

Chapter 8

The Role of Computational Intelligence in Experimental Design: A Literature Review

Erkan Işıklı and Seda Yanık

Abstract Experimental design (DOE) is a well-developed methodology that has been frequently adopted for different purposes in a wide range of fields such as control theory, optimization, and intelligent decision making. The main objective of DOE is to best select experiments to estimate a set of parameters while consuming as little resources as possible. The enrichment of literature on computational intelligence has supported DOE to extend its sphere of influence in the past two decades. Specifically, the most significant progress has been observed in the area of optimal experimentation, which deals with the calculation of the best scheme of measurements so that the information provided by the data collected is maximized. Nevertheless, determining the design that captures the true relationship between the response and control variables is the most fundamental objective. When deciding whether a design is better (or worse) than another one, usually a criterion is utilized to make an objective distinction. There is a wide range of optimality criteria available in the literature that has been proposed to solve theoretical or practical problems stemming from the challenging nature of optimal experimentation. This study focuses on the most recent applications of DOE related to heuristic optimization, fuzzy approach, and artificial intelligence with a special emphasis on the optimal experimental design and optimality criteria.

Keywords Optimal experimental design · Optimality criteria · Neural networks · Heuristic optimization · Fuzzy

E. Işıklı (✉) · S. Yanık
Department of Industrial Engineering, Istanbul Technical University,
34367 Macka, Istanbul, Turkey
e-mail: isiklie@itu.edu.tr

8.1 Introduction

Experimental design (DOE) aims to quantify the cause and effect relationship between the inputs (process variables) and outputs (responses) of a process as economically as possible. The process of interest could belong in any field where variance reduction or quality improvement is one of the main objectives. Since its introduction by R.A. Fisher in the 1930s, DOE has attracted much attention and has been applied in various areas ranging from manufacturing to biochemistry; service industry to quality control; and biomedical sciences to marketing. As the number of cases in which DOE approach has been adopted increased, interesting challenges, mostly related with the main assumptions of DOE or applicability and efficiency of traditional designs, have arisen. In order to deal with these challenges, newer designs such as Box–Wilson Central Composite Design, Doehlert Design, Box–Behnken Design, Plackett–Burman Design, Split Plot Design, and Rechtschaffner Design have been introduced. However, Box Behnken and Central Composite designs may not have performed well in case the process behaviour is more complex than a second order (Rollins and Bhandari 2004). Thus, the significant acceleration in the rise of new experimentation techniques was observed with the introduction of computer-aided designs in which one or more optimality criteria are used to construct optimal experimental designs. Lately, the ease of computation has propelled the use of computational intelligence based methods in optimal experimental design. In this study, four streams of the related literature are reviewed: optimization methods, heuristics, fuzzy techniques, and artificial intelligence. However, we should note that these streams are not clear-cut and it is highly likely to come across studies combining methods from different research streams when reviewing the DOE literature. The remainder of this chapter is organized as follows: Sect. 8.2 overviews the basic terminology of DOE and provides some insights on the extent of the use of DOE by reviewing its most recent applications, Sects. 8.3 and 8.4 focus on heuristic optimization methods, and artificial intelligence and fuzzy methods employed in DOE, respectively, Sect. 8.5 concludes with a discussion of potential research avenues.

8.2 The Fundamentals of Experimental Design

Experimental design (DOE) aims to reduce the experimental cost while observing how a response variable (output) is influenced by alternating one or more process variables (inputs). Traditional approaches in DOE include full factorial designs, fractional factorial designs, mixture designs, Taguchi designs, central composite

designs (Box–Wilson, Box–Behnken, etc.), and Latin hypercube designs. Lundstedt et al. (1998) provided an insightful review on DOE with a special emphasis on the screening methods along with central composite designs and the Doehlert design. Anderson-Cook et al. (2009) paid particular attention to robust parameter designs, split-plot designs, mixture experiment designs, and designs for generalized linear models. The authors underlined the importance of investing more in the analysis stage, before data collection, to obtain better results. Resolving issues regarding a design could drastically increase the related cost. Thus, attention should be paid in the earlier stages of the experimentation. Recently, a remarkable progress in optimal experimentation has been observed, especially due to new algorithmic approaches and a significant decrease in computation times; however, this stream of the literature is still developing and needs more attention even though its roots date back to 1920s.

Optimal experimental designs (sometimes also called computer-aided designs) are generated by an optimization algorithm that uses a design criterion to measure the quality of the experiment. There are several optimality criteria proposed in the literature which can be mainly classified into two groups: information-based criteria and distance-based criteria. The former are based on the Fisher information matrix, $X^T X$, whereas the latter are based on the distance $d(y, A)$ from a point (y) in the n -dimensional Euclidean space (R^n) to a subset (A) of R^n . These criteria play a vital role in optimal experimentation as they help experimenters choose between alternative designs—by calculating their efficiencies—without wasting too much resource, time, effort, and money. However, experimenters should also take into account the robustness of these candidate designs and the effect of missing data to make better conclusions (Anderson-Cook et al. 2009).

To provide the reader a background to better understand the details of various designs to be discussed in subsequent sections, we briefly cover the information-based criteria (also known as the alphabet criteria) below. Interested readers should refer to Das (2002) or Pukelsheim (1993) for a thorough review on this topic.

- A-optimality minimizes the trace of $(X^T X)^{-1}$, which is equivalent to minimizing the average variance of the parameter estimates. It is vulnerable to changes in the coding of the design variable(s) (Anderson-Cook et al. 2009).
- C-optimality minimizes the variance of the best linear unbiased estimator of a predetermined combination of model parameters (Harman and Jurik, 2008).
- D-optimality maximizes $\det(X^T X)$, which is equivalent to minimizing the inverse Fisher information matrix. This way, the volume of the confidence ellipsoid around the parameter vector is minimized. The higher the D-optimality criterion the smaller the confidence region for the parameter estimates (Balsa-Canto et al. 2007).

- E-optimality maximizes the minimum eigenvalue of $X^T X$, which implies the minimization of the maximum variance of all possible normalized linear combinations of parameter estimates. Modified E-optimality, which minimizes the ratio of the largest eigenvalue of $X^T X$ to the smallest one, represents the relationship between the longest and shortest semi-axes of the information hyper-ellipsoid (Balsa–Canto et al. 2007).
- T-optimality maximizes the trace of $X^T X$.
- G-optimality minimizes the maximum entry in the diagonal of $X(X^T X)^{-1} X^T$, which corresponds to the maximum variance of any predicted value over the design space.
- I-optimality (also known as Q-optimality, V-optimality, or I_V -optimality) minimizes the (normalized) average prediction variance over the region of interest.
- L-optimality, a modified version of A-optimality, minimizes the average variance of the parameter estimates (Wit et al. 2005).
- V-optimality minimizes the average prediction variance over a set of m specific points.

Fraleigh et al. (2003) mentioned that there were two particular optimal designs of interest: variance optimal and model discrimination designs. There are a range of variance optimal designs including A-, D-, E-, G- and Q-optimal approaches. D-optimal experimental design, which was developed to determine the experimental conditions that minimize the volume of the uncertainty region for the parameter estimates, has been very popular. T-optimal experimental designs are used to decide which experimental conditions to use so that one can discriminate between alternative models, and is based on the prediction error. The objective of T-optimal design is to maximize the sum of squares lack of fit between the observations and the model predictions. López-Fidalgo et al. (2007) proposed an extension to the conventional T-optimality criterion that considers the case of non-normal parametric regression models. Their criterion was further modified by Otsu (2008) to also cover the case of semi-parametric models as an assumption on the distribution of residuals may be restrictive in some cases. Fang et al. (2008) explored five different approaches to derive the lower bounds of the most common criteria employed in DOE.

As Anderson-Cook et al. (2009) informed, optimality criteria should not be the only aspects to consider for the estimation and/or prediction. Collecting data reasonably, estimating/interpreting model parameters carefully, and having Plan B are equally important aspects in DOE. Thus, creating a design that balances the pros and cons of each such aspect should be the first priority of an experimenter, which would result in a near optimal design in many occasions. Imhof et al. (2004) also discussed the pitfalls of an optimal experimental design methodology when some of the observations may not be available at the end of the experiment and showed how inefficient the experimentation could be if the anticipated missingness pattern was not accounted for at the design stage.

DOE is an efficient procedure for planning experiments such that the data obtained can be analyzed to yield valid and objective conclusions. Well-chosen experimental designs maximize the amount of information that can be obtained with a given amount of the experimental effort. The main goal of DOE is to plan a process in an optimal way with a single or multiple underlying objectives such as cost minimization, effective resource consumption, and reduced environment pollution. Therefore, it is natural that DOE can be viewed as an optimization technique (Siomina and Ahlinder 2008).

DOE is often used to select the significant factors that affect the output. Fraleigh et al. (2003) adopted DOE for this purpose in a sensor subsystem to ensure an effective real time optimization (RTO) system. The authors suggested a procedure that combines a modified D-optimal and a modified T-optimal design that fits the RTO problem geometry well and illustrated its use via a simulation study.

Rollins and Bhandari (2004) adopted DOE to determine the design points (to generate data) for sequential step tests in a new multiple input, multiple output (MIMO) constrained discrete-time modelling (DTM) approach for dynamic block-oriented processes. Their approach is essentially innovative as DOE provides the efficient information to estimate ultimate response and dynamic response behaviour. Similarly, Patana and Bogacka (2007) attempted to use DOE to properly design the data collection process and to avoid the noise in the parameter estimates for multi-response dynamic systems when one of its basic assumptions is violated: uncorrelated error terms.

Siomina and Ahlinder (2008) stressed one of the most important reasons to use DOE in practical applications: reducing the cost of experimental time and effort. The authors presented a lean optimization algorithm that sequentially uses super-saturated experimental designs for the optimization of a multi-parameter system in which the maximum number of experiments cannot exceed the number of factors. Their algorithm was proven to be computationally efficient and to significantly outperform the well-known Efficient Global Optimization (EGO) algorithm (Jones et al. 1998). EGO algorithm first fits a response surface to data collected by evaluating the objective function at a few points and then balances between finding the optimum point of the surface and improving the approximation by sampling where the prediction error may be high (Siomina and Ahlinder 2008).

Myers et al. (2004) observed that the response surface framework had become the standard approach for much of the experimentation carried out in industrial research, development, manufacturing, and technology commercialization. The Response Surface Methodology (RSM) has been originally designed to approximate an unknown or complex relationship between design variables and design functions by fitting a simpler model to a (relatively small) number of experimental points. In RSM, the direction of improvement is determined using the path of the steepest descent/ascent (for a minimization/maximization problem) based on the estimated first-order model or using ridge trace analysis for the second-order model (Siomina and Ahlinder 2008). Anderson-Cook et al. (2009) provided an insightful discussion on good response surface designs considering qualitative and quantitative characteristics.

RSM is a widely used technology also for rational experimental design and process optimization in the absence of mechanistic information. RSM initiates from design of experiments (DOE) to determine the factors' values for conducting experiments and collecting data. The data are then used to develop an empirical model that relates the process response to the factors. Subsequently, the model facilitates to search for better process response, which is validated through experiment(s). The above procedure iterates until an optimal process is identified or the limit on experimental resources is reached (Chi et al. 2012).

In traditional RSM, the first-order or second-order polynomial function is adopted for empirical modelling. However, the restrictive functional form of polynomials has long been recognized as ineffective in modelling complex processes. The non-traditional RSM is a stage-wise heuristic that searches for the input combination that maximizes the output (Kleijnen et al. 2004). Progress in adopting more flexible models in RSM includes artificial neural networks (ANN), support vector regression (SVR), and more recently Gaussian process regression (GPR). GPR, also known as kriging model with a slightly different formulation, has been accepted as a powerful modelling tool in various fields, especially in process systems engineering (Chi et al. 2012). The next two sections delve deeper into these topics.

8.3 The Use of Heuristic Optimization Methods in DOE

Lundstedt et al. (1998) compared the theoretical and practical aspects of two optimization approaches (simplex method and response surface methodology) in experimental design (DOE). It is possible to reach the optimal set of parameters using Response Surface Methodology (RSM); however, the experiments are performed one by one in simplex optimization and the global optimum is not guaranteed. Coles et al. (2011) emphasized the need for a comprehensive approach that compares the quality of the optimal experimental design and the computational efficiency of the algorithm used for parameter estimation. They claimed that it would not always be possible to find a unique algorithm that could perform well for different types of objective functions. This is one of the most important challenges in DOE: the trade-off between the optimum use of resources and the computational efficiency. Such challenges have usually been approached by using both linear and non-linear programming techniques. However, traditional algorithms may not work at some instances. This is where heuristic approach comes into play. A detailed discussion on the use of heuristic techniques is made in this section after providing a concise review on how, in general, optimization techniques have been employed in DOE.

On the linear side, Joutard (2007) proposed a large deviations principle for the least-squares estimator in a linear model and used its results to find optimal experimental designs. The author demonstrated the performance of this principle by estimating the whole parameter vector in a Gaussian linear model and one

component of the parameter vector in an arbitrary linear model in which certain assumptions on the distribution of errors were made. Harman and Jurík (2008) formulated the approximate C-optimal design for a linear regression model with uncorrelated observations and a finite experimental domain as a specific linear programming problem. The authors stated that the proposed algorithm can also be applied to *difficult* problems with singular C-optimal designs and relatively high dimension of β ; however, computing optimal designs with respect to other well-known criteria cannot be reduced to a linear programming problem. The algorithm the authors proposed (called SAC), which is based on the simplex method, identifies the design support points for a C-optimal design. It can also be applied to C-optimal design problems with a large experimental domain without significant loss of efficiency.

Research on the non-linear side has been more diversified. The related literature has mostly focused on the construction of D-optimal designs to estimate some fixed parameters. Recently, Loeza-Serrano and Donev (2014) drew attention to the lack of research on the estimation of variance components (or variance ratios) contributed to the literature by proposing a new algorithm for the construction of A- and D-optimal designs at such instances. Parameter estimation can get tedious for non-linear models in the sense of experimental effort and computational effectiveness. Sequential DOE has been proven to be very helpful in such cases to substantially reduce experimental cost. The experiments excluding the first one are run using the information on preceding experiments in order to optimize the design. Harman and Filová (2014) used a quadratic approximation of D-optimality criterion (DQ criterion) in the method they proposed when computing efficient exact experimental designs for linear regression models. They asserted that the main advantage of their method can be realized in case there are general linear constraints such as cost constraints on permissible designs. Bruwer and MacGregor (2006) extended the open-loop D-optimal design formulation of Koung and MacGregor (1994) for robust multi-variable identification. Their design formulations enable effective and efficient identification of robust models. The authors regarded that their design formulations also performed better in the presence of constraints using a two-input, two-output system as a case study. Even though the designs they proposed resulted in highly correlated physical input sequences as in the unconstrained case, the authors maintained that the designs would overcome this when highly unbalanced replications were used among the support-points to emphasize excitation of these directions. Similarly, Ucinski and Bogacka (2007) studied optimal experimental designs in the presence of constraints aiming to develop a theoretical background along with numerical algorithms for model discrimination design. The authors applied their numerical procedure in a chemical kinetic model discrimination problem in which some of the experimental conditions were allowed to continually vary during the experimental run.

Sagnol (2011) proposed an extension to Elving's theorem in the case of multi-response experiments and concluded that it would be possible to use second-order cone programming to compute the C-, A-, T-, and D-optimality

criteria when a finite number of experiments was to be run. The author also provided a way to avoid the complexity in the multi-response C-optimal designs.

DOE has often been employed in adaptive and optimal control. Pronzato (2008) underlined the role of DOE in the asymptotic behaviour of the parameters estimated. The author pointed out the strength of DOE as a tool to establish links between optimization, estimation, prediction and control problems. Pronzato (2008) presented a comprehensive review on the relationship between sequential design and adaptive control, the mathematical foundations of optimal experimental design when estimating parameters of dynamical models.

Mandal and Torsney (2006) proposed a way that makes it possible to calculate a probability distribution by first discretizing the (continuous) sample space and then using these disjoint clusters of points at each iteration until the algorithm converges.

According to Coles et al. (2011), the non-linear nature of most of the design criteria add too much complexity to the design algorithms. The authors also questioned the heuristic nature of many design algorithms and the lack of their convergence properties in the related literature. Thus, the design criteria should not be determined without taking the design algorithm into account as the individual choice for both the former and the latter could alter the final result. Goujot et al. (2012) proposed a method that does not require the use of a global optimization algorithm. In a similar study, a new method that blends results obtained from initial experimental design, empirical modelling, and model-based optimization to determine the most promising experiments that would be used as an input at the subsequent stage was introduced (Chi et al. 2012). The authors claimed that their approach could be used as an alternative to RSM, especially in case prediction uncertainty should be taken into account. The problem they were interested in can be classified as a multi-objective optimization problem. Balsa-Canto et al. (2007) formulated the optimal experimental design problem as a general dynamic optimization problem where the objective is to find those experimental variables that could be manipulated in order to achieve maximum information content (or minimum experimental cost), as measured by the Fisher information matrix. They illustrated their approach in the estimation of the thiamine degradation kinetic parameters during the thermal processing of canned tuna. Based on their results, the authors concluded that optimal dynamic experiments could both improve identifiability essentially and reduce the experimental effort. The authors employed a metaheuristic approach called scatter search method (SSm), which could guarantee convergence to the global solution, when simultaneously computing the system dynamics and the local parametric sensitivities.

Coles et al. (2011) presented an empirical formula for designing Bayesian experimental designs when D-optimality is employed. The authors considered the case of linearized experimental design and claimed that their approach can be generalized for both the case of non-linear experimental design and the case of Bayesian experimental design. They concluded that the choice of the design algorithm should be made by considering different aspects of the problem such as the experimental quality and the importance of computational efficiency.

Myunga et al. (2013) provided a thorough review on the use of Adaptive Design Optimization (ADO) in the construction of optimal experimental designs. ADO is a Bayesian statistical framework that can be employed to conduct maximally informative and highly efficient experiments. The authors compared the practicality of ADO and the traditional, non-adaptive heuristic approach to DOE and claimed that ADO combined with modern statistical computing techniques had high potential to lead the experimenter to better statistical inference while keeping the related cost at a minimum.

Even though DOE has been applied in a wide variety of areas, some problems are intrinsically ill-conditioned and/or very large and their solutions require the use of alternative methods such as metaheuristics that can reduce computation time while guaranteeing robustness in many occasions.

Kleijnen et al. (2004) focused on the non-traditional RSM, which searches for the input combination maximizing the output of a real system or its simulation. It is a heuristic that locally fits first-order polynomials, and estimates the corresponding steepest ascent paths. The authors proposed novel techniques that combined mathematical statistics with mathematical programming to solve issues stemming from the scale-dependence of the steepest ascent and the intuitive selection of its step size. One of the techniques, called adapted steepest ascent (ASA), accounts for the covariances between the components of the estimated local gradient. It is scale-independent; however, the step-size problem can only be solved tentatively. The other technique follows the steepest ascent direction using a step size inspired by ASA. Monte Carlo experiments showed that ASA would more likely lead to a better search direction than the steepest ascent would.

Box and Draper (1969) developed a heuristic approach called Evolutionary Operation, which iteratively builds a response surface around the optimum from the previous iteration. Torczon and Trosset (1998) defined and experimented with the use of merit functions chosen to simultaneously improve both the solution to the optimization problem and approximation quality. They used the distance between a possible new candidate point and an already evaluated point as a measure for the error of the metamodel (Bonte et al. 2010). A number of heuristic move-limit strategies have been developed for approximate design optimization. These methods vary the bounds of design variables in approximation iterations and differ from each other by different bound-adjustment strategies (Siomina and Ahlinder, 2008).

Alonso et al. (2011) proposed using simulated annealing to find the right permutations of levels of each factor in order to obtain uncorrelated main effects with a minimum number of runs. Factorial experiments are used in many scientific fields. As the number of factors increases, the number of runs required for a complete replica of the design grows exponentially. Usually, only a fraction of the full factorial is used. This is called a fractional factorial design. The key issue is to choose an appropriate fraction that verifies the desired properties, especially the orthogonal property.

When characterizing orthogonal fractional factorial, the following notation is used: $s_1^{k_1}; s_2^{k_2}; \dots; s_h^{k_h}(n)$ where n is the number of runs, s_i is the number of levels of the factors, and k_i is the number of factors with s_i levels. Let matrix d of dimension

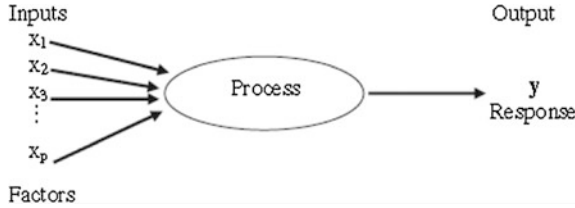


Fig. 8.1 Model of experimental design

$n \times p$ be built by factors as columns and runs as rows, with $p = k_1 + k_2 + \dots + k_h$. In the experimental design literature, d is known as the design matrix. An illustration of a typical DOE model is given in Fig. 8.1.

It is well known that the multiple linear regression model which represents an experimental design can be written as in Eq. (8.1).

$$y = \mu + X_1\beta + \varepsilon \tag{8.1}$$

where μ is the grand mean, y denotes the matrix of the response values, β denotes the matrix of the main effects coefficients, X_1 denotes the matrix of contrast coefficients for the vector of main effects, and ε denotes the vector of random errors.

If X^T is the transpose matrix of X , the correlations matrix is $X^T X$. The correlation matrix is an indicator of a good design. If the correlation matrix is diagonal, the computations will be simple and the estimators of all the regression coefficients are uncorrelated.

When orthogonal designs are not possible due to excessive runs and restricted budgets, it would be desirable to obtain a design as close as possible to an orthogonal one, with just a few runs. Such designs are called nearly-orthogonal, and generated by using several criteria. Alonso et al. (2011) employed a criterion based on Addelman frequencies that works with the design matrix. They applied simulated annealing to fractional factorial designs using the Addelman proportional frequencies criterion in order to obtain orthogonal designs.

Bates et al. (2003) used Genetic Algorithms (GA) to find the optimum points in the Audze–Eglais experimental design, which is achieved by distributing experimental points as uniformly as possible within the design domain. A uniform distribution results in the minimization of the potential energy of the points of a DOE. The potential energy is formulated in Eq. (8.2).

$$\min U = \min \sum_{p=1}^P \sum_{q=p+1}^P \frac{1}{L_{pq}^2} \tag{8.2}$$

where U is the potential energy and L_{pq} is the distance between the points p and q .

An example of Audze–Eglais Uniform Latin Hypercube (AELH) for two design variables and three points are given in Fig. 8.2.

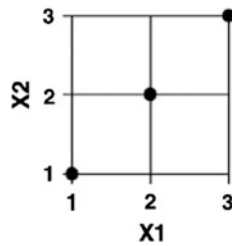


Fig. 8.2 An illustration of Audze-Eglais Latin Hypercube

Various experimental design combinations can be evaluated and the one with minimum objective function (i.e., Eq. (8.2) is minimized) is the AELH experimental design. Bates et al. (2003) carried out the search for the best DOE by minimizing the objective function in Eq. (8.2) using the Genetic Algorithm (GA). The fitness function of the GA is given in Eq. (8.2). For the encoding of two alternatives, the node numbers and the coordinates of the points are evaluated, and coordinates are chosen since it results in shorter length of chromosomes. Various numerical studies have been conducted varying the number of design variables and the number of points. The results indicate that the method works well and an improvement over previous results of Audze–Eglais Uniform Latin Hypercube experimental design and of random sampling Latin Hypercube experimental design has been achieved.

Chen and Zhang (2003) employed a GA for 2^{k-p} fractional factorial design. They used a MD-optimality criterion for optimizing the fractional factorial design. To select the optimal follow-up design using MD-optimality, traditionally the procedure below is used:

- (1) Identify potential regression models that can describe the response values in the initial experiment by using Bayesian analysis, and define all the factors appearing in these models as active factors.
- (2) Choose a set of runs (follow-up design) from all the experimental combinations of the active factors such that the best model can be discriminated from the potential regression models. Note that the effects of the factors and interactions included in the model are the significant effects in the experiment. Therefore, the confounded effects produced in the initial experiment are separated.

There is a weakness in this approach as the number of follow-up designs that needs to be examined significantly increases when the number of active factors increases, or the number of runs included in a follow-up design increases. Thus, Chen and Zhang (2003) developed a heuristic method based on an effective evolutionary algorithm and genetic algorithms (GA) for finding the optimal follow-up design. This heuristic is denoted as GA for maximum model-discrimination design (GAMMDD). In this GA, the encoding of a solution is represented as a follow-up design, U_i , which is described as a $n_1 \times k$ matrix, where n_1 is the number of experimental runs in the follow-up design, and k is the number of active factors. The fitness value is specified as the model-discrimination (MD) value of a design,

since the problem is to find a follow-up design that can identify a model with maximum model-discrimination value. Let X be a $n_1 \times k$ follow-up design matrix for k factors f_1, f_2, \dots, f_k ; and y denote the predicted vector under X , then the MD value for the design is calculated using Eq. (8.3).

$$MD = \frac{1}{2} \sum_{0 \leq i \neq j \leq m} p(M_i|y)p(M_j|y), \quad (8.3)$$

where $p(M_i|y)$ is the posterior probability of the model M_i , given y and considering a regression model M_i as in Eq. (8.4).

$$y = X_i\beta_i + \varepsilon_i \quad (8.4)$$

The computational results in their research show that the performance of GAMMDD was significantly better to that of the exchange algorithm, and would be able to enhance the strength of traditional two-step approach.

Lejeune (2003) implemented a one-exchange algorithm and used a generalized simulated annealing for the construction of D-optimal designs. The proposed method does not require to construct or to enumerate each point of the candidate set, whose size grows exponentially with the number of variables. In order to handle more complex problems, their procedure generates guided starting designs.

The focus is given to the D-optimality criterion, which requires the maximization of the determinant of the information matrix, $|X^T X|$, or, equivalently, its D-efficiency level formulated using Eq. (8.5)

$$D_{eff} = 100 \left(\frac{|X^T X|^{\left(\frac{1}{P}\right)}}{N} \right) \quad (8.5)$$

where P is the number of parameters, N is the number of experiments in the model. When a linear regression model is considered, $y = X_i\beta_i + \varepsilon_i$, any increase in the determinant of $X^T X$ reduces the error variances of the estimates.

The integrated algorithmic process presented to find the D-optimal designs has the following characteristics:

- The proposed algorithm selects a new point of the candidate space randomly and does not require such maximization operations.
- This is an important aspect in simulated annealing, which has also the advantage of preventing from premature convergence towards local optima and giving the possibility to escape from a sequence of local optima. In addition, the method does not involve the construction or enumeration of each point of the candidate set and is time-saving.
- The exchange algorithm is a one-exchange procedure.
- The algorithmic process includes a procedure for constructing guided starting designs. This procedure is implemented with, in mind, the objective of applying the algorithmic process for more complex models.

This procedure resulted in a highly D-efficient algorithmic process that could be applied for more complex models than those treated in the literature. The latter objective requires that the computing time does not rise exponentially with the number of factors. The time-saving property constitutes the third characteristic of the algorithmic process proposed.

Sanchez et al. (2012) focused on finding an experimental design that balances different competing criteria which is a multi-objective optimization problem. They tackled the problem by looking for the Pareto-optimal front in the competing criteria. They reported various criteria used in the literature such as A-, E-, and D-optimality criteria related to the joint estimation of the coefficients, or the I- and G-optimality criteria related to the prediction variance.

A design is said to be D-optimal when it achieves the maximum value of D in Eq. (8.5), which means the minimum volume of the joint confidence region, so the most precise joint estimation of the coefficients.

A- and E-optimality criteria are related to the shape of the confidence region (the more spherical the region, the less correlated the estimates). When the estimates are jointly considered, the $(1 - \alpha) \times 100$ % joint confidence ellipsoid for the coefficients is determined by the set of vectors β such that

$$(\beta - b)'X^T X(\beta * b) \leq P\hat{\sigma}^2 F_{\alpha, P, N-P} \tag{8.6}$$

where P is the number of estimated coefficients, N denotes the number of experiments in the design, $\hat{\sigma}^2$ is the variance of the residuals, (an estimate of σ^2) and $F_{\alpha, P, N-P}$ is the corresponding upper percentage point of an F distribution with P and $N - P$ degrees of freedom.

When using the I- and G-optimality criteria, the variance of the prediction is taken into account through the prediction variance. The variance of the response predicted for a given point x in the experimental domain, is given by Eq. (8.7) and the G-optimality criterion is shown in Eq. (8.8)

$$Var(\hat{y}(x)) = x'_{(m)}(X^T X)^{-1}x_{(m)}\hat{\sigma}^2 = d(x)\hat{\sigma}^2 \tag{8.7}$$

$$G = Nd_{max} = Nmax_x(d(x)) \tag{8.8}$$

A design is said to be G-optimal when it achieves the minimum value of G in Eq. (8.9), whereas I-optimality criterion uses the average value of $Nd(x)$ obtained by integrating it over the domain.

Sanchez et al. (2012) employed an evolutionary algorithm to compute the Pareto-optimal front for a given problem. The input for the algorithm is the number of factors (k), domain, model to be fit (that determines the number of coefficients, P) and number of experiments (N , $N \geq P$) to do so, and also the criteria to be taken into account. The evolutionary algorithm is designed such that each individual in the population is an experimental design ($N \times k$ design matrix), codified according to the search space and such that $\det(X^T X) \geq 0.01$. Every design is evaluated in terms of the criteria, so that the fitness associated to each individual is a vector.

The applicability and interpretability of the proposed approach was shown by an application to determine sulfathiazole in milk (substance that has a maximum residue limit established by the European Union) by using molecular fluorescence spectroscopy. Numerical results are presented and the results show that the proposed algorithmic approach makes it possible to address the computation of ad hoc experimental designs with the property of being optimal in one or several criteria stated by the user.

Fuerle and Sienz (2011) presented a procedure that creates Optimal Latin Hypercubes (OLH) for constrained design spaces. OLH in a constrained design space may result in infeasible points of experimental designs. Instead of omitting these infeasible points, a better mapping of the feasible space is generated using the same number of points by using permutation genetic algorithm. In the search procedure, the objective was set so that the Audze-Eglais potential energy of the points as shown in Eq. (8.2) is minimized.

8.4 The Use of Experimental Design in Artificial Intelligence and Fuzzy Methods

Experimental design (DOE) has been one of the most important tools to verify interactions and interrelations between parameters in the design of intelligent systems. Among these systems, artificial neural networks and fuzzy inference systems have been the most prominent ones to search for representations of the domain knowledge, reasoning on uncertainty, automatic learning and adaptation. Neuro-fuzzy system is an approach that can learn from the environment and then reason about its state. A neuro-fuzzy system is based on a fuzzy inference system, which is trained by a learning algorithm derived from artificial neural network theory.

The design of a neuro fuzzy system requires the tuning and configuration of the topology and many parameters. Setting the parameters such as the membership functions, number and shape of each input variable, learning rates is a difficult task. Zanchettin et al. (2010) used DOE for parameter estimation of two neuro-fuzzy systems—Adaptive Neuro Fuzzy Inference System (ANFIS) and Evolving Fuzzy Neural Networks (EFuNNs). A depiction of two intelligent systems (ANFIS and EFuNNs) is provided in Fig. 8.3.

The ANFIS architecture consists of a five-layer structure. In the first layer, the node output is the degree to which the given input satisfies the linguistic label associated to the membership functions named as premise parameters. In the second layer, each node function computes the firing strength of the associated rule. In the third layer, each node i calculates the ratio of the i th rule firing strength for the sum of firing strength of all rules. The fourth layer is the product of the normalized firing level and the individual rule output of the corresponding rule. Parameters in this layer are referred to as consequent parameters. EFuNNs also have a five-layer structure. Each input variable is represented by a group of spatially arranged

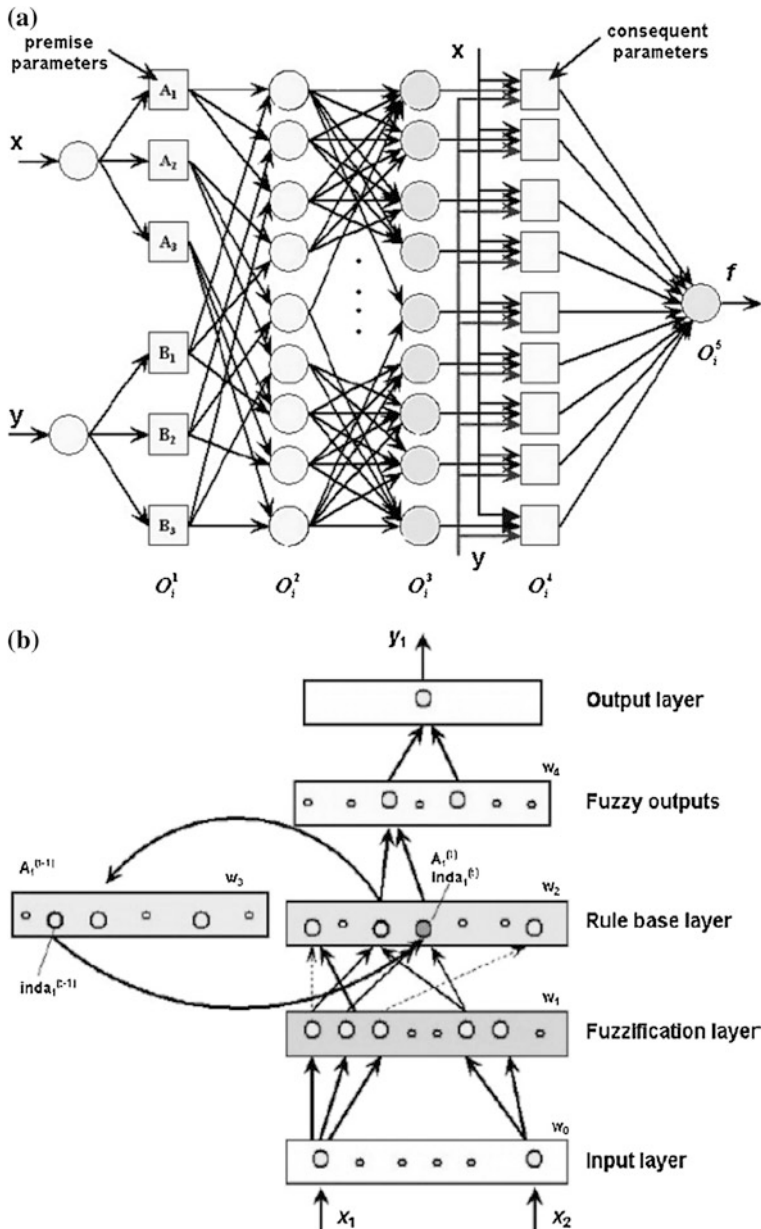


Fig. 8.3 A depiction of ANFIS (a) and EFuNN (b) (Zanchettin et al. 2010)

neurons to represent a fuzzy quantization of this variable. Fuzzy quantization in variable space is represented in the second layer of nodes. Different membership functions can be attached to these neurons (triangular, Gaussian, etc.). The

experiments for setting the parameters of the two intelligent systems are performed with four different prediction and classification problem datasets. The results show that for ANFIS, number of input membership functions and the shape of the output membership functions are usually the factors with the largest influence on the system's error measure. For the EFuNN, the membership function shape and the interaction between membership function shape and the number usually have the largest effect (Zanchettin et al. 2010).

Breban et al. (2013) used DOE for choosing the optimized parameters and determining the influence of the parameters of a fuzzy-logic supervision system in an embedded electrical power system. The fuzzy logic supervision system was developed to minimize the DC-link voltage variations, and to increase the system efficiency by reducing the dissipated power. In the experimental design step, first, the parameters and their variation range are chosen. Second task is to find the optimal ones and to test the system response to their changes. The most influential parameters are determined by testing the system response to each parameter extremity range modification. Breban et al. (2013) chose eight parameters, each with two extremity range values. Then, the influence of each parameter is tested on each optimization factor. For each parameter optimization, the low extremity range value becomes -1 , and the high extremity range value becomes $+1$. This assumption creates a matrix, called *test matrix*. Using relation (8.9), the influence E of each indicator is calculated as follows:

$$E = \frac{1}{n} M^t F \quad (8.9)$$

where n is the number of tests, M^t , the transpose of the test matrix and F , the indicators matrix of the parameters.

Basu et al. (2014) analyzed the process parameters of soap manufacturing industries. The process capability was determined using Fuzzy Inference System rule editor based on a set of justified "if-then" statements as applicable for the process. The data was collected in linguistic form to derive its process capability, using a set of justified rules and the effect of each factor was determined using DOE and ANOVA for improving the soap quality from the perspective of its softness. This article concludes that integrating fuzzy inference systems with DOE provides better results compared to those retrieved from DOE and Fuzzy Inference system in isolation.

Plumb et al. (2002) investigated the effect of experimental design strategy on the modelling of a film coating formulation by artificial neural networks (ANNs). Three different DOE approaches: (i) Box–Behnken, (ii) central composite and (iii) pseudo-random designs were used to train a multilayer perceptron (MLP). The structure of the ANN was optimized by training networks containing 3, 4, 5, 6, 7, or 9 nodes in the hidden layer. The predictive ability of each architecture was assessed by comparing the deviations mean square and R^2 from ANOVA analysis of the linear regression of predicted against observed property values. The architecture

with the lowest deviations mean square and highest R^2 was considered to be the most predictive one. Over-training was minimized by attenuated training.

Specifically, the onset of over-training was detected by setting a test error weight (W_T) calculated by Eq. (8.10):

$$W_T = \frac{N_{Test}}{(N_{Test} + N_{Train})} \quad (8.10)$$

where N_{Test} and N_{Train} are the number of records in the training and test sets, respectively.

As a result, ANN comprising six input and two output nodes separated by a single hidden layer of five nodes. The Box–Behnken and central composite models showed a poor predictive ability which is related to the high curvature of the response surfaces. In contrast, the pseudo-random design mapped the interior of the design space allowing improved interpolation and predictive ability. It was concluded that Box–Behnken and central composite experimental designs were not appropriate for ANN modelling of highly curved responses.

Alam et al. (2004) presented a case study which also investigated the experimental design on the development of artificial neural networks as simulation metamodels. The simulation model used in the study is a deterministic systems dynamics model. Six different DOE approaches which are the traditional full factorial design, random sampling design, central composite design, modified Latin Hypercube design and designs supplemented with domain knowledge are compared for developing the neural network metamodels. Various performance measures were used to evaluate the networks. The relative prediction error (RPE) which is commonly used for metamodels of deterministic simulations was used as a performance measure, which was defined as in Eq. (8.11)

$$RPE = \frac{\hat{Y}_r}{Y_r} \quad (8.11)$$

where Y_r is the known target value (simulation response) from the independent test data set, and \hat{Y}_r is the corresponding network output or prediction. Another measure of performance is the mean squared error of prediction (MSEP), defined as Eq. (8.12)

$$MSEP = \frac{1}{N} \sum (Y_r - \hat{Y}_r)^2 \quad (8.12)$$

The mean absolute percentage deviation (MAPD), which is used as the third performance measure is defined as Eq. (8.13)

$$MAPD = \frac{1}{N} \sum \left| \left[\frac{\hat{Y}_r - Y_r}{Y_r} \right] \right| \quad (8.13)$$

The neural network developed from the modified Latin Hypercube design supplemented with domain knowledge produced the best performance, outperforming networks developed from other designs of the same size.

Chang (2008) presented a case for the use of the Taguchi method for product design. Specifically, the aim was to optimize the parameter robust product design in terms of production time, cost, and quality as continuous control factors. They employed a four-stage approach based on artificial neural networks (ANN), desirability functions, and a simulated annealing (SA) algorithm to resolve the problems of dynamic parameter design with multiple responses. An ANN was employed to build a system's response function model. Desirability functions were used to evaluate the performance measures of multiple responses. AnSA algorithm was applied to obtain the best factor settings through the response function model.

Chang and Low (2008) also used Taguchi experiments to minimize various measures simultaneously (i.e., cost of the filter, its power loss, the total demand distortion of harmonic currents and the total harmonic distortion of voltages at each bus) of large-scale passive harmonic filters. Using the results of the Taguchi experiments as the learning data for an artificial neural network (ANN) model, an ANN was developed to predict the parameters at discrete levels. Then, the discrete levels were transformed into continuous scale using a genetic algorithm. Besides, the multiple objectives of the problem were tackled using the membership functions of fuzzy logic theory which were adopted in the algorithm for determining the weight of each single objective. The proposed approach significantly improves the performance of the harmonic filters when compared with the original design.

Balestrassi et al. (2009) applied DOE to find the optimal parameters of an Artificial Neural Network (ANN) in a problem of nonlinear time series forecasting. They presented a case study for six time series representing the electricity load for industrial consumers of a production company in Brazil. They employed an approach based on factorial DOE using screening, Taguchi, fractional and full factorial designs to set the parameters of a feed-forward multilayer perceptron neural network. The approach used classical factorial designs to sequentially define the main ANN parameters that a minimum prediction error could be reached. The main factors and interactions were identified using this approach and results suggest that ANNs using DOE can perform better comparably to the existent nonlinear autoregressive models.

Tansel et al. (2011) proposed using Taguchi Method and Genetically Optimized Neural Networks (GONNS) to estimate optimal cutting conditions for the milling of titanium alloy with PVD coated inserts. Taguchi method was used to determine the test conditions, the optimal cutting condition and influences of the cutting speed, feed rate and cutting depth on the surface roughness. GONNS was used to minimize or maximize one of the output parameters while the others were kept within a specified range.

Salmasnia et al. (2012) used DOE for data gathering to find the most valuable information used in a multiple response optimization problem. The multiple response optimization problem aims to find optimal inputs (design variables) to the system that yields in desirable values for stochastic outputs (responses).

Specifically, the problem of correlated multiple responses where relationship among response and design variables is highly nonlinear and the assumption that variance of each response is constant over the feasible region was tackled with a neuro-fuzzy (i.e. ANFIS) and principal component analysis derived desirability function. The resulting desirability functions were used to form a fitness function for optimization in GA. Effectiveness of the proposed method was presented through a numerical example.

Richard et al. (2012) proposed an alternative method to the classical response surface technique where the response surface was chosen as a support vector machine (SVM). An adaptive experimental design was used for the training of the SVM. As a result, the design can rotate according to the direction of the gradient of the SVM approximation leading to realistic samples. Furthermore, the precision of the probability of failure computation was improved since a closed form of expression of the Hessian matrix could be derived from the SVM approximation. This method was tested through a case study showing that high-dimensional problems can be solved with a fairly low computational cost and a good precision.

Hametner et al. (2013) dealt with the model based design of experiments for the identification of nonlinear dynamic systems. The aim of designing experiments was to generate informative data and to reduce the experimentation effort as much as possible as well as to comply with constraints on the system inputs and the system output. Two different modelling approaches, namely multilayer perceptron networks and local model networks were employed and the experimental design was based on the optimization of the Fisher information matrix of the associated model architecture. Deterministic data driven models with a stochastic component at the output were considered. The parameters of the considered models were denoted by θ . The measured output $y(k)$ at the time k was given by the model output $\hat{y}(k, \theta)$ plus some error $e(k)$. Then, the Fisher information matrix was formulated as in Eq. (8.14)

$$\tau = \frac{1}{\sigma^2} \sum_{k=1}^N \frac{\partial \hat{y}(k, \theta)}{\partial \theta} \frac{\partial \hat{y}(k, \theta)'}{\partial \theta} \quad (8.14)$$

The effects of the Fisher information matrix in the static and the dynamic configurations were discussed. Finally, the effectiveness of the proposed method was tested on a complex nonlinear dynamic engine simulation model. The presented model architectures for model based experiment design were compared.

Lotfi and Howarth (1997) proposed a novel technique named as the Experimental Design with Fuzzy Levels (EDFLs), which assigns a membership function for each level of variable factors. Traditionally, variable factors can be expressed with some linguistic terms such as low and high and they are converted into crisp values such as -1 , 0 , and $+1$. If some of the factor levels are not measurable, their values should be represented by equivalent fuzzy terms so that their importance is included in the system response. Using the fuzzy levels of factors, a set of fuzzy rules was used to represent the design matrix and observed responses. In this study, a number of examples were presented to clarify the proposed idea and

the results were compared with the conventional Taguchi methodology. In their study, they used a L_{18} orthogonal array EDFL for the application of the solder paste printing stage of surface mount printed circuit board assembly. For this case study, they provided a model for the process and optimized the selection of variable factors.

8.5 Conclusion

DOE is concerned with the selection of experimental settings that provide maximum information for the least experimental cost and can prove essential to successful modelling in an operating process application. According to the experimenter's objectives, DOE can dictate which variables should be measured, at which settings, and how many replicate measurements are needed to provide the required information (Fraleigh et al. 2003). The related literature offers very good examples for standard designs in case of fitting a first-order model; however, the choice of a response surface design for fitting a response surface design can be extremely challenging. Specifically, parameter estimation may not always be easy for non-linear models regarding experimental effort and computational effectiveness. Thus, the need for more flexible and/or specific designs is still viable (Anderson-Cook et al. 2009). Response surface methodology has seen the most significant progress in DOE-oriented research due to recent advances in metaheuristics and fuzzy techniques.

Coles et al. (2011) emphasized the need for a holistic approach that compares the quality of the optimal experimental design and the computational efficiency of the algorithm used for parameter estimation. They claimed that it would not always be possible to find a unique algorithm that could perform well for different types of objective functions. This is one of the most important challenges in DOE: maximizing the information to retrieve with scarce resources.

The number of avenues for future research is enormous. Bayesian techniques have been slightly touched in the literature. Active learning and nonlinear feedback control (NFC) are also available for further development according to Pronzato (2008). Computationally faster algorithms are still necessary especially for recently developed optimality criteria (Otsu, 2008). The derivation of lower (or upper) bounds or convergence properties of some algorithms should also be studied in more detail.

Another use of DOE is for tuning the parameters of artificial intelligence techniques such as neural networks, support vector machines or fuzzy inference systems. The literature shows that commonly traditional DOE methods are used to this aim. More sophisticated experimental design techniques (i.e. optimal DOE) for tuning the parameters of such systems present a new potential stream of research.

References

- Alam, F.M., McNaught, K.R., Ringrose, T.J.: A comparison of experimental designs in the development of a neural network simulation metamodel. *Simul. Model. Pract. Theory* **12**, 559–578 (2004)
- Alonso, M.C., Bousbaine, A., Llovet, J., Malpica, J.A.: Obtaining industrial experimental designs using a heuristic technique. *Expert Syst. Appl.* **38**, 10094–10098 (2011)
- Anderson-Cook, C.M., Borror, C.M., Montgomery, D.C.: Response surface design evaluation and comparison. *J. Stat. Plann. Infer.* **139**, 629–641 (2009)
- Balestrassi, P.P., Popova, E., Paiva, A.P., Lima, J.W.M.: Design of experiments on neural network's training for nonlinear time series forecasting. *Neurocomputing* **72**, 1160–1178 (2009)
- Balsa-Canto, E., Rodriguez-Fernandez, M., Banga, J.R.: Optimal design of dynamic experiments for improved estimation of kinetic parameters of thermal degradation. *J. Food Eng.* **82**, 178–188 (2007)
- Basu, S., Dan, P.K., Thakur, A.: Experimental design in soap manufacturing for optimization of fuzzified process capability index. *J. Manufact. Syst.* **33**, 323–334 (2014)
- Bates, S.J., Sienz, J., Langley, D.S.: Formulation of the Audze–Eglais Uniform latin hypercube design of experiments. *Adv. Eng. Softw.* **34**, 493–506 (2003)
- Bonte, M.H.A., Fourment, L., Do, T., van den Boogaard, A.H., Huétink, J.: Optimization of forging processes using finite element simulations: a comparison of sequential approximate optimization and other algorithms. *Struct. Multidisc Optim.* **42**, 797–810 (2010)
- Box, G.E.P., Draper, N.R.: *Evolutionary Operation: A Statistical Method for Process Management*. Wiley, Toronto (1969)
- Breban, S., Saudemont, C., Vieillard, S., Robyns, B.: Experimental design and genetic algorithm optimization of a fuzzy-logic supervisor for embedded electrical power systems. *Math. Comput. Simul.* **91**, 91–107 (2013)
- Bruwer, M.J., MacGregor, J.F.: Robust multi-variable identification: Optimal experimental design with constraints. *J. Process Control* **16**, 581–600 (2006)
- Chang, H.H.: A data mining approach to dynamic multiple responses in Taguchi experimental design. *Expert Syst. Appl.* **35**, 1095–1103 (2008)
- Chang, Y.P., Low, C.: Optimization of a passive harmonic filter based on the neural-genetic algorithm with fuzzy logic for a steel manufacturing plant. *Expert Syst. Appl.* **34**, 2059–2070 (2008)
- Chen, C.L., Lin, R.H., Zhang, J.: Genetic algorithms for MD-optimal follow-up designs. *Comput. Oper. Res.* **30**, 233–252 (2003)
- Chi, G., Hua, S., Yang, Y., Chen, T.: Response surface methodology with prediction uncertainty: a multi-objective optimisation approach. *Chem. Eng. Res. Des.* **90**, 1235–1244 (2012)
- Coles, D., Curtis, A., Coles, D., Curtis, A.: A free lunch in linearized experimental design? *Comput. Geosci.* **37**, 1026–1034 (2011)
- Das, A.: An Introduction to optimality criteria and some results on optimal block design. In: *Design Workshop Lecture Notes*, pp. 1–21. ISI, Kolkata 25–29 Nov 2002
- Fang, K.T., Tang, Y., Yin, J.: Lower bounds of various criteria in experimental designs. *J. Stat. Plann Infer* **138**, 184–195 (2008)
- Fraleigh, L.M., Guay, M., Forbes, J.F.: Sensor selection for model-based real-time optimization: relating design of experiments and design cost. *J. Process Control* **13**, 667–678 (2003)
- Fuerle, F., Sienz, J.: Formulation of the Audze–Eglais uniform Latin hypercube design of experiments for constrained design spaces. *Adv. Eng. Soft.* **42**, 680–689 (2011)
- Goujot, D., Meyer, X., Courtois, F.: Identification of a rice drying model with an improved sequential optimal design of experiments. *J. Process Control* **22**, 95–107 (2012)
- Hametner, C., Stadlbauer, M., Deregnaucourt, M., Jakubek, S., Winsel, T.: Optimal experiment design based on local model networks and multi-layerperceptron networks. *Eng. Appl. Artif. Intell.* **26**, 251–261 (2013)

- Harman, R., Filová, L.: Computing efficient exact designs of experiments using integer quadratic programming. *Comput. Stat. Data Anal.* **71**, 1159–1167 (2014)
- Harman, R., Jurík, T.: Computing c-optimal experimental designs using the simplex method of linear programming. *Comput. Stat. Data Anal.* **53**, 247–254 (2008)
- Imhof, L.A., Song, D., Wong, W.K.: Optimal design of experiments with anticipated pattern of missing observations. *J. Theor. Biol.* **228**, 251–260 (2004)
- Jones, D.R., Schonlau, M., Welch, W.J.: Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* **13**, 455–492 (1998)
- Joutard, C.: Applications of large deviations to optimal experimental designs. *Stat. Probab. Lett.* **77**, 231–238 (2007)
- Kleijnen, J.P.C., den Hertog, D., Angün, E.: Response surface methodology's steepest ascent and step size revisited. *Eur. J. Oper. Res.* **159**, 121–131 (2004)
- Koung, C.W., MacGregor, J.F.: Identification for robust multivariable control: the design of experiments. *Automatica* **30**(10), 1541–1554 (1994)
- Lejeune, M.A.: Heuristic optimization of experimental designs. *Eur. J. Oper. Res.* **147**, 484–498 (2003)
- Loeza-Serrano, S., Donev, A.N.: Construction of experimental designs for estimating variance components. *Comput. Stat. Data Anal.* **71**, 1168–1177 (2014)
- Lopez-Fidalgo, J., Trandafir, C., Tommasi, C.H.: An optimal experimental design criterion for discriminating between non-normal models. *J. Roy. Stat. Soc. B* **69**, 231–242 (2007)
- Lotfi, A., Howarth, M.: Experimental design with fuzzy levels. *J. Intell. Manuf.* **8**, 525–532 (1997)
- Lundstedt, T., Seifert, E., Abramo, L., Thelin, B., Nystrom, A., Pettersen, J., Bergman, R.: Experimental design and optimization. *Chemometr. Intell. Lab. Syst.* **42**, 3–40 (1998)
- Mandal, S., Torsney, B.: Construction of optimal designs using a clustering approach. *J. Stat. Plann. Infer.* **136**, 1120–1134 (2006)
- Myers, R.H., Montgomery, D.C., Vining, G.G., Borror, C.M., Kowalski, S.M.: Response surface methodology: A retrospective and literature survey, *J Qual Technol*, **36**(1), 53–77 (2004)
- Myunga, J.I., Cavagnaro, D.R., Pitt, M.A.: A tutorial on adaptive design optimization. *J. Math. Psychol.* **57**(3–4), 53–67 (2013)
- Otsu, T.: Optimal experimental design criterion for discriminating semiparametric models. *J. Stat. Plann. Infer.* **138**, 4141–4150 (2008)
- Patana, M., Bogacka, B.: Optimum experimental designs for dynamic systems in the presence of correlated errors. *Comput. Stat. Data Anal.* **51**, 5644–5661 (2007)
- Plumb, A.P., Rowe, R.C., York, P., Doherty, C.: The effect of experimental design on the modeling of a tablet coating formulation using artificial neural networks. *Eur. J. Pharm. Sci.* **16**, 281–288 (2002)
- Pronzato, L.: Optimal experimental design and some related control problems. *Automatica* **44**, 303–325 (2008)
- Pukelsheim, F.: *Optimal Design of Experiments*. Wiley, New York (1993)
- Richard, B., Cremona, C., Adelaide, L.: A response surface method based on support vector machines trained with an adaptive experimental design. *Struct. Saf.* **39**, 14–21 (2012)
- Rollins, D.K., Bhandari, N.: Constrained MIMO dynamic discrete-time modeling exploiting optimal experimental design. *J. Process Control* **14**, 671–683 (2004)
- Sagnol, G.: Computing optimal designs of multiresponse experiments reduces to second-order cone programming. *J. Stat. Plann. Infer.* **141**, 1684–1708 (2011)
- Salmasnia, A., Kazemzadeh, B., Tabrizi, M.M.: A novel approach for optimization of correlated multiple responses based on desirability function and fuzzy logics. *Neurocomputing* **91**, 56–66 (2012)
- Sánchez, M.S., Sarabia, L.A., Ortiz, M.C.: On the construction of experimental designs for a given task by jointly optimizing several quality criteria: Pareto-optimal experimental designs. *Analytica Chimica Acta* **754**, 39–46 (2012)

- Siomina, I., Ahlinder, S.: Lean optimization using supersaturated experimental design. *Appl. Numer. Math.* **58**, 1–15 (2008)
- Tansel, I.N., Gülmez, S., Demetgul, M., Aykut, S.: Taguchi Method-GONNS integration: Complete procedure covering from experimental design to complex optimization. *Expert Syst. Appl.* **38**, 4780–4789 (2011)
- Torczon, V., Trosset M.W.: Using approximations to accelerate industrial design optimization. In: Grandhi, R.V., Canfield, R.A. (eds.), *Proceedings of the 7th AAIA/NASA/USAF/ISSMO Symposium on Multidisciplinary Analysis and Optimization: A Collection of Technical Papers, Part 2*, pp. 738–748. American Institute of Aeronautics and Astronautics, Virginia, USA (1998)
- Ucinski, D., Bogacka, B.: A constrained optimum experimental design problem for model discrimination with a continuously varying factor. *J. Stat. Plan Infer.* **137**, 4048–4065 (2007)
- Wit, E., Nobile, A., Khanin, R.: Near-optimal designs for dual channel microarray studies. *Appl. Stat.* **54**, 817–830 (2005)
- Zanchettin, C., Minku, L.L., Ludermir, T.B.: Design of experiments in neuro-fuzzy systems. *Int. J. Comput. Intell. Appl.* **9**(2), 137–152 (2010)