Journal of the Taiwan Institute of Chemical Engineers 45 (2014) 1369–1388

Contents lists available at ScienceDirect

Journal of the Taiwan Institute of Chemical Engineers

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

Expected improvement in efficient experimental design supported by a global optimizer

Jyh-Shyong Chang *, Wei-Ling Liu, Jiun-Kai Tang, Wen-Chin Lin, Yuh-Jing Chiou

Department of Chemical Engineering, Tatung University, 40 Chungshan North Road, 3rd Section, Taipei 10452, Taiwan

ARTICLE INFO

Article history:
Received 16 September 2013
Received in revised form 23 January 2014
Accepted 5 February 2014
Available online 20 March 2014

Keywords:
Experimental design
Expected improvement
Global optimizer
Surrogate model

ABSTRACT

To develop an efficient experimental design method for real process applications, the well-known expected improvement infill criterion, which is usually adopted in achieving the surrogate-based optimization, and the global optimizer DIRECT are combined as the core of the developed experimental method. The method iterates through initial experimental design, empirical modeling and model-based optimization to allocate informative experiments for the next iteration. Specifically, the Kriging regression is adopted as the surrogate model due to its demonstrated prediction accuracy and reliable prediction uncertainty. Adopting a suitable threshold value of the initial expected improvement during the optimization process, the experiments located by the global optimizer could accelerate the optimization process to reach the defined target. Three termination criterions for stopping the iterating process are proposed to meet the requirements of both the simulation optimization problems and the experimental systems. Three simulation test problems demonstrate the efficiency of the developed experimental design method.

1. Introduction

The design of new products, or the improvement of existing ones, as well as the design and development of manufacturing processes to produce them, are crucial activities in most industrial organizations. It is beneficial to develop a new product and process design as well as to improve an existing process using statistically designed experiments when the input–output relationship of the process is unknown or partially known [1]. Most of the early development of experimental design was stimulated by applications in agriculture. Since its beginnings in agriculture, experimental design has applications across many sectors of industry. The statistical principles underlying design of experiments were largely developed by R.A. Fisher during his pioneering work at the Rothamsted Experimental Station in the 1920s and 1930s [2]. The use of experimental design methods in the chemical industry was promoted in the 1950s by the extensive work of Box and his collaborators on response surface designs [3]. Experimental design techniques are also becoming popular in the area of computer-aided design and engineering using computer/simulation models, including applications in manufacturing (the automobile and semiconductor industries), as well as in the nuclear industry [4]. Statistical issues in the design and analysis of computer/simulation experiments are discussed in [5].

Generally, experiments are used in studying or evaluating the performance of physically unknown or partially known systems. The objectives of the experiment may include the following [1]:

(a) to determine which variables are most influential in the response of the system;
(b) to determine where to set the influential variables so that the response or responses of the system are almost always near desired target values;
(c) to determine where to set the influential variables so that the variability in the response(s) is small; or
(d) to determine where to set the influential variables so that the effects of uncontrolled variable on the response(s) are small.

To achieve the objectives (a) and (b), which are the main concerns of this work, the following experimental design scheme is usually adopted [6]:

(a) first the variables to be optimized are chosen, often due to their importance, as determined by preliminary experiments;

* Corresponding author. Tel.: +886 2 1822928 6266; fax: +886 2 5861939.
E-mail address: jschang@ttu.edu.tw (J.-S. Chang).

http://dx.doi.org/10.1016/j.jtice.2014.02.005
1876-1070/

(b) some initial sample designs are analyzed according to some pre-defined plan;
(c) a surrogate model type is selected and used to build a model of the underlying problem;
(d) a search is carried out using the model to identify new design points for analysis;
(e) the new results are added to those already available and, provided further analyses are desired, the process returns to step (c).

The response surface methodology (RSM) addressing the above experimental design scheme has been applied successfully in various fields such as the chemical industry, biology, medicine, and economy [3,7–9]. The success or failure for applying the experimental design to a process system short of mechanistic information hinges on a correct choice of the surrogate model (metamodel) and efficient infill criteria. However, in traditional RSM, the first- or second-order polynomial response surface (PRS) approximation [7] is adopted for empirical modeling or as the surrogate model. The restrictive functional form of polynomials has long been recognized as ineffective in modeling complex processes [9,10]. The alternatives that could be adopted as the surrogate models or metamodels for the studied system are artificial neural networks (ANN) [10–14], multivariate adaptive regression splines [15], support vector regression (SVR) [16–19], radial basis functions (RBF) [20–22], Kriging (KR) [5,23,24] and Gaussian process regression (GPR) [25–27]. Fang et al. [28] found RBF gives accurate metamodels for highly nonlinear responses whereas. Simpson et al. [29] found Kriging to be most suitable for slightly nonlinear responses in high-dimension spaces. Jin et al. [30] proposed the use of PRS for slightly nonlinear and noisy responses, while Clarke et al. [17] found SVR metamodels to be the best in their study [31]. Queipo et al. [32] provide a good review of different metamodeling techniques for surrogate-based analysis and optimization.

The success of applying any global optimization method to experimental design based on a nonparametric surrogate model hinges on the effective provision of a balance between local and global searches [33–38]. Because the identified surrogate model is only an approximation of the true process we wish to optimize, it is tempting to enhance the accuracy of the model using further experimental responses (infill or update points), in addition to the initial sampling plan. On one hand, one may wish to improve the accuracy solely in the region of the optimum predicted by the surrogate to obtain an accurate optimal value quickly and make local exploitation; on the other hand one may, however, be unsure of the global accuracy of the surrogate and employ an infill strategy which enhances the general accuracy of the model and make global exploration [6]. Forrester et al. [6] reviewed several infill criteria including (a) prediction based exploitation; (b) error based exploration; (c) balanced exploitation and exploration; (d) conditional likelihood approaches and (d) other methods (parallel infill points, hybrid criteria) using the identified surrogate model that can provide the expected (mean) value and the variance of the process response. The most famous expected improvement infill criterion and its extensions [34,39] were frequently adopted in the computer/simulation experiments. Among the available surrogate models, Forrester et al. [6] suggested that the Gaussian bases (*e.g.*, KR and GPR) perform properly for both simple and complex landscapes for an unknown multi inputs-single output (MISO) system $y = f(\mathbf{x})$ with its dimension k of the design variable vector $\mathbf{x}$ less than 20, which is usually the cases for real processes. Furthermore, both these two surrogate models are quite efficient for finding the infill points in achieving the global optimum of the unknown process. An interactive approach for determining the infill points was proposed by Chia et al. [9] when used for investigating real processes. Their suggestion involves active and subjective decision of the experimenter and this human intervention brings in domain knowledge that is often difficult to be properly incorporated in the modeling framework.

Choosing a suitable convergence criterion to determine when to stop the surrogate infill process is rather subjective [6]. In the face of limited experimental resources for a real process, this is an important issue for a process engineer or researcher. In this work, an additional global optimizer such as DIRECT optimization algorithm [40] will be augmented to improve the efficiency of the expected improvement algorithm [34] in finding the infill points and to determine the convergence of the infill process in achieving a global optimum when dealing with a real process. The capability of the proposed experimental design via the data-driven global optimization methodology will be examined by three test examples. The first example is the modified Himmelblau function of two independent variables studied by Chang and Lin [10] and Chen et al. [11]. Fundamentally, this example is a static problem. The second example is concerned with a dynamic problem, which involves determining an optimal temperature trajectory of a batch reactor where a series reaction is carried out [10]. The last example is the simulated process based on the experimental data from the batch heterogeneous catalytic etherification reaction system [41]. For these three case studies, the proposed experimental design method in this work was applied to build a developing surrogate model and to locate the optimum of an unknown process simultaneously based on fewer experiments [10].

In the following section, the principle and implementation of the developed experimental design that adopts the expected improvement infill criterion supported by a global optimizer DIRECT are described. In Section 3, the capability of the developed experimental design method is presented through the test examples. The final section draws the conclusions for this study.

1.1. Experimental design via the data-driven global optimization methodology

Fig. 1 illustrates the proposed data-driven global optimization framework for the efficient experimental design. The proposed scheme is similar to the conventional experimental design scheme shown in the introduction section of Forrester and Keane [6]. Specially shown in Fig. 1 is a specific Gaussian process surrogate model (KR or GPR) adopted. Based on the identified surrogate model, the expected improvement algorithm supported by a global optimizer is used to allocate new experiments. Finally, when to stop the optimization process can be judged by visualizing whether the process response is approaching what is estimated by the global optimizer (DIRECT). The various aspects of the framework will be discussed in more detail below.

1.2. The objective function

The goal of the data-driven global optimization process is to minimize the objective (or loss) function, $f(\mathbf{x})$, within the feasible region,

$$\min_{x \subset \chi} f(\boldsymbol{x}) \tag{1}$$

where $f(\mathbf{x})$ represents the expected performance of a system and $\mathbf{x}$ is a k dimensional design variable vector to be adjusted. We consider the system as a black box that provides no information other than the measurements of system performance. We assume that the feasible region $\chi \subset R^k$ is continuous, connected, and compact. The measurement y of the objective function contains