Energy Conversion and Management 151 (2017) 630–640

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Review

# Experimental design and response surface methodology in energy applications: A tutorial review

Mikko Mäkelä

*Swedish University of Agricultural Sciences, Department of Forest Biomaterials and Technology, Division of Biomass Technology and Chemistry, Skogsmarksgränd, 90183 Umeå, Sweden*
*Aalto University, School of Chemical Engineering, Department of Bioproducts and Biosystems, PO Box 11000, 00076 Aalto, Finland*

ARTICLE INFO

*Keywords:*
Design of experiments
Multiple linear regression
Process modeling
Process optimization
Variance

ABSTRACT

Experimental design and response surface methodology are useful tools for studying, developing and optimizing a wide range of engineering systems. This tutorial provides a summary and discussion on their use in energy applications. The theory and relevant calculations are clearly presented and discussed along with model diagnostics and interpretation. This is followed by a review of recent reports within the energy field. Overall, this contribution will clarify many aspects of experimental design and response surface methodology that are often confusingly discussed in the academic literature and summarizes relevant applications where they have been found useful.

## 1. Introduction

Experimental design is a collection of tools used for studying the behavior of a system. Experimental design, or design of experiments, involves planning and performing a set of experiments to determine the effects of experimental variables on that system. The acquired data is separated into variation generated by the system itself and respective uncertainties or errors always present in empirical data. A statistically valid model is obtained, which by definition contains information on the effects of experimental conditions on the direction and magnitude of the measured response. The required experiments are also performed in a way that maximizes the information that can be extracted from a limited number of experiments. Once a satisfactory model has been determined, it can be used for predicting future observations within the original design range. Experimental design is thus useful for, not only studying, but also developing and optimizing a wide range of engineering systems. The method was originally developed by Fisher in the 1930s through factorial designs and analysis of variance for agricultural and biological research [1,2]. Response surface methodology was first discussed in the 1950s by Box and Wilson within chemical experimentation, and generally includes mathematical and statistical tools for both the design and analysis of response surfaces [3–5]. In practice, the methods are today closely related and the use of response surface methodology is without exception based on experimental designs. In this work, experimental design is used to refer to practices included in both topics.

Experimental design is closely related to the mentality of learning by experience and sequential experimentation. The effects and statistical significance of a larger group of experimental variables can be determined through factorial or screening designs, which enable choosing the relevant variables or conditions for the next set of experiments. Experimental designs are constructed in a way that eliminates or minimizes correlations between the chosen variables. This allows independent estimation of variable effects and their potential interactions. Here lies an important advantage of experimental design, as the variables are not varied one at a time while the others are being held constant. This approach assumes that the variables do not interact, i.e., the effect of one variable stays the same even though the others change. In many situations, this assumption can be unjustified. As an example, increasing the concentration of a catalyst might lower the temperature required for producing bio-oil of a specific quality. The effect of temperature thus changes based on catalyst concentration, indicating that the two variables interact.

Although experimental design is useful in many areas of energy research, it has no natural connection to the studied system. What is obtained is a simple mathematical approximation of the response based on empirical data. More simple designs and models are often easier to interpret, which increases their value in practical situations. The chosen design also determines the level of detail and complexity that can be described with the subsequent model. Factorial designs can be used for quantifying linear and interaction effects, whereas optimization designs allow describing more complex behavior by including higher order model components.

The mathematical and statistical procedures of experimental design

*E-mail addresses:* mikko.makela@slu.se, mikko.makela@aalto.fi.

http://dx.doi.org/10.1016/j.enconman.2017.09.021
Received 25 May 2017; Received in revised form 18 August 2017; Accepted 8 September 2017

have been well documented (e.g. [6–9]). In addition, several reviews have been published especially within analytical chemistry and food engineering [10–20]. As an example, Bezerra et al. [14] provided a good overview for the optimization of analytical methods and Leardi [15] included a discussion of mixture design for more specialized applications. Experimental design and response surface methodology are also frequently discussed in journals such as Technometrics published by the American Statistical Association [21]. What is missing is a practical tutorial that summarizes relevant reports on the use of experimental design within energy applications. This contribution aims to fulfill that knowledge gap. The theory and relevant equations are clearly presented and discussed along with model diagnostics and interpretation. The calculations are also illustrated based on an imaginary data set, which allows the reader to follow through the thought process. This is followed by a review of recent reports within the energy field. Overall, this contribution will clarify many aspects of experimental design that are often confusingly discussed in the academic literature and summarizes relevant applications where it has been found useful.

## 2. Materials and methods

This section describes the data set, relevant model calculations and diagnostics along with details on data compilation and review. Modeling results are then presented and discussed separately in Sections 3 and 4. The discussed calculations are also illustrated based on the data set. The calculations were performed using the Matlab R2016a (The Mathworks, Inc.) software package, but can be performed with any software capable of linear algebra. Open source alternatives are also available. Data plotting was performed with the OriginPro 2015 (OriginLab Corp.) software package. Once the modeling results have been presented and discussed, recently published work is reviewed to illustrate and discuss practical examples from the energy field.

### 2.1. The data set

The data set describes experiments that were performed to determine the effect of temperature and catalyst concentration on the molecular weight of bio-oil (Table 1). The data set was kept small to maintain simplicity. The temperature was varied within 160–320 °C and the catalyst concentration within 0.2–0.8%. The molecular weights of the attained oils were chromatographically determined and were in the range 0.59–2.0 kg mol$^{-1}$. A lower molecular weight was favorable to increase the performance of the oil in subsequent applications. The experimental order was randomized to minimize systematic errors.

The experiments were organized according to a face-centered central composite design with two variables or design factors and three replicated center-point experiments. A total of 11 experiments were performed. As illustrated in Table 1, the design included three levels for each variable and can be used for quantifying linear, interaction and higher-order model terms. The first four experiments in Table 1 equal a full factorial design on two variables. Factorial designs generally contain only two variable levels, a minimum and a maximum, and are described with the abbreviation $2^k$, where $k$ denotes the number of variables in the design. Experiments 5–8 are called star-points and are added to central composite designs for calculating higher order model components. Their distance α is expressed in coded units from the design center and is generally dictated by the problem at hand. Setting α as 1 or $\sqrt[4]{2^k}$ are secure choices. The value $\sqrt[4]{2^k}$ guarantees rotatability, i.e., spherical prediction variance around the design center. Experiments 9–11 are replicated center-points and enable estimating true replicate error.

**Table 1**
The data set based on a central composite design with two variables.

| Experiment | Temperature (°C) | Catalyst (%) | Molecular weight (kg mol$^{-1}$) |
|---|---|---|---|
| 1 | 160 | 0.2 | 2.0 |
| 2 | 320 | 0.2 | 0.85 |
| 3 | 160 | 0.8 | 1.8 |
| 4 | 320 | 0.8 | 1.0 |
| 5 | 160 | 0.5 | 1.7 |
| 6 | 320 | 0.5 | 0.59 |
| 7 | 240 | 0.2 | 1.4 |
| 8 | 240 | 0.8 | 1.2 |
| 9 | 240 | 0.5 | 0.89 |
| 10 | 240 | 0.5 | 1.2 |
| 11 | 240 | 0.5 | 0.94 |

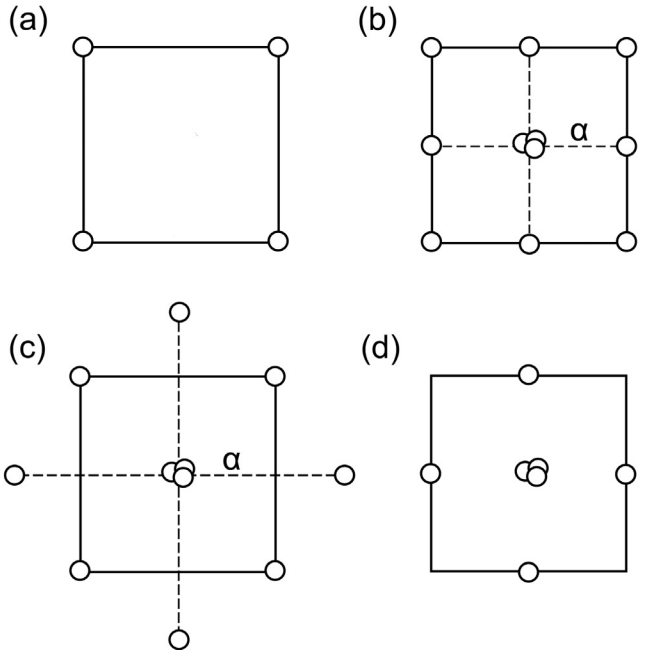


**Fig. 1.** Some common designs in two dimensions; (a) a factorial design, (b) a central composite design with a coded star-point distance α = 1, (c) a central composite design with α > 1 and (d) a Box Behnken design.

Central composite designs were first introduced by Box and Wilson in 1951 [3] and together with Box Behnken designs [22] have become one of the most common designs used for quadratic models. In general, many different designs are available in the literature, both for screening and optimization purposes, and will not be discussed here. The interested reader is recommended to turn to the many books or commercial software available in the field. Some common designs in two dimensions are illustrated in Fig. 1. The same logic applies in three or more dimensions.

### 2.2. Coding and model coefficients

Modeling is based on approximating the true behavior of a response:

$$y = f(\phi_1,\phi_2,\cdots,\phi_k) + \varepsilon \quad (1)$$

where $y$ denotes the measured response as a function of $(\phi_1,\phi_2,...,\phi_k)$ variables and other sources of variability $\varepsilon$. The variable values are coded to compare their effects within the design range:

$$x_i = \frac{(\phi_i - \phi_{min})}{\Delta\phi/2} - 1 \quad (2)$$

where $x_i$ denotes a coded value and $\phi_i$, $\phi_{min}$ and $\Delta\phi$ the respective variable value, minimum variable value and variable range, all in original units. In this way, the factorial design points in Table 1 range from −1 to 1 and the design center is situated at (0, 0). A quadratic regression equation is generally used to approximate $y$: