A Quantitative Measure to Evaluate Competing Designs for Nonlinear Dynamic Process Identification

Derrick K. Rollins*, Liza Pacheco, and Nidhi Bhandari

Department of Chemical Engineering, Iowa State University, Ames, Iowa 50011
*Author to whom correspondence should be sent

Abstract

The strategy for the collection of information (i.e., data) for model development is called experimental design. Optimal design seeks to maximize the information content under constraints of time and sampling. For the building of “gray-box” dynamic predictive nonlinear models, the dominant strategy has been the method of pseudo-random sequences (PRS). However, this work demonstrates the superiority of statistical design of experiments (SDOE) through a quantitative measure of information content, the D-optimal criterion.

Key words: Nonlinear, Hammerstein system, D-optimal criterion, experimental design

1. Introduction

One critical factor in modern day industrial survival is the ability to produce quality products. In seeking to meet this demands, many corporations have adapted the Six-Sigma management and operating philosophy, that requires quality and accountability of the highest degree from executives to plant operators. Process knowledge in the form of accurate modeling plays a major role in meeting these stringent quality and production objectives. However, to obtain cost effective accurate models, optimal experimental design is of critical necessity.

Experimental design is an intelligent process manipulation scheme that dictates the sizes and times of process changes over a pre-determined input space (called the design space) to
obtain information (i.e., data) for accurate model building. An optimal design gives maximum information to estimate true response behavior within a fixed amount of experimental time. The most popular and essentially exclusion method used by engineers to model dynamic processes has been the method of pseudo random sequences (PRS) (Brosilow and Joseph, 2002). A PRS is an experimental design that consists of a series of deterministic or random input changes at randomly determined times. In the time it would take the process to reach 99% of steady state for one set of input changes, several sets of input changes will be specified by a PRS design. If the changes in input levels are deterministic, they are usually just a few levels spaced over the input space. For example, for a pseudo random binary sequence (PRBS) uses only two levels for each input and they are the minimum and maximum values.

There are two major limitations of pseudo random sequence designs (PRSD) that contribute to its sub-optimal design properties. First, input changes are not based on optimal design theory and thus, are not mandated to be orthogonal to one another. When two inputs are not orthogonal they are said to confounded. Orthogonality or confounding of input can easily be checked by calculating their correlation coefficient. If it is non-zero, the set of inputs are non-orthogonal or confounded. When inputs are confounded, one cannot get a pure association of cause and effect. In the extreme case when two inputs are completely confounded (i.e., the correlation coefficient is 1 or -1), and only one input has an effect, there is no ability to determine which input from the data alone. Fitting a model to this case will equally share the causal effect of the two variables on the response which is incorrect and could be very undesirable in a context that depends strongly on correct causal models like model predictive control (MPC). The second major limitation of PRSD is the limited steady state information it
provides which leads to large estimation errors in the steady state model parameters (i.e., the static gain or ultimate response model). Steady state information is limited because PRSD typically have few or no changes with a long enough time duration that allows the process to get reasonably close to steady state. A similar observation is made by Tulleken (1990) in the case of identification of dynamic systems using random binary signals where the fast characteristics (e.g., dynamics) are identified better than the slower characteristics (e.g., gain or ultimate response) when using one-step-ahead prediction-error based methods. Rollins and Bhandari (2003) in the case of discrete time modeling (DTM) and Bhandari and Rollins (2003) in the case of continuous time modeling (CTM) demonstrate the importance of accurate estimation of the ultimate response portion of model on the overall performance of the dynamic model. Simply stated, a dynamic model will have limited accuracy if its ultimate response model is inaccurate, even when the process is strongly transient in its operation.

Based on the work of Rollins and Bhandari (2003) and Bhandari and Rollins (2003), we feel strongly that experimental design for dynamic modeling should be of a step test nature to allow processes to transition from one steady state to a new steady state to get adequate information to accurately estimate nonlinear ultimate response behavior. However, this requirement coupled with PRSD would require a prohibited amount of experimental time and would not address the confounding problem of PRSD. Thus, to keep experimental time to a minimum under this requirement, the number of step tests must be kept to a minimum. The approach that satisfies the requirements of optimality in the number of runs and orthogonality in the combination of input changes is statistical design of experiments (SDOE). However, SDOE has been developed in a strictly steady state context. Nonetheless, Rollins and Bhandari (2003)
and Bhandari and Rollins (2003) have demonstrated the superiority of SDOE over PRSD in simulation studies using a continuous stirred tank reactor (CSTR). Note that, in the dynamic context, the SDOE is selected and run as follows. The SDOE is selected as usual based on assumed ultimate response behavior. The selected design points are run as a sequential series of step tests. In their studies Rollins and Bhandari (2003) and Bhandari and Rollins (2003) used the same total run time and sampling time as the PRSD. They concluded that for equal cost of experimentation, the SDOE provided significantly more information than PRSD.

Although the simulations studies by Rollins and Bhandari (2003) and Bhandari and Rollins (2003) demonstrate better model building from SDOE, these studies do not give a quantitative measure of the relative greater information content of SDOE over PRSD. Hence, it is the objective of this work to demonstrate the superiority of SDOE over PRSD by introducing a quantitative measure of relative information content and conducting a study of comparison. The scope of this study consists of using a Hammerstein process with a polynomial nonlinear static model and first order dynamic model. The D-optimum criterion (Bates and Watts, 1988) is used as the measure of information content. This criterion seeks to minimize the confidence ellipsoid of the parameter vector. That is, it seeks to obtain confidence intervals for parameters with the smallest width. The relative measure that we use for comparison is the efficiency, which is the ratio of the value of the D-optimal criterion of the lesser design to the value of the D-optimal criterion to the superior design, i.e., the design with the greater information content.

Using this measurement of information efficiency, this work demonstrates very high SDOE superiority over PRSD for the cases studied. In addition, it demonstrates the usefulness of this tool to quantitatively evaluate competing design and thus, eliminates subjective opinions on
which of two competing design is best. For a single input, single output process, comparative results for efficiency have been presented by Rollins et al. (2003b). The article is organized as follows. Section 2 provides details on the formulation of the efficiency concept. Section 3 presents details on the Hammerstein system and the exact solution found by Rollins et al. (1998, 2003a). The solution is necessary to formulate analytical derivatives with respect to parameters which are used to formulate the information matrix in the measure of efficiency. Section 4 presents the design efficiency results for two, five, and seven input processes. Concluding remarks are given in Section 5.

2. Efficiency

Efficiency is widely used in the statistical literature to discriminate between experimental designs for the identification of nonlinear systems, mostly in the steady state setting (Bates and Watts, 1988; Atkinson and Donev, 1992). The most popular criterion used is the D-optimality criterion. The D-optimality criterion minimizes the general variance of the parameter estimates or in other words it minimizes the width of the confidence interval of the parameter estimates.

Eq. 1 below is used to calculate the percent efficiency between two competing designs, D1 and D2.

\[
\eta = \frac{\left(\frac{|\mathbf{V}^T \mathbf{V}|_{\text{D1}}}{N_{\text{D1}}^P}\right)^{1/P}}{\left(\frac{|\mathbf{V}^T \mathbf{V}|_{\text{D2}}}{N_{\text{D2}}^P}\right)^{1/P}} \times 100
\]

where P is the number of model parameters, N_{\text{D1}} and N_{\text{D2}} are the number of experimental points in Design 1 and Design 2, respectively and \( \mathbf{V} \) is the derivative matrix of the nonlinear model with (N x P) elements and it is defined in Eq. 2 below:
where  is the partial derivative of the output with respect to parameter \( \theta_i \) and \( i = 1, \ldots, P \) evaluated at each design point. These elements measure the sensitivity of the process model to each of the parameters.

The efficiency as described by Eqs. 1 and 2 gives a quantitative measurement of the information content of the competing designs under \( \alpha \) priori assumptions. The more efficient design is the one with the larger determinant of \( V^TV \). Design 1 is more efficient than Design 2, in terms of information to estimate the process parameters, if the ratio of their information matrix is larger than 1. To be able to apply Eqs. 1 and 2 analytical derivatives are needed, which requires a closed-form solution for the output. In this study, the nonlinear dynamic process that we have chosen is a Hammerstein system. The next section presents how the closed-form exact solution for a Hammerstein system developed by Rollins et al. (2003a) is used in the evaluation of the efficiency.

3. Hammerstein System

The Hammerstein system is classified as a block-oriented model that consists of a static nonlinear function followed by a linear dynamic block as shown in Figure 1. The advantage of this representation is that the identification process can be separated into the identification of the static nonlinear block and the linear dynamic block.

The use of Hammerstein models to represent nonlinear dynamic systems is a common
approach in the engineering systems community. The majority of the investigations have used discrete-time models (DTM) (Eskinat et al., 1991; Henson and Seborg, 1997; Pearson and Pottmann, 2000). However, DTM may not be practical in applications of slow or non-constant sampling (Chen and Rollins, 2000).

Figure 1. General structure of a Hammerstein system for a MIMO process.

3.1. Closed Form Exact Solution to Hammerstein Systems

Rollins et al. (2003a) presented a simple continuous-time approach to characterize a Hammerstein representation. Since this approach consists of an exact solution to the block-oriented system (Hammerstein system), it was called Hammerstein Block-Oriented Exact Solution Technique or H-BEST.
For an input sequence comprising of steps changes as shown in Eq. 3,

\[
\begin{align*}
0 \# t < t_1, & \quad u(t) \rightarrow u(0) \\
t_1 \# t < t_2, & \quad u(t) \rightarrow u(t_1) \\
t_2 \# t < t_3, & \quad u(t) \rightarrow u(t_2) \\
& \vdots \\
\end{align*}
\]

and given that the input vector \( u(t) \) and the output vector \( \xi(t) \) are expressed in deviation variables, the closed-form exact solution for the Hammerstein system proposed by Rollins et al. (2003a) is given by Eq. 4 below.

\[
y(t) \rightarrow y(0) \% \{ f(u(0)) \% (t) \% (s(t)) \}
\]

\[
\begin{align*}
\% \left[ y(t_1) \% \{ f(u(t_1) \& y(t_1)) \% (t \& t_1) \& (f(u(0)) \% (t)) \} \right] & \% (t \& t_1) \% \cdots \\
\% \left[ y(t_k) \% \{ f(u(t_k) \& y(t_k)) \% (t \& t_k) \& (y(t_{k+1}) \% \{ f(u(t_{k+1}) \& y(t_{k+1})) \} \right] \% (t \& t_k) \% \cdots
\end{align*}
\]

where \( V(t) \rightarrow f(u(t)) \), \( G(s) \rightarrow \frac{\xi(s)}{V(s)} \), \( g(t) \rightarrow \% s \{ \frac{G(s)}{s} \} \), \( s(t - t_k) \) is the shifted unit step function, and \( \%^{-1} \) is the inverse Laplace transform operator. The mathematical proof of the exactness of this technique can be found in Rollins and Bhandari (2002). In the following section we present the application of the efficiency technique to theoretical multiple-input examples.

4. Case Study

In order to demonstrate the ability of Eq. 1 to evaluate competing experimental designs, theoretical multiple input Hammerstein systems are used. We have considered three different cases; the first case has two input variables, the second has five input variables, and the third one has seven input variables.
4.1 Two Input Variables

The first multiple input system consists of two input variables, $u_1(t)$ and $u_2(t)$, representing the static gain (or steady state) function and a first order dynamics representing the dynamic function as shown in Eqs. 5 and 6 below:

$$f\{u(t)\} = a_1u_1(t) + a_2u_2(t) + a_3[u_1(t)]^2 + a_4[u_2(t)]^2 + a_5u_1(t)u_2(t)$$  \hspace{1cm} (5)

$$g(t) = \frac{a}{\xi}$$  \hspace{1cm} (6)

For this case the numerical values of $a_1$ to $a_5$ and $\tau$ are 2.25, 1.75, 0.25, 0.50, 1.5 and 5, respectively. The derivative matrix $V$ used to determinate the efficiency of the two designs is presented by Eq. 7 below.

$$V = \begin{bmatrix}
\frac{M_4(t)}{M_1} & \ddot{G} & \frac{M_4(t)}{M_5} & \frac{M_4(t)}{M_6} \\
\frac{M_4(t)}{M_1} & \ddot{G} & \frac{M_4(t)}{M_5} & \frac{M_4(t)}{M_6} \\
\frac{M_4(t)}{M_1} & \ddot{G} & \frac{M_4(t)}{M_5} & \frac{M_4(t)}{M_6} \\
\frac{M_4(t)}{M_1} & \ddot{G} & \frac{M_4(t)}{M_5} & \frac{M_4(t)}{M_6}
\end{bmatrix}$$  \hspace{1cm} (7)

The elements of the derivative matrix $V$ are obtained by partial differentiation of the exact solution and are shown in Eqs. 8 - 19.

For $0 < t \not= t_1$,

$$\frac{M_4(t)}{M_1} \ddot{u}_1(0) \begin{pmatrix} a_1 \\ 1 + \xi e^{\frac{\xi}{\tau}} \end{pmatrix}$$  \hspace{1cm} (8)

$$\frac{M_4(t)}{M_2} \ddot{u}_2(0) \begin{pmatrix} a_1 \\ 1 + \xi e^{\frac{\xi}{\tau}} \end{pmatrix}$$  \hspace{1cm} (9)
\[
\frac{M_\xi(t)}{M_3}, \begin{cases} 
M_1(0)u_1(0), \frac{\xi}{\tau}
\end{cases}
\] (10)

\[
\frac{M_\xi(t)}{M_4}, \begin{cases} 
M_2(0)u_2(0), \frac{\xi}{\tau}
\end{cases}
\] (11)

\[
\frac{M_\xi(t)}{M_5}, \begin{cases} 
u_1(0)u_2(0), \frac{\xi}{\tau}
\end{cases}
\] (12)

\[
\frac{M_\xi(t)}{M}, \begin{cases} 
(a_1u_1(0)a_2u_2(0)a_3u_1(0)^2a_4u_2(0)^2a_5u_1(0)u_2(0)) \left(\frac{\xi}{\tau}\right)(\frac{t}{\tau^2})
\end{cases}
\] (13)

For \(t_1 < t \neq t_2\),

\[
\frac{M_\xi(t)}{M_1}, \begin{cases} 
u_1(t_1), \frac{\xi(t_1\xi(t))}{\tau}
\end{cases}
\] (14)

\[
\frac{M_\xi(t)}{M_2}, \begin{cases} 
u_2(t_1), \frac{\xi(t_1\xi(t))}{\tau}
\end{cases}
\] (15)

\[
\frac{M_\xi(t)}{M_3}, \begin{cases} 
u_1(t_1), \frac{\xi(t_1\xi(t))}{\tau}
\end{cases}
\] (16)

\[
\frac{M_\xi(t)}{M_4}, \begin{cases} 
u_2(t_1), \frac{\xi(t_1\xi(t))}{\tau}
\end{cases}
\] (17)

\[
\frac{M_\xi(t)}{M_5}, \begin{cases} 
u_1(t_1)\nu_2(t_1), \frac{\xi(t_1\xi(t))}{\tau}
\end{cases}
\] (18)

\[
\frac{M_\xi(t)}{M}, \begin{cases} 
(a_1u_1(t_1)a_2u_2(t_1)a_3u_1(t_1)^2a_4u_2(t_1)^2a_5u_1(t_1)u_2(t_1)\xi(t_1)) \left(\frac{\xi}{\tau}\right)(\frac{t}{\tau^2})
\end{cases}
\] (19)

and so on.
Within this case three different scenarios were evaluated. The first scenario compares the efficiency of SDOE to PRSD. The second scenario evaluates the effect of experimental time length on the SDOE efficiency and the third scenario evaluates the effect of sampling time on the efficiency of a PRS design.

### 4.1.1 First scenario

A full factorial design was used to generate the input sequence for the SDOE in this scenario. The sequence has three inputs levels, coded from low to high as: -1, 0 and 1 as shown in Figure 2. Two methods were used for the generation of the PRS. The first method used a difference equation provided by Ljung (1999) for the generation of pseudo random binary sequences (PRBS) and we modified it to obtain three levels as shown in Eq. 20.

\[
\text{u}(t) \quad \text{rem}(A(q)u(t), 3) \quad \text{rem}\{a_1u(t \& 1) \% a_nu(t \& n), 3\} \tag{20}
\]

where \( u(t) \) is the level of the input, \( \text{rem}(x, 3) \) is the remainder as \( x \) is divided by 3, \( n \) is the order of the equation and decides the maximum length of the PRS. The coefficients \( a_1, ..., a_n \) depend on the order \( n \) and are also available in Ljung (1999). The PRS generated using Eq. 20 is shown in Figure 3. In the second method the input levels were randomly generated while the sequence switching time were fixed as shown in Figure 4. Since the PRSD is random, three different input sequences were used in the first method (with different values of \( n \)) while one hundred sequences were used in the second one to validate the efficiency results. For space considerations only one of the sequences from each method is presented graphically.

The time length of the sequences was based on the amount of time needed by the system to reach steady state for each step change in the SDOE design and was kept identical for both approaches (180 time units).
Figure 2.
The run time for a first-order process to reach 95% and 99.99% of steady state is $4\tau$ and $5\tau$, respectively. Although $5\tau$ is typically considered the run time to essentially reach steady state, we ran the step tests for the SDOE for $4\tau$ or 20 time units without much lost in model

accuracy as we shall illustrate in the second scenario.

The data were sampled every five time units, generating a derivative matrix $V$ with thirty-six rows. The efficiency was evaluated in two different ways: the first one included all the model parameters ($p = 6$) and the second one included just the steady state parameters ($p = 5$). The efficiency results are presented in Table 1 and it shows that since this ratio is always much less than one, the SDOE is a highly more efficient than the PRSD. The efficiencies for the modified Ljung method ranges between 4.4% and 5.1%. Thus, the SDOE, run as a series of step tests, provide a much higher level of information content (about 95% more) in estimating the model parameters. The mean and variance of a hundred sequences for the second PRSD method gave efficiencies of 1.4% and 0.07, respectively, which were even worse than the first PRSD method. For just the steady state parameters, both PRSD performed slightly worse than the cases when all the parameters are included indicating that a large amount of poor quality (i.e., lack of) information is associated with the ultimate response parameters. Figures 5 and 6 show how the sampling time relates to the input changes for both design methods. In the SDOE case, Figure 5 shows several samples over between input changes during transitions and leveling out periods. In contrast, the PRSD case, in Figure 6 shows no leveling out periods and the lack of sampling during a number of changes (i.e., a number of changes are occurring faster than the sampling rate). Later, in the third scenario, we demonstrate the impact of faster sampling on efficiency.

Table 1. Comparison of design efficiency results for the two input process.

<table>
<thead>
<tr>
<th>Design</th>
<th>% Efficiency: PRSD/SDOE</th>
</tr>
</thead>
<tbody>
<tr>
<td>All parameters (P = 6)</td>
<td>Steady State Parameters (P = 5)</td>
</tr>
<tr>
<td></td>
<td>PRSD/SDOE</td>
</tr>
<tr>
<td>4.4%</td>
<td>5.1%</td>
</tr>
<tr>
<td>PRSD</td>
<td>SDOE</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
</tr>
<tr>
<td>Modified Ljung</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Random Input Level</td>
<td>mean</td>
</tr>
<tr>
<td></td>
<td>variance</td>
</tr>
</tbody>
</table>

Figure 5. Process response to the SDOE input sequence shown in Fig. 2.
Figure 6. Process response to the PRS input sequence shown in Fig. 3 generated using the modified Ljung equation.

A theoretical explanation for the results given in this scenario will now be given by examining the contents in the information matrix in detail. For period representing the first input change and letting the true response be represented by $\xi(t)$, Eq. 4 becomes

$$\xi(t) = \begin{bmatrix} a_1 x_1(t) & a_2 x_2(t) & a_3 x_3(t) & a_4 x_4(t) & a_5 x_5(t) \end{bmatrix} \begin{bmatrix} 1 & e^{-t/\tau} \end{bmatrix}$$

(21)

where $x_1(t) = u_1(t)$, $x_2(t) = u_1(t)^2$, $x_3(t) = u_2(t)$, $x_4(t) = u_2(t)^2$ and $x_5(t) = u_1(t)u_2(t)$. Therefore, the partial derivatives for the steady state parameters ($a_1$ to $a_5$) are ($i = 1$ to 5):

$$\frac{M_i(t)}{M_i} \cdot x_i(t) \left( 1 \& e^{-t/\tau} \right)$$

(22)

Thus, we see from Eq. 22 that:
as $t \to 0$, \[ \frac{M_i(t)}{M_i} \to 0, \] and

as $t \to 4\frac{M_i(t)}{M_i}6\xi(t)$ \hspace{1cm} (24)

More specifically, Eq. 23 shows that the information context needed to estimate the steady state parameters vanishes as one approaches the time of change (i.e., at the initial steady state). In contrast, Eq. 24 shows that the information needed to estimate the steady state is maximized as the data are collected close to the final steady state. Note, that Eq. 24 also indicates that for high information content, it is not necessary to reach steady state, just to be reasonably close.

Now applying this theoretical analysis to the dynamic parameter, we first see that the partial derivative becomes:

\[
\frac{M_i(t)}{M} \left[ a_1 x_1(t) \% a_2 x_2(t) \% a_3 x_3(t) \% a_4 x_4(t) \% a_5 x_5(t) \right] \left( \frac{\xi}{\tau} \right) \left( \frac{t}{\tau^2} \right)
\] \hspace{1cm} (24)

Hence, Eq. 24 shows that as $t \to 0$ or as $t \to 4\frac{M_i(t)}{M_i}6\xi(t)$, the information for the dynamic parameters go to zero. More specifically, it shows that as $t \to 0$ or as $t \to 4\frac{M_i(t)}{M_i}6\xi(t)$,

\[
\frac{M_i(t)}{M} \to 0
\] \hspace{1cm} (25)

Thus, data close to initial or final steady states will not be very helpful in the estimation of the dynamic parameter. However, note that, for this dynamic system, Eq. 25 indicates that any good transient information should be sufficient in obtaining quality information to accurately estimate the dynamic parameters. Hence, either a step test or a PRSD will be adequate as long as the sampling is adequate (i.e., sufficiently fast enough).

**4.1.2 Second Scenario**
In this scenario the effect of shortening the experimental time between each step test on a SDOE design is studied. When the experimental time is reduced, the outputs are not allowed to reach steady state and there will be a penalty in the estimation of the steady state parameters. The next example will quantitatively show how much information is lost when the input change duration is reduced.

In this scenario, the time between step tests was taken as 15 (3τ), 20 (4τ) and 25 (5τ). Table 2 presents the efficiency results and shows that the efficiency is similar at any of the time lengths. It can be concluded that the reduction of input change time of the experiments can be reduced to 3τ without scarifying too much information to estimate the model parameters. In the next scenario, we examine how the efficiency is affected with different sampling times for PRSD.

Table 2. SDOE efficiency results for the two input process with different experimental time lengths

<table>
<thead>
<tr>
<th>Design</th>
<th>% Efficiency For SDOE #τ/SDOE 5τ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All parameters (P = 6)</td>
</tr>
<tr>
<td>SDOE</td>
<td></td>
</tr>
<tr>
<td>3τ</td>
<td>88.5</td>
</tr>
<tr>
<td>4τ</td>
<td>96.0</td>
</tr>
<tr>
<td>5τ</td>
<td>100</td>
</tr>
</tbody>
</table>

4.1.3 Third Scenario

This scenario will show how the efficiency can be used to quantitatively measure the amount of information that is gain when the sampling time rate is increased. In this scenario the modified Ljung method is the used to generate the PRSD. The efficiency is evaluated when the
data points are sampled every 0.5, 1 and 5 time units as shown in Figure 7. The graph shows that when the sampling time is every 0.5 time units the sequence can capture more dynamic and steady state information. Table 3 shows that as the sampling rate increases the efficiency increases too in agreement with Figure 7. In the next study, we similarly examine the efficiency for a five input system.

Table 3. PRSD efficiency results for the two input process when evaluated at different sampling rates.

<table>
<thead>
<tr>
<th>Design</th>
<th>% Efficiency: PRSD/SDOE</th>
<th>Sampling time interval (time units)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.5</td>
</tr>
<tr>
<td>SDOE</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>PRSD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modified Ljung</td>
<td>n = 6</td>
<td>30.0</td>
</tr>
<tr>
<td></td>
<td>n = 7</td>
<td>30.2</td>
</tr>
<tr>
<td></td>
<td>n = 8</td>
<td>28.3</td>
</tr>
</tbody>
</table>
Figure 7.
4.2 Five Input Variables

The second system consists of five input variables \( u_1(t), u_2(t), u_3(t), u_4(t) \) and \( u_5(t) \) representing the static gain (or steady state) function and first order dynamics representing the dynamic function as shown in Eqs. 26 and 27. As Eq. 26 shows, main effects, quadratic effects and interaction effects are present in this case.

\[
\begin{aligned}
f(u(t)) &= a_1 u_1(t) + a_2 \left[u_1(t)\right]^2 + a_3 u_2(t) + a_4 \left[u_2(t)\right]^2 + a_5 u_3(t) + a_6 \left[u_3(t)\right]^2 \\
&\quad + a_7 u_4(t) + a_8 \left[u_4(t)\right]^2 + a_9 u_5(t) + a_{10} \left[u_5(t)\right]^2 + a_{11} u_1(t)u_2(t) \\
&\quad + a_{12} u_1(t)u_3(t) + a_{13} u_1(t)u_4(t) + a_{14} u_1(t)u_5(t) + a_{15} u_2(t)u_3(t) \\
&\quad + a_{16} u_2(t)u_4(t) + a_{17} u_2(t)u_5(t) + a_{18} u_3(t)u_4(t) + a_{19} u_3(t)u_5(t) \\
&\quad + a_{20} u_4(t)u_5(t)
\end{aligned}
\]
For this process, the numerical values of $a_1$ through $a_{20}$ and $\tau$ are 2.4, 1.2, 2.6, 0.5, 1.5, 0.7, 3.2, 1.1, 4.8, 0.6, -3.0, 5.0, 6.7, -4.5, 2.3, 7.8, 5.4, 4.9, -2.3, 6.0, and 5.0, respectively. As in the previous process example, three different scenarios are studied. The first scenario compares SDOE with PRSD. The second scenario studies the effect of input change time on the SDOE efficiency and the third scenario studies the effect of sampling time on the PRSD efficiency.

$$g(t) = 1 \& e^{\frac{\xi}{\tau}}$$

(27)

4.2.1 First Scenario

In this scenario, a Box-Behnken design (46 runs) was used to generate the input sequence of the SDOE. The input sequence has three inputs levels, coded from low to high as: -1, 0 and 1. The two methods used for the generation of the PRSD’s were the modified Ljung (1999) method and the randomly generated input level method. The total experimental time for both design approaches was 920 time units. The data were sampled every five time units, generating a derivative matrix $V$ with one hundred and eighty-four rows.

Again, the overall efficiency was evaluated in two different ways: the first one included all the model parameters ($p = 21$) and the second one included the steady state parameters only ($p = 20$). Table 4 shows the per cent efficiency results and, as in the two factors case, it shows that SDOE is more efficient than PRSD. The efficiency for the PRSD when all the parameters were evaluated range from 20.9% to 21.4% for the first method; and for the second method the mean and the variance are 3.6% and 0.03%, respectively.

When only the steady state parameters were evaluated the efficiency for the modified Ljung equation method ranges between 15.3% and 15.5%. For the mean and variance for the
second method are 2.8% and 0.03% as shown in the last column of Table 4. These results are similar to the one in the first scenario and thus, give the same conclusions. The next scenario investigates at the effect of further reduction of test duration on the efficiency.

Table 4. Comparison of design efficiency results for the five input process.

<table>
<thead>
<tr>
<th>Design</th>
<th>% Efficiency: PRSD/SDOE</th>
<th>All parameters (P = 21)</th>
<th>Steady State Parameters (P = 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDOE</td>
<td>Box-Behnken</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>PRSD</td>
<td>Modified Ljung</td>
<td>1 20.9 15.3</td>
<td>2 21.2 15.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 21.4 15.5</td>
<td></td>
</tr>
<tr>
<td>Random Input Level</td>
<td>mean</td>
<td>3.7</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>variance</td>
<td>0.03</td>
<td>0.03</td>
</tr>
</tbody>
</table>

**4.2.2 Second Scenario**

In this scenario, the time between step tests was set to be 15 (3τ), 20 (4τ) and 25 (5τ) time intervals. Table 5 presents the efficiency results and shows that the efficiency is similar at least good at any of the time lengths. It can be concluded that the reduction of the duration of the experiments can be reduced to 3τ without greatly scarifying information to estimate the model parameters. The third scenario shows the effect of sampling rate on PRSD efficiency.

Table 5. SDOE efficiency results for the five input process with different experimental time lengths.
### 4.2.3 Third Scenario

As in the two factor case, this scenario will show the effect of sampling rate on PRSD efficiency. As previously, the Modified Ljung method is used to generate the PRSD’s. The efficiency is evaluated when the data points were sampled every 0.5, 1 and 5 time units. Table 6 shows that as the sampling time rate increases the PRSD efficiency increases but only gets to about 35% which is still much worse than SDOE. Next we examine experimental design efficiency for a seven input system.

#### Table 6. PRSD design efficiency results for the five input process at different sampling rates.

<table>
<thead>
<tr>
<th>Design</th>
<th>% Efficiency: PRSD/SDOE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sampling time rates (time units)</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
</tr>
<tr>
<td><strong>SDOE</strong></td>
<td>100</td>
</tr>
<tr>
<td><strong>PRSD</strong></td>
<td>Modified Ljung</td>
</tr>
<tr>
<td>n = 6</td>
<td>35.1</td>
</tr>
<tr>
<td>n = 7</td>
<td>35.2</td>
</tr>
<tr>
<td>n = 8</td>
<td>35.5</td>
</tr>
</tbody>
</table>

### 4.3 Seven Input Variables
The third system consists of seven input variables, \( u_1(t), u_2(t), u_3(t), u_4(t), u_5(t), u_6(t) \) and \( u_7(t) \), representing the static gain (or steady state) function and first order dynamics representing the dynamic function as shown in Eqs. 28 and 29.

\[
\begin{align*}
   f(u(t)) &= a_1 u_1(t) + a_2 u_1(t)^2 + a_3 u_2(t) + a_4 u_2(t)^2 + a_5 u_3(t) + a_6 u_3(t)^2 + a_7 u_4(t) + a_8 u_4(t)^2 + a_9 u_5(t) + a_10 u_5(t)^2 + a_11 u_6(t) + a_12 u_6(t)^2 + a_13 u_7(t) + a_14 u_7(t)^2 + a_15 u_1(t) u_7(t) + a_16 u_2(t) u_3(t) + a_17 u_2(t) u_4(t) + a_18 u_2(t) u_5(t) + a_19 u_2(t) u_6(t) + a_20 u_3(t) u_4(t) + a_21 u_3(t) u_5(t) + a_22 u_3(t) u_6(t) + a_23 u_3(t) u_7(t) + a_24 u_4(t) u_5(t) + a_25 u_4(t) u_6(t) + a_26 u_4(t) u_7(t) + a_27 u_5(t) u_6(t) + a_28 u_5(t) u_7(t) + a_29 u_5(t) u_7(t) + a_30 u_6(t) u_7(t) + a_31 u_6(t) u_7(t) + a_32 u_7(t) u_8(t) + a_33 u_7(t) u_9(t) + a_34 u_7(t) u_10(t) + a_35 u_7(t) u_11(t) + a_36 u_7(t) u_12(t) \tag{28}
   
   g(t) &= 1 + \frac{\delta}{\tau} \tag{29}
\end{align*}
\]

For this example the numerical values of \( a_1 \) through \( a_{35} \) and \( \tau \) are 2.4, 1.2, 2.6, 0.5, 1.5, 0.7, 3.2, 1.1, 4.8, 0.6, -3.0, 5.0, 6.7, -4.5, 2.3, 7.8, 5.4, 4.9, -2.3, 6.0, 0.5, 1.2, -0.7, 2.0, 1.0, 1.8, 1.3, 0.9, 3.4, 1.8, 2.9, -3.1, 0.4, and 5.0, respectively. As in the previous cases, SDOE is compared to PRSD. A Box-Behnken design (62 runs) was used to generate the input sequence of the SDOE. As before, two PRSD methods were used -- randomly generated input levels and randomly generated switching time. The data were sampled every five time units, generating a derivative matrix, \( V \), with two hundred and forty-eight rows.

For this scenario the overall efficiency was evaluated using all the model parameters \( (p = 36) \) only. Table 7 shows the results for the per cent efficiency and, as in the two and five input cases, it shows that SDOE is highly more efficient than PRSD. In both methods one hundred input sequences were evaluated, therefore only the mean and the variance of the results are
presented. For the randomly generated input level method the mean and variance are 2.8% and 0.01%, respectively; and for the random switching time method the mean and variance are 9.7% and 0.01%, respectively.

Table 7. Design efficiency results for the seven input study.

<table>
<thead>
<tr>
<th>Design</th>
<th>% Efficiency: PRSD/SDOE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All parameters</td>
</tr>
<tr>
<td></td>
<td>(P = 21)</td>
</tr>
<tr>
<td>SDOE</td>
<td>Box-Behnken</td>
</tr>
<tr>
<td>PRSD</td>
<td>Random Input Level</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Switching Time</td>
<td>mean 9.7</td>
</tr>
<tr>
<td></td>
<td>variance 0.01</td>
</tr>
</tbody>
</table>

5. Closing Remarks

For the dynamic parameters the pseudo random sequence design (PRSD) approach is more than adequate to obtain efficient information for model development. However, the accuracy of dynamic models depend very strongly on the accuracy of the ultimate response parameters. For estimating ultimate response parameters, PRSD is a poor approach for two critical reasons. First, to obtain quality information for estimating the ultimate response parameters, one needs quality ultimate response information. Since PRSD tends to require many input changes with few, to no, intervals between changes long enough to approach steady state
closely, it will be void of adequate ultimate response information to accurately estimate ultimate response parameters. This article demonstrated this limitation through very low values of PRSD efficiency for the ultimate response parameters and through the theoretical analysis given in Section 4.1. In addition, Rollins and Bhandari (2003) and Bhandari and Rollins (2003) demonstrate this limitation of PRSD in a study involving a mathematical MIMO continuous stirred tank reactor.

The second critical reason that PRSD suffers from poor ultimate response information is that the input levels are not controlled for optimality in terms of orthogonality. In contrast, statistical design of experiments (SDOE) not only provides orthogonality of inputs but the minimum number of input changes for adequate efficiency. This second property is critical since each design point (i.e., step test) should be run long enough to obtain adequate information to estimate the ultimate response parameters.

For a given situation, as the ratio of dynamic parameters to steady state parameters increases, the efficiency of PRSD to SDOE will increase but we do not believe that it will go above 100% because step tests are adequate to provide the necessary dynamic information and because of the limitations of PRSD for ultimate response information mentioned above. Thus, we feel very strongly that SDOE should be the design method of choice over PRSD for application of block-oriented nonlinear dynamic modeling. However, if one believes the contrary, we strongly suggest a comparative study like the ones presented here. Note that, there are other optimality criterions in the statistical literature if one prefers something other than D-optimality. However, given the limitations of PRSD mentioned above, we do not believe that a different criterion or block-oriented process will change our conclusion that SDOE is superior to
6. **Acknowledgments**

We would like to thank Dr. Max Morris and other attendees of the Engineering Statistics Working Group meetings for their helpful ideas and suggestions. We are also grateful to Brie Larson for her assistance with the simulations.

7. **Nomenclature**

- $a_i$: Parameters in the static gain function
- $f\{u(t)\}$: Steady state or ultimate response function
- $g(t)$: Dynamic function
- $P$: Number of model parameters
- $N_{D_i}$: Number of experimental runs
- $u(t)$: Input variable
- $y(t)$: Measured output variable
- $V$: Derivative matrix
- $v(t)$: Intermediate variable for a Hammerstein System
- $t$: Time

**Greek Letters**

- $\eta$: Efficiency
- $\theta_i$: Model parameters
- $\tau$: Dynamic parameter
- $\xi$: True output variable
Abbreviations

CSTR = Continuous Stirred Tank Reactor

H-BEST = Hammerstein Block-oriented Exact Solution Technique

MIMO = Multiple-Input, Multiple-Output

PRS = Pseudo Random Sequence

PRBS = Pseudo Random Binary Sequence

PRSD = Pseudo Random Sequence Design

SDOE = Statistical Design of Experiments

SISO = Single-Input, Single-Output

8. References


9. List of Figure Captions

Figure 1. General structure of a Hammerstein system for a MIMO process.

Figure 2. Input sequence based on SDOE for a two input (factor) system.

Figure 3. Input sequence generated by using the modified Ljung equation for a two input system.
Figure 4. Input sequence generated using random input levels for a two input system.

Figure 5. Process response to the SDOE input sequence shown in Fig. 2

Figure 6. Process response to the PRS input sequence shown in Fig. 3 generated using the modified Ljung equation.

Figure 7. Process response to the PRS input sequence shown in Fig. 3 generated using the modified Ljung equation. The data are sampled at intervals (Ts) of 0.5, 1, and 5 time units.