroft and Hubard (1942), Brancker et al. (1940), and Ruiz and Prats (1983).

ters regressed, all 1000 extraction performance evaluations were carried out using the ASPEN Plus process simulator.

The cumulative probability distributions obtained for the percentage of acetone extracted are presented in Fig. 7. We can see that in this case it is more difficult to distinguish among the distributions. The parameter λ (exponential tails represented a better option than Pareto) using the Hill-like (see Eq. (8)) estimator was computed for the different levels of systematic error for the lower and upper 10% of the tails. The results are presented in Table 6.

The 99% confidence intervals were calculated for the percentage of acetone extracted using the Gaussian normal and exponential models. The results are presented in Table 7. In this case, the Gaussian-based confidence intervals are more narrow than the ones calculated using the exponential model. Basically, this happens because the Gaussian probability density does not reproduce well the long flat tails present in the empirical distribution of Fig. 7.

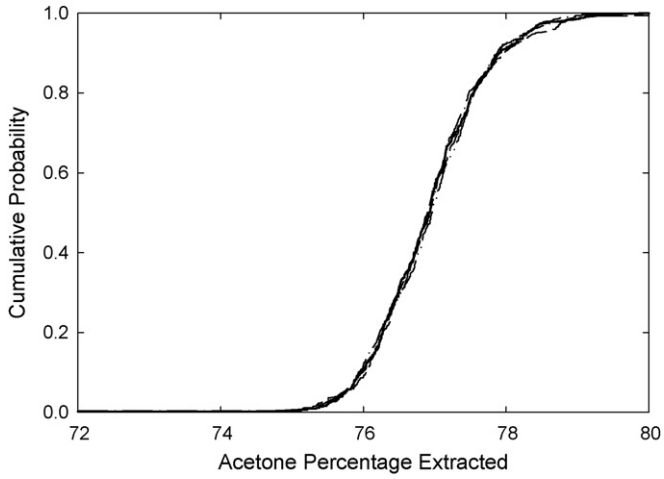

Fig. 7. Uncertainty of percentage of acetone extracted in the extractor under the effect of systematic and random errors.

Table 6
Parameters λ and *d* for different levels of systematic uncertainty in the percentage of acetone extracted. The parameters are for the lower and upper tails of the probability distributions presented in Fig. 7, respectively. The parameter *r* represents the order statistics at which λ and *d* were determined for the lower and upper 10% of the distributions. For the lower tails, the data were transformed by taking the inverse and then the Hill-like estimator was used. Notice from Eq. (8) that λ and *d* are unit dependent. Therefore, the inverse of the percentage acetone extracted was re-scaled by a factor of 10,000.

Bias limits	$\hat{\lambda}_r$				$\hat{d}_r$	
Level	*r*	Lower	*r*	Upper	Lower	Upper
60%	70	1.44	60	1.72	129.89	76.69
80%	60	1.83	60	2.25	130.25	76.95
100%	60	2.03	50	2.25	130.50	76.97
120%	50	2.21	50	2.56	130.65	77.10

Table 7
Lower and upper bounds for the 99% confidence intervals of the percentage of acetone extracted under the influence of both random and systematic errors. Confidence intervals are compared with Gaussian normal estimates.

Bias limits	Normal		Exponential	
Level	Lower	Upper	Lower	Upper
60%	76.11	77.81	74.87	79.74
80%	76.24	77.61	75.05	79.31
100%	76.32	77.55	75.17	79.32
120%	76.30	77.61	75.16	79.17

The exponential model takes into account the extension of the tails and therefore produces more accurate estimates of the confidence intervals.

6. Concluding remarks

An approach based on Monte Carlo simulation coupled with heavy and exponential probability tail analysis was presented to estimate confidence intervals under the effect of systematic and random errors on computer models. From the results it is observed that the cumulative probability distribution characteristics can be significantly affected by the error source or type. This suggests that traditional methods for estimating confidence intervals may not be appropriate under the presence of systematic and random errors. It was shown that the use of distribution tail analysis is a robust approach to estimate confidence intervals.

For the thermodynamics case studies presented, the approach for confidence intervals estimation is able to resolve problems such as under and over-estimation of inter-quantile ranges present in the Gaussian approach. This is particularly important for cases where the empirical distributions seem very Gaussian (see the liquid–liquid extraction of acetone case).

Acknowledgments

This work was supported, in part, by National Science Foundation grant CTS—96-96192.

References

Aban, I. B., & Meerschaert, M. M. (2004). Generalized least squares estimators for the thickness of heavy tails. *Journal of Statistical Planning and Inference*, *119*, 341–352.
Abrams, D., & Prausnitz, J. (1975). Statistical thermodynamics of liquid mixtures: A new expression for the excess Gibbs energy of partly or completely miscible systems. *AIChE Journal*, *21*(1), 116–128.
Anderson, P., & Meerschaert, M. (1998). Modeling river flows with heavy tails. *Water Resources Research*, *34*, 2271–2280.
ANSI. (1998). *ANSI/ASME PTC 19.1–1998: Measurement Uncertainty Part 1*. New York, USA: American Society of Mechanical Engineers.
Bancroft, W., & Hubard, S. (1942). A new method for determining dineric distribution. *Journal of the American Chemical Society*, *64*, 347–353.

Brancker, A., Hunter, T., & Nash, A. (1940). The quaternary system acetic acid–chloroform–acetone–water at 25 °C. *Journal of Physical Chemistry*, *44*, 683–698.

Helton, J. C. (1993). Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. *Reliability Engineering and System Safety*, *42*, 327–367.

Henrion, M., & Fischoff, B. (1986). Assessing uncertainty in physical constants. *American Journal of Physics*, *54*(9), 791–798.

Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, *3*(5), 1163–1173.

Hlavaty, K., & Linek, J. (1973). Liquid–liquid equilibriums in four ternary acetic acid–organic solvent–water systems at 24.6 °C. *Collection of Czechoslovak Chemical Communications*, *38*, 374.

ISO. (1995). *Guide to the expression of uncertainty in measurement*. Switzerland: International Organization for Standardization (ISO).

Jansen, D., & de Vries, C. (1991). On the frequency of large stock market returns: Putting booms and busts into perspective. *Review of Economics and Statistics*, *23*, 18–24.

Leadbetter, M., Lindgren, G., & Rootz, H. (1980). *Extremes and related properties of random sequences and processes*. New York: Springer-Verlag.

Leadbetter, M., Lingreen, G., & Rootzén, H. (1983). *Extremes and related properties of random sequences and processes*. New York: Springer.

Loretan, M., & Phillips, P. (1994). Testing the covariance stationarity of heavy tailed time series. *Journal of Empirical Finance*, *1*, 211–248.

Meerschaert, M., & Scheffler, H. (1998). A simple robust estimator for the thickness of heavy tails. *Journal of Statistical Planning and Inference*, *71*, 19–34.

Othmer, D., White, R., & Trueger, E. (1941). Liquid–liquid extraction data. *Industrial & Engineering Chemistry*, *33*, 1240–1248.

Reinders, W., & De Minjer, C. (1947). Vapor–liquid equilibria in ternary systems. VI. System water–acetone–chloroform. *Recueil des Travaux Chimiques des Pays-Bas*, *66*, 573–604.

Resnick, S., & Stărică, C. (1995). Consistency of Hill's estimator for dependent data. *Journal of Applied Probability*, *32*, 139–167.

Ruiz, B., & Prats, R. (1983). Quaternary liquid–liquid equilibria: Experimental determination and correlation of equilibrium data. Part I. System water–acetone–acetic acid–chloroform. *Fluid Phase Equilibria*, *10*, 77–93.

Shlyakhter, A. (1994). An improved framework for uncertainty analysis: Accounting for unsuspected errors. *Risk Analysis*, *14*(4), 441–447.

Smith, B. (1963). *Design of equilibrium stage processes*. New York, USA: McGraw-Hill.

Sokolov, I., & Klafter, J. (2005). A century after Einstein's Brownian motion. *Chaos*, *15*(2), 026103–026110.

Sørensen, J., & Arlt, W. (1980). *Liquid–liquid equilibrium data collection. Vol. 5 of chemistry data series*. Germany: DECHEMA, Frankfurt/Main.

Taylor, J. (1982). *An introduction to error analysis*. California, USA: Oxford University Press.

Thompson, K., Burmaster, D., & Crouch, E. (1992). Monte Carlo techniques for quantitative uncertainty analysis in public health risk assessments. *Risk Analysis*, *12*(1), 53–63.

Treybal, R. (1981). *Mass transfer operations* (Third edition). Singapore: McGraw-Hill.

Vasquez, V., & Whiting, W. (1998). Uncertainty of predicted process performance due to variations in thermodynamics model parameter estimation from different experimental data sets. *Fluid Phase Equilibria*, *142*, 115–130.

Vasquez, V., & Whiting, W. (1999). Evaluation of systematic and random error effects in thermodynamic models on chemical process design and simulation using Monte Carlo methods. *Industrial & Engineering Chemistry Research*, *38*(8), 3036–3045.