EXPERIMENTAL DESIGNS THAT MAXIMIZE INFORMATION FOR NONLINEAR DYNAMIC PROCESSES

Derrick K. Rollins, Nidhi Bhandari, and Liza Pacheco
Department of Chemical Engineering, Iowa State University
Ames, IA 50011

Abstract

Our discovery of a compact closed-form solution to Hammerstein systems has led to the ability to quantitatively compare competing experimental designs. This index is a quantitative measure of the information content in a design and is based on the D-optimality criterion. It is called efficiency in the statistical literature. This efficiency (ratio) shows the superiority of statistical design of experiments (SDOE) over the commonly used pseudo-random sequences (PRS) when modeling nonlinear dynamic processes. The ability of SDOE to control for confounding of input effects and the times of input changes is reflected in its higher efficiency.

Keywords

Nonlinear, D-optimal experimental design, Hammerstein systems

Introduction

The importance of accurate predictive models for optimal process operation and control is widely recognized. For example, if an inaccurate model is used in a control strategy it could cause deterioration of control resulting in inferior product quality. Process modeling can be broadly divided in two categories: theoretical and empirical. The complexity and cost of development of theoretical models has limited their applicability. Empirical models, as they are generated from experimental data, have become more popular. The usual procedure to develop empirical models has been to select a model structure and obtain estimates of the model parameters using data. The nonlinear and dynamic nature of chemical processes complicates the task of model development. In order to simplify the identification and estimation of nonlinear systems, block-oriented structures, which combine static nonlinear functions with a linear dynamics have been widely used (Billings, 1980 and Haber and Unbehauen, 1990).

Since, model identification relies heavily on data, ensuring high data quality is imperative. The experimental design has a direct bearing on data quality. The fundamental idea behind experimental design is to select an input sequence u(t) that minimizes some measure of the ultimate uncertainty in estimated results (Pearson and Ogunnike,1997). An optimal experimental design is one, which maximizes the information content of the output to estimate the unknown parameters.

A commonly used design in process identification is a Pseudo-Random Sequence or PRS. The pseudo-random signals are periodic and deterministic signals generated from pseudo-random sequences, which have similar characteristics to purely random signals (Godfrey et al. 1999). One most commonly known and widely used design in the identification of linear system is the pseudo-random binary sequence (PRBS). But since PRBS consists of only two levels, the resulting data does not provide sufficient information to identify nonlinear systems (Braun et al., 1999). To overcome this difficulty multilevel pseudo-random sequences are recommended. However, the PRS design with input changes at random times, cannot ensure that the collected information is sufficient for the estimation of all the significant effects in the model, e.g. bilinear effects.

The main objective of this paper is to demonstrate a quantitative measure of experimental design efficiency to evaluate competing designs. To this end, we present more details on the efficiency in the next section. The section after that presents the closed-form exact solution (to a Hammerstein system) which facilitates the quantitative comparison. A case study is presented in the following section and finally we have the conclusions.

Efficiency

Efficiency has been used in statistical literature to quantitatively compare competing experimental designs, for nonlinear models, mostly in a steady-state setting (Bates and Watts, 1988 and Atkinson and Donev, 1992). It has commonly been based on the D-optimality criterion, which 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) given below is used to calculate the efficiency:

$$\eta = \frac{\left[\frac{\left|\mathbf{V}^T\mathbf{V}\right|_{D1}}{(N_{D1})^p}\right]^{1/p}}{\left[\frac{\left|\mathbf{V}^T\mathbf{V}\right|_{D2}}{(N_{D2})^p}\right]^{1/p}} \tag{1}$$

where p is the number of parameters in the model, N_{D1} and N_{D2} are the number of experimental points in Design 1 and Design 2, respectively, and **V** is the derivative matrix for the nonlinear model with (N x p) elements and is defined in Eq. (2):

$$V_{N,p} = \begin{bmatrix} \left.\frac{\partial y}{\partial \theta_1}\right|_1 & \cdots & \left.\frac{\partial y}{\partial \theta p}\right|_1 \\ \vdots & \vdots & \vdots \\ \left.\frac{\partial y}{\partial \theta_1}\right|_N & \cdots & \left.\frac{\partial y}{\partial \theta_p}\right|_N \end{bmatrix} \quad (2)$$

where $\delta y/\delta\theta_i$ is the partial derivative of the output, y with respect to the model parameters and measures the sensitivity of the process model to each of the parameters.

Efficiency, given by Eq. (1), compares the information content of candidate designs under the *a priori* assumptions. The most efficient design will be the one with the largest determinant of V^TV. If the ratio is less that unity, Design 1 is inferior to Design 2 in terms of information content to estimate the process behavior. Taking the ratio of the determinant to the (1/p) power ensures that the efficiency calculated is based on the design size (N) and not on the model dimension, i.e. p.

The lack of continuous-time closed-form solutions for nonlinear dynamic systems has limited the applicability of Eq. (1). Our discovery of a compact closed-form exact solution to Hammerstein systems (Rollins et al., 2002) has provided a venue for utilizing efficiency to evaluate experimental designs. The details are presented next.

Hammerstein System

A true Hammerstein system is represented as a block-oriented model consisting of a nonlinear static function followed by a linear dynamic block as shown in Figure 1. Since the static nonlinear map is unrestricted, addressing nonlinear and interactive effects is possible (Rollins, 2002).

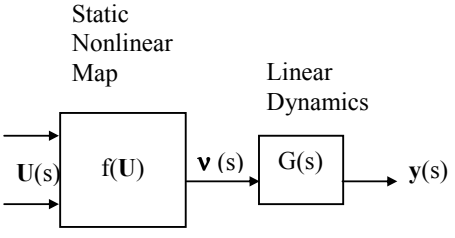

Figure 1. General structure of a Hammerstein model

The use of the Hammerstein representation to approximate physical systems in discrete-time has been widely studied (Eskinat and Johnson, 1991, Henson and Seborg, 1997, and Nelles, 2001). One disadvantage of using discrete-time models is that they are approximations of the Hammerstein system, which subsequently is an approximation of the underlying physical system. A continuous-time method would possibly eliminate one level of approximation.

Rollins et al. (2002) presented a continuous-time approach, to characterize a Hammerstein system, which is based on an exact solution to this block-oriented system. Therefore they called it the Hammerstein Block-oriented Exact Solution Technique or H-BEST.

For an input sequence comprising of step changes, given by Eq. (3) below,

$$u(t) = \begin{cases} u(0) & 0 \le t < t_1 \\ u(t_1) & t_1 \le t < t_2 \\ u(t_2) & t_2 \le t < t_3 \\ \vdots \end{cases} \quad (3)$$

and assuming that u(t) and y(t) are deviation variables, the Rollins et al. closed-form exact solution for the Hammerstein system is given by Eq. (4) below.

$$\left.\begin{array}{l} \text{For } 0 < t \le t_1, \\ \quad y(t) = f\{u(0)\}\, g(t) \\ \text{For } t_1 < t \le t_2, \\ \quad y(t) = y(t_1) + \left[f\{u(t_1)\} - y(t_1)\right] g(t - t_1) \\ \text{For } t_2 < t \le t_3, \\ \quad y(t) = y(t_2) + \left[f\{u(t_2)\} - y(t_2)\right] g(t - t_2) \\ \quad \vdots \qquad\qquad \vdots \end{array}\right\} \quad (4)$$

where v(t) = f(u(t)), G(s) = y(s)/v(s), $g(t) = L^{-1}\{G(s).\ 1/s\}$, and L^{-1} is the inverse Laplace transform operator.

A Simple Example

A theoretical single-input, single-output (SISO) Hammerstein system is used to demonstrate the procedure for utilizing Eq. (1) in comparing experimental designs. In this example the static gain (or steady state) function has quadratic behavior and the dynamic function has first-order dynamics as shown in Eq. (5).

$$\tau \frac{dy(t)}{dt} + y(t) = v(t) = a_1 u(t) + a_2 u(t)^2 \quad (5)$$

For step changes in the input (see Eq. (3)), Eq. (4) now becomes:

$$\left.\begin{array}{l} \text{For } 0 < t \le t_1, \\ \quad y(t) = \left(a_1 u(0) + a_2 u(0)^2\right)\left(1 - e^{\frac{-t}{\tau}}\right) \\ \text{For } t_1 < t \le t_2, \\ \quad y(t) = y(t_1) + \left[a_1 u(t_1) + a_2 u(t_1)^2 - y(t_1)\right]\left(1 - e^{\frac{-(t-t_1)}{\tau}}\right) \\ \text{For } t_2 < t \le t_3, \\ \quad y(t) = y(t_2) + \left[a_1 u(t_2) + a_2 u(t_2)^2 - y(t_2)\right]\left(1 - e^{\frac{-(t-t_2)}{\tau}}\right) \\ \quad \vdots \qquad\qquad \vdots \end{array}\right\} \quad (6)$$

where $f\{u(t)\} = a_1u(t) + a_2u(t)^2$, $G(s) = \frac{y(s)}{v(s)} = \frac{1}{\tau s + 1}$, and $g(t) = 1 - e^{-t/\tau}$.

For this example the numerical values for the model parameters a_1, a_2, and τ are chosen to be 1, 0.5, and 4, respectively. The derivative matrix **V** used to determine the efficiency of the two designs is presented in Eq. (7) below.

$$V = \begin{bmatrix} \left.\frac{\partial y}{\partial a_1}\right|_1 & \left.\frac{\partial y}{\partial a_2}\right|_1 & \left.\frac{\partial y}{\partial \tau}\right|_1 \\ \vdots & \vdots & \vdots \\ \left.\frac{\partial y}{\partial a_1}\right|_N & \left.\frac{\partial y}{\partial a_2}\right|_N & \left.\frac{\partial y}{\partial \tau}\right|_N \end{bmatrix} \quad (7)$$

The elements of the derivative matrix **V** are obtained by using the exact solution (i.e., Eq. (6)) and are shown in Eqs. (8) – (13).

For $0< t \le t_1$,

$$\frac{\partial y}{\partial a_1} = u(0)\left(1 - e^{\frac{-t}{\tau}}\right) \quad (8)$$

$$\frac{\partial y}{\partial a_2} = u(0)^2\left(1 - e^{\frac{-t}{\tau}}\right) \quad (9)$$

$$\frac{\partial y}{\partial \tau} = \left(a_1 u(0) + a_2 u(0)^2\right)\left(-e^{\frac{-t}{\tau}}\right)\left(\frac{t}{\tau^2}\right) \quad (10)$$

For $t_1 < t \le t_2$,

$$\frac{\partial y}{\partial a_1} = u(t_1)\left(1 - e^{\frac{-(t-t_1)}{\tau}}\right) \quad (11)$$

$$\frac{\partial y}{\partial a_2} = u(t_1)^2\left(1 - e^{\frac{-(t-t_1)}{\tau}}\right) \quad (12)$$

$$\frac{\partial y}{\partial \tau} = \left[a_1 u(t_1) + a_2 u(t_1)^2 - y(t_1)\right]\left(-e^{\frac{-(t-t_1)}{\tau}}\right)\left(\frac{t-t_1}{\tau^2}\right) \quad (13)$$

⋮ ⋮

The two design methods that will be compared: SDOE and PRS, shown respectively in Figures 2 and 3. Both design sequences have three input levels, coded from low to high as: –1, 0 and 1. The SDOE input sequences used in this example are implemented as step tests. For the generation of the PRS, the difference equation provided by Ljung (1987) for the generation of PRBS was modified to obtain three input levels. Since different PRS can result in different response, several sequences of different order (n) were used to ensure that the efficiency values obtained were similar. Only one PRS is shown for space considerations.

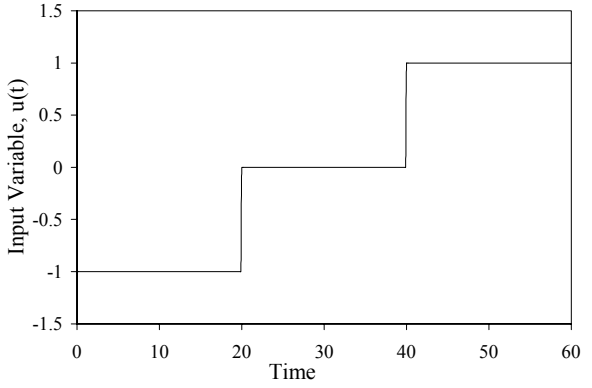

Figure 2. Input sequence for SDOE

The data were sampled every two minutes and the elements of each row of the derivative matrix **V** at evaluated at the successive sampling times. The derivative matrix for a design, which runs for 60 time units, would have 30 rows (excluding initial conditions). Two different scenarios have been considered in this study.

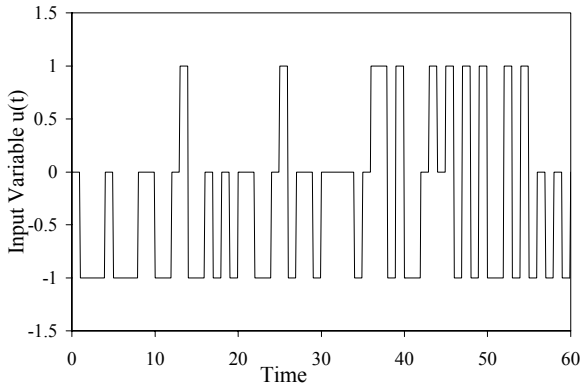

Figure 3. Input sequence for PRS (n=9)

Scenario 1

In the first comparison the time between step changes in the case of SDOE is sufficient to allow the process to essentially reach the new steady state. Steady state is assumed to be approximately reached at time 5τ, which is equal to 20 here. The length of each PRS was the same as that for SDOE (60 time units) to keep the experimental times equal. Table 1 presents the results for this scenario. This overall efficiency is evaluated by considering all the three columns in the derivative matrix and hence p = 3. The efficiencies for the PRS designs are quite low and lie between 6% – 15% as seen in the second column. Since the PRS designs never run long enough at any level for the response to reach steady state, they lack information to obtain accurate estimates of steady state parameters (a_1 and a_2). Their efficiency in estimating a_1 and a_2 is explicitly evaluated by considering only two columns in the derivative matrix and hence p = 2. The efficiency values are presented in the last column of Table 1. The efficiency of the PRS designs to estimate the steady state parameters lies between 5% – 10%. Thus, one can objectively evaluate competing designs and if one chose to use a PRS design, the best design (based on the a priori assumptions) could be selected from different candidate designs.

Table 1. Efficiency of the PRS designs relative to SDOE.

Design		**Efficiency = PRS/SDOE**	
		All parameters (p = 3)	*Steady-state only (p = 2)*
SDOE		1.0	1.0
PRS	n=6	0.133	0.084
	n=7	0.095	0.104
	n=8	0.155	0.098
	n=9	0.063	0.046

Note: n refers to the order of the PRS.

Scenario 2

In the second scenario, the effect of not allowing the outputs to reach steady state is studied for the SDOE designs. However, a shortening of the design time involves a penalty in terms of the loss of information needed to estimate model parameters and that is studied here. Often, when data are collected from a real experiment, the duration of the test is desired to be small. This example will quantitatively provide the extent of information lost when the test duration is reduced.

The efficiency is used to compare similar designs with different time interval between the step changes. The interval between the step changes corresponds to various fractions of 5τ. We looked at SDOE designs with an interval of 8, 12, and 16 between step changes and total test duration of 24, 36, and 48 respectively. The process response reaches, respectively, 86%, 95%, and 98% of the steady state response. These designs are compared with the SDOE design shown in Figure 2. The decline in efficiency as the experiment duration is reduced is expected, as seen in Table 2.

Table 2. Efficiency of SDOE designs of different duration

Design	Interval between Step Changes	Total Time	% of SS reached	Efficiency = D_i / D_4	
				All (p = 3)	*SS only (p = 2)*
D_1	2/5τ	24	86	0.607	0.604
D_2	3/5τ	36	95	0.813	0.792
D_3	4/5τ	48	98	0.934	0.916
D_4	5/5τ	60	99	1	1

Note: p refers to the number of columns in the derivative matrix, V.

It is interesting to note that the efficiency is roughly the same whether p = 3 or p = 2. The data needed to estimate the dynamic parameters are obtained early on in the experiment and therefore the estimation of dynamic parameters (e.g. τ in this example) is not significantly affected when the design length is shortened. The parameters in the steady-state model (e.g. a_1 and a_2) depend on steady-state data and if that data are not available, the loss of information is reflected in lower efficiencies of shorter tests. If experimental time were reduced to 60% of the longest design time, the efficiency would still remain at approximately 80%, as seen in Table 2. Thus the efficiency can also be used as a tool to come up with more practical designs without sacrificing too much information.

Closing Remarks

This work presented a methodology to evaluate different experimental designs and consequently their ability to provide accurate estimates of parameters in nonlinear models. The procedure was described specifically for a Hammerstein system, where an exact closed-form solution gave us the tools necessary to use this methodology, which is based on D-optimality.

It was demonstrated that SDOE is superior to PRS in the information it provides to estimate model parameters. This can be attributed to the fact that once the levels in a design have been fixed, they are maintained at that level so that process response reaches as close to the new steady state value as possible. On the other hand, PRS has the disadvantage that the input changes are random, so it cannot ensure if data close to steady state have been obtained.

This work presents the preliminary results of the application of this methodology and a more comprehensive analysis is underway. We propose to extend this framework to a multiple-input, multiple-output setting and to other types of block-oriented structures.

Acknowledgments

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

References

Atkinson, A. C., Donev, A. N. (1992) *Optimum Experimental Design*. Oxford University Press, New York, NY.

Bates, D. M., Watts, D. G.(1988). *Nonlinear regression analysis and its applications.* John Wiley & Sons, Inc., New York, NY.

Billings, S. A. (1980). Identification of nonlinear systems- a survey. *IEE Proc*., Vol. 127, Pt. D, No. 6.

Braun, M. W. et al. (1999). Multi-level Pseudo-Random Signal Design and "Model-on-Demand" Estimation Applied to Nonlinear Identification of a RTP Wafer Reactor. *In Proceedings of the American Control Conference*. San Diego, CA.

Eskinat, E., Johnson, S. H. (1991). Use of Hammerstein Models in Identification of Nonlinear Systems. *AIChE Journal*, Vol. 37, No. 2.

Godfrey, K. R., Barker, H. A., Tucker, A. J. (1999). Comparison of perturbation signals for linear system identification in the frequency domain. *IEE Proc.-Control Theory Appl*.,

Haber, R., Unbehauen, H. (1990). Structure identification of nonlinear dynamic systems – A survey on input/output approaches, *Automatica,* Vol. 26(4), pp. 651-677.

Henson, M. A., Seborg, D. E. (1997). Introduction in *Nonlinear Process Control*. Editors: M. A. Henson and D. E. Seborg Prentice-Hall PTR, Upper Saddle River, NJ.

Ljung, L (1987). *System identification: theory for the user*. Prentice-Hall, Englewood Cliffs, NJ.

Nelles, O. (2001). *Nonlinear System Identification - From Classical Approaches to Neural Networks and Fuzzy Models*. Springer-Verlag, Berlin, Germany. Vol. 146, No. 6.

Pearson, R. K., Ogunnaike, B. A. (1997). Nonlinear process identification, in *Nonlinear Process Control,* Editors: M. A. Henson and D. E. Seborg, Prentice-Hall PTR, Upper Saddle River, NJ.

Rollins, D. K., Bhandari, N., Bassily, A. M., Colver, G. M., and Chin, S. (2002). A continuous-time nonlinear dynamic predictive modeling method for Hammerstein processes. *Submitted to Industrial & Engineering Chemistry Research*

Rollins, D. K. (2002). A Continuous-Time Hammerstein Approach Working With Statistical Experimental Design," *Proc. "The Life of a Process Model: From Conception to Action" Workshop*, London, England.