Optimal Design of Large-Scale Chemical Processes Under Uncertainty: A Ranking-Based Approach

Sami S. Bahakim, Shabnam Rasoulian, and Luis A. Ricardez-Sandoval
Dept. of Chemical Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada

DOI 10.1002/aic.14515
Published online June 11, 2014 in Wiley Online Library (wileyonlinelibrary.com)

An approach for the optimal design of chemical processes in the presence of uncertainty was presented. The key idea in this work is to approximate the process constraint functions and model outputs using Power Series Expansions (PSE)-based functions. The PSE functions are used to efficiently identify the variability in the process constraint functions and model outputs due to multiple realizations in the uncertain parameters using Monte Carlo (MC) sampling methods. A ranking-based approach is adopted here where the user can assign priorities or probabilities of satisfaction for the different process constraints and model outputs considered in the analysis. The methodology was tested on a reactor–heat exchanger system and the Tennessee Eastman process. The results show that the present method is computationally attractive since the optimal process design is accomplished in shorter computational times when compared to the use of the MC method applied to the full plant model. © 2014 American Institute of Chemical Engineers *AIChE J*, 60: 3243–3257, 2014

Keywords: process design, mathematical modeling, optimization

Introduction

Chemical process design is an essential task performed to achieve the desired throughput and quality of the final products in the presence of safety, environmental, operational, and physical constraints. The design process involves a series of steps that aims to identify the most economically attractive design using steady-state optimization.[1,2] However, optimal steady-state designs may fail to comply with the process constraints when the process under analysis is subject to uncertainties in the model inputs (e.g., the composition of a reactant in a feedstream) or in the model parameters (e.g., the activation energy in a chemical reaction). Efforts to account for uncertainty at the design stage have, therefore, been suggested and widely studied in the literature.[3–5] The straightforward approach used to address this problem is to add overdesign factors; however, this practice of overdesigning a process to ensure feasibility under uncertainty has been proven to be costly, especially in the design of an expensive process unit or when uncertain parameters only affect specific equipment or process units. This has motivated the development of systematic methods that explicitly account for uncertainty in the calculation of the optimal process design. The resulting design represents the most economically attractive process that complies with the process constraints in the presence of uncertainty.

A popular method used to account for uncertainty in process design is referred to as the stochastic programming approach.[6–8] This method evaluates the system's optimal design by performing extensive simulations of the actual plant's model obtained from multiple realizations in the uncertain parameters. Different approaches that employ the stochastic programming approach are available in literature.[9–11] The multiscenario optimization approach is a popular stochastic programming method that has been proposed for optimal process design under uncertainty.[12–17] As discrete sampling is required in the multiscenario optimization approach, the more uncertainty realizations included in the analysis, the more accurate the results at the expense of higher computational costs. The latter is a key limitation of this approach to address the optimal design of large-scale process systems.[18]

The multiscenario approach is suitable when process reliability, that is, full compliance of the process constraints, is critical because it requires feasibility for all the possible uncertain scenarios at minimum cost. However, there are cases where the compliance of specific process constraints is critical (safety-related), whereas violation in other process constraints may be allowed with no actual risk to the process operation or products quality. For example, a slight variation in the liquid level of a storage tank away from its corresponding feasible limits due to parameter uncertainty may cause no serious implications to the plant's economics. Conversely, the variability in a reactor's working temperature outside its feasible limits due to uncertainty in the system parameters, for example, reaction rate kinetic parameters, does impact the plant's economics because it directly affects the products' formation and quality. As compliance of all the process constraints normally require larger (more expensive) designs, it may sometimes be more economical to allow violation of less critical constraints (e.g., tank liquid level) under uncertainty, than to design a robust process that

Correspondence concerning this article should be addressed to L. Ricardez-Sandoval at laricard@uwaterloo.ca.

satisfies the constraints at all times. In the latter case, it is, therefore, desired to develop a ranking-based design approach that ensures feasibility of the critical higher ranked constraints (e.g., reactor temperature) at all times but allows less ranked (noncritical) constraints to be partially violated with the aim of achieving more economical but yet feasible operational process designs.

A systematic method that implements the ranking-based approach is chance constrained programming.[19] In this method, the objective function usually aims to minimize the expected value and variance of the cost function, whereas the constraints are redefined as minimum probability of satisfaction held by the actual physical and process constraints under uncertainty. To calculate the probabilities of constraint satisfaction, monotonic relationships between individual uncertain variables and its corresponding constrained variables are used to map the output boundaries according to the region of uncertain inputs. This information is then used in a multivariate integration to compute the probabilities in the limited region of the uncertain inputs. This chance-constrained programming is then transformed into a deterministic equivalent optimization problem. The main challenge with this approach is the need to evaluate multiple integrals to compute expected values (and/or variances) for the objective function and constraints (at a given probability limit) in the presence of uncertainty. Li et al.[20–22] developed a chance-constrained based methodology that addressees several industrial problems, that is, production planning, process design and operation, optimal control. Ostrovsky et al.[23–25] developed a different approach to transform chance constraints into deterministic constraints using the concept of uncertainty regions. This method searches for the optimal form and location of the "uncertain space," which is a decision variable in the optimization formulation by assuring that the uncertain variables fall within this region with a high probability (close to unity). This approximation in the uncertainty region reduces the computational costs in the evaluation of multiple integrals. Despite the progress made in this area, the need to evaluate multivariate integrals, and the computational effort associated with this calculation, are the main challenges faced toward the application of these methodologies for large-scale industrial chemical plants.

This article presents a practical ranking-based methodology to address the optimal design and operation of large-scale processes under uncertainty. The approach used in this methodology uses power series expansions (PSE) to obtain analytical expressions of the actual process constraints and model outputs in terms of the uncertain parameters considered in the analysis. The resulting PSE analytical expressions are then used to compute the distributions of the process constraints and outputs (i.e., frequency histograms) due to the different (probabilistic-based) realizations in the uncertain parameters. Accordingly, the feasibility in the process constraints is evaluated at a given (user-defined) probability of satisfaction using their corresponding probability distributions. The effect of the system's uncertainty on the cost function can also be assessed in a similar manner to the process constraints. Different process designs alternatives can be assessed when the feasibility in the constraints are set to different probability of satisfaction limits, that is, a ranking-based design. Thus, the present approach aids the user in the decision-making process under uncertainty. The computational benefits and the accuracy in the calculations while using the present ranking-based design methodology are evaluated using a relatively simple case study, that is, a reactor–heat exchanger system, and the Tennessee Eastman (TE) process. The PSE method is utilized as it is a general and practical approach that can be readily implemented to approximate the behavior of nonlinear process models using sensitivity analysis. These features make this approach an attractive and practical alternative, especially for large-scale processes because their corresponding PSE-based functions can be readily estimated using established numerical methods, for example, finite differences. However, the key benefit of this approach is the significant reduction in the computational costs associated with running multiple simulations of the system to estimate the process outputs' distributions under uncertainty.

Pintarič et al.[26] recently proposed an approach to design flexible process flowsheets for systems under uncertainty by performing first-order sensitivity analysis to identify the critical scenarios that may produce the worst-case realizations in the uncertain parameters. These critical points, together with an identified central basic point, were used to evaluate the process constraints, which were found to be sufficient to ensure the design's flexibility. In that work, the objective function is evaluated only at the central basic point, that is, the efforts of evaluating a multidimensional integral are avoided. While that approach has been shown to be computationally attractive, especially when dealing with large number of uncertainties, their method aims to identify robust (conservative) process designs because they need to satisfy the process constraints for the entire space of the uncertain parameters. Thus, the approach is different to that presented here as the design attained by the proposed method in this article is subject to a probability of satisfaction in the process constraint functions and model outputs (i.e., a ranking-based approach) thus allowing the specification of more economically attractive designs. As it is shown in the Case Studies section, the present approach has been compared to Pintarič et al.'s method to evaluate its computational benefits.

The organization of this article is as follows: the next section presents the mathematical framework proposed in this study to compute the distributions in the constraints from knowledge of the uncertainty distribution in the model inputs and model parameters. The systematic method proposed in this work to address the optimal design under uncertainty is presented next. The method is initially tested using a case study that involves the design of a reactor and heat exchanger system. This case study was evaluated under different scenarios, which are aimed to analyze the benefits and limitations of the present approach. A second case study involving the optimal operation of the TE plant[27] demonstrates the computational benefits and accuracy of the present approach to address the optimal design and operation of large-scale systems under uncertainty. Concluding remarks are presented at the end of this article.

Process Design Under Probabilistic-Based Uncertainty

This section presents the methodology for the optimal design of process systems under uncertainty in the model parameters or in the model inputs. The present analysis assumes that a process model $\mathbf{z}$ describing the behavior of

the system under analysis is available for simulations and is described as follows

$$\mathbf{z}(\mathbf{d},\mathbf{p},\mathbf{x},\mathbf{y},\mathbf{u},\boldsymbol{\theta})=\mathbf{0} \tag{1}$$

where $\mathbf{d}$ is the vector of design variables, $\mathbf{p}$ represents the model parameters, whereas $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{u}$ are the state variables, the model outputs and inputs, respectively. Those model parameters in $\mathbf{p}$ and model inputs in $\mathbf{u}$ that are uncertain are defined as $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{u}}$ and will grouped in a single vector and referred from heretofore to as the system's uncertainty $\boldsymbol{\theta}$, that is

$$\boldsymbol{\theta}=[\tilde{\mathbf{p}},\tilde{\mathbf{u}}] \tag{2}$$

This work assumes that each of the system's uncertain model input or model parameter can be described with a particular probability density function (PDF), that is

$$\begin{gathered}\boldsymbol{\theta}=[\theta_1,\theta_2,...,\theta_l,...,\theta_L]\\ \theta_l\sim \mathrm{PDF}(\boldsymbol{\alpha}_l)\end{gathered} \tag{3}$$

where the lth uncertain variable included in $\boldsymbol{\theta}$ follows a specific PDF with distribution parameters $\boldsymbol{\alpha}_l$. The choice of the type of PDF to describe each uncertain parameter comes from process experience (when designing plants similar to an existing one) or historical data of the plant (when using the same input uncertainty source where such details are available). In the latter case, the distribution of the input uncertainty can be characterized by fitting the best PDF that describes the available data. When no such information is available, typically the PDFs of the uncertain variables are described using Gaussian or uniform probability distribution functions; however, the present method is not restricted to these functions and assumes that each uncertain parameter, such as θ_l, can be described using any symmetric or nonsymmetric probability distribution function, for example, lognormal, exponential. Note that the description presented in (3) does not assign specific values for the uncertain variables. Hence, an appropriate sampling technique such as the Monte Carlo (MC) sampling method is needed to obtain the different realizations in $\boldsymbol{\theta}$. MC sampling in the proposed approach chooses N independent sample points randomly from the known PDFs (with distribution parameters $\boldsymbol{\alpha}_l$) of the uncertain variables θ_l. This is typically a standard task using available off-the-shelf computing software. Unlike the stochastic programming method, the present approach does not simulate the plant model $\mathbf{z}$ for all N uncertain realizations to compute the variability in the constraints due to $\boldsymbol{\theta}$. Instead, the present method represents the process constraint functions and model outputs using a PSE function. Then, N MC sampling points representing the uncertain parameters' distribution are generated and used as inputs to simulate the corresponding PSE functions. The simulation results, describing the variability in the constraint functions and the model outputs due to θ_l, are then used to evaluate the feasibility and economics of the current design under analysis. As it will be shown in the next sections, the evaluation of the constraints and model outputs using PSE functions is a much less intensive task than simulating the actual process model ($\mathbf{z}$) N times, especially for large complex models. The procedure to obtain the PSE constraint functions is described next.

PSE method: Analytical approximation of the process constraints

The process constraints are typically described as a function of the system parameters, state variables as well as the model inputs and outputs, that is

$$\mathbf{h}(\mathbf{p},\mathbf{x},\mathbf{y},\mathbf{u},\boldsymbol{\theta})\le\mathbf{0} \tag{4}$$

where the vector of process constraints $\mathbf{h}$ usually impose a safety, physical or operational limitation on the process to be designed. The key idea on this method is to compute analytical expressions for each of the process constraint functions included in $\mathbf{h}$ due to the potential realizations in $\boldsymbol{\theta}$ using PSE functions. Therefore, the actual nonlinear constraint function h is represented in the present analysis as follows

$$\begin{aligned}&h(\boldsymbol{\theta})=h(\bar{\boldsymbol{\theta}})+\mathbf{M}^{(1)}(\boldsymbol{\theta}-\bar{\boldsymbol{\theta}})+\frac{1}{2}(\boldsymbol{\theta}-\bar{\boldsymbol{\theta}})^T\mathbf{M}^{(2)}(\boldsymbol{\theta}-\bar{\boldsymbol{\theta}})+...\\ &\mathbf{M}^{(1)}=\left.\frac{\partial h}{\partial\boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\bar{\boldsymbol{\theta}}}\\ &\mathbf{M}^{(2)}=\left.\frac{\partial^2 h}{\partial\boldsymbol{\theta}^2}\right|_{\boldsymbol{\theta}=\bar{\boldsymbol{\theta}}}\\ &E[\boldsymbol{\theta}]=\bar{\boldsymbol{\theta}}\end{aligned} \tag{5}$$

where $\mathbf{M}^{(i)}$ refers to the ith sensitivity term of the process constraint function h, that is, $\mathbf{M}^{(1)}$ and $\mathbf{M}^{(2)}$ represent the Jacobian and Hessian matrices of the process constraint function h, respectively. Similarly, $\bar{\boldsymbol{\theta}}$ represents the nominal (mean) value of the uncertain parameters $\boldsymbol{\theta}$. The PSE constraint function shown in (5) shows only up to an expansion order of 2, but it can be easily expanded for any higher order q. The constraint function h is assumed to be $(q+1)$ times totally differentiable with respect to $\boldsymbol{\theta}$. For the case, when a constraint h is a univariate function of a single uncertain variable θ, the resulting PSE approximation of h from (5) simplifies to the following expression

$$h(\theta)=h(\bar{\theta})+\sum_{j=1}^{q}\frac{1}{j!}\left(\frac{\partial^j h}{\partial\theta^j}\right)\Bigg|_{\theta=\bar{\theta}}(\theta-\bar{\theta})^j \tag{6}$$

As shown in (5), the PSE-based expansion model used to describe the constraint functions is easier to evaluate because it has an analytical expression, that is, explicitly defined in terms of the system's uncertain parameters $\boldsymbol{\theta}$. Therefore, the PSE constraint functions can be used to evaluate the process constraints $\mathbf{h}$ due to multiple realizations in the uncertain parameters $\boldsymbol{\theta}$ with minimum computational effort. By substituting the N sampled uncertainty realizations $\boldsymbol{\theta}_N\in\Re^{N\times L}$ into each PSE-based process constraint expression similar to that shown in (5), a set of estimates on each constraint function h corresponding to each realization of the N sampled uncertain variables is obtained, that is, $h(\boldsymbol{\theta}_N)\in\Re^{N\times 1}$. The set of estimates collected in $h(\boldsymbol{\theta}_N)$are then used to generate a frequency histogram that describes the distribution (variability) of the constraint function h due to $\boldsymbol{\theta}$. The accuracy of the estimated distribution of the process constraints (depicted by the histogram) improves as N becomes larger and as the number of expansion terms considered in the PSE-based constraint function for h is increased. While increasing the order of the expansion improves the analytical approximation of the process constraint, it also increases the computational

cost due to evaluation of higher order terms in the PSE-based expansion. Therefore, the choice of the expansion order in the PSE is problem specific because it depends on many aspects of the system under analysis, for example, the degree of nonlinearity of the system, the size of the process model, the method used to compute the terms in the expansion, that is, analytically or numerically,[28] the probability distribution function assigned to the system's uncertain parameters $\boldsymbol{\theta}$. These particular aspects of the method are further analyzed with the case studies presented in this work. Higher order PSE expansions may be required to represent systems with strong nonlinearities. In those cases, the computational costs of the present method will increase, especially when many uncertain parameters are considered in the analysis, because the proposed method relies on the calculation of the system's sensitivities to the uncertain parameters.

In the present approach, the distribution of each of the process constraints included in $\mathbf{h}$ will be used as a tool to implement a ranking-based design approach. A significance in the process constraints is specified by assigning a probability limit (Pb_h) to each constraint, that is, each constraint included in the analysis is ranked using Pb_h based on its significance. This probability of satisfaction indicates how often a particular constraint is expected to meet its corresponding feasible limits. Thus, to ensure a feasible process design, the variability in the process constraints $\mathbf{h}$ cannot violate its prespecified limits (Pb_h) due to the different realizations in $\boldsymbol{\theta}$. Using the assigned Pb_h limit and $h(\boldsymbol{\theta}_N)$ obtained from the PSE-based constraint function, an estimate of the extreme possible value of the process constraint function ξ_h that occurs ($100 \times \mathrm{Pb}_h$)% of the time can be evaluated as follows

$$\begin{aligned} \mathrm{Pb}_h &= F(\xi_h, h(\boldsymbol{\theta}_N)) = \mathrm{P}(h \le \xi_h) \\ \xi_h &= \{\xi_h : F(\xi_h, h(\boldsymbol{\theta}_N)) = \mathrm{Pb}_h\} = F^{-1}(\mathrm{Pb}_h, h(\boldsymbol{\theta}_N)) \end{aligned} \tag{7}$$

where F is the cumulative probability function of the process constraint function. Figure 1 presents a schematic representation of the PSE-based constraint function h evaluated at a given probability Pb_h. If the extreme possible value of the constraint ξ_h (calculated at a given Pb_h) satisfies the process restrictions specified in (4), a feasible process design is obtained with a ($100 \times \mathrm{Pb}_h$)% guarantee that constraint h is satisfied. Based on the above, the constraints (4) under system uncertainty are evaluated in the present analysis using the constraint function's extreme estimates (ξ_h) at a user-defined probability of satisfaction (Pb_h), that is

$$\xi_h(F, h(\boldsymbol{\theta}_N), \mathrm{Pb}_h) \le 0 \quad \Longleftrightarrow \quad \mathrm{P}(h(\mathbf{p}, \mathbf{x}, \mathbf{y}, \mathbf{u}, \boldsymbol{\theta}) \le 0) \ge \mathrm{Pb}_h \tag{8}$$

The expression on the right-hand side in (8) denotes the probability of the actual nonlinear constraint functions obtained with the actual nonlinear plant model $\mathbf{z}$, whereas the extreme estimate on the left-hand side is obtained from the PSE-based constraint function shown in (5) followed by a probabilistic inference. Because ξ_h is only an estimate of the constraint at a given probability limit Pb_h, there is also a probability of $(1 - \mathrm{Pb}_h)$ where h can take values beyond ξ_h. In the proposed approach, the probability of constraint satisfaction Pb_h, which determines the rank assigned to each constraint, represents an input to the present method, that is, it is a user-defined input parameter. Thus, higher probabilities should be assigned to those constraints that are considered to be critical. Setting $\mathrm{Pb}_h \to 1$ implies that the design will satisfy the constraints almost all the times, this is often termed as the worst-case scenario approach (see Figure 1). While the worst-case scenario ensures that the design remains feasible for almost surely all the realizations in $\boldsymbol{\theta}$, this robust design is typically conservative and expensive. The subscript in Pb_h suggests that different probabilities can be assigned for each process constraint enabling a ranking-based design. The latter will assist in achieving less conservative (economically attractive) designs but at the same time keeping the critical constraints within specification (at a given probability of occurrence). Therefore, the selection of a suitable ranking structure is problem-specific because it depends on the goals to be attained by the design, for example, process economics and process safety. The analysis described above for the process constraints can also be implemented in the same fashion to evaluate the variability in the model outputs and state variables that are included in the plant's cost function.

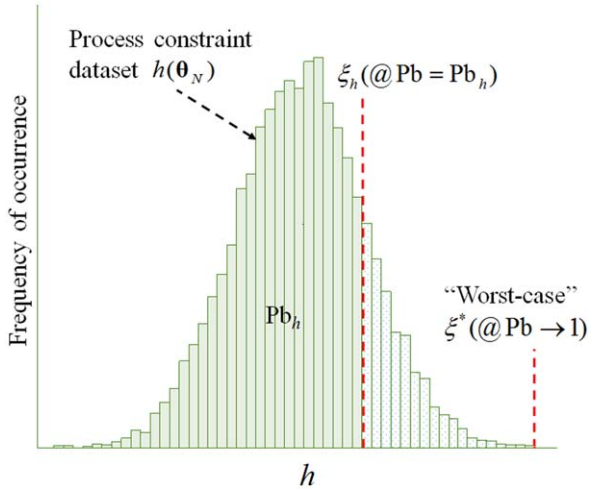

Figure 1. Schematic representation of the constraint function's distributional analysis.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Optimal design under uncertainty

Based on the above developments, the optimal design of a chemical system under uncertainty can be formulated as follows

$$\begin{aligned} \min_{\boldsymbol{\eta}=[\mathbf{d},\mathbf{u}]} \ & \phi = \mathrm{CC}(\mathbf{d}, \boldsymbol{\theta}) + \mathrm{OC}(\mathbf{x}, \mathbf{y}, \mathbf{u}, \boldsymbol{\theta}) \\ & \text{s.t.} \\ & \mathbf{z}(\mathbf{d}, \mathbf{p}, \mathbf{x}, \mathbf{y}, \mathbf{u}, \boldsymbol{\theta}) = \mathbf{0} \\ & \boldsymbol{\xi}_h(F, h(\boldsymbol{\theta}_N), \mathrm{Pb}_h) \le 0 \qquad h \in \mathbf{h} \\ & \mathbf{d}^l \le \mathbf{d} \le \mathbf{d}^u \\ & \mathbf{u}^l \le \mathbf{u} \le \mathbf{u}^u \end{aligned} \tag{9}$$

The above problem aims to minimize the economic cost of the process, usually described in terms of the capital cost (CC) and operating cost (OC), by selecting feasible process designs $\mathbf{d}$ and operating conditions (i.e., model inputs $\mathbf{u}$). The feasibility criteria follow a ranking-based approach where each constraint function h can take different values for Pb_h, which becomes the constraints' minimum probability of satisfaction. Despite the stochastic nature of this process design formulation, that is, each uncertain parameter is described with a specific PDF as shown in (3), the implementation of the PSE-

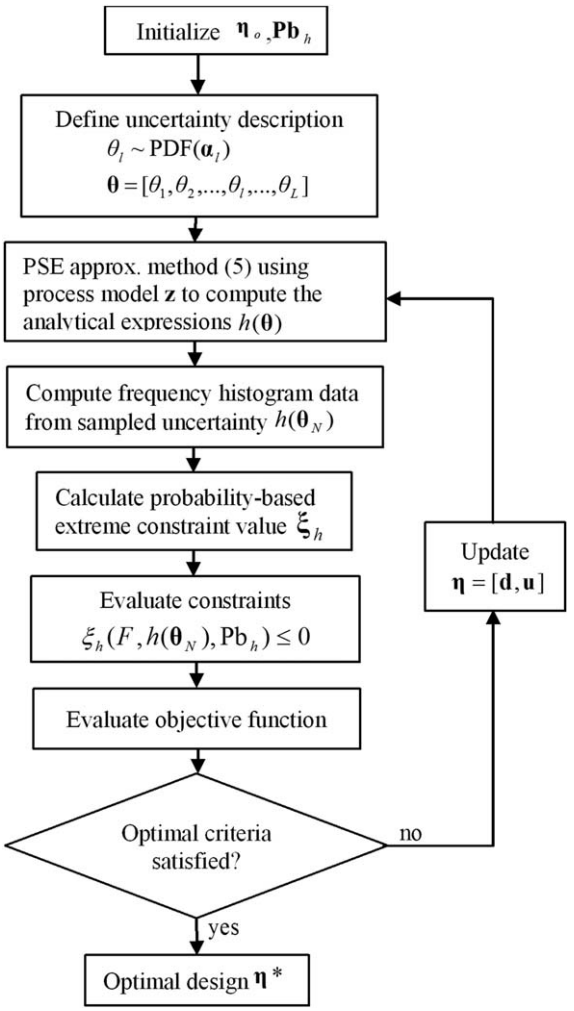

Figure 2. Algorithm proposed for the optimal design under uncertainty.

based approach proposed here to evaluate the process constraints and cost function under uncertainty in the parameters $\boldsymbol{\theta}$ reduces problem (9) into a deterministic nonlinear constrained optimization problem that can be solved using available NLP solvers such as sequential quadratic programming (SQP).[29] The outcome of the present formulation returns an optimal process design that accommodates uncertainty in the system's parameters $\boldsymbol{\theta}$ up to a user-defined probability of constraints satisfaction (Pb_h). Figure 2 shows the algorithm for the optimal plant design under uncertainty presented in this work. As shown in this figure, the plant model **z** is only required in the PSE approximation step to calculate the expansion terms in (5). Note that the present method uses the PSE-based model developed for each process constraint function h to compute the distributions on each constraint instead of simulating the nonlinear process model **z** for each uncertain realization in the parameters $\boldsymbol{\theta}$. The latter approach reduces the computational effort considerably especially for the optimal design of large-scale systems. Note from Figure 2 that the PSE-based constraint functions need to be recalculated at each iteration step in the optimization algorithm, which represents the most significant computational cost in the present approach.

Remarks

In the present approach, the equality constraints are satisfied for the entire space of the uncertain parameters. That is, the equality constraints, which are typically the process model equations in an optimal process design problem, that is, **z** in problem (9), need to be solved for the different realizations considered in the uncertain parameters. This specification enables the evaluation of the process outputs' variability under uncertainty using the process model equations, which are then used to compute the distribution of the process constraint functions and model outputs. The present approach can also account for structural (discrete) decisions in the analysis. The computation of the sensitivities in the present approach only requires that the process model is continuous. In problems involving discrete decisions, the discrete variables are fixed in advance and the subproblems to be solved have continuous process models that impose no restriction in the use of PSE to approximate and compute the process constraints and model outputs.

Note that in the case where the uncertainty is bounded but its distribution is unknown, a uniform distribution assumption would be adopted as this kind of distribution is the most pessimistic, that is, more realizations of the extreme values may occur. For a uniform distribution, any value within the specified bounds has equal probability of occurrence, unlike other distributions such as the normal distribution where the majority of realizations will be close to the mean value and only a rare occurrence for the extreme points. As a result, the uniform distribution assumption yields conservative designs and is adopted in the present methodology for those cases of unknown uncertainty distributions. In the next section, the application of the present approach to address the optimal process design and operation of two case studies under different scenarios is presented. The studies presented in the next section were performed on an Intel Core i7 3770 CPU @3.4GHz (8GB in RAM).

Case Studies

Case study 1: Reactor–heat exchanger system

The first case study considered in the present analysis is a plug-flow reactor coupled with a heat exchanger system as shown in Figure 3. This system has been previously studied for optimal process design.[30] A first-order exothermic reaction is assumed for the production of product B from reactant A in the direct reaction: A $\rightarrow$ B. F_0 is the reactor's feed flowrate with temperature T_0 and initial concentration C_{A0} of reactant A. The remaining concentration of the reactant in the product stream is denoted by C_{A1}.

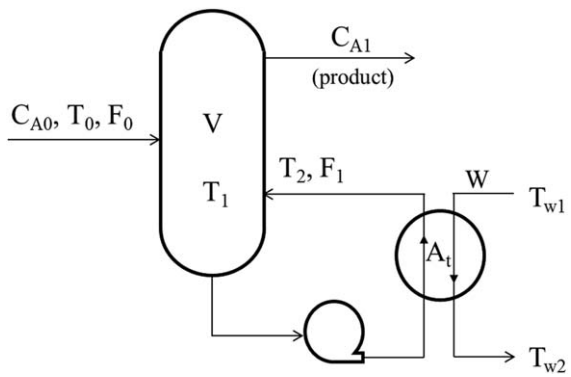

Figure 3. Flowsheet of a reactor–heat exchanger system.

The variables T_1, T_2, and F_2 are the reactor's temperature, and the temperature and flow rate of the recycled stream, respectively. The recycled stream is cooled down in the heat exchanger using cooling water supplied at a temperature T_{w1} and flow rate W to ensure that the reaction temperature T_1 does not exceed a maximum temperature limit. The material and energy balances of the reactor and heat exchanger represent the process model for this system and are as follows:

Reactor's mass and energy balance

$$F_0(C_{A0}-C_{A1})/C_{A0}=Vk_0\exp(-E/RT_1)C_{A1}$$
$$(-H)F_0(C_{A0}-C_{A1})/C_{A0}=F_0C_p(T_1-T_0)+Q_{HE} \tag{10}$$

Heat exchanger's energy balance

$$Q_{HE}=F_2C_p(T_1-T_2)=C_{pw}(T_{w2}-T_{w1})W$$
$$Q_{HE}=A_tU(0.5((T_1-T_{w2})+(T_2-T_{w1}))) \tag{11}$$

where k_0, E, and H are the rate constant, activation energy, and the heat of reaction, respectively. Q_{HE} is the heat transfer, U is the overall heat transfer coefficient, whereas the heat capacity of the recycled flow and cooling water are represented by C_p and C_{pw}, respectively. The design parameters are the reactor's volume V, and the heat exchanger's transfer area A_t. In addition, X represents the rate of conversion of reactant A. The state variables are C_{A1}, T_2, F_2, and W, which can be eliminated by analytic expressions using Eqs. 10 and 11 as follows

$$C_{A1}=C_{A0}(1-X)$$
$$X=\frac{Vk_0C_{A0}\exp(-E/RT_1)}{F_0+Vk_0C_{A0}\exp(-E/RT_1)}$$
$$T_2=\frac{2(-H)F_0X}{A_tU}-\frac{2F_0C_p(T_1-T_0)}{A_tU}-(T_1-T_{w2})+T_{w1} \tag{12}$$
$$F_2=\frac{Q_{HE}}{C_p(T_1-T_2)},\quad W=\frac{Q_{HE}}{C_{pw}(T_{w2}-T_{w1})}$$

The nominal values for the model parameters of this process are listed in Table 1.[30] In this case study, uncertainty is assumed in two model inputs (F_0 and C_{A0}) and one model parameter (k_0). These uncertain parameters were assumed to follow a normal probability distribution with specific mean and variances, that is

$$\boldsymbol{\theta}_{CS1}=[F_0,C_{A0},k_0]$$
$$F_0\sim N(45.36,1.506^2)$$
$$C_{A0}\sim N(32.04,1.266^2) \tag{13}$$
$$k_0\sim N(0.6242,0.079^2)$$

As shown in (13), the expected values of these distributions correspond to the nominal operating conditions shown in Table 1 for each of these parameters. A variance of 5% of its mean is assumed for the feed flow rate F_0 and concentration C_{A0}, respectively, whereas a variance of 1% of its nominal value was assigned for the more sensitive parameter k_0. The goal of this process is to achieve a minimum of 90% conversion of reactant A while maintaining temperature constraints in the process units (see Table 1). The decision variables for this case study consist of the design parameters

Table 1. Reactor-Heat Exchanger Case Study: Model Parameters and Process Constraints

Model Parameters		
Variable	Estimate	Units
E/R	555.6	K
H	−23,260	kJ/kg mol
U	1635	kJ/m^2 h K
C_p	167.4	kJ/kg mol
C_{pw}	75.327	kJ/kg mol
F_0	45.36	kg mol/h
k_0	0.6242	m^3/kg mol h
C_{A0}	32.04	kg mol/m^3
T_0	333	K
T_{w1}	300	K
Process Constraints		
$c1$	$0.9-X\le 0$	
$c2$	$X-1\le 0$	
$c3$	$T_2-T_1\le 0$	
$c4$	$311-T_2\le 0$	
$c5$	$T_2-389\le 0$	
$c6$	$T_{w1}-T_2+11.1\le 0$	
$c7$	$T_{w1}-T_{w2}\le 0$	
$c8$	$T_{w2}-T_1+11.1\le 0$	
$c9$	$311-T_1\le 0$	
$c10$	$T_1-389\le 0$	
$c11$	$301-T_2\le 0$	
$c12$	$T_{w2}-355\le 0$	

$\mathbf{d}=[V,A_t]$ and the operating variables $\mathbf{u}=[T_1,T_{w2}]$. Using the PSE approximation method presented in the previous sections, the variability in the process constraints due to uncertainty in $\boldsymbol{\theta}_{CS1}$ are calculated for each set of values in the design variables $\mathbf{d}$ and $\mathbf{u}$ tested by the optimization algorithm used to solve this process design problem. Using a probability of satisfaction of $Pb_h=0.6827$ for all constraints, the extreme possible values (ξ_h) at that probability level can be estimated as shown in (7). For example, for constraint $c1$ in Table 1, the probabilistic form as in (8) is as follows

$$\xi_{c1}\le 0\iff P(0.9-X\le 0)\ge Pb_h \tag{14}$$

The rest of the process constraints shown in Table 1 were reformulated in a similar fashion. Using the constraint formulation shown in (14) together with the method explained above to estimate ξ_h, the constraints are aimed to satisfy a probability of $Pb_h=0.6827$, that is, 68.27% of the time. Based on the above, the optimal design problem shown in (9) is reformulated for the present case study as follows

$$\min_{\boldsymbol{\eta}=[V,A_t,T_1,T_{w2}]}\ \phi=691.2V^{0.7}+873.6A_t^{0.6}+1.76W+7.056F_2$$
$$\text{s.t.}$$
$$\text{Process model eqns (11)}-\text{(12)}$$
$$\xi_w\le 0\qquad \forall w=c1,c2,...,c12$$
$$V^l\le V\le V^u\qquad A_t^l\le A_t\le A_t^u$$
$$T_1^l\le T_1\le T_1^u\qquad T_{w2}^l\le T_{w2}\le T_{w2}^u \tag{15}$$

where the objective function is a combination of the CC and OC of the plant that was taken from the literature.[30] Problem (15) was solved with MATLAB's SQP solver using different orders q for the PSE approximation and for different probability limits, Pb_h.

As shown in Table 2, it is clear that the design obtained from the first-order PSE approximation ($q=1$) is different

Table 2. Result for the Reactor-Heat Exchanger System, Optimal Designs for Different PSE Orders (Pb_h=0.6827)

Exp. Order	$q = 1$	$q = 2$	$q = 3$	$q = 4$
V (m^3)	90.77	92.13	91.65	91.70
A_t (m^2)	5.97	6.01	6.00	6.00
T_1 (K)	389.00	389.00	389.00	389.00
T_{w2} (K)	329.63	329.61	329.61	329.60
Costs ($/yr)	19,757	19,933	19,871	19,878
CPU time (s)	1.03	3.91	12.40	29.81

from those obtained with high orders in the PSE approximation function. Note that process designs obtained from a third-order ($q = 3$) and a fourth-order ($q = 4$) PSE approximation did not change significantly, which indicates convergence of the PSE approximation function at a given probability of satisfaction. This shows that the present methodology can account for the system's nonlinearity using higher order PSE approximations to achieve accurate optimal designs. The expense of using higher orders in the PSE approximation can be observed in the increase of the computational time needed to solve the optimal design problem for these scenarios. In this case study, the maximum difference between the solutions obtained from orders $q = 3$ and $q = 4$ is only about 0.05% though the lower order identifies an optimal design in half the time than that needed by the fourth-order PSE approximation. Therefore, it is reasonable to select the PSE order to $q = 3$ without losing accuracy in the results. Note that, although it is established that the solution of this case study converged using higher order PSE functions, the design obtained from a first-order PSE approximation is still feasible. An order of $q = 3$ PSE was selected solely for the purpose of demonstrating the convergence property of using higher orders for systems with strong nonlinearities.

Table 3 shows the optimal process design alternatives obtained when probability limits Pb_h were set to different values for all process constraints while using third-order PSE-based approximation functions ($q = 3$). As expected, when a higher probability of constraint satisfaction is chosen, a larger reactor and heat transfer area in the heat exchanger are required, which directly translates into higher plant costs. For example, a 15% increase in total costs is needed to design a plant that will satisfy the constraints with a probability of 99.73% as opposed to the design that satisfies all the process constraints 68.27% of the time. As explained above, the choice of Pb_h is user-defined; its direct relation to the plant's profitability can be clearly assessed using the present ranking-based method. Note that the computational times for different choices of Pb_h while using third-order PSE approximation functions is relatively the same. These designs were validated through simulations of the actual plant model, that is, Eqs. 10 and 11, using 100,000 MC realizations in the 3 uncertain variables included in $\boldsymbol{\theta}_{CS1}$. Figure 4 shows that the minimum conversion rate constraint ($c1$) complies with their corresponding prespecified probability limits. The rest of the process constraints were validated in the same fashion and are not shown here for brevity.

To verify the accuracy of the results and the computational costs obtained while using the present method, the present case study was also solved using a stochastic programming technique that uses the MC sampling method applied to the actual process model. That is, at each optimization step, random MC realizations of the uncertain parameters $\boldsymbol{\theta}_{CS1}$ are generated and used to simulate the complete process model. The results from the simulations are then used to obtain the output distributions and evaluate the compliance of the constraints and the cost function. For the present analysis, the actual plant model (10) and (11) is simulated for each realization and a histogram is obtained for the constraints' distribution due to uncertainty in $\boldsymbol{\theta}_{CS1}$. From these histograms, the constraints' extreme possible value at a given probability limit Pb_h is computed as shown in Figure 1; however, the distribution at each single function evaluation of the optimization algorithm is now obtained from actual simulations of the plant model rather than from the PSE approximation method. For the present analysis, 100,000 random MC sampling points were used in this approach. The optimal process design obtained at the user-defined probability limit Pb = 0.9545 for all constraints is presented in Table 4. This result is in close agreement to that obtained by the present method shown in Table 3, that is, the proposed method returned plant designs that are accurate within an error of less than 4%. Also, the present PSE-based methodology achieved a solution a few orders of magnitude faster than the MC-based stochastic approach. Sampling realizations of the uncertain parameters from their respective probability distributions is a requirement of stochastic programming, and thus the efficiency of the sampling method used in the analysis is a factor that contributes towards the computational demand costs.[31–33] To further demonstrate the computational benefits of the present approach, the present case study was also solved using the stochastic approach employing the Halton sampling technique,[31] which is known to be more efficient than the standard MC sampling method. As shown in Table 4, it is clear that the Halton-based stochastic approach is more efficient than the MC-based approach because the computational costs are reduced by an order of magnitude due to fewer simulations needed to attain the same convergence. However, the computational benefits of the present PSE-based approach are still apparent when compared with the Halton-based stochastic approach (i.e., at least by one order of magnitude). It is important to note that, while stochastic programming approaches require simulating the process model for each sampled point, different sampling methods differ only in the number of simulations N that is considered sufficient to obtain an accurate solution under the uncertainty conditions. Conversely, the present PSE-method is independent of the used sampling method, and simulates the process model only a few times ($<<$N) depending on the order used to calculate the sensitivities. As such, regardless of the efficiency of the sampling technique, the proposed method is less computationally intensive compared to stochastic programming approaches when handling large-scale problems.

To further demonstrate the ranking-based feature of the present approach, two additional cases (i.e., case A and B),

Table 3. Results for the Reactor-Heat Exchanger System, Optimal Designs for Different User-Input Probability Pb (PSE Method, $q = 3$)

Pb_h	0.5	0.6827	0.9545	0.9973
V (m^3)	86.56	91.65	104.69	116.42
A_t (m^2)	5.94	6.00	6.14	6.24
T_1 (K)	389.00	389.00	389.00	389.00
T_{w2} (K)	329.57	329.61	329.69	329.75
Costs ($/yr)	19,205	19,871	21,519	22,941
CPU time (s)	10.29	12.40	16.55	15.45

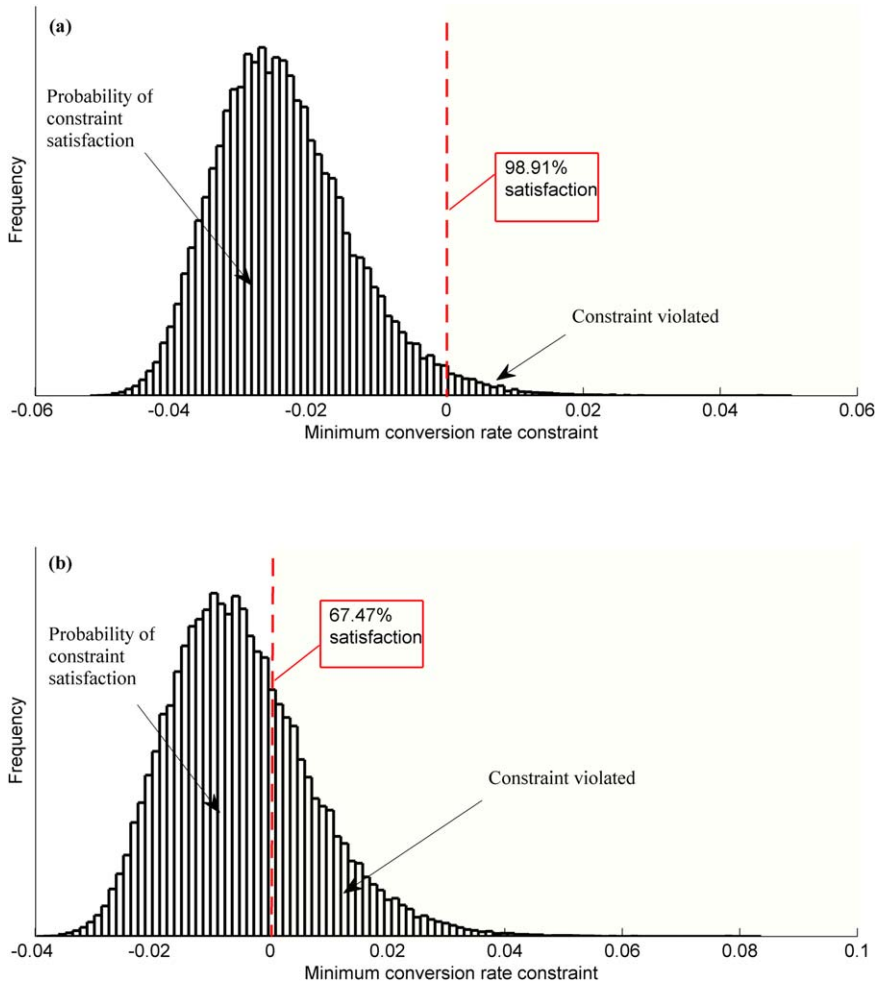

Figure 4. Frequency histogram of the minimum conversion rate constraint at probability limits (a) Pb = 0.9973, (b) Pb = 0.6827. Dashed line represents the maximum constraint limit.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

with different probability of constraint satisfaction assigned to the different process constraints, were considered. In this case, the constraints on the conversion rate (X) shown in Table 1 (i.e., $c1$ and $c2$) were assigned to 80% and 95% for cases A and B, respectively, whereas the recycled stream temperature (T_2) constraints, that is, $c3$–$c6$ in Table 1, were set to 68% for both case A and B. The remaining constraints (i.e., $c7$ to $c12$ in Table 1) were kept at high probabilities of satisfaction (99%). The optimal designs specified for cases A and B were obtained using third-order PSE approximation functions and are presented in Table 5. As shown in that table, a different set of optimal designs were obtained when compared to the case of assigning equal probabilities to all constraints, for example, see Table 4. The results show that increasing the probability of satisfaction for the conversion rate constraints from 80 to 95% is mainly achieved by increasing the volume of the reactor, which caused the total plant costs to increase by more than 5%. The histograms obtained from simulations of the actual plant model using different MC realizations in $\boldsymbol{\theta}_{CS1}$ indicate that the designs obtained for case A and B complies with the probability limit (Pb_{c1}) assigned to the constraints, for example, Figure 5 shows the compliance of the minimum conversion rate constraint at the corresponding probability limits specified for case A and B, respectively.

Table 4. Results for the Reactor-Heat Exchanger System, Optimal Design Obtained Using a Stochastic Approach that Implements a Different Sampling Technique ($Pb_h = 0.9545$)

Sampling Method	Monte Carlo	Halton
V (m^3)	108.06	107.50
A_t (m^2)	6.149	6.137
T_1 (K)	388.98	388.91
T_{w2} (K)	333.29	330.74
Costs ($/yr)	21,885	21,872
CPU time (s)	3,369	914

Table 5. Results for the Reactor-Heat Exchanger System, Ranking-Based Optimal Design Using PSE Method with Order ($q = 3$)

	Case A	Case B
V (m^3)	95.58	104.20
A_t (m^2)	6.05	6.14
T_1 (K)	389.00	389.00
T_{w2} (K)	329.63	329.69
Costs ($/yr)	20,376	21,459
CPU time(s)	12.61	15.19

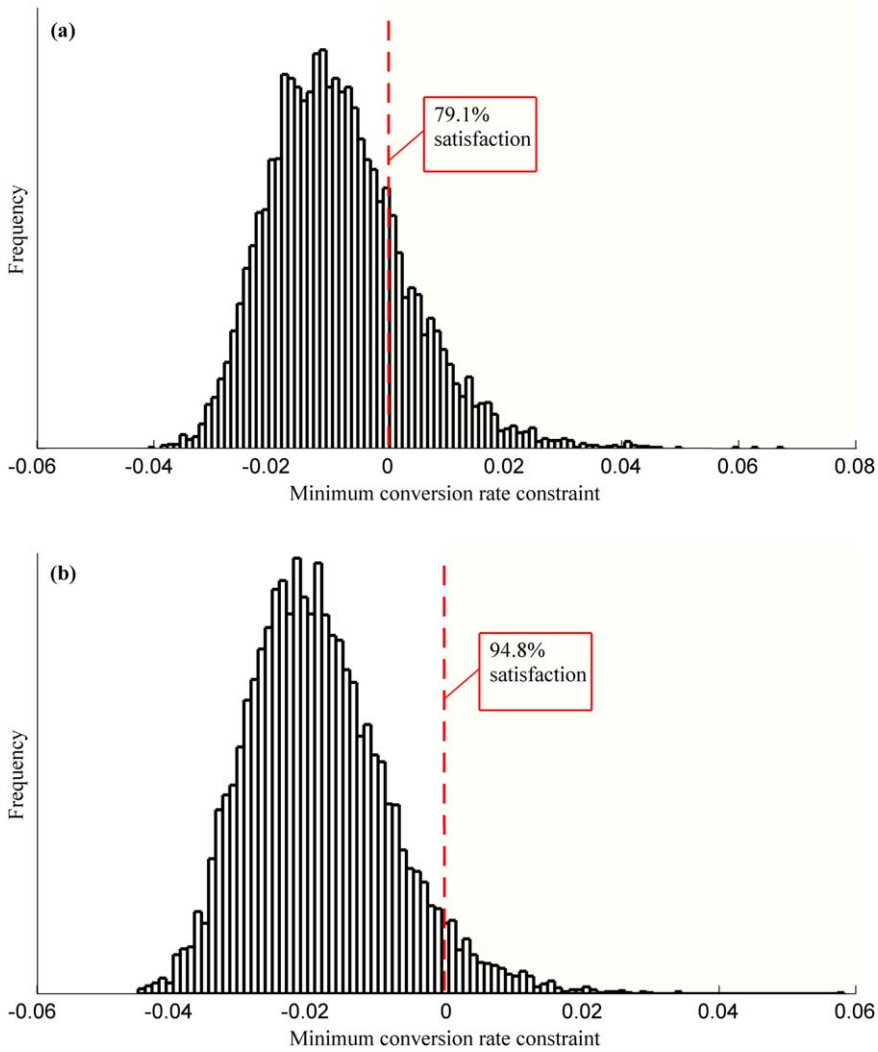

Figure 5. Frequency histogram for the minimum conversion rate constraint: (a) Case A, (b) Case B.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

As mentioned above, a recent computationally attractive approach presented by Pintarič et al. was presented in the literature to address the optimal process design under uncertainty. A comparison between the present approach and Pintarič's method can only be made for the worst-case scenario, that is, robust designs that satisfy process constraints all the time, with no ranking feature. One key feature of the methodology proposed in this work is that it can be used to approximate the robust optimization process design problem under uncertainty by setting the probability of satisfaction on the process constraints to values close to unity. To provide further insight and compare the computational efficiency of the proposed approach to recent computationally attractive approaches proposed in the literature, the present case study was solved for the worst-case scenario using both, the method proposed in this work (i.e., setting Pb $\rightarrow$ 1 for all constraints) and Pintarič's approach. As shown in Table 6, the computational times required by both approaches to achieve the optimal (robust) process design are comparable, that is, of the same order of magnitude, suggesting that the present approach is also a computationally attractive method for optimal (robust) process design under uncertainty. A first-order PSE was sufficient to obtain an optimal feasible design while using the present approach. That design was found to satisfy the constraints at a very high probability of satisfaction, that is, $Pb_h = 0.9999$, which approximates to the robust design obtained while using Pintarič et al.'s approach. The next case study aims to study the feasibility to use this PSE approach to address the optimal operation of a large-scale processing plant.

Case study 2: The TE process

The TE process is a widely used industrial problem that was proposed by Downs and Vogel[27] based on an actual process of the TE company. The process consists of a reactor, a recycle compressor, a partial condenser and a flash separator

Table 6. Results for the Reactor-Heat Exchanger System, Optimal Designs Using Different Approaches for the Worst-Case Problem

	Pintarič et al.	PSE ($q = 1$)
V (m^3)	171.2434	169.02
A_t (m^2)	6.5104	6.510
T_1 (K)	389.00	389.00
T_{w2} (K)	329.94	329.94
Costs ($/yr)	29,038	28,807
CPU time (s)	2.025	1.887

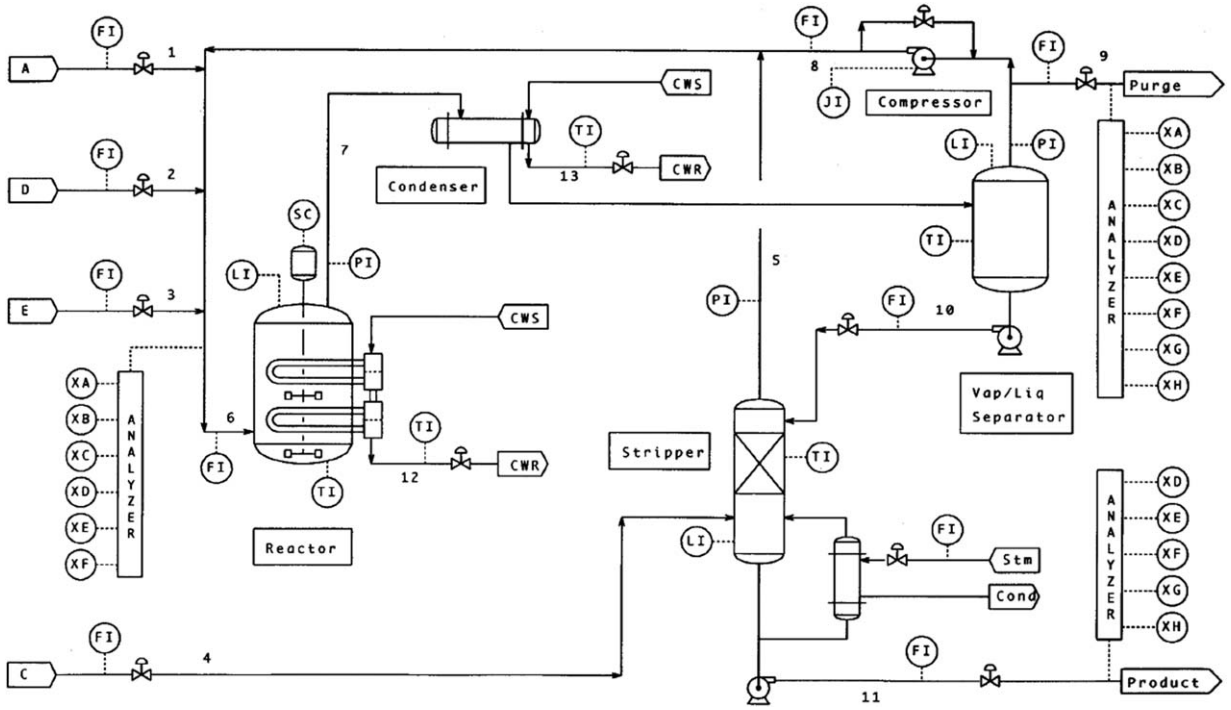

Figure 6. Schematic flowsheet of the TE process.

to produce two liquid products (G and H) using four gaseous reactants, that is, A, C, D, and E, from the following reactions

$$\begin{aligned} A\ (g) + C\ (g) + D\ (g) &\rightarrow G\ (liq) \\ A\ (g) + C\ (g) + E\ (g) &\rightarrow H\ (liq) \\ A\ (g) + E\ (g) &\rightarrow F\ (liq) \\ 3D\ (g) &\rightarrow F\ (liq) \end{aligned}$$

The feed has an inert component B, and F is a by-product. The plant (shown in Figure 6) can operate at different production mix rates of G and H depending on market fluctuations. The base case is a 50/50 production in both G and H at a production rate of 7038 kg/h; this is the case considered in the present study. The four reactions in the reactor are defined as irreversible exothermic reactions; the reaction rates are temperature-dependent and can be described by an Arrhenius-like function. The reactions are approximated by first-order kinetics with respect to the reaction concentrations. As shown in Figure 6, the reactants A, D, and E in the feed stream enter the reactor unit together with the gaseous recycled stream where they react to form the desired liquid products G and H. A nonvolatile catalyst dissolved in the liquid phase is used to drive the gas-phase reactions and the products exit the reactor along with some unreacted gases in a vapor phase. The liquid products are condensed and separated from the gaseous mixture in the partial condenser and flash separator respectively, while the noncondensable components (unreacted gases) are recycled back to the reactor through a centrifugal compressor. The liquids collected at the bottom of the separator is pumped to the stripper which helps recover the remaining unreacted species D and E which would otherwise be lost in the product stream. This separation in the stripper is achieved by using a mixture of A and C as the solvent stream entering the base of the stripper (stream 4 in Figure 6), sending the vapor stream leaving the top of the stripper back to the reactor through the mixed recycle stream. The liquid stream at the bottom of the stripper is refined by heating with steam to obtain an acceptable purity of the desired products G and H. The separation of these two products is carried out in a downstream separation unit not shown in Figure 6. Noncondensable inert species B enters through stream 4, and thus a purge stream is introduced that prevents the buildup of this species as well as the by-product F. This plant has 50 state variables, 12 manipulated (operating) variables, and 41 available output measurements. Six operational constraints are specified for the safe operation of the process. Detailed descriptions of the TE process are given in the original problem formulation presented by Downs and Vogel.[27] However, a mathematical model describing the behavior of this process was not explicitly provided; instead those authors made available a FORTRAN code that simulates this plant under various operating conditions, that is, a black-box model. That original FORTRAN code has been translated into several computing languages. This work makes use of a MATLAB version code provided by Ricker[34] to carry out the present analysis. The optimal steady-state operation of the TE problem has been previously studied by Ricker.[34] In that work, the optimization aimed to identify the nominal values in the system's states ($\mathbf{x}$) that minimizes the plant's operating costs. Uncertainty in the model parameters or the system's states was not considered by Ricker.[34] At steady state, the 12 manipulated variables $\mathbf{u}$ are represented by the last 12 states in $\mathbf{x}$ (i.e., x_{39} through x_{50}).

Scenario 1: Uncertainty in the model parameters

This scenario is similar to that presented by Ricker,[34] that is, determine the optimal steady-state operation of the TE process; however, the present analysis will explicitly account for uncertainty in one of the TE's model parameters. In this case, one of the states, that is, the number of moles of liquid product G inside the reactor, will be assumed to be uncertain due to some model errors or lack of information that prevents the availability of accurate data. Hence, this uncertain parameter is described as follows

$$x_7 \sim N(x_{7,\text{nom}}, {\sigma_7}^2) \tag{16}$$

where $x_{7,\text{nom}}$ and σ_7 represent the state's mean value and standard deviation (i.e., $x_{7,\text{nom}} = 135.363$, $\sigma_7 = 1.8396$),

respectively. The same objective function for the TE process as that presented by Downs and Vogel and used by Ricker will be used in the present analysis, that is

$$\varphi_{\mathrm{TE}} = 0.053m_3 + 0.0318m_{19} + 0.44791m_{10}[2.209m_{29} + 6.177m_{31} + 22.06m_{32} + 14.56m_{33} + 17.89m_{34} + 30.44m_{35} + 22.94m_{36}] + 4.541x_{46}[0.2206m_{37} + 0.1456m_{38} + 0.1789m_{39}] \tag{17}$$

The process constraints for the TE plant are as follows:

$$\begin{aligned} g_1 &= 1 - \frac{m_{12}}{50} = 0 \\ g_2 &= 1 - \frac{m_{15}}{50} = 0 \\ g_3 &= 1 - \frac{2.8153x_{46}m_{40}}{\psi_G} = 0 \\ g_4 &= 1 - \frac{3.4510x_{46}m_{41}}{\psi_H} = 0 \\ g_5 &= \frac{m_7}{P_{\max}} - 1 \le 0 \\ g_6 &= 1 - \frac{m_8}{L_{\min}} \le 0 \end{aligned} \tag{18}$$

where m_i represents the ith measurement in the plant (41 measured outputs total) whereas ψ_{G} and ψ_{H} are the desired production of products G and H, respectively. The first two constraints in (18) are the liquid level constraints in the flash separator and stripper, whereas g_3 and g_4 are production targets for products G and H, respectively. $P_{\max}$ and $L_{\min}$ represent the reactor's maximum allowable pressure and the reactor's minimum liquid level, respectively. Accordingly, g_5 and g_6 represent the reactor's maximum pressure and minimum liquid level constraints, respectively. Following the methodology presented in this work, the above constraints were reformulated in the form shown in (8). Applying the PSE approximation analysis, the distribution (variability) in the process constraints due to uncertainty in x_7 is obtained by substituting sampled uncertainty data into each of the PSE-based expressions developed for each of the process constraints shown in (18). Following (7) and (8), the TE process constraints shown in (18) were reformulated as follows:

$$\xi_{g(k)} \le 0 \qquad \forall k = 1, 2, ..., 6 \tag{19}$$

where $\xi_{g(k)}$ represents the kth process constraint in the TE process shown in (18) and that is evaluated using a PSE-based constraint function at a given probability of satisfaction Pb_h. Based on the above descriptions, the optimal operation problem of the TE process under uncertainty in x_7 can be formulated as follows

$$\begin{aligned} &\min_{\mathbf{x}} \ \varphi_{\mathrm{TE}}(\mathbf{x}, \mathbf{m}) \\ &\quad \text{s.t.} \\ &\quad \text{TE process model} \\ &\quad \xi_{g(k)} \le 0 \qquad \forall k = 1, 2, ..., 6 \\ &\quad x_i \ge 0 \qquad \forall i = 1, 2, ..., 38, i \ne 7 \\ &\quad 0 \le x_i \le 100 \qquad \forall i = 39, 40, ..., 50 \end{aligned} \tag{20}$$

where the TE process model was developed by Ricker[34] as a MATLAB code and not discussed here for brevity. The present analysis assumes that a suitable control scheme can be designed to maintain the feasible operation of the TE process. Problem (20) aims to identity the optimal operation of the 49 states in the TE process (which also include the nominal values in the manipulated variables) in the presence of uncertainty in x_7, which follows the description shown in (16). For comparison purposes, problem (20) was solved using the mean (nominal) value in x_7 (i.e., optimal design using only $x_{7,\mathrm{nom}}$) and using the uncertainty description shown in (16) for x_7. Also, the optimal operation of the TE plant was solved for different confidence levels in the process constraints and the model outputs. The results obtained for the available manipulated variables **u** and the output measurements **m**, which specify the TE's optimal steady-state operating conditions, are shown in Table 7. The present analysis shows that, for the nominal base case (i.e., $x_7 = x_{7,\mathrm{nom}}$), a total cost of 2% less than that reported by Ricker's[34] (114.31 \$/h) was obtained with the present method. As shown in Table 7, conservative (expensive) plant designs were specified by the present method to accommodate the uncertainty considered in x_7. A 13% increase in the total costs was observed for the case of compliance of the constraints under uncertainty ($\mathrm{Pb}_h = 0.9973$) as opposed to the case when x_7 is fixed to its nominal value. As shown in Table 7, the optimal TE process operation requires more purging when x_7 is assumed to be uncertain, which leads to high economic losses due to wasted products in the purge stream. Table 7 also shows that in the purged stream, the concentration of the more expensive components [according to (17)], which are reactant D, products G and H, are higher when the uncertainty in x_7 is considered in the analysis. This also results in higher losses in the purge stream than in the case of fixing x_7 to its nominal (mean) value. As x_7 represents the number of liquid moles of the product G in the reactor, this state has a direct effect on the reactor's pressure and liquid level, which results in a reduction in the condenser coolant flow rate. As shown in Table 7, the computational time needed to solve this problem was about half a minute when using the present PSE approach. However, the stochastic programming approaches using the standard MC sampling method and the efficient Halton-based sampling method required CPU times that are at least 500 and 240 times larger than that required by the present PSE-based method. This result shows the potential computational benefits of the present methodology to address the optimal design of large-scale processes under uncertainty. For the present analysis, a first-order PSE approximation was found to be sufficient to describe the output constraint distributions.

Scenario 2: Uncertainty in multiple parameters

This scenario aims to explore further the effect of using high-order terms in the PSE approximation functions due to the use of multiple uncertain parameters in the analysis. Accordingly, the uncertain states considered for this scenario follow a normal distribution PDF with the following characteristics:

$$\begin{aligned} x_7 &\sim N(135.363, 1.8396^2) \\ x_2 &\sim N(7.522, 0.6133^2) \\ x_9 &\sim N(2.580, 0.62204^2) \end{aligned} \tag{21}$$

where the means are the nominal values for the TE optimal base case[35] and a variance of 2.5, 5, and 15% were assumed as

Table 7. Optimal Operation of the TE Plant, Scenario 1

		Pb = 0.6827				
	$x_7 = x_{7,\text{nom}}$	PSE	MC	Halton	Pb = 0.9545	Pb = 0.9973
Manipulated Variables						
D feed flow (%)	62.781	62.781	62.782	62.782	62.796	62.816
E feed flow (%)	53.216	53.403	53.375	53.368	53.873	53.990
A feed flow (%)	27.594	27.786	27.851	27.844	27.442	26.933
A + C feed flow (%)	60.503	60.528	60.510	60.508	60.647	60.550
Recycle valve (%)	63.722	63.487	63.598	63.572	55.470	46.516
Purge valve (%)	20.978	20.836	20.816	20.825	21.601	22.624
Separator valve (%)	36.812	36.976	36.973	36.970	37.522	38.100
Stripper valve (%)	46.615	46.687	46.675	46.671	46.841	46.841
Steam valve (%)	1.000	1.000	1.000	1.000	1.000	1.000
Reactor coolant (%)	37.641	37.862	37.839	37.832	38.400	38.587
Condenser coolant (%)	100.000	100.000	100.000	100.000	52.640	38.300
Agitator speed (%)	48.623	48.702	48.676	48.673	49.054	49.172
Key measured outputs						
Recycle flow (ksmch)	20.834	20.909	20.896	20.892	21.708	22.387
Reactor pressure (kPa)	2800.000	2788.400	2788.200	2788.100	2757.400	2726.000
Reactor level (%)	65.001	65.241	65.226	65.220	65.846	66.377
Reactor temp. (°C)	125.180	124.770	124.800	124.820	123.930	123.400
Compressor work (kW)	327.860	330.090	330.120	330.050	338.200	340.160
Cond. cool. temperature (°C)	46.839	46.887	46.880	46.878	53.112	58.015
Purge %A (mol%)	39.525	40.011	40.115	40.087	38.359	35.838
Purge %B (mol%)	22.283	22.394	22.397	22.398	22.600	22.854
Purge %C (mol%)	15.327	14.461	14.299	14.315	13.720	12.920
Purge %D (mol%)	0.600	0.646	0.644	0.643	0.787	0.958
Purge %E (mol%)	11.554	12.109	12.096	12.091	14.181	16.767
Purge %F (mol %	4.858	4.636	4.688	4.699	4.396	4.383
Purge %G (mol%)	3.974	3.898	3.912	3.916	4.033	4.249
Purge %H (mol%)	1.879	1.845	1.851	1.853	1.924	2.031
Product %D (mol%)	0.013	0.014	0.014	0.014	0.017	0.020
Product %E (mol%)	0.807	0.875	0.870	0.869	0.999	1.136
Product %F (mol%)	0.329	0.325	0.327	0.327	0.300	0.288
Product %G (mol%)	53.629	53.547	53.560	53.564	53.370	53.371
Product %H (mol%)	43.749	43.767	43.756	43.754	43.844	43.719
Costs breakdown						
Purge losses ($/h)	56.824	56.347	56.362	56.388	58.378	61.221
Product losses ($/h)	37.971	39.980	39.911	39.864	43.146	47.044
Compressor ($/h)	17.575	17.693	17.694	17.691	18.128	18.233
Steam ($/h)	0.209	0.212	0.212	0.212	0.210	0.207
Total cost ($/h)	112.580	114.230	114.180	114.150	119.86	126.710
CPU time (s)	5.706	32.108	16,295.361	7721.978	36.651	36.875

parametric uncertainty for states 7, 2, and 9, respectively. The rest of the specifications are the same as in Scenario 1. This scenario was solved using different orders q in the PSE approximation at a constant user-defined probability for all the constraints ($Pb_h = 0.6$). As shown in Table 8, the computational time required for Scenario 2 is larger than for Scenario 1 even for the case of a first-order PSE approximation (Scenario 2, $q = 1$). This increase in the computational costs is mostly due to the simultaneous consideration of multiple uncertainties occurring in the plant. Table 8 also shows that the first-order PSE approximation does not provide accurate results when compared to the designs achieved with high-order PSE approximations, that is, using a first-order PSE approximation in the calculations resulted in plant costs that are 11% lower than that obtained from a second-order PSE approximation. Thus, high-order PSE approximation functions needed to be considered in the calculations to accurately describe the variability in the constraints due to the uncertainty in states x_2, x_7, and x_9. As shown in Table 8, the operating costs obtained for the second-order ($q = 2$) and the third-order PSE approximation ($q = 3$) have a relative error less than 1%. Hence, the current scenario will adopt a second-order PSE approximation as it identifies the optimal operating conditions of the TE plant in a CPU time that is about 5 times faster than using third-order PSE approximation functions for the process constraints and model outputs considered in the analysis.

As shown in Table 8 (Scenario 2, $q = 2$), a total operating cost of 140.71 $/h is needed to satisfy the TE process constraints at a probability limit of $Pb_h = 0.6$. This operating cost is more than 20% higher than that obtained for Scenario 1, that is, only one uncertain state (x_7) at $Pb_h = 0.687$, and about 25% higher than that obtained when x_7 is fixed to its nominal mean value (see Table 7: $x_7 = x_{7,\text{nom}}$). As shown in Table 8, the purge losses still dominates the total TE plant costs. Note that the concentration of the products lost in the purge stream is higher in Scenario 2 than in Scenario 1. Similarly, the estimates manipulated variables shown in Table 8 indicate that a larger purge valve opening (more purging) is needed when compared to the design obtained under perfect knowledge of all the system's states, that is, $x_7 = x_{7,\text{nom}}$ in Table 7. This result indicates that the specification of the optimal operation of the TE process under the assumption of system's states or model parameters that are assumed to be perfectly known results in inoperable plants under uncertainty. Conversely, the present method identified (at minimum computational cost) an optimal operating condition for this process that remains feasible (at given probability of satisfaction, Pb_h) in the presence of uncertainty in multiple system's states.

Table 8. Optimal Operation of the TE Plant, Scenarios 2 and 3

	Scenario 2 ($q = 1$)	*Scenario 2 ($q = 2$)*	*Scenario 2 ($q = 3$)*	*Scenario 3 ($q = 2$)*
Manipulated Variables				
D feed flow (%)	62.89	65.88	65.88	62.89
E feed flow (%)	53.72	54.34	54.35	56.72
A feed flow (%)	26.71	27.26	27.25	26.45
A + C feed flow (%)	60.50	62.03	62.03	61.62
Recycle valve (%)	45.13	43.90	43.82	1.00
Purge valve (%)	23.05	24.01	24.01	26.39
Separator valve (%)	38.03	40.07	40.10	41.33
Stripper valve (%)	46.75	48.16	48.17	47.78
Steam valve (%)	1.00	1.00	1.00	41.30
Reactor coolant (%)	38.33	38.90	38.91	40.80
Condenser coolant (%)	37.27	100.00	100.00	18.78
Agitator speed (%)	48.93	48.88	48.89	50.26
Key measured outputs				
Recycle flow (ksmch)	22.24	20.84	20.85	31.87
Reactor pressure (kPa)	2720.00	2678.70	2678.40	2800.00
Reactor level (%)	66.11	65.83	65.86	70.00
Reactor temperature (°C)	123.97	125.06	125.03	118.48
Compressor work (kW)	336.94	323.55	323.56	269.33
Cond. cool. temperature (°C)	58.30	47.20	47.20	81.65
Purge %A (mol %)	34.65	31.66	31.64	32.39
Purge %B (mol %)	22.90	23.85	23.85	21.31
Purge %C (mol %)	13.40	12.47	12.44	11.02
Purge %D (mol %)	0.92	1.05	1.06	1.80
Purge %E (mol %)	16.68	19.16	19.23	24.62
Purge %F (mol %)	4.86	5.47	5.44	2.61
Purge %G (mol %)	4.47	4.36	4.35	4.17
Purge %H (mol %)	2.13	1.98	1.98	2.08
Product %D (mol %)	0.02	0.02	0.02	0.03
Product %E (mol %)	1.06	1.35	1.36	1.45
Product %F (mol %)	0.30	0.38	0.37	0.15
Product %G (mol %)	53.53	54.44	54.43	52.33
Product %H (mol %)	43.62	42.35	42.34	44.57
Costs breakdown				
Purge losses ($/h)	62.44	64.25	64.24	74.74
Product losses ($/h)	44.85	58.91	59.23	53.33
Compressor ($/h)	18.06	17.34	17.34	14.44
Steam ($/h)	0.20	0.21	0.21	7.03
Total Cost ($/h)	125.54	140.71	141.02	149.54
CPU Time (s)	201.79	358.71	1,868.93	274.19

For this case study, it was not possible to compare the results presented in Table 8 with that of a stochastic programming approach because the solution of the optimal operation of the TE plant will require intensive calculations. That is, assessing the variability in the process constraints and model outputs at each optimization step using extensive simulations of the full TE plant model for a large set of realizations in the uncertain parameters may result in prohibitive computational times. To compare the computational costs while using the actual TE process model and different orders in the PSE approximation functions, the CPU time needed to obtain the distribution (frequency histogram) in the reactor's maximum pressure constraint was assessed. To perform this analysis, the specifications obtained for the second-order PSE approximation shown in Table 8 (Scenario 2, $q = 2$) were used to generate the frequency histogram for this constraint via the MC sampling method applied to the full plant model. The same frequency histogram was generated using a first-order, second-order, and a third-order PSE approximation function for that constraint. 10,000 realizations that follow the description in (21) for the uncertain parameters were used in this analysis. As shown in Figure 7, second-order and third-order PSE approximations accurately capture the distribution in the reactor's pressure due to the realizations considered for the uncertain parameters. The CPU time needed to generate this distribution using the full plant model and the different PSE approximations are shown in

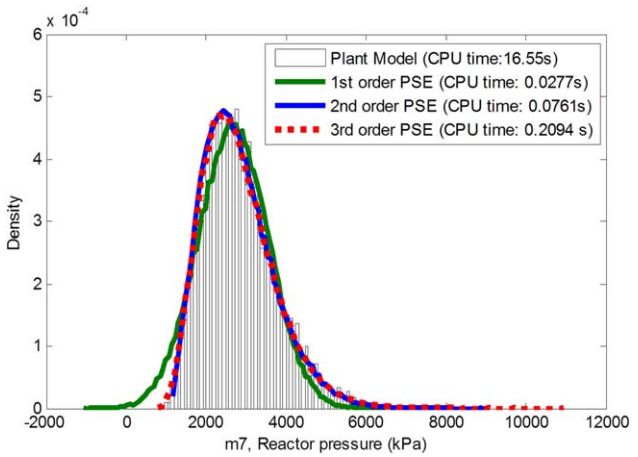

Figure 7. Frequency histogram for the reactor's pressure obtained via the MC method applied to the full TE plant model and the PSE-based model using different approximation orders.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

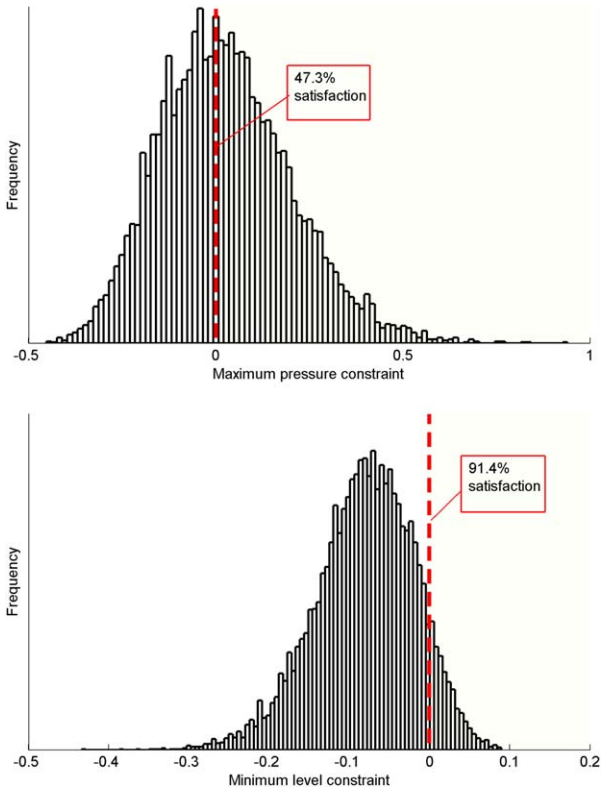

Figure 8. Frequency histogram for the reactor's maximum pressure constraint (a) and the reactor's minimum level constraint (b) obtained for Scenario 3 via the MC sampling method applied to the full TE process model.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

the legends of Figure 7. These results indicate that using a third-order PSE approximation ($q = 3$) returns accurate approximations in a CPU time that is about 80 times faster than using the complete TE plant model.

Scenario 3: Ranking-based designs

The present scenario aims to demonstrate the ranking-based feature of the proposed approach to address the optimal operation of the TE process. To perform this analysis, the results obtained from Scenario 2 will be compared to the case were the probability of satisfaction for the reactor's level and pressure constraints are set to 0.9 and 0.5, respectively, that is, $Pb_{g6}=0.9$ and $Pb_{g5}=0.5$. The probability of satisfaction for the rest of the constraints considered for this plant, that is, Pb_{g1}–Pb_{g4}, was set to 99.73%. The uncertainty in the parameters were assumed to be same as in Scenario 2 and shown in (21), whereas the order of the PSE approximation was set to $q = 2$. Table 8 shows the results obtained for the present scenario (Scenario 3). Figure 8 shows the validation of the results obtained for the present scenario by evaluating the distribution (variability) in the reactor's level and pressure constraints using the actual TE plant model. As shown in this figure, good approximations to the probability limits considered for the reactor's level and pressure are obtained while using the present ranking-based approach with as second-order PSE approximation. As shown in Table 8, the total costs specified for Scenario 3 are about 6% higher than those obtained Scenario 2 where the constraints were set to an equal probability of satisfaction of 60% ($Pb_h = 0.6$). To ensure the high probability of satisfaction assigned to the reactor level constraint in the present scenario ($Pb_{g6}=0.9$), the optimal operation requires the nominal value of the reactor level to be set at 70%, which is 4% higher than that specified for Scenario 2 ($q = 2$). This increase in the reactor's nominal liquid level also increased the reactor coolant's flow rate by almost 2% with respect to Scenario 2 ($q = 2$). Similarly, the reactor pressure constraint is not active at the solution for the Scenario 3 due to the low probability of satisfaction assigned to this constraint ($Pb_{g5}=0.5$). Hence, more violations were assumed to be allowed on this constraint for this Scenario. Moreover, the present scenario specified a nominal temperature in the reactor that is 7°C lower than that obtained for Scenario 2. This decrease in temperature, combined with a high reactor working pressure, resulted in a decrease in the coolant flow rate in the condenser unit. Note that the computational time required to achieve a solution for the present Scenario is comparable to that required by Scenario 2 ($q = 2$). These results show that the ranking-based approach proposed in this work is a computationally attractive practical tool that can be used to study different design alternatives that involve tradeoffs between profitability and robust (expensive) plant designs under uncertainty.

Conclusions

In this article, a new method that addresses the optimal process design under uncertainty was presented. A ranking-based approach is considered in the present study where the user can assign different probability limits to each of the safety, environmental, and operational constraints considered in the analysis. Thus, critical constraints are enforced to be satisfied all the times by setting a high probability limit, whereas less sensitive constraints to the process safety and economics may be allowed to be violated at an accepted level of confidence. This feature gives the flexibility to select between conservative (expensive) designs and economically attractive designs that allow constraint violations. An analytical expression for the process constraints in terms of the uncertain variables is obtained using a PSE approximation. The PSE expressions are then used to compute (at minimum computational cost) the distribution (variability) in the process constraints and model outputs due to multiple realizations in the uncertain parameters, which are assessed using the MC sampling method. The proposed approach is computationally attractive because it avoids the need to simulate the complete plant model for each realization in the uncertain parameters as it is the case in stochastic programming-based approaches. The key computational effort in the present method relies in the identification of the sensitivity terms for each of the PSE approximations that need to be developed for the process constraints and model outputs considered in the analysis. However, the two case studies examined in this work indicate that the computational times needed by the present method to address the optimal design of a large-scale system is orders of magnitude shorter than those required by the traditional stochastic programming-based methods, which rely on extensive simulations of the complete process model.

Acknowledgment

The authors would like to acknowledge the financial support provided by the United Mining Investments Co.

(UMIC) to carry out the research work presented in this article.

Notation

$\mathbf{z}$ = steady-state process model
$\mathbf{d}$ = design variables
$\mathbf{d}^l, \mathbf{d}^u$ = lower and upper bounds for design variables
$\mathbf{p}$ = model parameters
$\mathbf{x}$ = state variables
$\mathbf{y}$ = model outputs
$\mathbf{u}$ = model inputs
$\mathbf{u}^l, \mathbf{u}^u$ = lower and upper bounds for model inputs
$\boldsymbol{\theta}$ = system's uncertain variables
$\tilde{\mathbf{p}}$ = uncertain model parameters
$\tilde{\mathbf{u}}$ = uncertain model inputs
$\boldsymbol{\alpha}_l$ = probability distribution parameters
$\mathbf{h}$ = process constraints
$\mathbf{M}^{(i)}$ = ith order sensitivity term
N = number of sampled realizations in system's uncertainty
L = number of uncertain parameters
Pb_h = probability of constraint satisfaction
ξ_h = extreme possible value at user-defined Pb_h
F = cumulative distribution function
$\boldsymbol{\eta}$ = set of decision variables

Literature Cited

1. Biegler LT, Grossmann IE, Westerberg AW. Systematic Methods for Chemical Process Design. Prentice-Hall, New Jersey, 1997.
2. Seider WD, Seader JD, Lewin DR. Product & process design principles: synthesis, analysis and evaluation. John Wiley & Sons, New York, 2009.
3. Halemane KP, Grossmann IE. Optimal process design under uncertainty. *AIChE J*. 1983;29(3):425–433.
4. Grossmann IE, Sargent RWH. Optimum design of chemical plants with uncertain parameters. *AIChE J*. 1978;24(6):1021–1028.
5. Ierapetritou MG, Acevedo J, Pistikopoulos EN. An optimization approach for process engineering problems under uncertainty. *Comput Chem Eng*. 1996;20(6–7):703–709.
6. Birge JR, Louveaux FV. Introduction to Stochastic Programming. Springer, New York, 1997.
7. Liu ML, Sahinidis NV. Optimization in process planning under uncertainty. *Ind Eng Chem Res*. 1996;35(11):4154–4165.
8. Acevedo J, Pistikopoulos EN. Stochastic optimization based algorithms for process synthesis under uncertainty. *Comput Chem Eng*. 1998;22(4):647–671.
9. Ahmed S, Tawarmalani M, Sahinidis NV. A finite branch-and-bound algorithm for two-stage stochastic integer programs. *Math Program*. 2004;100(2):355–377.
10. Takriti S, Birge JR, Long E. A stochastic model for the unit commitment problem. *IEEE Trans Power Syst*. 1996;11(3):1497–1508.
11. Norkin VI, Ch G, Ruszczyliski A. A branch and bound method for stochastic global optimization. 1998;83:425–450.
12. Karuppiah R, Grossmann IE. Global optimization of multiscenario mixed integer nonlinear programming models arising in the synthesis of integrated water networks under uncertainty. *Comput Chem Eng*. 2008;32(1):145–160.
13. Paules G, Floudas C. Stochastic programming in process synthesis: a two-stage model with MINLP recourse for multiperiod heat-integrated distillation sequences. *Comput Chem Eng*. 1992;16(3):189–210.
14. Rooney WC, Biegler LT. Incorporating joint confidence regions into design under uncertainty. *Comput Chem Eng*. 1999;23(10):1563–1575.
15. Rooney WC, Biegler LT. Design for model parameter uncertainty using nonlinear confidence regions. *AIChE J*. 2001;47(8):1794–1804.
16. Rooney WC, Biegler LT. Optimal process design with model parameter uncertainty and process variability. *AIChE J*. 2003;49(2):438–449.
17. Raspanti C, Bandoni J, Biegler L. New strategies for flexibility analysis and design under uncertainty. *Comput Chem Eng*. 2000;24(9):2193–2209.
18. Zhu Y, Legg S, Laird CD. Optimal design of cryogenic air separation columns under uncertainty. *Comput Chem Eng*. 2010;34(9):1377–1384.
19. Charnes A, Cooper WW. Chance-constrained programming. *Manage Sci*. 1959;6(1):73–79.
20. Wendt M, Li P, Wozny G. Nonlinear chance-constrained process optimization under uncertainty. *Ind Eng Chem Res*. 2002;41(15):3621–3629.
21. Li P, Arellano-Garcia H, Wozny G. Chance constrained programming approach to process optimization under uncertainty. *Comput Chem Eng*. 2008;32(1–2):25–45.
22. Arellano-Garcia H, Wozny G. Chance constrained optimization of process systems under uncertainty: I. *strict monotonicity*. *Comput Chem Eng*. 2009;33(10):1568–1583.
23. Ostrovsky GM, Ziyatdinov NN, Lapteva TV, Zaitsev IV. Optimization of chemical processes with dependent uncertain parameters. *Chem Eng Sci*. 2012;83:119–127.
24. Ostrovsky GM. Optimization problem with normally distributed uncertain parameters. *AIChE J*. 2013;59(7):2471–2484.
25. Ostrovsky GM, Ziyatdinov NN, Lapteva TV. Optimal design of chemical processes with chance constraints. *Comput Chem Eng*. 2013;59:74–88.
26. Pintarič ZN, Kasaš M, Kravanja Z. Sensitivity analyses for scenario reduction in flexible flow sheet design with a large number of uncertain parameters. *AIChE J*. 2013;59(8):2862–2871.
27. Downs JJ, Vogel EF. A plant-wide industrial process control problem. *Comput Chem Eng*. 1993;17(3):245–255.
28. Maly T, Petzold L. Numerical methods and software for sensitivity analysis of differential-algebraic systems. *Appl Numer Math*. 1996;20(1):57–79.
29. Nocedal J, Wright SJ. Numerical Optimization, 2nd ed. Springer, New York, 2006.
30. Grossmann IE, Halemane KP. Decomposition strategy for designing flexible chemical plants. *AIChE J*. 1982;28(4):686–694.
31. Halton JH. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numer Math*. 1986;2(1):84–90.
32. Hammersley JM. Monte Carlo methods for solving multivariable problems. *Ann N Y Acad Sci*. 1960;86(3):844–874.
33. Diwekar UM, Kalagnanam JR. Efficient sampling technique for optimization under uncertainty. *AIChE J*. 1997;43(2):440–447.
34. Ricker N. Optimal steady-state operation of the tennessee eastman challenge process. *Comput Chem Eng*. 1995;19(9):949–959.

Manuscript received Feb. 5, 2014, and revision received Apr. 24, 2014.